

Near-surface Soil Moisture Assimilation in NWP

Clara S. Draper
Msc., B. Eng. (Hons), B. Sci.

Submitted in total fulfilment of the requirements of the degree of
Doctor of Philosophy

Department of Civil and Environmental Engineering
University of Melbourne
Parkville, Australia

August 4, 2011

Abstract

Root-zone soil moisture can exert a strong influence on atmospheric boundary layer forecasts. If not suitably constrained, the soil moisture in a Numerical Weather Prediction (NWP) model will drift away from the true climate, potentially resulting in erroneous atmospheric forecasts. In most NWP models this is prevented by constraining the model soil moisture to minimise errors in the screen-level (1.5 - 2 m) temperature and relative humidity forecasts. While assimilating screen-level observations can effectively decrease errors in boundary layer forecasts, it often generates unrealistic soil moisture, since the latter is adjusted to compensate for screen-level errors unrelated to the model surface state.

In response to these short-comings, this thesis seeks to establish whether NWP might benefit from assimilating remotely sensed soil moisture observations. The focus is on improving the model root-zone soil moisture, with the expectation that this will eventually lead to improved low-level atmospheric forecasts. Satellite remote sensors observe soil moisture in a thin near-surface soil layer, which interacts with the underlying soil moisture profile through diffusion processes, potentially offering a more direct relationship to root-zone soil moisture than the screen-level observations. Several studies have demonstrated that modeled root-zone soil moisture can be improved by assimilating near-surface soil moisture, although there has been very little work with NWP models. In particular, there have been no studies assimilating both near-surface soil moisture and screen-level observations at the continental scale.

First, the available remotely sensed soil moisture data sets are examined over Australia. At the time of this study, only the passive microwave Advanced Microwave Scanning Radiometer - Earth Observing System (AMSR-E) was providing soil moisture observations with sufficient coverage for use in NWP. Soil moisture data sets generated from AMSR-E brightness temperatures using four retrieval algorithms are assessed over 2006, principally by comparison to in situ

soil moisture observations from the Goulburn and Murrumbidgee Soil Moisture Monitoring Networks in southeast Australia. Based on this comparison the soil moisture data retrieved with the Vrije Universiteit Amsterdam - NASA (VUA-NASA) algorithm is identified as the most accurate. The accuracy of the VUA-NASA AMSR-E soil moisture data is then confirmed, by comparison to six years of in situ soil moisture observations from the Murrumbidgee, and to maps of observed precipitation and vegetation across Australia.

The VUA-NASA AMSR-E near-surface soil moisture data is then assimilated into the NWP land surface models used at the Australian Bureau of Meteorology (BoM) and Météo-France. The assimilation is performed with an Extended Kalman Filter (EKF) designed specifically to allow a computationally affordable assimilation of screen-level observations and near-surface soil moisture within operational NWP. The forward model itself (the land surface model) is used as the observation operator (to allow the assimilation of screen-level variables), and the assimilation is performed in an off-line version of the land surface model (to make the assimilation computationally affordable). The off-line assimilation is designed to be semicoupled to the NWP model, in that each soil moisture analysis is generated using the off-line land surface model forced with atmospheric forecasts generated from the NWP model updated with the previous land analysis update.

Using Météo-France's NWP land surface model (ISBA) over the European domain, the assimilation of AMSR-E near-surface soil moisture observations is compared to the assimilation of screen-level temperature and relative humidity over July 2006. The two observation types are first assimilated separately to determine how each is translated into root-zone soil moisture updates. These experiments showed no consistency between the root-zone soil moisture generated from each observation type, indicating that the screen-level observations could not have been substituted with AMSR-E data to achieve similar corrections to the low-level atmospheric forecasts. However, when both data types are assimilated together, the EKF is able to slightly improve the fit between the model forecasts and both observation types, although these improvements are extremely modest.

The AMSR-E observations are then assimilated into the BoM's NWP land surface model (MOSES) over a one year period, to test whether this improves the modeled root-zone soil moisture. The manner in which the assimilation constrains the root zone soil moisture from near-surface observations is first exam-

ined, and contrasted against the results from ISBA, to highlight the importance of model physics to soil moisture assimilation. Qualitatively, the AMSR-E data appears to have accurately detected errors in the (NWP forecast) precipitation used to force MOSES, however this could not be confirmed quantitatively (due either to inaccuracies in the AMSR-E data, or problems with the applied assessment technique). Nonetheless, comparison to the Murrumbidgee in situ soil moisture observations indicates that assimilating the AMSR-E data has improved the modeled near-surface and root-zone soil moisture at those sites, confirming the potential to improve modeled soil moisture by assimilating remotely sensed near-surface soil moisture observations. Combined with the results for the ISBA experiments, this suggests that forecasts of both soil moisture and the low-level atmosphere could be improved by assimilating both near-surface soil moisture and screen-level observations.

This is to certify that

1. this thesis comprises only my original work towards the PhD, except where indicated,
2. due acknowledgment has been made in the text to all other material used,
3. this thesis is less than 100,000 words in length, exclusive of tables, maps, bibliographies and appendices.

A handwritten signature in black ink, appearing to read 'Clara Draper', written in a cursive style.

Clara Draper, 1 August 2011.

Acknowledgments

First and foremost I would like to thank my supervisors, Jeffrey Walker, formerly of the University of Melbourne and now at Monash University, and Peter Steinle at the Centre for Australian Weather and Climate Research. I am extremely grateful to both of them for their hard work, insight, patience, and many hours of interesting conversation. Additionally, while not formally a supervisor, Jean-François Mahfouf of Météo-France / CNRM has very substantially shaped this thesis, and in particular the work relating to the ISBA/ALADIN models. I am extremely grateful to him for all of his input and for hosting my visits to Météo-France.

I would also like to thank the many scientists who have offered me advice, provided data, and/or shared their expertise with me along the way. In particular, I am grateful to staff at the National Snow and Ice Data Center, the Japan Aerospace Exploration Agency, and to Tom Jackson at USDA, and Richard de Jeu at the Free University of Amsterdam for providing me with AMSR-E soil moisture data. I am especially grateful to Richard and also Thomas Holmes for their input into the use of the VUA-NASA soil moisture data set. Additionally, I would like to acknowledge all of the staff and students at the Universities of Melbourne and Newcastle who have been involved in the provision of soil moisture monitoring data from the Murrumbidgee and Goulburn monitoring sites.

At Météo-France I would like to thank Patrick Le Moigne, Valéry Masson, Jean-Christophe Calvet, and Christoph Rüdiger for their assistance using SURFEX and navigating work at Météo-France. Karim Bergaoui of Institut National de la Météorologie in Tunisia is also acknowledged for coding an early version of the EKF assimilation system. At CAWCR I would like to thank Les Logon and Asri Sulaiman for assistance setting up the ACCESS model. Also, Imtiaz Dharssi at the UK Met Office has been extremely generous in sharing his wisdom regarding the intricacies of the JULES model, and I would also like

to thank Imtiaz and Bruce MacPherson for inviting me to the UK Met Office and providing valuable feedback on my work with JULES. Additionally, I am very grateful to Robert Gurney for inviting me to University of Reading, for what was a very interesting visit.

For financial support I am grateful to the Australian Research Council and the eWater CRC for funding my PhD Scholarships. I am also very grateful to the institutes that funded my overseas trips: the University of Melbourne for a Postgraduate Overseas Research Experience Scholarship to visit Météo-France, the eWater CRC for an international travel grant to attend the 2009 Joint Center for Satellite Data Assimilation Summer Colloquium on Data Assimilation, and the Fondation de Coopération Scientifique Sciences et Technologies pour l'Aéronautique et l'Espace for funding my second visit to Toulouse.

Contents

1	Introduction	1
1.1	Problem statement	1
1.2	Research objectives and outline	4
1.3	Outline of thesis	6
2	Literature Review and Research Outline	7
2.1	Overview of chapter	7
2.2	Soil moisture as an atmospheric control	7
2.3	Initialisation of soil moisture in NWP	9
2.3.1	Current initialisation methods: stand-alone land surface models	9
2.3.2	Current initialisation methods: assimilation of screen-level observations	10
2.3.3	Towards assimilation of land surface observations	12
2.4	Remote sensing of soil moisture	14
2.4.1	Microwave remote sensing	15
2.4.2	Future missions	20
2.4.3	Validation of remotely sensed soil moisture	20
2.5	Soil moisture assimilation	23
2.5.1	From near-surface to root-zone soil moisture	23
2.5.2	Soil moisture data assimilation techniques	26
2.5.3	Model - observation bias	29
2.5.4	Soil moisture assimilation in NWP	30
2.6	Outline of research	31
2.6.1	Assessment of remotely sensed soil moisture observations	33
2.6.2	Assimilation of remotely sensed soil moisture observations	34
2.7	Chapter summary	37

3	The EKF and the Land Surface Models	39
3.1	Overview of chapter	39
3.2	The EKF	40
3.3	Related assimilation approaches	42
3.4	The land surface models	43
3.4.1	The French ALADIN and ISBA models	43
3.4.2	The Australian ACCESS AND MOSES models	46
3.5	Chapter summary	51
4	Remotely Sensed Soil Moisture over Australia	53
4.1	Overview of chapter	53
4.2	Data and methods	54
4.2.1	AMSR-E soil moisture	54
4.2.2	In situ soil moisture	55
4.2.3	Methods to compare in situ and remotely sensed soil moisture	60
4.2.4	Precipitation data	63
4.3	Results	63
4.3.1	In situ soil moisture as a proxy for the area-average	63
4.3.2	Inter-comparison of the AMSR-E soil moisture retrievals	67
4.3.3	Assessment of VUA-NASA AMSR-E soil moisture	78
4.3.4	A filter for AMSR-E soil moisture	90
4.3.5	Inter-comparison of AMSR-E and ACCESS model soil moisture	93
4.4	Chapter summary	97
5	Assimilation Experiments with ISBA	99
5.1	Overview of chapter	99
5.2	Data and methods	100
5.2.1	AMSR-E soil moisture data	100
5.2.2	Screen-level data	102
5.2.3	The assimilation experiments and error selection	103
5.2.4	The relative information content	107
5.3	Implementing the EKF in ISBA	108
5.3.1	Testing the ISBA model Jacobians	108
5.3.2	The ISBA model Jacobians	111
5.4	Initial results: Assimilating AMSR-E soil moisture	117

5.4.1	Rescaling the AMSR-E data	118
5.4.2	Comparison of the EKF and SEKF assimilation of AMSR-E w_1	121
5.5	Comparison of the assimilation of AMSR-E near-surface soil moisture and screen-level observations	125
5.5.1	The observation increments	126
5.5.2	The analysis increments	127
5.5.3	Impact on subsequent observation increments	131
5.5.4	The relative information content	135
5.5.5	Comparison to SIM water balance	136
5.6	Chapter summary	141
6	Assimilation Experiments with MOSES	143
6.1	Overview of chapter	143
6.2	Data and methods	144
6.2.1	AMSR-E data	144
6.2.2	The Murrumbidgee Monitoring Network	145
6.2.3	Precipitation data	146
6.2.4	Köppen -Geiger climate classification	147
6.2.5	The assimilation experiments	147
6.3	Implementing the EKF in MOSES	149
6.3.1	MOSES model Jacobians	149
6.3.2	The model Jacobians	153
6.3.3	Model and observation error covariances	159
6.3.4	The state update vector	162
6.4	Testing the EKF: synthetic experiments	163
6.4.1	Time series examples	165
6.4.2	Net impact	173
6.5	Results: Assimilating AMSR-E data	176
6.5.1	Rescaling the AMSR-E observations	176
6.5.2	Time series examples	180
6.5.3	Net impact	189
6.5.4	Evaluation against in situ data	192
6.5.5	Comparison to observed precipitation	203
6.6	Chapter summary	206

7	Conclusions and Future Work	209
7.1	Chapter conclusions	210
7.1.1	Remotely sensed soil moisture over Australia	210
7.1.2	Assimilation experiments with ISBA	213
7.1.3	Assimilation experiments with MOSES	217
7.2	Main conclusions and future directions	221
	References	227
A	Related assimilation methods	247
A.1	Comparison to simplified 2-D Var	247
A.2	Comparison to standard EKF for assimilating near-surface soil moisture data	249
B	Comparing coupled MOSES and off-line JULES simulations	251
C	Statistics of fit between in situ data for non-surface layers at Kyeamba	257
D	Additional time series plots comparing original AMSR-E re- trieval algorithms to in situ data.	261
E	Additional time series plots comparing normalised AMSR-E retrieval algorithms to in situ data.	271
F	Additional time series plots comparing original VUA-NASA AMSR-E retrievals to in situ data.	281
G	Additional time series plots comparing normalised VUA-NASA AMSR-E retrievals to in situ data.	291

List of Figures

4.1	Location of the Murrumbidgee and Goulburn Monitoring Network sites, overlaid on the mean NDVI for 2006. See Tables 4.1 - 4.3 for details of each monitoring site.	56
4.2	Time series of soil moisture observations (m^3m^{-3}) for each layer over 2008, from five in situ soil moisture sensors at Kyeamba: K1 (blue), K2 (red), K3 (green), K5 (yellow), and K7 (purple), together with their mean (black).	64
4.3	Comparison of in situ (solid lines) and original (not normalised) AMSR-E (red diamonds) near-surface soil moisture (m^3m^{-3}) for each retrieval algorithm, at Kyeamba (M9) over 2006.	70
4.4	Comparison of in situ (solid lines) and original (not normalised) AMSR-E (red diamonds) near-surface soil moisture (m^3m^{-3}) for each retrieval algorithm, at Krui (G2) over 2006.	71
4.5	Comparison of in situ (solid lines) and normalised AMSR-E (red diamonds) near-surface soil moisture (m^3m^{-3}) for each retrieval algorithm, at Kyeamba (M9) over 2006.	72
4.6	Comparison of in situ (solid lines) and normalised AMSR-E (red diamonds) near-surface soil moisture (m^3m^{-3}) for each retrieval algorithm, at Krui (G2) over 2006.	73
4.7	Scatterplots of the AMSR-E soil moisture anomalies vs. the in situ anomalies over 2006 at Kyeamba (both in m^3m^{-3}), for the X-band retrieval algorithms.	77
4.8	Time series of in situ (black) and original (not normalised) AMSR-E (red) near-surface soil moisture (m^3m^{-3}) from the VUA-NASA retrieval algorithm at Kyeamba (M9), from 2003-2008.	82

4.9	Time series of in situ (black) and normalised AMSR-E (red) near-surface soil moisture (m^3m^{-3}) from the VUA-NASA retrieval algorithm at Kyeamba (M9), from 2003-2008.	83
4.10	Monthly mean AMSR-E soil moisture (m^3m^{-3}) across Australia, from the C-band ascending overpass (first row), C-band descending overpass (second row), X-band ascending overpass (third row), and the X-band descending overpasses (fourth row), for January (left) and July (right), 2006.	86
4.11	Monthly mean precipitation (in mm; first row), and NDVI (second row), for January (left) and July (right), 2006.	87
4.12	Maps of the C-band AMSR-E daily near-surface soil moisture anomaly (m^3m^{-3}), with 10 mm precipitation contours for 13 - 15 July, 2006, based on the average of the ascending and descending AMSR-E retrievals. Black indicates no AMSR-E data.	88
4.13	Time series of near-surface soil moisture (m^3m^{-3}) from C-band descending overpass AMSR-E (red), and Murrumbidgee Monitoring Network observations (black), for April 08 - April 09. The AMSR-E data have been normalised to the mean and variance of the in situ data, and filtered with an exponential moving average filter.	91
4.13	Time series of filtered and normalised C-band descending AMSR-E overpass soil moisture (red), normalised ACCESS near-surface soil moisture forecasts (blue), and Murrumbidgee Monitoring Network in situ observations (black), for April 08 - April 09.	95
4.14	Time series of precipitation ($mm\ day^{-1}$) from the Bureau of Meteorology's rain gauge analysis for the 0.25° grid surrounding M6, from April 2008 to March 2009.	97
5.1	Fraction of days in July 2006 for which the descending overpass AMSR-E soil moisture observations were available for assimilation (after quality control).	102
5.2	Screen-level observation network over the European domain. Black squares indicate GTS SYNOP observations, and small grey circles indicate French RADOME observations (reproduced from Figure 3 of Mahfouf et al. (2009)).	103

5.3	Comparison of \mathbf{M} (m^3m^{-3} / m^3m^{-3}) from 06:00 to 12:00 UTC on 1 July 2006, estimated using positive (x-axes) and negative (y-axes) perturbations of $10^{-4} \times (w_{fc} - w_{wilt})$	110
5.4	Comparison of \mathbf{M} (m^3m^{-3} / m^3m^{-3}) from 6:00 to 12:00 UTC on 1 July 2006, estimated using perturbations of $10^{-4} \times (w_{fc} - w_{wilt})$ (x-axes) and $10^{-1} \times (w_{fc} - w_{wilt})$ (y-axes).	112
5.5	Maps of the observation operator for modeled w_2 and observed T_{2m} (upper), RH_{2m} (middle), and w_1 (lower), for the analyses at 06:00 UTC (left) and 18:00 UTC (right) on 1 July 2006.	114
5.6	Surface wetness index at 18:00 on 1 July 2006 from the CTR simulation.	115
5.7	Time series of near-surface soil moisture (m^3m^{-3}) for a grid-cell in France (47.30 E/0.06 N) over 2006, from ALADIN (grey, solid), the original AMSR-E data (grey, dashed), and the seasonal-bias corrected (black, solid) and nonseasonal-bias corrected (black, dotted) CDF-matched AMSR-E data.	119
5.8	Net monthly w_2 increments (m^3m^{-3}) over July 2006, from the assimilation of a) seasonal-bias corrected AMSR-E data, and b) nonseasonal-bias corrected AMSR-E data (assuming approximately equal observation and background errors).	121
5.9	Kalman gain (m^3m^{-3}/m^3m^{-3}) for w_2 for the EKF at 18:00 on a) 2 July 2006 and b) 30 July 2006, and for the SEKF on c) 2 July 2006 and d) 30 July 2006.	122
5.10	Maps of the w_2 analysis increments (mm) accumulated over July 2006 by the a) EKF and b) SEKF assimilation of AMSR-E w_1 observations.	123
5.11	Histograms of the mean observation minus CTR forecast over July 2006 for a) T_{2m} (K), b) RH_{2m} (%), and c) w_1 (m^3m^{-3}).	126
5.12	Histograms of the w_2 analysis increments accumulated over July 2006 (mm) from the a) AMS, b) SLV, and c) CMB assimilation experiments.	128
5.13	Maps of the w_2 analysis increments (mm) accumulated over July 2006 from the a) AMS, b) SLV, and c) CMB assimilation experiments.	129

5.14	Maps of the w_2 analysis increments (mm) accumulated over July 2006 for each of the four daily assimilation cycles for the SLV assimilation. Each plot is labelled with the time of the analysis, and the colour scale is the same as used in Figure 5.13.	130
5.15	Diurnal cycle in the mean monthly observation minus model forecast over July 2006 for a) T_{2m} (K), b) RH_{2m} (%), and c) w_1 (m^3m^{-3}), for the CTR (black, solid), SLV (black, dashed), AMS (grey, solid), and CMB (grey, dashed) experiments. The horizontal axes are in hours.	132
5.16	Mean daily observation minus model forecast for each day in July 2006, for a) T_{2m} (K), b) RH_{2m} (%), and c) w_1 (m^3m^{-3}), for the CTR (black, solid), SLV (black, dashed), AMS (grey, solid), and CMB (grey, dashed) experiments.	133
5.17	Maps (left) and histograms (right) of the change in total soil moisture storage (mm) from 1 to 31 July 2006, from a) SIM, b) CTR, c) AMS, d) SLV, and e) CMB.	139
6.1	Coverage (% of days) of the AMSR-E descending overpass soil moisture observations over Australia from April 2008 - March 2009.	145
6.2	Köppen-Geiger climate zones over Australia (from Peel et al., 2007).	148
6.3	Dependence of estimated \mathbf{M}_{11} (y axis: %/%) on the perturbation used in the finite difference equation (x axis: fraction of soil moisture at saturation). The upper plot shows the mean of the Jacobians estimated with positive and negative perturbations, and the lower plot shows the absolute difference between the two estimates. All points plotted are the mean across the Australian domain, for the six hour forecast beginning from 09:00 UTC on 7 April 2008.	150
6.4	Scatterplots of the diagonal Jacobian terms (%/%), for the 6 hour forecast from 09:00 UTC on 6 April 2008, estimated using a perturbation of +0.1% (x-axis) and -0.1% (y-axis) of the soil moisture at saturation.	151
6.5	Maps of the linearised observation operator (%/%) for the 6 hour assimilation cycle from 09:00 UTC on 1 August 2008.	153

- 6.6 Maps of the MOSES soil moisture (as a fraction of the soil moisture at saturation) in each layer at 09:00 UTC on 1 August 2008. 154
- 6.7 Scatterplot of the $\partial S_1/\partial S_1$ (red, %/%) and $\partial S_1/\partial S_2$ (blue, %/%) verse S_1 (%) at Kyeamba A, for the twelve month period from April 2008. 156
- 6.8 Maps of the linearised observation operator (%/%) for the 6 hour assimilation cycle from 09:00 UTC on 1 February 2009. 157
- 6.9 Maps of the MOSES soil moisture (as a fraction of the soil moisture at saturation) in each layer at 09:00 UTC on 1 February 2009. 158
- 6.10 Soil moisture time series from the synthetic experiments at (137.375,-26.0) in arid (Bwh) central Australia, in units of % of saturation (left axis) and mm of soil moisture (right axis). The panels shows each soil layer (from top to bottom), from the EKF analyses with incorrect initialisation (solid, coloured), open-loop with incorrect initialisation (dashed, coloured), and the assumed truth from the reference open-loop (solid, black). Red (blue) lines indicate soil moisture initialised at 25% (75%) on 1 April 2008. 166
- 6.11 Time series of the soil moisture analysis increments (% of saturation) from the synthetic experiments at (137.375,-26.0), in arid (Bwh) central Australia, for layers one (upper) and two (lower). Red (blue) lines indicate soil moisture initialised at 25% (75%) on 1 April 2008. 167
- 6.12 Soil moisture time series from the synthetic experiments at (133.25,-13.625) in tropical (Aw) Northern Territory, in units of % of saturation (left axis) and mm of soil moisture (right axis). The panels shows each soil layer (from top to bottom), from the EKF analyses with incorrect initialisation (solid, coloured), open-loop with incorrect initialisation (dashed, coloured), and the assumed truth from the reference open-loop (solid, black). Red (blue) lines indicate soil moisture initialised at 25% (75%) on 1 April 2008. 168

- 6.13 Time series of the soil moisture analysis increments (% of saturation) from the synthetic experiments at (133.25, -13.625) in tropical (Aw) Northern Territory, for layers one (upper) and two (lower). Red (blue) lines indicate soil moisture initialised at 25% (75%) on 1 April 2008. 169
- 6.14 Soil moisture time series from the synthetic experiments at (147.125, -41.75), in temperate-oceanic (Cfb) Tasmania, in units of % of saturation (left axis) and mm of soil moisture (right axis). The panels shows each soil layer (from top to bottom), from the EKF analyses with incorrect initialisation (solid, coloured), open-loop with incorrect initialisation (dashed, coloured), and the assumed truth from the reference open-loop (solid, black). Red (blue) lines indicate soil moisture initialised at 25% (75%) on 1 April 2008. 170
- 6.15 Time series of the soil moisture analysis increments (% of saturation) from the synthetic experiments at (147.125, -41.75), in temperate-oceanic (Cfb) Tasmania, for layers one (upper) and two (lower). Red (blue) lines indicate soil moisture initialised at 25% (75%) on 1 April 2008. 171
- 6.16 Soil moisture time series from the synthetic experiments at (116.75, -34.25) in temperate-Mediterranean (Cfb) south West Australia, in units of % of saturation (left axis) and mm of soil moisture (right axis). The panels shows each soil layer (from top to bottom), from the EKF analyses with incorrect initialisation (solid, coloured), open-loop with incorrect initialisation (dashed, coloured), and the assumed truth from the reference open-loop (solid, black). Red (blue) lines indicate soil moisture initialised at 25% (75%) on 1 April 2008. 172
- 6.17 Time series of the soil moisture analysis increments (% of saturation) from the synthetic experiments at (116.75, -34.25) in temperate-Mediterranean (Cfb) south West Australia, for layers one (upper) and two (lower). Red (blue) lines indicate soil moisture initialised at 25% (75%) on 1 April 2008. 173
- 6.18 Maps of the soil moisture difference (% of saturation) from OPN_ACCESS for OPN_25 (left) and EKF_25 (right) on 1 July 2008, for S_1 (upper) and S_{RZ} (lower). 174

- 6.19 Near-surface soil moisture (as a fraction of saturation) on 1 August 2008 at 15:00 UTC from MOSES, and AMSR-E, before and after CDF-matching. 177
- 6.20 Time series of the assimilated AMSR-E observations minus the open-loop S_1 (% of saturation), averaged over the a) arid, b) tropical, c) temperate-oceanic, and d) temperate-Mediterranean climate zones. 178
- 6.21 Time series of precipitation (mm day^{-1}) from the BoM's rain gauge analysis (blue) and ACCESS forecasts (black), averaged over the a) arid, b) tropical, c) temperate-oceanic, and d) temperate-Mediterranean climate zones. 179
- 6.22 Time series of the modeled and observed daily precipitation and soil moisture at (137.375,-26.0) in arid (Bwh) central Australia. The upper panel shows the observed (blue) and modeled (red) precipitation (mm day^{-1}). The next four panels show the soil moisture in each layer (from top to bottom), from the EKF (red), the open-loop (black), and AMSR-E (blue), in units of % of saturation (left axis) and mm of soil moisture (right axis). 181
- 6.23 Time series of the soil moisture analysis increments (% of saturation) in layers one (upper) and two (lower) at (137.375,-26.0) in arid (Bwh) central Australia. 182
- 6.24 Time series of the modeled and observed daily precipitation and soil moisture at (133.25, -13.625) in tropical (Aw) Northern Territory. The upper panel shows the observed (blue) and modeled (red) precipitation (mm day^{-1}). The next four panels show the soil moisture in each layer (from top to bottom), from the EKF (red), the open-loop (black), and AMSR-E (blue), in units of % of saturation (left axis) and mm of soil moisture (right axis). 183
- 6.25 Time series of the soil moisture analysis increments (% of saturation) in layers one (upper) and two (lower) at (133.25, -13.625) in tropical (Aw) Northern Territory 184

6.26	Time series of the modeled and observed daily precipitation and soil moisture at (147.125, -41.75), in temperate-oceanic (Cfb) Tasmania. The upper panel shows the observed (blue) and modeled (red) precipitation (mm day^{-1}). The next four panels show the soil moisture in each layer (from top to bottom), from the EKF (red), the open-loop (black), and AMSR-E (blue), in units of % of saturation (left axis) and mm of soil moisture (right axis).	185
6.27	Time series of the soil moisture analysis increments (% of saturation) in layers one (upper) and two (lower) at (147.125, -41.75) in temperate-oceanic (Cfb) Tasmania.	186
6.28	Time series of the modeled and observed daily precipitation and soil moisture at (116.75, -34.25) in temperate-Mediterranean (Cfb) south West Australia. The upper panel shows the observed (blue) and modeled (red) precipitation (mm day^{-1}). The next four panels show the soil moisture in each layer (from top to bottom), from the EKF (red), the open-loop (black), and AMSR-E (blue), in units of % of saturation (left axis) and mm of soil moisture (right axis).	187
6.29	Time series of the soil moisture analysis increments (% of saturation) in layers one (upper) and two (lower) at (116.75, -34.25) in temperate-Mediterranean (Cfb) south West Australia.	188
6.30	Time series of the mean near-surface soil moisture (% of saturation) from the EKF (blue) and the open-loop (red), averaged over the a) arid, b) tropical, c) temperate-oceanic, and d) temperate-Mediterranean climate zones.	189
6.31	Time series of the mean root-zone soil moisture (% of saturation) from the EKF (blue) and the open-loop (red), averaged over the a) arid, b) tropical, c) temperate-oceanic, and d) temperate-Mediterranean climate zones. Note the different vertical axis for the tropical zone.	190
6.32	Time series of the mean daily evapotranspiration (mm day^{-1}) from the EKF (blue) and the open-loop (red) experiments, averaged over the a) arid, b) tropical, c) temperate-oceanic, and d) temperate-Mediterranean climate zones. Note the different vertical axis for the tropical time series.	191

- 6.33 Time series of modeled and observed precipitation and soil moisture at M1. The upper panel shows precipitation from observations (blue) and ACCESS forecasts (red). The middle and lower panels shows S_1 and S_{RZ} , respectively, from the EKF (blue), open-loop (red), and in situ observations (black), in units of % of saturation (left axis) and mm of soil moisture (right axis). The blue diamonds indicate the assimilated AMSR-E S_1 195
- 6.34 Time series of modeled and observed precipitation and soil moisture at M4. The upper panel shows precipitation from observations (blue) and ACCESS forecasts (red). The middle and lower panels shows S_1 and S_{RZ} , respectively, from the EKF (blue), open-loop (red), and in situ observations (black), in units of % of saturation (left axis) and mm of soil moisture (right axis). The blue diamonds indicate the assimilated AMSR-E S_1 196
- 6.35 Time series of modeled and observed precipitation and soil moisture at M5. The upper panel shows precipitation from observations (blue) and ACCESS forecasts (red). The middle and lower panels shows S_1 and S_{RZ} , respectively, from the EKF (blue), open-loop (red), and in situ observations (black), in units of % of saturation (left axis) and mm of soil moisture (right axis). The blue diamonds indicate the assimilated AMSR-E S_1 197
- 6.36 Time series of modeled and observed precipitation and soil moisture at M7. The upper panel shows precipitation from observations (blue) and ACCESS forecasts (red). The middle and lower panels shows S_1 and S_{RZ} , respectively, from the EKF (blue), open-loop (red), and in situ observations (black), in units of % of saturation (left axis) and mm of soil moisture (right axis). The blue diamonds indicate the assimilated AMSR-E S_1 198
- 6.37 Time series of modeled and observed precipitation and soil moisture at M8. The upper panel shows precipitation from observations (blue) and ACCESS forecasts (red). The middle and lower panels shows S_1 and S_{RZ} , respectively, from the EKF (blue), open-loop (red), and in situ observations (black), in units of % of saturation (left axis) and mm of soil moisture (right axis). The blue diamonds indicate the assimilated AMSR-E S_1 199

6.38	Time series of modeled and observed precipitation and soil moisture at KA. The upper panel shows precipitation from observations (blue) and ACCESS forecasts (red). The middle and lower panels shows S_1 and S_{RZ} , respectively, from the EKF (blue), open-loop (red), and in situ observations (black), in units of % of saturation (left axis) and mm of soil moisture (right axis). The blue diamonds indicate the assimilated AMSR-E S_1	200
6.39	Time series of modeled and observed precipitation and soil moisture at KB. The upper panel shows precipitation from observations (blue) and ACCESS forecasts (red). The middle and lower panels shows S_1 and S_{RZ} , respectively, from the EKF (blue), open-loop (red), and in situ observations (black), in units of % of saturation (left axis) and mm of soil moisture (right axis). The blue diamonds indicate the assimilated AMSR-E S_1	201
6.40	Time series of the root-zone soil moisture (mm) averaged over the a) arid, b) tropical, c) temperate-oceanic, and d) temperate-Mediterranean climate zones, from the EKF (blue), open-loop (red), and MOSES_PRECIP (black) simulations.	205
B.1	Difference between the layer one soil moisture (as a fraction of saturation) from the off-line and coupled MOSES models (off-line minus coupled), for the 24-hour forecast to 07:00 UTC on 2 November, 2008.	253
B.2	ACCESS forecast convective (upper) and large-scale (lower) rainfall (mm/day) for the 24-hours to 07:00 UTC on 2 November, 2008.	254
B.3	Difference between the layer one soil temperature from the off-line and coupled MOSES models (off-line minus coupled), for the 24-hour forecast to 07:00 UTC on 2 November, 2008.	255
D.1	Comparison of in-situ (solid lines) and original AMSR-E (red diamonds) near-surface soil moisture (m^3m^{-3}) for each retrieval algorithm, at Cooma Airfield (M1) over 2006.	262
D.2	Comparison of in-situ (solid lines) and original AMSR-E (red diamonds) near-surface soil moisture (m^3m^{-3}) for each retrieval algorithm, at Canberra Airport (M2) over 2006.	263

D.3	Comparison of in-situ (solid lines) and original AMSR-E (red diamonds) near-surface soil moisture (m^3m^{-3}) for each retrieval algorithm, at West Wyalong Airfield (M4) over 2006.	264
D.4	Comparison of in-situ (solid lines) and original AMSR-E (red diamonds) near-surface soil moisture (m^3m^{-3}) for each retrieval algorithm, at Balranald (M5) over 2006.	265
D.5	Comparison of in-situ (solid lines) and original AMSR-E (red diamonds) near-surface soil moisture (m^3m^{-3}) for each retrieval algorithm, at Hay (M6) over 2006.	266
D.6	Comparison of in-situ (solid lines) and original AMSR-E (red diamonds) near-surface soil moisture (m^3m^{-3}) for each retrieval algorithm, at Griffith Aerodrome (M7) over 2006.	267
D.7	Comparison of in-situ (solid lines) and original AMSR-E (red diamonds) near-surface soil moisture (m^3m^{-3}) for each retrieval algorithm, at Yanco (M8) over 2006.	268
D.8	Comparison of in-situ (solid lines) and original AMSR-E (red diamonds) near-surface soil moisture (m^3m^{-3}) for each retrieval algorithm, at Adelong (M10) over 2006.	269
D.9	Comparison of in-situ (solid lines) and original AMSR-E (red diamonds) near-surface soil moisture (m^3m^{-3}) for each retrieval algorithm, at Merriwa (G1) over 2006.	270
E.1	Comparison of in-situ (solid lines) and normalised AMSR-E (red diamonds) near-surface soil moisture (m^3m^{-3}) for each retrieval algorithm, at Cooma Airfield (M1) over 2006.	272
E.2	Comparison of in-situ (solid lines) and normalised AMSR-E (red diamonds) near-surface soil moisture (m^3m^{-3}) for each retrieval algorithm, at Canberra Airport (M2).	273
E.3	Comparison of in-situ (solid lines) and normalised AMSR-E (red diamonds) near-surface soil moisture (m^3m^{-3}) for each retrieval algorithm, at West Wyalong Airfield (M4) over 2006.	274
E.4	Comparison of in-situ (solid lines) and normalised AMSR-E (red diamonds) near-surface soil moisture (m^3m^{-3}) for each retrieval algorithm, at Balranald (M5) over 2006.	275

E.5	Comparison of in-situ (solid lines) and normalised AMSR-E (red diamonds) near-surface soil moisture (m^3m^{-3}) for each retrieval algorithm, at Hay (M6) over 2006.	276
E.6	Comparison of in-situ (solid lines) and normalised AMSR-E (red diamonds) near-surface soil moisture (m^3m^{-3}) for each retrieval algorithm, at Griffith Aerodrome (M7) over 2006.	277
E.7	Comparison of in-situ (solid lines) and normalised AMSR-E (red diamonds) near-surface soil moisture (m^3m^{-3}) for each retrieval algorithm, at Yanco (M8) over 2006.	278
E.8	Comparison of in-situ (solid lines) and normalised AMSR-E (red diamonds) near-surface soil moisture (m^3m^{-3}) for each retrieval algorithm, at Adelong (M10) over 2006.	279
E.9	Comparison of in-situ (solid lines) and normalised AMSR-E (red diamonds) near-surface soil moisture (m^3m^{-3}) for each retrieval algorithm, at Merriwa (G1) over 2006.	280
F.1	Time series of in-situ (black) and original AMSR-E (red) near-surface soil moisture (m^3m^{-3}) from the VUA-NASA retrieval algorithm at Cooma Airfield (M1), from 2003-2008.	282
F.2	Time series of in-situ (black) and original AMSR-E (red) near-surface soil moisture (m^3m^{-3}) from the VUA-NASA retrieval algorithm at Canberra Airport (M2), from 2003-2008.	283
F.3	Time series of in-situ (black) and original AMSR-E (red) near-surface soil moisture (m^3m^{-3}) from the VUA-NASA retrieval algorithm at West Wyalong Airfield (M4), from 2003-2008.	284
F.4	Time series of in-situ (black) and original AMSR-E (red) near-surface soil moisture (m^3m^{-3}) from the VUA-NASA retrieval algorithm at Balranald (M5), from 2003-2008.	285
F.5	Time series of in-situ (black) and original AMSR-E (red) near-surface soil moisture (m^3m^{-3}) from the VUA-NASA retrieval algorithm at Hay (M6), from 2003-2008.	286
F.6	Time series of in-situ (black) and original AMSR-E (red) near-surface soil moisture (m^3m^{-3}) from the VUA-NASA retrieval algorithm at Griffith Aerodrome (M7), from 2003-2008.	287

F.7	Time series of in-situ (black) and original AMSR-E (red) near-surface soil moisture (m^3m^{-3}) from the VUA-NASA retrieval algorithm at Yanco (M8), from 2003-2008.	288
F.8	Time series of in-situ (black) and original AMSR-E (red) near-surface soil moisture (m^3m^{-3}) from the VUA-NASA retrieval algorithm at Adelong (M10), from 2003-2008.	289
G.1	Time series of in-situ (black) and normalised AMSR-E (red) near-surface soil moisture (m^3m^{-3}) from the VUA-NASA retrieval algorithm at Cooma Airfield (M1), from 2003-2008.	292
G.2	Time series of in-situ (black) and normalised AMSR-E (red) near-surface soil moisture (m^3m^{-3}) from the VUA-NASA retrieval algorithm at Canberra Airport (M2), from 2003-2008.	293
G.3	Time series of in-situ (black) and normalised AMSR-E (red) near-surface soil moisture (m^3m^{-3}) from the VUA-NASA retrieval algorithm at West Wyalong Airfield (M4), from 2003-2008.	294
G.4	Time series of in-situ (black) and normalised AMSR-E (red) near-surface soil moisture (m^3m^{-3}) from the VUA-NASA retrieval algorithm at Balranald (M5), from 2003-2008.	295
G.5	Time series of in-situ (black) and normalised AMSR-E (red) near-surface soil moisture (m^3m^{-3}) from the VUA-NASA retrieval algorithm at Hay (M6), from 2003-2008.	296
G.6	Time series of in-situ (black) and normalised AMSR-E (red) near-surface soil moisture (m^3m^{-3}) from the VUA-NASA retrieval algorithm at Griffith Aerodrome (M7), from 2003-2008.	297
G.7	Time series of in-situ (black) and normalised AMSR-E (red) near-surface soil moisture (m^3m^{-3}) from the VUA-NASA retrieval algorithm at Yanco (M8), from 2003-2008.	298
G.8	Time series of in-situ (black) and normalised AMSR-E (red) near-surface soil moisture (m^3m^{-3}) from the VUA-NASA retrieval algorithm at Adelong (M10), from 2003-2008.	299

List of Tables

2.1	Summary of results from published evaluations of soil moisture retrieved from passive microwave brightness temperatures. Unless otherwise-stated, AMSR-E brightness temperature observations were used in each study. * indicates that remotely sensed data have been normalised to match the in situ soil moisture climatology.	18
4.1	Murrumbidgee Monitoring Network: First generation stations (installed in September 2001).	57
4.2	Murrumbidgee Monitoring Network: Second generation stations (installed in September 2003, with shallow soil moisture sensors installed in late 2006).	58
4.3	Goulburn Monitoring Network (installed in September 2002, with shallow soil moisture sensors installed in late 2005, data not available after 2007). Only stations used here are listed (there are another 13 stations within the network).	60
4.4	RMSD, correlation, and anomaly correlation between layer one soil moisture from the individual monitoring stations in the Kyeamba-A pixel, for April '08 to April '09. For each statistic the minimum and maximum off-diagonal values are in bold.	65
4.5	Descriptive statistics for the soil moisture time series retrieved from descending AMSR-E overpass brightness temperatures by each retrieval algorithm and the in situ data, for 2006. Thee bias values are calculated using only those days for which the AMSR-E data are available for that algorithm.	68

4.6	Statistics of fit between the soil moisture time series retrieved from descending AMSR-E overpass brightness temperatures by each retrieval algorithm, and the in situ data for 2006. The RMSD is provided for both the original and normalised AMSR-E time series. Bold correlations indicates significance at the 1% level.	76
4.7	Descriptive statistics for the VUA-NASA AMSR-E soil moisture and from the in situ data, for 2003-2008. The bias values are calculated using only those days for which the AMSR-E data are available.	79
4.8	Statistics of fit between the VUA-NASA AMSR-E soil moisture time series and the in situ data for 2003-2008. The RMSD is provided for both the original and normalised AMSR-E time series. Bold correlations indicates significance at the 1% level.	84
4.9	Statistics of fit between the filtered VUA-NASA AMSR-E soil moisture time series and the in situ data for 2006. The RMSD is provided for both the original and normalised AMSR-E time series. Bold correlations indicates significance at the 1% level.	92
4.10	Statistics of fit to the in situ observations, from the filtered C-band descending overpass VUA-NASA AMSR-E data, and the ACCESS forecast near-surface soil moisture.	96
5.1	Summary of the observation, model, and initial background error variances used in the assimilation experiments. All diagonal error covariances were assumed to be zero. Units are indicated in brackets.	105
5.2	Statistics of the 6-hour Jacobian terms (m^3m^{-3} / m^3m^{-3}) from 06:00 to 12:00 UTC on 1 July 2006, estimated using a perturbation size of $+10^{-4} \times (w_{fc} - w_{wilt})$ (pos), $-10^{-4} \times (w_{fc} - w_{wilt})$ (neg), and $+10^{-1} \times (w_{fc} - w_{wilt})$ (lrg).	111
5.3	Spatial mean \mathbf{M} (m^3m^{-3} / m^3m^{-3}) for a 6 hour forecast during the night (upper), and day(centre), and for the 24 hour period \mathbf{M} (lower) form 18:00 UTC on 1 July 2006.	113
5.4	Mean observation operator relating w_2 to each observation type from experiment CMB, averaged over July 2006 separately for each of the four daily assimilation cycles.	115

5.5	RMSD between the observations and model forecasts over July 2006 for each assimilation experiment. For the screen-level variables the RMSD of the daily averaged forecasts and observations is shown in brackets.	134
5.6	The relative information content as a fraction of the daily total (for experiment CMB with w_2 updated only), averaged over July 2006 separately for each of the four daily assimilation cycles.	137
5.7	First, second, and third quartile of the change in total soil moisture storage (mm) from 1 to 31 July 2006, for each experiment.	138
6.1	Comparison of the statistics for the MOSES Jacobian terms (%/%), for the 6 hour forecast from 09:00 UTC on 6 April 2008, estimated using a perturbation of +0.1% (+ve) and -0.1% (-ve) of the soil moisture at saturation.	152
6.2	MOSES model Jacobian terms in units of %/%, spatially averaged across Australia, for the 6 hour forecast from 09:00 UTC on 1 August 2008. Values below 1×10^{-5} are reported as 0.0.	155
6.3	MOSES model Jacobian terms in units of mm mm^{-1} , spatially averaged across Australia, for the 6 hour forecast from 09:00 UTC on 1 August 2008. Values below 1×10^{-5} are reported as 0.0.	155
6.4	The (spatially uniform) observation, model, and initial background error variances used in the assimilation of AMSR-E near-surface soil moisture, in units of $(\% \text{ of saturation})^{-2}$. All diagonal error covariances were assumed to be zero.	160
6.5	Summary of the root mean square of the background soil moisture errors ($m^3 m^{-3}$) across the Murrumbidgee Monitoring Network sites from April 2008 to March 2009. The errors are estimated from the in situ soil moisture observations (observed), and from the EKF background error covariance matrices (EKF), and are reported relative to the MOSES soil moisture climatology.	161

6.6	Summary of the root mean square of the AMSR-E observation errors (m^3m^{-3}) across the Murrumbidgee Monitoring Network sites from April 2008 to March 2009. The errors are estimated from the in situ soil moisture observations (observed), and from the EKF observations error covariance matrices (EKF), and are reported relative to the MOSES soil moisture climatology. . . .	161
6.7	Summary of the correlation between the background model soil moisture errors in layers 1, 2, and 3 across the Murrumbidgee Monitoring Network sites from April 2008 to March 2009. The correlations are estimated from the in situ soil moisture observations (observed) and the EKF background error covariance matrices (EKF).	162
6.8	Initial conditions and assimilated observations for the synthetic assimilation experiments.	164
6.9	Mean soil moisture difference from OPN_ACCESS on 1 July 2008 (% of saturation) for each synthetic experiment, averaged over each climate zone.	175
6.10	Mean near-surface soil moisture (mm), root-zone soil moisture (mm), and daily evapotranspiration ($mm\ day^{-1}$), averaged over April 2008 to March 2009 for each climate zone, from the open-loop and the EKF.	193
6.11	Statistics of fit between the near-surface soil moisture in situ observations from the Murrumbidgee Monitoring Network, and the assimilated AMSR-E data, the open-loop, and the EKF. The RMSD is calculated from model data normalised to match the mean and variance on the in situ data. For each statistic the best result is indicated in bold, and all correlations are significant at 1%.	202
6.12	Statistics of fit between the root-zone soil moisture in situ observations from the Murrumbidgee Monitoring Network, and the open-loop and EKF. The RMSD is calculated from model data normalised to match the mean and variance on the in situ data. For each statistic the best result is indicated in bold, and all correlations are significant at 1%.	202
6.13	Statistics of fit between the MOSES_PRECIP root-zone soil moisture, and each of the open-loop and EKF simulations.	204

C.1	RMSD, correlation, and anomaly correlation between layer 2 soil moisture from the individual monitoring stations in the Kyeamba-A pixel, for April '08 to April '09. For each statistic the minimum and maximum off-diagonal values are in bold.	257
C.2	RMSD, correlation, and anomaly correlation between layer 3 soil moisture from the individual monitoring stations in the Kyeamba-A pixel, for April '08 to April '09. For each statistic the minimum and maximum off-diagonal values are in bold.	258
C.3	RMSD, correlation, and anomaly correlation between layer 4 soil moisture from the individual monitoring stations in the Kyeamba-A pixel, for April '08 to April '09. For each statistic the minimum and maximum off-diagonal values are in bold.	258
C.4	RMSD, correlation, and anomaly correlation between the root-zone soil moisture from the individual monitoring stations in the Kyeamba-A pixel, for April '08 to April '09. For each statistic the minimum and maximum off-diagonal values are in bold. . .	259

Table of Symbols

d_i	[m]	Depth of i th soil layer
\mathcal{H}	[-]	Observation operator
\mathbf{H}	[-]	Linearised observation operator
\mathbf{K}	[-]	Kalman gain
\mathcal{M}	[-]	Forecast model
\mathbf{M}	[-]	Linearised forecast model
\mathbf{P}	[-]	Background error covariance matrix
\mathbf{Q}	[-]	Model forecast error covariance matrix
\mathbf{R}	[-]	Observation error covariance matrix
RH_{2m}	[%/%]	Screen-level relative humidity
r	[-]	Correlation (Pearson)
r_{abs}	[-]	Absolute correlation
r_{anom}	[-]	Anomaly correlation
S_i	[% of saturation]	Soil moisture in the i th model layer (MOSES)
S_{RZ}	[% of saturation]	Soil moisture in the root-zone (MOSES)
SWI	[-]	Surface Wetness Index (ISBA)
T_i	[K]	Soil temperature in the i th model layer
T_{2m}	[K]	Screen-level temperature
w_i	[m^3m^{-3}]	Soil moisture in the i th model layer (ISBA)
w_{fc}	[m^3m^{-3}]	Soil moisture at field capacity (ISBA)
w_{wilt}	[m^3m^{-3}]	Soil moisture at the wilting point (ISBA)
\mathbf{x}	[-]	Model state vector
\mathbf{y}	[-]	Observation vector

Table of Abbreviations and Acronyms

ACCESS	Australian Community Climate and Earth-System Simulator
ALADIN	Aire Limitée Adaptation Dynamique développement InterNational
AMSR-E	Advanced Microwave Scanning Radiometer - Earth Observing System (EOS)
ASCAT	Advanced SCATterometer
BoM	Bureau of Meteorology
CDF	Cumulative Distribution Frequency
CMC	Canadian Meteorological Centre
DWD	Deutscher WetterDienst
ECMWF	European Centre for Medium-range Weather Forecasting
EKF	Extended Kalman Filter
EnKF	Ensemble Kalman Filter
ERS	European Remote Sensing satellites
ESA	European Space Agency
FIFE	First International Satellite Land Surface Climatology Project (ISLSCP) Field Experiment
GEM	Global Environmental Multiscale
GLDAS	Global Land Data Assimilation System
HIRLAM	High Resolution Limited Area Model
ISBA	Interactions between Surface, Biosphere, and Atmosphere
JAXA	Japanese Aerospace Exploration Agency
LAPS	Limited Area Prediction System
MOSES	Met Office Surface Exchange Scheme
NASA	National Aeronautics and Space Administration
NCEP	National Centers for Environmental Protection

NDVI	Normalised Difference Vegetation Index
NLDAS	North American Land Data Assimilation System
NSIDC	National Snow and Ice Data Center
NWP	Numerical Weather Prediction
OI	Optimal Interpolation
RFI	Radio Frequency Interference
RMSD	Root Mean Square Difference
SEKF	Simplified Extended Kalman Filter
SAFRAN	Système d'Analyse Fournissant des Resneignements Atmosphériques à la Neige
SGP	Southern Great Plains experiment
SIM	SAFRAN ISBA Modcou
SMAP	Soil Moisture Active Passive
SMEX	Soil Moisture EXperiment
SMMR	Scanning Multichannel Microwave Radiometer
SMOS	Soil Moisture and Ocean Salinity
SURFEX	SURFace EXternalized
UKMO	UK Met Office
USDA	United States Department of Agriculture
VUA	Vrije Universiteit Amsterdam
2-D Var	Two-dimensional Variational assimilation

Chapter 1

Introduction

This thesis seeks to establish whether Numerical Weather Prediction (NWP) models might benefit from the assimilation of remotely sensed near-surface soil moisture data. It is motivated by the potential to improve the model root-zone soil moisture, with the expectation that this will eventually lead to improved low-level atmospheric forecasts.

It is first established that realistic near-surface soil moisture fields can be extracted over Australia from currently available remotely sensed data sets. Then remotely sensed near-surface soil moisture observations are assimilated into the land surface scheme of the NWP models used by Météo-France and the Australian Bureau of Meteorology (BoM). With Météo-France's model over Europe, the assimilation of near-surface soil moisture is contrasted with the assimilation of the low-level atmospheric temperature and relative humidity observations that are currently used to constrain soil moisture in most NWP models. It is then demonstrated with the BoM model over Australia that assimilating remotely sensed near-surface moisture observations can improve the modeled near-surface and root-zone soil moisture.

1.1 Problem statement

Soil moisture can have a strong influence on atmospheric forecasts, both at short (Baker et al., 2001; Drusch and Viterbo, 2007) and medium range (Zhang and Frederiksen, 2003; Fischer et al., 2007) time scales. If not suitably constrained, the soil moisture in an NWP model will drift from the true climate, resulting in erroneous boundary layer forecasts (Drusch and Viterbo, 2007). This is prevented in many NWP models, including those of the BoM (Best and Maisey,

2002) and Météo-France (Giard and Bazile, 2000), by constraining the model soil moisture according to errors in short-range low-level atmospheric forecasts, by assimilating observations of relative humidity and temperature taken at screen-level (1.5 - 2.0 m above the surface). The assimilation of screen-level observations has been shown to effectively decrease errors in boundary layer forecasts (Douville et al., 2000; Hess, 2001; Drusch and Viterbo, 2007), however from the inception of this approach it was recognised that the improvements were not necessarily due to improved model soil moisture (Douville et al., 2000; Hess, 2001). In fact, the assimilation of screen-level observations has been found to often degrade the modeled soil moisture, since it is often ‘corrected’ to compensate for screen-level errors that were not caused by the soil moisture itself (for example due to inaccuracies in the land surface flux parameterisations or the radiation physics) (Douville et al., 2000; Hess, 2001; Drusch and Viterbo, 2007). Ultimately, model soil moisture cannot be adjusted to exactly compensate for inaccurate model physics: a model with inaccurate soil moisture and inaccurate physics cannot accurately model the atmosphere across the full range of forecast lengths produced from NWP models. The use of screen-level observations is further limited by the strength of the local land surface - boundary layer coupling, and cannot be applied during periods of weak coupling, such as strong advection or weak radiative forcing. Its effectiveness is also limited by the availability of screen-level observations, which are particularly sparse across much of the Southern Hemisphere, including Australia.

In addition to the above, the lack of realism of the soil moisture in NWP models prevents its use in other applications. If available, realistic operationally supported soil moisture fields from NWP models would be extremely valuable to a range of applications. For example, accurate daily soil moisture information would be useful for hydrological prediction, including rainfall-runoff modelling (Merz and Blöschl, 2009) and flood forecasting (Komma et al., 2008). Knowledge of soil moisture conditions would also benefit environmental applications, such as monitoring and predicting drought (Sheffield et al., 2004; Todisco et al., 2009), and managing soil erosion (Fitzjohn et al., 1998). With appropriate downscaling, NWP-derived soil moisture might also be useful for targeted agricultural applications, such as crop yield prediction (Wagner et al., 2000) and irrigation scheduling (Campbell and Campbell, 1982).

A promising solution to improving the accuracy of the soil moisture in NWP models is to make use of novel remotely sensed observations of near-surface soil

moisture. Remote sensors observe the moisture across a thin surface layer (up to 5 cm depth, depending on the observation wavelength and the soil conditions), with close to global coverage. This thin surface layer directly interacts with the underlying soil moisture profile through diffusion processes, potentially offering a more direct relationship with the root-zone soil moisture (of greatest interest to NWP and most other applications) than the screen-level variables (Calvet and Noilhan, 2000). A number of studies have demonstrated improvements in model root-zone soil moisture from the assimilation of near-surface soil moisture information into general land surface models (Walker and Houser, 2001; Crow and Wood, 2003; Reichle et al., 2007) and NWP models (Seuffert et al., 2004; Drusch, 2007). Recent interest in using remotely sensed soil moisture in NWP is largely motivated by the first two purpose-designed soil moisture remote sensing missions: the European Space Agency's (ESA) Soil Moisture and Ocean Salinity mission (SMOS; Kerr et al., 2001), launched in November, 2009; and the National Aeronautics and Space Administration's (NASA) Soil Moisture Active Passive (SMAP; Entekhabi et al., 2004) mission, currently scheduled for launch in 2015. However, in addition to SMOS and SMAP, soil moisture information can already be retrieved from current microwave remote sensing missions (Wagner et al., 1999; Owe et al., 2001).

Since the assimilation of screen-level observations reduces low-level atmospheric forecast errors in part by adjusting the model soil moisture to compensate for model errors, it is not straight-forward that improving the realism of the model soil moisture will immediately improve atmospheric forecast skill. However, it is foreseen that improving model soil moisture will lead to improved atmospheric forecasts, since incorporating new observation types will also expose model deficiencies, thus identifying areas for future model improvements (Seuffert et al., 2004). Until these improvements occur, near-surface soil moisture observations may be best used in NWP together with screen-level observations, to target both the model soil moisture and low-level atmospheric forecasts. However, to date satellite derived near-surface soil moisture data and screen-level observations have not been assimilated together at large scales (i.e., beyond individual field sites), and the manner in which these two data types will interact in an assimilation is not well understood.

1.2 Research objectives and outline

The aim of this thesis is to determine whether current generation NWP modelling over Australia might benefit from the assimilation of remotely sensed near-surface soil moisture data. This includes identifying significant obstacles and/or shortcomings related to the use of this data. While the ultimate aim of assimilating remotely sensed soil moisture into NWP models is to improve land surface flux forecasts, the focus here was on the intermediate step of improving the model soil moisture. Additionally, realistic operationally supported soil moisture forecasts from NWP centres will have intrinsic value for applications beyond NWP (as identified in the previous section). This aim has been achieved through a series of related objectives:

- assessment of the realism of current remotely sensed soil moisture data sets over Australia
- demonstration of a soil moisture analysis system capable of assimilating both near-surface soil moisture and screen-level atmospheric observations, within the limitations of operational NWP
- identification of the physical processes by which root-zone soil moisture is constrained from near-surface soil moisture observations within this assimilation system
- assessment of the impact of assimilating remotely sensed near-surface soil moisture into an NWP land surface model, in terms of the model skill in forecasting near-surface and root-zone soil moisture
- comparison of the assimilation of near-surface soil moisture data to the assimilation of screen-level temperature and relative humidity observations, and investigation of the interaction between the two observation types when they are assimilated together

The above objectives were addressed through several stages of research. In the first stage the quality of the available remotely sensed soil moisture data sets was examined over Australia to select the best data set for assimilation, and to characterise its accuracy. At the time of this study only the passive microwave Advanced Microwave Scanning Radiometer - Earth Observing System (AMSR-E) sensor was providing observations with sufficient coverage over Australia,

hence this work focused on comparing soil moisture observations derived from AMSR-E brightness temperatures with different retrieval algorithms.

The selected AMSR-E near-surface soil moisture data set has been assimilated with an Extended Kalman Filter (EKF). This EKF is based on the simplified EKF that Mahfouf et al. (2009) developed to assimilate screen-level observations into Météo-France’s NWP land surface scheme. Following Balsamo et al. (2007) and Mahfouf et al. (2009), this land surface assimilation system is designed to be semicoupled to an NWP model, in that each assimilation is performed using an off-line version of the NWP land surface model, forced with atmospheric forecasts produced from the NWP model updated with the previous land surface analysis update. However for this thesis, the land surface assimilation has been completely decoupled from the atmospheric model, and the land surface model is forced by NWP forecasts generated once at the start of each experiment. While this simplification neglects any feedback between the soil moisture updates and subsequent atmospheric forecasts, it greatly reduces the computational cost and complexity of the experiments, while still allowing the first-order surface dynamics to be examined.

Two separate sets of assimilation experiments have been carried out, one over the Australian domain using the BoM’s NWP model and one over the European domain using Météo-France’s NWP model. For the Australian experiments, the EKF has been used to assimilate AMSR-E near-surface soil moisture into the land surface scheme of the BoM’s NWP model, the Australian Community Climate and Earth-System Simulator (ACCESS). ACCESS is a recently imported version of the UK Met Office’s NWP suite, and it was launched as the Australian operational NWP model in September 2009. The manner in which the assimilation translates the near-surface soil moisture observations into updates throughout the soil moisture profile in ACCESS has been examined, and the impact of the assimilation has been tested by assessing the analysed soil moisture fields against in situ soil moisture observations. Screen-level temperature and relative humidity observations suitable for use with the ACCESS model were not available for use in these experiments.¹ Consequently, experiments assimilating screen-level observations were instead conducted over the

¹The BoM archives screen-level observations in the form of a screen-level analysis, generated with the operational NWP model. Since ACCESS only recently became operational, the archived analyses were generated with the BoM’s previous NWP model, and will be strongly influenced by the biases in that model, and hence are not suitable for assimilation into ACCESS.

European domain with Météo-France’s NWP system, Aire Limitée Adaptation Dynamique développement InterNational (ALADIN). The use of two different land surface models, which have very different soil moisture physics, also allows additional insight into the role of the model physics in the assimilation process. Using Météo-France’s NWP system, the assimilation of AMSR-E near-surface soil moisture and screen-level observations has been compared over a one month period. The two different observations types were first assimilated separately, to determine how each is translated into root-zone soil moisture updates, as well as the level of agreement between these updates. They were then assimilated together to examine the interaction between the two data sets.

1.3 Outline of thesis

The research conducted towards the above-stated aim and objectives is presented in the following six chapters. In Chapter 2 the main literature supporting this thesis is reviewed, and based on this review the research methods used in this thesis are then formulated. An EKF-based land surface analysis was identified as the most suitable for use in NWP, and this land surface analysis is described in detail in Chapter 3, together with the land surface models from Météo-France and the BoM that have been used in the assimilation experiments. In Chapter 4 the available remotely sensed soil moisture data sets over Australia are assessed, and the most suitable product is selected for use in the assimilation experiments. The next two chapters then present the results of experiments assimilating this data. Chapter 5 introduces the assimilation of near-surface soil moisture data into Météo-France’s NWP land surface scheme, and this is then compared to the assimilation of screen-level observations. Having demonstrated how near-surface soil moisture observations can be combined with the preexisting assimilation of screen-level observations, Chapter 6 presents the assimilation of the AMSR-E near-surface soil moisture observations into the Australian BoM’s NWP suite, and the impact of the assimilation on the model soil moisture is assessed against available in situ soil moisture data. Finally, Chapter 7 presents the main findings and recommendations arising from the previous chapters.

Chapter 2

Literature Review and Research Outline

2.1 Overview of chapter

This chapter presents an overview of the literature relevant to this thesis. It begins with a review of the manner in which soil moisture influences atmospheric forecasts, followed by the main approaches that are currently used to initialise soil moisture for NWP forecasts. The relative advantages and disadvantages of the current assimilation techniques are identified, and the work done to date to assimilate remotely sensed land surface observations into NWP models is reviewed. The available remotely sensed soil moisture data sets are then identified, and briefly reviewed within the context of their potential application to NWP. The main research findings regarding the assimilation of remotely sensed soil moisture into generalised land surface models (rather than NWP models) are then surveyed. Based on the findings from this literature review, the methods used here to address the aims and objectives listed on page 4 of the Introduction are then described.

2.2 Soil moisture as an atmospheric control

The main process through which soil moisture influences the atmosphere is by limiting the amount of water available for evapotranspiration, and hence determining the partition of surface radiation into latent and sensible heat fluxes. This influence has been observed over adjacent wet and dry regions, with wet-

ter (greater latent heating), and shallower atmospheric boundary layers with reduced diurnal temperature range (reduced sensible heating) occurring over the wetter soils (e.g., Betts and Ball, 1995; Pielke, 2001).¹ Since the surface flux partition contributes to the temperature and humidity profile of the atmosphere, soil moisture can then have a profound impact the stability of the atmospheric boundary layer. While it is difficult to directly observe these effects, sensitivity studies with numerical models have shown that changes in the flux partition induced by changing the model soil moisture can have a profound impact on short-term atmospheric forecasts, particularly of precipitation. For example, studies of individual precipitation events have demonstrated that the model soil moisture can influence the initiation and intensity of precipitation associated with convection (Gallus and Segal, 2000; Findell and Eltahir, 2003), sea-breezes (Baker et al., 2001), and dry-lines (Grasso, 2000; Martin and Xue, 2006). In most of the above studies increased latent heating associated with increased soil moisture lead to increased precipitation, although under certain conditions the increased sensible heating associated with decreased soil moisture can also induce precipitation (Gallus and Segal, 2000; Findell and Eltahir, 2003). However, the influence of soil moisture is dependent on the atmospheric conditions, and under some conditions (strong large-scale atmospheric controls, or either very stable or unstable conditions) short range forecasts are insensitive to soil moisture conditions (Findell and Eltahir, 2003; Martin and Xue, 2006).

The influence of soil moisture on the atmosphere can also extend beyond individual events. Compared to atmospheric time scales, soil moisture has a relatively long memory, particularly where it is enhanced by a positive feedback loop (typically between evapotranspiration and precipitation), and soil moisture anomalies can persist over seasons or even years. Hence an initial soil moisture anomaly can have a significant impact on monthly to seasonal-scale forecasts. For example, the anomalously low soil moisture over Europe in spring 2003 contributed to the heat wave conditions the following summer (Fischer et al., 2007). Likewise, studies over the Mississippi River Basin have suggested while the anomalous (low) high precipitation in July (1998) 1993 was principally caused by anomalies in the large atmospheric circulation, these anomalies were enhanced by the antecedent soil moisture conditions (Trenberth and Guillemot,

¹Note that this occurs only in moisture limited conditions, so that above a certain surface wetness Betts and Ball (1995) observed that the atmosphere was no longer sensitive to the surface state.

1996; Seth and Giorgi, 1998). Additionally, in a statistical analysis of ensemble forecasts with different surface boundary conditions (Conil et al., 2007) concluded that in the extra-tropics during summer soil moisture contributes more to seasonal predictability than sea surface temperature does.

Since soil moisture influences the atmosphere through limiting evapotranspiration, it has the greatest influence on the atmosphere during warm conditions when evapotranspiration is most active. All of the results cited above regarding the influence of soil moisture on the atmosphere were related to warm season (summer and the adjacent months) events. During winter the importance of soil moisture is reduced due to lesser radiative forcing, and in many regions the presence of frozen cover creates a barrier between the soil moisture storage and the atmosphere. Statistics from multiple atmospheric models indicate that (at the seasonal scale) the strongest atmosphere - land surface coupling, and hence greatest sensitivity to soil moisture occurs in the transition zones between wet and dry climates during summer: in these regions the surface is sufficiently dry that the evapotranspiration is moisture limited, and there is sufficient humidity that precipitation can occur (Koster et al., 2004).

2.3 Initialisation of soil moisture in NWP

Soil moisture is a sink variable so that errors in an NWP model, for example in forecast precipitation, will accumulate in the model soil moisture, causing it to drift away from the true climate (e.g., Viterbo, 1995). The two main approaches that are currently used to prevent this drift are i) to initialise the land surface variables (usually soil temperature and moisture) from a stand-alone land surface model forced with observations, or ii) to constrain the surface states according to errors in the forecasts of low-level atmospheric variables. These two approaches are reviewed below, focusing on the initialisation of soil moisture (rather than temperature), and the main advantages and disadvantages of each are highlighted.

2.3.1 Current initialisation methods: stand-alone land surface models

The land surface states in an NWP model can be constrained by regularly updating it with fields from a stand-alone land surface model forced with ob-

servations. This approach is used by the Met Office for their regional model over a limited domain centred on the UK (Smith et al., 2006). The National Centers for Environmental Protection (NCEP) operate a similar system, called the North American Land Data Assimilation System (NLDAS; Mitchell et al., 2004), which is intended to supply initial conditions to their NWP model. In both of the above examples the stand-alone land surface model is forced with a mixture of ground-based observations (including precipitation), space-based observations (including radiation), and NWP model output (including low-level atmospheric temperature, humidity, and wind).

This approach prevents NWP forecast errors from accumulating in the surface state variables, removing one of the main error-sources in the NWP surface. At the same time it generates surface fields that are consistent with biases in the model's (imperfect) physics. However, for NWP applications accurate atmospheric forecasts are the first priority, and this approach cannot correct for errors in a model's land surface flux forecasts, since there is no feedback from the land surface to the atmosphere. Additionally, the effectiveness of this approach is limited by the availability of high quality observations. It is best suited to limited-domain models over regions with high-density observation networks, and so is not well suited to Australia (which has an unusually sparse observation network - see Section 2.3.3). While remote sensors can provide global forcing data (as is done for precipitation in the Global Land Data Assimilation System (GLDAS)), these data sets may not offer significantly greater accuracy than NWP forecasts do.

2.3.2 Current initialisation methods: assimilation of screen-level observations

A more common approach to constrain the land surface is to adjust the surface state variables at the beginning of each forecast by relating the errors in short-range forecasts of screen-level (1.5 - 2.0 m) variables (usually temperature and humidity) to errors in the surface states. Several different assimilation methods are used to achieve this, the most prominent of which are reviewed below.

The OI and related nudging schemes

Several centres use an Optimal Interpolation (OI) scheme developed at the European Centre for Medium Range Weather Forecasting (ECMWF) and Météo-

France (Mahfouf, 1991; Douville et al., 2000; Giard and Bazile, 2000; Mahfouf, 2000). At both ECMWF and Météo-France the OI updates model surface moisture every 6 hours based on errors in the forecast screen-level temperature and relative humidity. The OI coefficients relating the screen-level errors to soil moisture errors were derived from an ensemble of 100 single column simulations initialised with randomly perturbed soil moisture (Douville et al., 2000). The estimated coefficients were generated for clear-sky conditions with maximum insolation, and are based on the assumption that the screen-level errors are due only to soil moisture errors. Empirical reduction factors are then applied to the coefficients to account for periods of reduced surface-atmosphere interaction (such as strong advection, precipitation, frozen soils, snow cover, cloud cover, and nighttime). At Météo-France and ECMWF the soil temperature is also analysed based on screen-level temperature errors, using the simple linear relationship of Coiffier et al. (1987). The Canadian Meteorological Centre (CMC; Bélair et al., 2003) and the High Resolution Limited Area Model (HIRLAM; Rodriguez et al., 2003) consortium of European nations both use Météo-France’s land surface model, and hence the OI surface analysis scheme.

In addition to the ECMWF / Météo-France OI, several other NWP centres use variations of this approach, including the Australian BoM. As noted above, the Met Office initialise the land surface in their regional model using an off-line model forced with observations. However the necessary high-quality forcing data for this approach are not available globally, and the land surface variables in both the Australian ACCESS model and the Met Office global model are initialised according to errors in the screen-level forecasts. In place of the empirical OI coefficients the Met Office has developed a physically-based soil moisture nudging scheme, by estimating a linear relationship between screen-level errors and soil moisture corrections from the model physics (Best and Maisey, 2002). Additionally, in the BoM’s previous NWP model, the surface was initialised using an older ECMWF scheme, in which soil moisture was adjusted, or “nudged”, twice daily according to errors in the low level specific humidity forecasts (Pescod, 1994).

The simplified variational analysis

The German weather service (Deutscher WetterDienst; DWD) use a more sophisticated method to relate screen-level forecast errors to soil moisture updates. They update their model soil moisture once daily from daytime screen-level

temperature observations, using a simplified variational analysis, with Kalman filter-like cycling of the model background error matrix (Hess, 2001; Hess et al., 2008). As will be discussed in Section 2.5.2, to be computationally efficient variational assimilation requires model adjoints, which have not been developed for most land surface models. Other centres have previously tested a variational surface analysis from screen-level observations using a brute-force iterative minimisation of the cost function, but this option is too computationally expensive to implement within a full 3-D atmospheric model (Bouyssel et al., 1999). Instead, Hess (2001) avoid the cost of an iterative minimisation by linearising the forecast model over the assimilation window, using a first-order Taylor approximation, enabling a direct solution to minimising the cost function. With this linearisation the method further resembles an EKF (this similarity will be reviewed in Section 3.2 and Appendix A). The Taylor approximation requires an additional (perturbed) forecast for each element of the state vector, which is obtained with a single perturbed model run for each surface state variable, by neglecting the horizontal screen-level - surface feedback, and performing an independent (and one dimensional) analysis at each model grid.

2.3.3 Towards assimilation of land surface observations

From its inception it was recognised that updating model soil moisture from screen-level observations could improve atmospheric forecasts, but would not necessarily lead to realistic soil moisture. Hess (2001) states that this approach “does not attempt to provide accurate and realistic soil moisture, but it rather uses the soil moisture value to some extent to compensate for biases in the physical parameterisations”. In particular, Hess (2001) showed that the surface initialisation reduced atmospheric temperature errors during periods of strong radiative forcing, by compensating for errors in the radiative forcing. However, the initialisation degraded the model soil moisture compared to observations from Lindenberg Meteorological Observatory. Likewise, for the ECMWF model, both Douville et al. (2000) and Drusch and Viterbo (2007) found that the OI reduced errors in the screen-level observations, and prevented the model surface state from drifting towards unrealistic values. However, compared to in situ soil moisture observations from the First International Satellite Land Surface Climatology Project (ISLSCP) Field Experiment (FIFE) (Douville et al., 2000) and the Oklahoma Mesonet (Drusch and Viterbo, 2007), it did not generate

realistic soil moisture.

In addition to the lack of realism of the resulting model soil moisture, the use of screen-level observations is limited by the strength of the local soil moisture - boundary layer coupling, and cannot be applied during periods of weak coupling, such as strong advection or weak radiative forcing. This approach is also limited by the availability of screen-level observations, which are particularly sparse across much of the Southern Hemisphere, including Australia. For example, there are less than 250 observations available for each screen-level analysis over Australia (Mills, 2001), about 10% of the number available over west and central European (Mahfouf et al., 2009) which is about half the size of Australia.

A promising approach to addressing these short-comings is to assimilate observations that are more directly related to the land surface states than screen-level variables, such as remotely sensed observations of the land surface. Assimilation of remotely sensed observations of near-surface soil moisture have received the most attention to date (e.g., Balsamo et al., 2006; Drusch and Viterbo, 2007; Mahfouf et al., 2009), however observations of snow, soil temperature, and vegetation cover are also of interest (Drusch, 2007). To assimilate new data types several NWP centres, including Météo-France (Mahfouf et al., 2009), ECMWF (Drusch et al., 2009), and CMC (Balsamo et al., 2007), are working towards implementing a version of the simplified variational analysis used at DWD, which Balsamo et al. (2004) has dubbed “simplified 2-D Var” (where the dimensions refer to time and the vertical direction). The simplified 2-D Var can more easily assimilate new data types than the OI methods which require new coefficients to be calculated for each new observation type, using either Monte-Carlo experiments (for the ECMWF / Météo-France approach), or analytical means (for the Met Office approach). Additionally, the simplified 2-D Var also offers the advantage of responding more accurately to the synoptic and seasonal conditions at the time of each update (Hess, 2001; Mahfouf et al., 2009), and hence it reflects the model physics more accurately (Drusch et al., 2009). It can also better account for nonlinear relationships between the surface and screen-level atmospheric states (Hess, 2001).

In the form used by DWD (which is applied to their regional model), the computational cost of the additional model integrations required to linearise the model for the simplified 2-D Var is too expensive for operational implementation (e.g., see Drusch et al., 2009). This led Balsamo et al. (2007) to adapt the simplified 2-D Var of Balsamo et al. (2004) to perform the analysis using a

parallel off-line version of land surface model. In this set-up the (coupled) NWP model is updated with the land surface analysis update from the off-line assimilation, and the subsequent NWP forecasts are then used to force the off-line model for the next assimilation cycle. To allow the off-line assimilation of screen-level observations, the off-line model is forced at the level of the first atmospheric layer. This is higher than most land surface models to allow an interactive evolution of the atmospheric surface layer, producing the required feedback between the screen-level atmosphere and the land surface.²

Balsamo et al. (2007) showed that the gain terms and the information content derived from the 2-D Var assimilation of screen-level observations, and various remotely sensed land surface observations, was similar for the off-line system and the fully-coupled Canadian Global Environmental Multiscale (GEM) model. Météo-France have also developed a similar system, referred to by Mahfouf et al. (2009) as a “simplified EKF” (the “simplified EKF” of Mahfouf et al. (2009) and the “simplified 2-D Var” of Balsamo et al. (2007) are equivalent, as will be discussed in Section 3.2 and Appendix A). Mahfouf et al. (2009) showed that for the assimilation of screen-level variables the Kalman gain terms generated with this off-line system better capture the feedback between the land surface and the screen-level forecasts than the OI coefficients used in Météo-France’s current surface analysis system. With the use of an off-line model the simplified variational analysis is computationally affordable in global models, even at high resolution.

2.4 Remote sensing of soil moisture

The most widely accepted method for remote sensing of soil moisture is to use radiances in the microwave wavelengths (1 - 50 cm) (Schmugge and Jackson, 1996; Jackson, 2005; Pietroniro and Leconte, 2005), and to date the only continuously generated global soil moisture data sets with the spatial resolution

²Note the difference between the off-line simplified 2-D Var assimilation of screen-level observations and the use of the stand-alone model forced with observations in Section 2.3.1. For the simplified 2-D Var, the off-line model is used only to save computational cost, and it is intended to mimic the land surface in the coupled model as closely as possible (allowing the assimilation to correct the land surface for errors in the screen-level forecasts). In contrast, the stand-alone models in Section 2.3.1 are forced with observations, to prevent errors in the NWP forecasts (e.g., of precipitation) from accumulating in the land surface. The latter approach cannot account for feedback from the land-surface to the atmosphere, and so cannot correct for errors in the land surface flux forecasts.

required for NWP are derived from passive microwave radiometers and active microwave scatterometers. Each of the available remotely sensed soil moisture data sets is reviewed below, focusing on their potential for application in Australian NWP. Several other approaches to remote sensing soil moisture have also shown promise, including thermal infra-red (McVicar and Van Niel, 2005) and the Gravity Recovery and Climate Experiment (Tapley et al., 2004). However none of these provides global soil moisture data sets at the resolution required for NWP, and they are not considered here. Additionally, a major obstacle to the uptake of remotely sensed soil moisture data is that it is extremely difficult to validate, and the reasons for this are reviewed here.

2.4.1 Microwave remote sensing

Soil moisture is inferred from microwave observations by utilising the contrasting dielectric properties of soil and water. The main advantage of the microwave frequencies for sensing the land surface is that there is relatively little interference to the microwave signal from other factors. In particular, at the longer microwave wavelengths the atmosphere including (nonraining) clouds is transparent, and vegetation is semitransparent. In general, the longer microwave wavelengths are better suited to sensing soil moisture, since the observations relate to a deeper soil layer, there is less interference from vegetation, surface roughness, and the atmosphere, and the signal becomes more sensitive to moisture due to increases in the dielectric constant of water (Jackson, 1993). Microwave data are derived from both passive and active sensors; passive microwave (radiometers) observe natural thermal microwave emissions from the ground, and active microwave (scatterometers) observe the ratio between a pulse of microwave radiation that is emitted by the sensor and the back-scatter pulse received from the target surface. Each of these methods is reviewed below.

Passive microwave missions

The most commonly used passive microwave sensor for observing soil moisture at present is the Advanced Microwave Scanning Radiometer - Earth Observing System (AMSR-E), which has been orbiting on National Aeronautics and Space Administration's (NASA) Aqua satellite since May 2002. AMSR-E observes five dual-polarised frequency bands (centred at 6.92, 10.65, 18.7, 36.5, and 89.0 GHz), and until the recent launch of SMOS (see Section 2.4.2) it was the

lowest frequency radiometer in orbit.³ With the exception of regions of dense vegetation, snow, ice, or frozen soils, it provides global soil moisture coverage every two days, from two overpasses (Njoku et al., 2003): the ascending overpass (equator-crossing at about 1.30pm local time) and the descending overpass (equator-crossing at about 1.30 am). The two lowest frequencies, C-band (6.92 GHz) and X-band (10.65 GHz), are currently used to observe soil moisture. The C-band data have a nominal spatial resolution of 25 km (resampled from overlapping 45 x 75 km² swath data), and are sensitive to soil moisture in the upper-most ~ 1 cm of the surface, with the exact depth depending on the soil moisture content.

The utility of AMSR-E for remote sensing soil moisture has been limited in many regions by Radio Frequency Interference (RFI), principally from surface communication networks. RFI cannot be removed by post-processing (Njoku et al., 2003). Global maps of RFI produced by Njoku et al. (2005) show significant C-band RFI over the USA, Japan, and the middle East, as well as X-band RFI over Japan (Njoku et al., 2005) and some countries in Europe.

Soil moisture retrieval from radiometer observations

A radiative transfer model can relate the brightness temperatures observed by AMSR-E to soil moisture. The radiative transfer equation is under-determined, since the observed brightness temperature for each microwave frequency and polarisation depends on several geophysical parameters. At the scale of an AMSR-E footprint the primary surface geophysical variables influencing the brightness temperature are the volumetric soil moisture, vegetation opacity, and the vegetation and soil temperature. The surface roughness, single scattering albedo, and soil properties are also influential. It is unlikely that the parameters required to obtain soil moisture can be determined with sufficient accuracy from a single-frequency and polarisation retrieval algorithm (Crosson et al., 2005), necessitating the use of multiple frequency and/or polarisation data to minimise the use of ancillary data.

A range of different approaches has been developed to retrieve soil moisture from AMSR-E brightness temperatures, each of which frames the radiative transfer equations, and approaches the under-determination problem, differently (and so makes use of different ancillary data). The most prominent retrieval

³The radiometer on-board the Coriolis/WINDSAT Weather Satellite, launched in 2006, also observes similar frequencies to AMSR-E.

algorithms have been developed at;

- NASA, following Njoku et al. (2003) ;
- the Japanese Aerospace Exploration Agency (JAXA), following Koike et al. (2004);
- the United States Department of Agriculture (USDA), following Jackson (1993); and
- the Vrije Universiteit Amsterdam (VUA) in collaboration with NASA (referred to below as VUA-NASA), following Owe et al. (2001).

The NASA, JAXA, and USDA algorithms are all based on X-band AMSR-E brightness temperatures only, due to C-band RFI over north America and east Asia. The VUA-NASA algorithm is applied separately to C- and X-band AMSR-E data (recall that the longer wavelength C-band is theoretically better suited to observing soil moisture). RFI is associated with densely populated urban areas (Li et al., 2004), and is not expected to be problematic over Australia due to the low population density. In a global RFI survey for June 2002 to May 2003, Njoku et al. (2005) did not note any RFI over Australia (in any wavebands), while over Europe they observed C-band RFI over several urban regions and widespread X-band RFI over Italy and England.

The accuracy of each AMSR-E soil moisture data set is not well characterised, and the relative accuracy of the different retrieval algorithms is unknown. Validation of AMSR-E derived soil moisture has been hampered by the presence of RFI in many regions, the limited availability of ground truth data, and the differences in scale between remotely sensed soil moisture and other estimates (which will be discussed in Section 2.4.3). Table 2.1 lists a selection of studies that have quantitatively evaluated soil moisture derived from AMSR-E. A number of studies obtained better results from the VUA-NASA algorithm than from the NASA algorithm (Wagner et al., 2007; Rüdiger et al., 2009), however the JAXA and USDA soil moisture data sets have not been compared to the other retrievals. Results from different studies cannot be directly compared, due to differences between the evaluation methodologies and characteristics at each monitoring site. Very little work has evaluated AMSR-E soil moisture over Australian sites, and the work that has been done has used the VUA-NASA retrieval algorithm only. In general these results are encouraging:

Liu et al. (2009) have demonstrated that soil moisture and vegetation derived from VUA-NASA AMSR-E (and its passive microwave predecessors) contains a statistical signal of regional climate indices.

Table 2.1: Summary of results from published evaluations of soil moisture retrieved from passive microwave brightness temperatures. Unless otherwise-stated, AMSR-E brightness temperature observations were used in each study. * indicates that remotely sensed data have been normalised to match the in situ soil moisture climatology.

Publication: Reference Data Set	Algorithm	RMSD	Correlation
Wagner et al. (2007): Mean value from comparison to in situ data from 1 location (20 stations) in Spain, and three other remotely sensed soil moisture estimates.	VUA-NASA NASA	0.06* 0.08*	0.62 0.17
Rüdiger et al. (2009): In situ data from 1 location (4 stations) in France Land surface model forced with high quality observations over France	VUA-NASA NASA VUA-NASA NASA	0.06* 0.11* 0.08* 0.11*	0.78 0.11 0.49 -0.01
de Jeu et al. (2008): ERS soil moisture for all global regions with sparse vegetation	VUA-NASA	–	0.83
Gruhler et al. (2008): In situ data from one location (site with best fit) in France In situ data from one location (site with best fit) in Mali	NASA NASA	0.11 0.06	0.17 0.54
Bindlish et al. (2006): Intensive sampling from the Soil Moisture Experiment 2002 in the United States, compared to airborne C-band radiometer observations	USDA	0.03	–
Jackson et al. (2002): Intensive sampling from the Southern Great Plains 1999 Experiment in the United States, compared to airborne C-band radiometer observations	USDA	0.06	–

Active microwave missions

To date the only operationally-supported remotely sensed soil moisture data set is based on active microwave observations from the Advanced SCATterometer (ASCAT) on board ESA's MetOp satellite (Naeimi et al., 2009). ASCAT is a real aperture back-scatter radar at 5.255 GHz (C-band), and it is the lowest frequency scatterometer in orbit. It observes at a spatial resolution of 25 km, which is resampled onto a 12.5 km grid, and as with the passive microwave data, the C-band observations relate to soil moisture in about the uppermost 1 cm of soil. ASCAT was launched in late 2006 to replace the European Remote Sensing Satellites, ERS-1 (operational from 1991 to 1996), and ERS-2 (launched in April 1995), and ASCAT-derived soil moisture observations have been available since 2008. While the ERS satellites also observe C-band radar back-scatter, (at lower spatial resolution: 50 km, resampled to 25 km), the coverage of the ERS observations is rather poor. The ERS orbit enables global coverage approximately every three days, however the coverage of the scatterometer observations is much lower (Wagner et al., 2003), since the scatterometer is not operated continuously (the instruments are switched between a wind/wave mode and an imaging mode). Additionally, due to the loss of storage capacity on ERS-2 in June 2003, data are available only when the satellite has line-of-sight to a ground-based receiving station. For Australia, a ground station in Hobart has been receiving data since 30 November 2005, however this typically consists of just one or two (500 km wide) swaths over southeast Australia every three days (and coverage over Europe is similar). While the ASCAT soil moisture observations were not available in time to be used in this thesis, and the ERS observations do not have sufficient coverage, the soil moisture retrieval method used is briefly reviewed here for completeness.

Soil moisture retrieval from scatterometer observations

In contrast to the conceptual radiative transfer models applied to passive microwave, a semiempirical change detection approach has been developed to relate ASCAT and ERS observation to near-surface soil moisture. This method was developed at the Technische Universität Wien, initially based on eight years of ERS-1/2 data (Wagner et al., 1999), and has been extended to ASCAT (Bartalis et al., 2007). The approach is based on the assumption that over the long data record considered the highest observed reflectivity relates to soil moisture

at saturation, while the lowest represents a completely dry soil, and that a linear relationship can be used to interpolate the values in between. Soil moisture is provided as a surface wetness index (degree of saturation), which can be converted to a volumetric soil moisture if the local wet and dry-end soil moisture values are known.

Comparisons of soil moisture from ERS to in situ data (typically using locally observed upper and lower limits to convert the wetness index into volumetric values) suggest similar accuracy to AMSR-E. For example estimates of the Root Mean Square Difference (RMSD) and correlation (given in brackets) between ERS and ground data include $0.06 \text{ m}^3\text{m}^{-3}$ (0.43) in America (Drusch et al., 2004), $0.02 \text{ m}^3\text{m}^{-3}$ (0.87) in Spain (Ceballos et al., 2005), $0.07 \text{ m}^3\text{m}^{-3}$ (0.54) also in Spain (Wagner et al., 2007), and $0.07 \text{ m}^3\text{m}^{-3}$ (0.62) in France (Rüdiger et al., 2009). Early results for ASCAT are similar: Albergel et al. (2009) obtained RMSD (correlation) estimates of $0.06 \text{ m}^3\text{m}^{-3}$ (0.58) in France.

2.4.2 Future missions

Much of the present research applying remotely sensed soil moisture data is motivated by the emerging availability of soil moisture observations from ESA's SMOS Mission (Kerr et al., 2010), and NASA's SMAP mission (formally Hydros, Entekhabi et al., 2004), both of which operate in the microwave frequencies at L-band. SMOS was launched in November 2009, and is the first purpose designed soil moisture remote sensing mission. It carries an L-band (1.4 GHz) passive microwave radiometer, which is sensitive to soil moisture in the top 5 cm of the surface and is considered the ideal frequency for sensing soil moisture. It is anticipated that SMOS will provide global soil moisture observations at an accuracy of $0.04 \text{ m}^3\text{m}^{-3}$ every three days (Kerr et al., 2001). The SMAP instrument will also observe at L-band, using both a radiometer and a synthetic aperture radar (providing 10km resolution), and is designed to measure soil moisture and its freeze/thaw states. Both of the SMAP and SMOS missions should provide soil moisture observations with enhanced accuracy and utility, and these data sets will likely supersede the present data sets.

2.4.3 Validation of remotely sensed soil moisture

The traditional approach to validating remotely sensed soil moisture has been by comparison to in situ soil moisture data, however there are several short-

comings associated with this approach. Most notably, remotely sensed and in situ soil moisture observations relate to quite different quantities. Horizontally, remote sensors observe an area-average soil moisture, typically with a resolution of tens of km, while in situ sensors observe the soil moisture at a point, yet soil moisture variability is controlled by different processes across these two scales (Vinnikov et al., 1999; Entin et al., 2000; Robock et al., 2000). At the coarse scale of remote sensing observations, soil moisture variability is driven by atmospheric forcing, predominantly precipitation, while point-based observations are controlled by fine scale processes, due to the land surface characteristics (such as soil properties, vegetation, and topography) controlling the infiltration and gravitational drainage of precipitated water (Vinnikov et al., 1999). These differences prevent direct comparison between remotely sensed and ground-based estimates of soil moisture. However, at the coarse scale of remotely sensed observations the within-pixel fine scale variability can be regarded as random noise (Vinnikov et al., 1999; Robock et al., 2000), and Schmugge and Jackson (1996) recommend estimating the area-average by averaging a large number (>50) of point-based observations. More recent studies indicate a modest number of carefully chosen in situ sensors (< 10) may suffice (Bosch et al., 2006; Cosh et al., 2006). Additionally, even though their absolute values differ substantially, point-based and area-average soil moisture estimates show a strong temporal agreement (Vachaud et al., 1996). Consequently comparisons between the two are best based on their temporal dynamics (Reichle et al., 2004b). To focus on measures of association rather than absolute difference it is common to normalise the remotely sensed data to fit the distribution of the in situ data before comparing the two (e.g., Wagner et al., 2007; Rüdiger et al., 2009).

In addition to differences in their horizontal scale, there are often differences in the depths of the two observation types, since current remote sensors observe only a thin surface layer (approx. 1 cm), while ground-based sensors tend to be deeper (typically 5 - 10 cm). The thin surface layer responds more rapidly to atmospheric forcing, and the vertical soil moisture gradient is often steep close to the surface, introducing systematic differences between near-surface soil moisture observed at different depths. The ERS and ASCAT soil moisture observations are often extrapolated to a deeper layer soil wetness from time series of near-surface wetness observations using an exponentially weighted moving average filter (e.g., Wagner et al., 1999; Ceballos et al., 2005; Albergel et al., 2009). The exponential filter has been shown to improve the agreement between

observations of the near-surface wetness and deeper layer observations, for both remotely sensed (Wagner et al., 1999) and in situ data (Albergel et al., 2008). While the exponential filter was developed from a simple vertical soil moisture flow model (Wagner et al., 1999), it is as possible that the improvements it generates are due simply to the filtering of noise.

To overcome the difference in scale between in situ and remotely sensed soil moisture, several field campaigns have attempted to observe coarse scale area-average soil moisture using intensive ground observations over limited time periods (typically one month or less). These campaigns often make use of air-plane-mounted sensors, which can be used to test soil moisture retrieval algorithms under controlled conditions. For example, see Jackson et al. (2002) (Southern Great Plains (SGP) 1999 Experiment), McCabe et al. (2005a) (Soil Moisture EXperiment (SMEX) 2002), Jackson et al. (2005) (SMEX03), and de Jeu et al. (2009) (National Airborne Field Experiment 2005). Also, airborne data can provide a useful intermediate scale observation, which can be scaled up to the satellite footprint (e.g., Drusch et al. (2004) (SGP99) and Bindlish et al. (2006) (SMEX02)).

Intensive field campaigns and in situ monitoring networks can only provide observations at limited locations. To assess remotely soil moisture data sets globally, other data must be used. One option is to assess remotely sensed soil moisture through comparison to other global soil moisture estimates, such as alternative remotely sensed soil moisture data (Wagner et al., 2007; de Jeu et al., 2008), or modeled soil moisture (Wagner et al., 2003; Naeimi et al., 2009; Rüdiger et al., 2009). An even better option would be to assess remotely sensed data against two independent soil moisture estimates, which allows root mean square error estimates (useful for data assimilation) to be calculated for each data set, via triple collocation (Stoffelen., 1998). However, this strategy is limited to situations where three *independent* soil moisture estimates are available, and hence to date it has only been applied at continental scales to past time periods during which ERS soil moisture data were available, or to recent periods during which ASCAT data were available Dorigo et al. (2010). Alternatively, since precipitation is the principal driver of soil moisture variability, information regarding the realism of remotely sensed soil moisture can be extracted by comparison to precipitation observations, either qualitatively (McCabe et al., 2005b), or quantitatively using correlation anomalies (Wagner et al., 2003), or more complex data assimilation methods (Crow et al., 2009). While these ap-

proaches certainly yield useful information regarding the accuracy of remotely sensed soil moisture, it is difficult to assign errors between the soil moisture data sets when two uncertain soil moisture data sets are compared, and it is also difficult to derive a meaningful soil moisture error from comparison to precipitation data.

2.5 Soil moisture assimilation

The recent developments in remote sensing of near-surface soil moisture outlined above introduce the potential to improve model root-zone soil moisture by assimilating this data. To date, most of the work assimilating soil moisture has used stand-alone land surface models (rather than NWP models), within projects such as the GLDAS and NLDAS. These studies are briefly reviewed below, however much of this work uses assimilation methods too complex and costly to be implemented in NWP, with experiments performed at sites with more data available than would be the case for NWP (in terms of the assimilated data, model parameters, or validation / tuning data). Those studies that have specifically focussed on assimilating soil moisture data into NWP models are then discussed separately in Section 2.5.4.

2.5.1 From near-surface to root-zone soil moisture

Remote sensors observe only the near-surface soil layer, yet the root-zone soil moisture is of greater importance to NWP (and most other applications). The model root-zone soil moisture can potentially be constrained using near-surface observations by either updating the near-surface soil moisture layer and using the model to propagate the updates through the soil profile, or by directly updating the deeper layers based on the near-surface observations (or a combination of both). The efficiency of both of these processes depends on the frequency of the observations (e.g., Calvet and Noilhan, 2000), and the coupling strength between the model near-surface and root-zone soil moisture, which is model-dependent (e.g., Kumar et al., 2009). In general the root-zone soil moisture is only weakly constrained by the surface layer, and updating only the model surface layer may not generate substantial improvements in the deeper soil layers (e.g., Walker et al., 2001b). Most assimilation systems aim to use the near-surface observations to update the model soil moisture state throughout

the profile. The viability of this approach was initially proven with synthetic data experiments, in which synthetic “data” are generated by the model, and perturbed in some way to synthesise the model error, before being assimilated back into the model and compared against an unperturbed modeled reference (the “truth”). Synthetic studies have shown that assimilation of near-surface soil moisture, or related microwave brightness temperatures, can improve both the model near-surface soil moisture, and root-zone soil moisture relative to the unperturbed model truth (Walker and Houser, 2001; Reichle et al., 2002b; Balsamo et al., 2006; Zhou et al., 2006).

Experiments assimilating nonsynthetic soil moisture observations have also demonstrated root-zone soil moisture improvements, relative to independent observations from ground based sensors. Initial experiments using in situ observations demonstrated that assimilating near-surface soil moisture observations could correct the root-zone soil moisture for errors in the initial soil moisture conditions (Calvet and Noilhan, 2000) and/or atmospheric forcing (Walker et al., 2001a). Similar results have also been obtained by assimilating microwave data collected during field experiments by airplane-mounted remote sensors. By assimilating L-band brightness temperature observations from the SGP97 field experiment, Crow and Wood (2003) improved the root-zone soil moisture and latent heat flux forecasts from the TOPMODEL-based Land Surface-Atmosphere Transfer Scheme, using an Ensemble Kalman Filter (EnKF). Also, by assimilating the SGP97 data, Dunne and Entekhabi (2006) improved estimates of the latent heat flux and near-surface and root-zone soil moisture from the Noah model with an Ensemble Kalman Smoother (with the exception of an extended period of missing data). In studies using satellite data, (Reichle and Koster, 2005) assimilated C-band Scanning Multichannel Microwave Radiometer (SMMR) data, and (Reichle et al., 2007) assimilated SMMR C-band and AMSR-E X-band data, into the NASA Catchment Land Model with an EnKF. In both of these studies the assimilation improved the near-surface soil moisture estimates relative to both the original modeled and observed values, however this had only a modest impact on the root-zone: Reichle and Koster (2005) obtained small root-zone soil moisture improvements by assimilating the SMMR data, while for Reichle et al. (2007) only the AMSR-E data generated significant root-zone improvements. Also using the Catchment Land Model and an EnKF, Ni-Meister et al. (2007) improved the model root-zone soil moisture by assimilating C-band SMMR data globally, and in the only study focused

on the Australian continent, Walker et al. (2003) assimilated C-band SMMR to improve the fit between the model root-zone soil moisture and Normalised Difference Vegetation Index (NDVI) data over Australia. In the first large scale studies to assimilate satellite-based soil moisture data into NWP models Drusch (2007) and Scipal et al. (2008) used a simple nudging method, in which only the near-surface soil layer is updated, to assimilate soil moisture into ECMWF's NWP model. While both reported improved correlations to in situ soil moisture data, the bi-weekly ERS data assimilated by Scipal et al. (2008) was insufficient to fully constrain the root-zone soil moisture, while the daily TMI data used by Drusch (2007) was more effective. Additionally, the UK Met Office recently implemented a similar nudging scheme to assimilate ASCAT derived soil moisture into their operational NWP model, resulting in a small positive impact on forecasts of screen-level temperature and humidity for the tropics, north America and Australia (Dharssi et al., 2010).

Our capacity to correctly update model root-zone soil moisture from near-surface observations depends on the accuracy of the model physics and the strength of the coupling between the model near-surface and root-zone soil moisture (Kumar et al., 2009). Many of the studies assimilating remotely sensed soil moisture data have used the same model: NASA's Catchment Land Model (e.g., Walker et al., 2003; Reichle and Koster, 2005; Ni-Meister et al., 2007; Reichle et al., 2007), and there is a clear need to experiment with a broader range of models. In particular very little work has been done with NWP models: the only studies to have assimilated remotely sensed data into NWP models (or their land surface schemes) are Drusch (2007), Scipal et al. (2008), and Dharssi et al. (2010), all of which used a simple assimilation scheme (and all recommend developing more advanced methods). Additionally, much of the work to date has focused on North America, and all validations against in situ observations have used data from the USA (Crow and Wood, 2003; Reichle and Koster, 2005; Dunne and Entekhabi, 2006; Drusch, 2007; Reichle et al., 2007; Scipal et al., 2008), or (and) the Global Soil Moisture Data Bank (Reichle and Koster, 2005; Ni-Meister et al., 2007; Reichle et al., 2007), using Northern Hemisphere sites only⁴. Assimilating soil moisture across a more diverse range of locations will advance our understanding of how the assimilation is influenced by differences in climate, surface conditions, and observation coverage and quality (including

⁴The GSMDB includes two Australian sites, however neither has been used to validate soil moisture assimilation experiments.

forcing).

Radiance vs. retrieval assimilation

Some of the studies cited above assimilated soil moisture fields retrieved from microwave brightness temperatures (retrievals; Reichle and Koster, 2005; Drusch, 2007; Ni-Meister et al., 2007), while others assimilated the brightness temperatures (radiances; Crow and Wood, 2003; Seuffert et al., 2004; Dunne and Entekhabi, 2006), by coupling the land surface model to a radiative transfer model. In general radiance assimilation is superior, since it solves the (under-determined) microwave emission equations using ancillary fields consistent with the model. Additionally, the error correlation structure of radiance observations is easier to define, although bias correction is more difficult (Balsamo et al., 2006).

2.5.2 Soil moisture data assimilation techniques

A range of methods have been applied to soil moisture assimilation, including variational smoothers (Castelli et al., 1999; Calvet and Noilhan, 2000; Reichle et al., 2001), and sequential Kalman filter-based approaches (Walker and Houser, 2001; Reichle et al., 2002b; Crow and Wood, 2003). Variational smoothing seeks to minimise a cost function measuring the net disagreement between the observations and the model. The assimilation is performed over a time-window, and the cost function is minimised across all observations within that window. To be computationally affordable this requires an adjoint of the forecast model (which is derived from the tangent linear model), to translate the model state backwards in time. In contrast to smoothers, sequential assimilation methods update the model state each time new data are available. The Kalman filters update the model states by minimising the expected error covariance of the analysed variables. For linear systems, the optimal Kalman filter (Kalman, 1960) explicitly propagates the model error covariance matrices through time using the forecast model. For nonlinear models, the EKF approximates the Kalman filter by using a tangent linear model to propagate the model error, while for the EnKF (Evensen, 1994) the forecast model error is diagnosed using an ensemble of forecasts to sample the error distribution.

The method used for soil moisture assimilation must be tailored to the unique characteristics of the land surface. Most notably, land surface models

are nonlinear, requiring nonlinear assimilation techniques. Additionally, land surface models contain switches that are nondifferentiable, and tangent linear models and model adjoints are not available for most land surface models. Also, the land surface is dissipative rather than chaotic, and smoothers are better suited to assimilating noisy (remotely sensed) data into dissipative models (Dunne and Entekhabi, 2006). However, variational smoothers have not been widely used for soil moisture assimilation, due to difficulty in obtaining model adjoints (although ensemble Kalman smoothers have been successfully applied in reanalysis type problems (Dunne and Entekhabi, 2005)). Those studies that have used variational methods have either; used approximate methods (Houser et al., 1998; Balsamo et al., 2006); or circumvented the need for an adjoint by using extremely expensive brute-force iterative methods to minimise the cost function (Calvet and Noilhan, 2000); or made use of simplified land surface models, specifically designed to be differentiable (Castelli et al., 1999; Reichle et al., 2001).

Due to lack of model adjoints for use in variational assimilation, the nonlinear approximations to the Kalman filter are more commonly used in land surface assimilation problems. For the EKF, the tangent linear model is typically estimated for each assimilation cycle using a first-order Taylor approximation about the model state⁵ (Walker and Houser, 2001; Reichle et al., 2002b). This requires an additional model forecast for each element of the state update vector for each assimilation cycle. These additional integrations, as well as the explicit propagation of the model error covariance, make the EKF very expensive for large-dimension problems (Reichle et al., 2002b). However, most land surface models (including those in NWP models) do not account for lateral sub-surface flow, and are essentially a series of independent 1-D models. This greatly reduces the horizontal error correlations, and in particular “errors of the day” will not develop significant flow-dependent horizontal structures, as occurs in the atmosphere. The EKF is then made affordable over large domains by neglecting horizontal error correlations, and performing an independent 1-D (vertical) assimilation at each grid (Walker and Houser, 2001; Reichle et al., 2002b). This 1-D assumption is currently made when screen-level temperature and relative humidity are used to initialise the NWP land surface states (e.g, Hess, 2001; Balsamo et al., 2004).

⁵This is the same method used by Hess (2001) to linearise the forecast model for the simplified variational analysis.

However, Reichle and Koster (2003) point out that even without lateral flow in models, horizontal model error correlations still exist for example, due to approximation or neglect of physical processes by the model, large scale errors in soil parameters, and errors in meteorological forcing data. In a synthetic experiment with an EnKF, Reichle and Koster (2003) demonstrated that the inclusion of horizontal error covariance (3-D assimilation) improved the analyses, with the greatest improvement occurring where data coverage was sparse (due to information spreading into regions with no observations): the average soil moisture error was reduced by 47% by the 1-D assimilation, compared to 58% for the 3-D assimilation. However, (even for an EnKF) the inclusion of the horizontal error correlation increased the cost of the assimilation by 60%, and the benefits obtained by Reichle and Koster (2003) are also contingent on the horizontal error correlations having been (approximately) correctly specified.

The EnKF is the most common method for assimilating soil moisture data, and Reichle et al. (2002a) demonstrated that the (3-D) EnKF generated similar error reductions to the optimal (but prohibitively expensive) 4-D Var, with half of the computational cost. A major advantage of the EnKF over the EKF is that it can more easily account for horizontal error correlations, and it does not require an accurate model linearisation (although the Gaussian assumption under-pinning the error sampling will break down under strongly nonlinear conditions). The EnKF also allows more flexible specification of errors, so that errors can be specified more intuitively (for example, through errors in forcing, or model parameters), than the additive model forecast error required by the EKF. The main cost of the EnKF is the ensemble generation. While large ensembles reduce the sampling uncertainty in the error correlations, reasonable results have been obtained with modest ensemble sizes, typically between 10 to 100 (Reichle et al., 2002a; Crow and Wood, 2003; Reichle and Koster, 2003). Reichle et al. (2002b) compared the performance of the EKF and EnKF (both 1-D), and found that the computational cost and performance of the two methods was similar for an EKF with four ensemble members, although when the ensemble size was increased, the EnKF outperformed the EKF. In contrast, Muñoz Sabater et al. (2007) found that the simplified 2-D Var of Balsamo et al. (2006) (which is very similar to an EKF - see Section 3.2 and Appendix A) outperformed an EnKF at an experimental site in France, even for very large ensemble sizes (200). While the EnKF is flexible, it is not straight-forward to implement. In particular, for dissipative land surface model, maintaining the

ensemble spread can be difficult, and Muñoz Sabater et al. (2007) suggest this as the reason for the lesser performance of the EnKF in their experiment (since the EnKF did outperform the other assimilation methods in the year used to tune the ensemble inflation factor).

2.5.3 Model - observation bias

Data assimilation is based on the assumption that the model and observations are unbiased. If biases are present, the assimilation will be sub-optimal, and will generate biased results (Dee, 2005). It is well established that soil moisture estimates from models and observations are systematically different (Reichle et al., 2004b; Ni-Meister et al., 2007), and so soil moisture assimilation must address these biases. One option is to use a “bias aware” assimilation, in which the assimilation adjusts for both bias and random error (Dee, 2005). This is an attractive idea for land surface assimilation, since the nonchaotic physics increase the susceptibility of the surface to biases, and also reduce the importance of random errors. However, bias aware techniques require the source of the (model and/or observation) biases to be identified and then modeled, and there is no established truth against which remotely sensed and model soil moisture can be referenced (Reichle et al., 2004b). In particular, model soil moisture is a model dependent property (Dirmeyer et al., 2004; Schaake et al., 2004; Koster et al., 2009). Koster et al. (2009) argue that model soil moisture is better thought of as a model specific index, rather than a physical quantity that can be observed in the field, since the model soil moisture is defined by model-specific soil parameters (such as soil texture) and parameterisations (such as the evaporation). Hence even with perfect observations of soil moisture at the same resolution as the model, the representativity error in the observations would still be significant and unknown. Consequently when de Lannoy et al. (2007) assimilated soil moisture data using a bias aware assimilation to correct for model soil moisture biases (relative to in situ land observations), the model soil moisture was improved (relative to the same observations), however the flux forecasts were degraded, since the model was tuned to give (hopefully) unbiased correct fluxes from its “biased” soil moisture climatology.

An alternative option to addressing the biases, which is more common, is to assume that the model is unbiased, and then rescale the observations to match the model climatology. This removes the observation representativity

error discussed above. The observations can be rescaled based on the range (Muñoz Sabater et al., 2007), the mean and variance (Scipal et al., 2008), or the mean, variance, and higher order modes (Reichle et al., 2004a) of the model climatology. The latter is achieved by matching the Cumulative Distribution Frequency (CDF) of observations to that of the model (Reichle et al., 2007).

2.5.4 Soil moisture assimilation in NWP

As mentioned in Section 2.3.2 several NWP centres are working towards implementing a land surface assimilation based on the simplified variational analysis scheme of Balsamo et al. (2007), which utilises an off-line version of the NWP land surface model to generate the necessary model Jacobians. These systems will be capable of assimilating both screen-level observations and remotely sensed soil moisture data. However, they are still under development, and very few published studies have assimilated soil moisture data (or associated radiances) into NWP models (Balsamo et al. (2007) assimilated synthetic microwave brightness temperatures in an information content study). In an early feasibility study with a single column version of ECMWF's NWP model and data from the SGP97 experiment, Seuffert et al. (2004) showed that assimilating L-band brightness temperatures improved the modeled root-zone soil moisture, but this degraded forecasts of screen-level temperature and humidity compared to the assimilation of screen-level variables. However, the brightness temperature assimilation did improve the screen-level forecasts relative to an open-loop, and the best overall results were obtained by assimilating both the screen-level data and near-surface soil moisture information (due to the soil moisture information compensated for a period of missing screen-level data during the experiment). As discussed in Section 2.5.1, Drusch (2007) and Scipal et al. (2008) obtained modest improvements to the model soil moisture by assimilating remotely sensed soil moisture data. However, in both cases the improved soil moisture slightly degraded the screen-level forecasts, relative to the assimilation of screen-level observations. These results, and also that of Seuffert et al. (2004), suggest that assimilating soil moisture information does not minimise screen-level forecast errors as effectively as assimilating screen-level data. This is consistent with the tendency for assimilation of screen-level data to generate unrealistic soil moisture (in part by tuning the soil moisture to compensate for unrelated errors).

2.6 Outline of research

The review of the literature presented in the previous sections highlights that root-zone soil moisture has a demonstrable influence on the evolution of the atmospheric boundary layer. If not suitably constrained the soil moisture in NWP models will drift from the true climate, resulting in a detrimental impact on atmospheric forecasts. This can result in substantial errors in seasonal forecasts, where there is a positive feedback between soil moisture and the atmosphere. Many NWP centres prevent this by analysing land surface variables from screen-level observations (typically of relative humidity and temperature). While this approach effectively minimises screen-level forecast errors, it frequently generates unrealistic soil moisture states. Ultimately the unrealistic model soil moisture will generate errors in atmospheric forecasts. To improve the realism of soil moisture in their models, several NWP centres are pursuing an alternative approach that involves assimilating remotely sensed near-surface soil moisture data. The potential value for near-surface soil moisture observations to improve weather forecasts has also motivated the development of the first purpose-built soil moisture remote sensing missions, SMOS and SMAP. Consequently, there has been significant interest in exploring the best use of that data using (sub-optimal) soil moisture data sets obtained from preexisting satellite missions.

While several studies have demonstrated that assimilating near-surface soil moisture observations can improve the root-zone soil moisture prediction in land surface models, these studies have been limited in scope. In particular, there have been few studies to date that have used actual soil moisture observations, due in part to the RFI that exists in current remotely sensed soil moisture data sets, and the uncertain accuracy of the data. Additionally, very little work has been done assimilating near-surface soil moisture (or associated brightness temperatures) into NWP models, and yet the effectiveness with which the model root-zone soil moisture can be updated from near-surface observations is model-specific, depending on the accuracy and strength of the coupling between the model near-surface and root-zone soil moisture. Consequently there is a need for assimilation experiments to be conducted with (nonsynthetic) remotely sensed observations within the framework of NWP to identify practical hurdles to an operational assimilation, for example due to time/computational constraints, shortage of quality data for validation, tuning, and forcing the model.

The assimilation of near-surface soil moisture observations cannot be expected to correct the low-level atmosphere as effectively as the assimilation of screen-level observations, since the latter targets the atmosphere directly and thus compensates for errors occurring elsewhere in the model. Since accurate atmospheric forecasts are of the most immediate interest to NWP, soil moisture data is unlikely to (immediately) replace the screen-level observations, and would instead be best to complement the screen-level observations. However, to date remotely sensed soil moisture and screen-level observations have not been assimilated together over regions larger than individual field sites, and the manner in which the information from the two data sources will interact across different regions is not well understood.

In response to these issues, this thesis seeks to better our understanding of the potential improvements to weather forecasts from assimilating remotely sensed soil moisture data into NWP models. This work is particularly motivated by the potential benefit that such an assimilation system might offer Australian NWP. Australia is well placed to benefit from the assimilation of remotely sensed soil moisture. In general, the use of satellite data has benefited atmospheric forecasts over the Southern Hemisphere more than the Northern Hemisphere, due to the lower coverage of ground-based observations in the Southern Hemisphere (Le Marshall et al., 2009), and a similar result is expected for land surface assimilation. Additionally, Australia consists of a large arid in-land region surrounded by humid coastal regions, and so has extensive areas in the transition-zone between arid and humid climates, for which soil moisture has been shown to have an enhanced impact on the atmosphere at seasonal scales (Koster et al., 2004). Finally, the factors that have limited the use of soil moisture from current microwave platforms, including AMSR-E, are not prevalent over Australia, giving an unusually complete coverage of high-quality satellite data. In particular, Australia has only a small corridor of dense vegetation, very limited occurrence of snow, frozen cover, or complex terrain, and no significant RFI.

The potential benefit of using remotely sensed soil moisture data in NWP is examined in this thesis by first assessing the quality of the available remotely sensed soil moisture observations over Australia, and then assimilating this data into the land surface of several NWP models. Each of these investigations is described in detail below.

2.6.1 Assessment of remotely sensed soil moisture observations

The first stage of the research was to select the most suitable remotely sensed soil moisture data set for assimilation over Australia, and to characterise the accuracy of those data. The above review of remotely sensed soil moisture products found that the only soil moisture data sets that were continuously available with global coverage for the time period of this study were being derived from active (ASCAT / ERS) and passive (AMSR-E) microwave sensors. However, only AMSR-E had sufficient spatial/temporal coverage for use over Australia (and most other regions including Europe), since the ERS coverage is limited to just one or two (500 km wide) swaths over southeast Australia every three days and ASCAT observations were not available at the time of this study. Four prominent algorithms to retrieve soil moisture from AMSR-E were identified in Section 2.4 : the VUA-NASA, NASA, JAXA, and USDA algorithms. The relative accuracy of these retrieval algorithms is not known (previously only the VUA-NASA and NASA retrieval algorithms had been compared), and so they have been assessed here to identify which is most accurate. In total, five remotely sensed soil moisture data sets were assessed: separate C-band (since RFI is not problematic over Australia) and X-band data sets for the VUA-NASA algorithm, and X-band data for the other retrieval algorithms, for which C-band products are not produced.

The five AMSR-E data sets identified above were assessed by comparison to in situ soil moisture data from two monitoring networks in southeast Australia, over a one year period. The most realistic of these data sets was then assessed in greater detail to better characterise its uncertainty, before being used in the assimilation experiments described below. This assessment included temporal comparison to in situ soil moisture time series over the full six year AMSR-E record, and spatial comparison against maps of related hydrological observations. The AMSR-E soil moisture observations were also bench-marked against soil moisture forecasts from the Australian ACCESS NWP model, to establish whether the model might benefit from the assimilation of AMSR-E data. As discussed in Section 2.4.3, comparisons between in situ and remotely sensed soil moisture data are hampered by the differences between the horizontal and vertical resolutions of each data type, and so the suitability of the in situ data sets for assessing large scale (modeled or remotely sensed) soil moisture has also

been examined.

The results of the assessment of the AMSR-E soil moisture are presented in Chapter 4, and have been published in Draper et al. (2009b).

2.6.2 Assimilation of remotely sensed soil moisture observations

The Assimilation Approach

The next stage of the research involved experiments assimilating the selected AMSR-E near-surface soil moisture data set. The assimilation technique, which is described in detail in Chapter 3, was based on the simplified EKF used by Mahfouf et al. (2009) to assimilate screen-level observations into Météo-France's NWP model. Following the simplified 2-D Var approach of Balsamo et al. (2007), the simplified EKF uses an off-line version of the NWP land surface model to perform the land surface analysis. In this study, the simplified EKF of Mahfouf et al. (2009) has been extended to include the assimilation of near-surface soil moisture observations: previously Mahfouf et al. (2009) had assimilated only screen-level observations, although Balsamo et al. (2007) had assimilated synthetic passive microwave brightness temperatures into the Canadian ISBA model. Both Balsamo et al. (2007) and Mahfouf et al. (2009) recommend that the full EKF be used for assimilating remotely sensed data, since it is expected to better account for the irregular availability of remotely sensed data. Hence, the simplified EKF has been expanded into a full EKF in this thesis by introducing the temporal evolution of the model error covariance matrix.

Note that most of the literature relating to the assimilation of near-surface soil moisture uses an EnKF assimilation, in preference to an EKF. However, much of this literature uses land surface models forced with reanalyses or observations, often run over small scale field sites, consequently avoiding many of the restrictions imposed on NWP modeling. Most notably, the computational cost of generating an ensemble for use in an EnKF is problematic within an NWP system. Additionally, one of the main attractions of the EnKF is its ability to include horizontal error correlations, which are (for computational reasons) neglected by the EKF method used here. However, this is not expected to be a major limitation in this thesis, since horizontal error correlations are

also neglected by the current surface analysis schemes used in NWP.⁶ Consequently, simpler and less costly methods, such as the EKF-based approaches, are favoured by the NWP community.

The review of the literature also highlighted that systematic differences are expected between soil moisture from remote sensors and models (Reichle et al., 2004b), and that these biases must be addressed by the assimilating system (Dee, 2005). While a bias aware assimilation of soil moisture would be more consistent with the nonchaotic nature of land surface models, at the continental scale there is no established soil moisture truth against which the biases could be defined. Additionally, the model-dependent nature of soil moisture introduces significant representativity errors between modeled and observed soil moisture. Consequently, correcting model soil moisture to an observed “true” climatology would be unlikely to improve land surface flux predictions, since the fluxes are tuned to interpret soil moisture consistent with the model’s climatology. Hence, this thesis has adopted the common approach of normalising the soil moisture observations to be consistent with the the model climatology, using the CDF-matching technique of Reichle and Koster (2004). Finally, for simplicity, soil moisture retrievals were assimilated rather than radiances.

The assimilation experiments

The EKF has been used to assimilate near-surface soil moisture from AMSR-E into the land surface models in the NWP suites used at the Australian BoM and Météo-France, and each of these experiments is outlined below. Since screen-level observations were not available for the recently launched Australian model⁷, the experiments comparing assimilation of screen-level and near-surface soil moisture observations were conducted with Météo-France’s model. Con-

⁶Note that with appropriate localisation, horizontal error correlations could be included in the EKF, by concatenating the state-update vectors from multiple grids, and specifying the assumed off-diagonal cross-correlations. This would not require any additional perturbed model runs to linearise the model with the expanded state update vector, since the model Jacobians for flow between adjacent grids in NWP models will be zero, due to the absence of horizontal flows.

⁷Before screen-level data are assimilated into the land surface component of an NWP model they must be analysed onto the model grid. Within an NWP suite this is done during the atmospheric assimilation cycle, using the model screen level forecasts as the background state. The only screen-level archives that the Bureau of Meteorology has for the experiment period were generated using background information from the BoM’s previous NWP model, the Limited Area Prediction System (LAPS;Puri et al., 1998). These analyses are not suitable for use in other models, since they will contain LAPS-specific biases.

ducting assimilation experiments with two models, which have very different soil moisture physics (see Chapter 3), offers an additional opportunity to highlight the role of the model physics in the assimilation.

While the EKF is designed to be semicoupled to the NWP model, the land surface assimilation experiments conducted here have been completely decoupled from the atmospheric model, and static short-range (24-hour) forecasts (generated once) from the NWP model were used to force off-line versions of the NWP land surface schemes. This neglects feedback between the soil moisture updates and the atmospheric forecasts. However, it greatly reduces the computational cost and complexity of the experiments conducted here, while still allowing the main dynamics of the assimilation to be examined.

Assimilation of near-surface soil moisture and screen-level observations into the Météo-France’s ISBA model

Near-surface soil moisture from AMSR-E and screen-level observations have been assimilated into the land surface model used in Météo-France’s ALADIN NWP model, over the European domain. The land surface model in ALADIN is the two-layer version of the Interactions between Surface, Biosphere, and Atmosphere (ISBA) model, which is described in Chapter 3. Satellite observations of soil moisture have not been assimilated into ISBA before, beyond the scale of individual field sites, and the mechanisms by which the near-surface soil moisture observations are related to the model root-zone soil moisture were not well understood. Consequently, the manner in which the near-surface soil moisture observations are translated into root-zone soil moisture increments is first established, by examining the model Jacobians generated by the assimilation. Then, the assimilation of AMSR-E observations is compared to the assimilation of screen-level observations, before assimilating the AMSR-E soil moisture and screen-level data together, to determine the level of agreement between the two observation types, and the manner in which the information from each is combined by the assimilation. Also, the potential benefit of using the full EKF was tested by comparing the EKF and simplified EKF assimilation. The ISBA assimilation experiments were limited to the month of July 2006 by the availability of forcing data from Météo-France. While one month is too short to confidently validate results, the emphasis of these experiments was on understanding the dynamics of the assimilation and the interactions between the near-surface soil moisture and the screen-level observations.

The results of the above investigations are presented in Chapter 5, and the experiments assimilating near-surface soil moisture have been published in Draper et al. (2009a), while the assimilation of near-surface soil moisture and screen-level observations have been published in Draper et al. (2011). This work also contributed to the simplified EKF assimilation of screen-level data reported in Mahfouf et al. (2009).

Assimilation of near-surface soil moisture in the Australian MOSES model

AMSR-E soil moisture observations have been assimilated over Australia into the land surface model used in the Australian ACCESS model, the Met Office Surface Exchange Scheme (MOSES), which is described in Chapter 3. The manner in which the model relates the near-surface soil moisture observations to moisture in each soil layer is first identified, again by examining the model Jacobians. The AMSR-E observations are then assimilated into MOSES, and the net impact on the model near-surface and root-zone soil moisture is examined, before the impact on the model forecast skill is assessed by comparison to in situ soil moisture soil moisture observations. The assimilation was conducted over a one year period from April 2008 to take advantage of an important bug-fix to the derivation of soil properties in MOSES. The ACCESS model was only recently ported to the BoM from the Met Office, and was still under development during this research. Consequently the BoM had no archived ACCESS forecasts, and atmospheric forecasts (for forcing MOSES) were specifically generated for these experiments using a research version of ACCESS.

The results of the above investigations are presented in Chapter 6.

2.7 Chapter summary

This chapter has presented an overview of the literature supporting this thesis, and then based on this literature review the research methods used throughout this thesis have been formulated. The literature review highlighted that model root-zone soil moisture can significantly impact the accuracy of weather forecasts. In most NWP models, including those used at the BoM and Météo-France, soil moisture is initialised by assimilating screen-level observations of temperature and relative humidity. While this approach improves low-level at-

ospheric forecasts, it does not generally result in realistic soil moisture, since soil moisture is often adjusted to compensate for errors occurring elsewhere in the model. A promising approach to improving the realism of NWP model soil moisture is to assimilate remotely sensed observations of soil moisture in a thin near-surface layer. Several studies have demonstrated that assimilating near-surface soil moisture data can improve model root-zone soil moisture. However assimilating near-surface soil moisture observations is not expected to correct the low-level atmosphere as effectively as the assimilation of screen-level observations, since the latter accounts for feedback from the surface to the atmosphere (thus accounting for model errors unrelated to soil moisture). Since accurate atmospheric forecasts are the most immediate concern for NWP, remotely sensed soil moisture data would be best used together with the screen-level observations (in current-generation models at least).

This thesis examines the potential value of assimilating remotely sensed near-surface soil moisture into NWP models over Australia. The impact of assimilating soil moisture data has been examined when it is assimilated on its own, and together with screen-level observations. An EKF is used for the assimilation experiments, which is based on the simplified 2-D Var/EKF approach of Balsamo et al. (2007) and Mahfouf et al. (2009). The thesis experiments are divided into three broad investigations:

- The available remotely sensed soil moisture data sets over Australia have been assessed, and the most accurate was selected for use in the assimilation experiments.
- The selected AMSR-E soil moisture data set has been assimilated into Météo-France ISBA land surface model over one month, together with screen-level temperature and relative humidity, to determine the level of agreement between the two data sets, and how they interact when used together.
- The AMSR-E soil moisture has then been assimilated into the Bureau of Meteorology's MOSES land surface model over one year, and the resulting near-surface and root-zone soil moisture forecasts have been assessed against in situ soil moisture observations.

Chapter 3

The EKF and the Land Surface Models

3.1 Overview of chapter

This chapter describes the land surface models used in the assimilation experiments in this thesis, together with the method used to assimilate near-surface soil moisture (and screen-level) observations into these models. The assimilation experiments have been performed with an EKF-based land surface analysis scheme, described in Section 3.2, which is derived from the simplified EKF of Mahfouf et al. (2009). In Chapter 5 AMSR-E near-surface soil moisture and screen-level observations are assimilated into the ISBA land surface model, which is used in Météo-France’s ALADIN NWP. The experiments with the ISBA model were conducted in a preexisting experimental environment, called SURFEX, and the ISBA model and the SURFEX modelling environment are described below in Section 3.4.1. In Chapter 6 AMSR-E near-surface soil moisture is assimilated into the MOSES land surface model, which is used in the Australian ACCESS NWP model. There was no previous land surface assimilation research with the ACCESS model at the Bureau of Meteorology, and an off-line modelling environment for MOSES has been developed for use in this thesis; both the MOSES model and the environment in which it has been run are described below in Section 3.4.2.

3.2 The EKF

An EKF has been used in this thesis, which is based on the simplified EKF of Mahfouf et al. (2009) and the equivalent simplified 2-D Var of Balsamo et al. (2006) and Balsamo et al. (2007). Both Balsamo et al. (2007) and Mahfouf et al. (2009) recommended extending their “simplified” approaches into a full EKF, in which the background error covariances are propagated through time. It is expected that this will better account for the different observation availabilities when remotely sensed soil moisture and screen-level data are assimilated together, particularly given that the temporal availability of the remotely sensed data is variable. For the assimilation of near-surface soil moisture, an additional advantage of the full EKF is that it will introduce a long-term memory into the model error cross-covariance relating the near-surface and root-zone soil moisture model errors (increasing the accuracy of the background cross-covariances used by the EKF). Hence, the simplified approaches have been extended into a full EKF in this thesis, by introducing the temporal evolution of the background error covariances.

Using the notation of Ide et al. (1997), the Extended Kalman Filter (Ghil and Malanotte-Rizzoli, 1991) equations for the i th model state forecast and update at time t_i are:

$$\mathbf{x}^b(t_i) = \mathcal{M}_{i-1}[\mathbf{x}^a(t_{i-1})] \quad (3.1)$$

and

$$\mathbf{x}^a(t_i) = \mathbf{x}^b(t_i) + \mathbf{K}_i (\mathbf{y}_i^o - \mathcal{H}_i[\mathbf{x}^b(t_i)]) \quad (3.2)$$

where \mathbf{x} indicates the model state vector and \mathbf{y} is the observation vector. The superscripts a , b , and o indicate the analysis, background, and observations, respectively. \mathcal{M} is the nonlinear state forecast model, and, \mathcal{H} is the nonlinear observation operator (to map the model state into the observation state space). \mathbf{K} is the Kalman gain, given by:

$$\mathbf{K}_i = \mathbf{P}^f(t_i) \mathbf{H}_i^T (\mathbf{H}_i \mathbf{P}^f(t_i) \mathbf{H}_i^T + \mathbf{R}_i)^{-1} \quad (3.3)$$

where \mathbf{P} and \mathbf{R} are the covariance matrices of the model background and observation errors, respectively, and \mathbf{H} is the linearisation of \mathcal{H} . The defining feature of the EKF is that the background model error is also evolved through a series of model forecasts and updates:

$$\mathbf{P}^f(t_i) = \mathbf{M}_{i-1}\mathbf{P}^a(t_{i-1})\mathbf{M}_{i-1}^T + \mathbf{Q}(t_{i-1}) \quad (3.4)$$

$$\mathbf{P}^a(t_i) = (\mathbf{I} - \mathbf{K}_i\mathbf{H}_i)\mathbf{P}^f(t_i) \quad (3.5)$$

where \mathbf{Q} is the error covariance matrix for the (additive) model forecast error, and \mathbf{M} is the linearisation of \mathcal{M} .

For the EKF as applied here, the observations are 6 hours later than the analysis time, and the observation operator is a 6 hour model forecast followed by conversion to the observed state, following Mahfouf (1991) and Hess (2001). For example, for the ISBA model from Chapter 5, in which a two-layer soil moisture state is updated from observations of the upper-most layer, the observation operator is:

$$\mathcal{H}_i = \begin{pmatrix} 1 & 0 \end{pmatrix} \mathcal{M}_{t_i \rightarrow t_i+6} \quad (3.6)$$

A 6 hour forecast was chosen to match the availability of the screen-level observations. Due to the continuous evolution of the background error matrix through time, the forecast-length used in the observation operator has little influence on the assimilation results for the full EKF.

The linearisation of \mathcal{H} and \mathcal{M} is obtained by finite differences, using a first-order Taylor expansion about \mathbf{x} . The n th element of the observation vector, and the m th element of the control vector, \mathbf{H}_{nm} is given by:

$$\mathbf{H}_{nm,i} = \frac{\mathcal{H}[\mathbf{x}(t_i) + \delta x_m(t_i)]_n - \mathcal{H}[\mathbf{x}(t_i)]_n}{\delta x_m(t_i)} \quad (3.7)$$

Calculating the linearisation requires an additional (perturbed) 6 hour model integration for each element of the state vector. Following Balsamo et al. (2007)

the computational cost of the additional model forecasts required by the observation operator are reduced by performing the assimilation in an off-line version of the NWP model's land surface scheme. The off-line assimilation is designed to be semicoupled to the NWP model, such that the NWP model is updated with each land surface analysis, before the forcing for the subsequent off-line assimilation cycle is generated by the NWP model. To assimilate the screen-level observations in this off-line system, the land surface model is forced at the height of the first atmospheric model layer (higher than a typical land surface model) to allow the screen-level observations to be assimilated off-line, again following Balsamo et al. (2007).

Since \mathbf{H} is used in equation (3.3) before the state vector is updated and \mathbf{M} is used in equation (3.4) after the update, two sets of additional model forecasts are required for each assimilation cycle. However, for the assimilation experiments with the ISBA model in Chapter 5 an approximation has been introduced in which the perturbed model forecasts made before the state vector update (for estimating \mathbf{H}) were used to estimate \mathbf{M} (see Section 5.3.2)¹.

3.3 Related assimilation approaches

The simplified EKF of Mahfouf et al. (2009) differs from the EKF described above in that equations (3.4) and (3.5) for the update and evolution of the background model error covariance matrix are not applied. Instead the same \mathbf{P} is used at the beginning of each assimilation cycle, based on the assumption that the reduction in \mathbf{P} induced by each analysis (equation (3.5)) is balanced by the subsequent increase in \mathbf{P} during the next model forecast (equation (3.4)). Note that the simplified EKF still accounts for some temporal evolution of \mathbf{P} , due to the inclusion of a model forecast in the observation operator.

Additionally, the EKF is also closely related to a number of other land surface analysis techniques, as is reviewed in Appendix A. Specifically, the similarity between the simplified variational approach of Balsamo et al. (2007) and the EKF is demonstrated. It is also shown that for assimilating of near-surface soil moisture the soil moisture analyses generated by the EKF presented here differs only slightly from those generated by the more standard EKF formula-

¹This approximation was not introduced for the MOSES model in Chapter 6, since the data were assimilated (and the observation operator calculated) only once every 24-hours, or every fourth assimilation cycle.

tion (e.g, as used by Walker and Houser, 2001; Reichle et al., 2002b; Muñoz Sabater et al., 2007), in which the observation operator for updating the model near-surface soil moisture is equal to the identity (rather than using a model forecast).

3.4 The land surface models

The modelling systems used in the assimilation experiments are described below. While the off-line land surface analysis described above is designed to be semicoupled to the NWP model, in the experiments conducted here static short-range forecasts (generated once) from the NWP model have been used to force the land surface models. Apart from this, the off-line land surface models are run in an environment that mimics the coupled NWP models as closely as possible, and the modeling environment used for each is also described below.

3.4.1 The French ALADIN and ISBA models

The ALADIN NWP model

The NWP suite used at Météo-France was developed in collaboration with ECMWF. It includes the Action de Recherche Petite Echelle Grande Echelle (ARPEGE) global model, and its limited area version, ALADIN/France (Bubnova et al., 1995), referred to throughout this thesis as “ALADIN”. ALADIN and ARPEGE are hydrostatic models, with hybrid terrain-following height coordinates, semi-Lagrangian advection, and a 4D-VAR atmospheric assimilation. ALADIN is nested within ARPEGE, and it has an (irregular) stretched grid centred over France, with a resolution of approximately 9 km. The 2006 operational configuration used here has 46 vertical levels.

The ISBA land surface model

ALADIN uses a two-layer version of the ISBA land surface scheme (Noilhan and Planton, 1989; Noilhan and Mahfouf, 1996). ISBA is a force-restore model and has eight prognostic variables: surface temperature, mean (deep-layer) soil temperature, near-surface water content (liquid / frozen), total (deep-layer) water content (liquid / frozen), vegetation intercepted water content, and snow

water. The (liquid) near-surface soil moisture, w_1 , is defined as the reservoir from which moisture is extracted by bare soil evaporation, and is arbitrarily assigned a depth (d_1) of 0.01 m; while the total soil moisture, w_2 , is defined as the reservoir from which moisture is extracted by bare soil evaporation and transpiration, with the depth (d_2) depending on the local soil type and climate (ranging between 0.24 to 3.8 m over the ALADIN domain). In the two layer ISBA model there is no distinction between the root-zone and total layer depth, and so the root-zone is effectively modeled over the total soil layer. Horizontal sub-surface flows are neglected (each grid is modeled independent from its neighbours), and the vertical liquid moisture dynamics are described by the force restore method of Deardorff (1977):

$$\partial w_1 / \partial t = \frac{C_1}{\rho_w d_1} (P_g - E_g) - \frac{C_2}{\tau} (w_g - w_{eq}); 0 \leq w_1 \leq w_{sat} \quad (3.8)$$

$$\partial w_2 / \partial t = \frac{1}{\rho_w d_2} (P_g - E_g - E_{tr}) - \frac{C_3}{d_2 \tau} \max[0, (w_2 - w_{fc})]; 0 \leq w_2 \leq w_{sat} \quad (3.9)$$

The first term in each equation represents the forcing: precipitation (P_g) and bare-soil evaporation (E_g) affect both layers, while transpiration (E_{tr}) also acts on w_2 (equation 3.9). The second set of terms in each equation are the restore action. In equation 3.8 for w_1 the restore term represents the balance between capillary rise from w_2 and gravitation drainage from w_1 : w_1 is restored towards an equilibrium value (w_{eq}), calculated as a function of w_2 and the saturation water content, w_{sat} . In equation 3.9 for w_2 the restore term represents gravitational drainage, which is proportional to the volume of moisture exceeding field capacity (w_{fc}). ρ_w is the density of liquid water, and the coefficients (C_1 , C_2 , and C_3) are calibrated for different soil textures, with C_1 and C_2 also depending on soil moisture.

Bare soil evaporation is dependent on the soil type, surface roughness, wind speed and humidity deficit, while transpiration also depends on vegetation type, temperature, and soil moisture availability. For a given energy input the evapotranspiration from a grid varies according to w_1 and w_2 : w_1 controls the bare soil evaporation by driving the surface humidity, while the w_2 controls the stomatal resistance to transpiration. Following Noilhan and Planton (1989), the resis-

tance to transpiration depends on a normalised soil moisture, called the Surface Wetness Index (SWI):

$$SWI = \frac{w_2 - w_{wilt}}{w_{fc} - w_{wilt}} \quad (3.10)$$

Transpiration occurs at the potential rate for SWI above 1 ($w_2 > w_{fc}$), while it effectively ceases for SWI less than 0 ($w_2 < w_{wilt}$), and for intermediate soil moisture values transpiration is nonlinearly related to SWI. The soil temperature dynamics are determined by the force-restore method proposed by Bhumralkar (1975), and a single energy budget is applied to both the vegetation and bare-soil surface for each grid.

All of the required soil and vegetation parameters in ISBA are calculated from the soil and vegetation classification, which are obtained from the ECOCLIMAP data base (Masson et al., 2003). Each nonurban land grid is classified as a mixture of nine land cover types (broadleaf trees, coniferous trees, C3 and C4 crops, flat bare soil, grassland, irrigated crops, parks and marshes, permanent snow, rocks, tropical grass, and tropical trees), and the surface parameters are aggregated to a single parameter set for each grid. The Köppen-Geiger climate classification (Peel et al., 2007) classes most of the ALADIN Europe domain as temperate, with arid regions in Spain and north Africa, and some isolated area with continental climates, associated with the Alps. The land cover is quite diverse, and is predominantly made of up broadleaf and needleleaf trees and nonirrigated crops, with isolated regions of irrigated crops, permanent ice, rocks, and parks and marshes.

Running ISBA off-line in the SURFEX environment

The assimilation experiments conducted with the ISBA model used the same experimental set-up as was used by Mahfouf et al. (2009) for the simplified EKF assimilation of screen-level observations into ISBA. ISBA was run within the Surface EXternalisée (SURFEX) Environment (Le Moigne et al, 2009) , in which it has been decoupled from the atmospheric model using the implicit coupling scheme of Best et al. (2004). All ISBA model parameters were read from the same ancillary files used by ALADIN, and all model physics options were set the same as for ALADIN. The atmospheric forcing for ISBA (of precipitation, temperature, specific humidity, pressure, wind components, and short- and

long-wave radiation) was derived from ALADIN forecasts, interpolated onto the ISBA time-step (300 s). The atmospheric forcing is applied at the first atmospheric model layer in ALADIN (17 m), and prognostic screen-level forecasts are obtained by interpolating between the relevant quantities from the atmospheric forcing and the forecast surface values, using Monin-Obukhov similarity theory. It was demonstrated by Mahfouf et al. (2009) that this method provides sufficient sensitivity between the model surface and screen-level variables to accurately estimate the model Jacobians required to assimilate the screen-level observations in the SURFEX off-line environment.

3.4.2 The Australian ACCESS AND MOSES models

The ACCESS NWP model

The Australian ACCESS NWP model is a coupled climate and Earth-system simulator, based on the UK Met Office's Unified Model (UM). The UM (Davies et al., 2005) was designed to model the land surface, ocean, and atmospheric dynamics, across a range of scales, from operational weather forecasting to long range (centuries plus) climate research. The UM is a nonhydrostatic model, with a hybrid height formulation, a regular latitude-longitude grid, semi-Lagrangian advection, and a 4D-VAR atmospheric assimilation system (Rawlins et al., 2007). The UM was ported to the BoM in 2006 (Puri, 2006) and adapted for local use, before being launched as the BoM's operational NWP model in September 2009 (National Meteorological and Oceanographic Centre, 2010). The initial operational implementation of ACCESS consisted of a hierarchy of nested models with similar spatial resolution to the legacy NWP suite, LAPS. The Australian-domain ACCESS simulation (ACCESS-A) is nested in a regional simulation (ACCESS-R), which is in turn nested in the global ACCESS model (ACCESS-G), all of which have 50 vertical model levels, and horizontal resolutions of approximately 12, 37, and 80 km, respectively.

As discussed in Section 2.3 soil moisture and temperature are nudged in the Met Office's global model according to errors in low-level temperature and humidity (Best and Maisey, 2002). In their regional models, the Met Office initialise soil moisture by either interpolating global fields for regions outside the UK, or by using fields from the UK NWP Post-Processing (UKPP) system over the UK. The UKPP generates nowcasts of the soil states and surface hydrology at 2 km resolution over the UK, using the same land surface scheme as the

UM, coupled to a probability distributed model for generating surface runoff, as well as a river flow model (Smith et al., 2006). In the Australian ACCESS modelling suite, soil moisture and temperature are initialised in the global and regional models using the nudging scheme of Best and Maisey (2002) (the second initialisation in the regional model is required since the globally nudged fields are not yet available when the regional model forecast is commenced, and the high quality observations required to run a UKPP-type scheme are not available over Australia).

The MOSES land surface model

The land surface scheme in ACCESS is the Met Office Surface Exchange Scheme (MOSES) 2.2 (Cox et al., 1999; Essery et al., 2001). The soil moisture dynamics in MOSES are very different to those used in ISBA. In its standard formulation (as used in NWP), MOSES has four soil layers, of thickness, 0.1, 0.25, 0.65, 2.0 m. For the vegetation covers present in Australia, the root-zone is largely contained within the three uppermost model layers (covering 1.0 m)². MOSES has 14 prognostic variables: moisture (frozen and liquid) and temperature in each layer, snow on the ground, and canopy water storage. Horizontal sub-surface flows are neglected (each grid is modeled independent from its neighbours), and vertical moisture flow is a finite difference form of the Richards equation (Richards., 1931):

$$\partial M_i / \partial t = W_{i-1} - W_i - E_i \quad (3.11)$$

where, M_i is the net volume (mm) of soil moisture in the i th layer, W_i is the diffusive flux flowing from layer i to $i + 1$, and E_i is the evapotranspiration flux from layer i , calculated from the total evapotranspiration based on the soil moisture and root density profiles. The diffusive flux for each layer depends on the soil moisture and soil moisture gradient between neighbouring layers, and is calculated according to the Darcy equation (Darcy, 1856) extended to

²With the exception of some small regions of broadleaf vegetation, the root-zone in MOSES has an e-folding root-depth of 0.5 m, so that 90% of the roots are within 1.0 m of the surface.

unsaturated soils:

$$W = K\left(\frac{\partial\psi}{\partial z} + 1\right) \quad (3.12)$$

where ψ is the soil suction, z is the depth from the surface, and K is the hydraulic conductivity. These are calculated using the Clapp and Hornberger relations:

$$\psi = \psi_s S^{-b} \quad (3.13)$$

$$K = K_s S^{2b+3} \quad (3.14)$$

where K_s , ψ_s , and b are empirical soil dependent constants, and S is the unfrozen soil moisture as a fraction of the volumetric soil moisture at saturation (θ_s):

$$S = \frac{\theta}{\theta_s} \quad (3.15)$$

There is free drainage from the lower boundary (i.e. $W_4 = K_4$), and the surface flux is precipitation throughfall minus runoff. Evapotranspiration is the sum of evaporation from the canopy and bare soil, and transpiration via vegetation. Canopy evaporation occurs at the potential rate, and depends on surface roughness, wind speed, and surface layer stability. Bare soil evaporation and transpiration experience an additional surface resistance, dependent on soil moisture in layer one, while transpiration has an additional canopy resistance, dependent on vegetation type, temperature, humidity deficit, radiation, and soil moisture availability. The soil thermodynamics consist of diffusive heat exchange between soil layers, together with heat advection generated by the flux of moisture. The surface heat flux is calculated to balance the surface energy budget, and the lower boundary has zero heat flux.

MOSES 2.2 is a tiled model (Essery et al., 2003), and each model grid is divided between nine surface cover types (the ‘‘tiles’’): urban, inland water, bare soil, ice, and five plant function types (broadleaf, needleleaf, temperate grass,

tropical grass, shrubs). According to the Köppen-Geiger climate classification (Peel et al., 2007), Australia is predominantly arid in the center and west, with a narrow temperate strip along the east coast, and also in the south, and a tropical climate in the far north. The arid centre is classed as bare soils and shrubs, while the north is tropical grasses, and the temperate regions have a mixture of broadleaf trees (concentrated along the east coast), temperate grass, bare soil, and shrubs. There are no grids categorised as ice or needleleaf trees, and only isolated instances of lakes and urban environments. MOSES has the capability to calculate a separate energy and water balance for each tile (although all tiles interface with the same sub-surface soil layers), however this heterogeneity is not fully enabled in the ACCESS model and the surface parameters are instead aggregated across all tiles in each grid to give a single value for use in the surface fluxes calculations (as is done in ISBA).

Running MOSES off-line: the JULES model

An off-line version of MOSES, called the Joint UK Land Environment Simulator (JULES), was developed at the Joint Centre for Hydrometeorological Research to enable the use of MOSES outside of atmospheric modelling applications (Joint Centre for Hydro-Meteorological Research, 2001). For the experiments in this thesis, JULES was configured to be as close as possible to the coupled MOSES model, and all model physics options and parameters were taken directly from ACCESS. There were no archived ACCESS forecasts available for the period of these experiments, and so forecasts were specifically generated to provide the required forcing (and also initial conditions) for the off-line model. Since the ACCESS modelling suite was still under development at the BoM, a research version of the UM, based on version 6.4 – Parallel Suite 18 (PS18), was used. The research suite consisted of a development version of ACCESS-R (run at the same resolution: 0.375° with 50 vertical layers), run once daily and nested within a development version of the global model, (run at the resolution as the Met Office global model: 0.5625 degrees meridionally, and 0.375 degrees zonally, and 50 vertical levels). The atmospheric assimilation in ACCESS was still under development at the time of these experiments, and so the global ACCESS model was initialised each day with global UM atmospheric analyses downloaded from the Met Office, to incorporate the updates from their assimilation cycle.

The ACCESS modeling suite described above was used to generate one year of consecutive 24 hour forecasts for forcing MOSES. These forecasts, of surface

pressure, downward short- and long-wave radiation, and low-level atmospheric temperature, humidity, wind, and large scale and convective precipitation (separated into snow and rain), were saved hourly and then interpolated onto the 15 minutes time-step used for MOSES. Following the Assistance for Land surface Modelling Activities guidelines developed for the Global Land/Atmosphere System Study (Assistance for Land-Surface Modelling Activities, 2004) the fluxes were specified as hourly mean values, while the remaining forcing variables were specified instantaneously. The fluxes were applied at a constant rate each hour (rather than being interpolated), to ensure that the net quantity across the diurnal cycle was maintained for each. Finally, since screen-level observations have not been assimilated into MOSES, the forcing is applied at the standard reference heights used by MOSES within ACCESS (10 m for wind, 1.5 m for temperature and humidity).³

JULES was launched in October 2006, with the same programming as in MOSES at that time. However, the Met Office updates the UM code four times a year, and the MOSES code in ACCESS now differs from that in JULES. In particular, there have been two significant updates to MOSES since the launch of JULES, including an update for PS18 in April 2008 to amend a long standing error in the soil hydraulic properties. The PS18 update had a positive impact on the atmospheric forecast skill of the UM, including a significant reduction in the screen-level temperature and humidity errors (Dharssi et al., 2009). Hence the ACCESS model experiments conducted here have been limited to the period after April 2008 to take advantage of this update, and the JULES model code has been updated to better reflect the current MOSES code (and to amend several model short-comings encountered during this research) as is described in Appendix B. It is also demonstrated in the Appendix that the off-line JULES model environment developed here provides a sufficiently accurate representation of the MOSES model within ACCESS to be used in these experiments (although there is a discrepancy between the soil temperature forecast by the coupled and off-line MOSES models which would need to be resolved before the off-line land surface analysis could be coupled to the ACCESS model).

³Note that height of the forcing can be specified in JULES, so that the forcing could be easily changed to the height of the first atmospheric layer for the assimilation of screen-level observations, however the sensitivity of screen-level forecasts in the off-line system would need to be carefully investigated before the screen-level observations were assimilated.

3.5 Chapter summary

The land surface assimilation method and modelling systems used in this thesis have been described in this chapter. The land surface assimilation is based on the EKF, and has been adapted from the simplified EKF of Mahfouf et al. (2009), and the equivalent simplified 2-D Var of Balsamo et al. (2007). The EKF land surface analysis is designed to assimilate both remotely sensed near-surface soil moisture (and other novel observation types) and screen-level observations, within the constraints of operational NWP. It uses an off-line version of the land surface model to perform the analysis, with the resultant soil moisture update fed-back to the NWP model before the forcing for the subsequent off-line assimilation is generated. To allow the assimilation of screen-level observations the EKF uses the forecast model itself as the observation operator, and the atmospheric forcing is applied at the first atmospheric model level. While the EKF land surface analysis is designed to be semicoupled to the NWP model, in the experiments conducted here the land surface models have been completely decoupled.

In Chapter 5 the EKF is used to assimilate near-surface soil moisture and screen-level observations into ISBA, the land surface model in Météo-France's ALADIN/ARPEGE NWP suite. Then in Chapter 6 near-surface soil moisture observations are assimilated into an off-line version of MOSES, the land surface model in the Australian ACCESS NWP suite. Both models have been configured to mimic their coupled counterparts as closely as possible. Météo-France has developed an off-line environment, called SURFEX, as a research test-bed for ISBA, and the ALADIN environment was easily mimicked within SURFEX. For the MOSES model no such environment existed, and the off-line version of MOSES, called JULES, was specifically configured for use in this thesis (including several code updates). Additionally, the ACCESS model was under development at the time of this work, and the forecasts to force MOSES were specifically generated with a research version of the limited area ACCESS model.

The main soil moisture dynamics in the MOSES and ISBA land surface models have also been described in this Chapter. These two models have very different soil moisture physics and vertical resolution: MOSES is a four-layer model with explicitly modeled diffusive flow between layers, while ISBA is a two-layer force-restore model. These differences will be important in determining

how each model responds to the assimilation of near-surface soil moisture in Chapters 5 and 6, since they affect the definition of soil moisture in each model (and specifically the soil moisture in the near-surface layer, to which the AMSR-E observations are equated during the assimilation), as well as the strength and nature of the coupling between the soil moisture in the near-surface layer and the underlying soil profile.

Chapter 4

Remotely Sensed Soil Moisture over Australia

4.1 Overview of chapter

This chapter assesses the value of current remotely sensed near-surface soil moisture data sets over Australia for use in data assimilation. Observations from the passive microwave AMSR-E instrument only are considered, since until recently this was the only instrument offering sufficient coverage over Australia. Soil moisture data sets derived from AMSR-E observations using four different retrieval algorithms are compared to in situ soil moisture observations from southeast Australia over 2006. Based on this comparison, the VUA-NASA retrieval algorithm is selected for the assimilation experiments in Chapters 5 - 6. The VUA-NASA soil moisture products are then assessed in greater detail to better characterise their accuracy, and to determine their best use in the assimilation experiments. This is achieved by temporally comparing the VUA-NASA soil moisture to the longest available in situ soil moisture data record (six years), and spatially comparing it against maps of observed precipitation and vegetation. Finally, the AMSR-E data is bench-marked against soil moisture forecasts from the BoM's ACCESS model, to establish whether the model is likely to benefit from the assimilation of the AMSR-E data in Chapter 6. Additionally, the limitations of comparing area-averaged remotely soil moisture observations against point-based in situ data are highlighted by examining the consistency of time series of in situ observations from within a single AMSR-E pixel.

Much of the work presented in this chapter has been published in Draper et al. (2009b).

4.2 Data and methods

Three different types of data have been used in this Chapter. The AMSR-E near-surface soil moisture data sets are described in Section 4.2.1. These data sets have been assessed by comparison to in situ soil moisture observations from the Murrumbidgee and Goulburn Monitoring Networks, which are described in Section 4.2.2. Additionally, the methods used to compare the in situ and remotely sensed soil moisture time series are also reviewed in Section 4.2.3. Finally, maps of the AMSR-E soil moisture have also been qualitatively compared to maps of precipitation observations, which are described in Section 4.2.4.

4.2.1 AMSR-E soil moisture

Recall from Section 2.4 that the four most prominent soil moisture retrieval algorithms that have been applied to AMSR-E have been developed by:

- NASA, following Njoku et al. (2003);
- JAXA, following Koike et al. (2004);
- USDA, following Jackson (1993); and
- VUA-NASA, following Owe et al. (2001).

The defining feature of the VUA-NASA retrieval, which will be used throughout this thesis, is the expression of the vegetation optical depth as a function of the dielectric constant and the passive microwave polarisation ratio. This function is substituted into the radiative transfer equation for the H-polarised brightness temperature, together with ancillary soil temperature. The radiative transfer equation is then solved for the soil dielectric constant, and subsequently soil moisture content. The algorithm is designed to minimise reliance on ancillary data and requires only ancillary soil temperature, which is currently derived from the 36.5 GHz V-polarised AMSR-E observations.

Due to RFI in the C-band observations across much of the globe, NASA, JAXA, and USDA retrieve soil moisture from X-band AMSR-E brightness temperatures only, while VUA-NASA generate separate soil moisture products from C- and X-band brightness temperature. However, Njoku et al. (2005) did not note any RFI over Australia, while over Europe they observed C-band RFI only over isolated urban regions (although X-band RFI was widespread over Italy and England). For Australia, the lack of RFI has been confirmed by applying the spectral difference method of Li et al. (2004) to AMSR-E observations (pers. comm. Thomas Holmes): in 2006 there was no problematic C-band RFI (defined as RFI on more than 30 days over the year), and only a small region in northeast Australia with problematic X-band RFI. Consequently, both the C- and X-band data sets have been assessed here, giving a total of five soil moisture products that have been inter-compared below.

Soil moisture data sets retrieved with each of these algorithms have been obtained from the relevant institutions for 2006. While the resolution of the C-band data is $45 \times 75 \text{ km}^2$ (Njoku et al., 2003), the AMSR-E swaths (level 1) are over-sampled at approximately every 5 km, and the (level 2 and 3) data are typically reported on a 0.25° grid. Each of the AMSR-E data sets was provided on either a regular 0.25° grid (JAXA), or at the original swath level (NASA, USDA, VUA-NASA). The latter have been projected onto the regular 0.25° grid using a nearest neighbour approach, before any further processing is performed.

Passive microwave soil moisture retrievals are thought to be more accurate during the nighttime than during the daytime, due to the greater (vertical and horizontal) homogeneity of the surface temperatures at nighttime (Owe et al., 2001). Consequently the soil moisture derived from the descending (nighttime) and ascending (daytime) AMSR-E overpasses have been assessed separately.

4.2.2 In situ soil moisture

Time series of AMSR-E derived soil moisture have been compared to in situ soil moisture data from the Murrumbidgee and Goulburn River Basins in southeast Australia, each of which are indicated in Figure 4.1. The Murrumbidgee Monitoring Network is maintained by the University of Melbourne, and currently consists of 38 monitoring stations at which surface hydrological and thermodynamic variables are observed every 20 to 30 minutes. The Goulburn Monitoring Network (Rüdiger et al., 2007) is approximately 200 km north of the Mur-

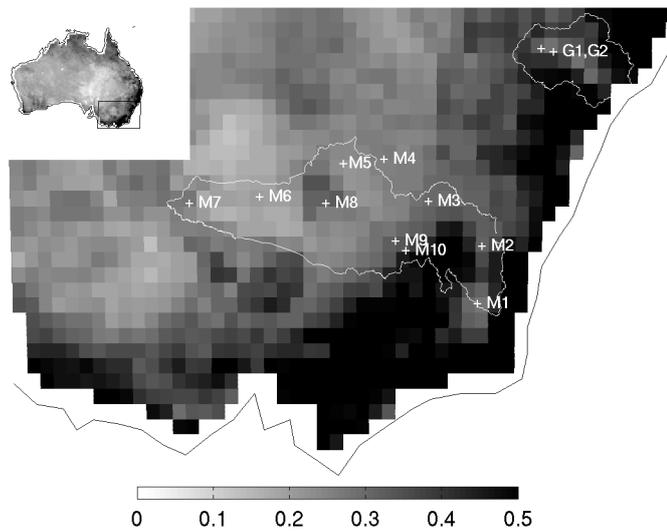


Figure 4.1: Location of the Murrumbidgee and Goulburn Monitoring Network sites, overlaid on the mean NDVI for 2006. See Tables 4.1 - 4.3 for details of each monitoring site.

urrumbidgee, and currently has 26 monitoring stations operated by the University of Newcastle.

Vegetation is the main limitation to observing soil moisture using the AMSR-E wavelengths, and the mean 2006 NDVI (from the Advanced Very High Resolution Radiometer) is plotted in Figure 4.1 to indicate the potential attenuation of the soil moisture signal by vegetation across the monitoring sites. All of the monitoring stations in both networks are grazed or grassland, although some are very close to dense vegetation (which may fall within the AMSR-E swaths). In particular, the AMSR-E pixels adjacent to Adelong (M10) are forested. The vegetation at the Goulburn and Murrumbidgee monitoring sites is dense relative to Australian conditions: a mean 2006 NDVI of 0.29 (the mean across the monitoring sites) or 0.42 (the maximum across the sites) represents the 80th and 93rd percentile of the mean 2006 NDVI across Australia.

The Murrumbidgee Monitoring Network

The Murrumbidgee Monitoring Network consists of two generations of monitoring stations. The first generation of 18 stations was installed in 2001 (see Table 4.1). Each station observes precipitation, and soil moisture at four depths (0 - 8 cm, 0 - 30 cm, 30 - 60 cm, and 60 - 90 cm), as well as soil temperature and suction at similar depths, every 20 to 30 minutes. Five stations are grouped

Table 4.1: Murrumbidgee Monitoring Network: First generation stations (installed in September 2001).

Code	Name	Latitude	Longitude	Coverage Notes
M1	Cooma Airfield	-36.293	148.971	
M2	Canberra Airport	-35.305	149.201	Missing Jun.-Nov. '08
M3	Cootamundra Aerodrome	-34.630	148.037	Missing Oct. '06 to Oct. '08
M4	West Wyalong Airfield	-33.938	147.196	
M5	Balranald	-34.658	143.549	
M6	Hay	-34.547	144.867	
M7	Griffith Aerodrome	-34.249	146.070	
M8	YANCO REGION			
Y3	Yanco Research Station	-34.621	146.424	
M9	KYEAMBA REGION			
K1	Waitara	-35.493	147.559	
K2	Kyeamba Downs	-35.435	147.530	
K3	Kyeamba Station	-35.434	147.569	Very poor coverage (\approx 50%) Dec.'06 - Apr. '08
K4	Ginnindera Flat	-35.427	147.600	
K5	Ginnindera Slope	-35.419	147.604	
M10	ADELONG REGION			
A1	Keenan	-35.497	148.106	
A2	Strathvale	-35.428	148.132	Destroyed by fire in Jan '05
A3	Weeroona	-35.400	148.101	Missing Dec.'07 - Oct. '08
A4	Rochedale	-35.373	148.066	Missing Jun. '07 - Jun. '08
A5	Crawford	-35.360	148.085	

Table 4.2: Murrumbidgee Monitoring Network: Second generation stations (installed in September 2003, with shallow soil moisture sensors installed in late 2006).

Code	Name	Latitude	Longitude	Coverage Notes
M8	YANCO REGION			
Y1	Uri Park	-34.629	145.849	
Y2	Banandra	-34.655	146.110	
Y4	Eulo	-34.719	146.020	
Y5	Dry Lake	-34.728	146.293	
Y6	S. Coleambally	-34.843	145.867	
Y7	Yamma Rd.	-34.852	146.115	
Y8	Wynella	-34.847	146.414	
Y9	Yammacoona	-34.968	146.016	
Y10	Cheverelis	-35.005	146.310	
Y11	Bundure	-35.110	145.935	
Y12	Spring Bank	-35.070	146.169	
Y13	Widgiewa	-35.090	146.306	
M9	KYEAMBA REGION			
K6	Cox	-35.390	147.457	
K7	Wollumbi	-35.394	147.566	
K8	Benwerrin	-35.316	147.344	
K10	Alabama	-35.324	147.535	
K11	Silver Springs	-35.272	147.429	
K12	Samarra	-35.227	147.485	
K13	Evergreen	-35.239	147.533	
K14	Kyeamba Mouth	-35.125	147.497	

together at Kyeamba, and four are grouped at Adelong. The other eight stations are colocated with Australian Weather Service atmospheric observation stations, which regularly monitor air temperature, humidity, wind direction and strength, and air pressure. The second generation stations were installed in 2003 (see Table 4.2), with nine stations added at Kyeamba and twelve at Yanco (where there was previously one station). These stations observed precipitation, soil temperature at 15 cm, and soil moisture at three depths (0 - 30 cm, 30 - 60 cm, and 60 - 90 cm) until late 2006, when near-surface soil moisture (0 - 5 cm) and temperature (2.5 cm) sensors were added.

Soil moisture is measured with Campbell Scientific water content reflectome-

ters (CS615 sensors for first generation sites, and CS616 sensors for the second generation sites), except for the second generation near-surface sensors which are Stevens Hydraprobe soil dielectric sensors. Calibration studies estimate an RMSE between $0.02 \text{ m}^3\text{m}^{-3}$ and $0.05 \text{ m}^3\text{m}^{-3}$ for the Campbell Scientific sensors (Young et al., 2008), and $0.03 \text{ m}^3\text{m}^{-3}$ for the Stevens Hydraprobes (Merlin et al., 2007).

The Goulburn Monitoring Network

The Goulburn Monitoring Network was established in September 2002 in the Goulburn River Basin. Initially it consisted of 26 stations, each observing soil moisture at three depths (0 - 30 cm, 30 - 60 cm, and 60 - 90 cm) with Campbell Scientific water content reflectometers. In October 2005 Stevens Hydraprobes were installed at most sites to observe the near-surface (0 - 5cm) soil moisture. While the Goulburn monitoring network is still operational, soil moisture data are currently available only until the end of 2007, due to delays in data-processing and telemetry failures. The Goulburn data have been used only for the 2006 comparison of the different retrieval algorithms (Section 4.3.2) due to their limited availability in later years. Further to this, of the 26 Goulburn Monitoring Network stations, only the 13 stations listed in Table 4.3 have been used here since the remaining stations either had no available near-surface soil moisture data or significant data quality problems. The 13 stations that were used are located close together, with less than 50 km separation between them, and seven of the stations are within a 1 km^2 focus area at Stanley.

Preparation of in situ time series

The in situ soil moisture data from the above monitoring networks have been used in three separate comparisons below, each of which covers a different time period. In Section 4.3.2 the AMSR-E soil moisture products from each of the retrieval algorithms listed above have been assessed over 2006. Then in Section 4.3.3, the VUA-NASA soil moisture data have been evaluated in greater detail from 2003-2008. In Section 4.3.5 the VUA-NASA soil moisture has been benchmarked against the near-surface soil moisture from the ACCESS model, over the experimental period used in Chapter 6, from April 2008 to March 2009. To maximise the amount of data used, the in situ data have been separately processed for each of the above-listed comparisons. In each case:

Table 4.3: Goulburn Monitoring Network (installed in September 2002, with shallow soil moisture sensors installed in late 2005, data not available after 2007). Only stations used here are listed (there are another 13 stations within the network).

Code	Name	Latitude	Longitude	Notes
G1	MERRIWA REGION			
	Maram Park	-32.25	150.32	
	Cullingral	-32.17	150.33	
	Merriwa Park	-32.1	150.42	
G2	KRUI REGION			
	Illogan	-32.15	150.07	
	Roscommon	-32.17	150.15	
	Pembroke N	-31.98	150.18	
	Stanley	-32.1	150.13	7 stations in a 1 km ² area

- The in situ data were subsampled at the approximate time of the AMSR-E overpasses (1:30 am/pm local time), for comparison to the colocated AMSR-E pixel.
- For evaluation of near-surface soil moisture, only sensors within 8 cm of the surface were included.
- If data were available from an in situ sensor for less than 80% of the evaluation period, that sensors was excluded.
- Where there is more than one in situ sensor in an AMSR-E pixel, the average has been calculated (after checking that all the stations reflect similar temporal dynamics). The mean was not calculated at a given time if any of the selected stations were missing data (to prevent the station availability from influencing the mean).

4.2.3 Methods to compare in situ and remotely sensed soil moisture

The AMSR-E soil moisture products have been principally assessed by comparison to time series of near-surface soil moisture observations from in situ monitoring stations. However, as was reviewed in Section 2.4.3, the soil moisture

quantities observed by in situ and remote sensors are fundamentally different (see also Section 4.3.1), and comparisons between them are best based on their temporal dynamics, rather than their absolute values. Consequently, to focus on comparing the temporal behaviour of the AMSR-E and in situ soil moisture time series, the AMSR-E data have been normalised to match the statistics of the in situ time series. This is consistent with the use of remotely sensed soil moisture in data assimilation, since the data are typically rescaled to the model climatology prior to assimilation (to account for the model-observation biases).

Each AMSR-E observation (θ_A) has been normalised to match the mean ($\bar{\theta}$) and variance ($s(\theta)^2$) of the in situ data (θ_i), according to:

$$\theta'_A = (\theta_A - \bar{\theta}_A) \times (s(\theta_i)/s(\theta_A)) + \bar{\theta}_i \quad (4.1)$$

For each AMSR-E pixel the colocated in situ data are unlikely to reflect the absolute value of the pixel-average soil moisture. For example, it will be demonstrated in Section 4.3.1 that there is considerable spread between the absolute values from different monitoring stations within an AMSR-E pixel (see Figure 4.2). Consequently, the inter-pixel differences in the in situ data do not necessarily represent the expected inter-pixel differences in the remotely sensed data, and the in situ data cannot be sensibly used to assess the spatial variation in the remotely sensed data. The normalisation has then been applied independently at each AMSR-E pixel, and the spatial patterns in the in situ and remotely sensed data have not been compared.

Since the comparison between the in situ and remotely sensed soil moisture seeks to establish the degree of association between them, the (Pearson product moment) correlation coefficient (r) is the main statistic used here. The correlations between the absolute values of each time series, r_{abs} , as well as their anomaly correlations, r_{anm} , have been calculated. The r_{anm} is calculated from the anomaly between the soil moisture on each day and the surrounding 31 day moving average. r_{anm} measures the agreement in the short-term variability of each time series, including the response to individual rain events, while r_{abs} also measures agreement between their seasonal cycles. For many data assimilation applications, r_{anm} is a better metric for assessing remotely sensed data, since data are typically normalised to match the model climatology before assimilation, removing much of the seasonal scale information (for an extreme example, see Chapter 5).

The significance of the estimated sample correlations has been tested against the null hypothesis that the population correlation (ρ) is zero, using the test statistic:

$$t = r \times \frac{\sqrt{N-2}}{\sqrt{1-r^2}} \quad (4.2)$$

Where t has a Student's t -distribution, with $N - 2$ degrees of freedom for a sample size of N (Dawdy and Matalas, 1964).

In several instances the difference between two estimated sample correlations has been tested to determine whether one data set yields significantly higher correlations than the other. Since the above statistic has a t -distribution only for ρ equal to zero, it cannot be used to test the significance of the difference between two estimates of (nonzero) ρ , and instead, a Fisher r to z transform has been used (Yevdjovich, 1964):

$$z = 1/2 \times \ln[(1+r)/(1-r)] \quad (4.3)$$

Where z is (approximately) normal, with standard error $1/\sqrt{N-3}$. The significance of the difference between the two correlation estimates has been tested using the difference between the two z estimates (which has a normal distribution, with $\sqrt{1/(N_1-3) + 1/(N_2-3)}$ standard error).

The above statistics assume that the sample correlations have been estimated from random samples, yet soil moisture time series are not random, and exhibit strong serial correlation. For example, for the Adelong time series used in Section 4.3.2, the lag-1 auto-correlations of the in situ and VUA-NASA C-band soil moisture were 0.964 and 0.774, respectively. Serial correlation reduces the number of independent data in the sample, generating a bias in the statistical inference, although not in the estimation of the correlation itself (Yevdjovich, 1964). To test the significance of the correlations between two serially correlated time series (with lag-1 auto-correlations of r_x and r_y), the sample size, N , must be reduced to account for the lack of independence between the samples. Soil moisture time series can be approximated by a first-order Markov processes (Vinnikov and Yeserkepova, 1991), and for first-order Markov processes, an "effective sample size", N_{eff} , can be estimated from (Dawdy and Matalas,

1964):

$$N_{eff} = N(1 - r_x r_y) / (1 + r_x r_y) \quad (4.4)$$

N_{eff} has been used in all statistical inference for correlation between soil moisture time series. This reduces the sample size to about 10 - 30% of the number of observations. For example, for the VUA-NASA C-band descending pass and in situ data at Adelong over 2006, N is reduced from 293 to an N_{eff} of 43.

4.2.4 Precipitation data

In the absence of in situ soil moisture observations for most of Australia, maps of AMSR-E derived soil moisture have been qualitatively compared to precipitation maps to check the realism of the continental scale spatial patterns in the AMSR-E data. Since precipitation is the dominant forcing of soil moisture at atmospheric time scales, a strong spatial correlation is expected between near-surface soil moisture and precipitation. The precipitation maps have been prepared from the BoM's daily rain-gauge analysis (Weymouth et al., 1999), which analyses daily precipitation observations (to 9 am) from approximately 6000 rain-gauges across Australia onto a 0.25° grid.

4.3 Results

4.3.1 In situ soil moisture as a proxy for the area-average

As reviewed in Section 2.4.3 remote sensors observe an area-average soil moisture, typically with a resolution of tens of km, while in situ sensors observe the soil moisture at a point, and yet soil moisture variability is controlled by different processes across these two scales (Vinnikov et al., 1999; Entin et al., 2000; Robock et al., 2000). To highlight the importance of this issue, and specifically its consequences for comparing in situ and area-average (modeled or remotely sensed) soil moisture estimates, the variation between the in situ soil moisture observations from different monitoring stations within an AMSR-E pixel is examined here. Figure 4.2 shows time series of the in situ observations from the five monitoring stations at Kyeamba that will be used to estimate the Kyeamba-A pixel average for comparison to the ACCESS model soil moisture in Section

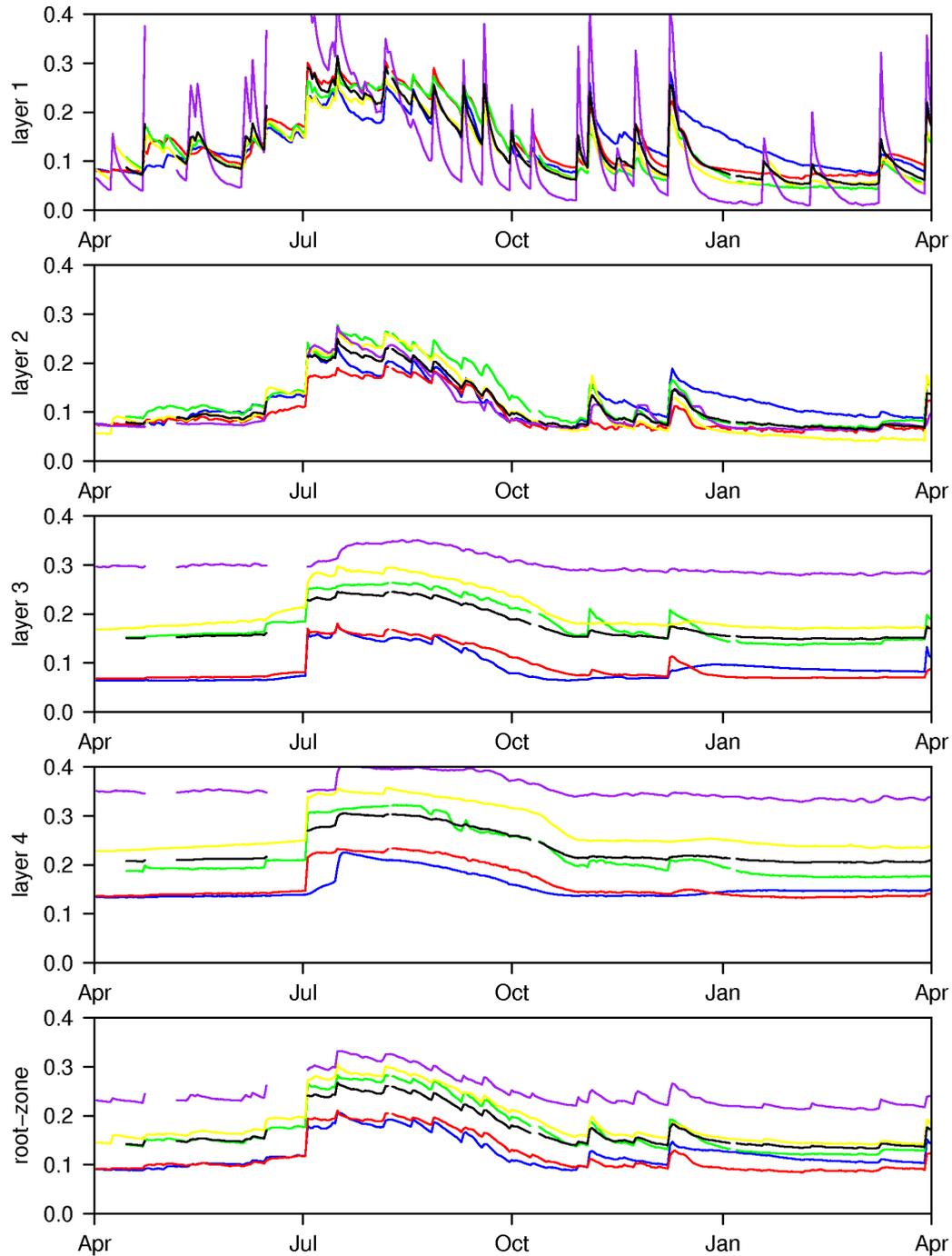


Figure 4.2: Time series of soil moisture observations (m^3m^{-3}) for each layer over 2008, from five in situ soil moisture sensors at Kyeamba: K1 (blue), K2 (red), K3 (green), K5 (yellow), and K7 (purple), together with their mean (black).

4.3.5. In layer one, the timing of the precipitation signal was similar at all stations, since they were all subject to similar (large scale) atmospheric forcing. However the precipitation responses recorded at each station were different, in terms of the magnitude and persistence of the precipitation induced spikes, due to (small scale) variability between the land surface characteristics at the different stations. For example, K7 showed a greater response and more rapid dry-down after each rain event. At longer time scales, the different stations had similar seasonal cycles, although there were still differences. Most notably, K7 dried more rapidly after winter, and K1 was elevated relative to the other stations from August onwards. Table 4.4 presents statistics describing the fit between the layer one soil moisture observations from all possible combinations of the stations plotted in Figure 4.2, using the same statistics that will be used to compare the AMSR-E and in situ soil moisture observations. For these different combinations r_{abs} ranged between 0.71 - 0.97, and r_{anm} ranged between 0.69 - 0.92, while the RMSD was between 0.021 - 0.081 m^3m^{-3} . These statistics provide an indication of the upper performance limit that might be expected from comparing time series of individual point-based observations of soil moisture to

Table 4.4: RMSD, correlation, and anomaly correlation between layer one soil moisture from the individual monitoring stations in the Kyeamba-A pixel, for April '08 to April '09. For each statistic the minimum and maximum off-diagonal values are in bold.

	RMSD (m^3m^{-3})				
	K1	K2	K3	K5	K7
K1	0.000	-	-	-	-
K2	0.036	0.000	-	-	-
K3	0.045	0.021	0.000	-	-
K5	0.039	0.028	0.021	0.000	-
K7	0.081	0.073	0.075	0.067	0.000

	r_{abs}						r_{anm}				
	K1	K2	K3	K5	K7		K1	K2	K3	K5	K7
K1	1.00	-	-	-	-	K1	1.00	-	-	-	-
K2	0.85	1.00	-	-	-	K2	0.84	1.00	-	-	-
K3	0.81	0.97	1.00	-	-	K3	0.85	0.89	1.00	-	-
K5	0.81	0.96	0.97	1.00	-	K5	0.81	0.84	0.92	1.00	-
K7	0.71	0.75	0.72	0.81	1.00	K7	0.70	0.69	0.79	0.87	1.00

area-averages. Additionally, previous studies have indicated that the expected RMSD between soil moisture at the point scale and the typical satellite footprint scale is closer to the upper-limit of the RMSD values obtained here: by comparison of a very large (36,000) number of in situ observations Famiglietti (2008) obtained a RMSD estimate of $0.07 \text{ m}^3\text{m}^{-3}$, while Miralles et al. (2010) used triple collocation to compare soil moisture from remote sensing, an in situ sensor, and a land surface model, and obtained an estimate of $0.06 \text{ m}^3\text{m}^{-3}$.

Soil moisture observations from the nonsurface soil layers have not been used in this chapter, however they will be used in Chapter 6, and so the in situ observations throughout the soil profile are also considered here. Comparing the soil moisture time series from each observation depth in Figure 4.2, demonstrates that the rapid fluctuations at the surface are filtered as the surface moisture infiltrates through the soil profile. The rate of infiltration will depend on the (small scale) soil properties. While there is good agreement between the layer two soil moisture time series from the different monitoring sites, there is substantial spread in their absolute values in layers three and four ($> 0.2 \text{ m}^3\text{m}^{-3}$ in some cases). For the two deepest layers there was good agreement in the main temporal dynamics at each station, although there were occasional diversions from this. For example, K1 and K7 show a delayed response to the rain in July, despite both having responded to the same rain events in layer one. Also layers three and four at K2 and K3 show a signal of two rain events in August and September that are not evident in these layers at the other sites.

The statistics of fit between the different in situ observations (equivalent to Table 4.4) for the nonsurface layers are included in Appendix C. Since the RMSD is influenced by biases, it decreases from layers one to two, before increasing in layers three and four. The r_{abs} was consistently very high in each layer, and in most instances it increased with depth (since short-term variability is decreased), and was between 0.86 and 0.99 in layer four. In contrast r_{anm} decreased with depth, to the extent that it was negative at M7 in layers three and four, and also at M1 in layer four, due to the different temporal behaviour in these layers noted above. In layer four r_{abs} was between -0.48 to 0.97. These results suggest that in situ soil moisture observations are of limited value for assessing the anomaly correlation (arguably the most important of the evaluation statistics used here) of large scale soil moisture estimates in individual layers below the surface.

The lowest panel in Figure 4.2 shows the estimated root-zone soil moisture

at each station, calculated by aggregating the observations in all layers (giving an observation of the uppermost 90 cm). There is still considerable spread in the absolute root-zone soil moisture at the different monitoring sites, however the similarity in the temporal behaviour at each site appears stronger than in any of the individual layers. Consequently, in nearly all instances r_{abs} for the root-zone soil moisture is better than for the individual layers, with a range of 0.89 to 0.99. The r_{annm} is also improved in most instances, compared to the values from individual soil layers. In particular, K1 and K7 do not have negative r_{annm} as had occurred in some of the individual layers, although the values are still rather low for K7 (between 0.39 and 0.58). At the remaining sites r_{annm} is consistently high, and is generally higher than in the individual layers, and is between 0.88 and 0.93. From these results, and particularly the improved r_{annm} statistics, it is concluded that evaluation of large scale estimates of nonsurface soil moisture using in situ observations should be based on the bulk soil moisture over the root-zone (or some other depth of the order of 1 m). Additionally, the poor agreement between the nonsurface soil moisture from K7 and the other monitoring sites, even after a bulk root-zone estimates are compared, highlights that in some instances the soil moisture from an individual monitoring site is simply not representative of the area-averaged behaviour, and comparisons between in situ and area-averaged soil moisture estimates should be interpreted with this in mind.

4.3.2 Inter-comparison of the AMSR-E soil moisture retrievals

Soil moisture from each retrieval algorithm has been compared to in situ observations over 2006 for the descending AMSR-E overpass. After screening and processing the in situ data as described in Section 4.2.2, in situ time series were obtained for the 11 AMSR-E pixels listed in Table 4.5. For the Murrumbidgee network the second generation sites did not yet have near-surface soil moisture sensors, A2 was no longer operating, and M3 had no data for the last three months of 2006, and these stations were excluded. The remaining stations included five (K1, K2, K3, K4, K5) at Kyeamba (M9), four (A1, A3, A4, A5) at Adelong (M10), and eight others throughout the Murrumbidgee basin. For the Goulburn monitoring network, there were three stations within a single pixel at Merriwa (G1), and ten within a nearly adjacent pixel at Krui (G2; see Ta-

Table 4.5: Descriptive statistics for the soil moisture time series retrieved from descending AMSR-E overpass brightness temperatures by each retrieval algorithm and the in situ data, for 2006. Three bias values are calculated using only those days for which the AMSR-E data are available for that algorithm.

	in situ		VUA-NASA C		VUA-NASA X	
	No.	Mean (bias) Stddev.	No.	Mean (bias) Stddev	No.	Mean (bias) Stddev
M1	352	0.12 (0.034)	260	0.28 (0.16) 0.129	260	0.22 (0.10) 0.097
M2	352	0.08 (0.030)	267	0.23 (0.15) 0.122	267	0.03 (-0.04) 0.066
M4	360	0.11 (0.060)	268	0.11 (0.00) 0.102	268	0.13 (0.02) 0.102
M5	360	0.09 (0.027)	267	0.14 (0.05) 0.088	267	0.18 (0.09) 0.084
M6	360	0.07 (0.040)	267	0.11 (0.04) 0.094	267	0.15 (0.07) 0.084
M7	360	0.09 (0.025)	267	0.12 (0.02) 0.094	267	0.14 (0.05) 0.093
M8	360	0.09 (0.040)	267	0.15 (0.06) 0.083	267	0.16 (0.07) 0.086
M9	327	0.13 (0.039)	271	0.12 (0.00) 0.109	271	0.11 (-0.013) 0.112
M10	360	0.12 (0.035)	293	0.23 (0.12) 0.123	293	0.11 (-0.014) 0.109
G1	365	0.05 (0.050)	252	0.19 (0.14) 0.082	252	0.16 (0.12) 0.088
G2	365	0.10 (0.081)	270	0.20 (0.10) 0.083	270	0.16 (0.06) 0.089

	NASA		JAXA		USDA	
	No.	Mean (bias) Stddev	No.	Mean (bias) Stddev	No.	Mean (bias) Stddev
M1	259	0.13 (0.01) 0.015	194	0.13 (0.02) 0.048	254	0.11 (-0.01) 0.055
M2	205	0.14 (0.05) 0.017	230	0.05 (-0.03) 0.012	265	0.07 (-0.01) 0.051
M4	244	0.08 (-0.03) 0.023	224	0.07 (-0.04) 0.026	243	0.13 (0.02) 0.114
M5	247	0.07 (-0.02) 0.016	229	0.08 (-0.01) 0.033	258	0.17 (0.08) 0.090
M6	266	0.05 (-0.02) 0.015	230	0.06 (-0.01) 0.023	267	0.20 (0.13) 0.154
M7	266	0.09 (0.00) 0.021	190	0.06 (-0.03) 0.019	264	0.16 (0.06) 0.097
M8	266	0.10 (0.01) 0.017	203	0.07 (-0.01) 0.026	266	0.19 (0.10) 0.103
M9	270	0.12 (0.00) 0.021	150	0.13 (0.00) 0.106	270	0.12 (0.00) 0.096
M10	273	0.13 (0.01) 0.019	163	0.10 (-0.01) 0.042	280	0.12 (0.00) 0.061
G1	247	0.15 (0.10) 0.016	190	0.13 (0.08) 0.043	247	0.13 (0.08) 0.046
G2	270	0.14 (0.04) 0.015	174	0.13 (0.03) 0.043	268	0.13 (0.03) 0.051

ble 4.3), including seven at Stanley. To prevent local conditions in the Stanley focus area from dominating the pixel average at Krui, the average of the seven Stanley stations was treated as a single observation when calculating the Krui pixel average time series.

The AMSR-E soil moisture time series have been compared to the in situ time series over 2006, both before and after normalising the AMSR-E data with equation 4.1. Examples of the original soil moisture time series are given in Figure 4.3 for Kyeamba in the Murrumbidgee network, and in Figure 4.4 for Krui in the Goulburn network; plots for the remaining monitoring stations are in Appendix D. Additionally, the equivalent plots for the normalised time series are given in Figures 4.5 and 4.6, and Appendix E, respectively. Finally, descriptive statistics for each data set are provided in Table 4.5, and statistics describing the fit between the AMSR-E and in situ data are provided in Table 4.6.

There were substantial differences between the various AMSR-E soil moisture data sets, in terms of both their climatology and their agreement with the in situ data. The VUA-NASA and USDA time series had quite different climatologies to the in situ data. Both of these data sets had a greater seasonal amplitude than the in situ data (see Figures 4.3 and 4.4), and the VUA-NASA data sets also had more short-term variability. As a result, the standard deviations in Table 4.5 for VUA-NASA (range: C-band $0.08 - 0.13 \text{ m}^3\text{m}^{-3}$, X-band $0.07 - 0.11 \text{ m}^3\text{m}^{-3}$) and USDA (range: $0.05 - 0.15 \text{ m}^3\text{m}^{-3}$) are consistently much larger than for the in situ data (range: $0.03 - 0.08 \text{ m}^3\text{m}^{-3}$). Also there were often large biases in the VUA-NASA (max C-band: $0.16 \text{ m}^3\text{m}^{-3}$ at M1, X-band 0.12 at G1) and USDA (max: $0.13 \text{ m}^3\text{m}^{-3}$ at M6) data sets. In contrast, the NASA and JAXA data sets had smaller biases (max: NASA $0.05 \text{ m}^3\text{m}^{-3}$ at M2, JAXA $-0.04 \text{ m}^3\text{m}^{-3}$ at M4), and lower standard deviations, particularly for NASA, which consistently had standard deviations lower than the in situ data (range: $0.01 - 0.02 \text{ m}^3\text{m}^{-3}$). The temporal coverage differed for the AMSR-E data sets, and was particularly low for JAXA, with a mean of 198 days (compared to 268 days for VUA-NASA (C- and X-band), 262 days for USDA, and 256 days for NASA). It is not known why the JAXA coverage was lower, however it is likely due to technical problems, lack of algorithm convergence, or quality control procedures.

At the Murrumbidgee sites, the VUA-NASA data sets qualitatively reflected the temporal behaviour of the in situ time series better than the other retrieval

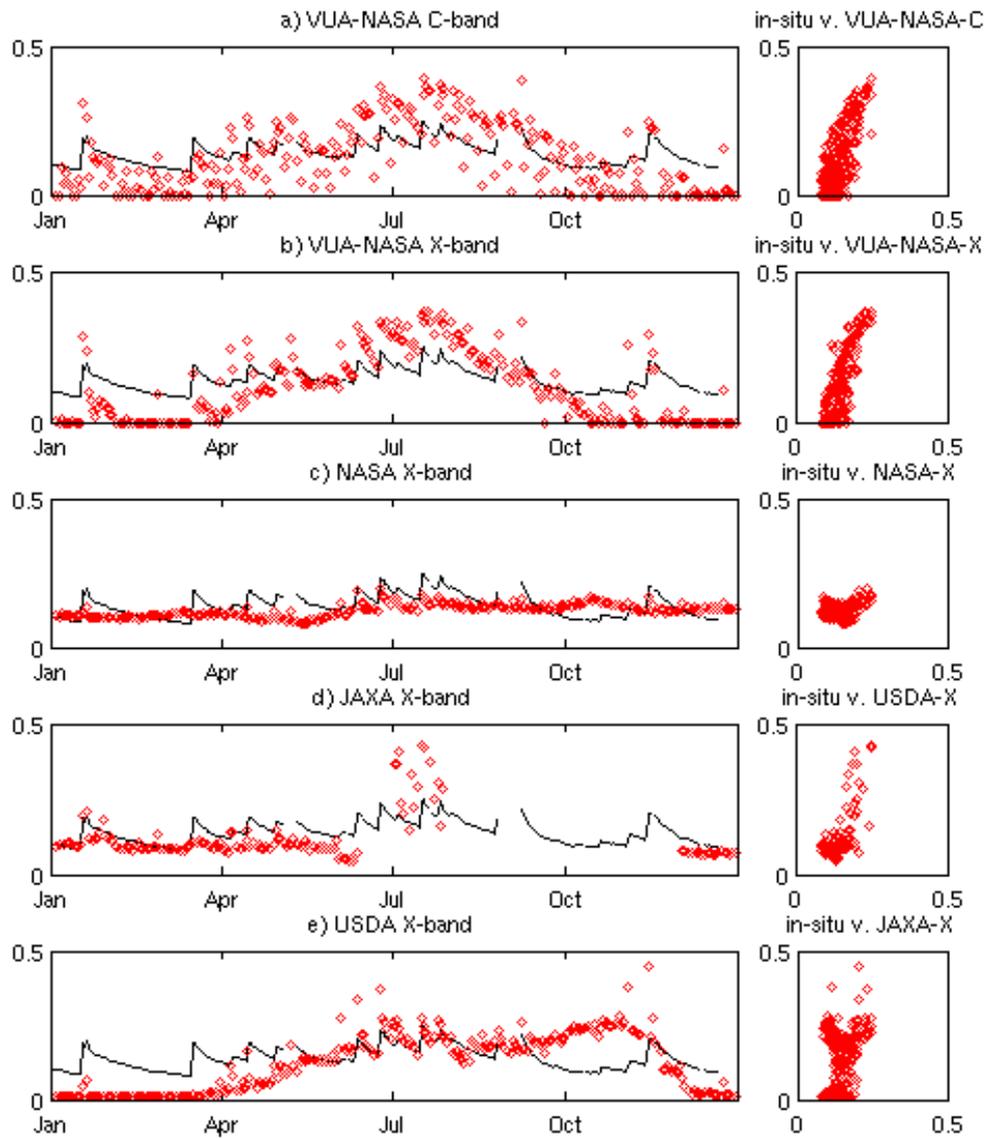


Figure 4.3: Comparison of in situ (solid lines) and original (not normalised) AMSR-E (red diamonds) near-surface soil moisture (m^3m^{-3}) for each retrieval algorithm, at Kyeamba (M9) over 2006.

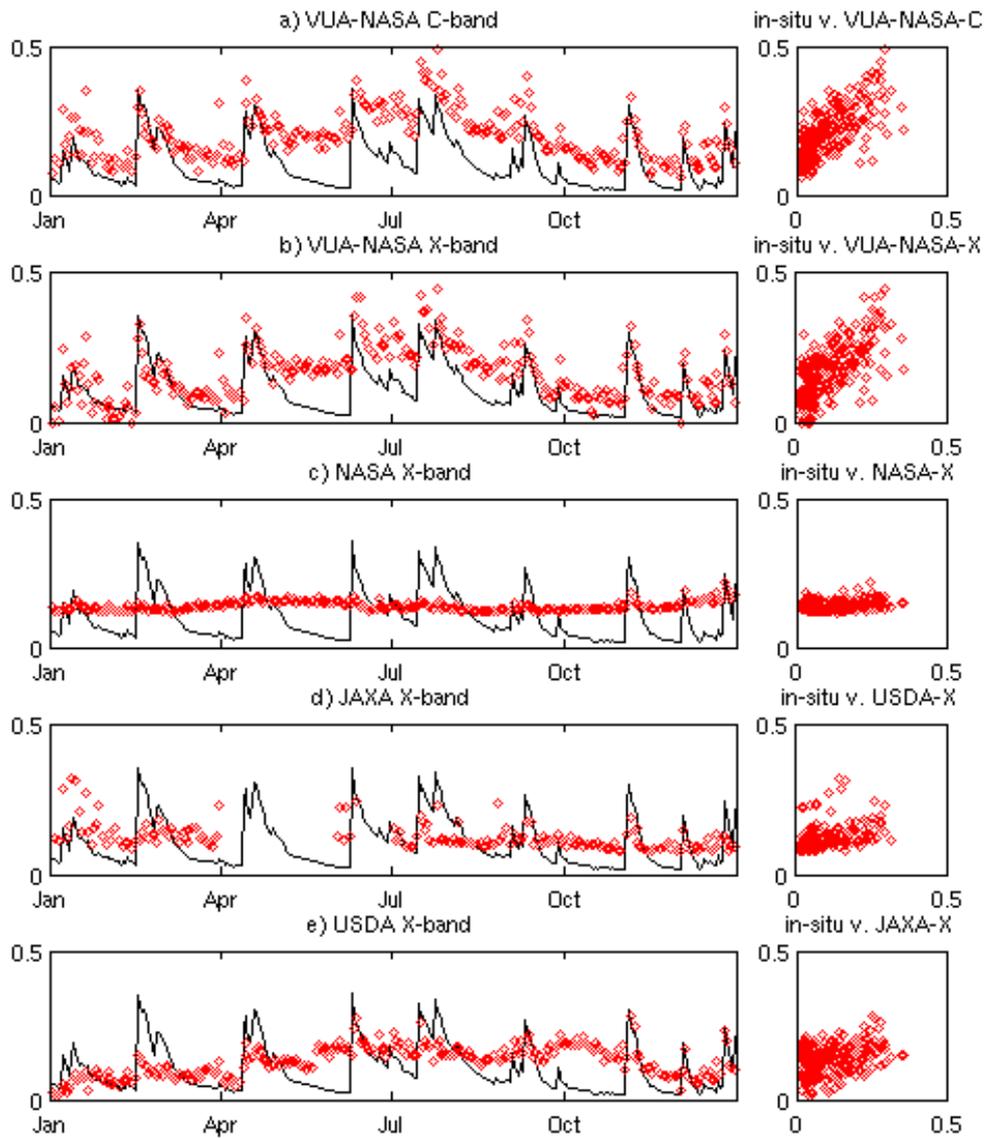


Figure 4.4: Comparison of in situ (solid lines) and original (not normalised) AMSR-E (red diamonds) near-surface soil moisture (m^3m^{-3}) for each retrieval algorithm, at Krui (G2) over 2006.

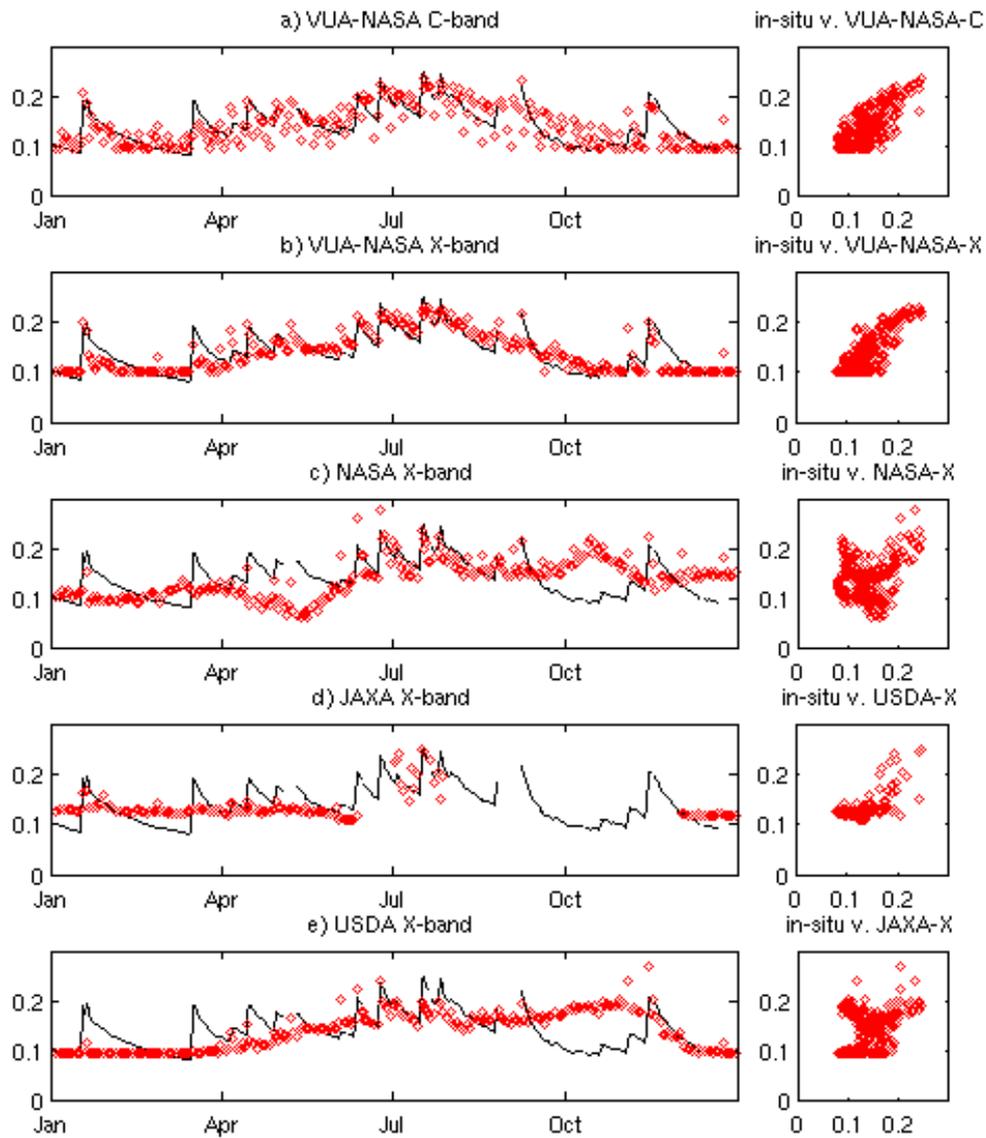


Figure 4.5: Comparison of in situ (solid lines) and normalised AMSR-E (red diamonds) near-surface soil moisture (m^3m^{-3}) for each retrieval algorithm, at Kyeamba (M9) over 2006.

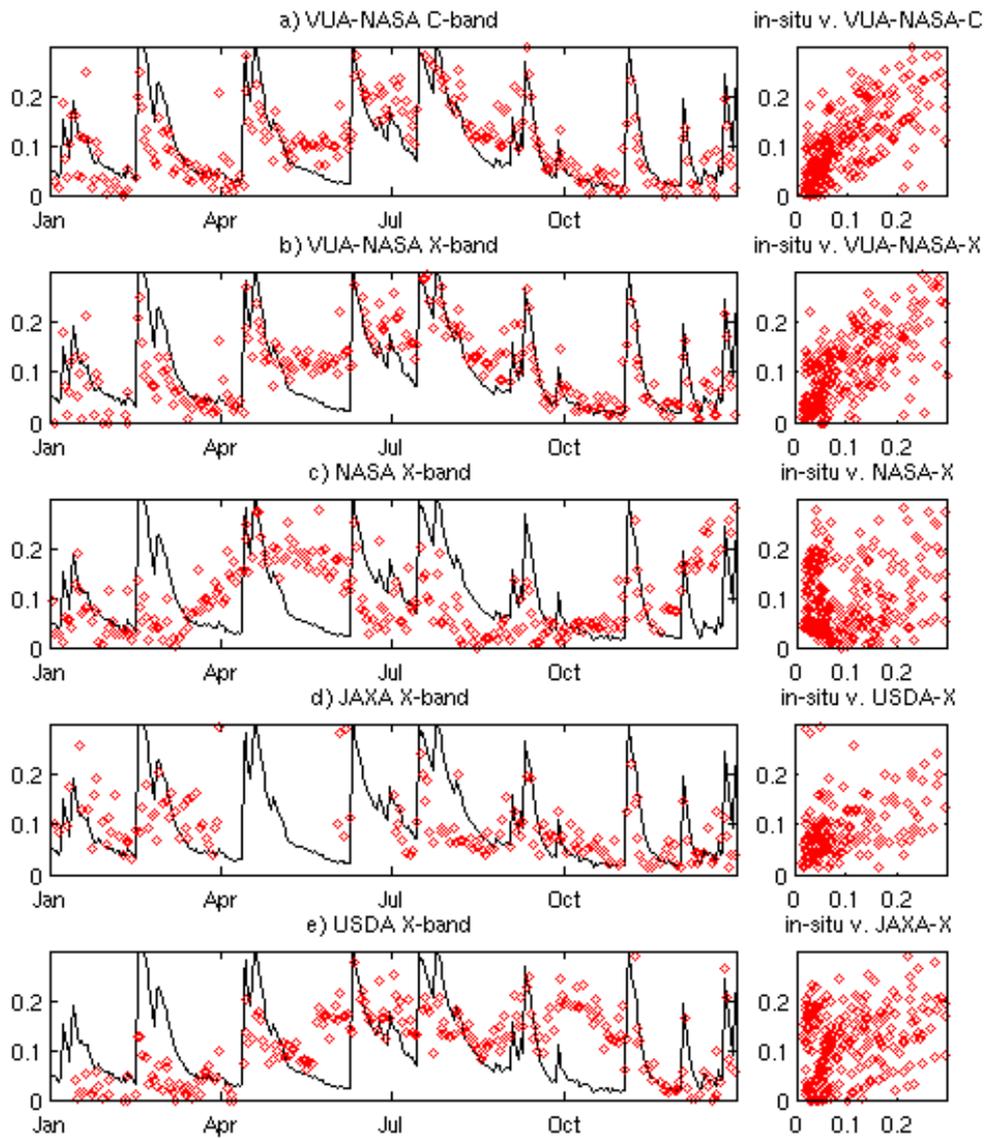


Figure 4.6: Comparison of in situ (solid lines) and normalised AMSR-E (red diamonds) near-surface soil moisture (m^3m^{-3}) for each retrieval algorithm, at Krui (G2) over 2006.

algorithms did. The timing of the seasonal cycle, and the response to precipitation, in both the C- and X-band VUA-NASA data sets agreed closely with the in situ time series in each of the Murrumbidgee plots (see Figure 4.5 and Figures E.1 to E.8 in Appendix E). In contrast, the other data sets often diverged from the in situ data, and in most cases they did not capture the timing of the seasonal cycle. Additionally, there is a clear monotonic relationship between scatterplots of the VUA-NASA and in situ soil moisture in Figure 4.3 and Appendix D. In these plots the gradient of the scatterplots is steeper than the 1:1 line, indicating the VUA-NASA retrievals were more sensitive to changes in soil moisture than the in situ data (recall the enhanced seasonal cycle in the VUA-NASA time series). It is interesting to note that the gradient at most sites was slightly steeper for drier soils, which is likely a consequence of the shallower depth of the AMSR-E observations, since soil drying occurs more rapidly close to the surface in dry conditions (Li and Islam, 2002). At M4 and M6 the VUA-NASA scatterplots showed a clear departure from linearity, with more than one linear regime over the year.

The statistics in Table 4.6 confirm that the VUA-NASA retrievals had the best fit to the in situ data across the Murrumbidgee sites. For both VUA-NASA data sets r_{abs} was significant at all sites, including values of 0.77 (C-band) and 0.84 (X-band) for Figure 4.3 at Kyeamba, and a range of 0.69 - 0.91 (C-band) and 0.61- 0.91 (X-band) across the Murrumbidgee sites. In contrast the other data sets did not have consistent monotonic relationships to the in situ data, giving lower r_{abs} of between -0.03 - 0.57 for NASA, 0.07 - 0.77 for JAXA, and 0.30 - 0.75 for USDA (including values of 0.29, 0.64, and 0.39, respectively at Kyeamba). Additionally, r_{abs} was not always significant for the other data sets, particularly for NASA and USDA.

For all of the AMSR-E data sets the agreement between the AMSR-E and in situ soil moisture was not as good at the Goulburn sites as it was at the Murrumbidgee sites, although the VUA-NASA data sets again had the best agreement. At both Goulburn sites (see Figure 4.6, and Figure E.9 in Appendix E for Merriwa) the in situ time series consisted of a series of rain-induced peaks, with low base values in between. Both VUA-NASA time series detected the timing of the rain-induced maxima well, although the maxima themselves were underestimated. Also in the first half of the year the VUA-NASA time series drifted upwards in between the precipitation events (by $\sim 0.15 \text{ m}^3 \text{ m}^{-3}$ in Figure 4.4), while the in situ data remained low. There is a reasonably linear relationship in

the scatterplots comparing the VUA-NASA and in situ data at the Goulburn sites, although the AMSR-E data has more spread (C-band: $r_{abs}=0.73$, 0.73 , X-band: $r_{abs}=0.71$, 0.69 , at Krui and Merriwa respectively). The remaining algorithms did not agree with the in situ time series as well. Before normalisation, the NASA time series had little detectable variability in Figure 4.6 ($r_{abs} = 0.36$, 0.22), while JAXA ($r_{abs} = 0.46$, 0.27) and USDA ($r_{abs} = 0.43$, 0.35) both detected some, although not all, of the rain induced soil moisture maxima.

The anomaly correlation statistics also indicate that the VUA-NASA data sets had the best agreement with the in situ time series. Scatterplots comparing the AMSR-E and in situ anomalies at Kyeamba are shown in Figure 4.7 (the C-band VUA-NASA plot is not included since it similar to the X-band VUA-NASA plot). For all of the retrieval algorithms the data are clustered about the origin, however the data points further from the origin for both VUA-NASA ($r_{anm}=0.62$) and USDA ($r_{anm}=0.40$) show a weak linear relationship. In contrast NASA ($r_{anm}=0.29$) and JAXA ($r_{anm}=0.24$) had a much weaker relationship. Across all of the sites r_{anm} ranged between 0.34 - 0.77 for C-band VUA-NASA, and from 0.26 - 0.76 for X-band VUA-NASA, compared to 0.32 - 0.68 for USDA, 0.09 - 0.51 for JAXA, and 0.21 - 0.54 for NASA. While r_{anm} was significant at almost all locations for all retrieval algorithms, the correlations were generally quite small, indicating a significant but weak relationship between the remotely sensed and in situ data. Except for the Goulburn sites and for the NASA data set, r_{anm} was generally less than r_{abs} . For VUA-NASA, r_{anm} was about 0.3 lower than r_{abs} at each Murrumbidgee site, indicating that a substantial portion of r_{abs} between the VUA-NASA and in situ soil moisture was due to their similarities at time scales beyond one month. In contrast, at the Goulburn sites, where none of the retrievals detected the seasonal cycle accurately, the calculation of the anomalies improved the agreement between the in situ and remotely sensed data, and r_{anm} was slightly higher than r_{abs} .

The RMSD has also been included in Table 4.6 to give a measure of the absolute agreement between the AMSR-E and in situ time series. However, for the original AMSR-E time series the RMSD does not necessarily measure the accuracy of the AMSR-E observations, since it will be influenced by biases, which are as likely due to the mismatch in the horizontal and vertical scales of the two different observations types as to errors in the remotely sensed data (recall the spread between the different in situ stations in Figure 4.2 for a

Table 4.6: Statistics of fit between the soil moisture time series retrieved from descending AMSR-E overpass brightness temperatures by each retrieval algorithm, and the in situ data for 2006. The RMSD is provided for both the original and normalised AMSR-E time series. Bold correlations indicates significance at the 1% level.

	VUA-NASA C		VUA-NASA X		NASA	
	r_{abs}	r_{amm}	r_{abs}	r_{amm}	r_{abs}	r_{amm}
M1	0.77	0.42	0.78	0.59	0.32	0.35
M2	0.79	0.36	0.61	0.26	0.28	0.29
M4	0.91	0.54	0.91	0.56	0.19	0.21
M5	0.77	0.49	0.79	0.47	0.55	0.49
M6	0.69	0.34	0.68	0.36	0.13	0.23
M7	0.84	0.45	0.85	0.55	-0.03	0.38
M8	0.87	0.63	0.88	0.63	0.57	0.33
M9	0.77	0.43	0.84	0.62	0.29	0.29
M10	0.81	0.47	0.81	0.48	0.08	0.20
G1	0.73	0.77	0.71	0.76	0.36	0.54
G2	0.73	0.71	0.69	0.71	0.22	0.48

	JAXA		USDA	
	r_{abs}	r_{amm}	r_{abs}	r_{amm}
M1	0.35	0.36	0.30	0.50
M2	0.07	0.09	0.63	0.39
M4	0.77	0.12	0.75	0.32
M5	0.56	0.27	0.72	0.42
M6	0.51	0.32	0.72	0.48
M7	0.73	0.17	0.69	0.49
M8	0.73	0.38	0.60	0.48
M9	0.64	0.24	0.39	0.40
M10	0.42	0.42	0.36	0.35
G1	0.46	0.51	0.43	0.68
G2	0.37	0.49	0.35	0.61

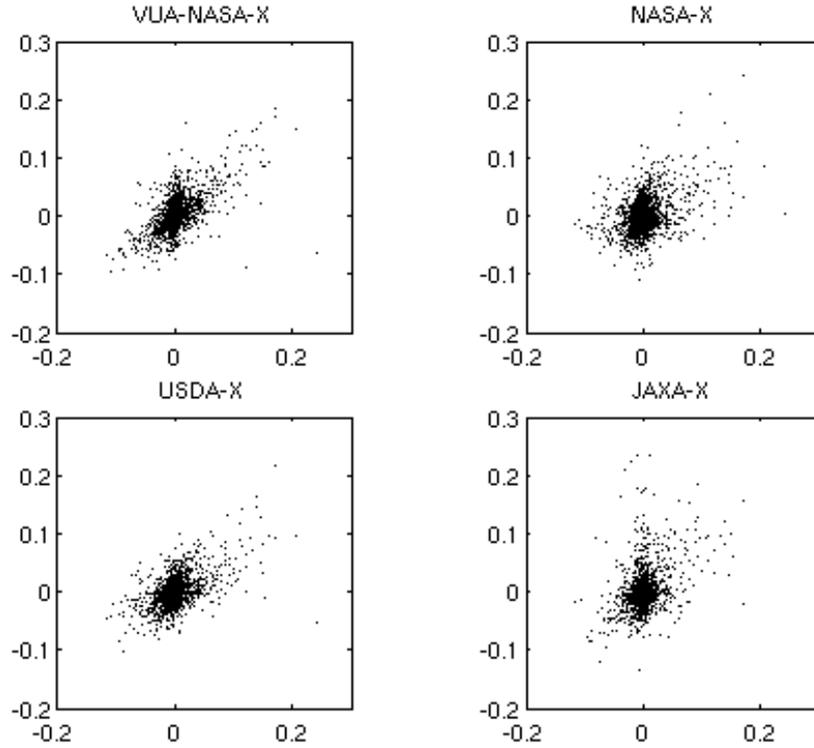


Figure 4.7: Scatterplots of the AMSR-E soil moisture anomalies vs. the in situ anomalies over 2006 at Kyeamba (both in m^3m^{-3}), for the X-band retrieval algorithms.

single AMSR-E pixel). The normalised RMSD must also be interpreted with caution, since it contains a signal of the variance of the in situ time series. Despite this the normalised RMSD is useful for inter-comparing the different remotely sensed data sets at each location. Again, the best results were obtained for VUA-NASA, and the two lowest normalised RMSD at each site were for the two VUA-NASA data sets, which had a value between $0.02 - 0.03 m^3m^{-3}$ at most Murrumbidgee sites, with larger values (up to $0.06 m^3m^{-3}$) at the Goulburn sites. The VUA-NASA normalised RMSD were also lower than the standard deviation of the in situ data at all sites¹. The other data sets had higher normalised RMSD statistics, of between $0.03 - 0.04 m^3m^{-3}$ at most Murrumbidgee sites, and higher again at the Goulburn sites. Additionally, the normalised RMSD for NASA was larger than the in situ standard deviation at most sites, while the JAXA and USDA retrieval algorithms had RMSD higher

¹The standard deviation gives the RMSD that would be obtained from a time series equal to the mean value at all times.

than the in situ standard deviation at about half the sites (interestingly these are mostly the same sites for both data sets: M1, M2, M10, G1, and G2 - and also at G9 for JAXA).

In summary, of the five AMSR-E data sets examined here the two VUA-NASA retrievals (C- and X- band) had the best fit to the in situ data from the Murrumbidgee and Goulburn Monitoring Networks over 2006. This finding is consistent with previous inter-comparisons between the VUA-NASA and NASA data sets; for example Wagner et al. (2007) and Rüdiger et al. (2009) in Table 2.1. All of the AMSR-E data sets had a stronger fit to the in situ observations from the Murrumbidgee than from the Goulburn Network. Qualitatively, both VUA-NASA data sets could detect the step-changes in the in situ data associated with precipitation events, and the correlation between the in situ data and the VUA-NASA time series were generally above 0.7 and consistently significant. While the other data sets often had significant correlations (with the exception of NASA), the portion of explained variance was less than for VUA-NASA. Additionally, the normalised RMSD between the AMSR-E and in situ soil moisture was smaller for the VUA-NASA data sets (except at M2 for VUA-NASA X-band). For brevity, results were presented here for the descending (nighttime) AMSR-E overpass only, however an assessment of the soil moisture retrieved from each algorithm for the ascending AMSR-E overpass produced the same conclusions.

Based on the above inter-comparison, the VUA-NASA product has been selected for use in the assimilation experiments in the following chapters. The remainder of this chapter is focused on understanding the capabilities of the VUA-NASA data sets, including the identification of any differences between the soil moisture estimated from the C- and X-band observations, and from the ascending and descending AMSR-E overpasses.

4.3.3 Assessment of VUA-NASA AMSR-E soil moisture

Temporal assessment against in situ observations

While one year of data was sufficient to establish the superior performance of the VUA-NASA retrieval algorithm, comparison over a longer period will allow a more robust evaluation of the data, and the comparison is now extended to a six year period. After screening the data as described in Section 4.2.2, in situ time series were obtained over 2003-2008 for the nine AMSR-E pixels in the

Table 4.7: Descriptive statistics for the VUA-NASA AMSR-E soil moisture and from the in situ data, for 2003–2008. The bias values are calculated using only those days for which the AMSR-E data are available.

	in situ			VUA-NASA C - Dsc.			VUA-NASA X - Dsc.		
	No.	Mean (m^3m^{-3})	Stdev. (m^3m^{-3})	No.	Mean (bias) (m^3m^{-3})	Stdev (m^3m^{-3})	No.	Mean (bias) (m^3m^{-3})	Stdev (m^3m^{-3})
M1	2163	0.12	0.031	1592	0.27 (0.15)	0.116	1593	0.23 (0.11)	0.090
M2	1959	0.09	0.030	1301	0.22 (0.13)	0.118	954	0.13 (0.03)	0.097
M4	2102	0.13	0.056	1426	0.15 (0.01)	0.086	1504	0.16 (0.03)	0.090
M5	2169	0.10	0.030	1611	0.16 (0.07)	0.085	1638	0.19 (0.09)	0.084
M6	2134	0.10	0.051	1448	0.16 (0.06)	0.091	1550	0.18 (0.08)	0.089
M7	2170	0.11	0.029	1578	0.16 (0.05)	0.078	1596	0.17 (0.07)	0.083
M8	2169	0.11	0.052	1568	0.17 (0.06)	0.090	1590	0.18 (0.07)	0.095
M9	2002	0.17	0.073	1268	0.20 (0.02)	0.118	1319	0.21 (0.03)	0.121
M10	2023	0.15	0.064	1596	0.29 (0.14)	0.143	1382	0.22 (0.05)	0.135

	in situ			VUA-NASA C - Asc.			VUA-NASA X - Asc.		
	No.	Mean (m^3m^{-3})	Stdev. (m^3m^{-3})	No.	Mean (bias) (m^3m^{-3})	Stdev (m^3m^{-3})	No.	Mean (bias) (m^3m^{-3})	Stdev (m^3m^{-3})
M1	2163	0.12	0.031	1700	0.27 (0.15)	0.072	1691	0.23 (0.11)	0.087
M2	1958	0.08	0.028	1349	0.24 (0.16)	0.124	1098	0.14 (0.05)	0.099
M4	2101	0.13	0.054	1515	0.16 (0.03)	0.061	1516	0.15 (0.02)	0.070
M5	2168	0.09	0.030	1607	0.19 (0.09)	0.053	1608	0.18 (0.09)	0.061
M6	2134	0.10	0.052	1535	0.16 (0.06)	0.070	1542	0.16 (0.06)	0.074
M7	2170	0.10	0.028	1577	0.17 (0.07)	0.049	1578	0.17 (0.07)	0.061
M8	2166	0.11	0.052	1573	0.18 (0.07)	0.058	1572	0.17 (0.06)	0.069
M9	2006	0.16	0.073	1420	0.23 (0.06)	0.092	1441	0.19 (0.03)	0.103
M10	2019	0.15	0.064	1463	0.32 (0.16)	0.117	1416	0.20 (0.04)	0.135

Murrumbidgee Monitoring Network listed in Table 4.7. The Goulburn data are available for less than two years, and were excluded, together with the second generation Murrumbidgee sites, which did not have near-surface soil sensors for most of the this period. Additionally A2 and M3 were also excluded due to insufficient data coverage. Of the remaining stations five (K1, K2, K3, K4, K5) are at Kyeamba (M9), four (A1, A3, A4, A5) are at Adelong (M10), and six others are spread throughout the Murrumbidgee basin.

Descriptive statistics for the in situ and VUA-NASA time series at each site are provided in Table 4.7, for both the ascending and descending AMSR-E overpass, and for the C- and X-band observations. The coverage of the AMSR-E data was unusually poor at M2, where the lowest coverage over 2003-2008 was 49% for the X-band descending overpass (and the highest was 69% for the C-band ascending overpass), compared to a mean of 70% for the descending overpass and 71% for the ascending overpass. This low coverage was likely due to localised RFI, as M2 is at a major airport. The slightly reduced coverage of the in situ data at M2 (close to 90%) is unrelated and was due to an equipment failure in mid-June 2008.

The in situ and original VUA-NASA time series are plotted in Figure 4.8 for Kyeamba, and in Appendix F for the remaining sites. The most striking features of these plots is that the VUA-NASA data sets had much more short-term variability than the in situ data, and also a more exaggerated seasonal cycle, particularly for the descending overpass. The short-term variability will be partly due to the shallower soil layer observed by AMSR-E (1 cm for C-band, and less for X-band) compared to the in situ sensors (5 - 8 cm), since the shallow surface layer will respond more rapidly to atmospheric forcing. However, it is likely that a substantial portion of the short-term variance in the AMSR-E data is from noise in the observations. As well as the usual sampling uncertainties of remotely sensed data, the method used to map the swath data onto the 0.25° grid introduces small errors. The AMSR-E value for each grid cell is the mean of all (NSIDC Level 1) swath data for which the foot-print is centred on that grid cell (Owe et al., 2008). Due to the progression of the AMSR-E orbit, the land area (and hence soil moisture) contributing to each grid cell varies from day to day, with a 16-day cycle.

The normalised VUA-NASA time series for Kyeamba are plotted in Figure 4.9, and the remaining Murrumbidgee sites are plotted in Appendix G. These plots highlight that in general the AMSR-E time series predicted the timing and

relative amplitude of the in situ seasonal cycle in each year, with the exception of M6, where the onset and relative amplitude of the winter maxima differ. There were other differences between the two, including the tendency for the AMSR-E time series to drift upwards in the spring of 2005, when the in situ data remained low. Due to these differences r_{abs} and the normalised RMSD for the 6 year period in Table 4.8 suggest a poorer fit than the 2006 statistics in Table 4.6. For the descending overpass, r_{abs} over the 6 years was about 0.1 to 0.2 lower than over 2006, and ranged between 0.53 - 0.82 for C-band, and 0.52 - 0.80 for X-band. The ascending overpass had similar r_{abs} , of between 0.51 - 0.79 for C-band, and 0.52 - 0.79 for X-band. For both overpasses all estimated correlations were significant at 1%, and the RMSD were again lower than the in situ standard deviation at all sites (except M2, for the X-band descending overpass). In contrast to r_{abs} , the anomaly correlations for the 6 year period and for 2006 were similar, since r_{anm} is robust to interannual differences in seasonal behaviour. For the descending overpass r_{anm} was between 0.31 - 0.65 for C-band, and 0.29 - 0.65 for X-band, and the ascending overpass values were similar, between 0.27 - 0.55 for C-band, and 0.30 - 0.63 for X-band.²

C-band passive microwave data are expected to yield more accurate soil moisture than the shorter-wavelength X-band data. However, it is difficult to discern a consistent difference between them in the plots in Figure 4.9 and Appendix G. In Table 4.7 the mean soil moisture (and hence the biases) in the C- and X-band AMSR-E data differ, most notably at M2 and M10. These differences suggest possible errors, since the C- and X-band observations occur at the same time and use the same retrieval parameters. However, they could also be due the different depth of soil observed by each wavelength. The statistics in Table 4.8 slightly favour the C-band overpass for the descending AMSR-E overpass only. For the descending overpass, r_{abs} was higher and the RMSD was lower for the C-band data at seven of the nine sites, but for the ascending overpass the C-band time series correlations were higher and the RMSD was lower at just two sites, with another two that were equal. For both overpasses, the difference between the correlations for the C- and X-band overpasses was not

²With the longer data record available in this section, the anomaly correlation could have been reported relative to the climatological mean 31 day moving average (rather than the 31 day mean for the present year only). This method was tested, and yielded slightly higher r_{abs} across the Murrumbidgee, indicating that in general the in situ and remotely sensed soil moisture data have similar interannual variability in their seasonal cycles. However, this method has not been reported here for consistency with the remainder of the thesis, where only one year of data is available.

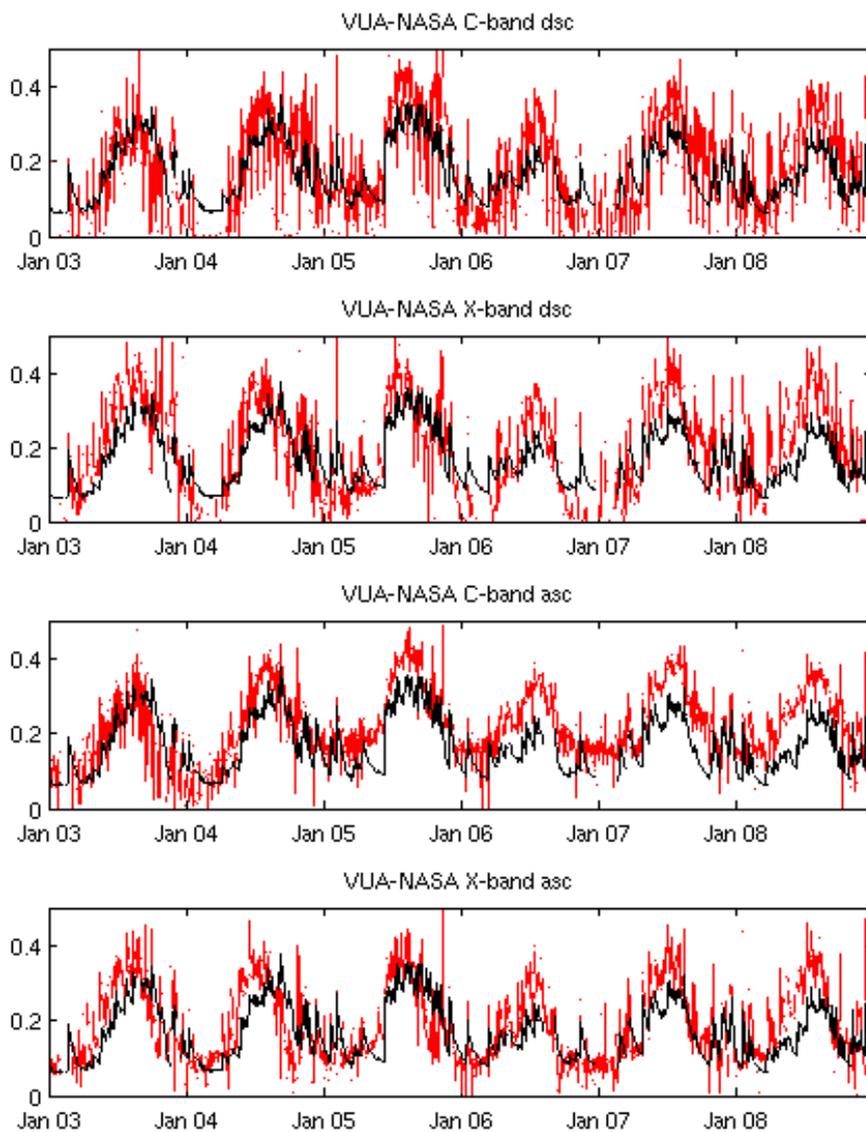


Figure 4.8: Time series of in situ (black) and original (not normalised) AMSR-E (red) near-surface soil moisture (m^3m^{-3}) from the VUA-NASA retrieval algorithm at Kyeamba (M9), from 2003-2008.

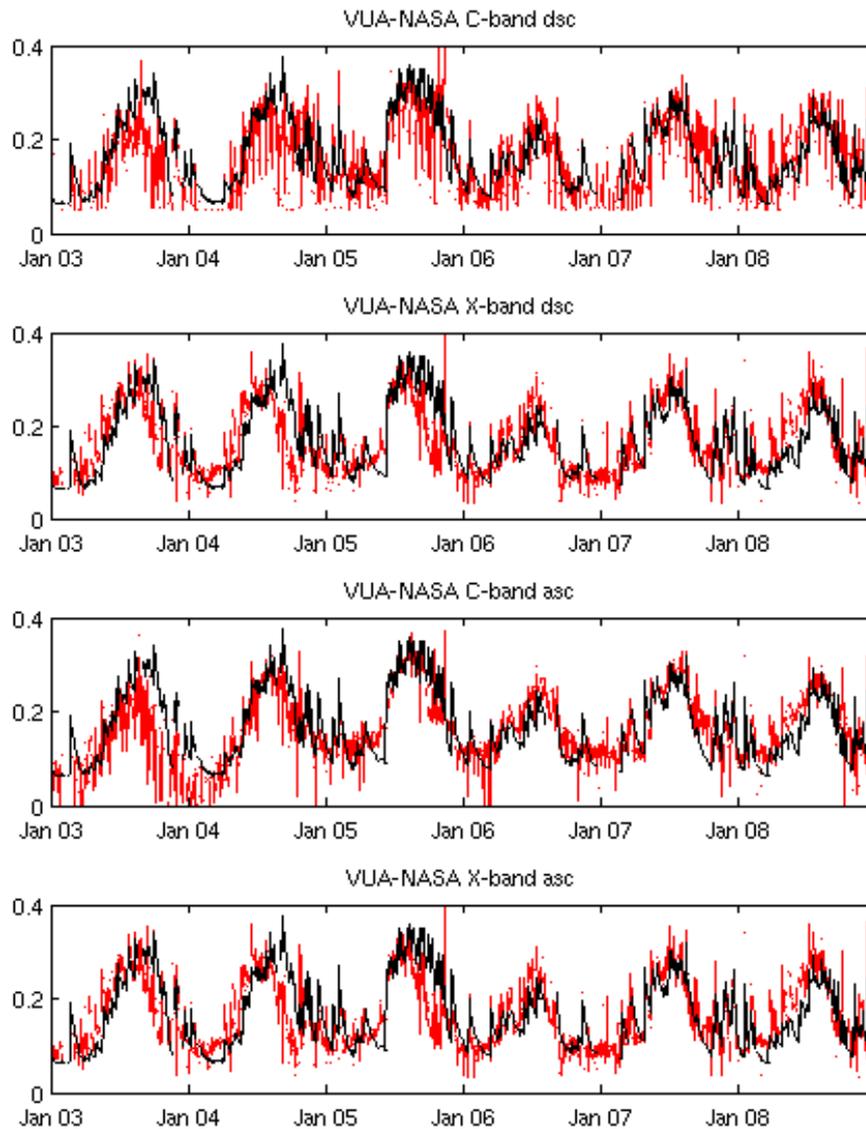


Figure 4.9: Time series of in situ (black) and normalised AMSR-E (red) near-surface soil moisture (m^3m^{-3}) from the VUA-NASA retrieval algorithm at Kyeamba (M9), from 2003-2008.

Table 4.8: Statistics of fit between the VUA-NASA AMSR-E soil moisture time series and the in situ data for 2003-2008. The RMSD is provided for both the original and normalised AMSR-E time series. Bold correlations indicates significance at the 1% level.

	VUA-NASA C - Dsc.				VUA-NASA X - Dsc.			
	r_{abs}	r_{anm}	RMSD (m^3m^{-3})		r_{abs}	r_{anm}	RMSD (m^3m^{-3})	
			raw	norm			raw	norm
M1	0.62	0.43	0.180	0.028	0.66	0.51	0.135	0.026
M2	0.56	0.33	0.169	0.028	0.52	0.33	0.090	0.031
M4	0.82	0.62	0.053	0.034	0.80	0.60	0.065	0.035
M5	0.72	0.60	0.094	0.023	0.65	0.59	0.117	0.025
M6	0.53	0.31	0.096	0.049	0.52	0.29	0.111	0.050
M7	0.66	0.49	0.080	0.024	0.61	0.50	0.094	0.026
M8	0.81	0.65	0.083	0.033	0.80	0.65	0.097	0.034
M9	0.65	0.40	0.092	0.061	0.74	0.62	0.090	0.053
M10	0.77	0.43	0.170	0.043	0.70	0.53	0.115	0.050

	VUA-NASA C - Asc.				VUA-NASA X - Asc.			
	r_{abs}	r_{anm}	RMSD (m^3m^{-3})		r_{abs}	r_{anm}	RMSD (m^3m^{-3})	
			raw	norm			raw	norm
M1	0.61	0.46	0.164	0.028	0.67	0.54	0.134	0.025
M2	0.60	0.27	0.193	0.026	0.70	0.47	0.097	0.023
M4	0.79	0.55	0.045	0.035	0.79	0.54	0.049	0.035
M5	0.57	0.40	0.102	0.028	0.62	0.44	0.103	0.026
M6	0.51	0.30	0.086	0.051	0.52	0.30	0.089	0.050
M7	0.55	0.50	0.084	0.026	0.58	0.55	0.085	0.026
M8	0.78	0.55	0.082	0.034	0.78	0.59	0.077	0.034
M9	0.78	0.49	0.085	0.049	0.75	0.63	0.073	0.051
M10	0.79	0.40	0.181	0.042	0.73	0.48	0.107	0.047

significant at 1%. Additionally, comparison of r_{ann} did not consistently favour either wavelength.

The descending (nighttime) AMSR-E overpass was expected to produce more accurate soil moisture than the ascending (daytime) overpass. For the original AMSR-E data plotted in Figure 4.8, the descending overpass had more variance, and had a more exaggerated seasonal cycle, and in particular the summer minima were much lower (contrary to expectation that bare-soil evaporation would generate a greater tendency toward dry-end values during the day). This suggests that the descending overpass was more sensitive to changes in near-surface soil moisture. However, the descending overpass also had more short-term variability, which was likely due to both noise and a true signal. This behaviour occurred across all of the Murrumbidgee sites, resulting in higher standard deviations for the descending overpass than for the ascending overpass for both wavelengths in Table 4.7. Despite these differences, there are no obvious differences between the normalised time series plots for each wavelength (see Figure 4.9 and Appendix G). Also, the statistics in Table 4.8 suggest that for C-band at least, the descending overpass may be slightly more accurate than the ascending overpass: r_{abs} is higher, and the RMSD is lower at six of nine sites. However, for the ascending overpass the C-band has better statistics at just four locations (with one site having equal statistics). Again, the differences between the absolute correlations were not significant, and the anomaly correlation did not consistently favour either overpass.

Spatial assessment against related observations

In this section the realism of the spatial patterns in the AMSR-E soil moisture is checked by qualitative comparison against observed precipitation and vegetation greenness across Australia. Maps of the mean monthly soil moisture from AMSR-E in January and June are shown in Figure 4.10, and maps of the equivalent monthly precipitation and mean NDVI are shown in Figure 4.11. At the continental scale, the mean soil moisture reflects the precipitation patterns in each month. In January, the extremely high monsoonal rain (~ 200 to ~ 600 mm/month) in tropical north Australia is evident in the mean monthly soil moisture, as are the smaller regions of elevated precipitation (~ 100 mm) along the east coast. However, the rainfall (~ 100 mm) in Western Australia is not reflected in the soil moisture maps, likely due to the precipitation signal being reduced by the episodic nature of rainfall there (most of the monthly total fell

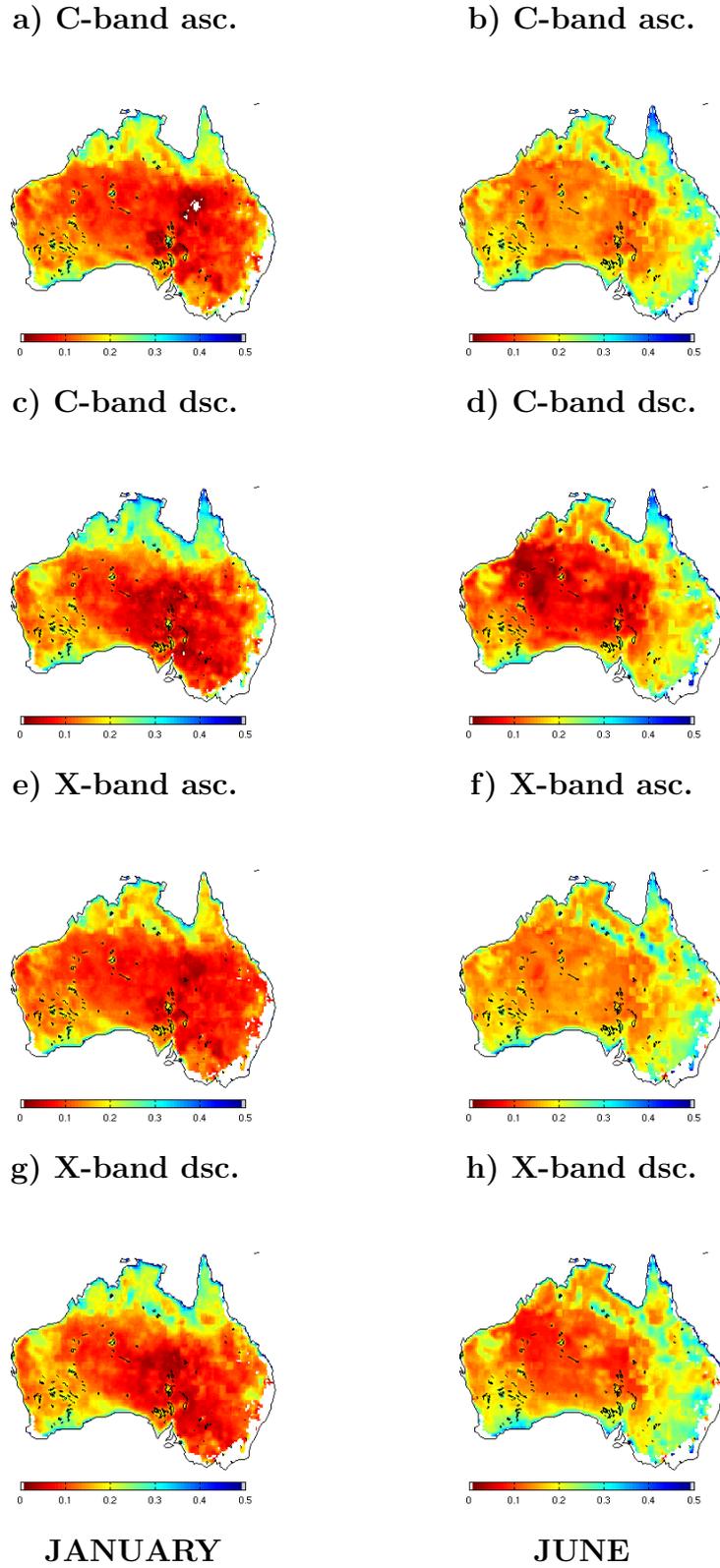


Figure 4.10: Monthly mean AMSR-E soil moisture (m^3m^{-3}) across Australia, from the C-band ascending overpass (first row), C-band descending overpass (second row), X-band ascending overpass (third row), and the X-band descending overpasses (fourth row), for January (left) and July (right), 2006.

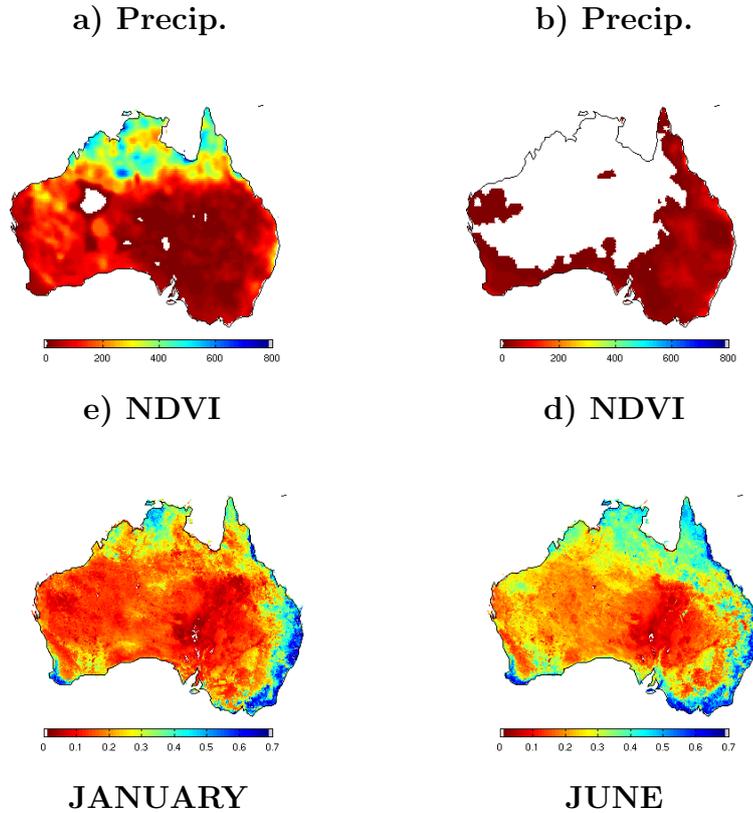


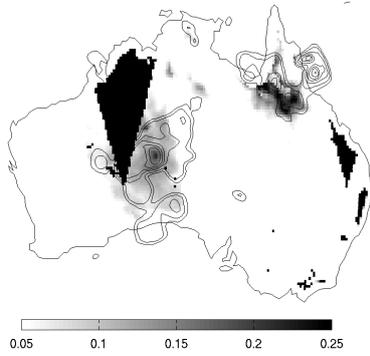
Figure 4.11: Monthly mean precipitation (in mm; first row), and NDVI (second row), for January (left) and July (right), 2006.

in the first 2 weeks of the month), combined with the extremely high potential evaporation³. In June, the winter precipitation across east Australia is also evident in the soil moisture maps. As with the time series plots above, the descending overpass soil moisture has a stronger contrast between wet and dry regions in Figure 4.10, and in particular it has a stronger signal of the rain in tropical north Australia in January, particularly for C-band. The mean monthly X-band and C-band maps are quite similar, although there is less contrast between the wet- and dry-end conditions in the X-band maps.

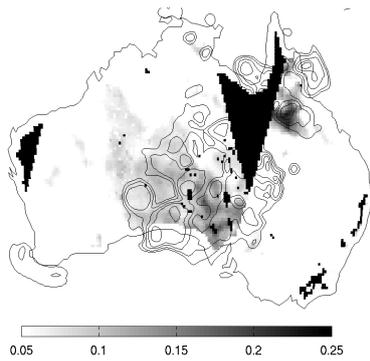
At finer spatial scales, two other features are evident in the AMSR-E soil moisture maps; ephemeral salt-lakes and vegetation cover. Inland water bodies have been outlined in black in Figure 4.10, and many arid-zone lakes are identifiable as regions of elevated soil moisture. Most of these are ephemeral salt lakes, which fill only during flood events, and the elevated soil moisture may be due to ground-water discharge, possibly combined with a surface salt crust

³The BoM estimates the annual mean average pan evaporation across inland West Australia to be between 2000 and 4000 mm/year.

13 July



14 July



15 July

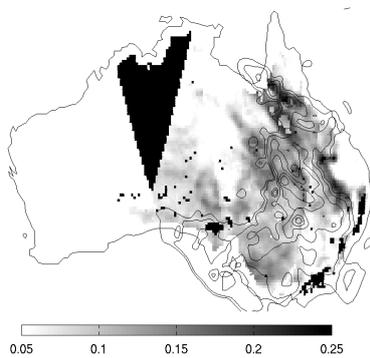


Figure 4.12: Maps of the C-band AMSR-E daily near-surface soil moisture anomaly (m^3m^{-3}), with 10 mm precipitation contours for 13 - 15 July, 2006, based on the average of the ascending and descending AMSR-E retrievals. Black indicates no AMSR-E data.

acting to reduce surface moisture evaporation. There is also a strong correspondence between the maps of monthly soil moisture and monthly NDVI in Figure 4.11. This correspondence could be due to either a true relationship between elevated soil moisture and vegetation vigour, or to vegetation artefacts in the soil moisture retrievals. Even if these are vegetation artefacts, in Section 4.3.3 the AMSR-E data still showed a strong soil moisture signal in the presence of dense vegetation (recall that the Murrumbidgee Monitoring Network includes some relatively densely vegetated sites).

In addition to the monthly averages in Figure 4.10, there is also a strong spatial relationship between the AMSR-E soil moisture and precipitation at shorter time scales. Figure 4.12 shows examples of positive daily soil moisture anomalies, indicating daily wetting, together with 10 mm daily precipitation contours, from 13-15 July, 2006. This period was selected as an example of widespread precipitation preceded by a dry spell. In each panel of the Figure there is a clear pattern of elevated soil moisture in the regions of precipitation. There is some mismatch between the locations of each, some of which will be due to timing differences between the two observations: precipitation is observed over the 24 hours to 9 am, while the AMSR-E maps are an average of the anomaly at 1:30 am and 1:30 pm (the mean was used to maximise the spatial coverage of the soil moisture maps).

Implications of the VUA-NASA soil moisture assessment

For both the C- and X-band AMSR-E data, the positive assessment of the VUA-NASA soil moisture over 2006 in Section 4.3.2 has been confirmed by extending the temporal comparison to the in situ data from Murrumbidgee monitoring network to a 6 year period, and by qualitatively comparing maps of the AMSR-E soil moisture against observed precipitation and vegetation. While the C-band passive microwave data were expected to yield more accurate estimates of soil moisture than the longer X-band wavelengths, neither wavelength yielded consistently better results. Despite this, the C-band VUA-NASA product is selected for use in the assimilation experiments, since it is theoretically favoured and RFI is not problematic in Australia. However, in regions where RFI prevents the use of C-band data, these results indicate that the X-band soil moisture product could be confidently used. Likewise, the (nighttime) descending AMSR-E overpass was theoretically favoured over the (daytime) ascending overpass data. While the descending overpass showed more sensitivity

to changes in soil moisture, both temporally (Figures 4.8 and Appendix F) and spatially (Figure 4.10), neither overpass consistently had better agreement with the in situ data. The descending overpass has been selected for use in the assimilation experiments in preference to the ascending overpass data, since it is theoretically favoured. However, these results suggest that the ascending overpass data could also be used without loss of accuracy.

Note that the best statistics for the association between the AMSR-E and in situ observations in Table 4.8 are similar to the statistics comparing the in situ time series from different sensors within a single pixel in Table 4.4. Consequently, it is possible that the above inter-comparisons of the soil moisture time series from different AMSR-E wavelengths and overpasses did not produce differing levels of skill (despite differences being expected) due to the limited accuracy with which the in situ data represents the area-average soil moisture observed by AMSR-E.

4.3.4 A filter for AMSR-E soil moisture

The VUA-NASA soil moisture time series consistently had much greater short-term variability than the in situ data. While some of this short-term variability is likely due to noise in the AMSR-E signal, much of it will be due to the shallower depth of AMSR-E observations (compared to the in situ data). Consequently, an exponentially weighted moving average filter is presented here to filter the noise in the AMSR-E data, and hence reduce the systematic differences between soil moisture from AMSR-E and from (observed or modeled) soil moisture over a deeper layer. This is the same as the “exponential filter” that is commonly applied to ERS and ASCAT data (Wagner et al., 1999) to extrapolate the near-surface surface observations to a deeper soil moisture layer, however it is used here principally to reduce temporal noise. Qualitatively, the filtered AMSR-E time series at Kyeamba in Figure 4.13 has a better fit to the in situ data, since it is less noisy than the original data, while still accurately detecting precipitation events. The quantitative agreement with the in situ data is also improved (note that applying a filter will in general improve correlation statistics by reducing variability about the mean). Comparing the statistics for the original and filtered AMSR-E data sets in Tables 4.8 and 4.9 shows that the filter increased r_{abs} by about 0.1 in most cases. Since r_{anm} is more susceptible to noise, it increased more dramatically, by 0.1 to 0.3 in most cases, giving

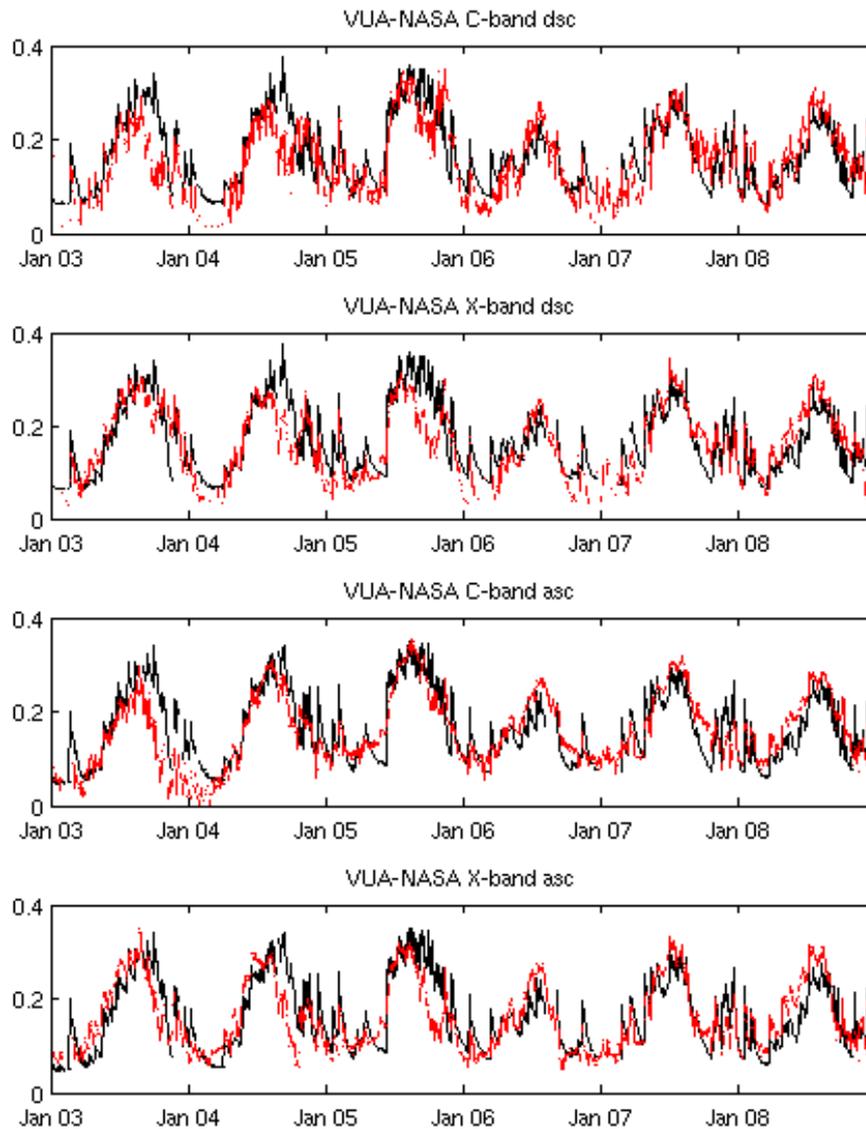


Figure 4.13: Time series of near-surface soil moisture (m^3m^{-3}) from C-band descending overpass AMSR-E (red), and Murrumbidgee Monitoring Network observations (black), for April 08 - April 09. The AMSR-E data have been normalised to the mean and variance of the in situ data, and filtered with an exponential moving average filter.

Table 4.9: Statistics of fit between the filtered VUA-NASA AMSR-E soil moisture time series and the in situ data for 2006. The RMSD is provided for both the original and normalised AMSR-E time series. Bold correlations indicates significance at the 1% level.

	VUA-NASA C - Dsc.				VUA-NASA X - Dsc.			
	r_{abs}	r_{anm}	RMSD (m^3m^{-3})		r_{abs}	r_{anm}	RMSD (m^3m^{-3})	
			raw	norm			raw	norm
M1	0.77	0.72	0.161	0.021	0.79	0.77	0.123	0.020
M2	0.69	0.55	0.153	0.024	0.68	0.57	0.064	0.026
M4	0.87	0.77	0.042	0.029	0.83	0.75	0.056	0.032
M5	0.76	0.76	0.086	0.021	0.68	0.75	0.111	0.024
M6	0.63	0.55	0.084	0.044	0.59	0.53	0.102	0.046
M7	0.75	0.70	0.072	0.021	0.67	0.70	0.089	0.024
M8	0.82	0.75	0.078	0.031	0.81	0.73	0.092	0.033
M9	0.77	0.55	0.061	0.050	0.79	0.70	0.075	0.048
M10	0.87	0.59	0.155	0.033	0.77	0.69	0.096	0.044

	VUA-NASA C - Asc.				VUA-NASA X - Asc.			
	r_{abs}	r_{anm}	RMSD (m^3m^{-3})		r_{abs}	r_{anm}	RMSD (m^3m^{-3})	
			raw	norm			raw	norm
M1	0.71	0.69	0.160	0.024	0.78	0.79	0.125	0.021
M2	0.74	0.40	0.176	0.021	0.78	0.59	0.084	0.020
M4	0.87	0.76	0.038	0.028	0.87	0.77	0.038	0.027
M5	0.67	0.66	0.097	0.024	0.71	0.70	0.098	0.023
M6	0.61	0.56	0.078	0.046	0.63	0.56	0.080	0.045
M7	0.64	0.67	0.080	0.024	0.66	0.72	0.079	0.023
M8	0.85	0.73	0.078	0.028	0.85	0.75	0.071	0.028
M9	0.83	0.57	0.075	0.043	0.81	0.69	0.061	0.045
M10	0.84	0.48	0.175	0.037	0.81	0.65	0.088	0.040

values between 0.55 - 0.77 (0.53 - 0.77) for the descending overpass C-band (X-band) time series, and 0.40 - 0.76 (0.59 - 0.79) for the equivalent ascending overpass time series. Note that applying the filter did not change the relative performance of the data sets from different overpasses or different wavelengths.

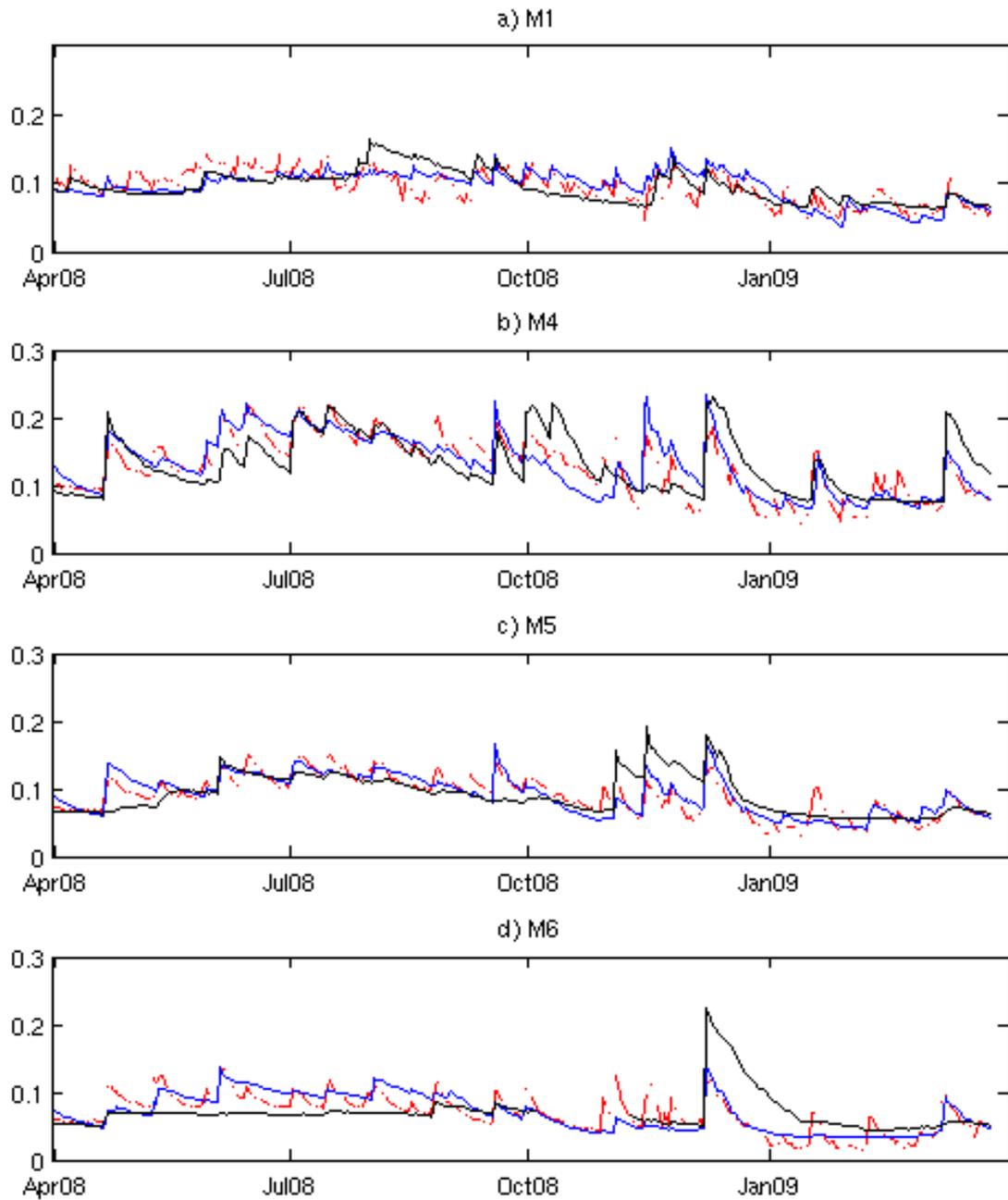
The qualitative and quantitative improvement in the fit to the in situ data that is generated by filtering the AMSR-E data suggests that the filter has successfully reduced the systematic differences between the AMSR-E and in situ data. For data assimilation into models which have a deeper near-surface layer than the AMSR-E observation depth, applying such a filter is expected

to be beneficial, and the AMSR-E data will be filtered before being assimilated into the ACCESS model (for which the near-surface model layer has a depth of 10 cm), although it will not be filtered for use in the ISBA model (which has a near-surface layer depth of 1 cm).

4.3.5 Inter-comparison of AMSR-E and ACCESS model soil moisture

Finally the AMSR-E soil moisture data have been bench-marked against the near-surface soil moisture forecasts from the BoM's ACCESS model (used in the Australian assimilation experiments in Chapter 6). The ACCESS soil moisture forecasts were generated by forcing an off-line version of its land surface model with short-range atmospheric forecasts from ACCESS, from April 2008 to April 2009 (this is the open-loop simulation introduced in Chapter 6). There were no data available from the Goulburn Monitoring Network for this period, and so only Murrumbidgee in situ data have been used. The second generation Murrumbidgee stations had near-surface sensors for this period, and have been included, however there were widespread data outages across the Murrumbidgee network in late 2008, reducing the number of monitoring sites to eight: one at Yanco (M8: mean of Y1, Y3-Y13), two at Kyeamba (M9) – Kyeamba A (mean of K1, K2, K3, K5, & K7), and Kyeamba B (mean of K8, K11, & K14) – and five individual stations across the Murrumbidgee. The ACCESS forecasts have been compared to the AMSR-E soil moisture for the C-band descending overpass, and the AMSR-E data has been regridded onto the 0.375° ACCESS grid (rather than the 0.25° grid used earlier in this Chapter), using a nearest-neighbour approach. The AMSR-E time series at each grid-cell has then been filtered, as described in Section 4.3.4.

Time series of the ACCESS and AMSR-E soil moisture have been compared to the in situ data in Figure 4.13, and statistics describing the agreement with the in situ data are given in Table 4.10. Both AMSR-E and ACCESS agree well with the in situ data. Excluding M6, r_{abs} (r_{anm}) was between 0.60 - 0.87 (0.57 - 0.86) for AMSR-E, and 0.64 - 0.86 (0.64 - 0.88) for ACCESS. AMSR-E and ACCESS have comparable agreement with the in situ data, and the AMSR-E time series generated higher r_{abs} and r_{anm} , and lower (normalised) RMSD at four sites (and ACCESS generated better statistics at the other four sites). Both had unusually poor fit at M6, and Figure 4.13d suggests this might be due to



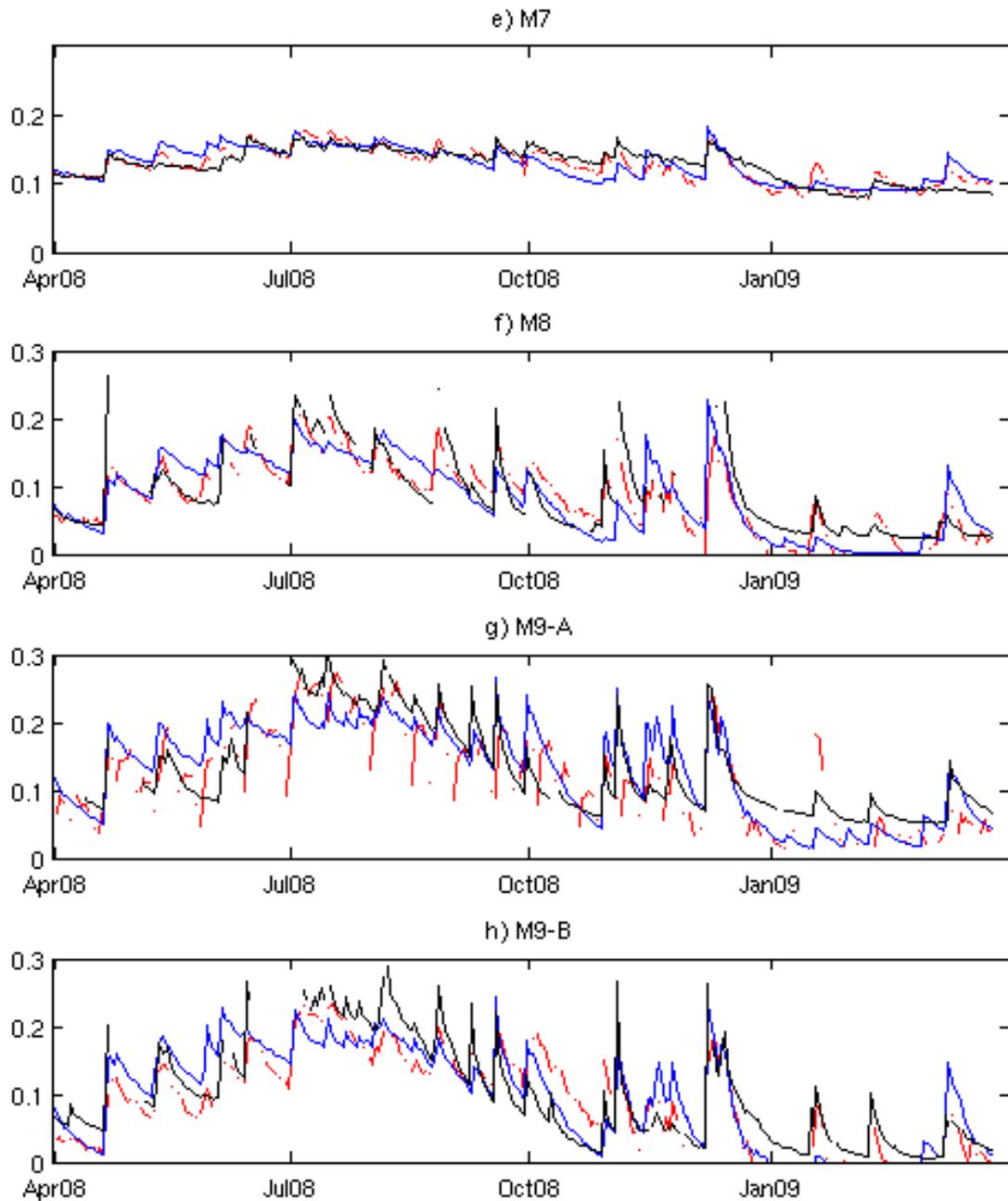


Figure 4.13: Time series of filtered and normalised C-band descending AMSR-E overpass soil moisture (red), normalised ACCESS near-surface soil moisture forecasts (blue), and Murrumbidgee Monitoring Network in situ observations (black), for April 08 - April 09.

Table 4.10: Statistics of fit to the in situ observations, from the filtered C-band descending overpass VUA-NASA AMSR-E data, and the ACCESS forecast near-surface soil moisture.

	AMSR-E				UM			
	r_{abs}	r_{anm}	RMSD (m^3m^{-3})		r_{abs}	r_{anm}	RMSD (m^3m^{-3})	
			raw	norm			raw	norm
M1	0.60	0.59	0.230	0.020	0.68	0.50	0.429	0.018
M4	0.71	0.73	0.049	0.033	0.64	0.71	0.199	0.036
M5	0.60	0.48	0.069	0.026	0.69	0.62	0.203	0.023
M6	0.27	0.44	0.071	0.034	0.35	0.74	0.220	0.031
M7	0.86	0.73	0.055	0.013	0.76	0.65	0.202	0.017
YA	0.87	0.72	0.069	0.030	0.78	0.56	0.186	0.038
KA	0.84	0.53	0.051	0.040	0.83	0.76	0.208	0.041
KB	0.84	0.77	0.072	0.045	0.86	0.77	0.194	0.040

a problem with the in situ data: both the AMSR-E and ACCESS soil moisture increased during the winter of 2008, while the in situ measurement remained low. The soil moisture at the surrounding in situ stations (M5, M7 and M8; see Figure 4.1) also increased through winter, with similarly timed step changes (suggesting large scale precipitation events) as shown by ACCESS and AMSR-E at M6. Figure 4.14 shows the time series of observed precipitation from the BoM's rain gauge analysis over an 0.25° grid surrounding the M6 site, confirming the occurrence of precipitation, consistent with the signals from AMSR-E and ACCESS at M6. This strongly suggests that the in situ data at M6 does not accurately reflect the area-average soil moisture, either because of measurement errors, or because the precipitation events in question were localised, and did not occur at the location of the M6 in situ sensor (although the detection of similarly timed precipitation at the surrounding in situ sensors suggests that this was not the case).

Assimilating near-surface soil moisture data with similar accuracy to the model soil moisture forecasts is expected to generate modest improvements in the model soil moisture (Reichle et al., 2008b), hence the above inter-comparison suggests that assimilating the C-band VUA-NASA AMSR-E soil moisture into the ACCESS model in Chapter 6 will generate improved soil moisture forecasts.

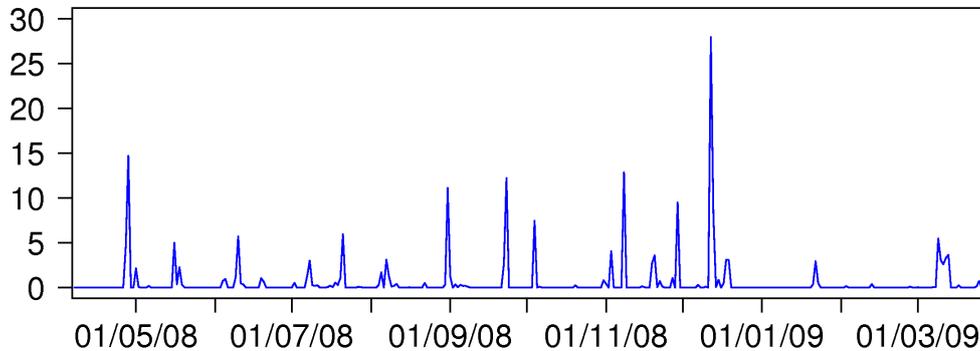


Figure 4.14: Time series of precipitation (mm day^{-1}) from the Bureau of Meteorology's rain gauge analysis for the 0.25° grid surrounding M6, from April 2008 to March 2009.

4.4 Chapter summary

The currently available remotely sensed near-surface soil moisture data sets over Australia have been assessed to identify which data set should be used in the assimilation experiments in this thesis, and to establish the best use of that data. At the time of this study the only soil moisture data sets with sufficient coverage over Australia for use in NWP model were being derived from the passive microwave AMSR-E instrument. Soil moisture data retrieved from AMSR-E with the algorithms developed at VUA-NASA, NASA, USDA, and JAXA were compared to in situ soil moisture observations from the Murrumbidgee and Goulburn Monitoring Networks over 2006, and based on this comparison the VUA-NASA retrieval algorithm was identified as the most realistic, and consequently it will be used in the assimilation experiments in Chapters 5 - 6.

The VUA-NASA soil moisture data was then examined in more detail, in principal by comparison to six years of in situ data from the Murrumbidgee Monitoring Network, to identify which observation wavelength (C-band or X-band) and which overpass time (descending or ascending) should be used in these assimilation experiments, and to further characterise the accuracy of that data. Neither of the observation wavelengths or observation times was identified as yielding superior near-surface soil moisture estimates, and consequently the

C-band descending overpass data will be preferentially used, since it is theoretically favoured. However, in situations where more frequent data are warranted, or where RFI prevents the use of C-band data, the results obtained here suggest that the ascending pass and/or X-band data could also be used confidently. Additionally, comparison over the six year period confirmed that the VUA-NASA observations can detect a strong signal of temporal changes in soil moisture.

The above evaluation of remotely sensed soil moisture is based on in situ soil moisture observations, and so is limited by both the systematic differences between the soil moisture quantities observed by in situ and remote sensors, and also by the limited coverage of the in situ data. However, it is not yet clear how a more robust evaluation of remotely sensed soil moisture can be attained. Consequently, the AMSR-E data were also compared visually to both precipitation and vegetation observations, confirming that the soil moisture data had the expected large-scale spatial patterns.

The AMSR-E observations has much greater short short-term variability than the in situ data, due to the much shallower surface layer observed by C-band passive microwave sensors (1 cm, compared to 5-8 cm for the in situ data), although there was likely also some noise in the remotely sensed signal. An exponential filter was then presented to reduce the short-term variability in the AMSR-E soil moisture time series, for better comparison to soil moisture estimates over deeper surface layers. Applying this filter improved the statistics of fit between the AMSR-E and in situ soil moisture observations, and so it will be applied to the AMSR-E data before it is assimilated into the MOSES model (which has a 10 cm near-surface layer) in Chapter 6, however it will not be applied for use with the ISBA model (which has a 1 cm near-surface layer) model in Chapter 5.

Finally, inter-comparison of the filtered C-band descending overpass AMSR-E soil moisture and near-surface soil moisture forecasts from the MOSES model, showed that they have similar accuracy (compared to in situ observations from the Murrumbidgee Monitoring Network). It is then expected that assimilating the AMSR-E data into the MOSES model in Chapter 6 will yield small improvements in the modeled soil moisture.

Chapter 5

Assimilation Experiments with ISBA

5.1 Overview of chapter

Chapter 4 demonstrated that useful near-surface soil moisture information can be retrieved from AMSR-E brightness temperatures using the retrieval algorithm developed by VUA-NASA. Additionally, Section 3.2 outlined an EKF-based land surface analysis method capable of assimilating both screen-level observations and novel remotely sensed observations of the land surface. In this chapter the EKF is used to assimilate the AMSR-E near-surface soil moisture into Météo-France’s NWP land surface model (ISBA), to obtain updates to the model root-zone soil moisture.

There has been little previous work assimilating near-surface soil moisture data into ISBA, and this is the first study to assimilate satellite observations over a large domain. Consequently, this chapter begins by exploring how the EKF relates the near-surface soil moisture observations to the underlying root-zone soil moisture in the model. Additionally, the assimilation of the AMSR-E near-surface soil moisture data is compared to the assimilation of screen-level temperature (T_{2m}) and relative humidity (RH_{2m}). Specifically, a series of experiments is conducted to examine whether the AMSR-E data might be useful, either as a substitute for the screen-level data (to improve the low-level atmospheric forecasts) or as a complimentary data source (to improve both the root-zone soil moisture and low-level atmospheric forecasts). The relative impact of each data type is measured, in terms of the net magnitude

and direction of the root-zone soil moisture increments generated by each, as well as the relative information content contributed by each data source when they are assimilated together.

The development of the EKF assimilation of AMSR-E near-surface soil moisture into ISBA has been published in Draper et al. (2009a), while the combined assimilation of near-surface soil moisture and screen-level observations has been published in Draper et al. (2011). Additionally, the work presented here also contributed to the simplified EKF assimilation of screen level variables reported in Mahfouf et al. (2009).

5.2 Data and methods

This section describes the assimilated data sets, together with the main experiments undertaken to assimilate these data into ISBA.

5.2.1 AMSR-E soil moisture data

Near-surface soil moisture data retrieved from AMSR-E brightness temperatures with the VUA-NASA retrieval algorithm of Owe et al. (2008) have been used in the assimilation experiments, following the findings in Chapter 4. Since Njoku et al. (2005) showed that C-band RFI is not widespread across Europe (with the exception of isolated pockets over some urban areas), the C-band data have been used, with RFI-contaminated data screened out. Both the ascending and descending overpass data have been assimilated to maximise the potential impact of the AMSR-E data, since the impact of assimilating the AMSR-E data was relatively low with the assimilation strategy used here (and the assessment in Chapter 4 did not reveal a significant difference between the accuracy of the two overpasses). The EKF uses a 6 hour assimilation cycle (starting at 0:00, 6:00, 12:00 and 18:00 UTC each day) with observations assimilated at the end of each cycle, for consistency with the observation times of the screen-level observations. The AMSR-E data have then been assimilated at the closest 6 hour interval, so that the descending overpass observations (observed at approximately 1:30 am local time) were assumed to occur at 0:00 UTC, and the ascending overpass observations (observed at approximately 1:30 pm) were assumed to occur at 12:00 UTC.

The level 1 AMSR-E swath-data have been directly mapped onto the irregu-

lar ALADIN grid using a nearest neighbour approach. All observations contaminated by RFI or open water, and where dense vegetation or frozen ground-cover conceals the near-surface microwave signal have been screened out. RFI contamination was identified based on the RFI index of Li et al. (2004), which is provided together with the soil moisture data. Frozen ground-cover is identified and removed during the moisture retrieval, although this was not significant over Europe in July. Densely vegetated regions were screened out according to the vegetation optical depth that is retrieved from the AMSR-E brightness temperatures together with the soil moisture data (Owe et al., 2001). The vegetation optical depth is linearly proportional to the vegetation water content, and the sensitivity of the microwave brightness temperature to soil moisture decreases with increasing vegetation optical depth (e.g., de Jeu et al., 2008). Owe et al. (2001) show that the soil moisture sensitivity is quite low for optical depths above about 0.75, and so a mean monthly optical depth threshold of 0.8 has been adopted, following de Jeu et al. (2008). Finally, to prevent outliers in the AMSR-E data from generating large analysis updates, an additional quality control has been applied immediately prior to the assimilation, by discarding data for which the observation increment was more than twice the standard deviation of the increments over the month. This rather strict threshold was chosen arbitrarily in response to the presence of occasional suspect outliers in the AMSR-E data (likely resulting from the absence of strict quality controls during the soil moisture retrieval), since these outliers can have a substantial detrimental impact on the quality of the analyses.

After the above quality control was applied, the average spatial coverage of the AMSR-E data over July 2006 was 49% of the European domain, for both overpasses. Figure 5.1 shows the temporal coverage of the assimilated AMSR-E data for the descending overpass over July 2006 (the ascending overpass coverage is similar). In the figure, the daily coverage over Europe is nearly 100% at 58°N, decreasing towards the equator, with 70% of days covered at 33°N. Superimposed on this general pattern are several isolated regions of reduced coverage due to local RFI, usually over urban regions (for example, London). Additionally, there was widespread significant RFI over Italy, and roughly 50% of the data over the Italian peninsula was removed. This result differs from the initial RFI investigation over 2003 of Njoku et al. (2005), who found limited C-band RFI over Europe, and X-band RFI over Italy, also using the index of Li et al. (2004).

Finally, the quantity observed by AMSR-E is assumed to be equivalent to the model near-surface soil moisture (w_1), since both represent the soil moisture in approximately the upper-most 10 mm of soil. Additionally, before being assimilated the AMSR-E data were rescaled to reduce the systematic differences between the observed and modeled w_1 . The bias removal strategy that was applied, and the resulting soil moisture fields, will be described in Section 5.4.1.

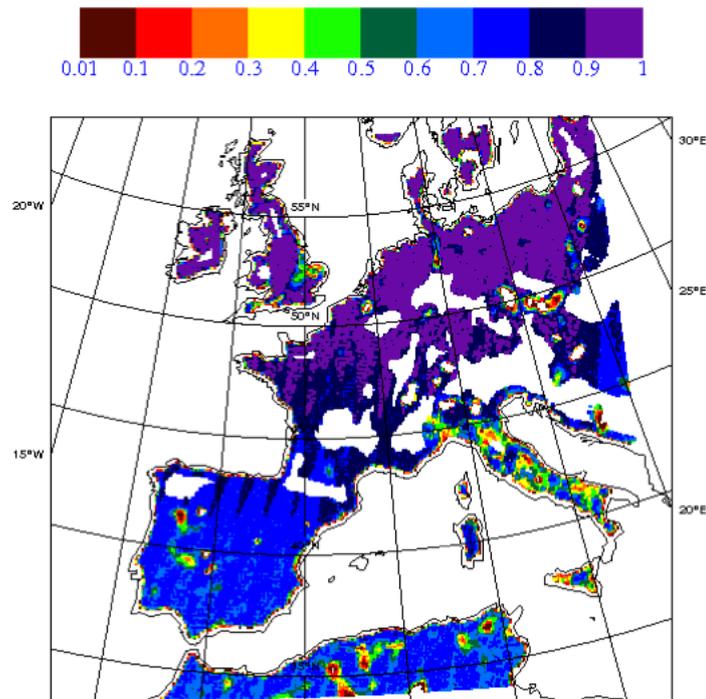


Figure 5.1: Fraction of days in July 2006 for which the descending overpass AMSR-E soil moisture observations were available for assimilation (after quality control).

5.2.2 Screen-level data

The assimilated screen level T_{2m} and RH_{2m} fields were taken from the 6-hourly bidimensional OI scheme used to analyse the screen-level fields in ALADIN (Taillefer, 2002). This is the same data that is currently used in the OI land surface analysis in ALADIN (described in Section 2.3.2).¹ The screen-level OI analyses the screen-level variables in ALADIN every 6 hours, based on screen-level observations from GTS SYNOP, SHIP, BUOY reports, and the French

¹In practise, it is actually the observation increments from the screen-level analysis that are used in the land surface analysis.

RADOME network. Approximately 2000 observations each of T_{2m} and RH_{2m} are ingested per cycle, although the density of the observation network, shown in Figure 5.2, varies widely across Europe. In particular, observations are reasonably dense across central Europe, and particularly over France where there are about 1000 observations available per analysis (due to the RADOME data). In contrast, the coverage over Spain and north Africa is much sparser, and there are only about 50 stations across Spain, which has a similar area to France. The background field for each analysis is an ALADIN forecast, so that the screen-level analysis for model grids not affected by observations will equal the background forecast. Since the current land surface analysis does not include bias correction of the screen-level data, they have not been bias-corrected here either.

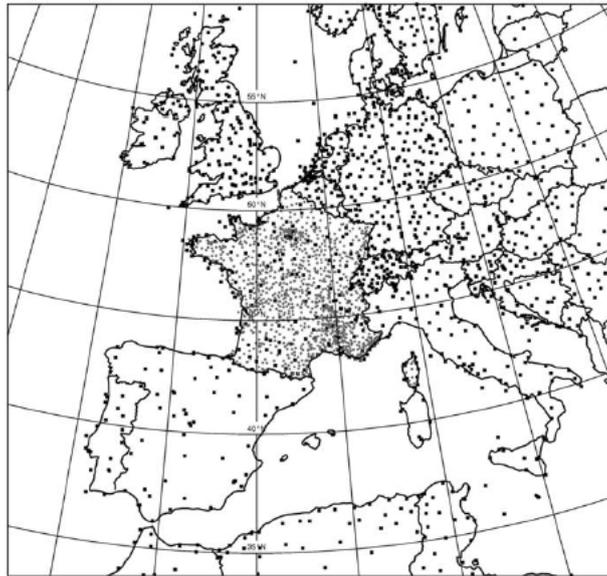


Figure 5.2: Screen-level observation network over the European domain. Black squares indicate GTS SYNOP observations, and small grey circles indicate French RADOME observations (reproduced from Figure 3 of Mahfouf et al. (2009)).

5.2.3 The assimilation experiments and error selection

To compare the impact of assimilating the screen-level observations and AMSR-E near-surface soil moisture, four main assimilation experiments have been conducted:

- **CTR** - an open-loop simulation with no assimilation (the control)
- **SLV** - assimilation of screen-level variables (T_{2m} and RH_{2m})
- **AMS** - assimilation of AMSR-E w_1
- **CMB** - a combined assimilation of T_{2m} , RH_{2m} , and AMSR-E w_1 .

The EKF described in Section 3.2 has been used for each experiment (note that an approximation to reduce the number of model Jacobians required for each assimilation cycle will be introduced in Section 5.3.2). In ISBA, w_1 is defined over a thin surface layer, giving it a very short time scale (as will be shown in Section 5.3.2, perturbations to w_1 are largely lost within a day), and limiting its influence on other model variables. As a result, the accurate initialisation of w_1 is not very important in ISBA, and in their experiments assimilating screen-level observations into ISBA, Balsamo et al. (2007) and Mahfouf et al. (2009) updated the root-zone soil moisture (w_2) only. In contrast, for the EKF assimilation used here, w_1 has been included in the state update vector, to allow the EKF to evolve long term error covariance between w_1 and w_2 . However the analysis results will only be reported for w_2 .

The same observation and model error covariances have been prescribed for all experiments (for the relevant variables) conducted in this chapter. Since the true soil moisture is not well understood at the large scales modeled by NWP, accurate specification of the error covariances required to assimilate and/or analyse soil moisture is extremely difficult. Consequently, a fairly simplistic error selection has been used here, based on the assumption that these errors can be refined in the future if the assimilation proves useful. The selection of the error covariances is discussed below, and the selected covariances are summarised in Table 5.1.

The observation errors

The observation error covariances (\mathbf{R}) for the screen-level observations were taken directly from Mahfouf et al. (2009), and spatially uniform error variances of $(10\%)^2$ and $(1K)^2$ have been used for RH_{2m} and T_{2m} , respectively. These are the errors used in Météo-France’s screen-level OI analyses, from which the assimilated observations were taken. For the AMSR-E observations, a spatially uniform error variance of $(0.05 \text{ m}^3\text{m}^{-3})^2$ has been used. This is close to the

Table 5.1: Summary of the observation, model, and initial background error variances used in the assimilation experiments. All diagonal error covariances were assumed to be zero. Units are indicated in brackets.

Observation Errors		
T_{2m}	1.0^2	(K ²)
RH_{2m}	10^2	(% ²)
w_1	0.05^2	((m^3m^{-3}) ²)
Forecast Model Errors*		
w_1	$(0.2 \times (w_{fc} - w_{wilt}))^2$	((m^3m^{-3}) ²)
w_2	$(0.02 \times (w_{fc} - w_{wilt}))^2$	((m^3m^{-3}) ²)
Initial Background Errors*		
w_1	$(0.2 \times (w_{fc} - w_{wilt}))^2$	((m^3m^{-3}) ²)
w_2	$(0.2 \times (w_{fc} - w_{wilt}))^2$	((m^3m^{-3}) ²)

*The mean $(w_{fc} - w_{wilt})$ is $0.086 m^3m^{-3}$.

AMSR-E error estimates from Chapter 4: for example, the mean RMSD between the VUA-NASA descending overpass AMSR-E data and the Australian in situ data in Table 4.8 is $0.04 m^3m^{-3}$. It is also consistent with published root mean square error estimates from Europe, specifically from Wagner et al. (2007) ($0.06 m^3m^{-3}$ from one location in Spain), and Rüdiger et al. (2009) ($0.06 m^3m^{-3}$ at one location in France). The errors for each observation type were assumed to be independent from each other (i.e., \mathbf{R} is diagonal). While this assumption is clear for the w_1 data, it will neglect some correlation between T_{2m} and RH_{2m} , for example due to representativity errors (note that T_{2m} and RH_{2m} are observed at the same locations by different sensors, hence their instrument errors will not be aliased).

Background and model errors for w_2

For the EKF, both the additive model forecast error covariances (\mathbf{Q}) and the initial background model error covariances (\mathbf{P}) must be defined, although beyond the first few assimilation cycles \mathbf{P} is largely determined by \mathbf{Q} . Following Mahfouf et al. (2009) and Balsamo et al. (2007), the initial \mathbf{P} and \mathbf{Q} were assumed to be diagonal and spatially uniform, and both have been defined as a fraction of the difference between the soil moisture at field capacity (w_{fc}) and at the wilting point (w_{wilt}). The w_2 error variance in \mathbf{Q} has been set at $(0.02 \times (w_{fc} - w_{wilt}))^2$. Over the mean depth of w_2 (2.3 m) this amounts to applying an error

of 4 mm every 6 hours, giving 16 mm day^{-1} applied across the soil layer, which is slightly more than the approximate NWP surface water budget error of 10 mm day^{-1} quoted by Douville et al. (2000).

The initial w_2 error variance in \mathbf{P} was set rather high at $(0.2 \times (w_{fc} - w_{wilt}))^2$, equivalent to an error of approximately 40 mm. The EKF rapidly (within a few assimilation cycles) reduced the mean w_2 error variance in \mathbf{P} to approximately $(20 \text{ mm})^2$, which is consistent with empirical error estimates for w_2 . For example, comparisons of ISBA w_2 forecasts to in situ soil moisture observations have yielded root mean square error estimates of 28 mm (Muñoz Sabater et al., 2007), 20.0 mm (Calvet and Noilhan, 2000), and 22.6 mm, 32.2mm, and 14.5mm (Calvet et al., 1998), at various locations in France. Additionally, 20 mm is slightly larger than the (static) w_2 errors used by Mahfouf et al. (2009) (recall that the screen-level observation errors have also been taken from Mahfouf et al. (2009)), who in turn based their model and observation errors on those used to derive the update coefficients for the current operational OI surface analysis in ALADIN.

The mean w_2 error of 20 mm in \mathbf{P} is equivalent to $0.01 \text{ m}^3\text{m}^{-3}$, which is much smaller than the errors typically expected from a land surface model, and five-times smaller than the assumed AMSR-E observation errors. This apparently small error is due to the unrealistic nature of the soil moisture in the ALADIN ISBA model, associated with its coarse vertical resolution and poorly specified soil parameters. Since there is no distinction between the root-zone and total soil depth, the root-zone is defined as the entire soil column, which has a mean depth over the ALADIN domain of 2.3 m. To compensate for the potentially large water reservoir defined by this overly-thick root-zone, the water holding capacity $(w_{fc} - w_{wilt})$ in ISBA is specified to be too small (mean $(w_{fc} - w_{wilt})$: $0.086 \text{ m}^3\text{m}^{-3}$). As a comparison, the ECMWF H-TESEL land surface model has a 1 m deep root-zone, and a $(w_{fc} - w_{wilt})$ close to $0.2 \text{ m}^3\text{m}^{-3}$ for a medium soil texture (Balsamo et al., 2009). The misspecification of soil parameters in ISBA reduces the soil moisture variability, since a given volume of moisture forcing (P-E), or any errors in that forcing, is dispersed across the overly thick layer. For example applying a 20 mm error to the ECMWF root-zone would result in a volumetric error of $0.02 \text{ m}^3\text{m}^{-3}$ (note that this is still equivalent to $0.1 \times (w_{fc} - w_{wilt})$), rather than the $0.01 \text{ m}^3\text{m}^{-3}$ value obtained here.

Background and model errors for w_1

The model errors for w_1 are even more difficult to define than those for w_2 , due to the difficulty of taking observations over such a thin layer. The w_1 errors (in m^3m^{-3}) are expected to be greater than those in w_2 , since the surface layer has more active soil moisture dynamics and a greater exposure to the atmospheric forcing (errors). The model w_1 errors have then been set to give a mean w_1 error standard deviation approximately double the mean w_2 error standard deviation in \mathbf{P} . This was achieved using an error variance of $(0.2 \times (w_{fc} - w_{wilt}))^2$ for both the initial \mathbf{P} and \mathbf{Q} . For \mathbf{Q} this is equivalent to adding an error of 0.8 mm day^{-1} , which is approximately one third of the mean daily (bare soil) evaporation forecasts. w_1 has a very short memory, and the mean w_1 error variance in \mathbf{P} was $(0.024 \text{ m}^3\text{m}^{-3})^2$ ($(0.24 \text{ mm})^2$), only slightly larger than the added \mathbf{Q} .

While the error variance assumed for w_1 is less certain than the previously specified error variances, the w_1 error has less impact on the analysis results than the other errors. To confirm this, a sensitivity test was conducted in which the w_1 error variance in \mathbf{Q} was replaced with estimates of the minimum and maximum of its likely values: $0.1 \times (w_{fc} - w_{wilt})^2$ which gives \mathbf{P}_{11} close to \mathbf{P}_{22} , and $(0.5 \times (w_{fc} - w_{wilt}))^2$ which gives \mathbf{P}_{11} close to \mathbf{R} for w_1 . Using these values resulted in a variation in the monthly mean \mathbf{K} for updating w_2 from w_1 of +3% and -16% of the original value, respectively.² In comparison, varying the w_2 model error over the same range resulted in a factor of five difference in \mathbf{K} . Hence, the uncertainty in the specification of the w_1 error in these experiments is not expected to have a significant impact on the conclusions of this study.

5.2.4 The relative information content

To quantify the contribution of the AMSR-E observations to the CMB experiments the relative information content of the assimilated data has been calculated. The relative information content, calculated following Balsamo et al. (2007), measures the sensitivity of the analysis update to each observed variable, scaled by the net sensitivity of the analysis to all of the assimilated observations. From Cardinali et al. (2004), the sensitivity of the analysis to the

²Note that increasing the model error variance for w_1 decreases the Kalman gain for updating w_2 , since this reduces the confidence of the comparison between the observed and modeled w_1 .

observations, mapped into observation space, is given by the derivative:

$$\frac{\partial \mathbf{Hx}_a}{\partial \mathbf{y}} = \mathbf{K}^T \mathbf{H}^T \quad (5.1)$$

Additionally, the *ith* diagonal of $\mathbf{K}^T \mathbf{H}^T$ gives the sensitivity of the model analysis to the *ith* observation, and $Tr(\mathbf{K}^T \mathbf{H}^T)$ gives the total sensitivity, or information content, from all of the assimilated observations (where Tr is the trace of the matrix).³ The relative information content of the *ith* observation can then be calculated as:

$$rIC_i = (\mathbf{K}^T \mathbf{H}^T)_{ii} / Tr(\mathbf{K}^T \mathbf{H}^T) \quad (5.2)$$

5.3 Implementing the EKF in ISBA

Before the assimilation results are presented in Section 5.5, the preliminary work testing and refining the assimilation strategy is presented here. This includes testing the accuracy of the estimates of the linearised model used by the EKF, examining the manner in which the root-zone soil moisture is related to near-surface soil moisture in ISBA, and comparing this to the relationship between the screen-level variables and the root-zone soil moisture.

5.3.1 Testing the ISBA model Jacobians

The EKF described in Section 3.2 requires linear estimates of the ISBA model, for both the observation operator (\mathcal{H} in equations 3.3 and 3.5) and the forward model (\mathcal{M} in equation 3.4). As outlined in Section 3.2, the linearised model has been approximated using the Jacobians of the model. The perturbation used to estimate the model Jacobians must be chosen carefully to ensure accurate and stable results (Balsamo et al., 2004). For the simplified EKF assimilation of screen-level observations, Mahfouf et al. (2009) demonstrated that a perturbation of $10^{-4} \times (w_{fc} - w_{wilt})$ was most suitable for estimating the linearised observation operator, and the suitability of using the same perturbation size for the EKF assimilation of near-surface soil moisture is tested here. Following Balsamo et al. (2004), the accuracy of estimating \mathbf{M} with a given perturbation size

³Note the relationship to equation 3.5: in response to the information gained from the analysis, \mathbf{P}^f is reduced by the transpose of the observation sensitivity.

has been tested by examining the agreement between the Jacobians estimated using both positive and negative perturbations, since the difference between the two estimates gives a measure of the nonlinearity of \mathcal{M} for perturbations of that size. Both \mathbf{H} for near-surface soil moisture and \mathbf{M} must be tested. However since \mathbf{H} is included in \mathbf{M} , only \mathbf{M} will be explicitly considered here.

The Jacobian terms estimated with positive and negative perturbations of $10^{-4} \times (w_{fc} - w_{wilt})$ were compared for each of the four assimilation cycles on the first day of the assimilation experiments. For each assimilation cycle, the two estimates were very close, consistent with \mathcal{M} being well approximated by \mathbf{M} within the range of the applied perturbation. The greatest disagreement occurred for the 06:00-12:00 UTC forecast, and the Jacobians at this time are compared in Figure 5.3 and summarised in Table 5.2. In Table 5.2 there is little difference between the mean, standard deviation, and extreme values of the two estimates of \mathbf{M} . Additionally, the scatterplot in Figure 5.3 shows very good agreement between the two estimates of \mathbf{M} , with nearly all points aligned on the one-to-one line. For $\partial w_1(t+6)/\partial w_2(t)$ several data points stand out to the right of the one-to-one line, indicating that the positive perturbation generated larger estimates. However, each plot contains 39417 data, the majority of which are aligned close to the one-to-one line: less than 0.2% of the $\partial w_1(t+6)/\partial w_2(t)$ estimates from the positive and negative perturbations differed by more than $0.01 \text{ m}^3\text{m}^{-3} / \text{m}^3\text{m}^{-3}$.

The above analysis indicates that \mathcal{M} is well approximated by \mathbf{M} within the range of the very small perturbations that were applied. However, this does not guarantee that \mathbf{M} approximates \mathcal{M} well when applied to the errors in \mathbf{P} , since these errors are typically much larger than the applied perturbations. To test the potential error generated when \mathbf{M} (or \mathbf{H}) is used to propagate \mathbf{P} , the model Jacobians estimated using perturbations with magnitude similar to the expected model error ($10^{-1} \times (w_{fc} - w_{wilt}) \sim O(10^{-2})$) have been compared to the above estimates, with the summary statistics again included in Table 5.2. Scatterplots in Figure 5.4 compare the Jacobian estimates derived from this larger perturbation to the original estimates (with perturbation $+10^{-4} \times (w_{fc} - w_{wilt})$) for the 06:00-12:00 UTC assimilation cycle, showing that there is now more scatter about the one-to-one line. Additionally, for $\partial w_1(t+6)/\partial w_2(t)$ there are data-points for which the smaller perturbation generated values close to zero, while the larger perturbation generated much larger values, ranging up to one; and for $\partial w_2(t+6)/\partial w_2(t)$ there are points for which the smaller perturbation

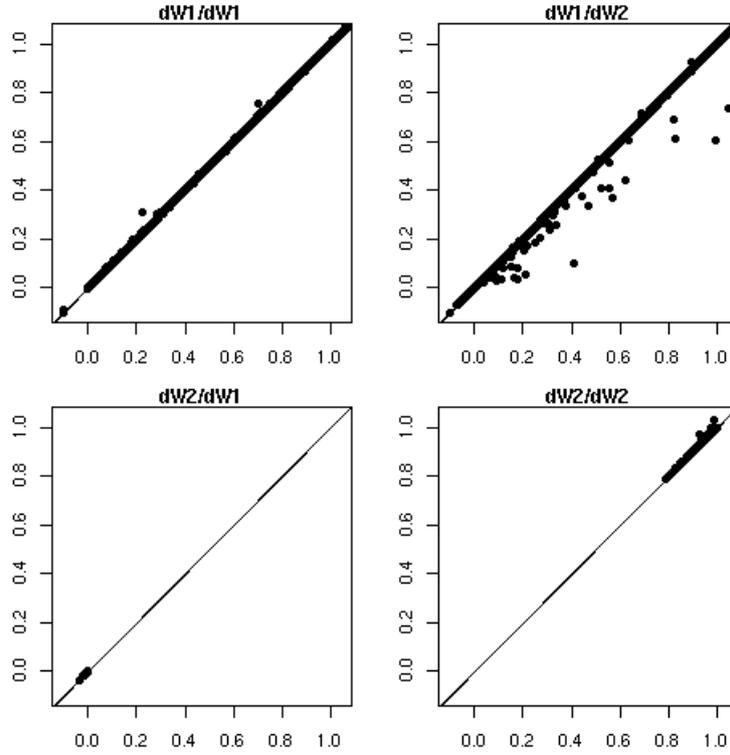


Figure 5.3: Comparison of \mathbf{M} (m^3m^{-3} / m^3m^{-3}) from 06:00 to 12:00 UTC on 1 July 2006, estimated using positive (x-axes) and negative (y-axes) perturbations of $10^{-4} \times (w_{fc} - w_{wilt})$.

generated values close to one, while the larger perturbation generated smaller values, down to 0.16. While these points stand out in the scatterplots, they represent only a small proportion of the data, and less than 2% (1%) of the $\partial w_1(t+6)/\partial w_2(t)$ ($\partial w_2(t+6)/\partial w_2(t)$) estimates from the smaller and larger perturbations differed by more than $0.01 m^3m^{-3} / m^3m^{-3}$. These differences are not large enough to impact the statistics in Table 5.2, which are very similar for the two estimates. In conclusion, while the larger perturbation was outside the model's linear regime in more instances, the Jacobian estimates compare favourably over the majority of grid cells, indicating that \mathbf{M} estimated with the smaller perturbation of $10^{-4} \times (w_{fc} - w_{wilt})$ led to an acceptable approximation of nonlinear \mathcal{M} for propagating \mathbf{P} through each assimilation cycle.

Table 5.2: Statistics of the 6-hour Jacobian terms (m^3m^{-3} / m^3m^{-3}) from 06:00 to 12:00 UTC on 1 July 2006, estimated using a perturbation size of $+10^{-4} \times (w_{fc} - w_{wilt})$ (pos), $-10^{-4} \times (w_{fc} - w_{wilt})$ (neg), and $+10^{-1} \times (w_{fc} - w_{wilt})$ (lrg).

		mean	stdev	min	max
$\partial w_1(t+6)/\partial w_1(t)$	+ve	0.35	0.42	-0.10	1.1
	-ve	0.35	0.42	-0.10	1.1
	lrg	0.36	0.42	-0.10	1.1
$\partial w_2(t+6)/\partial w_1(t)$	+ve	-0.0017	0.0015	-0.036	0.0
	-ve	-0.0017	0.0015	-0.036	0.0
	lrg	-0.0017	0.0015	-0.037	0.0
$\partial w_1(t+6)/\partial w_2(t)$	+ve	0.34	0.25	-0.10	1.1
	-ve	0.34	0.25	-0.10	1.1
	lrg	0.34	0.24	-0.072	1.1
$\partial w_2(t+6)/\partial w_2(t)$	+ve	0.99	0.014	0.78	1.0
	-ve	0.99	0.014	0.78	1.0
	lrg	0.99	0.015	0.16	1.0

5.3.2 The ISBA model Jacobians

It is well established that screen-level variables are related to root-zone soil moisture via transpiration. However the relationship between near-surface and root-zone soil moisture is not as well understood (for either reality, or specific models such as ISBA). Consequently, this section examines the manner in which the EKF translates near-surface soil moisture observations into root-zone soil moisture increments in ISBA. This is achieved by examining the ISBA model Jacobians, to first examine the inter-dependencies between w_1 and w_2 in ISBA, and then to compare the observation operators for assimilating w_1 and screen-level observations.

Soil moisture coupling

Table 5.3 shows the spatial mean model Jacobian terms for 1 July 2006, for a 6 hour period during both the day and night, and for a 24 hour period. Since the superficial soil layer responds rapidly to atmospheric forcing, w_1 has less memory during the day when the surface is more strongly forced, and the nighttime spatial mean $\partial w_1/\partial w_1$ of $0.80 m^3m^{-3} / m^3m^{-3}$ is reduced to $0.35 m^3m^{-3} / m^3m^{-3}$ during the day. Over the 24 hour forecast from 00:00 UTC,

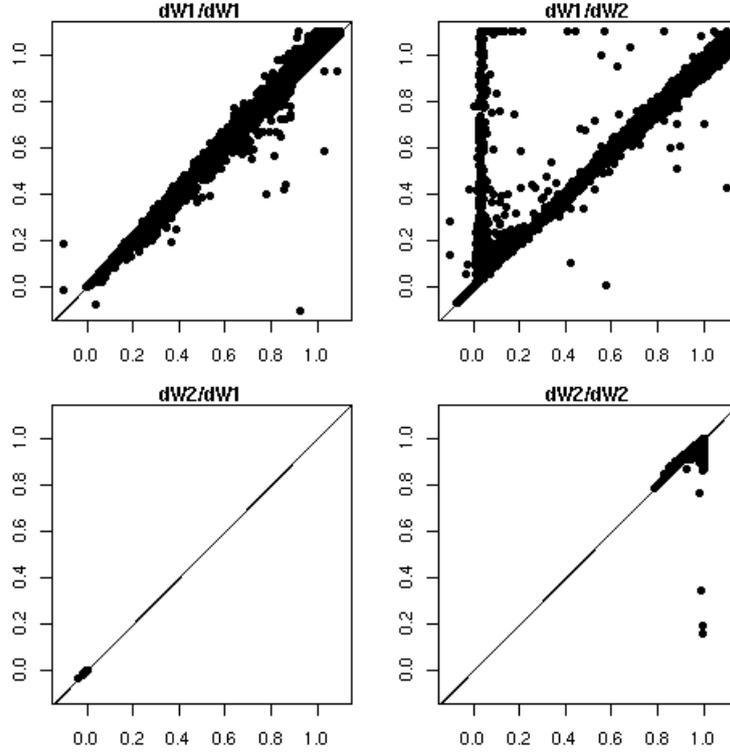


Figure 5.4: Comparison of M (m^3m^{-3} / m^3m^{-3}) from 6:00 to 12:00 UTC on 1 July 2006, estimated using perturbations of $10^{-4} \times (w_{fc} - w_{wilt})$ (x-axes) and $10^{-1} \times (w_{fc} - w_{wilt})$ (y-axes).

w_1 has little memory, and the spatial mean $\partial w_1 / \partial w_1$ is $0.25 m^3m^{-3} / m^3m^{-3}$. In addition to its short time scale, the small reservoir of w_1 cannot exert much influence on the more substantial w_2 , so that $\partial w_2 / \partial w_1$ is extremely small (mean $< |0.01| m^3m^{-3} / m^3m^{-3}$) over all of the time periods considered.

In contrast to w_1 , the atmospheric forcing is applied more slowly to w_2 , which has a time scale of about ten days. Over the comparatively shorter periods considered here, the w_2 perturbations were largely retained (mean $\partial w_2(t + 24) / \partial w_2(t)$: $0.95 m^3m^{-3} / m^3m^{-3}$ over 24 hours). Additionally, w_2 has a clear influence on w_1 , with a mean $\partial w_1(t + 6) / \partial w_2(t)$ of 0.20 and $0.34 m^3m^{-3} / m^3m^{-3}$ for the 6 hours from 18:00 and 06:00 UTC, respectively. The influence of w_2 increased over time, giving $\partial w_1(t + 24) / \partial w_2(t)$ of $0.60 m^3m^{-3} / m^3m^{-3}$ for the 24 hour forecast. Comparing this to $\partial w_1(t + 24) / \partial w_1(t)$ highlights that the initial w_2 has a stronger influence on the 24 hour forecasts of w_1 than the initial w_1 does.

The above results have several consequences for analysing w_2 from w_1 observations. The short memory and limited influence of the ISBA w_1 was noted

by Balsamo et al. (2007) and Mahfouf et al. (2009), who did not update w_1 . As noted by Calvet and Noilhan (2000), an additional consequence is that assimilating w_1 into ISBA relies on updating w_2 based on the sensitivity of w_1 forecasts to w_2 , in much the same way that assimilating screen-level observations makes use of the influence of w_2 on the screen-level forecasts.

Table 5.3: Spatial mean \mathbf{M} (m^3m^{-3} / m^3m^{-3}) for a 6 hour forecast during the night (upper), and day(centre), and for the 24 hour period \mathbf{M} (lower) from 18:00 UTC on 1 July 2006.

$$M_{18 \rightarrow 00UTC} = \begin{pmatrix} 0.80 & 0.20 \\ 0.00 & 0.99 \end{pmatrix}$$

$$M_{06 \rightarrow 12UTC} = \begin{pmatrix} 0.35 & 0.34 \\ -0.002 & 0.99 \end{pmatrix}$$

$$M_{18 \rightarrow 18UTC} = \begin{pmatrix} 0.25 & 0.60 \\ 0.00 & 0.95 \end{pmatrix}$$

The observation operator

Figure 5.5 shows the Jacobians of the observation operator for each observation type for observations at 00:00 and 12:00 UTC on 1 July 2006. For the screen-level observations the Jacobian maps show the expected dependency on evapotranspiration. As outlined by Mahfouf et al. (2009), in ISBA the surface canopy resistance to transpiration is nonlinearly dependent on the root-zone soil moisture, with stronger dependencies close to the wilting point, so that the screen-level variables are most sensitive to w_2 in clear sky regions with a high fraction of vegetation cover and dry soils. For example, comparing the model Jacobians in Figure 5.5 to the map of the Surface Wetness Index in Figure 5.6 shows that in north central Europe and England, where the vegetation fraction is greater than 0.75, the daytime Jacobians are higher where the SWI is close to zero (recall from Section 3.4.1 that $SWI = (w_2 - w_{wilt}) / (w_{fc} - w_{wilt})$). In contrast, in sparsely vegetated Spain where the vegetation fraction is less than 0.5, the SWI is similarly low and yet there is no response in the Jacobians. Additionally, in those regions where w_2 is below the wilting point (most prominently in Africa, inland Spain, and close to the French Atlantic coast), transpiration

ceases so that the screen-level Jacobians in Figure 5.5 are close to zero during the day.

The role of evapotranspiration is reflected in the mean Jacobians for each of the four daily assimilation cycles in Table 5.4. The screen-level Jacobians are reduced at 18:00 to 00:00 UTC and 00:00 to 06:00 UTC, although there was still

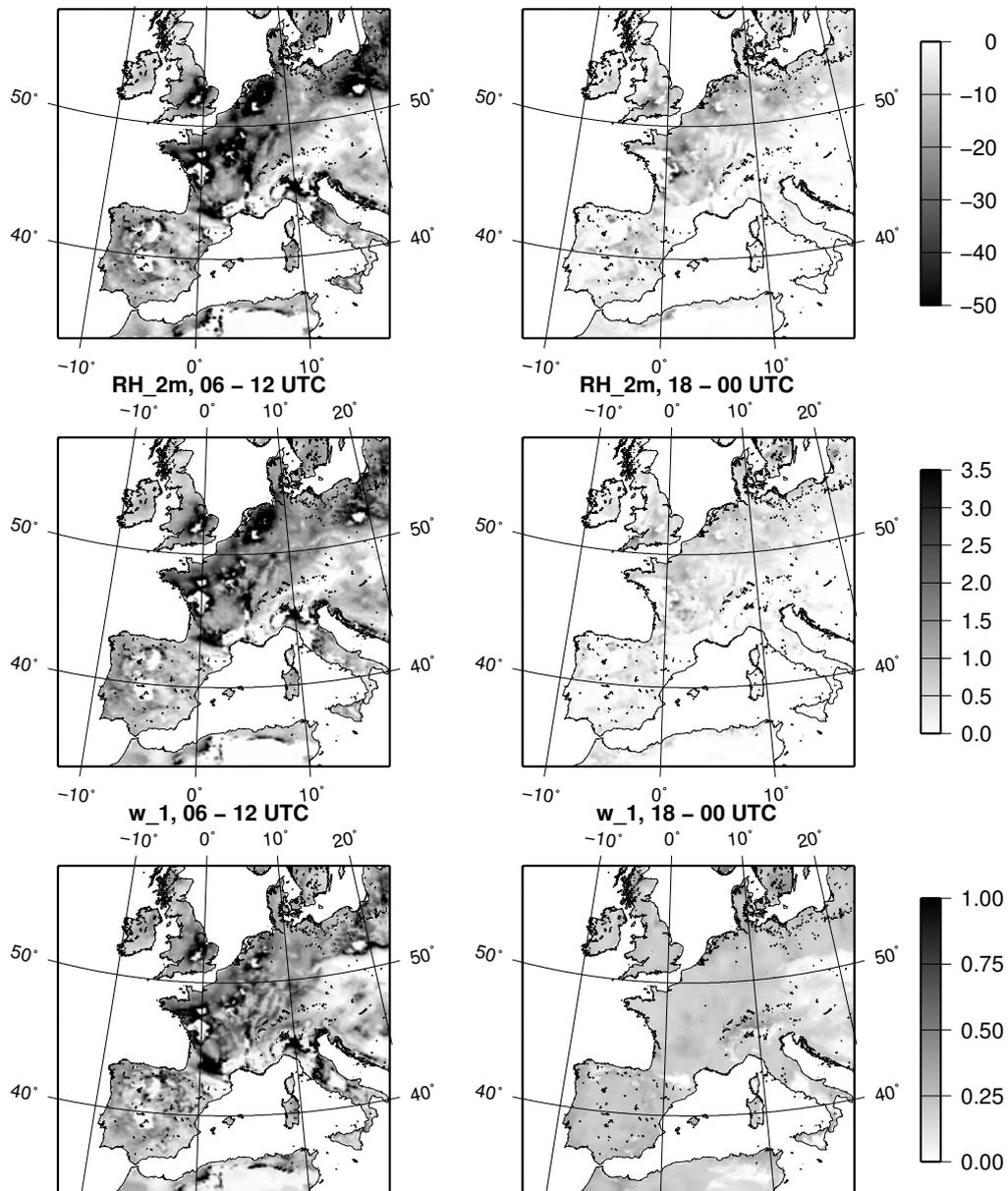


Figure 5.5: Maps of the observation operator for modeled w_2 and observed T_{2m} (upper), RH_{2m} (middle), and w_1 (lower), for the analyses at 06:00 UTC (left) and 18:00 UTC (right) on 1 July 2006.

Table 5.4: Mean observation operator relating w_2 to each observation type from experiment CMB, averaged over July 2006 separately for each of the four daily assimilation cycles.

Time	T_{2m} ($\text{K}/\text{m}^3\text{m}^{-3}$)	RH_{2m} ($\%/ \text{m}^3\text{m}^{-3}$)	w_1 ($\text{m}^3\text{m}^{-3}/\text{m}^3\text{m}^{-3}$)
00-06	-2.6	0.8	0.16
06-12	-17.3	1.1	0.30
12-18	-19.1	1.7	0.33
18-00	-2.9	0.1	0.17

a small response since both time periods include some day-light. The exception is the 00:00 to 06:00 UTC Jacobian for RH_{2m} , which had a relatively large mean of $80 \%/(\text{m}^3\text{m}^{-3})^{-1}$, compared to $110\text{-}170\%/(\text{m}^3\text{m}^{-3})^{-1}$ for the subsequent assimilation cycles, suggesting a rapid humidity response at sunrise.

For the w_1 observation operator, w_2 directly influences w_1 in equation 3.8 via the w_1 restore term, representing the balance between gravitational drainage and capillary rise. However, there is a strong and unexpected similarity between the observation operators for w_1 and those for the screen-level variables in Figure 5.5. Again the highest values (up to $1.0 \text{ m}^3\text{m}^{-3} / \text{m}^3\text{m}^{-3}$) occur in dry vegetated regions, while regions with negative surface wetness index have very low values. This suggests that the greatest influence of w_2 on w_1 does not arise from the direct restore term as expected, but from an indirect relationship via evapotranspiration. Tests on the forecast impact of switching off various as-

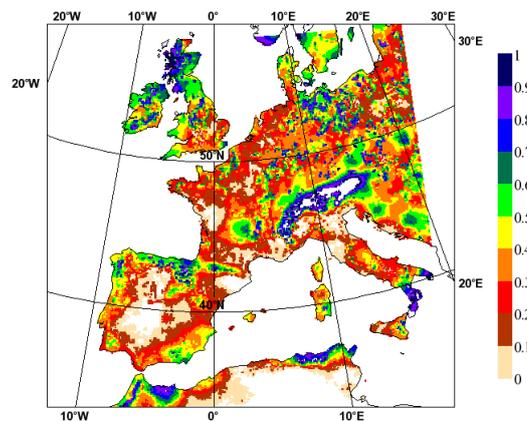


Figure 5.6: Surface wetness index at 18:00 on 1 July 2006 from the CTR simulation.

pects of the model physics have confirmed this finding, and also revealed that the dependency arises from the influence of transpiration on the surface temperature. The mechanism by which this occurs is that increasing w_2 increases transpiration, which reduces the surface temperature, thus reducing the depletion of w_1 by bare-ground evaporation, giving a relative increase in w_1 , and a positive dependency of w_1 and w_2 (pers. comm Jean-François Mahfouf).

During the night the restore term introduces a very weak sensitivity of w_1 to w_2 , with the mean $\partial w_1(t+6)/\partial w_2(t)$ close to half the daytime value in Table 5.4. In Figure 5.5 $\partial w_1(t+6)/\partial w_2(t)$ is less than $0.3 \text{ m}^3\text{m}^{-3} / \text{m}^3\text{m}^{-3}$ across most (93%) of the domain, although there are a few regions with slightly higher values of up to $0.4 \text{ m}^3\text{m}^{-3} / \text{m}^3\text{m}^{-3}$. All of these regions have low clay content (<10%), so that the soil texture dependent coefficient of the w_1 restore term (C_2 in equation 3.8) will restore w_1 towards w_2 more rapidly.

As noted above, the assimilation of w_1 into ISBA relies on updating w_2 rather than w_1 . Consequently, the effectiveness with which the ISBA w_2 can be updated from w_1 observations is limited by the strength of $\partial w_1/\partial w_2$, since this is the observation operator and it also controls the evolution of error cross covariance between w_1 and w_2 in \mathbf{P} . One of the main motivations for assimilating near-surface soil moisture was the expectation that it will provide a more direct observation of total soil moisture than the screen-level observations, since the latter must rely on the model flux parameterisations to link the root-zone to the screen-level atmosphere. However the above analysis has indicated that this is not the case for ISBA, since the dominant link between w_2 and w_1 is provided by evapotranspiration during the day. This relationship is more likely to be a consequence of the use of a single surface temperature for both the soil and vegetation surface in ISBA, than due to a true physical relationship between transpiration and bare soil evaporation. Given this indirect relationship between w_1 and w_2 during the day, ISBA will likely benefit more from the assimilation of nighttime w_1 data than daytime data. Despite this, both the ascending and descending overpass AMSR-E observations have been assimilated in these experiments, in response to the relatively weak relationship between w_1 and w_2 in ISBA.

An approximation to reduce to number of model Jacobians calculated

An approximation to the EKF has been introduced for the ISBA experiments to reduce the computational cost of the assimilation. For each assimilation cy-

cle the EKF requires two separate estimates of the linearised forecast model, necessitating two sets of perturbed forecasts: one made before the analysis to estimate \mathbf{H} for use in equation 3.2, and one made after the analysis to estimate \mathbf{M} for use in equation 3.4. The cost of the assimilation can then be reduced by estimating \mathbf{M} based on the perturbed model simulations made prior to the analysis (neglecting changes to \mathbf{M} due to the analysis update). To test the impact of this approximation, assimilation results obtained with the above approximation have been compared to a reference run in which \mathbf{M} is estimated using perturbed model integrations made after the analysis update. For 1 July 2006, the mean absolute difference between the approximate and reference \mathbf{M} was small ($O(10^{-3})$ and $O(10^{-5})$ for the daytime and nighttime assimilation cycles, respectively). These differences had a limited impact on the analyses, and after two weeks of assimilating T_{2m} , RH_{2m} , and AMSR-E w_1 (using the CMB experimental set-up introduced in section 5.2.3), the mean absolute difference between the two estimates of w_2 was $1.0 \times 10^{-4} \text{ m}^3\text{m}^{-3}$, with an absolute difference less than $1.0 \times 10^{-3} \text{ m}^3\text{m}^{-3}$ across 99% of the model domain. This is several orders of magnitude less than the impact of the assimilation: the mean absolute difference between w_2 from the reference assimilation and an open-loop simulation after two weeks was close to $1.0 \times 10^{-2} \text{ m}^3\text{m}^{-3}$. Hence this approximation is considered acceptable, and has been adopted for the ISBA assimilation experiments in this Chapter.

5.4 Initial results: Assimilating AMSR-E soil moisture

This section presents the initial results from the EKF assimilation of AMSR-E near-surface soil moisture data. First, the rescaling of the AMSR-E observations to reduce the systematic differences between the modeled and observed soil moisture is presented. Then the EKF assimilation of AMSR-E into ISBA is demonstrated, before the EKF and Simplified EKF (SEKF) assimilation of the AMSR-E observations are compared to determine the role of evolving the background error covariances in the assimilation. The results from assimilating the AMSR-E data are examined in further detail, and also compared to the assimilation of screen-level observations, in the subsequent section (Section 5.4.2).

5.4.1 Rescaling the AMSR-E data

The CDF-matching strategy

Since the EKF is derived from the assumption that the model and observations are not biased relative to each other, the AMSR-E w_1 observations have been normalised to better match the model climatology, by matching the observed CDF (Reichle and Koster, 2004; Drusch et al., 2005) to that of the superficial soil moisture forecasts from ALADIN. Ideally, a long data set is used to sample the model and observation climatology, and the CDF-matching is performed on as localised a scale as possible. However, for this study only one year of ALADIN soil moisture fields were available. To compensate for the truncated temporal sample, ergodic substitution has been used to expand the sample for estimating the CDF-matching operator, with the operator at each grid estimated using data from the surrounding one-degree window, following Reichle and Koster (2004). In Chapter 4 the climatology of ascending and descending AMSR-E data differed (for example in Figure 4.10). The ISBA w_1 also has a substantial diurnal cycle, and to ensure that the CDF-matched observations do not include a diurnal bias, the CDF-matching has been performed separately for each overpass.

CDF-matching is based on the assumption that the systematic differences between the model and observations are stationary. However, for ALADIN and AMSR-E this is not the case. For example, Figure 5.7 shows a time series of the descending overpass AMSR-E data before and after CDF-matching at a location in north France. While both the original AMSR-E and ALADIN time series have a similar range of short-term (up to several days) variability, with amplitude between 0.1 and 0.2 m^3m^{-3} , the seasonal cycle in the AMSR-E data has a greater amplitude ($>0.2 m^3m^{-3}$) than that in the ALADIN data ($\sim 0.1 m^3m^{-3}$). To compensate for the variance generated by the exaggerated seasonal cycle in the AMSR-E data, the CDF-matching has overly dampened the short-term variability, resulting in a lessened response to rain events in the CDF-matched time series. Note that there are also several monthly to seasonal scale periods of consistent bias in the CDF-matched time series (e.g., around day 100 in Figure 5.7), which cannot be attributed to either the observations or the model.

To avoid these problems, the CDF-matching was repeated using seasonal-bias corrected AMSR-E data. If sufficient data were available, the AMSR-E seasonal cycle would be best corrected based on the climatological seasonal

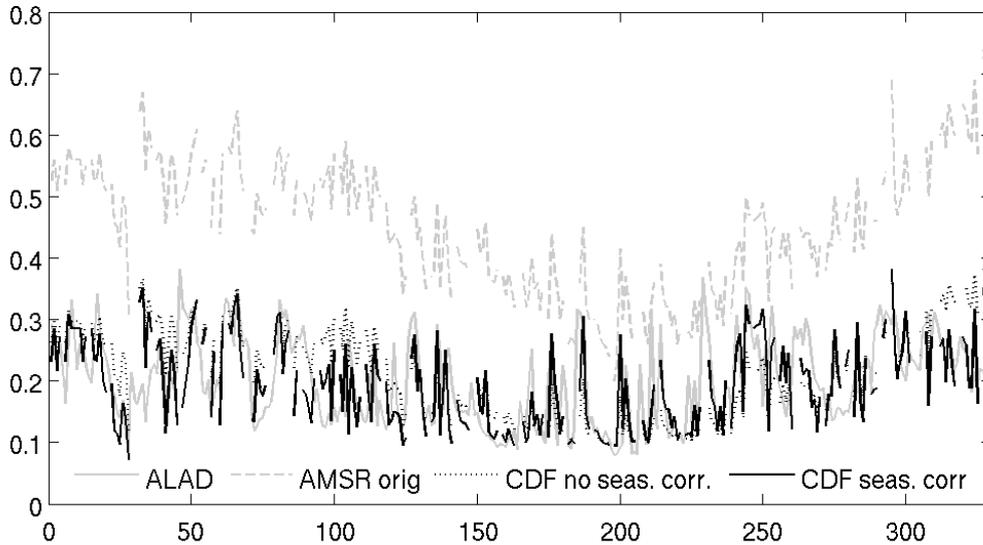


Figure 5.7: Time series of near-surface soil moisture (m^3m^{-3}) for a grid-cell in France (47.30 E/0.06 N) over 2006, from ALADIN (grey, solid), the original AMSR-E data (grey, dashed), and the seasonal-bias corrected (black, solid) and nonseasonal-bias corrected (black, dotted) CDF-matched AMSR-E data.

cycle of the model and the observations, as this would retain any seasonal scale anomalies in the AMSR-E data. However, with only one year of ALADIN w_1 fields available, seasonal scale bias anomalies cannot be detected, regardless of the method used to rescale the data. Consequently, a conservative approach was taken here, by assuming the ISBA 2006 seasonal cycle to be correct, and subtracting the difference between the 31 day moving average of the modeled and observed w_1 from the AMSR-E time series. Figure 5.7 includes the time series of the seasonal-bias corrected and then CDF-matched time series, showing that it has retained an appropriate response to precipitation (while also reducing the monthly biases).

The CDF-matching reduced the mean bias between the AMSR-E and ALADIN w_1 for July 2006, from $-0.14 m^3m^{-3}$ in the original AMSR-E data, to $0.014 m^3m^{-3}$ in the seasonal-bias corrected and CDF-matched data for the descending overpass, and from $-0.12m^3m^{-3}$ to $0.013m^3m^{-3}$ for the ascending AMSR-E overpass. For 2006, the RMSD between the CDF-matched AMSR-E and ALADIN soil moisture was $0.058 m^3m^{-3}$ and $0.072 m^3m^{-3}$, for the descending and ascending overpasses respectively (compared to standard deviations of $0.09 m^3m^{-3}$ and $0.10 m^3m^{-3}$ for the ALADIN w_1 at the corresponding times).

Impact of the rescaling strategy

To highlight the importance of the strategy adopted for bias correcting the assimilated w_1 data, preliminary experiments have compared the assimilation of the seasonal-bias corrected and nonseasonal-bias corrected AMSR-E observations. For these experiments, the observation and background soil moisture errors (in all layers) were assumed to be equal (rather than using the errors described for the AMS experiment in Table 5.1). For \mathbf{R} , the same w_1 observation error variance of $(0.05 \text{ m}^3\text{m}^{-3})^2$ as for AMS was used, and the diagonal elements of \mathbf{P} were initialised at $(0.6 \times (w_{fc} - w_{wilt}))^2$ (giving a mean error of $0.052 \text{ m}^3\text{m}^{-3}$). The diagonal elements of \mathbf{Q} were then tuned to maintain \mathbf{P} at close to this value, giving $(0.3 \times (w_{fc} - w_{wilt}))^2$ and $(0.2 \times (w_{fc} - w_{wilt}))^2$ for the w_1 and w_2 error variances, respectively. The off-diagonal elements of \mathbf{P} and \mathbf{Q} were again assumed to be zero. While these model errors are clearly too large (see for example, the discussion in Section 5.2.3), this experimental set up is sufficient to highlight the impact of the bias-correction technique for AMSR-E.

Figure 5.8 compares the net monthly soil moisture increments generated by assimilating the seasonal-bias corrected and nonseasonal-bias corrected AMSR-E observations over July 2006. The net monthly increments added by each were very different. The seasonal-bias corrected data generating net positive increments across most of the domain, with negative increments in Spain, north Africa, and small regions in the east, while the nonseasonal-bias corrected data generated net negative increments almost everywhere except north Europe. Consequently, assimilating the seasonal-bias corrected data added a mean monthly total of 55 mm of water, while assimilating the nonseasonal-bias corrected data removed a mean of 6.2 mm of water from w_2 .

The substantial differences between the two plots in Figure 5.8 highlight that the outcome of an assimilation is limited by the quality of the assimilated data, and consequently the strategy used to bias-correct soil moisture data can have a profound impact on the analysis results.

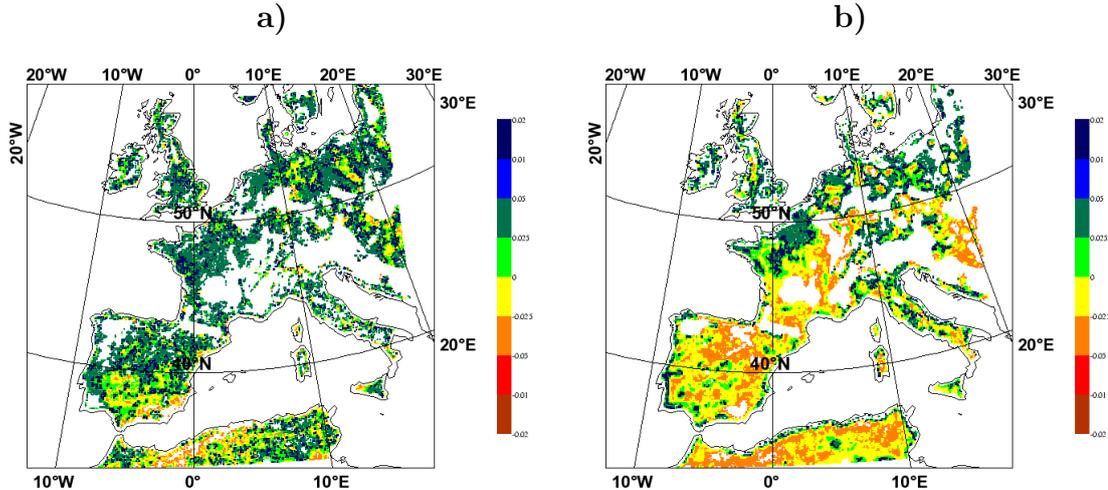


Figure 5.8: Net monthly w_2 increments (m^3m^{-3}) over July 2006, from the assimilation of a) seasonal-bias corrected AMSR-E data, and b) nonseasonal-bias corrected AMSR-E data (assuming approximately equal observation and background errors).

5.4.2 Comparison of the EKF and SEKF assimilation of AMSR-E w_1

The EKF assimilation of AMSR-E w_1

Here, the results of the AMS experiment to assimilate the AMSR-E data are briefly presented to establish that the EKF can extract sufficient information from observations of w_1 to make significant updates to w_2 in ISBA.

The upper panels of Figure 5.9 show maps of the Kalman gain for w_2 (\mathbf{k}_2) and the resulting net monthly analysis increments generated by the AMS experiment, respectively. The Kalman gain is plotted close to the beginning and end of the one month assimilation, at 18:00 UTC on 2 and 30 July 2006. In both cases the EKF gain was consistently below $0.2 m^3m^{-3} / m^3m^{-3}$, with a mean of $0.09 m^3m^{-3} / m^3m^{-3}$ on 2 July, and $0.11 m^3m^{-3} / m^3m^{-3}$ on 30 July. Given the ratio of the depths of the w_1 and w_2 soil layers, these gains will result in reasonably large volumes of moisture being added to w_2 . The depth of w_1 is 0.4% of the average depth of the w_2 layer, so that a Kalman gain of $0.1 m^3m^{-3} / m^3m^{-3}$ will convert an error of 1 mm in w_1 into an average update of 23 mm to w_2 .

The mean monthly increment to w_2 from assimilating the AMSR-E data was 24 mm, which is very close to the mean monthly surface water budget (precipitation minus evaporation and runoff) of 30 mm. There is considerable

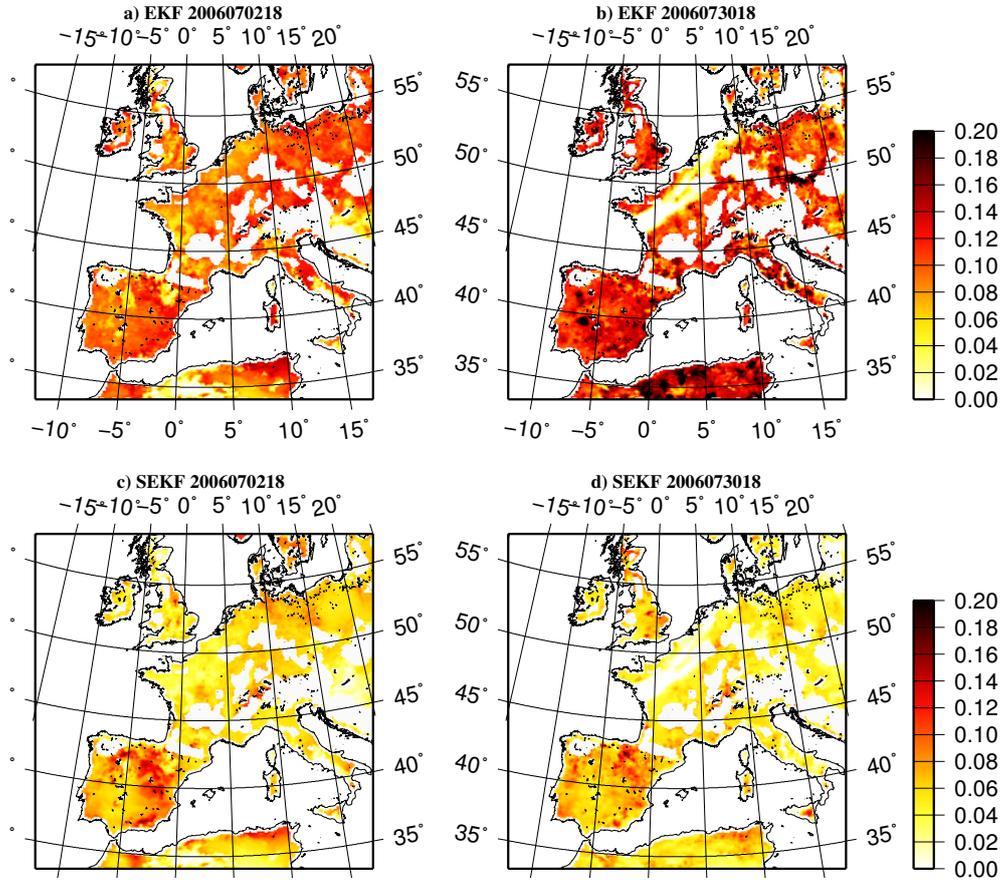


Figure 5.9: Kalman gain (m^3m^{-3}/m^3m^{-3}) for w_2 for the EKF at 18:00 on a) 2 July 2006 and b) 30 July 2006, and for the SEKF on c) 2 July 2006 and d) 30 July 2006.

spatial variability in the increments shown in Figure 5.10, giving a standard deviation of the net monthly increments of 49 mm. The analysis increments generated by the AMSR-E assimilation are not examined further here, and will instead be explored in greater detail in Section 5.5.

The SEKF assimilation of AMSR-E w_1

The SEKF of Mahfouf et al. (2009) has been extended into an EKF for the assimilation of w_1 data, to allow the evolution of long-term memory in the background model error covariance between w_1 and w_2 . For the AMS experiment, the square root of the w_1 - w_2 error covariance was typically about half the w_2 error standard deviation: the mean values for the square root of each element of \mathbf{P} over the experiment (excluding the first few days) were 0.024 for

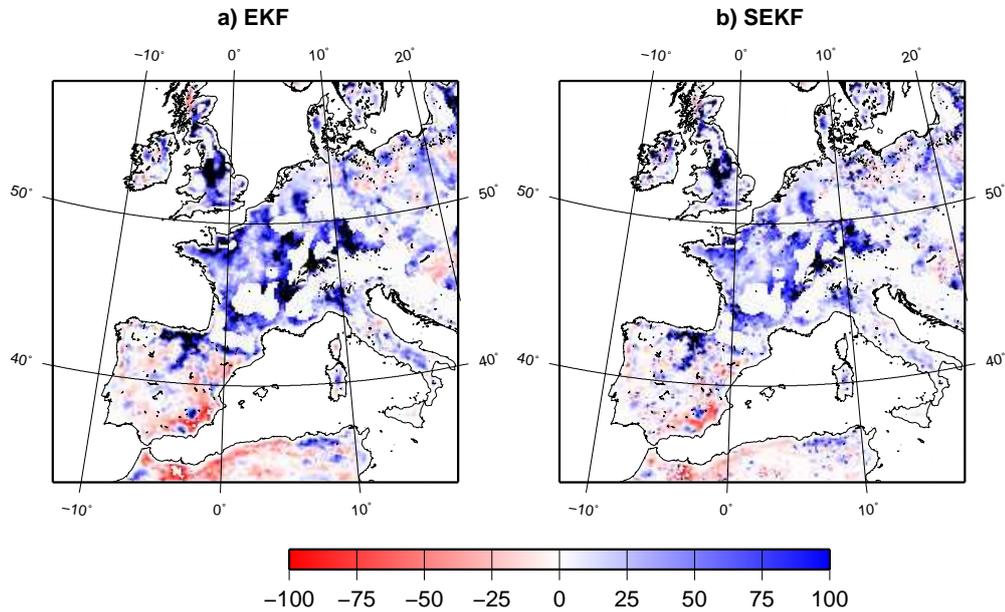


Figure 5.10: Maps of the w_2 analysis increments (mm) accumulated over July 2006 by the a) EKF and b) SEKF assimilation of AMSR-E w_1 observations.

\mathbf{P}_{11} , 0.009 for \mathbf{P}_{22} , and 0.005 for \mathbf{P}_{21} . In contrast to the EKF, the SEKF of Balsamo et al. (2007) and Mahfouf et al. (2009) does not use equations 3.4 and 3.5 to evolve and update \mathbf{P} , and instead uses the same (diagonal) \mathbf{P} at the beginning of each assimilation cycle. While this neglects the long term evolution of the model errors, it still allows some temporal evolution of \mathbf{P} through the inclusion of the (6 hour) forecast model in the observation operator (see Appendix A). To directly test the impact of the off-diagonal error covariance on the EKF assimilation, its impact has been isolated by repeating the AMS experiment with w_1 excluded from the update vector. This reduced the mean monthly increment to 18 mm month^{-1} , with a standard deviation of 43 mm month^{-1} , indicating that the error covariance between w_1 and w_2 does enhance the correction of w_2 from w_1 observations.

To further test the impact of using the full EKF, the AMS experiment has been repeated with a SEKF. For the SEKF, the diagonal error variances in \mathbf{P} were set at $(0.2 \times (w_{fc} - w_{wilt}))^2$ and $(0.1 \times (w_{fc} - w_{wilt}))^2$ for w_1 and w_2 , respectively (close to the mean diagonal \mathbf{P}^f elements for AMS quoted above). The off-diagonal elements of \mathbf{P} were assumed to be zero, following Balsamo et al. (2007) and Mahfouf et al. (2009). The Kalman gains for the SEKF assimilation

experiments are included in Figure 5.9. The SEKF gain was consistently smaller than the EKF gain, with means of 0.05 and 0.06 m^3m^{-3}/m^3m^{-3} on 2 and 30 July. Apart from the magnitude, the most apparent difference between the Kalman gains is that evolving \mathbf{P} increased the fine scale spatial heterogeneity in the EKF gain. This was evident after just two days of assimilation, although it was more pronounced on 30 July. At larger scales the maps of the EKF and SEKF Kalman gain show similar regions of comparatively high and low values, particularly on 2 July when both are relatively low over much of north Africa, Spain, and north France. However, the evolution of \mathbf{P} has introduced some differences, and on 30 July the EKF gain was comparatively high in the arid zones of north Africa and Spain (the prominent region of reduced gain in both plots over north France on 30 July was caused by precipitation). As a result of the smaller SEKF gain, the net analysis increments in Figure 5.10 are slightly smaller than for the EKF, with a mean and standard deviation of 18 and 44 mm, respectively (compared to 24 mm and 49 mm, respectively for the EKF). However, consistent with the similarity in their Kalman gains, the large scale patterns of moisture addition and subtraction for the EKF and the SEKF were very similar.

In summary, using an EKF rather than a SEKF allows for the development of long-term memory in the background model error covariance between w_1 and w_2 , which for the AMS experiment was close to half the magnitude of the w_2 error variance. Comparing the EKF assimilation of AMSR-E data to the assimilation with w_1 excluded from the update vector indicates that this off-diagonal error covariance has enhanced \mathbf{k}_2 . While the Kalman gain, and subsequent analysis updates, were smaller for the SEKF assimilation of AMSR-E, the broad spatial patterns in the analysis updates were similar to those from the EKF assimilation. This similarity between the SEKF and EKF assimilation is a consequence of the use of a 1-D assimilation. For a 1-D assimilation of near-surface soil moisture, the temporal evolution of \mathbf{P} can only affect the vertical profile of the analysis increments. In contrast, for (3-D) atmospheric data assimilation the main benefit of evolving \mathbf{P} to account for “errors of the day” is the ability to capture the spatial correlations generated by horizontal flow (Kalnay, 2003). For the 2-layer ISBA model, evolving \mathbf{P} can only influence the magnitude of the update to w_2 from a given w_1 observation, and these differences in \mathbf{k}_2 are unlikely to affect the broad scale patterns in the addition and subtraction of moisture. In contrast, the processing of the observations can

have a more profound impact on the net monthly analysis increments than the evolution of \mathbf{P} : note that the method used to rescale the observations had a much greater impact on the spatial patterns in the analysis increments in Figure 5.8 than the model error covariance evolution did in Figure 5.10.

Despite the minimal impact of evolving the background error covariances in this experiment, the EKF is recommended for future work and it has been used throughout this thesis, since with a 6 hour assimilation cycle it does not incur any extra cost if \mathbf{M} is approximated from the same perturbed model simulations used to estimate \mathbf{H} (Section 5.3.2). Additionally, the evolution of \mathbf{P} is expected to generate more realistic error covariances, while also allowing greater flexibility in the specification of errors, which should further improve the background error covariances (for example, Mahfouf (2010a) present a parameterisation of \mathbf{Q} , based on precipitation).

5.5 Comparison of the assimilation of AMSR-E near-surface soil moisture and screen-level observations

In this section the results of the experiments described in Section 5.2.3 to assimilate the AMSR-E w_1 , the screen-level observations, and both, are presented and compared. The focus is on determining whether there is any consistency between the soil moisture analyses generated by assimilating each data type, and to examine how the information from each data type interacts when they are assimilated together. First, the distributions of the mean monthly observation increments for each observed variable, and then the analysis increments for each assimilation experiment, are compared. Then, the impact of assimilating each observation type on the subsequent observation increments⁴ for all observed variables is examined. Finally, the results of each experiment are compared to the surface water balance generated by forcing the ISBA model with high quality observations.

⁴The term “observation increment” has been used loosely in this chapter, to refer to the difference between the model forecasts and observations, including for observations not assimilated in a given experiment.

5.5.1 The observation increments

Histograms of the mean monthly difference across the ALADIN domain between the open-loop forecasts and observations of T_{2m} , RH_{2m} , and w_1 are shown in Figure 5.11 (for the assimilated observations this is the mean observation increment). For consistency, the screen-level increments have been excluded at locations where AMSR-E data were screened-out⁵. While the spatial mean observation increment for the screen-level variables was close to zero (-0.005 K and -0.3 %), suggesting that there was no significant net model - observation bias, there were some relatively large increments (giving standard deviations of 0.5 K and 3.6 %), indicating significant monthly biases at some locations. Qualitatively, the distributions of the observation increments for T_{2m} and RH_{2m} are approximately mirror-images, suggesting a consistent signal from w_2 . Both are skewed, with T_{2m} (RH_{2m}) having a longer positive (negative) tail, suggesting a cool-moist model bias.

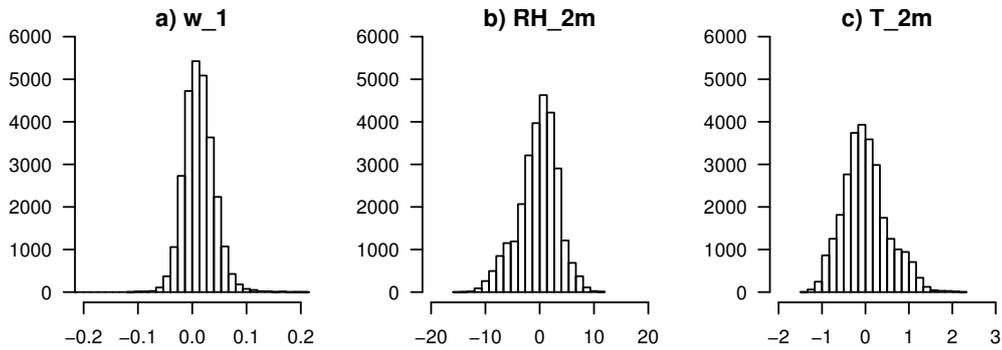


Figure 5.11: Histograms of the mean observation minus CTR forecast over July 2006 for a) T_{2m} (K), b) RH_{2m} (%), and c) w_1 ($m^3 m^{-3}$).

While the long term (monthly) mean observation increments for the screen-level observations were small, there was a significant diurnal cycle in the increments. The largest increments occurred in the early morning, when the observations suggested a cool-moist model bias (mean monthly increments at 06:00 UTC: 0.56 K and -3.5 %). As each day proceeded, this cool-moist bias was initially reduced (mean increments at 12:00 UTC: 0.09K & -1.6 %), and then reversed in the evening (mean increments at 18:00 UTC: -0.29 K & 0.6 %, and at 00:00 UTC: -0.33 K & 3.2 %). Time series at individual locations show that

⁵If all available T_{2m} and RH_{2m} data are used, the results are very similar for the subset of observations presented here.

this diurnal cycle occurs throughout the month, rather than being an artefact of biases occurring at different periods during the month. This diurnal cycle in the observation increments cannot be easily attributed to errors in w_2 , and almost certainly has another cause. In particular, the largest biases occurred early in the morning, when only RH_{2m} had an intermediate sensitivity to w_2 (and T_{2m} has very little sensitivity) in Table 5.4.

Compared to the screen-level observations, the AMSR-E observation increments in Figure 5.11 were distributed more symmetrically, although the mean was slightly positive ($0.013 \text{ m}^3\text{m}^{-3}$), indicating a tendency for the model to be drier than the observations. Additionally, there was limited spread in the AMSR-E observation increments (giving a standard deviation of $0.028 \text{ m}^3\text{m}^{-3}$). Since the rescaling of the AMSR-E data explicitly removed any diurnal cycle in the model-observation bias, the means of the observations increments were similar at 00:00 and 12:00 UTC (0.014 and $0.013 \text{ m}^3\text{m}^{-3}$, respectively).

5.5.2 The analysis increments

Histograms of the total volume of moisture added to w_2 over July 2006 from the assimilation of AMSR-E (AMS), screen-level variables (SLV), and both (CMB) are shown in Figure 5.12. For all experiments the analysis increments can be quite large, with the tails of the distribution reaching $\pm 200 \text{ mm month}^{-1}$ (recall as a comparison that the mean monthly surface water budget over the ALADIN domain was 30 mm). For SLV, the skewed T_{2m} and RH_{2m} observation increments generated an asymmetrical distribution of analysis increments with a slightly negative mode, giving a small negative mean reduction in w_2 ($-5.5 \text{ mm month}^{-1}$), with considerable spread (the standard deviation was 57 mm month^{-1}). The monthly AMS increments were generally smaller than those for SLV, although the strong positive skew generated a net addition of moisture (the mean was 24 mm month^{-1} , and the standard deviation was 49 mm month^{-1}). While the mode of the AMS increments is centred on zero, the positive tail reaches approximately $200 \text{ mm month}^{-1}$, double the negative tail ($\sim -100 \text{ mm month}^{-1}$).

The maps of the net monthly w_2 analysis increment for the AMS and SLV experiments in Figure 5.13 show little spatial agreement between the increments generated by assimilating each data set. While they both removed moisture across most of arid Spain and north Africa, there is little agreement over central

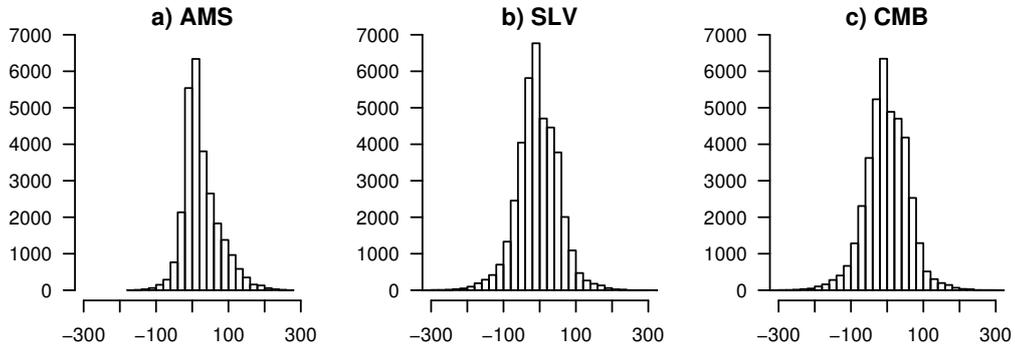


Figure 5.12: Histograms of the w_2 analysis increments accumulated over July 2006 (mm) from the a) AMS, b) SLV, and c) CMB assimilation experiments.

and east Europe, where SLV tended to decrease w_2 and AMS increased it. It is interesting to note that the regions of strongest moisture reduction by SLV correspond to locations where the AMSR-E data was screened out due to mountainous terrain and/or dense vegetation. Additionally, AMS generated large positive increments along the northern edge of the Alps, which may reflect errors in the model and/or the observations, since both have known problems in regions of steep terrain (Rüdiger et al., 2009). The lack of association between the net monthly SLV and AMS analysis increments has also been confirmed by a scatterplot (not shown), for which the correlation coefficient between the net SLV and AMS increments was just 0.1. Additionally, at individual grid cells the temporal correlation between the increments applied by SLV and AMS was also consistently very low (<0.25 across the model domain).

The disparity between the analysis increments generated by the AMS and SLV experiments could have several causes. It may indicate a conflict between the assimilated data sets; the treatment of the biases in each is certainly inconsistent. The AMSR-E data have been strongly bias-corrected to remove all systematic differences to the model, including any diurnal or seasonal cycles in the model-observation bias, based on the assumption that the model is bias-free (as is common practise in the off-line land surface assimilation community). Even though the AMSR-E data were heavily adjusted towards the model climatology, there is still a small positive bias in the AMSR-E data, which could indicate a failure of the CDF-matching algorithm used to rescale the observations. In contrast the screen-level observations have not been bias corrected at all (and currently they are not bias corrected in Météo-France’s operational

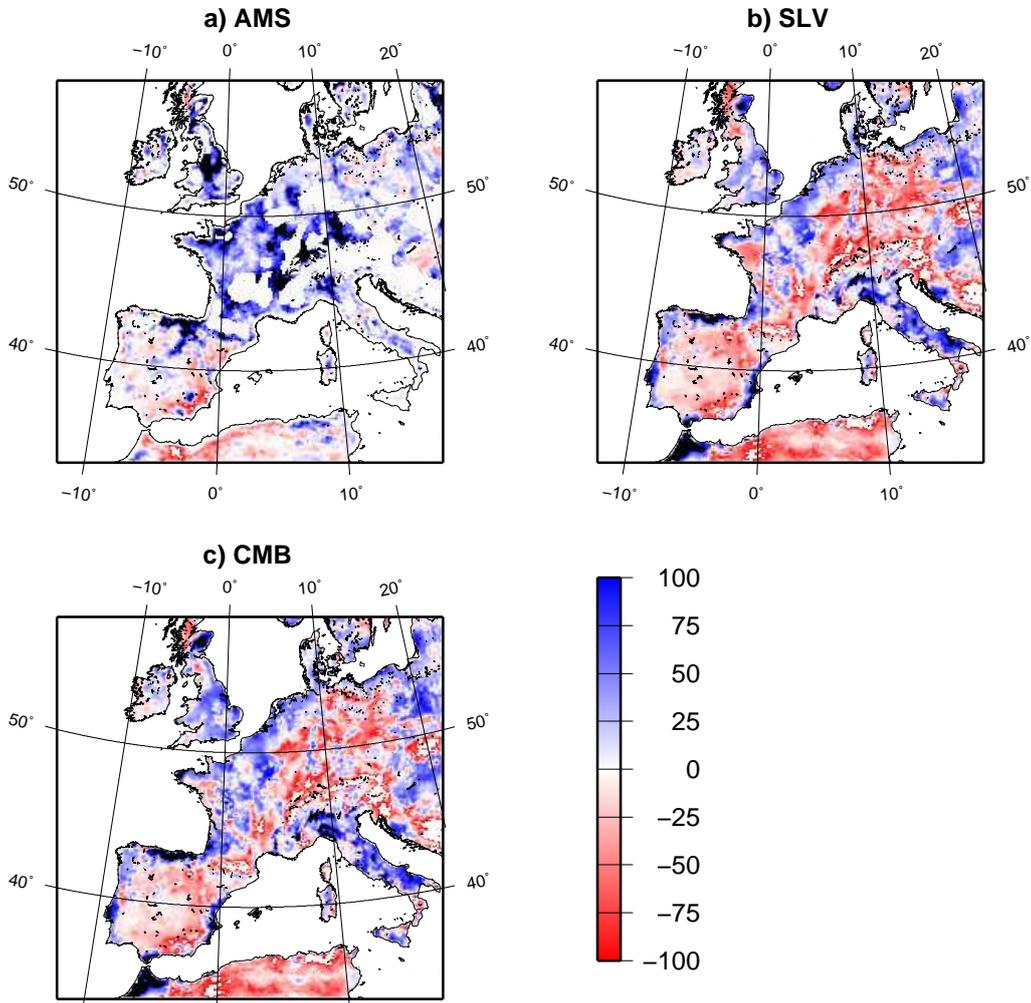


Figure 5.13: Maps of the w_2 analysis increments (mm) accumulated over July 2006 from the a) AMS, b) SLV, and c) CMB assimilation experiments.

land surface scheme). As discussed in Section 5.5.1, there is a strong diurnal cycle in the bias, as well as large biases at some locations, although the monthly averaged model-observation bias for the screen-level variables is close to zero. Alternatively, the disparity between the SLV and AMS experiments may also be caused by systematic errors in the model forecasts of the observed variables, particularly since several studies over heavily instrumented sites have pointed towards this result (Douvillle et al., 2000; Hess, 2001; Drusch and Viterbo, 2007).

Even though the net monthly analysis increments generated by the AMS and SLV assimilation experiments are of the same order of magnitude, the analysis increments for the CMB experiment are very similar to those from SLV. The inclusion of the positively skewed AMSR-E data in CMB slightly reduced the

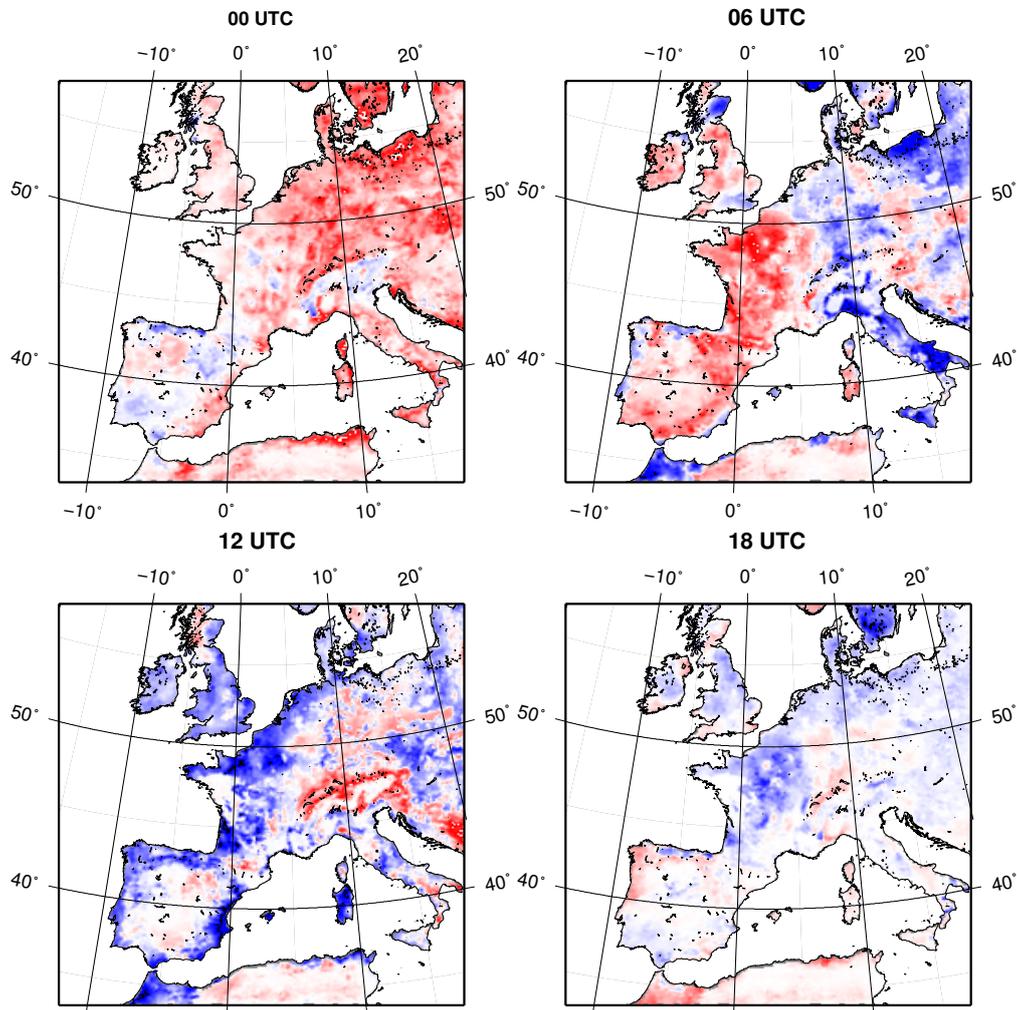


Figure 5.14: Maps of the w_2 analysis increments (mm) accumulated over July 2006 for each of the four daily assimilation cycles for the SLV assimilation. Each plot is labelled with the time of the analysis, and the colour scale is the same as used in Figure 5.13.

symmetry of the SLV distribution, by both decreasing the frequency of the negative mode and increasing the occurrence of positive increments. Compared to SLV, this resulted in a slight increase in the monthly mean increment to $-1.1 \text{ mm month}^{-1}$, as well as a larger spread in the increments (standard deviation: 59 mm month^{-1}). However, the difference between the SLV and CMB analysis increments was slight enough that it is difficult to distinguished between the maps of each in Figure 5.13.

The diurnal cycle in the screen-level increments

The diurnal cycle in the observation increments for the screen-level variables resulted in a diurnal cycle in the analysis increments from the SLV (and CMB) experiments. This is demonstrated by the maps of the net SLV analysis increments for each of the four daily assimilation cycles (at 00:00, 06:00, 12:00, and 18:00 UTC) in Figure 5.14. There was a strong tendency for SLV to remove moisture in the morning, initially in the east of the domain at 00:00 UTC, and then in the west at 06:00 UTC, giving net negative monthly increments of -22 mm and -1.2 mm at 00:00 UTC and 06:00 UTC, respectively. Then as each day progressed, SLV added moisture back into w_2 , giving net positive increments of 17 mm and 5 mm at 12:00 UTC and 18:00 UTC, balancing out the earlier reduction of w_2 . Additionally, the regions where the strongest early morning moisture reductions occurred correspond to regions of strong wetting 6 hours later. For example, in north Poland moisture was removed at 00:00 UTC, and then added back in at 06:00 UTC, with the same pattern occurring six hours later in north France.

5.5.3 Impact on subsequent observation increments

To examine how assimilating each data set impacts the subsequent forecasts of all observed variables, the observation increments for each experiment have been compared. Time series of the mean diurnal cycle and the daily mean of the observation increments are plotted separately in Figures 5.15 and 5.16. Supporting the earlier assertion that the diurnal cycle in the screen-level observation increments was not related to w_2 , the SLV assimilation did not significantly affect this diurnal cycle, and it is difficult to distinguish between the mean diurnal cycle of the observation increments from the SLV and CTR experiments in Figure 5.15. At many individual locations the SLV assimilation reduced the screen-level biases during one phase of the diurnal cycle, while enhancing the biases during the opposite phase. At the daily time scale, the SLV assimilation did generate reductions in the screen-level biases, as seen in Figure 5.16. For the first two thirds of the month the observations were warmer and drier than the model and SLV removed moisture from w_2 , while at the end of the month the screen-level biases were reversed and SLV added moisture to the surface. During both periods SLV generated a clear reduction in the daily mean biases in Figure 5.16, by up to 0.1 K and 1.0%, which is approximately half of the

mean daily CTR bias for both variables.

In contrast to the screen-level variables, the w_1 observation increments were consistently positive (both spatially and temporally), and AMS had a strong tendency to add moisture to w_2 . The AMS assimilation reduced the daily mean w_1 observation increments by close to $0.005 \text{ m}^3\text{m}^{-3}$ on most days (which is again about half of the mean daily bias for CTR). However, there were two periods of mean negative daily observation increments, and AMS was slow to respond to these, resulting in larger negative observation increments than for the CTR simulation.

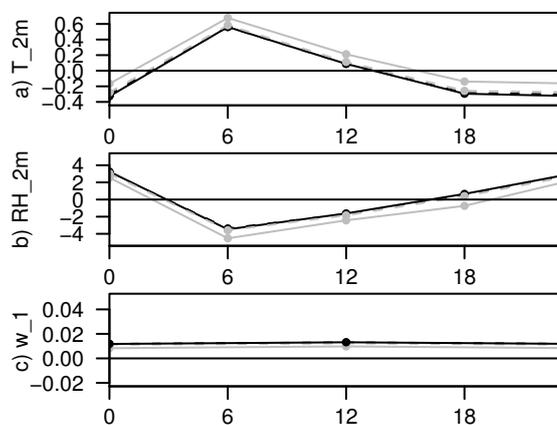


Figure 5.15: Diurnal cycle in the mean monthly observation minus model forecast over July 2006 for a) T_{2m} (K), b) RH_{2m} (%), and c) w_1 (m^3m^{-3}), for the CTR (black, solid), SLV (black, dashed), AMS (grey, solid), and CMB (grey, dashed) experiments. The horizontal axes are in hours.

The cool-moist screen-level model bias for the first two thirds of the month was reduced by SLV and increased by AMS, while the warm-dry bias at the end of the month was reduced by both AMS and SLV. Even though the net monthly w_2 analysis increments for AMS were generally smaller than those from SLV, the AMS assimilation had a greater impact on the forecasts of the screen-level variables than SLV, due to the more consistent direction of the AMS increments. For the mean daily time series, AMS moistened and cooled the screen-level forecasts, by a maximum of approximately 0.2 K and 1.5% (double the impact of SLV). For the diurnal cycle in Figure 5.15, the addition of moisture by AMS decreased the warm-dry model biases at the end of the day, while increasing the cool-moist biases at the start of the day. Likewise, the w_1 observation increments were consistently reduced by AMS, while SLV

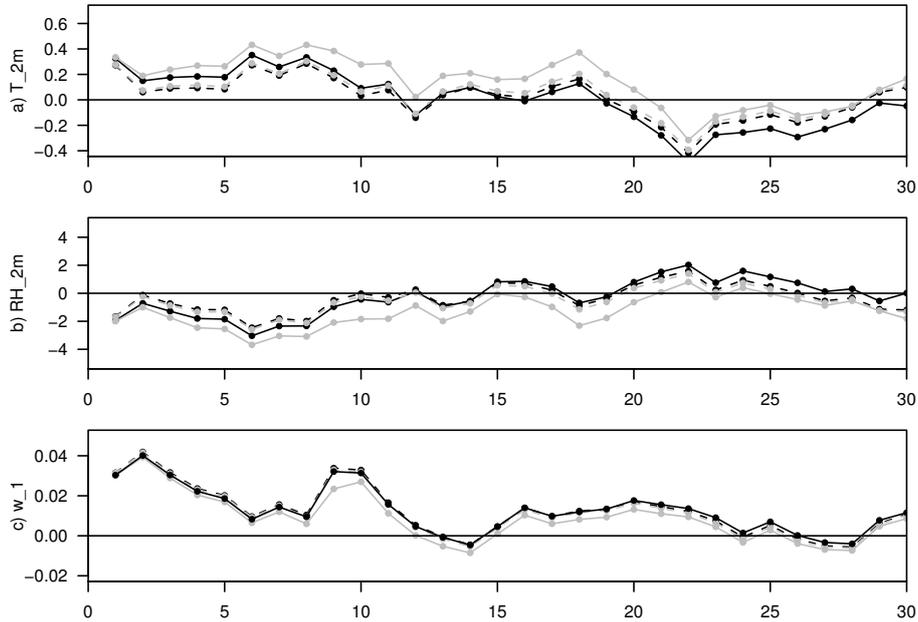


Figure 5.16: Mean daily observation minus model forecast for each day in July 2006, for a) T_{2m} (K), b) RH_{2m} (%), and c) w_1 ($m^3 m^{-3}$), for the CTR (black, solid), SLV (black, dashed), AMS (grey, solid), and CMB (grey, dashed) experiments.

had minimal impact, slightly increasing (decreasing) the biases at the start (end) of the month. For the combined assimilation of AMSR-E and screen-level observations it was again difficult to distinguish the CMB and SLV results in the plot of the diurnal observation increments in Figure 5.15, while the daily mean observation increments in Figure 5.16 are drawn slightly towards the AMS results.

Table 5.5 shows the monthly RMSD between the observations and the model background for each experiment (for the assimilated observations this is the standard deviation of the observation increments). The rather small RMSD for the screen-level observations is partly due to the use of the ALADIN screen-level analysis (rather than the direct use of observations) in these assimilation experiments, since in regions with sparse observations the analysis will strongly resemble the model. For example, about half of the screen-level observations ingested by the screen-level analysis were over France, where the RMSD of the CTR observation increments is 1.62 K and 10.3%, which is somewhat higher than the values in Table 5.5. As an aside, the root mean squares of the observation increments in Table 5.5 are consistent with the observation errors used in

Table 5.5: RMSD between the observations and model forecasts over July 2006 for each assimilation experiment. For the screen-level variables the RMSD of the daily averaged forecasts and observations is shown in brackets.

Time	T_{2m} (K)	RH_{2m} (%)	w_1 (m^3m^{-3})
CTR	1.26 (0.8)	9.41 (6.0)	0.0749
SLV	1.21 (0.7)	9.05 (5.5)	0.0748
AMS	1.25 (0.8)	9.60 (6.2)	0.0732
CMB	1.21 (0.7)	9.05 (5.5)	0.0743

the assimilation, lending support to the selected observation errors: if the observation increments are partitioned equally between the model and observations (i.e. the background and observation errors are assumed to be independent and have equal mean squares), the error standard deviation for each would be 0.9 K for T_{2m} , 7% for RH_{2m} (slightly lower than the RH_{2m} observation error used), and $0.05 m^3m^{-3}$ for w_1 .

The SLV assimilation only slightly reduced the T_{2m} and RH_{2m} RMSD, by about 5% of the original CTR RMSD. This relatively small impact is thought to be a consequence of the diurnal cycle in the observation increments, since this was the greatest contributor to the screen-level RMSD, and it was not amended by the assimilation. Taking the RMSD of the daily average of the forecast and observed screen-level observations (shown in brackets in Table 5.5) slightly increases the relative impact of the assimilation, giving an improvement of close to 8% of CTR. The SLV assimilation had little impact on the w_1 RMSD. In contrast the AMS assimilation slightly reduced the w_1 RMSD, while also very slightly reducing the T_{2m} RMSD and increasing the RH_{2m} RMSD. The disparity between the T_{2m} and RH_{2m} results for AMS is likely due to the sensitivity of humidity to w_2 in the early morning (recall that AMS increased the early morning cool-moist model bias). The CMB assimilation retained the small reductions in the screen-level RMSD generated by SLV, while also slightly reducing the w_1 RMSD.

Despite the modest improvements in the RMSD obtained here, these results are believed to reflect a true reduction in the observation increments, since the statistics were based on a very large sample size (for the screen-level observations, approximately 250,000 observations were used). If this result can be substantiated once the diurnal cycle in the screen-level observations has been addressed (and ideally using a longer experiment period), this implies that as-

simulating remotely sensed soil moisture together with screen-level observations has the potential to help improve the realism of the NWP land surface without degrading the low-level atmospheric forecasts.

Updating the soil temperature

In the current ALADIN surface OI scheme the surface temperature and moisture states are both updated from the screen-level observations. Since the deep-layer soil temperature (T_2) is sensitive to the screen-level observations (specifically T_{2m}) at night (Mahfouf et al., 2009), the persistence of the diurnal biases in these experiments could be explained by the exclusion of T_2 from the update vector. To rule out this possibility, the SLV experiment has been repeated with T_2 included in the update vector (with T_2 error standard deviations of 1K for \mathbf{Q} and 2 K for the initial \mathbf{P}). While the inclusion of T_2 improved the overall observation increment statistics for the screen-level observations, it did not amend the diurnal biases. In fact it further enhanced the warm dry morning departures (to 0.61 K and -3.9% at 06:00 UTC, and 0.15 K and -2.0% at 12:00 UTC), while decreasing the cool moist departures later in the day (-0.22 K and 0.002% at 18:00 UTC, and -0.24 K and 2.9% at 00:00 UTC).

5.5.4 The relative information content

Table 5.6 compares the relative information content of the screen-level observations and AMSR-E data, calculated as described in Section 5.2.4. To isolate the impact on w_2 (rather than w_1), the relative information content was calculated from an additional experiment in which only w_2 was analysed (otherwise identical to CMB). Recall from Section 5.4 that excluding w_1 from the update vector slightly reduces the impact of the EKF assimilation of w_1 data, by removing the long-term evolution of error covariance between w_1 and w_2 . The experiments with w_1 excluded have then underestimated the impact of the AMSR-E data. However, it should still provide a reasonable approximation of the sensitivity of the CMB assimilation to each observation type.

The relative information content in Table 5.6 shows that most of the information assimilated by the CMB experiment was derived from the screen-level observations during periods of active evapotranspiration. In total, 96% of the daily information content was derived from T_{2m} and RH_{2m} data during the day (12:00 and 18:00 UTC) together with RH_{2m} in the early morning (6:00 UTC).

Just 2% of the information was derived from the AMSR-E data, explaining the strong similarities between the SLV and CMB assimilation results above. These statistics are aggregated over all available observations, yet the spatial and temporal coverage of the AMSR-E data was less than that of the screen-level observations. On average, the number of AMSR-E observations per day was 24% of coverage of the screen-level observations (assuming that the grided observations are independent data; in reality the number of individual data was much less than this for all observation types). If the information content of each observed variable is normalised by the data coverage, the information content per observation from the AMSR-E data was still an order of magnitude less than that from the screen-level observations. This result differs from Balsamo et al. (2007), who found that the information content of a single C-band brightness temperature observation was greater than that of the screen-level observations, when assimilated into the ISBA model coupled to the Canadian GEM system. This difference is likely due to the very low AMSR-E observation error of 3 K used in that study, which is equivalent to a volumetric soil moisture error of approximately $0.01 \text{ m}^3\text{m}^{-3}$ over bare soil (e.g., Jackson and Schmugge, 1991), while the remaining error statistics were similar to those used here.

It was noted in Section 5.2.3 that the misspecification of the soil parameters in ISBA suppresses the variability of w_2 , so that the mean background model error of 20 mm for w_2 corresponds to a volumetric error of only $0.01 \text{ m}^3\text{m}^{-3}$, which is five-times less than the assumed observation error. To test whether the low information content of the w_1 data is related to the large ratio of the observed and modeled errors, an additional experiment has been conducted in which the observation error standard deviation for AMSR-E was artificially reduced to $0.02 \text{ m}^3\text{m}^{-3}$ (close to the model w_1 error). In this experiment the analysed w_2 was still dominated by the signal from the screen-level variables, and the information content of the AMSR-E data was increased to 11%. The low information content of the w_1 observations is then principally related to the relative sensitivity of each observed variable to w_2 , rather than the particular error covariance matrices that have been used here.

5.5.5 Comparison to SIM water balance

Here, the results of the AMS, SLV, and CMB experiments are compared to the surface water budget simulated by SAFRAN-ISBA-MODCOU (SIM; Habets et

Table 5.6: The relative information content as a fraction of the daily total (for experiment CMB with w_2 updated only), averaged over July 2006 separately for each of the four daily assimilation cycles.

Analysis time	T_{2m}	RH_{2m}	w_1
00:00	0.01	0.10	-
06:00	0.25	0.10	0.01
12:00	0.27	0.24	-
18:00	0.01	0.00	0.01

al., 2008), which is a 3-layer version of ISBA forced with high-quality observations (Quintana-Seguí et al., 2008) over France. The surface water budget (the change in soil moisture over the month), provides an integration of the surface-moisture inputs (precipitation), outputs (evapotranspiration and runoff), and soil moisture increments where an assimilation is performed. Due to the use of high quality forcing data, the surface water budget from SIM is expected to reflect the true surface water budget more accurately than that generated by forcing ISBA with ALADIN forecasts (Mahfouf et al., 2009). Additionally, assimilating observations into the latter should ideally correct for errors in the ALADIN forecasts, bringing the surface water budget closer to that generated by SIM. Consequently, Mahfouf et al. (2009) compared the surface water budget over July 2006 from SIM to that from an open-loop ISBA simulation (identical to the CTR simulation), and that from the assimilation of screen-level observations. They concluded that the SEKF assimilation of screen-level observations had some skill in replenishing the surface water storage in regions where the ALADIN forecast precipitation was biased low, however in other regions the assimilation further enhanced the soil moisture deficits in the open-loop. Here the surface water budgets generated by the EKF assimilation of AMSR-E w_1 , the screen-level observations, and both, are compared to that from SIM to determine whether assimilating the AMSR-E data can improve upon the results of Mahfouf et al. (2009).

Figure 5.17 shows maps and histograms of the surface water budget for July 2006 from SIM, CTR, and each of the main assimilation experiments (SLV, AMS, and CMB). As noted by Mahfouf et al. (2009), the open-loop CTR had a tendency toward excessive drying and insufficient wetting compared to SIM, which was associated with a low bias in the ALADIN precipitation forecasts over

Table 5.7: First, second, and third quartile of the change in total soil moisture storage (mm) from 1 to 31 July 2006, for each experiment.

	Q_1	Q_2	Q_3
SIM	-25	-10	0.2
CTR	-41	-27	-11
SLV	-55	-28	-10
AMS	-30	-6	19
CMB	-48	-23	-5

this period. In particular, the maps in Figure 5.17 show that CTR simulated excessive drying (in spatial extent and magnitude) along the English Channel coast and in central France (Massif Central), with a region of insufficient moistening in between. Additionally, CTR did not moisten the regions along the Atlantic Coast and south of the Alps as indicated by SIM. The exceptions to this tendency for excessive drying were the mountainous regions along the Italian, Swiss and Spanish borders, where CTR substantially increased the surface water storage (by between 50 and 200 mm). The histograms in Figure 5.17 also reflect the tendency for excessive drying in the open-loop, and the CTR distribution is to the left of the SIM distribution, resulting in more negative quartile statistics (-41, -27, and -11 mm, for the first, second, and third quartiles) than for the SIM experiment (-25, -10, and 0.2 mm, respectively) in Table 5.7. The addition of moisture along the mountainous borders also produced a long positive tail that is not present in the SIM distribution.

Consistent with the results for the SEKF assimilation of screen-level variables reported by Mahfouf et al. (2009), the SLV assimilation correctly reduced the region of drying along the English Channel and the south Atlantic Coast in Figure 5.17, while also further enhancing and expanding the excessive drying in central France. As a result, the most obvious difference between the SLV and CTR histograms in Figure 5.17 is the introduction of a long negative tail for SLV, while the same overly negative mode is retained. Consequently, in Table 5.7 the first quartile is much more negative for SLV (-55 mm), while the median and third quartile are not greatly changed from the open-loop (-28 and -10 mm, respectively).

In contrast to SLV, assimilating the AMSR-E data added moisture across most of France. This reduced the regions of excessive drying in CTR (particu-

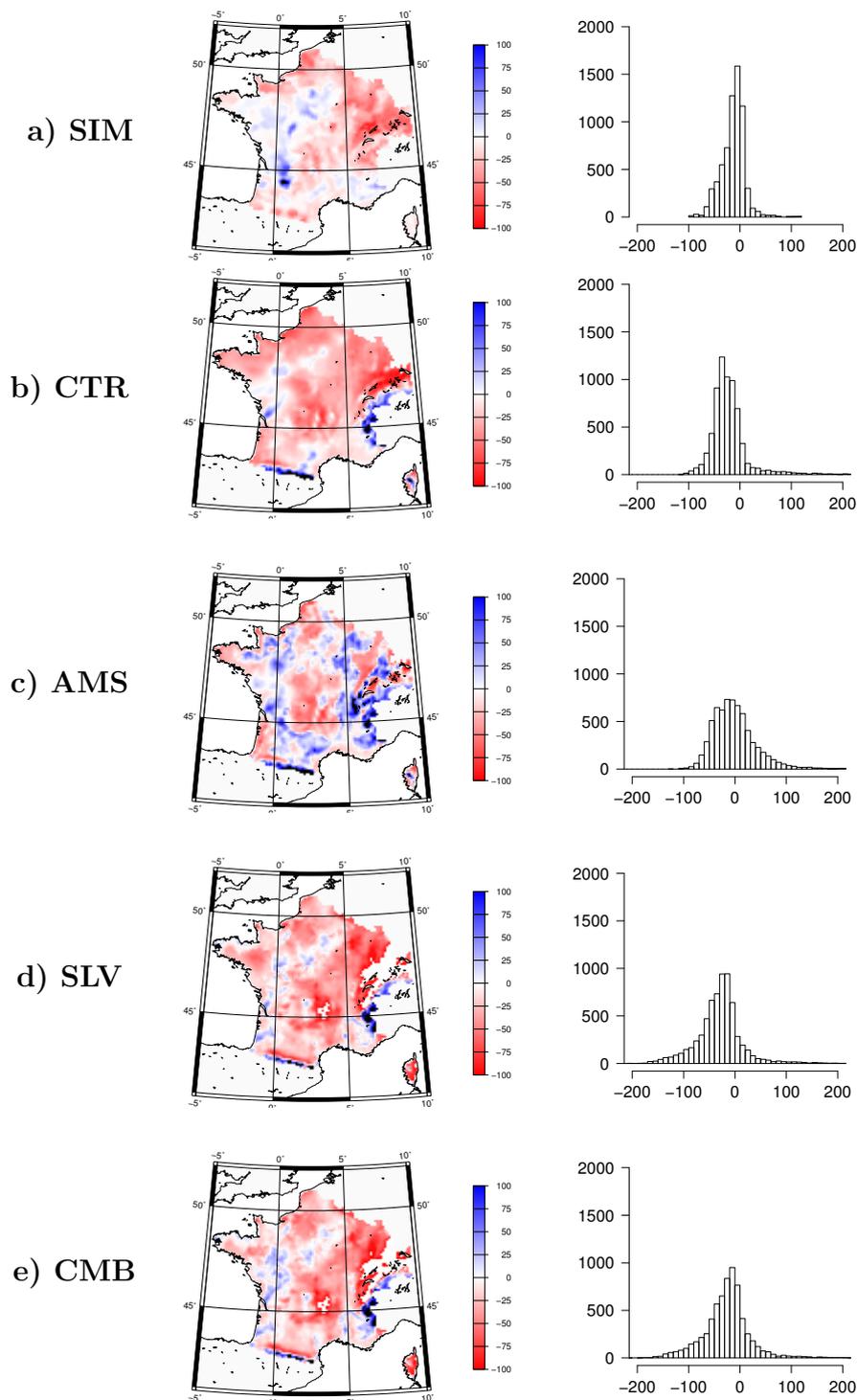


Figure 5.17: Maps (left) and histograms (right) of the change in total soil moisture storage (mm) from 1 to 31 July 2006, from a) SIM, b) CTR, c) AMS, d) SLV, and e) CMB.

larly in the north), and corrected the excessive drying in central France. As a result, AMS shifted the surface water budget distribution to the right (towards SIM), so that the first quartile and median for AMS are reasonably close to the SIM values in Table 5.7 (-30 mm and -6 mm, respectively). However, in many regions AMS has added too much moisture, resulting in an excessive positive tail, with the third quartile (19 mm) much larger than that for SIM.

Consistent with the earlier results, the CMB surface water budget is very similar to that from SLV. However, the inclusion of the AMS data increased the positive tail of the distribution, and compared to SLV, the CMB maps in Figure 5.17 show a slight reduction in the regions of excessive moisture loss, particularly in the northeast. This has brought the quartile statistics slightly closer to SIM (giving -48, -23, and -5 mm, respectively).

In summary, none of the assimilation experiments have generated a surface water budget with a clearly superior fit to the SIM water budget. The addition of moisture across the domain by AMS correctly compensated for the low-biased ALADIN precipitation in many regions, resulting in the best match to the surface water budget statistics for SIM. However, AMS also added too much moisture in some regions, and it added moisture in regions where SIM did not indicate that it was required. As a result, maps of the surface water budget for SIM and AMS do not show any qualitative similarity, and it is not clear from this comparison that the AMS assimilation has responded to the same errors that are indicated by the difference between the CTR and SIM simulations. That is, given the failure of AMS to distinguish between those regions where SIM was wetter than CTR and those where it was not, it is quite possible that AMS generated a better fit to SIM by chance, rather than a true ability to correct for errors in the surface moisture.

Finally, compared to CTR the assimilation experiments (in particular AMS) have added a lot of fine scale variability into the surface water budget. Given that the AMSR-E data (nominally at 25 km) has a much coarser resolution than ISBA (approx. 9 km), assimilating the AMSR-E data was not expected to provide useful information at the resolution of ISBA. The surface water budgets from Figure 5.17 have then also been compared after aggregating them up to various resolutions of up to 50 km, however this did not change the comparative performance of the three assimilation experiments.

5.6 Chapter summary

The continental scale assimilation of remotely sensed near-surface soil moisture into the ISBA model has been presented for the first time. The assimilation was performed with an EKF, which was based on the SEKF developed by Mahfouf et al. (2009) to assimilate screen-level observations into ISBA. Extending the SEKF into a full EKF enhanced the impact of the assimilated near-surface soil moisture data on the model root-zone soil moisture, due to the evolution of the long-term error covariance between the model near-surface and root-zone soil moisture errors by the EKF. However, at the broad temporal and spatial scales examined here, the difference between the results of the EKF and SEKF assimilation of near-surface soil moisture was relatively minor. In fact the method used to rescale the AMSR-E data (to reduce the systematic difference between the observed and modeled soil moisture prior to the assimilation) had a much more profound impact on the root-zone soil moisture analyses than the choice of using the SEKF or the EKF. Despite these results, the EKF will be used in the assimilation experiments throughout the remainder of this thesis, since it has the potential to provide a more realistic specification of the background model errors, while incurring no additional computational cost.

Examining the soil moisture Jacobians for the ISBA model revealed that its two-layer force restore physics are not well suited to assimilating near-surface soil moisture. The near-surface soil moisture (w_1) in ISBA has little influence on the root-zone soil moisture (w_2), and so to constrain w_2 using w_1 observations the sensitivity of the model w_1 forecasts to the initial w_2 must be utilised (in the same way that assimilating screen-level observations relies on model forecasts of the observation equivalent). Further to this, w_1 and w_2 are only weakly coupled in ISBA, and the strongest influence of w_2 on w_1 is an indirect link, via transpiration. Consequently, w_1 observations provide a much weaker constraint of the model w_2 , than the screen-level variables do.

A series of experiments has been conducted to compare the assimilation of AMSR-E derived near-surface soil moisture to the assimilation of screen-level observations. While the experiments were for too short a period for robust validation (due to the limited availability of forcing fields), the focus was on contrasting the impact of assimilating each data type. When the AMSR-E w_1 and screen-level observations were assimilated separately, there was no clear consistency between the resultant root-zone soil moisture, indicating that for

these experiments the screen-level observations could not have been substituted with AMSR-E data to achieve similar corrections to the low-level atmospheric forecasts. However, when both data types were assimilated together, the EKF was able to slightly improve the fit between the model and both observation types, although the improvements were extremely modest, being just 1% of the CTR RMSD for w_1 , and close to 5% of the CTR RMSD for both screen-level variables.

This work also revealed a number of problems with the use of screen-level observations for analysing the root-zone soil moisture in ISBA. Most notably, the screen-level observation increments were dominated by a diurnal cycle, which was not related to the model soil moisture. The resulting diurnal cycle in the analysis increments demonstrates how assimilating screen-level observations can lead to unrealistic soil moisture updates, reinforcing the need to assimilate alternative data sets. Consequently, in the next chapter the AMSR-E near-surface soil moisture will be assimilated into the Australian NWP land surface model (MOSES) over the Australian continent for a one year period, and the resulting root-zone soil moisture estimates will be validated against in situ observations. Since MOSES is a multi-layer model with explicit parameterisation of infiltration through each soil layer, it is expected to provide a more direct relationship between the near-surface and deeper soil moisture than exists in ISBA.

Chapter 6

Assimilation Experiments with MOSES

6.1 Overview of chapter

In this Chapter remotely sensed near-surface soil moisture observations are assimilated into an off-line version of the land surface model from the Australian BoM's NWP model (MOSES¹), to test whether this can improve the modeled root-zone soil moisture. As with the ISBA assimilation experiments in Chapter 5, the EKF described in Section 3.2 is used. This is the first study to assimilate near-surface soil moisture observations into MOSES, and so the physical mechanisms by which MOSES relates the near-surface and root-zone soil moisture, and the strength with which the EKF can constrain the soil moisture throughout the soil profile from near-surface observations, are first examined.

Experiments assimilating the near-surface soil moisture data are then presented, using soil moisture observations retrieved from C-band descending AMSR-E overpass brightness temperatures using the VUA-NASA retrieval algorithm, following the findings of Chapter 4. The AMSR-E data are assimilated over the Australian domain for a one year period from 1 April 2008, and the results of the assimilation are examined in terms of both the net impact on the model soil moisture and evapotranspiration forecasts, and the impact on the model skill in forecasting soil moisture. The latter is assessed at the Murrumbidgee Monitoring Network sites by comparison to in situ soil moisture observations, and then across Australia by examining whether the assimilation has corrected

¹As discussed in Section 3.4.2, the off-line version of MOSES is called JULES, however it will be referred to as MOSES throughout this chapter for simplicity.

for errors in the precipitation forecasts used to force MOSES.

6.2 Data and methods

This section reviews the different data sets used in this chapter, including the data used in the assimilation, to evaluate the results of the assimilation, and to categorise the results into dominant climatic zones. Following this, the EKF assimilation strategy is outlined.

6.2.1 AMSR-E data

In response to the findings of Chapter 4, near-surface soil moisture observations derived from C-band descending AMSR-E brightness temperatures with the VUA-NASA retrieval algorithm have been used in the assimilation experiments in this chapter. This is the same data set that was assimilated into ISBA in Chapter 5, and AMSR-E data has been processed in this chapter using the same methods that were applied for ISBA (see Section 5.2.1 for details). In summary, the AMSR-E near-surface soil moisture observations have been screened to remove contamination from RFI based on the index of Li et al. (2004), and to remove contamination from dense vegetation based on a mean monthly optical depth threshold of 0.8. Following this screening, the level one AMSR-E swath-data have been directly mapped onto the 0.375° ACCESS model grid using a nearest neighbour approach. Also, a final quality control has been applied immediately prior to the assimilation, by excluding all data for which the observation increment is more than twice the standard deviation of the observation increments over the experiment.

A map of the temporal coverage of the assimilated AMSR-E observations is shown in Figure 6.1. As noted in Chapter 4, RFI, dense vegetation, and frozen soil conditions are rare in Australia, giving an unusually high coverage of usable passive microwave observations. Consequently, just 2% of the model grid cells in Figure 6.1 have no data. On average the AMSR-E descending overpass soil moisture data were available for 58% of days in the experimental period, with the coverage ranging from close to 65% in the north (10°S) to 85% over Tasmania (41°S). The coverage is reduced to less than 50% (and as low as 10%) over a small vegetated region in southeast Australia, which is roughly collocated with the Australian Alpine mountain ranges. The coverage is also low (down to 10%)

in a large area of central arid Australia, which is associated with convergence problems in the VUA-NASA retrieval algorithm in dry environments (Crow et al., 2009).

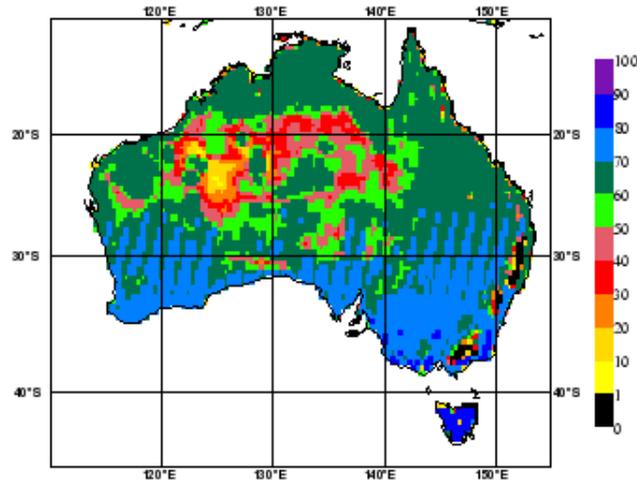


Figure 6.1: Coverage (% of days) of the AMSR-E descending overpass soil moisture observations over Australia from April 2008 - March 2009.

The near-surface soil layer in MOSES is 10 cm thick, much thicker than the (approximate) 1 cm layer to which C-band microwave observations are sensitive. In Chapter 4 it was demonstrated that temporally filtering the AMSR-E soil moisture data can reduce the systematic differences between soil moisture estimates from AMSR-E and from modeled or observed soil moisture that relates to a thicker surface layer. Consequently, the AMSR-E soil moisture time series for each model grid cell has been filtered using the exponential moving average filter described in Section 4.3.4 before being assimilated into MOSES.

Finally, the (filtered) AMSR-E soil moisture observations have been rescaled to better match the soil moisture climatology from MOSES, using the same CDF-matching method that was applied in Chapter 5 for ISBA. The results of this bias correction are presented in Section 6.5.1.

6.2.2 The Murrumbidgee Monitoring Network

In Chapter 4, near-surface soil moisture from AMSR-E was evaluated by comparison to in situ soil moisture observations from the Murrumbidgee Monitoring Network. The Murrumbidgee observations have also been used in this chapter

to assess the impact of assimilating the AMSR-E data on the MOSES forecast near-surface and root-zone soil moisture. Following the recommendations from Section 4.3.1, the root-zone soil moisture forecasts from MOSES have been evaluated by aggregating the in situ observations to a root-zone soil moisture estimate, rather than using observations from the individual soil layers. For the Australian vegetation classes, the root-zone in MOSES is approximately 1 m thick (see Section 3.4.2). Consequently, the MOSES root-zone soil moisture (S_{RZ}) has been calculated by aggregating the moisture content of the three uppermost soil layers (which cover depths of 0.00 - 0.10, 0.10 - 0.35, and 0.35 - 1.0 m, respectively). Similarly, the observed root-zone soil moisture values have been calculated from the in situ observations at three depths (covering 0.00 - 0.30, 0.30 - 0.60, and 0.60 - 0.90 m).

The Murrumbidgee Monitoring Network was introduced in Section 4.2.2: in particular the monitoring stations are located in Figure 4.1 and described in Tables 4.1 and 4.2. For the experiments conducted in this chapter, in situ observations were available for only seven MOSES grid cells. In the Kyeamba region the available data were spread across two MOSES grid cells, with five stations (K1, K2, K3, K5, & K7) in the grid cell centred at (147.5,-35.375), referred to as Kyeamba A (KA), and three stations (K8, K11, & K14) in the grid cell centred at (146.75,-36.125), referred to as Kyeamba B (KB). Additionally, there were ten stations (Y3-Y13) with available data in the grid cell over Yanco (M8), and another four stations each in separate grid cells at Cooma Airfield (M1), West Wyalong (M4), Balranald (M5), and Griffith (M7). However, reliable time series of root-zone soil moisture were not available at M1 and M5, due to poor coverage of the nonsurface observations (in both cases more than half of the data were missing at some depths). However, the near-surface observations were available at these sites, and have been used. Additionally, while in situ observations were available from M6, they are unrealistic and so have not been used (see Section 4.3.5; the nonsurface observations at M6 had similar unrealistic features).

6.2.3 Precipitation data

Daily precipitation observations from the BoM's rain gauge analysis have been used to determine whether the EKF can correct the MOSES soil moisture in response to errors in the ACCESS model precipitation forecasts that have been

used to force MOSES. The observed precipitation data set used in this Chapter (for 2008-2009) is different from that used in Chapter 4 (for 2003-2008), following an update to the method used to generate spatial climate data sets at the BoM. The updated data sets use the same input rain gauge data as was used by the previous precipitation analyses, however the analysis technique has been refined, giving a substantial accuracy improvement (Jones et al., 2009). The estimated mean absolute error of the updated precipitation analysis is 0.9 mm day^{-1} (Jones et al., 2009). The precipitation analyses are generated daily on an 0.05° grid from approximately 5760 rain gauge observations Jones et al. (2007). These analyses have been aggregated to the 0.375° MOSES grid for comparison to the model.

6.2.4 Köppen -Geiger climate classification

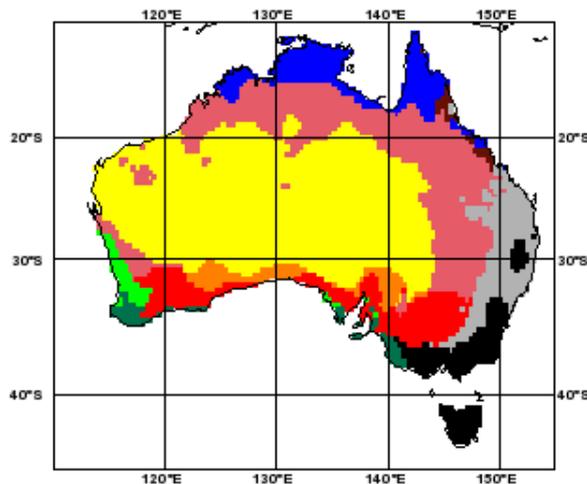
The Australian continent covers a diverse range of climate regimes and the EKF has been examined separately for each of these regimes, to prevent the characteristics of the arid zone (which covers approximately 70% of the continent) from dominating the results. Note that this was not necessary for ISBA in Chapter 5, since the climate across the European domain is more uniform. The Australian climate has been classified according to the Köppen-Geiger classification, as derived by Peel et al. (2007). The map of the Köppen-Geiger classifications in Figure 6.2 shows that Australia can be divided into three main climate regimes:

- Tropical in the north (exclusively Aw; see Figure 6.2 for an explanation of the Köppen-Geiger codes)
- Arid in the centre and west (including Bwh, Bwk, Bsh, and Bsk)
- Temperate along the east coast (Cfa, and Cfb) and westward facing coasts in the south (Csa and Csb).

The results of the EKF are presented below separately for each of the above classifications, with the temperate zone separated into oceanic (Cfa and Cfb) and Mediterranean (Csa and Csb) zones, since these are geographically distinct.

6.2.5 The assimilation experiments

The AMSR-E data have been assimilated into MOSES using the EKF described in Section 3.2, and also used in Chapter 5. The soil moisture variable in MOSES

**Tropical**

Blue Aw - Tropical savannah

Temperate – Mediterranean

Light green Csa - Warm Mediterranean

Dark green Csb - Temperate Mediterranean

Arid

Yellow Bwh - Warm desert

Orange Bwk - Cool desert

Pink Bsh - Warm semiarid

Red Bsk - Cool semiarid

Temperate – Oceanic

Gray Cfa - Warm oceanic

Black Cfb - Temperate oceanic

Figure 6.2: Köppen-Geiger climate zones over Australia (from Peel et al., 2007).

upon which the physics operate is the fraction of saturation, \mathbf{S} (e.g, see equation 3.12). Consequently the EKF has been performed in \mathbf{S} space, and for clarity \mathbf{S} has been reported here as a percentage, while the volumetric soil moisture, θ_u , has been exclusively reported as m^3m^{-3} . As discussed in Section 3.4.2, \mathbf{S} includes four soil moisture layers, and for the vegetation types in Australia the upper three layers approximate the root zone soil moisture. The fourth layer is a slowly varying moisture storage, representing inter-annual variability. Since the fourth layer will be very weakly coupled to the near-surface soil layer, it has been excluded from the state update vector. Additionally, in Section 6.3.4 it will be demonstrated that S_3 can also be excluded without significantly affecting the soil moisture analyses, hence only S_1 and S_2 have been updated by the EKF. The (exponentially filtered) AMSR-E observations were assumed to be observation-equivalent to the (10 cm deep) near-surface soil moisture layer in

MOSES. Additionally, a 6 hour assimilation cycle has been used for consistency with the eventual assimilation of screen-level observations, and the AMSR-E data (observed at roughly 12:00 UTC) have been assimilated at the end of the assimilation window (at 15:00 UTC).

The error covariance matrices required by the assimilation have been specified using a similar approach as in Chapter 5 for ISBA. Spatially uniform values have been used for all specified errors, and the off-diagonal background and model error covariances were assumed to be zero. The AMSR-E observation error and the initial background model errors were based on the evaluation of AMSR-E and MOSES soil moisture in Chapter 4. The additive model forecast error was based on the approximate NWP surface water budget error of 10 mm day⁻¹ quoted by Douville et al. (2000) (as was done for ISBA). Full details of the error selection, and an evaluation of the resulting error covariance matrices at the Murrumbidgee Monitoring Network sites, will be presented in Section 6.3.3.

6.3 Implementing the EKF in MOSES

Before the assimilation results are presented in Sections 6.4 and 6.5, the preliminary testing and refining of the assimilation strategy is presented here. This includes testing the accuracy of the model Jacobians used to estimate the linearised forecast model, and examining the manner in which the near-surface soil moisture is related to the underlying soil moisture profile in MOSES. Based on these results, the state update vector for the EKF is refined to include only S_1 and S_2 , and the impact of this is tested. Finally, the model and observation error covariances used in the assimilation are presented, and checked against in situ soil moisture observations.

6.3.1 MOSES model Jacobians

The magnitude of the perturbations used to estimate the model Jacobians for the EKF has been chosen by testing the similarity of the model Jacobians estimated with positive and negative perturbations for a range of perturbation sizes. Figure 6.3 shows a scatterplot comparing the difference between the two estimates of \mathbf{M}_{11} , for a range of perturbation sizes. As the magnitude of the perturbation decreased, the two estimates converged until the perturbation was

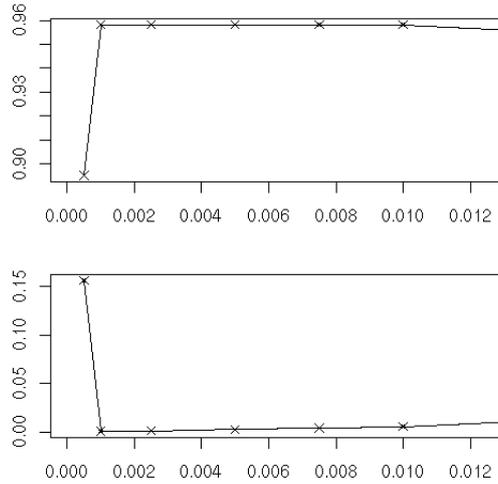


Figure 6.3: Dependence of estimated M_{11} (y axis: %/%) on the perturbation used in the finite difference equation (x axis: fraction of soil moisture at saturation). The upper plot shows the mean of the Jacobians estimated with positive and negative perturbations, and the lower plot shows the absolute difference between the two estimates. All points plotted are the mean across the Australian domain, for the six hour forecast beginning from 09:00 UTC on 7 April 2008.

below 0.1%, beyond which the two estimates diverged. This divergence was due to the signal being dominated by numerical and dynamical noise for overly small perturbations (even though linear tangent theory would suggest that infinitesimal perturbations are best (Balsamo et al., 2004)). Similar findings were obtained for the other elements of \mathcal{M} , and so a perturbation of 0.1% has been used to estimate the Jacobians for the EKF.

Scatterplots of the Jacobian terms estimated with positive and negative perturbations of 0.1% are shown in Figure 6.4. All points are aligned along the one-to-one line, consistent with \mathcal{M} being well approximated by \mathbf{M} within the range of the applied perturbation. Only the diagonal terms of the Jacobian matrix are included in Figure 6.4 for brevity, however the off-diagonal terms estimated from the positive and negative perturbations also agreed strongly. For example Table 6.1 shows excellent agreement between the statistics describing the Jacobians estimated with the positive and negative perturbations, for both the diagonal and off-diagonal terms. In most cases the statistics for each estimate were the same (to two decimal places), even for the extreme values, which can represent instances of strong nonlinearity.

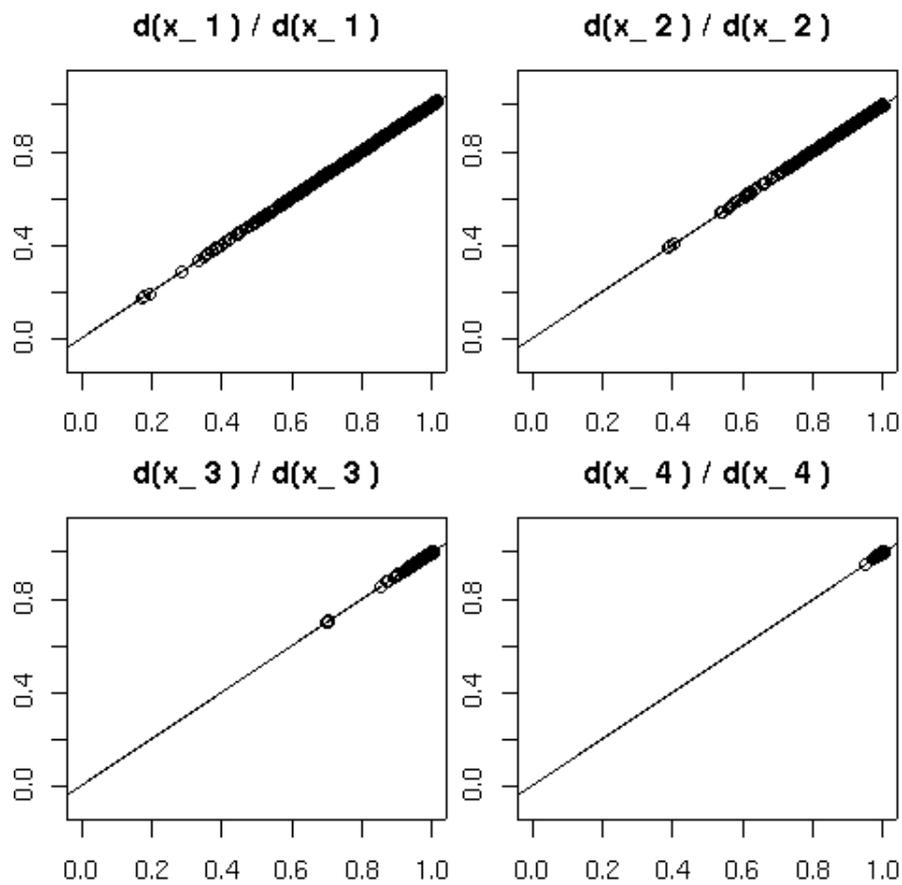


Figure 6.4: Scatterplots of the diagonal Jacobian terms (%/%), for the 6 hour forecast from 09:00 UTC on 6 April 2008, estimated using a perturbation of +0.1% (x-axis) and -0.1% (y-axis) of the soil moisture at saturation.

Table 6.1: Comparison of the statistics for the MOSES Jacobian terms (%/%), for the 6 hour forecast from 09:00 UTC on 6 April 2008, estimated using a perturbation of +0.1% (+ve) and -0.1% (-ve) of the soil moisture at saturation.

		mean	stdev	min	max
$\frac{\partial \mathbf{S}_1}{\partial \mathbf{S}_1}$	+ve	0.92	0.077	0.16	1.0
	-ve	0.92	0.076	0.16	1.0
$\frac{\partial \mathbf{S}_2}{\partial \mathbf{S}_1}$	+ve	0.0072	0.028	-0.019	0.27
	-ve	0.0071	0.027	-0.018	0.27
$\frac{\partial \mathbf{S}_3}{\partial \mathbf{S}_1}$	+ve	-0.00004	0.0017	-0.0040	0.053
	-ve	-0.00005	0.0016	-0.0040	0.053
$\frac{\partial \mathbf{S}_4}{\partial \mathbf{S}_1}$	+ve	-0.00008	0.00028	-0.0010	0.0020
	-ve	-0.00007	0.00027	-0.0010	0.0010
$\frac{\partial \mathbf{S}_1}{\partial \mathbf{S}_2}$	+ve	0.022	0.071	-0.034	0.49
	-ve	0.022	0.070	-0.034	0.49
$\frac{\partial \mathbf{S}_2}{\partial \mathbf{S}_2}$	+ve	0.97	0.052	0.43	1.0
	-ve	0.97	0.051	0.43	1.0
$\frac{\partial \mathbf{S}_3}{\partial \mathbf{S}_2}$	+ve	0.0042	0.0098	-0.0010	0.14
	-ve	0.0041	0.0097	-0.0010	0.14
$\frac{\partial \mathbf{S}_4}{\partial \mathbf{S}_2}$	+ve	0.00004	0.00028	-0.0010	0.0060
	-ve	0.00004	0.00029	-0.0010	0.0050
$\frac{\partial \mathbf{S}_1}{\partial \mathbf{S}_3}$	+ve	-0.0028	0.011	-0.039	0.14
	-ve	-0.0029	0.010	-0.039	0.14
$\frac{\partial \mathbf{S}_2}{\partial \mathbf{S}_3}$	+ve	0.0090	0.016	-0.0010	0.20
	-ve	0.0090	0.016	0	0.20
$\frac{\partial \mathbf{S}_3}{\partial \mathbf{S}_3}$	+ve	0.99	0.013	0.72	1.00
	-ve	0.99	0.013	0.73	1.00
$\frac{\partial \mathbf{S}_4}{\partial \mathbf{S}_3}$	+ve	0.00069	0.0023	-0.0010	0.055
	-ve	0.00069	0.0022	-0.0010	0.054
$\frac{\partial \mathbf{S}_1}{\partial \mathbf{S}_4}$	+ve	-0.0012	0.0038	-0.014	0.0019
	-ve	-0.0013	0.0036	-0.014	0.0019
$\frac{\partial \mathbf{S}_2}{\partial \mathbf{S}_4}$	+ve	0.00036	0.00065	-0.0010	0.0050
	-ve	0.00037	0.00065	-0.0010	0.0050
$\frac{\partial \mathbf{S}_3}{\partial \mathbf{S}_4}$	+ve	0.0014	0.0026	0.0	0.029
	-ve	0.0014	0.0026	0.0	0.029
$\frac{\partial \mathbf{S}_4}{\partial \mathbf{S}_4}$	+ve	1.00	0.0022	0.95	1.0
	-ve	1.00	0.0023	0.95	1.0

6.3.2 The model Jacobians

The soil moisture dynamics in the MOSES model are very different to those from the two-layer force-restore ISBA model (see Section 5.3.2). MOSES has finer vertical resolution, and the soil moisture in each layer directly influences the moisture in adjacent layers through diffusive flow (equation 3.11). Figure 6.5 shows maps of the first row ($\partial S_1/\partial \mathbf{S}$) of the MOSES model Jacobian for a typical winter period, in this case for the 6 hour forecast beginning at 09:00 UTC on 1 August 2008 (note that this is equivalent to the linearised observation operator for the assimilation of descending overpass AMSR-E S_1 observations on that day). These maps show that the 6 hour forecasts of near-surface moisture had very little sensitivity to the moisture in the third and fourth layers, and $\partial S_3/\partial S_1$ and $\partial S_4/\partial S_1$ were close to zero almost everywhere. Additionally, across most of Australia an anomaly in S_1 largely persisted over the 6 hours, giving $\partial S_1/\partial S_1$ close to one and $\partial S_2/\partial S_1$ close to zero (less than 0.01 %/%) . It is only along the

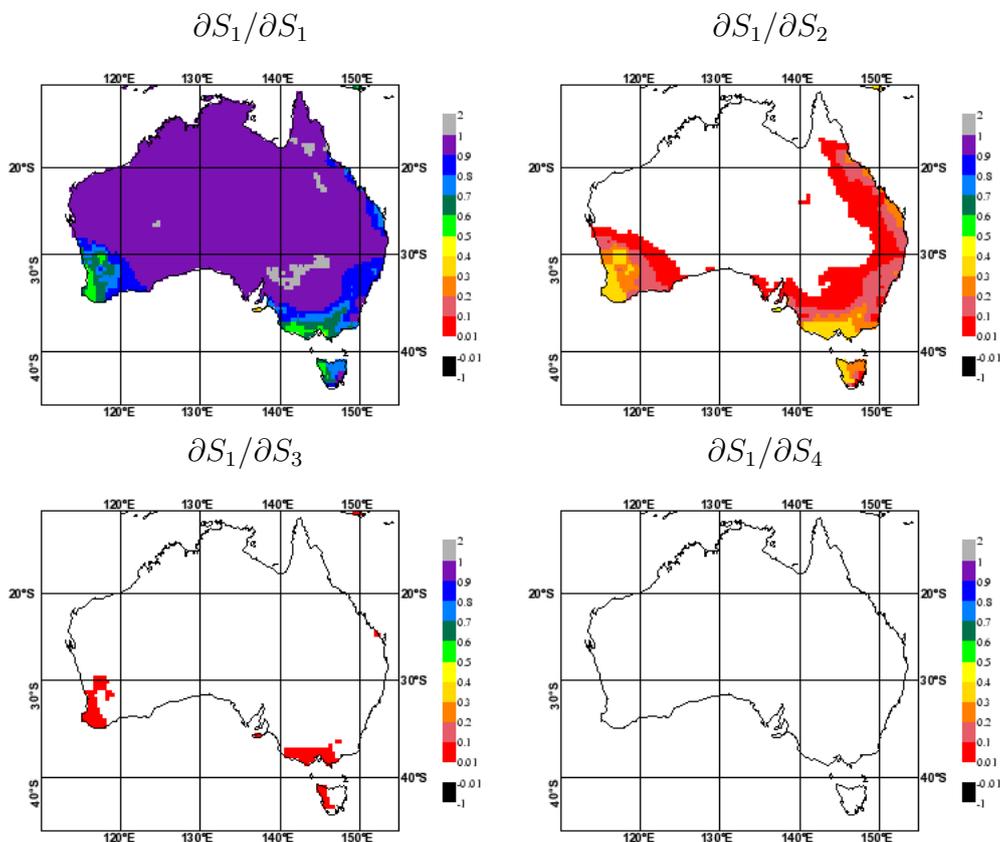


Figure 6.5: Maps of the linearised observation operator (%/%) for the 6 hour assimilation cycle from 09:00 UTC on 1 August 2008.

southwest and east coasts that the perturbations applied to S_1 influenced the other layers: in these regions $\partial S_1/\partial S_1$ was reduced to as low as 0.5 %/%, and $\partial S_2/\partial S_1$ was increased to as high as 0.4 %/%, (although it was more typically less than 0.2 %/%).

Figure 6.6 shows maps of the soil moisture in MOSES at 09:00 UTC on 1 August 2008. Comparison to Figure 6.5 highlights that those regions mentioned above as having lower $\partial S_1/\partial S_1$ (less persistence) and higher $\partial S_1/\partial S_2$ (more rapid flow of soil moisture information) correspond to wetter soils. This relationship is caused by the dependency of hydraulic conductivity on soil moisture in equation 3.13. This relationship is also demonstrated by the scatterplot comparing the model Jacobians to S_1 for the KA model grid cell in Figure 6.7. When S_1 was below about 40% of saturation at KA, $\partial S_1/\partial S_2$ was close to zero, while for S_1 above 40% $\partial S_1/\partial S_2$ increased rapidly (with inverse behaviour from $\partial S_1/\partial S_1$). Spatially, the relationship between soil moisture and the Jacobians in

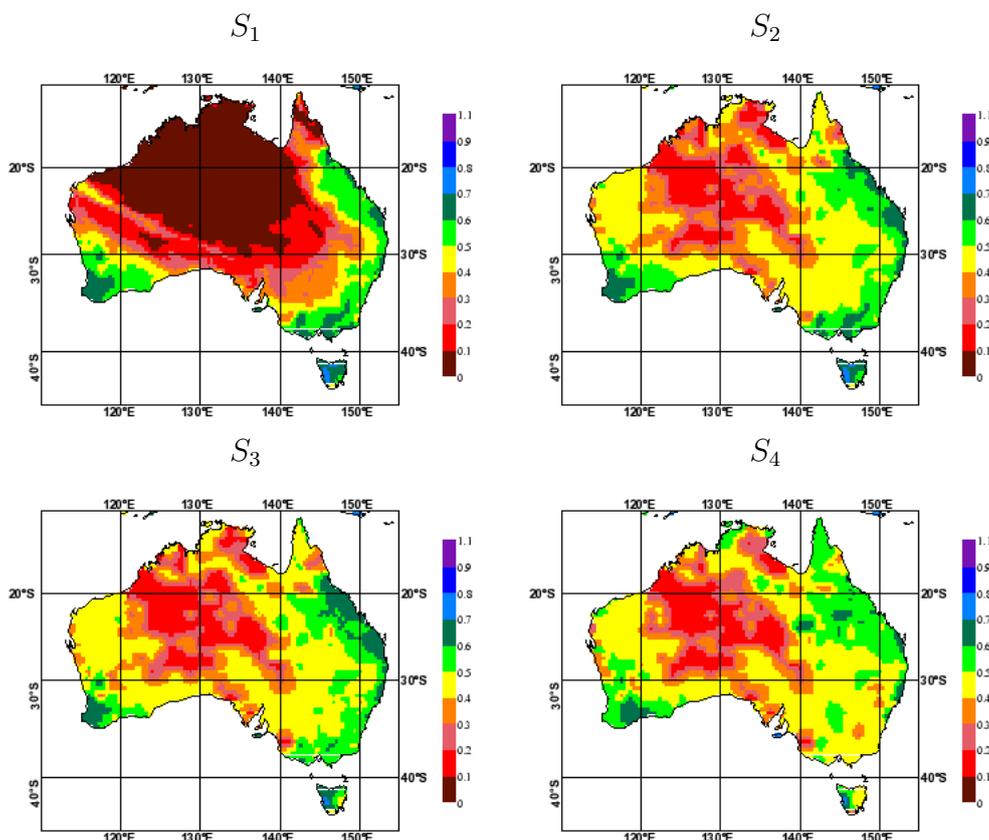


Figure 6.6: Maps of the MOSES soil moisture (as a fraction of the soil moisture at saturation) in each layer at 09:00 UTC on 1 August 2008.

Table 6.2: MOSES model Jacobian terms in units of %/%, spatially averaged across Australia, for the 6 hour forecast from 09:00 UTC on 1 August 2008. Values below 1×10^{-5} are reported as 0.0.

$$M_{9 \rightarrow 15UTC} = \begin{pmatrix} 0.96 & 2.5 \times 10^{-2} & 9.4 \times 10^{-4} & 3.5 \times 10^{-5} \\ 8.1 \times 10^{-3} & 0.98 & 8.1 \times 10^{-3} & 3.9 \times 10^{-5} \\ 1.7 \times 10^{-4} & 4.1 \times 10^{-3} & 0.99 & 1.2 \times 10^{-3} \\ 0.0 & 1.3 \times 10^{-5} & 7.0 \times 10^{-4} & 1.0 \end{pmatrix}$$

Table 6.3: MOSES model Jacobian terms in units of mm mm^{-1} , spatially averaged across Australia, for the 6 hour forecast from 09:00 UTC on 1 August 2008. Values below 1×10^{-5} are reported as 0.0.

$$M_{9 \rightarrow 15UTC} = \begin{pmatrix} 0.96 & 1.1 \times 10^{-2} & 1.4 \times 10^{-4} & 0.0 \\ 2.0 \times 10^{-2} & 0.98 & 3.1 \times 10^{-3} & 0.0 \\ 1.2 \times 10^{-3} & 1.1 \times 10^{-2} & 0.99 & 3.9 \times 10^{-4} \\ 1.3 \times 10^{-5} & 1.1 \times 10^{-4} & 2.2 \times 10^{-3} & 1.0 \end{pmatrix}$$

Figure 6.5 is similar, in that $\partial S_1 / \partial S_2$ was close to zero across most of Australia, and increased sharply in wetter areas.

For brevity only the first row of the MOSES soil moisture Jacobians is plotted in Figure 6.5. However the nonsurface soil moisture layers showed similar behaviour. Table 6.2 lists the spatial average of the model Jacobians for each soil layer in MOSES. In each case, the applied perturbation mostly persisted in the layer to which it was applied (giving spatial mean $\partial S_i / \partial S_i$ typically between 0.95 and 1.0 %/%), with a small flow of moisture into adjacent layers ($\partial S_i / \partial S_{i \pm 1}$ typically $O(10^{-2})$ to $O(10^{-3}$ %/%), and much smaller responses in the remote layers (10^{-4} %/%, or smaller). As an aside, Table 6.3 shows the same model Jacobians converted to units of mm mm^{-1} for consistency with Darcy's diffusive flow equation (equation 3.12). This table highlights that the soil moisture in each model layer has a similar impact on the moisture in the layers above and below (since S_i similarly influences the flow out of (W_i) and into (W_{i-1}) the i th layer).

Since Darcian flow does not explicitly depend on surface forcing, the MOSES soil moisture Jacobians do not show a diurnal cycle as occurred for ISBA. However, the dependency on the soil moisture state itself leads to a seasonal cycle in the model Jacobians. As discussed above, Figures 6.5 and 6.6 provide a typ-

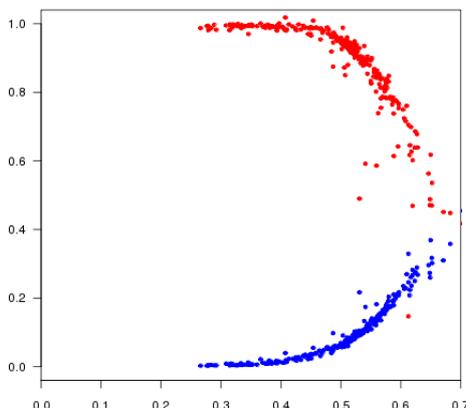


Figure 6.7: Scatterplot of the $\partial S_1/\partial S_1$ (red, %/%) and $\partial S_1/\partial S_2$ (blue, %/%) verse S_1 (%) at Kyeamba A, for the twelve month period from April 2008.

ical example of winter conditions, during which the temperate southeast and southwest experience their wet season. Conversely, Figures 6.8 and 6.9 show the corresponding soil moisture and model Jacobians for summer conditions, during which the tropical north experiences its wet season. As in the winter example, the regions of increased soil moisture correspond to increased $\partial S_1/\partial S_2$ and decreased $\partial S_1/\partial S_1$, while S_1 has little sensitivity to S_3 and S_4 .

In summary, there is limited vertical exchange of soil moisture information through the MOSES soil moisture profile. This limited exchange is expected in arid and semiarid conditions (which occur across most of Australia) due to the dependence of hydraulic conductivity on soil moisture. However, it is possible that it is being exaggerated by the MOSES model, due to the model having insufficient vertical soil moisture flow. As a consequence, the soil moisture in each layer of MOSES will only have a significant influence on the 6 hour forecasts of soil moisture in other layers in humid conditions, specifically during the winter in the temperate south and during the wet season in the tropical north of Australia. Even at these times this influence is restricted to the adjacent soil layers.

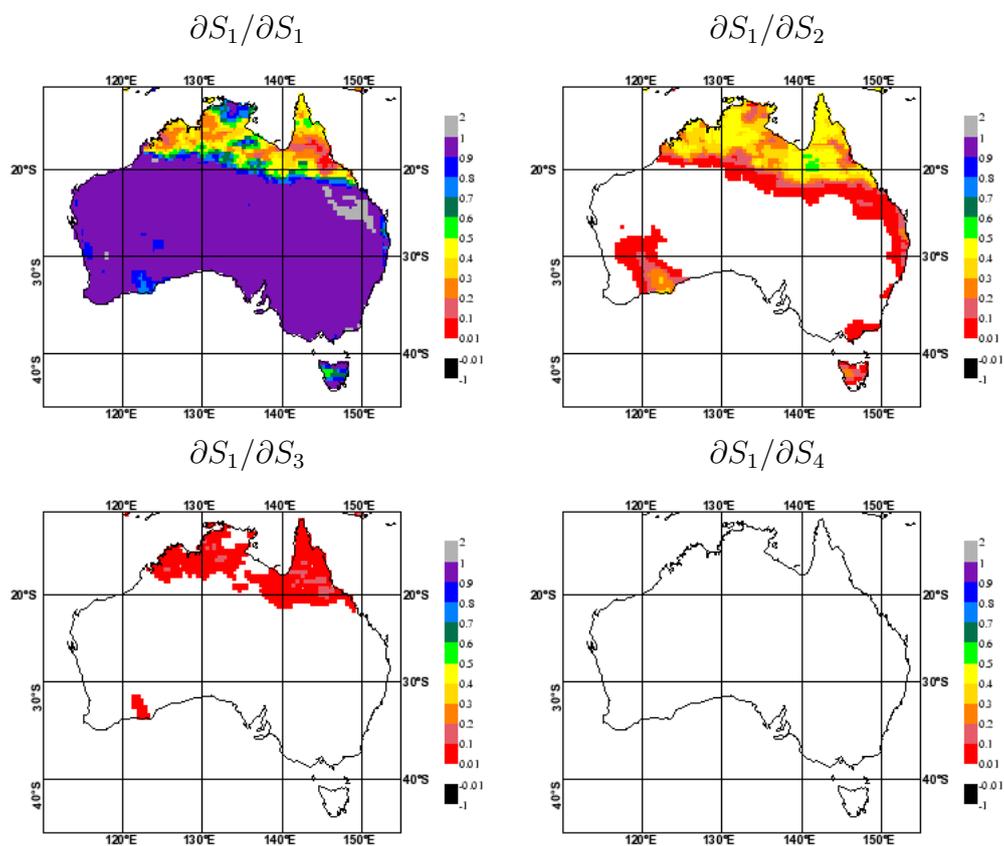


Figure 6.8: Maps of the linearised observation operator (%/%) for the 6 hour assimilation cycle from 09:00 UTC on 1 February 2009.

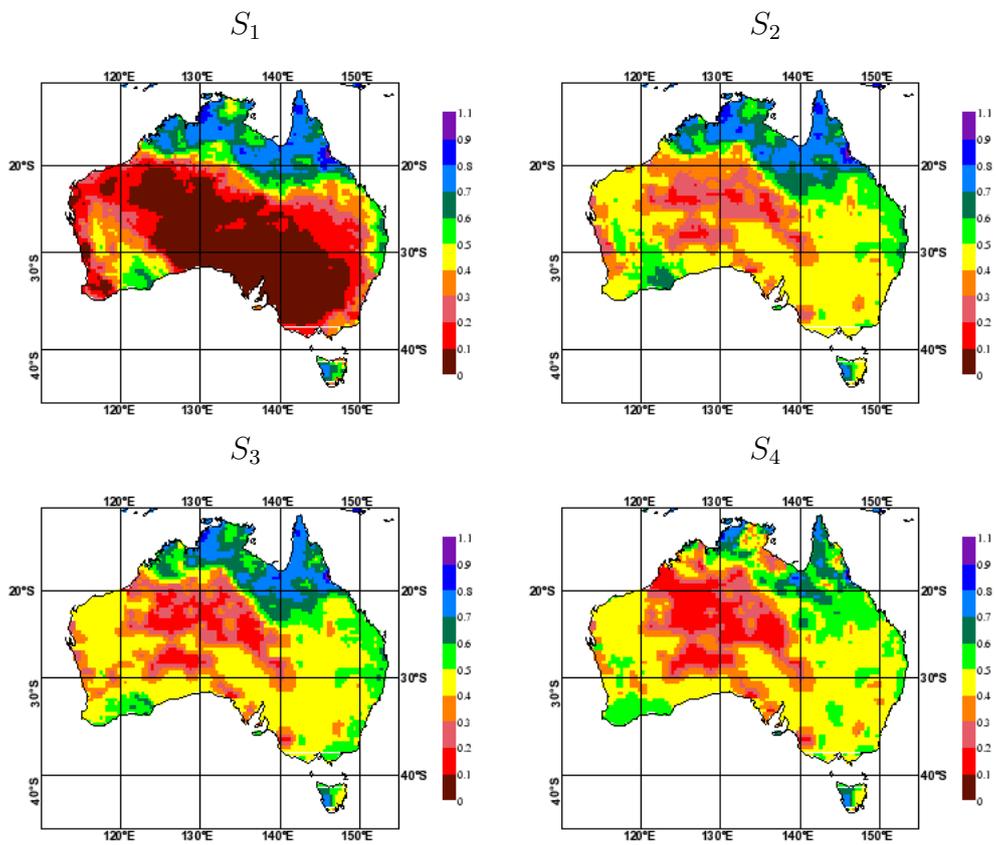


Figure 6.9: Maps of the MOSES soil moisture (as a fraction of the soil moisture at saturation) in each layer at 09:00 UTC on 1 February 2009.

6.3.3 Model and observation error covariances

In this section the EKF error covariance matrix selection is outlined, and then evaluated against in situ soil moisture observations at the Murrumbidgee Monitoring Network sites. For the near-surface soil moisture, the initial observation and background model error variances were based on the comparison to the in situ observations in Chapter 4. Both error variances were set to $(10\%)^2$, equivalent to a root mean square error of $0.04 \text{ m}^3\text{m}^{-3}$ assuming the mean volumetric soil moisture at saturation in MOSES ($0.44 \text{ m}^3\text{m}^{-3}$). This is slightly larger than the mean RMSD of $0.03 \text{ m}^3\text{m}^{-3}$ for AMSR-E and ACCESS near-surface soil moisture in Table 4.10 of Chapter 4. The background model error variances for the nonsurface soil moisture layers were set to the same value.

As in Chapter 5, the model error matrix (\mathbf{Q}) was selected so that the total error applied to the soil moisture profile was close to the NWP surface water budget error of 10 mm day^{-1} quoted by Douville et al. (2000). Additionally, the standard deviation of the error in the near-surface layer was assumed to be double that of the underlying layers to account for its greater sensitivity to forcing errors. For S_1 , a model error variance of $(0.4/d_1 \%)^2$ has been used, while for the underlying layers a variance of $(0.2/d_i \%)^2$ has been used, where d_i is the depth of the i th model soil layer. These values are equivalent to adding an error of $(2 \text{ mm})^2$ every 6 hours to S_1 , and $(1 \text{ mm})^2$ every 6 hours to the deeper layers. While the error covariance matrices are tested in this section using a state update vector that includes the three upper-most soil moisture layers, the AMSR-E assimilation experiments presented below include only S_1 and S_2 in the state update vector (see Section 6.3.4). With $\mathbf{S} = (S_1, S_2)^T$ the above-defined values of \mathbf{Q} result in a total error of 12 mm being added to the soil profile each day, slightly larger than the target value of 10 mm day^{-1} .

The error covariances used for the EKF are summarised in Table 6.4. The realism of the applied model error covariances has been tested by comparing the EKF background model error covariance matrices against model errors estimated from the Murrumbidgee Monitoring Network observations. For this comparison the AMSR-E soil moisture was assimilated for one year, using the EKF to update $\mathbf{S} = (S_1, S_2, S_3)^T$. From this experiment, the ‘‘EKF’’ background error covariances were estimated as the mean of the \mathbf{P}^b used in each analysis. The ‘‘observed’’ background error covariances were estimated as the mean square difference between the model background state and the normalised in situ ob-

Table 6.4: The (spatially uniform) observation, model, and initial background error variances used in the assimilation of AMSR-E near-surface soil moisture, in units of $(\% \text{ of saturation})^{-2}$. All diagonal error covariances were assumed to be zero.

Observation Errors (\mathbf{R})	
S_1	10^2
Forecast Model Errors (\mathbf{Q})	
S_1	4.0^2
S_2	0.8^2
S_3 *	0.3^2
Initial Background Errors (\mathbf{P})	
S_1	10^2
S_2	10^2
S_3	10^2

* For Sections 6.3.3 and 6.3.4 only.

servations at the time of each analysis (based on the assumption that the in situ observations represent the truth). For both estimates, the first week of the assimilation was excluded from the mean, to reduce the impact of errors in the initial \mathbf{x} and \mathbf{P} . Before being compared to the model background soil moisture, the in situ observations were interpolated onto the model soil moisture layers, and then normalised to the model climatology based on the mean and variance (as in equation 4.1). Recall from Section 4.3.1 that the in situ observations are not expected to reflect the area-average soil moisture dynamics for individual sub-surface soil layers well (hence they have been aggregated to a single root-zone observation elsewhere in this chapter). Consequently the results of this comparison have considerable uncertainty.

Tables 6.5 and 6.6 compare the observed and EKF estimates of the diagonal elements of \mathbf{P} (i.e., the variance of the background errors in each soil moisture layer) and \mathbf{R} , respectively. The results suggest that for the surface layer the error variances in \mathbf{P} and \mathbf{R} are both approximately correct, while for the non-surface layers the EKF model error variance in \mathbf{P} has been overestimated. The observed model error variance decreased with the depth of the soil layer, from a mean of $(0.04 \text{ m}^3\text{m}^{-3})^2$ in layer one to a mean of $(0.01 \text{ m}^3\text{m}^{-3})^2$ in layers two and three. The errors were very small in the nonsurface layers due to the limited variability of the MOSES soil moisture in those layers (recall that the in situ data were normalised to match this variance). While the EKF error

variances also decreased with depth, the rate of decrease was not as rapid as for the observed errors. Additionally, the observed errors showed more variation across the Murrumbidgee sites than the EKF errors, which could be due to the (overly simplistic) use of spatially uniform model errors in the EKF.

Table 6.7 compares the observed and EKF estimates of the off-diagonal elements of \mathbf{P} , presented as the correlation between the errors for each pair of model layers (i.e., the error covariances between each pair of model layers, normalised by the error standard deviation for each layer). There was considerable spread between the observed vertical error correlations at the different sites, including some negative values (which are possibly incorrect). Despite this large spread, the mean vertical error correlation across the Murrumbidgee sites was consistently close to 0.4, for each combination of soil layers. In contrast, the vertical error correlations from the EKF \mathbf{P} were several orders of magnitude smaller (since \mathbf{Q} was assumed diagonal and the model Jacobians described above will not generate significant vertical error correlations).

In summary, comparing the observed and EKF error covariance terms es-

Table 6.5: Summary of the root mean square of the background soil moisture errors (m^3m^{-3}) across the Murrumbidgee Monitoring Network sites from April 2008 to March 2009. The errors are estimated from the in situ soil moisture observations (observed), and from the EKF background error covariance matrices (EKF), and are reported relative to the MOSES soil moisture climatology.

	Observed			EKF		
	min	max	mean	min	max	mean
S_1	0.035	0.055	0.045	0.035	0.047	0.043
S_2	0.010	0.018	0.013	0.017	0.032	0.027
S_3	0.006	0.025	0.012	0.016	0.027	0.023

Table 6.6: Summary of the root mean square of the AMSR-E observation errors (m^3m^{-3}) across the Murrumbidgee Monitoring Network sites from April 2008 to March 2009. The errors are estimated from the in situ soil moisture observations (observed), and from the EKF observations error covariance matrices (EKF), and are reported relative to the MOSES soil moisture climatology.

	Observed			EKF		
	min	max	mean	min	max	mean
S_1	0.027	0.062	0.043	0.042	0.046	0.045

Table 6.7: Summary of the correlation between the background model soil moisture errors in layers 1, 2, and 3 across the Murrumbidgee Monitoring Network sites from April 2008 to March 2009. The correlations are estimated from the in situ soil moisture observations (observed) and the EKF background error covariance matrices (EKF).

	Observed			EKF		
	min	max	mean	min	max	mean
$r_{1,2}$	0.037	0.74	0.37	0.0023	0.013	0.005
$r_{1,3}$	0.030	0.63	0.39	0.0005	0.005	0.001
$r_{2,3}$	-0.23	0.74	0.41	0.0065	0.0095	0.008

timated above suggests that the EKF has overestimated the diagonal model background errors in the nonsurface layers, and underestimated the off-diagonal error covariance terms. Within the assimilation these misrepresented errors will to some extent compensate for each other, since the former will lead to an overestimate of the corrections to the deeper layers, while the latter will lead to an underestimate of the corrections.

If the above result regarding the off-diagonal errors is correct, it implies that the model forecast does in fact introduce vertical model error correlations, and \mathbf{Q} should then include nonzero off-diagonal elements. However, due to the considerable uncertainty associated with using in situ data to estimate the soil moisture for individual model layers, as well as the limited spatial coverage of the in situ data, the off-diagonal elements of \mathbf{Q} have been kept at zero. This conservative choice was made since it was considered more detrimental to overestimate the vertical model error correlations (thus overestimating the corrections to the deeper soil layers), than to underestimate them.

6.3.4 The state update vector

The capacity of the EKF to update the nonsurface soil moisture layers from near-surface soil moisture observations depends on the vertical exchange of soil moisture information in the model, since this will determine the observation operator and the evolution of off-diagonal terms in the background error matrix. The above analysis of the MOSES model Jacobians suggests that the EKF is unlikely to make significant updates to S_3 from the S_1 observations, and that it will only make significant updates to S_2 under humid conditions. Consequently,

the computational cost of the analysis could be reduced by excluding S_3 from the state update vector, thus reducing the number of perturbed simulations required to calculate the model Jacobians for \mathbf{H} and \mathbf{M} .

To confirm that the EKF does not generate significant updates to S_3 a three month test was undertaken comparing the assimilation of AMSR-E observations with S_3 included and excluded from the state update vector (starting on 1 April 2008 and using the error covariances from Table 6.4). For these experiments the analysis increments added to each layer decreased rapidly with depth from the surface. For the experiment with S_3 updated, the mean net monthly increment to S_1 ($1.1 \times 10^{-1} \text{ m}^3\text{m}^{-3} \text{ month}^{-1}$) was an order of magnitude greater than the mean increment to S_2 ($1.4 \times 10^{-2} \text{ m}^3\text{m}^{-3} \text{ month}^{-1}$), which was another order of magnitude greater than the mean increment to S_3 ($2.9 \times 10^{-3} \text{ m}^3\text{m}^{-3} \text{ month}^{-1}$). Comparing the resulting soil moisture analysis to the experiment in which S_3 was excluded from the state update vector shows including S_3 had only a small impact on the model state. After three months of assimilation, the difference between the soil moisture from the two assimilation experiments was less than $10^{-3} \text{ m}^3\text{m}^{-3}$ in all soil layers. Conditions in east Australia were reasonably wet ($S_{RZ} > 0.5$) during the three months of this experiment, yet the difference between the two experiments was not significantly greater in these wetter regions (recall that the exchange of moisture information between the different model layers increases with soil wetness). Finally, the results from Section 6.3.3 suggest that the limited analysis updates added to S_3 were not due to the choice of model error covariances used here (within the constraint of zero off-diagonal \mathbf{Q}). Section 6.3.3 suggested that the nonsurface model error variances used by the EKF were too high, which if anything would cause the EKF to overestimate the updates to the nonsurface soil layers.

Since it was shown above that the impact on the analyses of including S_3 in the update vector is small, the cost of the analyses has been reduced by including only S_1 and S_2 for the remainder of this chapter. The impact of the assimilation on S_3 will then be limited to the infiltration of updates from above.

6.4 Testing the EKF: synthetic experiments

In response to the limited coupling between the near-surface and underlying soil moisture in MOSES, the effectiveness with which the EKF can correct the soil moisture throughout the model profile from near-surface observations is ex-

amined in this section through a series of synthetic experiments. The model simulations used in these synthetic experiments are listed in Table 6.8. The first was an open-loop simulation (OPN_ACCESS), run for twelve months from 1 April 2008. OPN_ACCESS was assumed to represent the true model state, and the synthetic observations were generated by taking the S_1 at 15:00 UTC each day from this simulation. The initial conditions on 1 April 2008 were then perturbed to unrealistic values, and the synthetic observations were assimilated, to test whether the EKF could retrieve the true (OPN_ACCESS) soil moisture profile. Two perturbed assimilation experiments were conducted; in the first (EKF_25) the initial S was set to 25% of saturation in all layers, while in the second (EKF_75) the initial S was set to 75% of saturation. The same error covariance matrices from Table 6.4 were used in these experiments, with the initial \mathbf{P} doubled to account for the much greater error in the initial conditions. For each assimilation experiment, the rate at which the EKF converged toward OPN_ACCESS was benchmarked against the corresponding open-loop simulation, initialised with the same perturbed soil moisture (OPN_25 and OPN_75, respectively).

Note that these synthetic experiments have been designed to examine only whether the vertical coupling in MOSES is sufficient for the EKF to correct the soil profile based on near-surface soil moisture observations. Consequently, no error terms have been added to the synthetic truth, model, or forcing, and so the results provide an upper-limit for how effectively the assimilation can constrain the model soil moisture profile from near-surface observations. In reality, the ability of the EKF to correct the model soil moisture is greatly reduced by inaccuracies in the observations and the forward model.

Table 6.8: Initial conditions and assimilated observations for the synthetic assimilation experiments.

Experiment	Assimilated observations (\mathbf{y})	Initial \mathbf{x}
OPN_ACCESS	None	From ACCESS
EKF_25	S_1 from OPN_ACCESS	25%
OPN_25	None	25%
EKF_75	S_1 from OPN_ACCESS	75%
OPN_75	None	75%

6.4.1 Time series examples

Since the model Jacobians, and hence the EKF, are dependent on the local soil moisture conditions the results of the synthetic experiments are examined below at four locations, selected to sample each of the main Australian climate regimes identified in Section 6.2.4.

An arid location

Figure 6.10 shows the soil moisture time series from each of the perturbed EKF and open-loop experiments at an arid location in central Australia. At this location, the OPN_ACCESS soil moisture in all model layers remained close to its lower limit throughout the year (recall from Section 3.4.2 that in layers two to four the model soil moisture is bounded below at 10% of the wilting point). Superimposed on this, S_1 shows a series of precipitation induced spikes, after which the soil moisture rapidly dried back to its lower limit. The larger precipitation events also generated a small response in S_2 , which then gradually dried to its lower bound, while S_3 and S_4 show no variability (at the plotted scale).

Both of the perturbed initial conditions were wetter than the initial conditions in OPN_ACCESS, and in both S_1 dried rapidly towards OPN_ACCESS. In both cases S_1 was within 1% of OPN_ACCESS within 10 days, after which the S_1 time series were indistinguishable from the synthetic truth until precipitation occurred and the S_1 time series briefly diverged. The deeper soil layers dried more gradually, so that even after a year S_3 and S_4 did not converge to OPN_ACCESS. With the exception of brief periods following precipitation, the S_1 from the reference and perturbed open-loop experiments were very similar. Since the S_1 from the perturbed experiments had no signal of the wetter soils in the underlying layers, the assimilation had little impact, and the EKF_25 and EKF_75 time series were similar to the corresponding perturbed open-loop experiments. Figure 6.11 shows the analysis increments at the same location, demonstrating that increments were added to the model only during the first few days of the experiment (reducing the difference to the OPN_ACCESS errors more rapidly than for the open-loop simulations), and immediately after precipitation events.

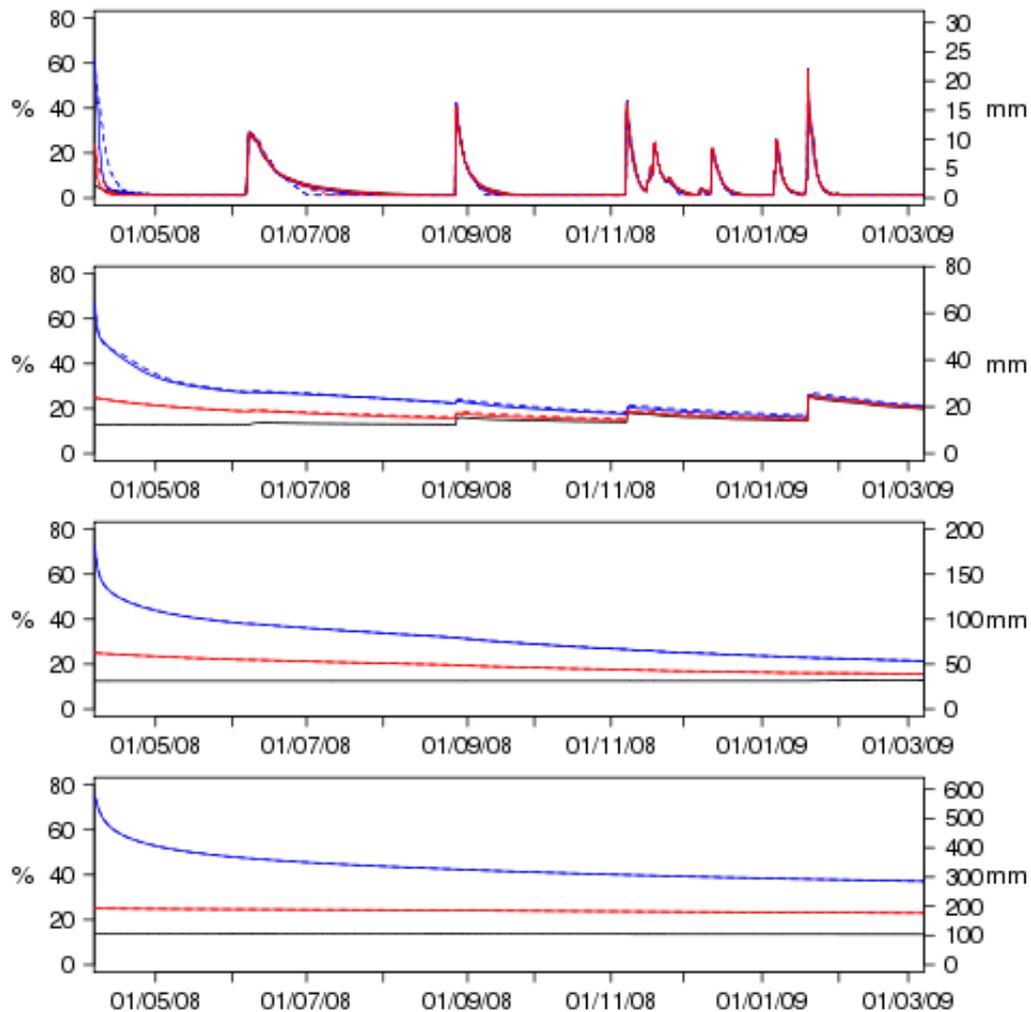


Figure 6.10: Soil moisture time series from the synthetic experiments at (137.375,-26.0) in arid (Bwh) central Australia, in units of % of saturation (left axis) and mm of soil moisture (right axis). The panels shows each soil layer (from top to bottom), from the EKF analyses with incorrect initialisation (solid, coloured), open-loop with incorrect initialisation (dashed, coloured), and the assumed truth from the reference open-loop (solid, black). Red (blue) lines indicate soil moisture initialised at 25% (75%) on 1 April 2008.

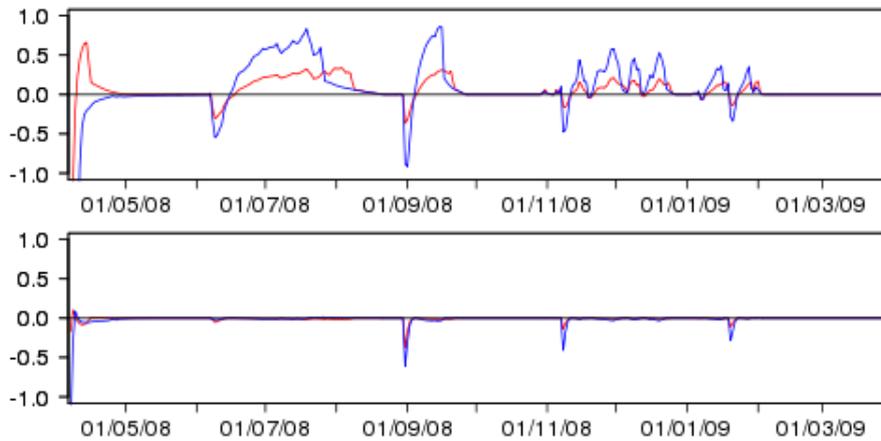


Figure 6.11: Time series of the soil moisture analysis increments (% of saturation) from the synthetic experiments at (137.375,-26.0), in arid (Bwh) central Australia, for layers one (upper) and two (lower). Red (blue) lines indicate soil moisture initialised at 25% (75%) on 1 April 2008.

A tropical location

Figures 6.12 and 6.13 show time series plots of the soil moisture and soil moisture increments for the synthetic experiments at a tropical location in the Northern Territory. During the dry season (May - October) the climate was similar to the arid example discussed above, and the soil moisture time series in Figure 6.12 show a similar decoupling between S_1 and the underlying soil moisture, leading to the same lack of impact of the assimilation as occurred at the arid location. At the onset of the wet season in November, the soil moisture throughout the root-zone (and also in S_4) increased rapidly to about 60% of saturation, regardless of the preexisting soil moisture states. The small increments added to S_1 during the wet season in Figure 6.13 show that in between precipitation events there was some divergence between S_1 from the perturbed assimilation experiments and the assimilated OPN_ACCESS values, however this divergence was small (less than 0.5% for both experiments), and the assimilation had little impact on the model soil moisture during the wet season.

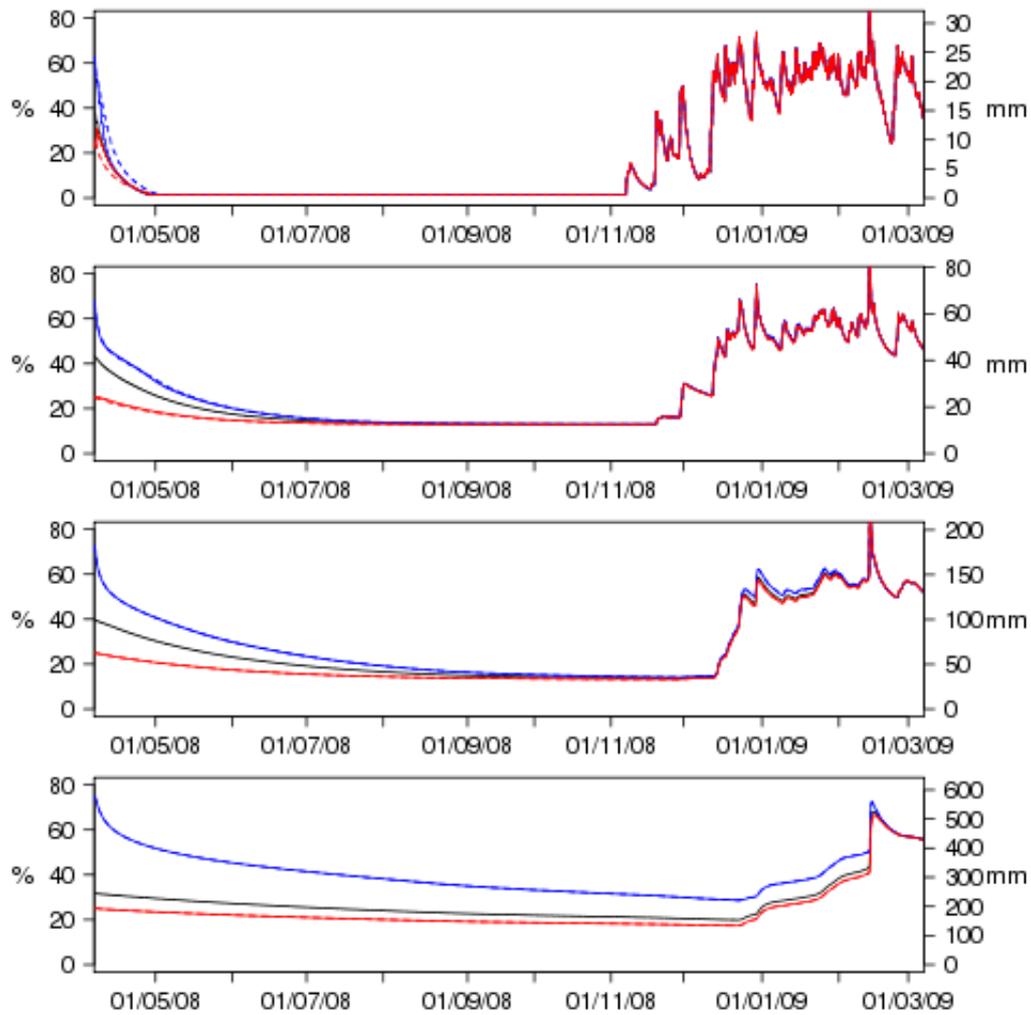


Figure 6.12: Soil moisture time series from the synthetic experiments at (133.25, -13.625) in tropical (Aw) Northern Territory, in units of % of saturation (left axis) and mm of soil moisture (right axis). The panels shows each soil layer (from top to bottom), from the EKF analyses with incorrect initialisation (solid, coloured), open-loop with incorrect initialisation (dashed, coloured), and the assumed truth from the reference open-loop (solid, black). Red (blue) lines indicate soil moisture initialised at 25% (75%) on 1 April 2008.

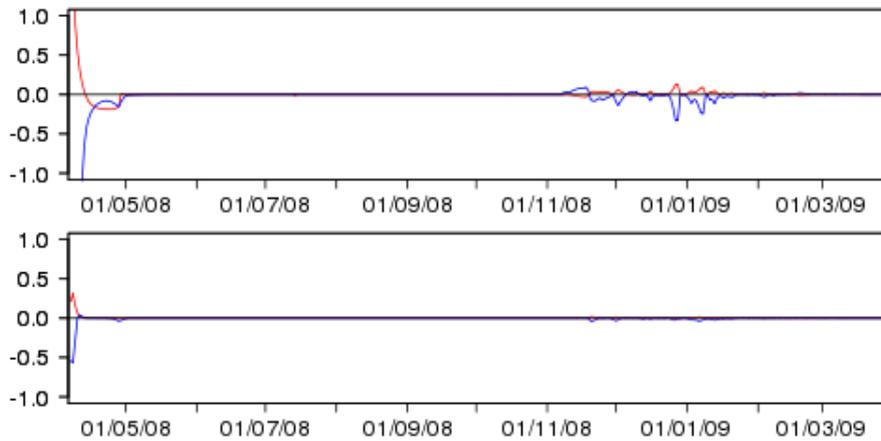


Figure 6.13: Time series of the soil moisture analysis increments (% of saturation) from the synthetic experiments at (133.25, -13.625) in tropical (Aw) Northern Territory, for layers one (upper) and two (lower). Red (blue) lines indicate soil moisture initialised at 25% (75%) on 1 April 2008.

Temperate locations

Figures 6.14 and 6.15 show time series plots from the synthetic experiments for a temperate-oceanic location in central Tasmania, while Figures 6.16 and 6.17 shows the equivalent plots for a temperate-Mediterranean location in south West Australia. At both locations S_1 was much higher than in the previous two examples, and it had a smooth seasonal cycle with maxima in winter, overlaid with a series of precipitation induced spikes. S_2 and S_3 show a filtered response to the variability in S_1 , while S_4 gradually increased throughout the year. The initial soil moisture in OPN_ACCESS was between 40 - 50% at both locations, so that the perturbed experiments were initialised to either side of OPN_ACCESS. The wetter conditions caused the open-loop S_1 to converge more slowly towards OPN_ACCESS than in the previous examples, while the underlying layers converged more rapidly.

The assimilation of S_1 was more effective at the temperate locations than in the previous examples. Each analysis update corrected S_1 towards the observed S_1 , however the differences in the underlying soil layers caused the subsequent forecasts to drift away from OPN_ACCESS in between the observations. As a result, it took EKF_75 just 9 days to correct S_1 to within 5% of OPN_ACCESS in Figure 6.14, although it took 190 days for the error to be reduced to less than 1%, and during this time both S_1 and S_2 were being regularly corrected.

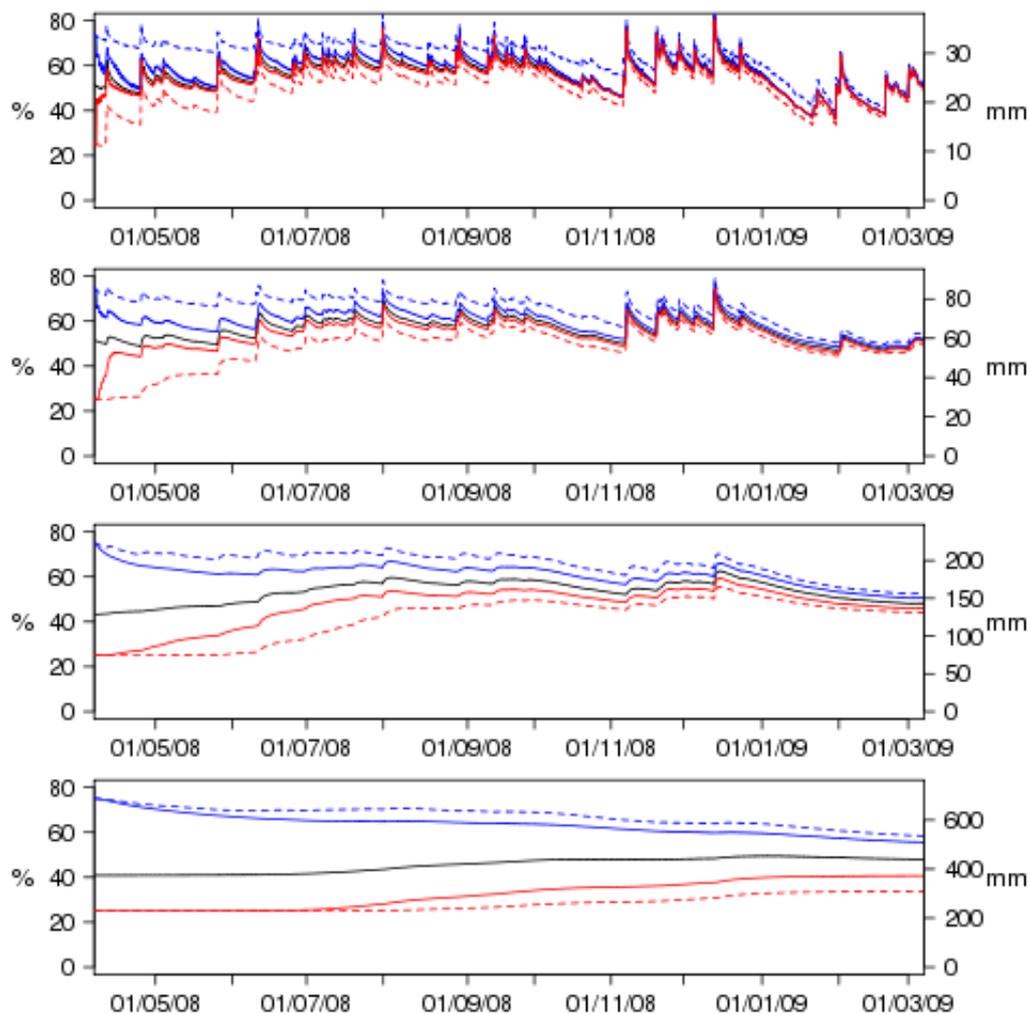


Figure 6.14: Soil moisture time series from the synthetic experiments at (147.125, -41.75), in temperate-oceanic (Cfb) Tasmania, in units of % of saturation (left axis) and mm of soil moisture (right axis). The panels shows each soil layer (from top to bottom), from the EKF analyses with incorrect initialisation (solid, coloured), open-loop with incorrect initialisation (dashed, coloured), and the assumed truth from the reference open-loop (solid, black). Red (blue) lines indicate soil moisture initialised at 25% (75%) on 1 April 2008.

Compared to the previous examples, a larger net volume of water was added to the surface, due to the persistence of the observation increments (and the larger analysis increments added to S_2). Combined with the enhanced rate of drainage in wetter conditions, this resulted in the assimilation having a significant impact on S_2 and S_3 , and even on the S_4 .

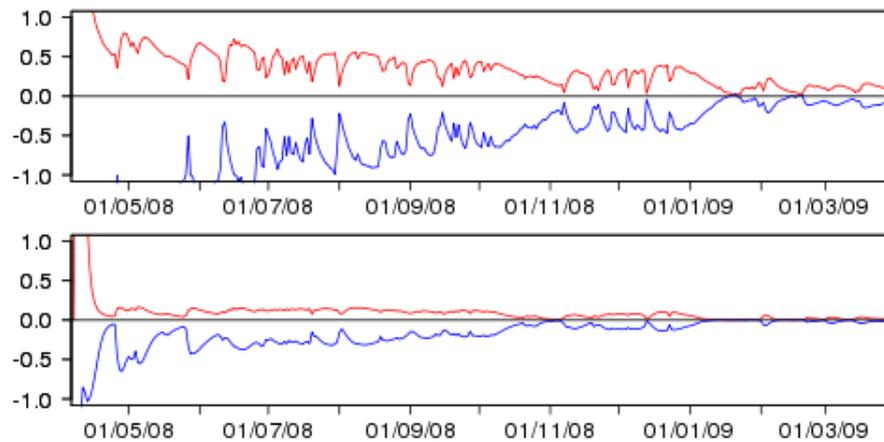


Figure 6.15: Time series of the soil moisture analysis increments (% of saturation) from the synthetic experiments at $(147.125, -41.75)$, in temperate-oceanic (Cfb) Tasmania, for layers one (upper) and two (lower). Red (blue) lines indicate soil moisture initialised at 25% (75%) on 1 April 2008.

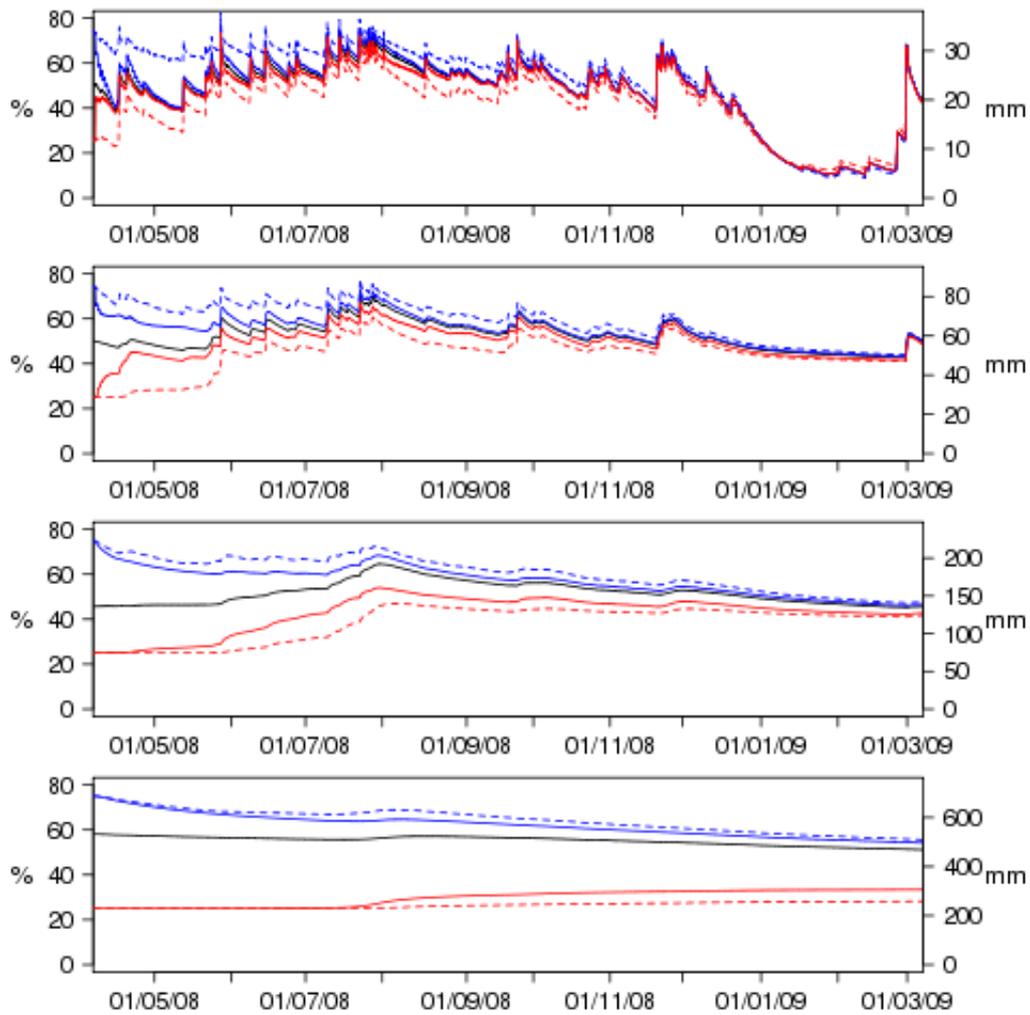


Figure 6.16: Soil moisture time series from the synthetic experiments at (116.75, -34.25) in temperate-Mediterranean (Cfb) south West Australia, in units of % of saturation (left axis) and mm of soil moisture (right axis). The panels shows each soil layer (from top to bottom), from the EKF analyses with incorrect initialisation (solid, coloured), open-loop with incorrect initialisation (dashed, coloured), and the assumed truth from the reference open-loop (solid, black). Red (blue) lines indicate soil moisture initialised at 25% (75%) on 1 April 2008.

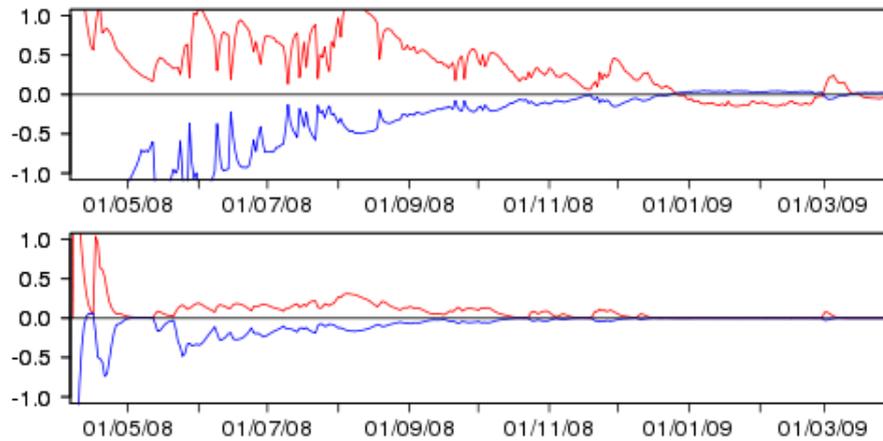


Figure 6.17: Time series of the soil moisture analysis increments (% of saturation) from the synthetic experiments at (116.75, -34.25) in temperate-Mediterranean (Cfb) south West Australia, for layers one (upper) and two (lower). Red (blue) lines indicate soil moisture initialised at 25% (75%) on 1 April 2008.

6.4.2 Net impact

Figure 6.18 compares the soil moisture from OPN_ACCESS to the soil moisture from the EKF_25 and OPEN_25 simulations after three months of assimilating S_1 , demonstrating that the above results can be generalised for each climate regime (maps for EKF_75 and OPEN_75 indicate similar conclusions). For the arid and tropical regions (since the three month period considered is in the dry season), the OPEN_25 S_1 was already very close to the assumed truth. In the temperate zones, the OPN_25 S_1 was further away from the assumed truth, and the assimilation substantially reduced this difference. Additionally, the only corrections (evident at the plotted scale) to the MOSES root-zone soil moisture occurred in the temperate regions of southeast and southwest Australia, where the assimilation corrected S_{RZ} towards OPN_ACCESS (although the EKF_25 S_{RZ} errors were still between 5 and 10% of saturation).

The statistics in Table 6.9 quantitatively confirms these results. In the arid and tropical zones, the S_1 from all of the experiments were similar, while the differences between the S_{RZ} from the perturbed experiments and OPN_ACCESS were much greater, and these difference was not substantially reduced by the assimilation. In contrast, in the temperate zones the differences between the perturbed open-loops and OPN_ACCESS S_1 were larger, and assimilating S_1

reduced these differences, while also reducing the differences between the S_{RZ} forecasts and the OPN_ACCESS. Consequently, in temperate zones assimilating S_1 reduced the mean difference between the S_{RZ} from the perturbed simulations and OPN_ACCESS by between 30 - 40 % of the differences for the perturbed open-loops.

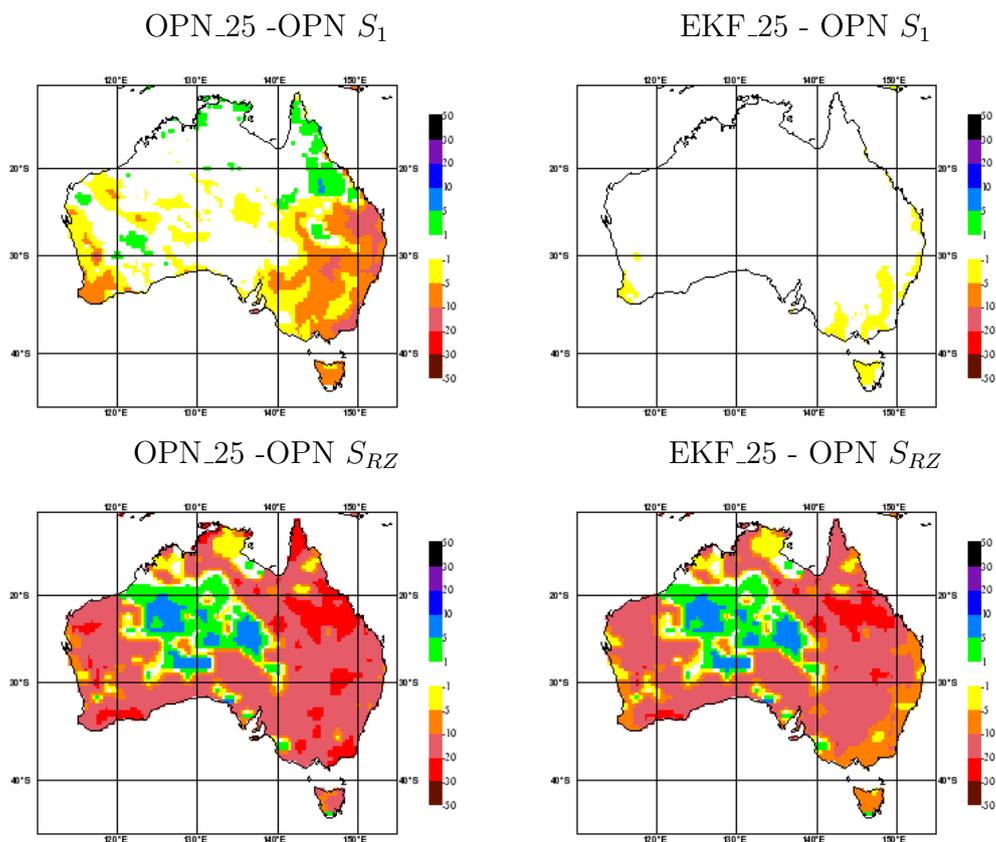


Figure 6.18: Maps of the soil moisture difference (% of saturation) from OPN_ACCESS for OPN_25 (left) and EKF_25 (right) on 1 July 2008, for S_1 (upper) and S_{RZ} (lower).

In summary, the synthetic experiments demonstrated that for the MOSES model over Australia assimilating S_1 observations can only constrain the model S_{RZ} in temperate conditions. In the tropics, the intense wet season precipitation overrides any memory of the prior soil moisture (or soil moisture updates), so that the assimilation has little impact. Additionally, in arid conditions the near-surface soil moisture becomes decoupled from the underlying soil moisture, preventing the signal from the nonsurface layers from reaching S_1 , and preventing any updates to S_1 from diffusing through the nonsurface layers. Capehart

Table 6.9: Mean soil moisture difference from OPN_ACCESS on 1 July 2008 (% of saturation) for each synthetic experiment, averaged over each climate zone.

	Arid			
	OL_25	EKF_25	OL_75	EKF_75
S_1	-1.6	-0.16	6.1	0.016
S_{RZ}	-10.3	-9.3	13.4	12.3
	Tropical			
S_1	0.3	0.1	0.9	0.01
S_{RZ}	-12.3	-9.8	3.8	3.5
	Temperate-Oceanic			
S_1	-4.2	-1.1	10.7	2.1
S_{RZ}	-12.7	-9.1	13.9	10.0
	Temperate-Mediterranean			
S_1	-8.7	-1.2	10.1	1.4
S_{RZ}	-16.6	-9.7	11.1	6.4

and Carlson (1997) propose that this decoupling occurs due to an enhanced rate of soil drying close to the surface (due to bare soil evaporation), so that over time the hydraulic conductivity in the surface layer decreases, until the near-surface layer is effectively cut-off from the soil below. It has been noted previously that this decoupling will limit the effectiveness of near-surface soil moisture assimilation in arid regions (e.g., Li and Islam, 2002). Li and Islam (2002) also note that the decoupling mechanism described above will be more pronounced where vegetation cover is lower (such as in arid Australia). However, as noted in Section 6.3.2 it is possible that the decoupling observed in these experiments is unrealistically severe, due to unrealistic soil physics in MOSES (specifically insufficient vertical soil moisture exchange). Note that the soil moisture Jacobians maps in Section 6.3.2, as well as the results of the synthetic experiments (Figure 6.18), indicate that this decoupling is not limited to extreme arid conditions, since the hydraulic conductivity in MOSES does not increase significantly until the soil moisture is reasonably wet (above about 40% in Figure 6.5).

In contrast, for temperate regions neither of the above phenomena occur, and the assimilation of S_1 has a much greater impact on the model root-zone soil moisture. Since the EKF updates only the top two soil moisture layers (covering 35 cm), with most of this in the 10 cm surface layer, the impact on

the root-zone evolves rather slowly, although it is still substantial. For both of the perturbed initial conditions, assimilating the S_1 data for three months reduced the mean difference from OPN_ACCESS by 30-40% of the perturbed open-loop values in the temperate zones in Table 6.9. While these reductions are insufficient to overcome grossly incorrect initial values, this is generally not required for NWP due to the continuous nature of those models.

6.5 Results: Assimilating AMSR-E data

This section presents the results of assimilating the AMSR-E near-surface soil moisture observations into MOSES. First, the rescaling of the AMSR-E data to remove the systematic differences between the modeled and assimilated soil moisture is presented. The results of the assimilation are then examined at the same four locations presented above for the synthetic experiments. Additionally, the net impact of assimilating the AMSR-E data is tested by examining the difference between the open-loop and EKF forecasts of soil moisture and evapotranspiration across Australia. Finally, the impact of the assimilation on the soil moisture forecast skill is tested against the Murrumbidgee Monitoring Network in situ soil moisture observations, and against a model simulation forced with high quality observed precipitation.

6.5.1 Rescaling the AMSR-E observations

The AMSR-E data have been rescaled to better match the model climatology by matching the CDF of the observations to the CDF of the ACCESS S_1 forecasts. Only one year of soil moisture forecasts were available from ACCESS, and as was done for the ISBA experiments, ergodic substitution has been used to increase the sample size for estimating the CDF, following Reichle and Koster (2004). Specifically, all of the data from the surrounding 1° area were used to estimate the CDF at each model grid cell. The AMSR-E data were converted from m^3m^{-3} to % of saturation, based on the soil moisture at saturation in MOSES, before the CDF-matching operator was applied. The problems encountered during the AMSR-E / ISBA rescaling, associated with the inconsistent magnitude of the responses to the seasonal cycle and precipitation in the modeled and observed soil moisture, were not detected for MOSES. Consequently, the seasonal cycle of the AMSR-E data did not need to be corrected before applying

the CDF-matching, as was done in Chapter 5.

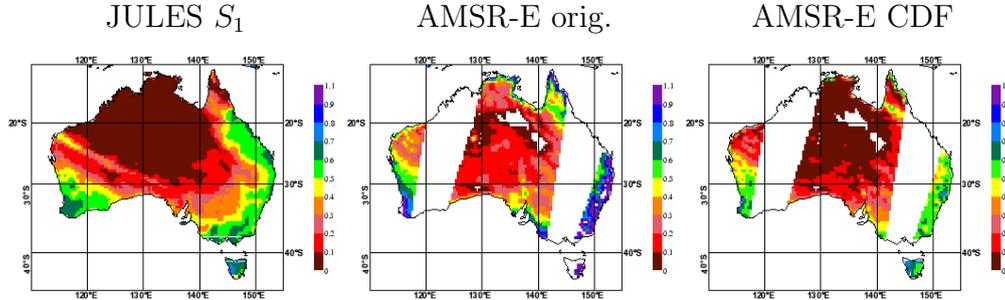


Figure 6.19: Near-surface soil moisture (as a fraction of saturation) on 1 August 2008 at 15:00 UTC from MOSES, and AMSR-E, before and after CDF-matching.

The CDF-matching successfully reduced the magnitude of the bias (RMSD) between the model and the observations over the one year experiment from -6% (13%) in the original data to 0.007% (10%) in the CDF-matched data. Figure 6.19 compares the AMSR-E near-surface soil moisture before and after the CDF-matching on 1 August 2008. The large scale features of the bias corrected AMSR-E soil moisture appear similar to those from MOSES, while some of the smaller scale features from the original AMSR-E data are still evident.

Figure 6.20 shows time series of the mean difference between the assimilated AMSR-E data and the open-loop forecasts of S_1 , averaged over each climate zone (for an assimilation experiment this would be the mean observation increment). The arid zone had the smallest biases, consistent with the limited soil moisture variability, while the tropical zone had the largest biases, which were as high as 10% of saturation. In the tropics the observation increments were persistently positive in the dry season, and persistently negative in the wet season. The time series of the mean precipitation for each climate zone from the BoM's daily rain gauge analyses and the ACCESS forecasts in Figure 6.21 show that the wet season precipitation in ACCESS was biased low, suggesting that the negative observation increments in the tropics during the wet season were incorrect. The poor performance of the bias removal in tropical conditions is likely associated with the tropics having two distinct climate modes, since this cannot be accurately described by a single set of statistics within the CDF-matching operator (more so than for extra-tropical regions with four less contrasted seasons). A better approach to rescaling the remotely sensed soil moisture in the tropics

would be to calculate a separate rescaling operator for each of the wet and dry seasons. However, this was not done in this study, since the data record available for the rescaling was already very short. Additionally, the limited impact of the S_1 assimilation in the tropics in the synthetic experiments above (and also in the AMSR-E experiments below) indicates that inaccuracies in the assimilated AMSR-E data in the tropics will be of little consequence.

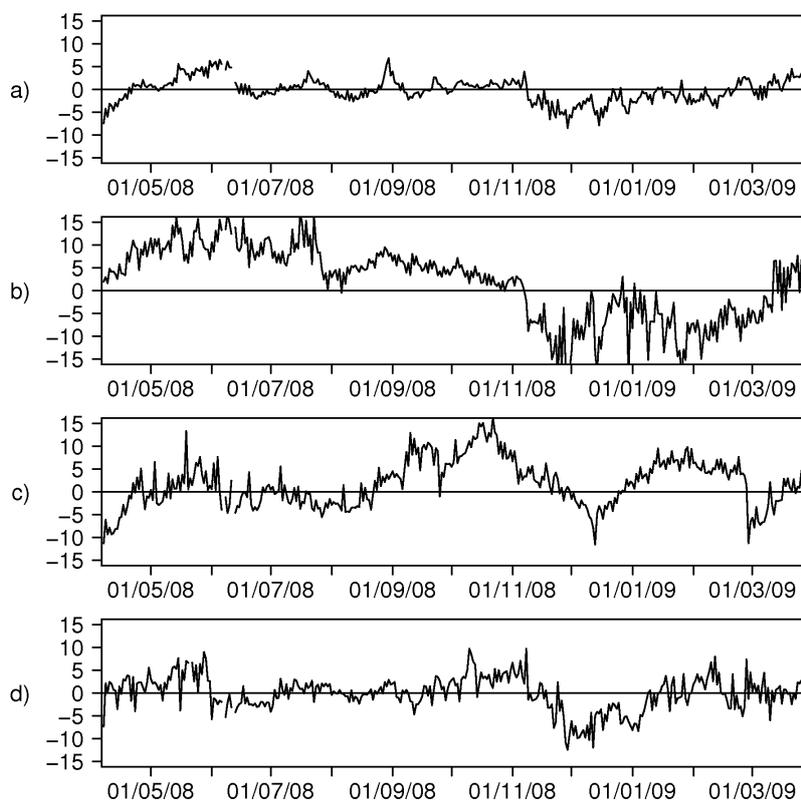


Figure 6.20: Time series of the assimilated AMSR-E observations minus the open-loop S_1 (% of saturation), averaged over the a) arid, b) tropical, c) temperate-oceanic, and d) temperate-Mediterranean climate zones.

Both of the temperate zones also showed periods of extended bias between the AMSR-E and MOSES soil moisture in Figure 6.20, although these biases were less extreme than for the tropics. Again, the soil moisture observation increments were not supported by the precipitation time series in Figure 6.21. For both temperate zones, the model precipitation was biased low for the first half of the year, while the soil moisture observation increments had a slight tendency to be negative during this period. Additionally, towards the end of winter a large positive soil moisture observation increment developed, which

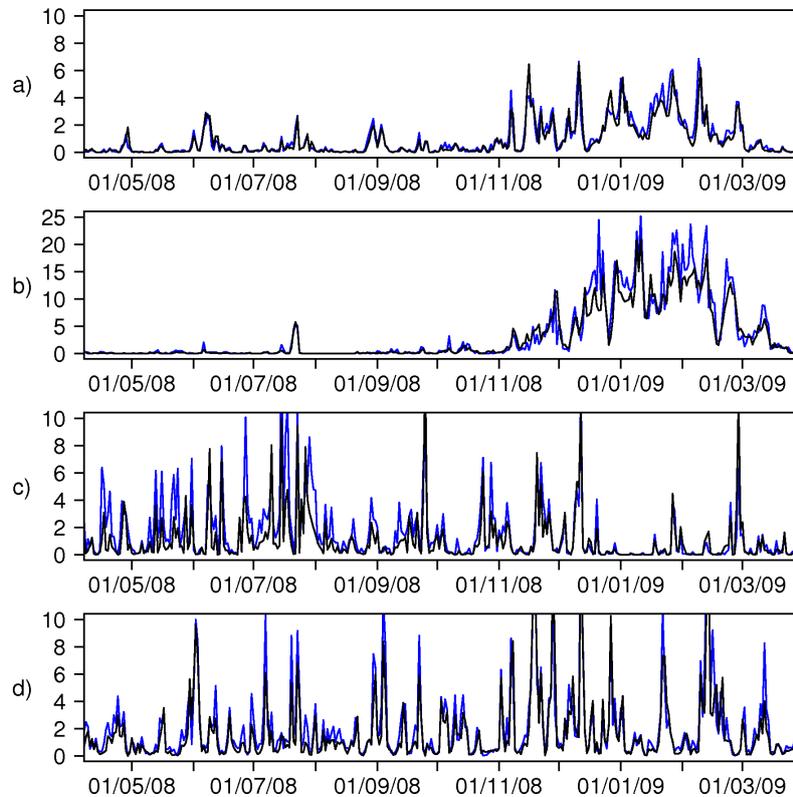


Figure 6.21: Time series of precipitation (mm day^{-1}) from the BoM's rain gauge analysis (blue) and ACCESS forecasts (black), averaged over the a) arid, b) tropical, c) temperate-oceanic, and d) temperate-Mediterranean climate zones.

was not reflected by the precipitation time series. Examining the soil moisture at individual locations (not shown) suggests that this was caused by differences in the timing of the seasonal cycle from AMSR-E and MOSES.

Since the AMSR-E observations have been bias-corrected to the model soil moisture simulated by forcing the model with the (biased) ACCESS precipitation forecasts, the bias-corrected AMSR-E observations were not expected to reflect the annual precipitation biases. However, it was hoped that the assimilated AMSR-E data would include a signal of any seasonal-scale errors in the precipitation forcing, and yet there are few similarities between the soil moisture biases in Figure 6.20 and the precipitation biases in Figure 6.21. This could indicate errors in the assimilated AMSR-E data, due to representativity errors between the modeled and observed soil moisture (that are not resolved by statistical rescaling), or due to inaccuracies in the observations, or both.

Alternatively, the precipitation-induced soil moisture errors may have been obscured by other model errors, such as errors in the radiation forcing or model evapotranspiration parameterisation.

6.5.2 Time series examples

An arid location

Figure 6.22 shows time series of the soil moisture from the open-loop and the EKF assimilation of AMSR-E, together with the modeled and observed precipitation, at the arid location in central Australia that was presented above for the synthetic experiments. The precipitation time series indicate that the model forecast precipitation was biased high: while the timing of each event was generally correct, the magnitude was consistently too large. In general, the AMSR-E observations had a better correspondence to the observed precipitation magnitude than the model forecast soil moisture.

Consistent with the decoupling that occurred in the synthetic experiments, the assimilation of AMSR-E very quickly constrained the model S_1 to fit the observations, while having little impact on the underlying soil moisture. The exception is immediately after precipitation events, when the temporarily wet soils had much higher hydraulic conductivity. Figure 6.23 shows the corresponding plot of analysis increments from the EKF experiment. In several instances relatively large increments were added to S_2 in response to S_1 errors associated with the precipitation errors. For example S_2 was reduced by nearly 10% in early November, and yet the annual range of S_2 in the open-loop was just 12%. These unrealistically large updates occurred because the EKF calculates the update to S_2 based on the assumption that the observed error in S_1 is caused by incorrect S_2 , rather than erroneous forcing (although this is moderated by the S_1 error covariance in \mathbf{P}).

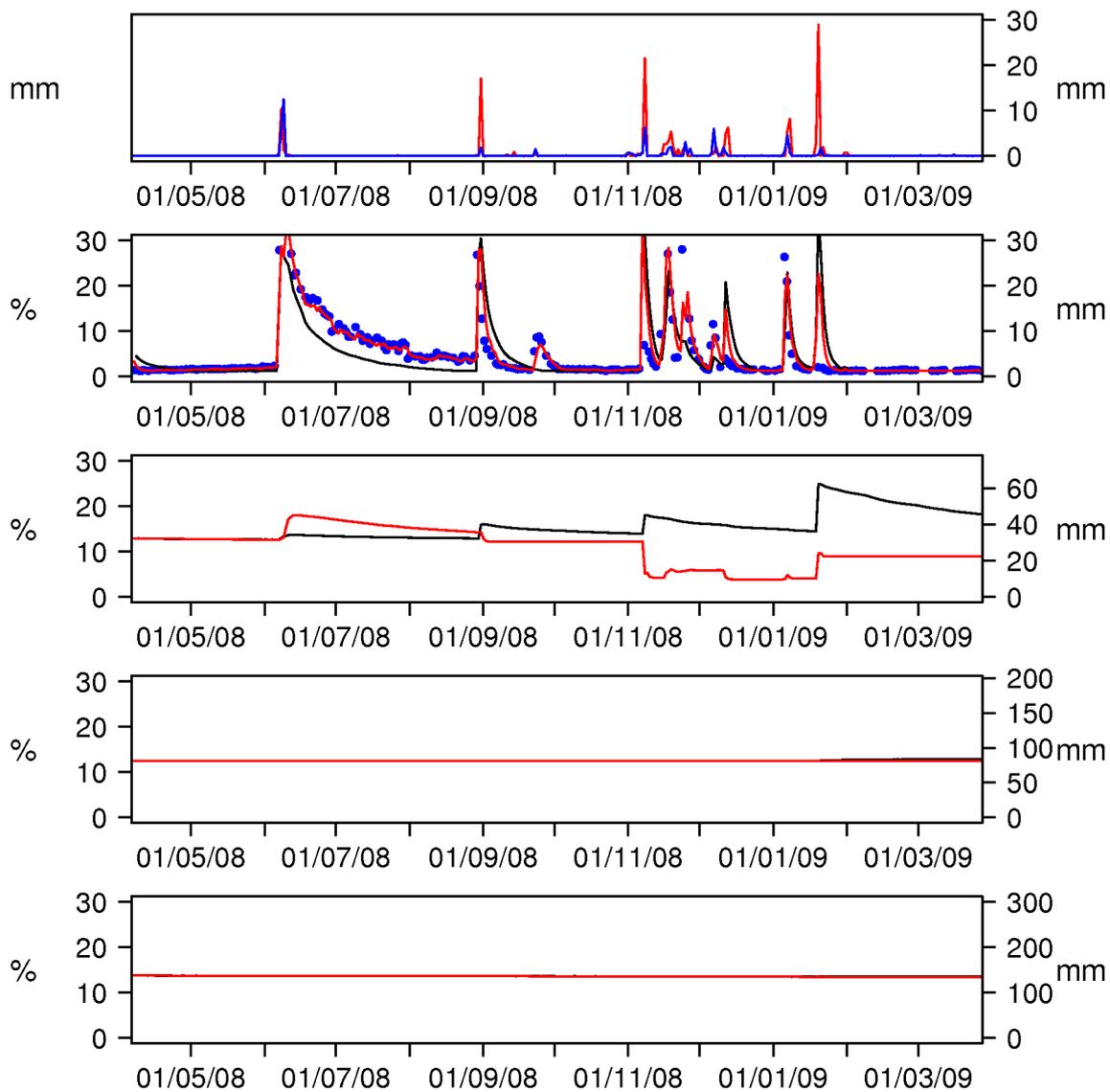


Figure 6.22: Time series of the modeled and observed daily precipitation and soil moisture at (137.375,-26.0) in arid (Bwh) central Australia. The upper panel shows the observed (blue) and modeled (red) precipitation (mm day^{-1}). The next four panels show the soil moisture in each layer (from top to bottom), from the EKF (red), the open-loop (black), and AMSR-E (blue), in units of % of saturation (left axis) and mm of soil moisture (right axis).

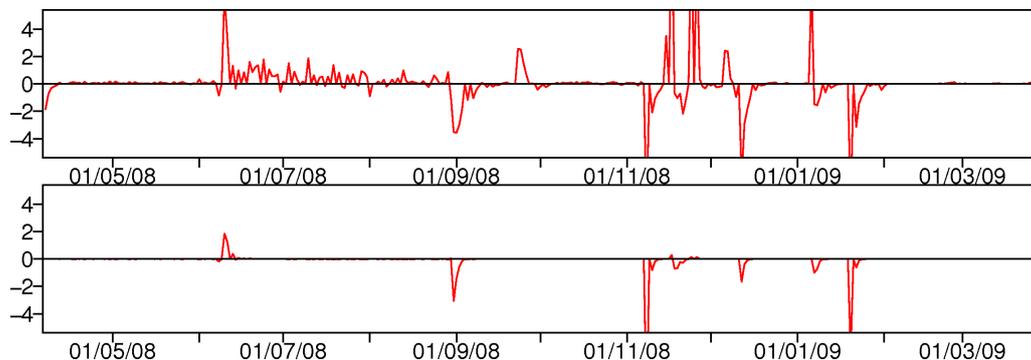


Figure 6.23: Time series of the soil moisture analysis increments (% of saturation) in layers one (upper) and two (lower) at (137.375,-26.0) in arid (Bwh) central Australia.

A tropical location

Figures 6.24 and 6.25 show time series of the soil moisture and analysis increments, respectively, at the tropical location in the Northern Territory. As with the arid example above, the timing of the model precipitation was reasonably accurate, while the magnitude was not. In this case the precipitation observations suggest that the forecast precipitation was initially too high at the onset of the wet season, and then too low in the middle of the wet season (around January). The AMSR-E observations accurately detected the lower precipitation at the onset of the wet season, as well as the large event at the end of the wet season that was underestimated by the model. However, the AMSR-E observations were consistently lower than the model for the remainder of the wet season, in contradiction with the precipitation observations. This inconsistency is likely associated with the poor performance of the AMSR-E bias-correction in the tropical regions mentioned in Section 6.5.1. Additionally, despite the lack of precipitation during the dry season, the AMSR-E S_1 time series had some noise in April and May 2008.

Again the results of assimilating the AMSR-E data were consistent with the findings from the synthetic experiments. The assimilation had a significant impact on the model soil moisture only at the onset of the wet season, when the EKF drew S_1 down and these reductions were transmitted through to the deeper soil layers. However, as the wet season progressed the impact of these reductions was lost, since the soil moisture was quickly dominated by precipitation, so that the EKF soil moisture time series converged to the open-loop time

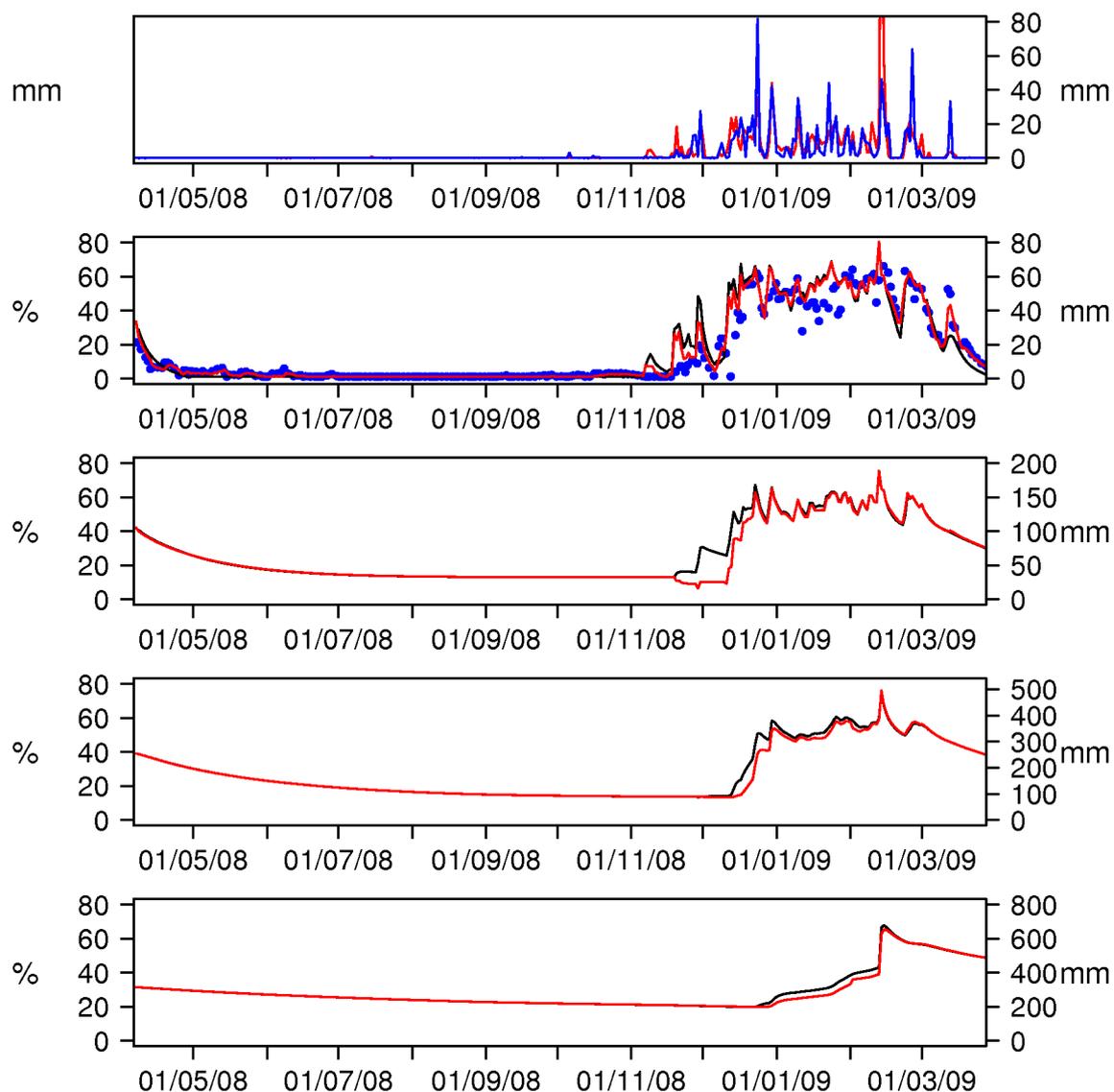


Figure 6.24: Time series of the modeled and observed daily precipitation and soil moisture at (133.25, -13.625) in tropical (Aw) Northern Territory. The upper panel shows the observed (blue) and modeled (red) precipitation (mm day^{-1}). The next four panels show the soil moisture in each layer (from top to bottom), from the EKF (red), the open-loop (black), and AMSR-E (blue), in units of % of saturation (left axis) and mm of soil moisture (right axis).

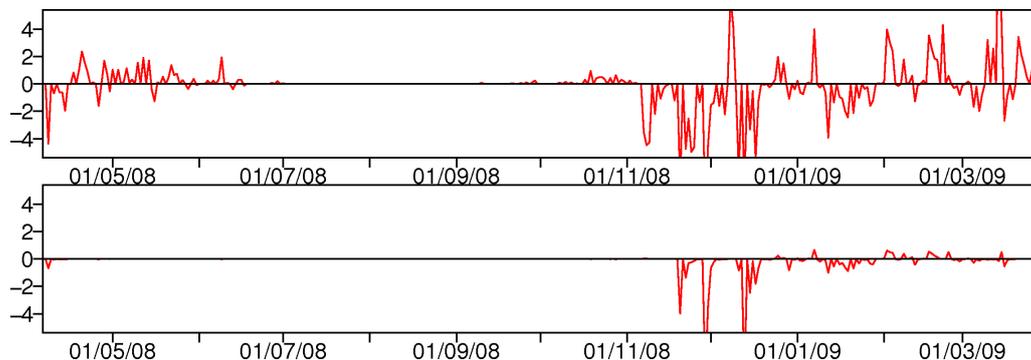


Figure 6.25: Time series of the soil moisture analysis increments (% of saturation) in layers one (upper) and two (lower) at (133.25, -13.625) in tropical (Aw) Northern Territory

series. Note that significant updates were added to the model soil moisture in Figure 6.25 during this period, however the impact of these updates was quickly overwhelmed by the subsequent precipitation.

Temperate locations

Figures 6.26 and 6.27 show the soil moisture and analysis increment time series, respectively, for the cool oceanic location in central Tasmania, while Figures 6.28 and 6.29 show the equivalent plots for the temperate-Mediterranean location in south West Australia. In Figure 6.26 for the Tasmanian location, the AMSR-E observations are consistent with the observed precipitation in some instances (e.g., the reduced peak on 1 August 2008), although not in others (e.g., there is no AMSR-E peak corresponding to the observed rain on 3 February 2009). There were several extended periods during which the AMSR-E time series diverged from the open-loop (e.g., January 2008), and there were also some anomalously low AMSR-E observations which were likely incorrect (e.g., 11 March 2009). As with the synthetic experiments, assimilating S_1 had a much greater net impact on the model soil moisture at the temperate locations than at the arid or tropical locations, and in Figure 6.26 the assimilation effectively drew the model S_1 and S_2 down, resulting in slight reductions in S_3 and S_4 .

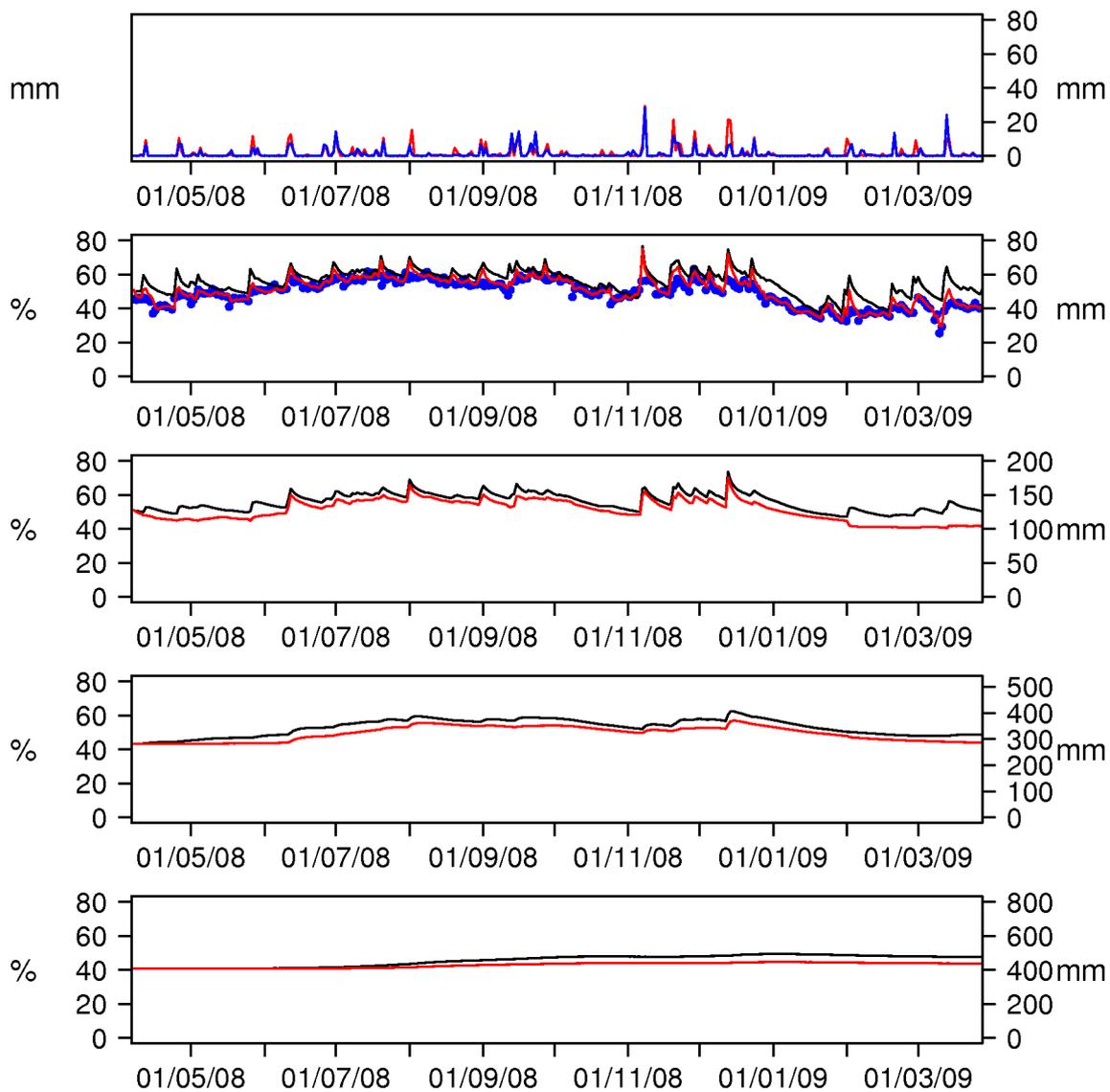


Figure 6.26: Time series of the modeled and observed daily precipitation and soil moisture at (147.125, -41.75), in temperate-oceanic (Cfb) Tasmania. The upper panel shows the observed (blue) and modeled (red) precipitation (mm day^{-1}). The next four panels show the soil moisture in each layer (from top to bottom), from the EKF (red), the open-loop (black), and AMSR-E (blue), in units of % of saturation (left axis) and mm of soil moisture (right axis).

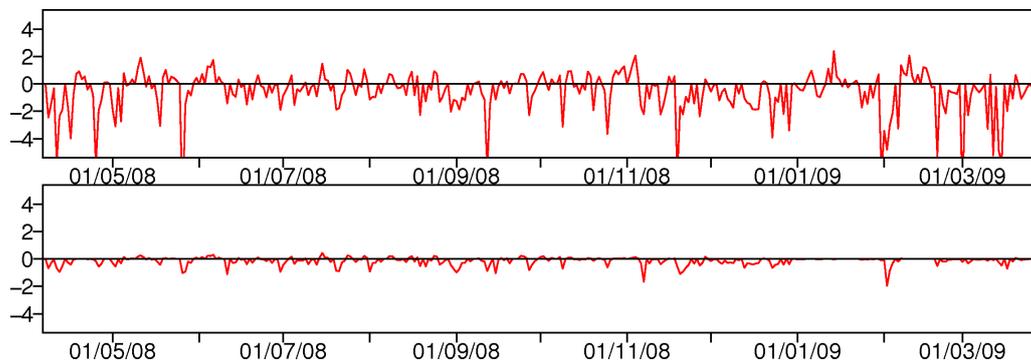


Figure 6.27: Time series of the soil moisture analysis increments (% of saturation) in layers one (upper) and two (lower) at (147.125, -41.75) in temperate-oceanic (Cfb) Tasmania.

In contrast to the above examples, the AMSR-E observations for the West Australian location in Figure 6.28 appear to be unrealistic, with little correspondence to the observed precipitation. In particular the AMSR-E time series was rather smooth, and did not reflect the precipitation events indicated by both the modeled and observed precipitation time series. The AMSR-E time series also had a different seasonal cycle to the model, with the winter maxima occurring several months later than in the model. The EKF drew the model soil moisture down, generating slight reductions in the soil moisture in layer three and a very small impact in layer four (which had limited annual variability).

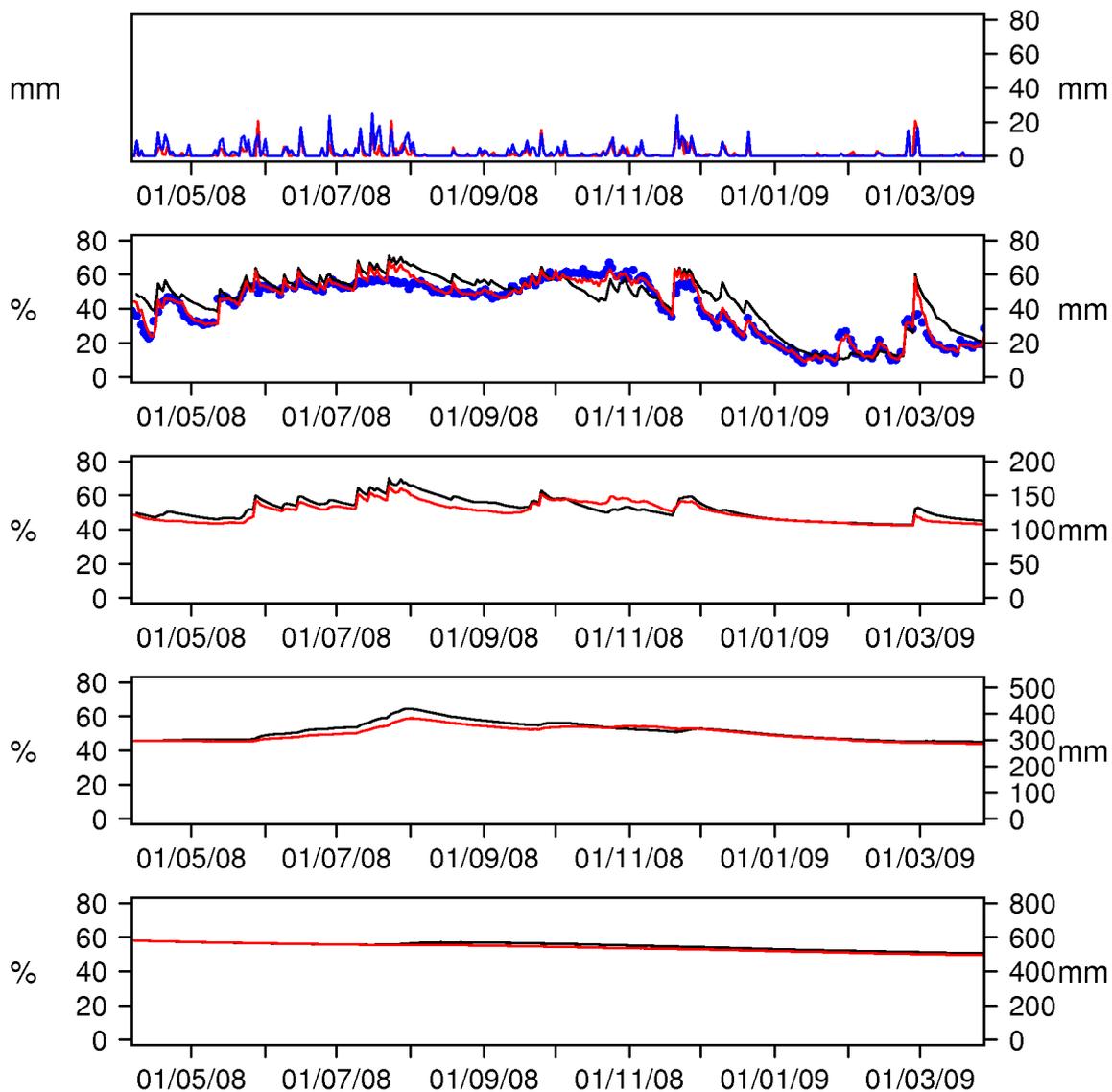


Figure 6.28: Time series of the modeled and observed daily precipitation and soil moisture at (116.75, -34.25) in temperate-Mediterranean (Cfb) south West Australia. The upper panel shows the observed (blue) and modeled (red) precipitation (mm day^{-1}). The next four panels show the soil moisture in each layer (from top to bottom), from the EKF (red), the open-loop (black), and AMSR-E (blue), in units of % of saturation (left axis) and mm of soil moisture (right axis).

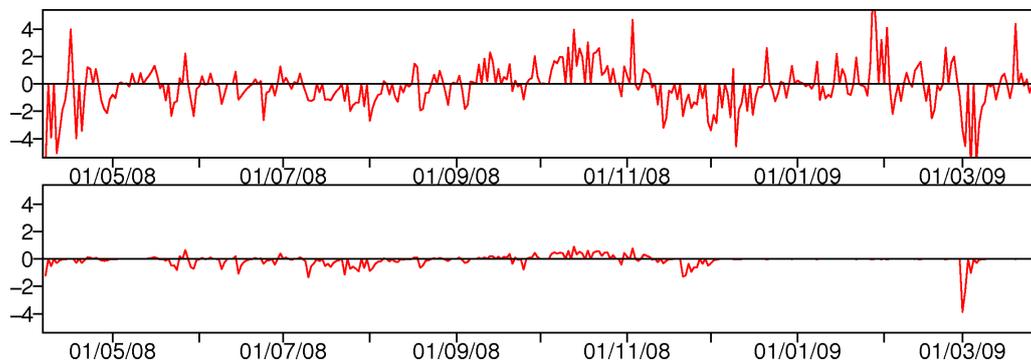


Figure 6.29: Time series of the soil moisture analysis increments (% of saturation) in layers one (upper) and two (lower) at (116.75, -34.25) in temperate-Mediterranean (Cfb) south West Australia.

In summary, the above examples extend the findings from the synthetic experiments. The assimilation cannot constrain the root-zone soil moisture in dry conditions (due to the decoupling of the vertical soil moisture profile), or during the wet season in the tropics (due to the lack of soil moisture memory associated with the intense precipitation). However, in temperate conditions there were periods of persistent divergence between the modeled and observed S_1 , resulting in substantial corrections to the root-zone soil moisture (due both to updates to S_2 , and infiltration of updates to S_1). From the synthetic experiments, this divergence between the observed and modeled S_1 could be a signal of incorrect soil moisture in the nonsurface layers. However, it could also indicate systematic differences between the modeled and observed soil moisture: recall from Chapter 4 that the AMSR-E observations often diverged from the in situ observations for extended time periods (for example due to the shallower layer observed by AMSR-E drying more rapidly).

While the EKF assimilation of S_1 cannot always correct for errors in the underlying soil moisture, it will be more effective at correcting errors in the forecast precipitation, which enter the model soil moisture through S_1 . Qualitatively, comparison between the AMSR-E and precipitation observations in Figures 6.22-6.28 suggest that in many instances AMSR-E accurately detected errors in the ACCESS precipitation forecasts. Additionally, the EKF effectively adjusted S_1 in response to the detected precipitation errors. However, there were several instances when unrealistically large updates were applied to S_2 in response to the precipitation signal from AMSR-E, since the EKF updates were derived from the assumption that errors in S_1 were due to errors in S_2 . While

this is made obvious in arid conditions by the suppressed soil moisture dynamics, it also occurred at the temperate sites (for example on 2 February in Figure 6.26 and 6.27). This issue could be possibly be addressed by parameterising the S_1 model error (in \mathbf{Q}) so that it is enhanced during precipitation events, since this would better account for the greater uncertainty in the model background S_1 after precipitation (leading to increased S_1 updates, and decreased S_2 updates). Note that this occurrence provides an additional argument for excluding S_3 from the state update vector; if the S_1 observation increments are not consistently representative of the errors in S_2 , then they certainly won't be representative of errors in the seasonal-scale S_3 (risking that incorrect updates will be made to S_3).

6.5.3 Net impact

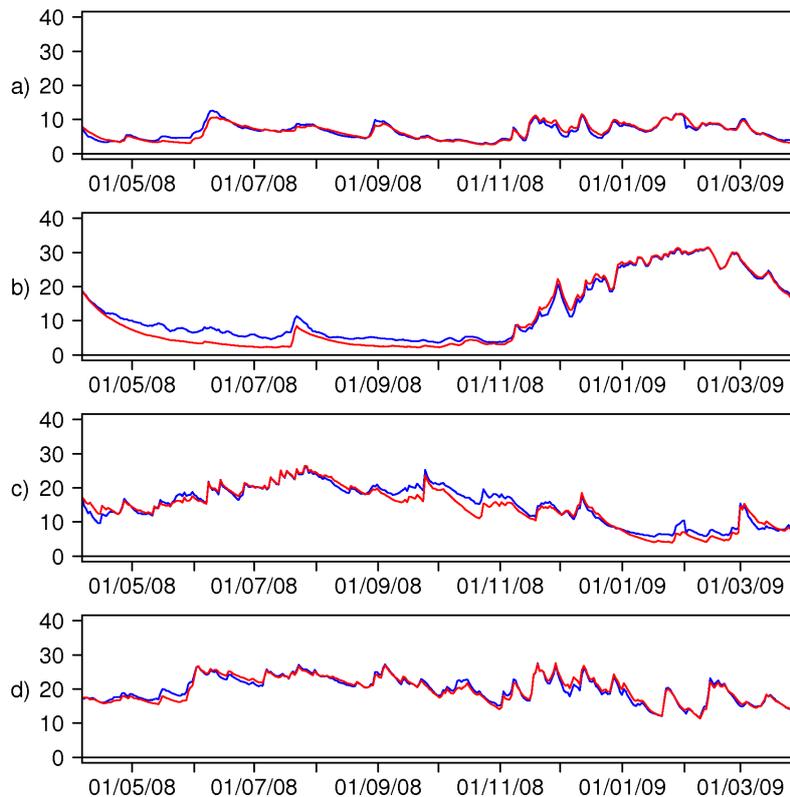


Figure 6.30: Time series of the mean near-surface soil moisture (% of saturation) from the EKF (blue) and the open-loop (red), averaged over the a) arid, b) tropical, c) temperate-oceanic, and d) temperate-Mediterranean climate zones.

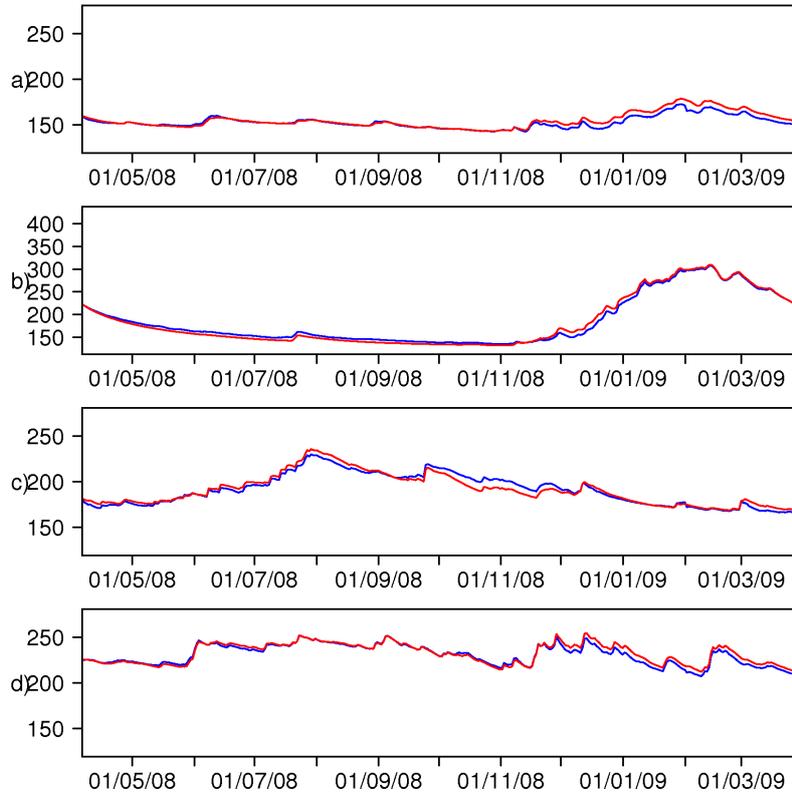


Figure 6.31: Time series of the mean root-zone soil moisture (% of saturation) from the EKF (blue) and the open-loop (red), averaged over the a) arid, b) tropical, c) temperate-oceanic, and d) temperate-Mediterranean climate zones. Note the different vertical axis for the tropical zone.

This section examines the net impact of the EKF assimilation of AMSR-E S_1 on the near-surface soil moisture, root-zone soil moisture, and evapotranspiration forecasts. The assimilation added a reasonable volume of moisture to the model surface, and the mean absolute daily increments were 0.45 mm day^{-1} for the arid zone, 0.72 mm day^{-1} for the tropical zone, and 0.48 and 0.50 mm day^{-1} , respectively, in the temperate-Mediterranean and oceanic zones. This is approximately half the mean daily precipitation in the arid zone (0.98 mm day^{-1}), and one quarter of the mean daily precipitation in the other climate zones (3.6 , 1.6 , and 2.1 mm day^{-1} , respectively).

For both the open-loop and EKF simulations, the mean S_1 , S_{RZ} , and daily evapotranspiration forecasts for each climate zone are listed in Table 6.10, together with the mean absolute differences between the two simulations. Figure 6.30 shows the time series of the mean S_1 forecasts for each climate zone. For S_1

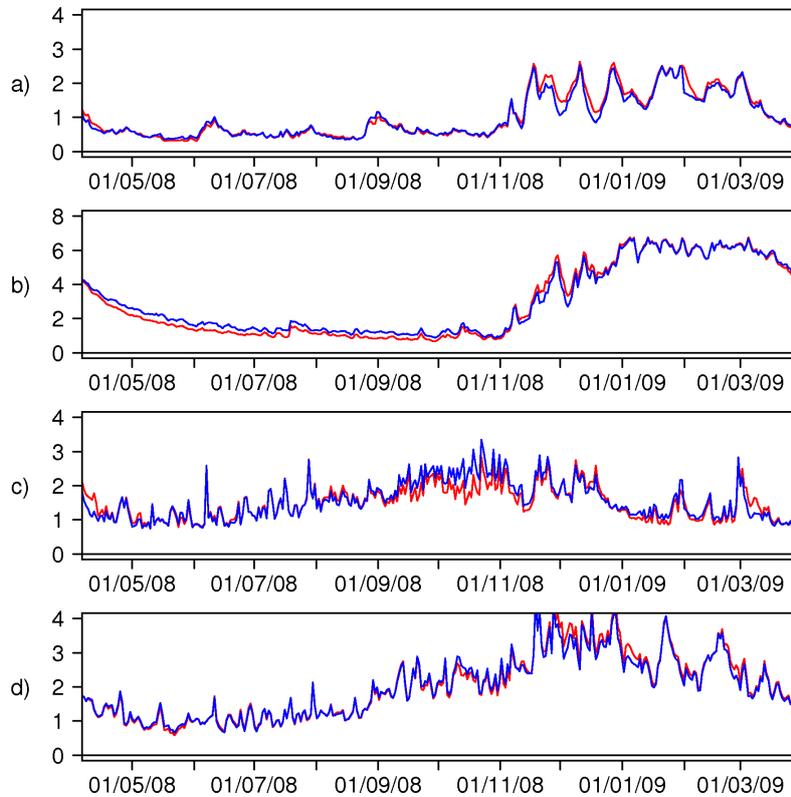


Figure 6.32: Time series of the mean daily evapotranspiration (mm day^{-1}) from the EKF (blue) and the open-loop (red) experiments, averaged over the a) arid, b) tropical, c) temperate-oceanic, and d) temperate-Mediterranean climate zones. Note the different vertical axis for the tropical time series.

there were extended periods during which the EKF differed from the open-loop by up to 5 mm, compared to mean S_1 values of the order of 10 mm in Table 6.10. The greatest S_1 divergence occurred during the dry season in the tropics and during October in the temperate-Mediterranean zone, corresponding to the greatest observation increments in Figure 6.20.

Figure 6.31 shows the equivalent time series for S_{RZ} . In general the divergence between the S_{RZ} time series lagged behind periods of sustained divergence in S_1 , due to the S_1 (and S_2) updates gradually dispersing through the soil root-zone. The greatest differences for S_{RZ} were close to 10 mm (compared to mean values of 150 mm in Table 6.10), and occurred during the wet season in the arid zone (which showed a small response to wet season precipitation), at the onset of the wet season in the tropics (before the differences were eliminated later in the wet season), and in November and January in the temperate-Mediterranean

and oceanic zones, respectively.

The equivalent time series of the mean daily evapotranspiration forecasts from the open-loop and EKF experiments are plotted in Figure 6.32. In each of the climate zones, the impact of the assimilation on the evapotranspiration was modest but significant, and the greatest differences between the open-loop and the EKF were close to 1 mm day^{-1} (compared to mean values of 1-4 mm day^{-1} in each zone). For each climate zone, the mean absolute difference in the evapotranspiration forecasts from the open-loop and the EKF over the year was approximately 10-20% of the open-loop values. In general, the divergence between the evapotranspiration time series (Figure 6.32) was more strongly correlated to the divergence between the S_1 time series (Figure 6.30) than the S_{RZ} time series (Figure 6.31). This suggests that the impact of the assimilation on the flux forecasts might not be primarily via transpiration.

In summary, while the EKF assimilation of AMSR-E near-surface soil moisture observations did not have a dramatic impact on the model soil moisture forecasts, the impact of the assimilation on the model was still substantial. In particular, the absolute impact on the mean daily evapotranspiration fluxes was around 10% of the open-loop forecasts. There were also several prolonged periods during which the assimilation had a much greater impact on the forecast evapotranspiration (for example in November 2008 in Figure 6.32), which would be expected to have a significant impact on screen-level forecasts from a coupled model.

6.5.4 Evaluation against in situ data

In this section the soil moisture forecasts generated by assimilating the AMSR-E observations are evaluated against the Murrumbidgee Monitoring Network in situ data, to determine whether the assimilation can improve the realism of the near-surface and root-zone soil moisture at these sites. Figures 6.33 to 6.39 compare time series of the near-surface and root-zone soil moisture from the open-loop and EKF (background 6 hour forecasts) to the in situ data at each of the Murrumbidgee sites. The observed precipitation and AMSR-E data show similar behaviour to the examples in Section 6.5.2. Again the precipitation time series show that ACCESS accurately forecast the occurrence of precipitation, although not necessarily its magnitude. Additionally, the AMSR-E data often detected these model precipitation errors: for example, at M7 assimilating

Table 6.10: Mean near-surface soil moisture (mm), root-zone soil moisture (mm), and daily evapotranspiration (mm day⁻¹), averaged over April 2008 to March 2009 for each climate zone, from the open-loop and the EKF.

	OPN	EKF	Mean absolute difference
Arid			
S_1	6.3	6.3	1.5
S_{RZ}	152.0	150.0	4.3
ET	1.03	0.99	0.21
Tropical			
S_1	11.6	12.7	2.3
S_{RZ}	181.8	182.9	7.3
ET	2.93	3.05	0.35
Temperate-Mediterranean			
S_1	14.2	14.7	2.3
S_{RZ}	189.5	189.3	7.7
ET	1.49	1.55	0.24
Temperate-oceanic			
S_1	19.8	19.9	1.9
S_{RZ}	231.9	230.4	7.5
ET	2.03	1.99	0.21

the AMSR-E data corrected both the over-forecast precipitation in November 2008 and the under-forecast precipitation event in February 2009. At each of the plotted Murrumbidgee locations, the EKF effectively drew S_1 towards the AMSR-E data, so that it is difficult to distinguish between the EKF and AMSR-E time series in Figures 6.33 to 6.39. At each site, assimilating the AMSR-E data also generated reasonable divergence between the EKF and open-loop S_{RZ} time series, including at the three western-most sites (M5, M7, and M8), which are classified as arid (likely because the soil moisture at these sites was no drier during these experiments than at the easterly sites, which are classified as temperate-oceanic).

Tables 6.11 and 6.12 compare the statistics of fit to the Murrumbidgee observations for the near-surface and root-zone soil moisture from the open-loop and the EKF, as well as for the AMSR-E data. In Chapter 4, the near-surface soil moisture from AMSR-E and MOSES had similarly good agreement to the in situ observations, and this result was repeated here. The AMSR-E r_{abs} (r_{anm}) ranged between 0.59 and 0.85 (0.52 and 0.75) with a mean of 0.75 (0.64), while

for the open-loop r_{abs} (r_{anm}) ranged between 0.64 and 0.86 (0.50 and 0.77), with a mean of 0.74 (0.59). In general assimilating the AMSR-E data improved the fit between the MOSES forecasts and the in situ near-surface soil moisture observations, and in Table 6.11 the EKF had better statistics than the open-loop at five out of the seven sites, with r_{abs} (r_{anm}) ranging between 0.65 and 0.86 (0.59 and 0.80) with a mean of 0.78 (0.71). In particular, the EKF had a better fit to the in situ data than either of the open-loop or the AMSR-E data at all of the Murrumbidgee sites for which multiple monitoring stations were used to estimate the in situ time series (M8, KA, and KB), and hence the in situ time series are more certain. Additionally, the two sites for which assimilating the AMSR-E data did not improve the model S_1 had some of the lowest correlations with the in situ data for both AMSR-E and MOSES. This hints that the poor statistics for the assimilation at these sites may have been due to problems with the in situ data. However, it is also possible that the lack of improvement was due to systematic errors in the model that could not be corrected by the assimilation (although this does not explain the low correlations for the AMSR-E observations at these sites).

For the root-zone soil moisture, reliable in situ observations were available at only five sites, since the observations at M1 and M5 had limited temporal coverage (see Figures 6.33 and 6.35). These are the same five sites for which the assimilation improved the model S_1 statistics in Table 6.11, and consistent with the S_1 results, the EKF S_{RZ} had a better fit to the in situ data than the open-loop at each of these sites in Table 6.12. The r_{abs} (r_{anm}) increased from a range of 0.34 to 0.87 (0.61 to 0.78), with a mean of 0.60 (0.67) for the open-loop, to a range of 0.37 to 0.92 (0.63 to 0.79), with a mean of 0.71 (0.72) for the EKF.

In summary, comparison to in situ soil moisture observations from the Murrumbidgee Monitoring Network showed that assimilating the AMSR-E near-surface soil moisture in general improved the MOSES soil moisture forecasts at those locations. The statistics of fit between the modeled and in situ soil moisture for the near-surface soil moisture were improved by the assimilation at most (five out of seven) sites, while the statistics for the root-zone soil moisture were improved at all locations (five) for which data were available. Additionally, the improvements in the correlations were substantial, and for the root-zone soil moisture the mean r_{abs} (r_{anm}) were increased from 0.63 (0.68) for the open-loop to 0.71 (0.72) for the EKF.

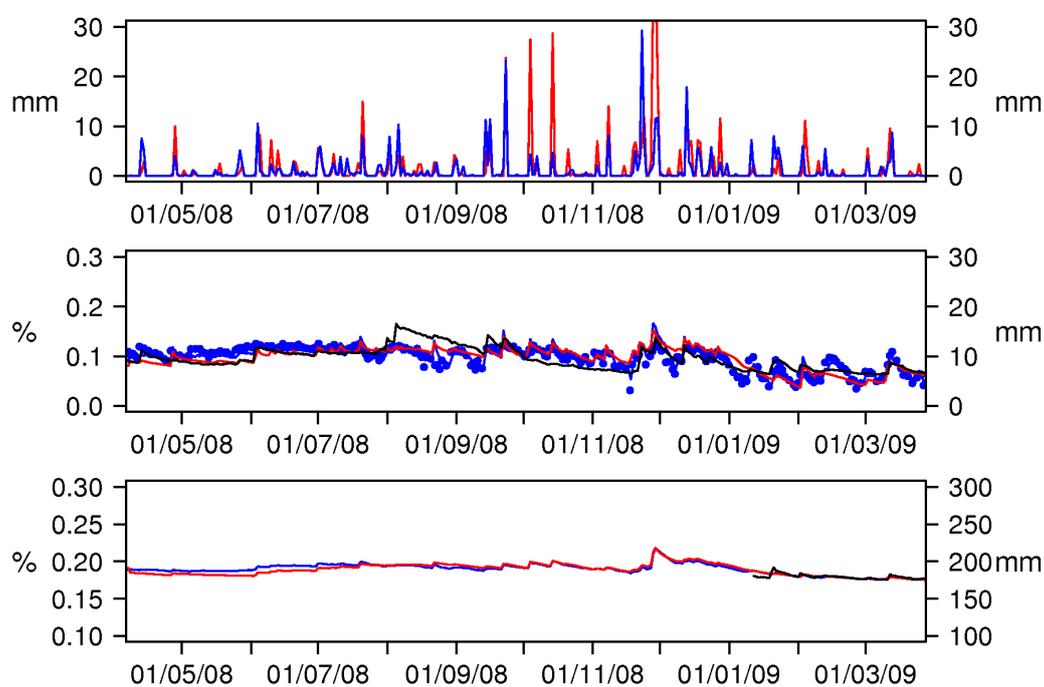


Figure 6.33: Time series of modeled and observed precipitation and soil moisture at M1. The upper panel shows precipitation from observations (blue) and ACCESS forecasts (red). The middle and lower panels show S_1 and S_{RZ} , respectively, from the EKF (blue), open-loop (red), and in situ observations (black), in units of % of saturation (left axis) and mm of soil moisture (right axis). The blue diamonds indicate the assimilated AMSR-E S_1 .

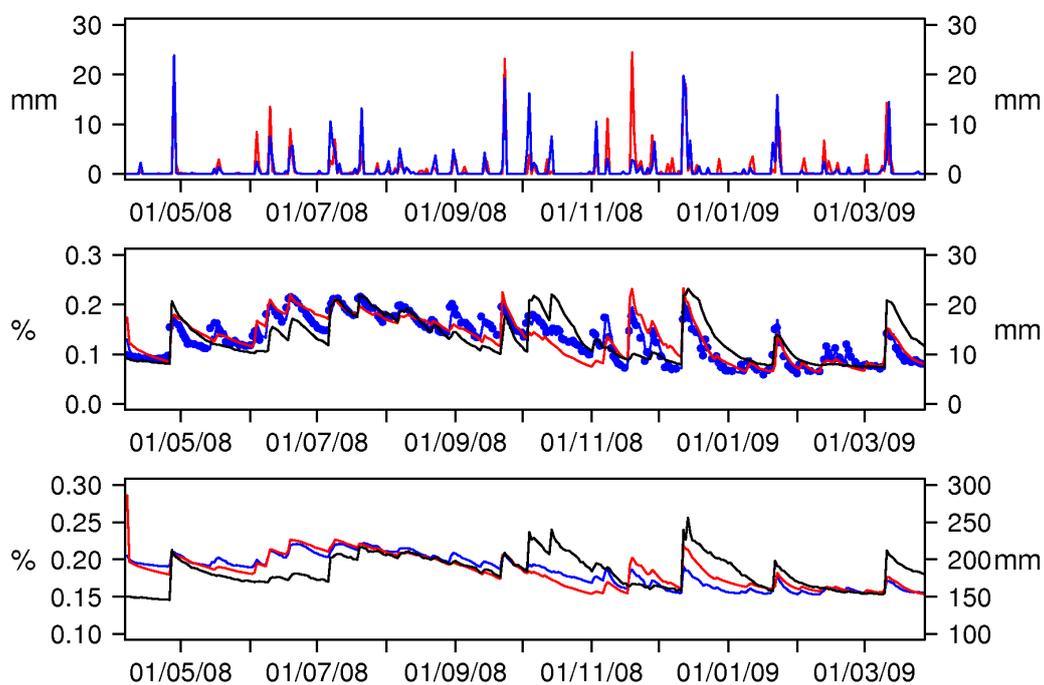


Figure 6.34: Time series of modeled and observed precipitation and soil moisture at M4. The upper panel shows precipitation from observations (blue) and ACCESS forecasts (red). The middle and lower panels shows S_1 and S_{RZ} , respectively, from the EKF (blue), open-loop (red), and in situ observations (black), in units of % of saturation (left axis) and mm of soil moisture (right axis). The blue diamonds indicate the assimilated AMSR-E S_1 .

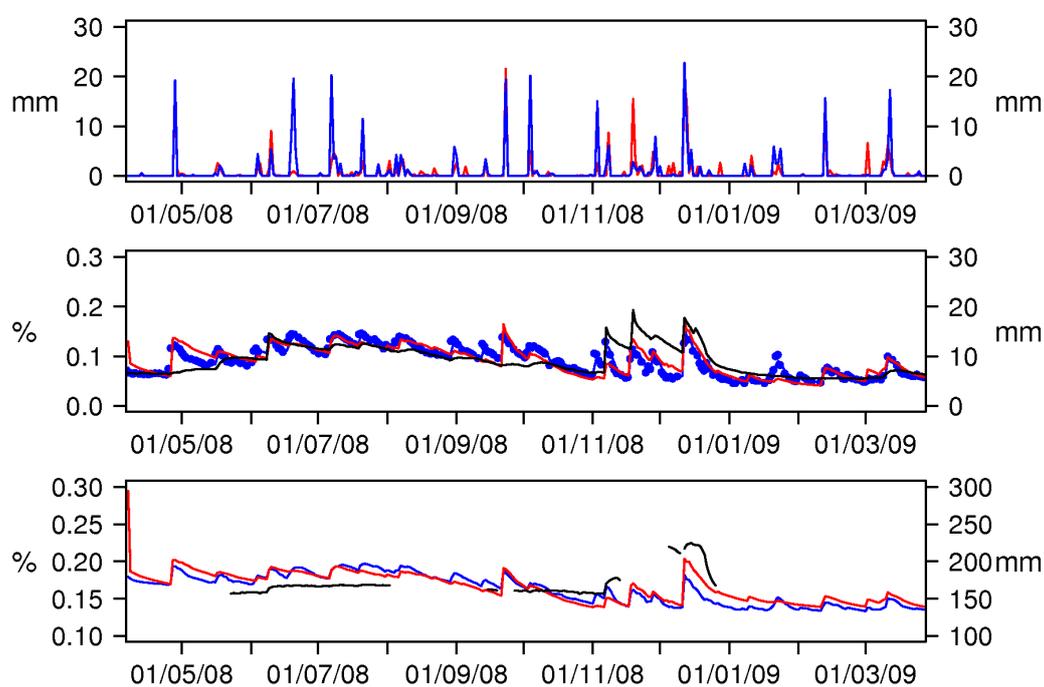


Figure 6.35: Time series of modeled and observed precipitation and soil moisture at M5. The upper panel shows precipitation from observations (blue) and ACCESS forecasts (red). The middle and lower panels show S_1 and S_{RZ} , respectively, from the EKF (blue), open-loop (red), and in situ observations (black), in units of % of saturation (left axis) and mm of soil moisture (right axis). The blue diamonds indicate the assimilated AMSR-E S_1 .

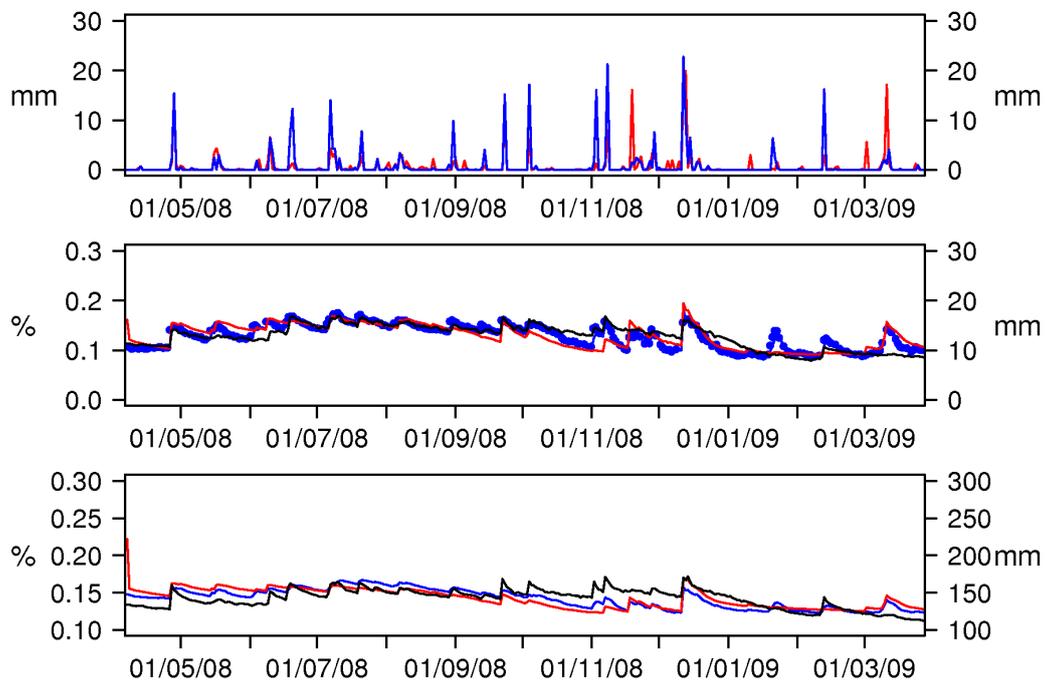


Figure 6.36: Time series of modeled and observed precipitation and soil moisture at M7. The upper panel shows precipitation from observations (blue) and ACCESS forecasts (red). The middle and lower panels show S_1 and S_{RZ} , respectively, from the EKF (blue), open-loop (red), and in situ observations (black), in units of % of saturation (left axis) and mm of soil moisture (right axis). The blue diamonds indicate the assimilated AMSR-E S_1 .

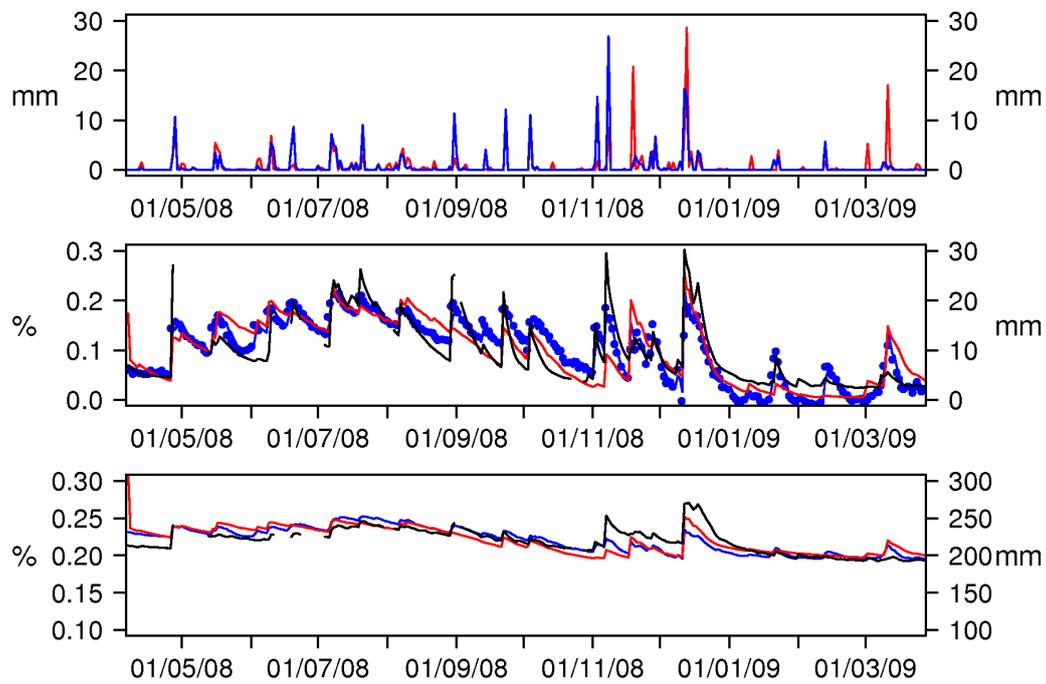


Figure 6.37: Time series of modeled and observed precipitation and soil moisture at M8. The upper panel shows precipitation from observations (blue) and ACCESS forecasts (red). The middle and lower panels show S_1 and S_{RZ} , respectively, from the EKF (blue), open-loop (red), and in situ observations (black), in units of % of saturation (left axis) and mm of soil moisture (right axis). The blue diamonds indicate the assimilated AMSR-E S_1 .

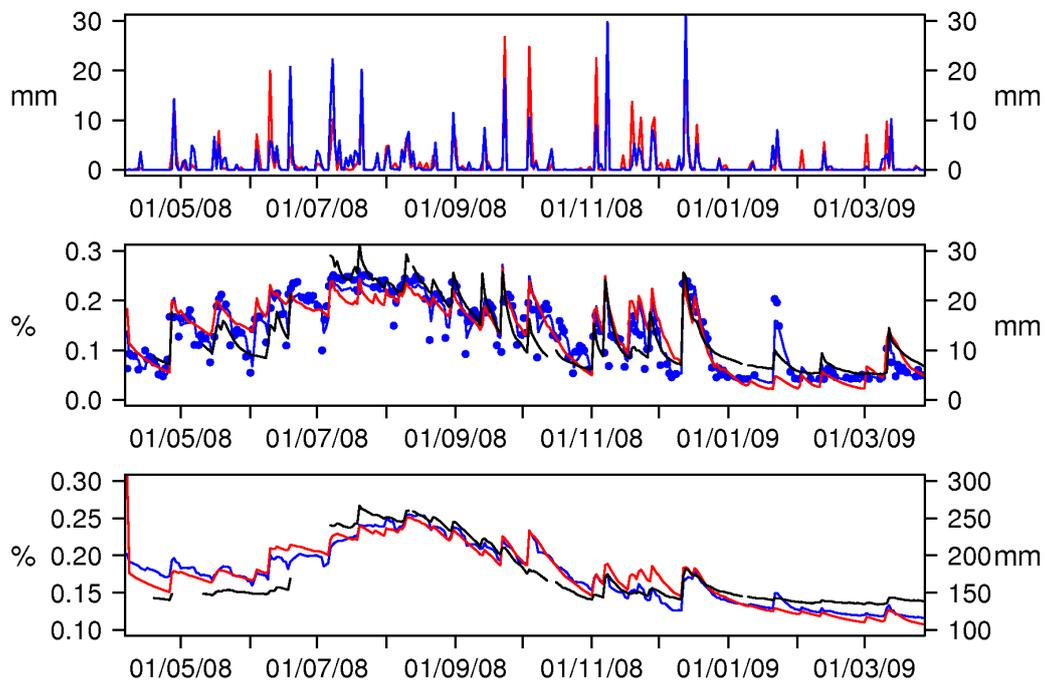


Figure 6.38: Time series of modeled and observed precipitation and soil moisture at KA. The upper panel shows precipitation from observations (blue) and ACCESS forecasts (red). The middle and lower panels shows S_1 and S_{RZ} , respectively, from the EKF (blue), open-loop (red), and in situ observations (black), in units of % of saturation (left axis) and mm of soil moisture (right axis). The blue diamonds indicate the assimilated AMSR-E S_1 .

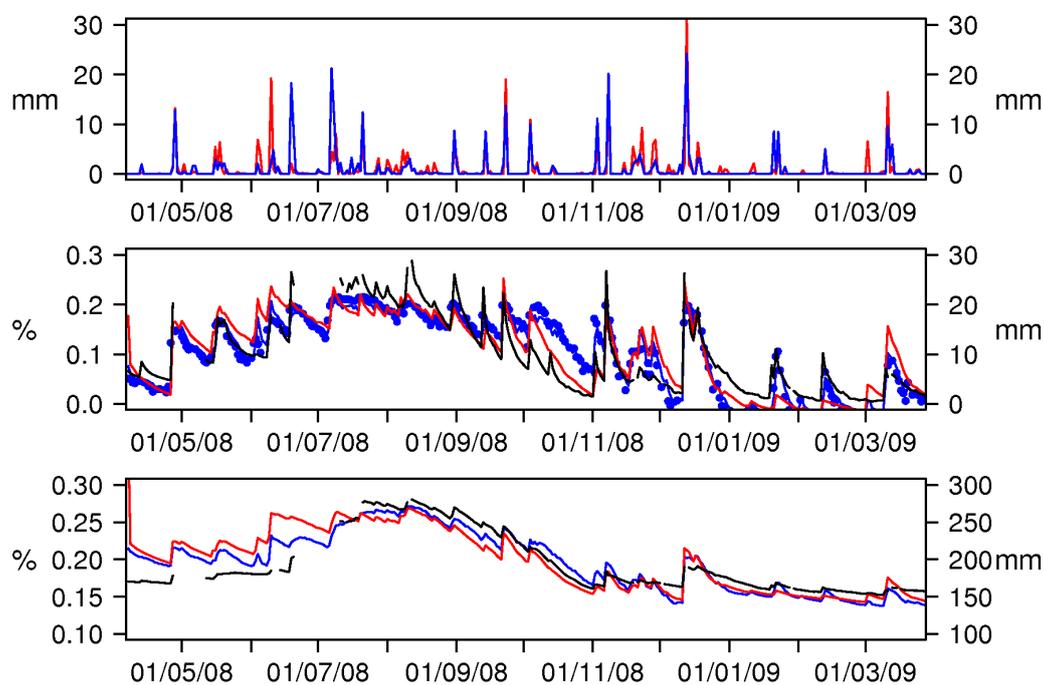


Figure 6.39: Time series of modeled and observed precipitation and soil moisture at KB. The upper panel shows precipitation from observations (blue) and ACCESS forecasts (red). The middle and lower panels show S_1 and S_{RZ} , respectively, from the EKF (blue), open-loop (red), and in situ observations (black), in units of % of saturation (left axis) and mm of soil moisture (right axis). The blue diamonds indicate the assimilated AMSR-E S_1 .

Table 6.11: Statistics of fit between the near-surface soil moisture in situ observations from the Murrumbidgee Monitoring Network, and the assimilated AMSR-E data, the open-loop, and the EKF. The RMSD is calculated from model data normalised to match the mean and variance on the in situ data. For each statistic the best result is indicated in bold, and all correlations are significant at 1%.

	AMSR-E			OPN			EKF		
	r_{abs}	r_{anm}	RMSD (m^3m^{-3})	r_{abs}	r_{anm}	RMSD (m^3m^{-3})	r_{abs}	r_{anm}	RMSD (m^3m^{-3})
M1	0.59	0.53	0.021	0.68	0.50	0.018	0.65	0.60	0.019
M4	0.71	0.73	0.032	0.64	0.71	0.036	0.72	0.76	0.032
M5	0.60	0.52	0.025	0.69	0.62	0.022	0.65	0.59	0.024
M7	0.83	0.70	0.014	0.72	0.64	0.018	0.83	0.74	0.014
M8	0.85	0.71	0.030	0.78	0.56	0.037	0.85	0.72	0.030
KA	0.82	0.52	0.041	0.83	0.76	0.040	0.88	0.75	0.033
KB	0.82	0.75	0.045	0.86	0.77	0.039	0.86	0.80	0.039

Table 6.12: Statistics of fit between the root-zone soil moisture in situ observations from the Murrumbidgee Monitoring Network, and the open-loop and EKF. The RMSD is calculated from model data normalised to match the mean and variance on the in situ data. For each statistic the best result is indicated in bold, and all correlations are significant at 1%.

	OPN			EKF		
	r_{abs}	r_{anm}	RMSD (m^3m^{-3})	r_{abs}	r_{anm}	RMSD (m^3m^{-3})
M4	0.36	0.69	0.024	0.37	0.71	0.024
M7	0.36	0.61	0.015	0.60	0.75	0.012
M8	0.70	0.78	0.013	0.73	0.79	0.012
KA	0.87	0.64	0.021	0.91	0.63	0.018
KB	0.84	0.64	0.023	0.92	0.70	0.016

6.5.5 Comparison to observed precipitation

It is extremely difficult to evaluate soil moisture forecasts at the continental scale. While the above comparison to in situ soil moisture observations is encouraging, it does not guarantee that the assimilation had a similarly positive impact at other sites. Since the atmospheric forcing is a major cause of error in NWP land surface states, the results of the assimilation could be assessed by comparison to soil moisture forecasts generated by forcing the land surface model with observations. Australia is not well covered by in situ observations, and the only atmospheric forcing variable for which high quality observations are available is precipitation. However, precipitation is thought to be the largest single source of soil moisture errors in NWP models. For example, van den Hurk et al. (2008) found that about half of the soil moisture increments added to the ECMWF model by the assimilation of screen-level observations at 36 locations throughout Europe were attributable to precipitation biases in the model. Consequently, replacing the NWP forecast precipitation used to force MOSES with observed precipitation is expected to significantly improve the accuracy of the model soil moisture, potentially providing a useful benchmark for testing the impact of assimilating the AMSR-E observations.

Consequently, the accuracy with which assimilating the AMSR-E data has corrected for errors in the precipitation forcing from the ACCESS model has been tested by comparing the assimilation results to a simulation of the MOSES model in which the precipitation forcing was adjusted to better match the BoM's rain gauge analyses (see Section 6.2.3). The BoM precipitation analysis is performed daily, whereas MOSES has been forced with hourly precipitation forecasts from ACCESS. Consequently, these hourly forecasts were multiplied by the ratio of the observed to forecast daily precipitation each day, to produce the "observed" precipitation forcing. This approach is consistent with the earlier finding that ACCESS can forecast the occurrence of precipitation reasonably well, although not necessarily the magnitude of each event. However, the long term mean of the observed forcing (1.28 mm day^{-1} , compared to 1.09 mm day^{-1} in the original forecasts) was slightly lower than the mean of the BoM's rain gauge analyses (1.33 mm day^{-1}), since the method used to adjust the precipitation forcing cannot account for observed precipitation events that were not originally forecast by ACCESS.

MOSES has been forced with the observed precipitation forcing, with all

Table 6.13: Statistics of fit between the MOSES_PRECIP root-zone soil moisture, and each of the open-loop and EKF simulations.

	Open-loop			EKF		
	r_{abs}	r_{anm}	RMSD (mm)	r_{abs}	r_{anm}	RMSD (mm)
M4	0.80	0.77	9.6	0.86	0.76	13.4
M7	0.80	0.80	11.0	0.92	0.82	6.1
M8	0.90	0.80	6.3	0.94	0.84	4.4
KA	0.95	0.76	13.3	0.95	0.70	21.3
KB	0.90	0.76	10.7	0.93	0.74	8.6

other forcing, initial conditions, and model settings identical to the open-loop; this simulation is referred to as the MOSES_PRECIP simulation. Ideally, assimilating the AMSR-E observations should correct the MOSES soil moisture for errors in the ACCESS precipitation, drawing the EKF soil moisture towards the MOSES_PRECIP values. To test this approach, the results are first considered at the Murrumbidgee Monitoring Network sites, where comparison to the in situ observations indicated that the assimilation improved the model soil moisture (see Section 6.5.4). Table 6.13 shows the statistics of fit between the root-zone soil moisture from MOSES_PRECIP, and each of the open-loop and EKF experiments, at the same locations that were compared to the in situ data in Table 6.12. The comparison to MOSES_PRECIP was inconclusive at these locations: at three out of five sites the EKF had a better fit to MOSES_PRECIP than the open-loop. The average r_{anm} and RMSD across the five sites was decreased by the assimilation, while the mean r_{abs} was increased. This lack of consistency between the results obtained by assessing the EKF S_{RZ} against the in situ observations and the MOSES_PRECIP simulations could be due to errors (including representativity errors) in the in situ observations. Alternatively, it could also be due to the EKF (and the in situ data) reflecting soil moisture errors that were not caused by precipitation.

Consistent with the results at the Murrumbidgee sites, assimilating the AMSR-E observations did not in general bring the MOSES forecast soil moisture closer to the MOSES_PRECIP simulations across Australia. Figure 6.40 shows the time series of the root-zone soil moisture averaged over each climate zone from the open-loop, the EKF, and MOSES_PRECIP: with a few exceptions (such as at the onset of the wet season in the tropics), assimilating the

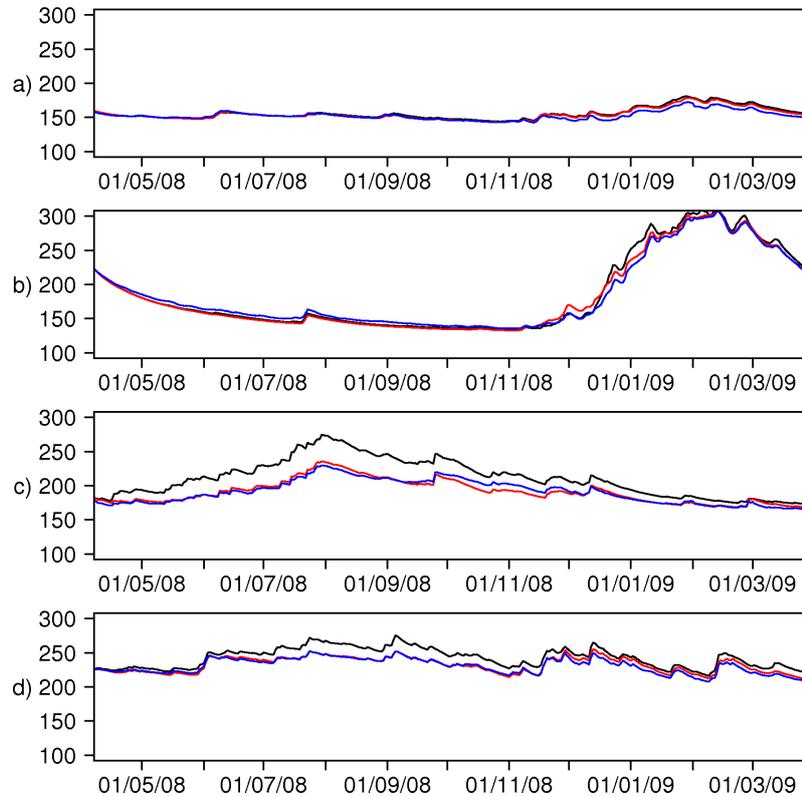


Figure 6.40: Time series of the root-zone soil moisture (mm) averaged over the a) arid, b) tropical, c) temperate-oceanic, and d) temperate-Mediterranean climate zones, from the EKF (blue), open-loop (red), and MOSES_PRECIP (black) simulations.

AMSR-E data did not reduce the S_{RZ} bias relative to MOSES_PRECIP. This result is not surprising, given that the difference between the AMSR-E and open-loop soil moisture in Figure 6.20 does not reflect the seasonal-scale errors in the forecast precipitation in Figure 6.21. Additionally, the statistics of fit to the MOSES_PRECIP root-zone soil moisture were slightly degraded by the assimilation. Averaged across Australia the absolute (anomaly) correlation was improved by the assimilation at 42% (41%) of the grid cells, while the mean correlation across the domain was reduced from 0.78 (0.64) for the open-loop to 0.76 (0.63) for the EKF. Additionally, the assimilation reduced the RMSD at 38% of the grid cells, with the mean RMSD increasing from 14.7 mm for the open-loop to 16.3 mm for the EKF. The tendency for the statistics of fit to be degraded by the assimilation was consistent across the model domain, and in nearly all instances the individual statistics for each climate zone were

degraded by the assimilation (not shown). Given the inconclusive results at the Murrumbidgee sites in Table 6.13, it is unclear whether these findings are due to the assimilation having not improved the model soil moisture, or due to short-comings of the strategy of assessing the EKF against MOSES_PRECIP (most notably, that the assimilation may be responding to soil moisture errors not associated with precipitation).

6.6 Chapter summary

In this chapter remotely sensed near-surface soil moisture observations have been assimilated into the Australian BoM's MOSES model. Near-surface soil moisture observations from AMSR-E have been assimilated over a one year time period to determine whether the AMSR-E data can improve the realism of the modeled near-surface and root-zone soil moisture. The same EKF assimilation strategy that was used to assimilate near-surface soil moisture and screen-level observations into Météo-France's ISBA model in Chapter 5 has been used.

Examining the model Jacobians for MOSES highlights that it has very different soil moisture physics to ISBA, leading to differences in the way in which the EKF updates the root-zone soil moisture from the near-surface observations. In MOSES the vertical flow of moisture through the soil is controlled by explicitly parameterised Darcian flow, giving a more direct relationship between the near-surface and root-zone soil moisture than in ISBA. However, for these experiments this relationship was very weak, due in part to the prevalence of dry surface conditions across most of Australia. Consequently, assimilating the near-surface soil moisture observations made significant updates only to the near-surface soil moisture layer (and also to the second layer in more humid conditions). Any impact on the deeper soil layers was limited to changes that diffuse through the soil profile from the near-surface layers.

A series of synthetic experiments showed that the EKF assimilation of near-surface soil moisture (S_1) observations could reduce the errors throughout the model soil moisture profile in temperate regions only, since in arid and tropical regions S_1 is only very weakly coupled to the underlying soil moisture. For these experiments, the mean error in the root-zone soil moisture across Australia's temperate zones were reduced by 30-40% of their initial value, after three months of assimilating synthetic S_1 data. However, for NWP the soil moisture errors are more likely to be associated with incorrect atmospheric forc-

ing, rather than gross errors in the initial conditions. By making corrections to the near-surface layers, the EKF should be sufficient to correct the soil moisture profile for forcing errors detected by the observations, in particular for errors entering the soil profile through S_1 , such as precipitation.

Even though the AMSR-E assimilation qualitatively appears to have corrected some of the errors in the NWP forecast precipitation used in these experiments, this could not be confirmed quantitatively. This could be due to the AMSR-E observations having not accurately detected soil moisture errors associated with incorrect precipitation, or it could be due to short-comings in the evaluation strategies that were applied. Supporting the latter possibility assimilating the AMSR-E observations did improve the fit between the model root-zone soil moisture and the in situ observations at the five Murrumbidgee Monitoring Network sites for which the in situ observations were available (while the assessment based on the precipitation observations was inconclusive at these sites). Additionally, the improvements against the in situ observations were substantial, with the mean absolute (anomaly) correlation to the in situ data being increased from 0.60 (0.67) for the open-loop to 0.71 (0.72) for the EKF.

Looking forward, these results suggest that for the MOSES model the assimilation of near-surface soil moisture could compliment the assimilation of screen-level observations very well. Assimilating remotely sensed near-surface soil moisture data has the potential to improve the model soil moisture forecasts, while the screen-level observations could more tightly constrain the nonsurface soil moisture layers according the errors in the flux forecasts.

Chapter 7

Conclusions and Future Work

This thesis has demonstrated the potential for NWP models to benefit from the assimilation of remotely sensed near-surface soil moisture data. While the ultimate aim of assimilating the soil moisture observations is to improve low-level atmospheric forecasts, the focus here is on the intermediate step of improving the model root-zone soil moisture. It was first established that realistic near-surface soil moisture observations can be retrieved over Australia from existing remote sensing satellites, specifically from the C-band passive microwave AMSR-E instrument. Following this, near-surface soil moisture observations from AMSR-E were assimilated into the NWP land surface models used at the Australian Bureau of Meteorology (BoM) and Météo-France. The AMSR-E soil moisture observations were assimilated into the BoM's land surface model over Australia for a one year period. At those locations where in situ soil moisture observations were available, the assimilation improved the fit between the in situ and modeled soil moisture, in both the near-surface and root-zone soil layers. This demonstrates that assimilating near-surface soil moisture observations has the potential to improve modeled root-zone soil moisture. Additionally, since the use of screen-level observations is known to improve low-level atmospheric forecasts, the assimilation of AMSR-E soil moisture was compared to the assimilation of screen-level observations, using Météo-France's land surface model. Over the one month experimental period, assimilating each data type separately generated contrasting root-zone soil moisture increments, while assimilating them together very slightly improved the model fit to both data sets. This supports the notion that assimilating the near-surface soil moisture and screen-level observations together can benefit NWP by improving the realism of the model soil moisture, while maintaining the low-level atmospheric forecasts

skill.

7.1 Chapter conclusions

The research for this thesis consisted of three separate investigations, reported in Chapters 4 to 6. The main findings, as well as the limitations and implications arising from the work in each of these chapters, are reviewed below.

7.1.1 Remotely sensed soil moisture over Australia

In Chapter 4 the available remotely sensed near-surface soil moisture data sets over Australia were assessed, to select the most suitable product for assimilation into land surface models, and to better characterise its accuracy.

Selecting a remotely sensed near-surface soil moisture data set

At the time of this study, the only instrument providing remotely sensed soil moisture observations with sufficient coverage for use in Australian NWP was the passive microwave AMSR-E. Soil moisture data sets obtained from AMSR-E brightness temperatures using the four most prominent retrieval algorithms (from JAXA (Koike et al., 2004), NASA (Njoku et al., 2003), USDA (Jackson, 1993), and VUA-NASA (Owe et al., 2001)) were compared, to select the most accurate. Due to widespread RFI in the C-band AMSR-E brightness temperatures (particularly over north America and east Asia), the NASA, USDA, and JAXA algorithms are applied to X-band brightness temperatures only, while the VUA-NASA algorithm is separately applied to C- and X-band brightness temperatures. Since RFI is rare in Australia, both the C- and X-band VUA-NASA products were used here.

The five remotely sensed near-surface soil moisture data sets identified above were compared to in situ soil moisture observations from 12 locations from the Murrumbidgee (ten sites) and Goulburn (two sites) Monitoring Networks, over 2006. Based on this comparison, the VUA-NASA soil moisture was identified as the most realistic remotely sensed soil moisture data set. For both the C- and X-band wavelengths, the VUA-NASA soil moisture product consistently had strong temporal correlations to the in situ soil moisture, although results from the Goulburn sites were not as good as those from the Murrumbidgee sites. Consequently the VUA-NASA soil moisture retrieval algorithm is recommended

over the other algorithms considered here, and the VUA-NASA soil moisture data set was used in the assimilation experiments in Chapters 5 and 6.

Assessing the VUA-NASA near-surface soil moisture data

The VUA-NASA soil moisture observations were examined in more detail, principally by comparison to six years of in situ soil moisture observations from the Murrumbidgee Monitoring Network, to better characterise its accuracy and to establish the best use of the data. The longer wavelength C-band brightness temperatures were expected to yield more accurate soil moisture observations than the X-band brightness temperatures, due to increased transparency of vegetation and deeper penetration of the soil at longer wavelengths. However, neither waveband generated soil moisture with a consistently better fit to the in situ observations. Since they are theoretically favoured, the C-band observations have been used in this thesis, and are recommended where RFI is not problematic. The good performance of the VUA-NASA X-band soil moisture data set indicates that the poor performance of the other retrieval algorithms was not due to the use of X-band brightness temperatures. Likewise, the (nighttime) descending AMSR-E overpass was expected to produce better soil moisture estimates than the (daytime) ascending overpass, due to the greater homogeneity of the soil conditions at night. Again, neither yielded a superior fit to the in situ soil moisture observations, and the theoretically favoured descending overpass AMSR-E soil moisture observations are recommended in preference to the ascending overpass. However, if more frequent data are preferred (such as in Chapter 5) these results suggest the ascending overpass AMSR-E soil moisture could also be used with confidence.

Over the six year period the VUA-NASA AMSR-E soil moisture detected the temporal variability of the Murrumbidgee in situ soil moisture observations very well. For the C-band descending AMSR-E overpass, the absolute (anomaly) correlation between the remotely sensed and in situ observations ranged from 0.56 to 0.82 (0.31 to 0.65) across the Murrumbidgee sites, while the normalised RMSD was consistently less than the standard deviation of the in situ time series. Additionally, maps of the VUA-NASA soil moisture anomalies over Australia showed a good correspondence to precipitation at the daily time scale, while maps of monthly mean soil moisture showed the expected correspondence to observed precipitation and vegetation growth.

Limitations of using in situ soil moisture data

The above evaluation of remotely sensed soil moisture was largely based on in situ soil moisture observations, and is limited by the shortcomings of those data. Most notably, in situ observations are available at only a few locations, and results from these locations can not necessarily be extrapolated to other sites (as was highlighted by the differing results from the Murrumbidgee and Goulburn Monitoring Networks over 2006). Additionally, there are systematic differences between the soil moisture quantities observed by in situ and remote sensors, due to differences in the horizontal and vertical scales observed by each. Horizontally, remote sensors observe an area-average soil moisture (45-75 km² for C-band AMSRE), while in situ sensors observe at a single point, and are sensitive to small scale variability in the land surface characteristics that is averaged out at the scale of a remotely sensed pixel. To quantify the impact of this small scale heterogeneity on soil moisture, the in situ soil moisture timeseries from five soil moisture sensors within a single AMSR-E pixel were compared. The absolute (anomaly) correlation between the near-surface soil moisture observations from the five stations ranged from 0.71 to 0.97 (0.69 to 0.92). These statistics are similar to the best results obtained for AMSR-E, and it is possible that the above comparison of the different AMSR-E wavelengths and overpasses did not yield the expected differences in skill due to the limited accuracy with which the in situ observations represent the area-average soil moisture observed by AMSR-E.

Given the above shortcomings of in situ soil moisture observations, there is a clear need for more robust methods to assess large scale (modeled and remotely sensed) soil moisture. One promising option is to use related variables that are observed at the continental scale, such as precipitation. However it is not clear how meaningful quantitative results can be generated from such a comparison. While the AMSR-E soil moisture data were qualitatively compared to maps of precipitation and vegetation in this study, the strongest conclusion that could be drawn from this was that the remotely sensed soil moisture have no obvious errors. A more likely approach to achieving an objective and quantitative validation would be to demonstrate that assimilating remotely sensed soil moisture data into a models can improve forecasts of soil moisture-dependent variables, such as river discharge.

7.1.2 Assimilation experiments with ISBA

Chapter 5 presented the assimilation of the AMSR-E near-surface soil moisture observations into Météo-France’s ISBA land surface model over the European domain. This was then compared to the assimilation of screen-level temperature and relative humidity, which are currently used to initialise the soil moisture in Météo-France’s NWP suite (and most other NWP models). A series of experiments was conducted to determine whether near-surface soil moisture observations might be used either in place of, or as a complement to, the screen-level observations.

The land surface assimilation scheme

The assimilation experiments used an off-line EKF-based land surface analysis scheme. To allow the flexibility to assimilate both remotely sensed land surface observations and screen-level variables, the EKF used a short-range forecast of the land surface model as the observation operator. Additionally, the EKF is made affordable within operational NWP by using an off-line version of the land surface model, forced with short-range forecasts from the atmospheric model. This off-line assimilation is intended to be semicoupled to the NWP model, in that the land surface state in the NWP model is updated with each land surface analysis, before the forcing for the subsequent off-line assimilation cycle is generated. However, in this thesis the land surface assimilation was completely decoupled from the atmospheric model, and the forcing was generated once at the beginning of each experiment. Finally, to allow the off-line assimilation of screen-level observations, the atmospheric forcing is applied higher than is standard for a land surface model, at the height of the first atmospheric model layer.

The EKF land surface analysis scheme was based on a simplified EKF used previously to assimilate screen-level observations into ISBA (Balsamo et al., 2007; Mahfouf et al., 2009). The simplified EKF differs from the EKF in that the background error covariances are not evolved through each model forecast cycle, and instead the same background error covariance matrix is used at the start of each assimilation cycle. The full EKF was used here since it was hoped that the long term evolution of covariance between the near-surface and underlying soil moisture errors would enhance the analysis updates to the nonsurface soil moisture layers. The EKF is also better suited to the irregular availability of

remotely sensed observations, since the simplified EKF relies on the assumption that the reduction in the model errors generated by each analysis is exactly balanced by the increase in the errors during the subsequent forecast cycle.

Comparison of the EKF and simplified EKF assimilation of AMSR-E soil moisture into ISBA revealed that while the Kalman gains for each differed, this had little impact on the resulting root-zone soil moisture analyses (at the broad spatial and temporal scales that were considered). This is principally a consequence of the use of a 1-D assimilation, since with no horizontal aspect to the assimilation the temporal evolution of \mathbf{P} can only affect the vertical profile of the analysis increments generated from a given observation increment. Despite this result, the EKF was used throughout this thesis since it incurs no additional computational cost (once a small approximation is introduced to the estimation of the linear tangent model), while allowing the possibility of specifying \mathbf{P} more realistically through the use of the forecast model (and potentially through the parameterisation of \mathbf{Q}).

Assimilating near-surface soil moisture into ISBA

This is the first study to assimilate near-surface soil moisture into ISBA at large spatial scales, and the results revealed that ISBA is not well suited to near-surface soil moisture assimilation. ISBA is a two layer force-restore model, with soil moisture defined by a near-surface soil moisture (spanning the depth of bare soil evaporation) and a root-zone soil moisture (spanning the depth of evapotranspiration). In ISBA the near-surface soil moisture does not influence forecasts of other variables, and consequently the root-zone soil moisture can only be constrained from near-surface observations by inverting the influence of the root-zone soil moisture on near-surface soil moisture forecasts. However, the strongest influence of the root-zone soil moisture on the near-surface soil moisture is via an indirect link through the impact of transpiration on the surface temperature. This is most likely caused by the use of a single surface temperature for both soil and vegetation, rather than a true physical relationship. A major motivation for assimilating near-surface soil moisture was the expectation that it would provide a more direct observation of the profile soil moisture, and yet the above described mechanism is no more direct than the influence of the root-zone soil moisture on the screen-level observations (which is also driven by evapotranspiration). This spurious link between near-surface and root-zone soil moisture could be avoided by assimilating the near-surface

soil moisture only during the night time, when the root-zone soil moisture influences the near-surface soil moisture via the model restore term (representing capillary and gravitational drainage).

The above findings are specific to the two layer force-restore physics in ISBA. As demonstrated with the Australian MOSES model in Chapter 6, multi-layer models with explicit diffusive flow between soil layers are better suited to near-surface soil moisture assimilation. Consequently, it is recommended that the near-surface soil moisture assimilation be ported into the new diffusive multi-layer version of ISBA (Boone et al., 2000) once this model has been coupled to ALADIN.

The systematic differences between modeled and observed soil moisture (see Reichle et al., 2004b), present a major obstacle to assimilating observed soil moisture into land surface models. The most common strategy used to address these differences is to normalise the observations to match the model soil moisture climatology prior to assimilating the data. In this thesis the normalisation was achieved using the CDF-matching strategy of Reichle and Koster (2004). However, for AMSR-E and ISBA the CDF-matching was hampered by the limited data record available to sample the model soil moisture climatology (one year), as well as the substantial seasonal cycle in the bias between the modeled and observed soil moisture (since statistical rescaling techniques assume the biases are stationary). Consequently a conservative approach was adopted here, by assuming that the model seasonal cycle was correct, and then adjusting the observations to match this seasonal cycle before applying the CDF-matching. To highlight the importance of applying an appropriate bias correction strategy, the AMSR-E observations were also assimilated without the seasonal bias correction (but with CDF-matching). The impact of excluding the seasonal bias correction on the soil analyses was profound, and much greater than the selection of the error covariance evolution approach (i.e., the use of the simplified or full EKF).

Assimilating screen-level observations into ISBA

In addition to the above short-comings of assimilating near-surface soil moisture into ISBA, significant problems were also encountered when assimilating the screen-level observations. Most notably, there was a persistent diurnal cycle in the model-observation bias for the screen-level variables, which cannot be easily attributed to model soil moisture errors. Consequently, the root-zone soil

moisture increments generated by the screen-level assimilation had a substantial diurnal cycle, and yet the assimilation did not amend the diurnal cycle in the forecast screen-level biases (supporting the assertion that the diurnal cycle was not caused by the model soil moisture). This demonstrates how assimilating screen-level observations can adjust the land surface states to compensate for errors unrelated to the surface, which is one of the main motivations for exploring the use of near-surface soil moisture observations.

Extending the assimilation window to 24 hours (as was done by Balsamo et al. (2007)) may allow the EKF to better synthesise these diurnal biases. Alternatively, the screen-level observations could be bias corrected to match the model climatology, as is typically done for soil moisture observations. However, it is not clear how such a scheme could be implemented in an operational NWP land-surface analysis, and nor is it clear that this would be advantageous to NWP. While allowing the assimilation to correct the root-zone soil moisture for systematic errors in the screen-level forecasts clearly violates the assumptions of the EKF, from a purely pragmatic perspective it ensures better low-level forecasts (although in this example, the assimilation generated only very modest improvements). It is not yet clear how screen-level observation-model biases should be handled within NWP, although ultimately the best option would be to identify the cause of the bias, and address any associated model problems. The assimilation of remotely sensed land-surface observations, such as near-surface soil moisture, will likely assist in the identification of such errors.

Assimilating near-surface soil moisture and screen-level observations

When the AMSR-E near-surface soil moisture and screen-level observations were assimilated separately, there was no clear consistency between the resulting soil moisture analyses increments generated by each. Consequently, in these experiments the screen-level observations could not have been substituted with the AMSR-E near-surface soil moisture to achieve the same corrections to the boundary layer forecasts. If this is extrapolated to regions with scarce screen-level observations it implies that remotely sensed soil moisture may not be immediately useful for Météo-France's NWP model. The lack of consistency could have been caused by systematic errors in the model forecasts of the observed variables (as in other studies (Douville et al., 2000; Hess, 2001; Drusch and Viterbo, 2007)), or by errors in the observations (since both data sets have considerable uncertainties). Validating the assimilation results against indepen-

dent data (specifically from flux towers and in situ soil moisture sensors) would help to resolve this issue, although such a validation would require a longer experimental period than was possible for this study (and would likely benefit from coupling the off-line land data assimilation to the NWP model).

When the near-surface soil moisture and screen-level observations were assimilated together, the sensitivity of the near-surface soil moisture forecasts to the model root-zone soil moisture was much less than the sensitivity of the screen-level observations. Consequently, the AMSR-E observations contributed only a small fraction (2%) of the information content gained from the assimilation, and the root-zone soil moisture analyses retained a strong similarity to the analyses generated by assimilating the screen-level observations alone. However, introducing the AMSR-E observations still had a significant impact on the soil moisture analyses, since the AMSR-E observation increments were more consistent over time (in contrast to the screen-level observations increments, which were dominated by the diurnal cycle).

Despite the conflicting information from the two observation data sets, when the AMSR-E and screen-level observations were assimilated together, the EKF was able to slightly improve the fit between the model forecasts and both observed data sets. However, these improvements were extremely modest, and the RMSD between the ISBA forecasts and the assimilated observations was reduced by less than 5% of the open loop values, for all assimilated variables. If this result can be substantiated with larger improvements once the diurnal cycle in the screen-level observation increments has been addressed, and by coupling the land data assimilation system to the NWP model, this would confirm that assimilating remotely sensed soil moisture together with screen-level observations has the potential to improve the realism of the NWP land surface without degrading the low-level atmospheric forecasts.

7.1.3 Assimilation experiments with MOSES

In Chapter 6 the root-zone soil moisture in the Australian BoM's NWP land surface model (MOSES) was analysed by assimilating the AMSR-E near-surface soil moisture observations, for a one year period from April 2008. The results were then compared to in situ soil moisture observations from the Murrumbidgee Soil Moisture Monitoring Network, and to observed precipitation across Australia, to determine whether the assimilation improved the model soil moisture.

Assimilating near-surface soil moisture into MOSES

The soil moisture physics in MOSES are very different to those in ISBA. MOSES describes soil moisture in four layers (three of which are in the root-zone), with Darcian flow between layers. This gives a more direct relationship between the soil moisture in the near-surface and underlying soil layers than occurs in ISBA. However, vertical soil moisture information is exchanged between layers very slowly in MOSES, particularly for arid conditions, since hydraulic conductivity is low in dry soils. Consequently, across most of Australia MOSES does not evolve significant vertical correlations between the background soil moisture errors in each model layer (i.e., off-diagonal terms in \mathbf{P}), and the observation operator does not translate the near-surface observations into updates of soil moisture in the underlying layers either. Consequently, the EKF in general updated only the near-surface soil moisture layer from the AMSR-E observations, and the impact of the assimilation on the underlying soil moisture was limited to changes that diffused down from the near-surface updates. The exception is in humid conditions, specifically during winter in the temperate south and during the wet season in tropical north Australia, when the coupling between the soil moisture in the first and second layers of MOSES is stronger, resulting in the EKF updates to the second soil layer also.

Assimilating synthetic observations

Given the limited vertical extent of the analysis updates from the near-surface soil moisture assimilation, synthetic experiments were used to explore how effectively the soil moisture profile in MOSES can be constrained from near-surface soil moisture observations. These experiments revealed that assimilating near-surface soil moisture observations is ineffective in arid conditions and in the tropical wet season. In arid conditions the near-surface soil moisture becomes decoupled from the underlying soil moisture profile, so that the soil moisture in the nonsurface layers cannot be corrected from near-surface observations (both because corrections to the surface layer do not infiltrate into the underlying layers, and because the surface layer does not contain a signal of the underlying model errors). However, the errors in the model soil moisture are likely very low in arid conditions, and this result may be of little consequence. At the opposite extreme, during the tropical wet season soil moisture is controlled by intense precipitation, and has no memory of the previous soil moisture states,

or soil moisture updates. Consequently, assimilating near-surface soil moisture observations only affected the model soil moisture briefly at the onset of the wet season. The ineffectiveness of the assimilation in arid and tropical regions will occur regardless of the assimilation strategy used, since in both instances it arises from the near-surface soil moisture being decoupled from the (historical or forecast) root-zone soil moisture.

In the temperature regions of Australia, assimilating the (synthetic) near-surface soil moisture observations was more effective. At these locations the EKF updated the model soil moisture to reduce the errors arising from the incorrect initial conditions prescribed in the synthetic experiments. For example, after three months of assimilating the synthetic near-surface soil moisture observations, the EKF reduced the error in the root-zone soil moisture by 30-40% of the error in the openloop simulations initialised with the same incorrect values.

Assimilating AMSR-E observations

The AMSR-E observations were bias corrected prior to assimilation into MOSES using the same CDF-matching strategy that was used for ISBA. The problems encountered with ISBA, associated with the nonstationary model-observation bias, did not generally occur for MOSES. The exception was the tropics, where soil moisture experiences two very distinct climate regimes, associated with the dry and wet seasons, and a single CDF-matching operator was unable to correct the biases in both seasons. If a sufficient observation record were available, a better approach would be to calculate separate CDF-matching operators for the dry and wet seasons. Note that correcting the seasonal cycle before applying the CDF-matching operator (as was done for ISBA) would not be effective in this instance, since the (precipitation-driven) variance also differs between the dry and wet seasons.

While the impact of assimilating the AMSR-E observations was modest, due to the limited vertical extent of the model updates, the impact was still significant. For example, the assimilation resulted in a mean absolute difference in the daily evapotranspiration forecasts of 10% of the openloop values. Additionally, at the Murrumbidgee Monitoring Network sites, the assimilation improved the fit between the MOSES forecasts and in situ observations of near-surface soil moisture at five of the seven sites with in situ data. Overall, the mean absolute (anomaly) correlation across the Murrumbidgee sites was increased from 0.74 (0.59) for the open-loop, to 0.78 (0.71) for the EKF. Of greater importance, the

assimilation also improved the root-zone soil moisture forecasts relative to the in situ data at all five sites for which in situ observations were available, increasing the mean absolute (anomaly) correlation from 0.60 (0.67) for the openloop, to 0.71 (0.72) for the EKF.

Errors in precipitation forecasts are a major source of error in NWP modeled soil moisture. While the EKF assimilation of near-surface soil moisture into MOSES cannot always correct for errors in the underlying soil moisture, it should be more effective at correcting errors in the forecast precipitation, which enter the model soil moisture through the near-surface layer. Examining individual time series highlights clear instances when assimilating the AMSR-E observations corrected the model for errors in the precipitation forcing (generated by the ACCESS NWP model). However, a quantitative relationship could not be found between the precipitation forcing errors and the analysis updates, including between the soil moisture analyses and model simulations generated by forcing the model with observed (rather than ACCESS) precipitation. This could be due to the precipitation errors being obscured by other errors in the model soil moisture, or it could also be due to errors in the assimilated AMSR-E observations, including the applied rescaling strategy. The fact that the assimilation improved the model soil moisture at the Murrumbidgee sites supports the former explanation.

Looking forward, these results indicate that the assimilation of near-surface soil moisture could benefit the ACCESS/MOSES models. The assimilation improved the model soil moisture forecasts at those sites where in situ data were available, by responding to errors in the near-surface layers, which are most sensitive to errors in the atmospheric forcing. This would compliment the assimilation of screen-level observations quite well, since the latter could more tightly constrain the deeper soil moisture layers to reduce errors in the transpiration forecasts. The assimilation strategy used here was specifically designed to allow the assimilation of both near-surface soil moisture and screen-level observations. To introduce the assimilation of screen-level observation into ACCESS/MOSES, the height at which the atmospheric forcing is applied to MOSES should be raised to the height of the first atmospheric layer in ACCESS (20 m). Additionally, the accuracy with which the off-line MOSES can estimate the screen-level forecasts from the coupled model would need to be confirmed.

7.2 Main conclusions and future directions

Potential of assimilating near-surface soil moisture and screen-level observations

Taken together, the results summarised above strongly suggest that near-surface soil moisture assimilation could benefit NWP. It has been demonstrated that near-surface soil moisture can be accurately sensed from current microwave sensors, and that assimilating these observations into the Australian NWP land surface model can improve the model root-zone soil moisture relative to the available in situ soil moisture observations. Additionally, while assimilating near-surface soil moisture and screen-level observations separately into Météo-France's NWP land surface model generated quite different root-zone soil moisture increments, assimilating the two data types together improved the model fit to both observation types. These results suggest that forecasts of both the low-level atmosphere and soil moisture could be slightly improved by assimilating screen-level observations together with remotely sensed near-surface soil moisture.

The next stage of this research is to couple the off-line land data assimilation system to the atmospheric model, so that each surface analysis is fed-back to the NWP model to generate the forcing for the next off-line assimilation cycle. This will likely enhance the impact of the assimilation, by introducing feedback between the land surface and the atmosphere. It will also allow further assessment of the impact of the soil moisture analyses, including the interaction between the different observation types, by allowing evaluation of the impact on forecasts of soil moisture-dependent variables.

Recently launched soil moisture remote sensors

Remotely sensed soil moisture from the AMSR-E instrument was used in this thesis, since until recently AMSR-E was the only instrument observing soil moisture with sufficient coverage for NWP. However, the ASCAT and SMOS satellites have since been launched, and both are recommended as potential alternatives to using AMSR-E data. SMOS was purpose designed for sensing soil moisture, and is expected to provide more accurate observations than AMSR-E. Additionally, the deeper (5 cm) surface soil layer observed by the L-band SMOS sensor is more strongly coupled to the root-zone soil moisture than the very

thin layer observed by AMSR-E. While ASCAT observes at the same C-band microwave frequencies as AMSR-E, and offers similar accuracy (Albergel et al., 2009), it has the advantage of being operationally supported. Consequently, ASCAT soil moisture observations can be obtained in near-real time, and are supported by sophisticated data quality flags and error estimates that are not currently available for AMSR-E.

Dependency on the model and model design

The MOSES and ISBA models treat the vertical resolution and flow of soil moisture quite differently, and the contrasting results obtained from assimilating near-surface soil moisture into each model highlights that the effectiveness (and likely accuracy) with which root-zone soil moisture can be constrained from near-surface soil moisture observations is model-dependent. For Météo-France's two-layer force-restore ISBA model, the near-surface soil moisture layer is so thin (about 1 cm) that updates to this layer have little impact on the subsequent model forecasts. Consequently, the assimilation must rely on deriving updates to the root-zone soil moisture from the near-surface observations. Since the root-zone soil moisture updates are applied to the entire soil column, assimilating the AMSR-E observations into ISBA added a reasonably large volume of moisture to the surface (1.8 mm day^{-1} on average). In contrast, the BoM's MOSES model describes soil moisture in four layers, with Darcian moisture flow between them. Since the coupling between the different soil moisture layers is quite weak, assimilating the near-surface soil moisture observations into MOSES only generated significant updates close to the surface (in the uppermost 35 cm, with most of this occurring in the uppermost 10 cm of the surface). In contrast to ISBA, the near-surface soil moisture layer in MOSES is sufficiently thick (10 cm) that analysis updates to this layer infiltrate through the soil profile, with a significant impact on the model root-zone soil moisture. However, this mechanism results in much smaller volumes of moisture being added and subtracted than for ISBA (0.5 mm day^{-1} on average, for similarly derived error covariances).

It is not known which of these models is more realistic (if either), and both compare well to other estimates of soil moisture (Calvet et al., 1998; Rooney and Claxton, 2006). In terms of selecting a model that is well suited to near-surface soil moisture assimilation, while the MOSES assimilation did not constrain the model root-zone soil moisture as strongly, it was more consistent with the

information that can be reliably extracted from remotely sensed near-surface soil moisture observations. The temporal dynamics of near-surface soil moisture are dominated by short-term (sub-daily) variability associated with atmospheric forcing, while the root-zone soil moisture evolves much more slowly (at monthly to seasonal scales). For the ISBA model, there is a dramatic difference in the time scales of the two model soil moisture layers, and the assimilation must rely on being able to extract information regarding errors in the slower mode (the root-zone) from observations that are dominated by the the faster mode (the near-surface layer). Given the systematic differences between modeled and observed soil moisture, it may not be possible to detect errors in the slowly varying mode from observations of the near-surface layer. In contrast, the inclusion of soil moisture layers with intermediate time scales in MOSES, allows the model itself to filter updates to the near-surface soil moisture through the soil profile. This approach is certainly better suited to correcting the NWP land surface in response to errors that enter the land surface via the near-surface layer, such as precipitation.

However, one advantage of ISBA is that it defines the near-surface soil moisture over a similar depth than is observed by C-band microwave brightness temperatures from AMSR-E (and ASCAT). In contrast, the near-surface layer in MOSES is 10 cm deep, introducing substantial representativity errors between the assimilated and modeled near-surface soil moisture. In this thesis these differences were reduced by temporally filtering the AMSR-E data with an exponential moving average filter. However, the benefit of this filter is not thought to extend beyond reducing the noisiness of the shallower observations (and the filter is certainly less accurate than using a sophisticated soil moisture model, such as ISBA or MOSES, to extrapolate near-surface observations over a deeper soil layer). The ideal model design (for assimilation purposes) would include both a near-surface layer of equivalent depth to the observations (for an accurate observation equivalent), and a second layer with a depth of 10 - 30 cm (for a longer memory). Additionally, once SMOS observations of the ~ 5 cm soil layer are available, this will enhance the coupling between the observed near-surface layer and the underlying soil moisture, and hence the ability to detect root-zone soil moisture errors from remotely sensed observations.

Refining the assimilation strategy

The EKF-based land surface assimilation scheme used in this thesis is an effective and affordable method of analysing the soil moisture in NWP models from observations of both near-surface soil moisture and screen-level variables, and it is recommended for use at NWP centres. The EKF offers several advantages over the current OI and nudging strategies used to assimilate screen-level observations at most NWP centres, including a more accurate representation of the relationship between model soil states and screen-level forecasts. However, its greatest advantage is the flexibility to easily assimilate new observation types, including near-surface soil moisture, although in the future other observations should also be investigated, including surface temperature, snow cover, albedo, and vegetation indices. Additionally, the state update vector in the EKF can easily be expanded, and in particular soil temperature should be included, since it is currently updated from screen-level observations, and has been neglected in this thesis.

While it is a clear improvement over the current techniques used to constrain soil moisture in NWP models, the EKF assimilation used here is very simplistic, and could likely be improved. Firstly, the spatially uniform and strictly diagonal error covariance matrices used here are too simplistic, and specifying more realistic errors would likely improve the resulting soil moisture analyses (e.g., Wilker et al., 2006). For the MOSES model, comparison of the model soil moisture in each layer to in situ observations suggested that there are vertical correlations between the model background errors in each soil layer, which were neglected here. For MOSES (and similarly structured models) the impact of assimilating the near-surface soil moisture with the EKF could then be enhanced by introducing vertical correlations into the additive model error term (off-diagonal terms in \mathbf{Q}), if the correlations could be accurately estimated.

While several recent papers have explored the use of adaptive filtering for improving the accuracy of model errors (Crow and Reichle, 2008; Reichle et al., 2008a), these adaptive approaches are based on the assumption that inaccuracies in the assimilation diagnostics are due to inaccuracies in the specified error covariance (rather than other causes, such as model-observation biases, or weak coupling between the observed and analysed variables). This assumption needs to be confirmed before these adaptive approaches can be pursued. Alternatively, using an EnKF rather than the EKF should improve the back-

ground error covariances, since this would allow the model errors to be defined more intuitively through the generation of the ensemble. An EnKF with an appropriately defined ensemble would likely also generate more substantial vertical model error correlations, since perturbing the model parameters and/or physics would likely produce vertically correlated errors. Additionally, Reichle and Koster (2003) showed that the impact of the near-surface soil moisture observations can be enhanced by specifying horizontal error covariances, which could be more easily achieved in the EnKF.

Addressing the model-observation systematic differences

The greatest difficulties encountered during this thesis were associated with the lack of an established truth for large scale soil moisture, and the systematic differences between remotely sensed, in situ observed, and modeled soil moisture. In particular, the soil moisture defined by a model is dependent on that model's physics, parameters, and resolution, and regardless of the accuracy achieved from soil moisture remote sensors, significant representativity errors (or systematic differences) will remain between model predicted and remotely sensed soil moisture. These representativity errors are exacerbated in NWP modeling by the use of globally defined, often inaccurate, model parameters. Near-surface soil moisture assimilation must then address these systematic differences between modeled and observed soil moisture. This was achieved in this thesis using a statistical bias correction prior to the assimilation, however for both ISBA and MOSES the bias correction was hampered by the limited data record available to sample the model soil moisture climatology, as well as the nonstationary nature of the model-observation biases. It was demonstrated in Chapter 5 that the assimilation results are very sensitive to the method used to remove the model-observation bias, and consequently using an inappropriate method will be detrimental. Obtaining a sufficient data record to accurately perform the bias correction then presents a major hurdle for the early uptake of recently launched soil moisture remote sensing missions, such as SMOS. It is also problematic for NWP modeling (excepting at those centres with reanalyses), since NWP model physics are frequently updated, making it difficult to obtain a long record of soil moisture forecasts. For regions with sufficient atmospheric observations, the model soil moisture data record could be increased by generating the model climatology using observed forcing; this approach offers the additional advantage over using archived NWP forecast soil moisture, since

it matches the soil moisture observations to a more accurate climatology.

Additionally, addressing the issues arising from the systematic difference between modeled and observed soil moisture will yield more immediate benefits than refining the data assimilation strategy, certainly until 3-D land surface assimilation becomes operationally feasible. This includes developing robust methods to evaluate soil moisture at continental to global scales, and improved methods to address the model-observation biases in soil moisture. In particular, the benefit of any refinements to the assimilation strategy cannot be judged until the impact of the assimilation can be confidently evaluated over large spatial and temporal domains. Consequently, it is recommended that addressing the systematic differences between modeled and observed soil moisture be given higher priority than introducing the above-mentioned refinements to the assimilation strategy.

References

- Albergel, C., Rüdiger, C., Carrer, D., Calvet, J.-C., Fritz, N., Naeimi, V., Bartalis, Z. and Hasenauer, S. (2009). An evaluation of ASCAT surface soil moisture products with in-situ observations in southwestern France, *Hydrology and Earth System Sciences*, **13**, 115–124.
- Albergel, C., Rüdiger, C., Pellarin, T., Calvet, J.-C., Fritz, N., Froissard, F., Suquia, D., Petitpa, A., Pignatelli, B. and Martin, E. (2008). From near-surface to root-zone soil moisture using an exponential filter: an assessment of the method based on in-situ observations and model simulations, *Hydrology and Earth System Sciences*, **12**, 1323–1337.
- Assistance for Land-Surface Modelling Activities (2004). The ALMA data exchange convention, Global Land Atmosphere System Study. At http://web.lmd.jussieu.fr/~polcher/ALMA/convention_input_3.html [last accessed May 2010].
- Baker, R., Lynn, B., Boone, A., Tao, W.-K. and Simpson, J. (2001). The influence of soil moisture, coastline curvature, and land-breeze circulations on sea-breeze initiated precipitation, *Journal of Hydrometeorology*, **2**, 193–211.
- Balsamo, G., Bouyssel, F. and Noilhan, J. (2004). A simplified bi-dimensional variational analysis of soil moisture from screen-level observations in a mesoscale numerical weather-prediction model, *Quarterly Journal of the Royal Meteorological Society*, **130**, 895–915.
- Balsamo, G., Mahfouf, J.-F., Bélair, S. and Deblonde, G. (2006). A global root-zone soil moisture analysis using simulated L-band brightness temperature in preparation for the hydros satellite mission, *Journal of Hydrometeorology*, **7**, 1126–1146.

- Balsamo, G., Mahfouf, J.-F., Bélair, S. and Deblonde, G. (2007). A land data assimilation system for soil moisture and temperature: An information content study, *Journal of Hydrometeorology*, **8**, 1225–1242.
- Balsamo, G., Viterbo, P., Beljaars, A., van den Hurk, B., Hirschi, M., Betts, A. and Scipal, K. (2009). A revised hydrology for the ECMWF Model: Verification from field site to terrestrial water storage and impact in the Integrated Forecast System, *Journal of Hydrometeorology*, **10**(3), 623–643.
- Bartalis, Z., Wagner, W., Naeimi, V., Hasenauer, S., Scipal, K., Bonekamp, H., Figa, J. and Anderson, C. (2007). Initial soil moisture retrievals from the METOP-A Advanced Scatterometer (ASCAT), *Geophysical Research Letters*, **34**, L20401.
- Bélair, S., Crevier, L.-P., Mailhot, J., Bilodeau, B. and Delage, Y. (2003). Operational implementation of the ISBA land surface scheme in the Canadian regional weather forecast model. Part I: Warm season results, *Journal of Hydrometeorology*, **4**, 352–370.
- Best, M. and Maisey, P. (2002). A physically based soil moisture nudging scheme, Hadley Centre Technical Note 35, 22 pp., Met Office Hadley Centre, Exeter, UK.
- Best, M., Baljaars, A., Polcher, J. and Viterbo, P. (2004). A proposed structure for coupling tiled surfaces with the planetary boundary layer, *Journal of Hydrometeorology*, **5**, 1271–1278.
- Betts, A. and Ball, J. (1995). The FIFE surface diurnal cycle climate, *Journal of Geophysical Research*, **100**, 25679–25693.
- Bhumralkar, C. (1975). Numerical experiment on the computation of ground surface-temperature in an atmospheric general circulation model, *Journal of Applied Meteorology*, **14**, 1246–1258.
- Bindlish, R., Jackson, T., Gasiewski, A., Klein, M. and Njoku, E. (2006). Soil moisture mapping and AMSR-E validation using PSR in SMEX02, *Remote Sensing of Environment*, **103**, 127–139.
- Boone, A., Masson, V., Meyers, T. and Noilhan, J. (2000). The influence of the inclusion of soil freezing on simulations by a soil-vegetation-atmosphere transfer scheme, *Journal of Applied Meteorology*, **39**, 1544–1569.

- Bosch, D., Lakshmi, V., Jackson, T., Choi, M. and Jacobs, J. (2006). Large scale measurements of soil moisture for validation of remotely sensed data: Georgia soil moisture experiment of 2003, *Journal of Hydrology*, **323**, 120–137.
- Bouyssel, F., Casse, V. and Pailleux, J. (1999). Variational surface analysis from screen level atmospheric parameters, *Tellus*, **51A**, 453–468.
- Bubnova, R., Hello, G., Bénard, P. and Geleyn, J.-F. (1995). Integration of the fully elastic equations cast in the hydrostatic-pressure terrain-following coordinate in the framework of the ARPEGE/ALADIN NWP system, *Monthly Weather Review*, **123**, 515–535.
- Calvet, J.-C. and Noilhan, J. (2000). From near-surface to root-zone soil moisture using year-round data, *Journal of Hydrometeorology*, **1**, 393–411.
- Calvet, J.-C., Noilhan, J., Roujean, J.-L., Bessemoulin, P., Cabelguenne, M., Olioso, A. and Wigneron, J.-P. (1998). An interactive vegetation SVAT model tested against data from six contrasting sites, *Agricultural and Forest Meteorology*, **92**, 73–95.
- Campbell, G. and Campbell, M. (1982). Irrigation scheduling using soil moisture measurements: Theory and practice, *Advances in Irrigation*, **1**, 25–42.
- Capehart, W. and Carlson, T. (1997). Decoupling of surface and near-surface soil water content: A remote sensing perspective, *Water Resources Research*, **33**, 1383–1395.
- Cardinali, C., Pezzulli, S. and Andersson, E. (2004). Influence-matrix diagnostic of a data assimilation system, *Quarterly Journal of the Royal Meteorological Society*, **130**, 2767–2786.
- Castelli, F., Entekhabi, D. and Caporali, E. (1999). Estimation of surface heat flux and an index of soil moisture using adjoint-state surface energy balance, *Water Resources Research*, **35**, 3115–3125.
- Ceballos, A., Scipal, K., Wagner, W. and Martínez-Fernández, J. (2005). Validation of ERS scatterometer-derived soil moisture data in the central part of the Duero Basin, Spain, *Hydrological Processes*, **19**, 1549–1566.
- Coiffier, J., Ernie, Y., Geleyn, J.-F., Clochard, J., Hoffman, J. and Dupont, F. (1987). The operational hemispheric model at the French meteorological

- service, *Journal of the Meteorological Society of Japan, Special Issue on Short and Medium Range Numerical Weather Prediction*, pp. 337–345.
- Conil, S., Douville, H. and Tyteca, S. (2007). The relative influence of soil moisture and SST in climate predictability explored within ensembles of AMIP type experiments, *Climate Dynamics*, **28**, 125–145.
- Cosh, M., Jackson, T., Starks, P. and Heathman, G. (2006). Temporal stability of surface soil moisture in the Little Washita River watershed and its applications in satellite soil moisture product validation, *Journal of Hydrology*, **323**, 168–177.
- Cox, P., Betts, R., Bunton, C., Essery, R., Rowntree, P. and Smith, J. (1999). The impact of new land surface physics on the GCM simulation of climate and climate sensitivity, *Climate Dynamics*, **15**, 183–203.
- Crosson, W., Limaye, A. and Laymon, C. (2005). Parameter sensitivity of soil moisture retrievals from airborne C- and X-band radiometer measurements in SMEX02, *IEEE Transactions on Geoscience and Remote Sensing*, **43**, 2842–2853.
- Crow, W. and Reichle, R. (2008). Comparison of adaptive filtering techniques for land surface data assimilation, *Water Resources Research*, **44**, W08423.
- Crow, W. and Wood, E. (2003). The assimilation of remotely sensed soil brightness temperature imagery into a land surface model using Ensemble Kalman filtering: a case study based on ESTAR measurements during SGP97, *Advances in Water Resources*, **26**, 137–149.
- Crow, W., Miralles, D. and Cosh, M. (2009). A quasi-global evaluation system for satellite surface soil moisture retrievals, *IEEE Transactions on Geoscience and Remote Sensing*, **48**, 2516–2527.
- Darcy (1856). *Les fontaines publiques de la ville de Dijon*, 647pp, Dalmont, Paris.
- Davies, T., Cullen, M., Malcolm, A., Mawson, M., Staniforth, A., White, A. and Wood, N. (2005). A new dynamical core for the Met Office's global and regional modelling of the atmosphere, *Quarterly Journal of the Royal Meteorological Society*, **131**, 1759–1782.

- Dawdy, D. and Matalas, N. (1964). Statistical and probability analysis of hydrologic data, part III: Analysis of variance, covariance and time series, in *Handbook of Applied Hydrology, A Compendium of Water-Resources Technology*, pp. 8.68–8.91, McGraw-Hill Book Company, New York.
- de Jeu, R., Holmes, T., Panciera, R. and Walker, J. P. (2009). Parameterization of the Land Parameter Retrieval Model for L-Band observations using the NAFE05 data set, *IEEE Geoscience and Remote Sensing Letters*, **6**, 630–634.
- de Jeu, R., Wagner, W., Holmes, T., Dolman, A., van de Giesen, N. and Friesen, J. (2008). Global soil moisture patterns observed by space borne microwave radiometers and scatterometers, *Surveys in Geophysics*, **29**, 399–420.
- de Lannoy, G., Reichle, R., Houser, P., Pauwels, V. and Verhoest, N. (2007). Correcting for forecast bias in soil moisture assimilation with the ensemble Kalman filter, *Water Resources Research*, **43**, W09410.
- Deardorff, J. (1977). A parameterization of the ground surface moisture content for use in atmospheric prediction models, *Journal of Applied Meteorology*, **16**, 1182–1185.
- Dee, D. (2005). Bias and data assimilation, *Quarterly Journal of the Royal Meteorological Society*, **131**, 3323–3343.
- Dharssi, I., Vidale, P., Verhoef, A., Macpherson, B., Jones, C. and Best, M. (2009). New soil physical properties implemented in the Unified Model at PS18, Met Office Meteorology and Development, Technical Report No. 528, 34 pp., Met Office, Exeter, UK.
- Dharssi, I., Bovis, K., and Macpherson, B. (2010). Assimilation of ASCAT surface soil wetness, Met Office Meteorology and Development, Technical Report No. 548, 24 pp., Met Office, Exeter, UK.
- Dirmeyer, P., Guo, Z. and Gao, X. (2004). Comparison, validation, and transferability of eight multiyear global soil wetness products, *Journal of Hydrometeorology*, **5**, 1011–1033.
- Dorigo, W., Scipal, K., Parinussa, R., Liu, Y., Wagner, W., de Jeu, R., and Naeimi, V. (2010). Error characterisation of global active and pas-

- sive microwave soil moisture datasets, *Hydrology and Earth System Sciences*, **14**, 2605–2616.
- Douville, H., Viterbo, P., Mahfouf, J.-F. and Beljaars, A. (2000). Evaluation of the optimum interpolation and nudging techniques for soil moisture analysis using FIFE data, *Monthly Weather Review*, **128**, 1733–1756.
- Draper, C., Mahfouf, J.-F. and Walker, J. (2009a). An EKF assimilation of AMSR-E soil moisture into the ISBA land surface scheme, *Journal of Geophysical Research*, **114**, D20104.
- Draper, C., Mahfouf, J.-F. and Walker, J. (2011). Root-zone soil moisture from the assimilation of screen-level variables and remotely sensed soil moisture, *Journal of Geophysical Research*, **116**, D02127.
- Draper, C., Walker, J., Steinle, P., de Jeu R and Holmes, T. (2009b). An evaluation of AMSR-E derived soil moisture over Australia, *Remote Sensing of Environment*, **113**, 703–710.
- Drusch, M. (2007). Initializing numerical weather prediction models with satellite-derived surface soil moisture: Data assimilation experiments with ECMWF’s Integrated Forecast System and the TMI soil moisture data set, *Journal of Geophysical Research*, **112**, D03102.
- Drusch, M. and Viterbo, P. (2007). Assimilation of screen-level variables in ECMWF’s Integrated Forecast System: A study on the impact on the forecast quality and analyzed soil moisture, *Monthly Weather Review*, **135**, 300–314.
- Drusch, M., Scipal, K., de Rosnay, P., Balsamo, G., Andersson, E., Bougeault, P. and Viterbo, P. (2009). Towards a Kalman Filter based soil moisture analysis system for the operational ECMWF Integrated Forecast System, *Geophysical Research Letters*, **36**, L10401.
- Drusch, M., Wood, E. and Gao, H. (2005). Observation operators for the direct assimilation of TRMM microwave imager retrieved soil moisture, *Geophysical Research Letters*, **32**, L15403.
- Drusch, M., Wood, E., Gao, H. and Thiele, A. (2004). Soil moisture retrieval during the Southern Great Plains Hydrology Experiment 1999: A comparison between experimental remote sensing data and operational products, *Water Resources Research*, **40**, W02504.

- Dunne, S. and Entekhabi, D. (2005). An ensemble-based reanalysis approach to land data assimilation, *Water Resources Research*, **41**, W02013.
- Dunne, S. and Entekhabi, D. (2006). Land surface state and flux estimation using the ensemble Kalman smoother during the Southern Great Plains 1997 field experiment, *Water Resources Research*, **42**, W01407.
- Entekhabi, D., Njoku, E., Houser, P., Spencer, M., Doiron, T., Kim, Y., Smith, J., Girard, R., Bélair, S., Crow, W., Jackson, T., Kerr, Y., Kimball, J., Koster, R., McDonald, K., O'Neill, P., Pultz, T., Running, S., Shi, J., Wood, E. and van Zyl, J. (2004). The hydrosphere state (Hydros) satellite mission: An earth system pathfinder for global mapping of soil moisture and land freeze/thaw, *IEEE Transactions on Geoscience and Remote Sensing*, **42**, 2184–2195.
- Entin, J., Robock, A., Vinnikov, K., Hollinger, S., Liu and Namkhai, A. (2000). Temporal and spatial scales of observed soil moisture variations in the extra-tropics, *Journal of Geophysical Research*, **105**, 11865 – 11877.
- Essery, R., Best, M. and Cox, P. (2001). MOSES 2.2 Technical Documentation, Hadley Centre Technical Note 30, 30 pp., Met Office Hadley Centre, Exeter, UK.
- Essery, R., Best, M., Betts, R., Cox, P. and Taylor, C. (2003). Explicit representation of subgrid heterogeneity in a GCM land-surface scheme, *Journal of Hydrometeorology*, **4**, 530–543.
- Evensen, G. (1994). Sequential data assimilation with a nonlinear quasi-geostrophic model using Monte-Carlo methods to forecast error statistics, *Journal of Geophysical Research*, **99**, 10143–10162.
- Famiglietti, J., Ryu, D., Berg, A., Rodell, M., and Jackson, T. (2008). Field observations of soil moisture variability across scales, *Journal of Geophysical Research*, **44**, W01423.
- Findell, K. and Eltahir, E. (2003). Atmospheric controls on soil moisture-boundary layer interactions. Part 1: Framework development, *Journal of Hydrometeorology*, **4**, 552–569.
- Fischer, E., Seneviratne, S., Vidale, P., Lüthi, D. and Schär, C. (2007). Soil moisture - atmosphere interactions during the 2003 European summer heat wave, *Journal of Climate*, **20**, 5081–5099.

- Fitzjohn, C., Ternan, J. and Williams, A. (1998). Soil moisture variability in a semi-arid gully catchment: implications for runoff and erosion control, *CATENA*, **32**, 55–70.
- Gallus, W. and Segal, M. (2000). Sensitivity of forecast rainfall in a Texas convective system to soil moisture and convective parameterization, *Weather and Forecasting*, **15**, 509–525.
- Ghil, M. and Malanotte-Rizzoli, P. (1991). Data assimilation in meteorology and oceanography, *Advances in Geophysics*, **33**, 141–266.
- Giard, D. and Bazile, E. (2000). Implementation of a new assimilation scheme for soil and surface variables in a global NWP model, *Monthly Weather and Forecasting*, **128**, 997–1015.
- Grasso, L. (2000). A numerical simulation of dryline sensitivity to soil moisture, *Monthly Weather Review*, **128**, 2816–2834.
- Gruhler, C., de Rosnay, P., Kerr, Y., Mougin, E., Ceschia, E., Calvet, J.-C. and Richaume, P. (2008). Evaluation of AMSR-E soil moisture product based on ground measurements over a temperate and semi-arid regions, *Geophysical Research Letters*, **35**, L10405.
- Habets, F., Boone, A., Champeaux, J.-L., Etchevers, P., Franchisteguy, L., Leblois, E., Ledoux, E., Le Moigne, P., Martin, E., Morel, S., Noilhan, J. and Quintana-Seguí, P. (2008). The SAFRAN-ISBA-MODCOU hydrometeorological model applied over France, *Journal of Geophysical Research*, **113**, D06113.
- Hess, H. (2001). Assimilation of screen-level observations by variational soil moisture analysis, *Meteorology and Atmospheric Physics*, **77**, 145–154.
- Hess, R., Lange, M. and Wergen, W. (2008). Evaluation of the variational soil moisture assimilation scheme at Deutscher Wetterdienst, *Quarterly Journal of the Royal Meteorological Society*, **134**, 1499 – 1512.
- Houser, P., Shuttleworth, W., Famiglietti, J., Gupta, H., Syed, K. and Goodrich, D. (1998). Integration of soil moisture remote sensing and hydrologic modeling using data assimilation, *Water Resources Research*, **34**, 3405–3420.

- Ide, K., Courtier, P., Ghil, M. and Lorenc, A. (1997). Unified notation for data assimilation: Operational, sequential and variational, *Journal of the Meteorological Society of Japan*, **75**, 181–189.
- Jackson, T. (1993). Measuring surface soil moisture using passive microwave remote sensing. III, *Hydrological Processes*, **7**, 139–152.
- Jackson, T. (2005). Satellite remote sensing of soil moisture, in *Advances in Water Science Methodologies*, pp. 91–97, Balkema, Leiden, The Netherlands.
- Jackson, T. and Schmugge, T. (1991). Vegetation effects on the microwave emission of soils, *Remote Sensing of Environment*, **36**, 203–212.
- Jackson, T., Bindlish, R., Gasiewski, A., Stankov, B., Klein, M., Njoku, E., Bosch, D., Coleman, T., Laymon, C. and Starks, P. (2005). Polarimetric Scanning Radiometer C- and X- band observations during SMEX03, *IEEE Transactions on Geoscience and Remote Sensing*, **43**, 2418–2430.
- Jackson, T., Gasiewski, A., Oldak, A., Klein, M., Njoku, E., Yevgrafov, A., Christiani, S. and Bindlish, R. (2002). Soil moisture retrieval using the C-band polarimetric scanning radiometer during the Southern Great Plains 1999 Experiment, *IEEE Transactions on Geoscience and Remote Sensing*, **40**, 2151–2841.
- Joint Centre for Hydro-Meteorological Research (2001). JULES: Joint UK Land Environment Simulator, Joint Centre for Hydro-Meteorological Research, Wallingford, UK. At <http://www.jchmr.org/jules/index.html> [last accessed May 2010].
- Jones, D., Wang, W. and Fawcett, R. (2007). Climate data for the Australian Water Availability Project, Australian Water Availability Project Final Milestone Report, 36 pp., Australian Government Bureau of Meteorology, Melbourne, Australia. At <http://adl.brs.gov.au/brsShop/data/awapfinalreport200710.pdf> [last accessed December 2010].
- Jones, D., Wang, W. and Fawcett, R. (2009). High-quality spatial climate data-sets for Australia, *Australian Meteorological and Oceanographic Journal*, **58**, 233–248.

- Kalman, R. (1960). A new approach to linear filtering and prediction problems, *Transactions of the ASME—Journal of Basic Engineering*, **82**(Series D), 35–45.
- Kalnay, E. (2003). *Atmospheric Modeling, Data Assimilation, and Predictability*, Cambridge University Press, Cambridge, UK.
- Kerr, Y., Waldteufel, P., Wigneron, J., Cabot, J., Boutin, J., Escoriheula, M.-J., Font, J. Reul, N., Gruhier, C., Juglea, S., Delwart, S., Drinkwater, M., Hahne, A., Martin-Neira, M., and Mecklenburg, S. (2010). The SMOS mission: new tool for monitoring key elements of the global water cycle, *Proceedings of the IEEE*, **95**, 666–687.
- Kerr, Y., Waldteufel, P., Wigneron, J., Martinuzzi, J., Font, J. and Berger, M. (2001). Soil moisture retrieval from space: The Soil Moisture and Ocean Salinity (SMOS) mission, *IEEE Transactions on Geoscience and Remote Sensing*, **39**, 1729–1735.
- Koike, T., Nakamura, Y., Kaihotsu, I., Davva, G., Matsuura, N., Tamagawa, K. and Fujii, H. (2004). Development of an Advanced Microwave Scanning Radiometer (AMSR-E) algorithm of soil moisture and vegetation water content, *Annual Journal of Hydraulic Engineering, JSCE*, **48**, 217–222.
- Komma, J., Blöschl, G. and Reszler, C. (2008). Soil moisture updating by Ensemble Kalman Filtering in real-time flood forecasting, *Journal of Hydrology*, **357**, 228–242.
- Koster, R., Dirmeyer, P., Guo, Z., Bonan, G., Chan, E., Cox, P., Gordon, C., Kanae, S., Kowalczyk, E., Lawrence, D., Liu, P., Lu, C., Malyshev, S., McAvaney, B., Mitchell, K., Mocko, D., Oki, T., Oleson, K., Pitman, A., Sud, Y., Taylor, C., Versegny, D., Vasic, R., Xue, Y. and Yamada, T. (2004). Regions of strong coupling between precipitation and soil moisture, *Science*, **305**, 1138–1140.
- Koster, R., Guo, Z., Yang, R., Dirmeyer, P., Mitchell, K. and Puma, M. (2009). On the nature of soil moisture in land surface models, *Journal of Climate*, **22**, 4322–4335.
- Kumar, S., Reichle, R., Koster, R., Crow, W. and Peters-Lidard, C. (2009). Role

- of subsurface physics in the assimilation of surface soil moisture observations, *Journal of Hydrometeorology*, **10**, 1534–1547.
- Le Marshall, J., Jung, J., Riishojgaard, L.-P., Lord, S., Derber, J., Xiao, Y., Seecamp, R., Steinle, P. and Sims, H. (2009). Satellite data assimilation, *5th WMO Data Assimilation Symposium, Extended Abstract, Melbourne, Australia, 5 - 9 October 2009*. At http://www.cawcr.gov.au/staff/pxs/wmoda5/ExtendedAbstracts/LeMarshall_et.al.pdf [last accessed December 2010].
- Le Moigne, P. (2009). SURFEX Scientific Documentation, Note de Travail du Groupe de Meteorologie a Moyenne Echelle, No 87., 211 pp., Météo-France CNRM, Toulouse, France.
- Li, J. and Islam, S. (2002). Estimation of root zone soil moisture and surface fluxes partitioning using near surface soil moisture measurements, *Journal of Hydrology*, **259**, 1–14.
- Li, L., Njoku, E., Im, E., Chang, P. and Germain, K. (2004). A preliminary survey of radio-frequency interference over the US in Aqua AMSR-E data, *IEEE Transactions on Geoscience and Remote Sensing*, **42**, 380–390.
- Liu, Y., van Dijk, A., de Jeu, R. and Holmes, T. (2009). An analysis of spatiotemporal variations of soil and vegetation moisture from a 29-year satellite-derived data set over mainland Australia, *Water Resources Research*, **45**, W07405.
- Mahfouf, J.-F. (1991). Analysis of soil moisture from near-surface parameters: A feasibility study, *Journal of Applied Meteorology*, **30**, 1534–1547.
- Mahfouf, J.-F. (2000). A revised land-surface analysis scheme in the Integrated Forecast System, *ECMWF newsletter 88, Summer-Autumn 2000*.
- Mahfouf, J.-F. (2010a). Amelioration des statistiques des erreurs d'ébauche pour l'analyse de l'état hydrique des sols, *Ateliers de Modélisation de l'Atmosphère 2010*, abstract, Toulouse, 26-28 January 2010. At http://www.cnrm.meteo.fr/ama2010/resumes_courts/resume_69.html [last accessed June 2010].

- Mahfouf, J.-F. (2010b). Assimilation of satellite derived soil moisture from ASCAT in a limited area NWP model, *Quarterly Journal of the Royal Meteorological Society*, **136**, 784–798.
- Mahfouf, J.-F., Bergaoui, K., Draper, C., Bouyssel, C., Taillefer, F. and Taseva, L. (2009). A comparison of two off-line soil analysis schemes for assimilation of screen-level observations, *Journal of Geophysical Research*, **114**, D08105.
- Martin, W. and Xue, M. (2006). Sensitivity analysis of convection of the 24 May 2002 IHOP case using very large ensembles, *Monthly Weather Review*, **134**, 192–207.
- Masson, V., Champeaux, J.-L., Chauvin, F., Meriguet, C. and Lacaze, R. (2003). A global database of land surface parameters at 1-km resolution in meteorological and climate models, *Journal of Climate*, **16**, 1261 – 1282.
- McCabe, M., Gao, H. and Wood, E. H. (2005a). Evaluation of AMSR-E-derived soil moisture retrievals using ground-based and PSR airborne data during SMEX02, *Journal of Hydrometeorology*, **6**, 864–877.
- McCabe, M., Wood, E. and Gao, H. (2005b). Initial soil moisture retrievals from AMSR-E: Multiscale comparison using in situ data and rainfall patterns over Iowa, *Geophysical Research Letters*, **32**, L06403.
- McVicar, T. and Van Niel, T. (2005). Deriving moisture availability from time series remote sensing for drought assessment, Report to the Reference Group of the Australian Water Availability Project, 29 pp., CSIRO Land and Water, Canberra, Australia. At http://www.clw.csiro.au/publications/consultancy/2005/deriving_moisture_availability_AWAP.pdf [last accessed December 2010].
- Merlin, O., Walker, J., Panciera, R., Young, R., Kalma, J. and Kim, E. (2007). Soil moisture measurement in heterogeneous terrain, *MODSIM 2007 International Congress on Modelling and Simulation. Modelling and Simulation Society of Australia and New Zealand*, pp. 2604–2610.
- Merz, R. and Blöschl, G. (2009). A regional analysis of event runoff coefficients with respect to climate and catchment characteristics in Austria, *Water Resources Research*, **45**, W01405.

- Mills, G. (2001). Impact of screen-level moisture observations in a regional data assimilation system, Bureau of Meteorology Research Centre Research Report No. 82, 42 pp., Australian Government Bureau of Meteorology, Melbourne, Australia.
- Miralles, D., Crow, W., Cosh, M. (2010). Estimating Spatial Sampling Errors in Coarse-Scale Soil Moisture Estimates Derived from Point-Scale Observations, *Journal of Hydrometeorology*, **11**, 1423–1429.
- Mitchell and Mitchell, K., Lohmann, D., Houser, P., Wood, E., Schaake, J., Robock, A., Cosgrove, B., Sheffield, J., Qingyun, D., Lifeng, L., Higgins, R., Pinker, R., Tarpley, D., Lettenmaier, D., Curtis, M., Entin, J., Ming, P., Wei, S., Koren, V., Meng, J., Ramsay, B. and Bailey, A. The multi-institution North American Land Data Assimilation System (NLDAS): Utilizing multiple GCIP products and partners in a continental distributed hydrological modeling system, *Journal of Geophysical Research*, **109**, D07S90.
- Muñoz Sabater, J., Jarlan, L., Calvet, J.-C., Bouyssel, F. and De Rosnay, P. (2007). From near-surface to root-zone soil moisture using different assimilation techniques, *Journal of Hydrometeorology*, **8**, 194–206.
- Naeimi, V., Scipal, K., Bartalis, Z., Hasenauer, S. and Wagner, W. (2009). An improved soil moisture retrieval algorithm for ERS and METOP scatterometer observations, *IEEE Transactions on Geoscience and Remote Sensing*, **47**, 1999–2013.
- National Meteorological and Oceanographic Centre (2010). Operational implementation of the ACCESS Numerical Weather Prediction systems, NMOC Operations Bulletin No. 83, At <http://www.bom.gov.au/australia/charts/bulletins/apob83.pdf> [last accessed January 2011].
- Ni-Meister, W., Houser, P. and Walker, J. (2007). Soil moisture initialization for climate prediction: Assimilation of scanning multifrequency microwave radiometer soil moisture data into a land surface model, *Journal of Geophysical Research*, **111**, D20102.
- Njoku, E., Ashcroft, P., Chan, T. and Li, L. (2005). Global survey of statistics of radio-frequency interference in AMSR-E land observations, *IEEE Transactions on Geoscience and Remote Sensing*, **43**, 938–947.

- Njoku, E., Jackson, T., Lakshmi, V., Chan, T. and Nghiem, S. (2003). Soil moisture retrieval from AMSR-E, *IEEE Transactions on Geoscience and Remote Sensing*, **41**, 215–229.
- Noilhan, J. and Mahfouf, J. (1996). The ISBA land surface parameterisation scheme, *Global and Planetary Change*, **13**, 145–159.
- Noilhan, J. and Planton, S. (1989). A simple parameterization of land surface processes for meteorological models, *Monthly Weather Review*, **117**, 536–549.
- Owe, M., de Jeu, R. and Holmes, T. (2008). Multisensor historical climatology of satellite-derived global land surface moisture, *Journal of Geophysical Research*, **113**, F01002.
- Owe, M., de Jeu, R. and Walker, J. (2001). A methodology for surface soil moisture and vegetation optical depth retrieval using the microwave polarization difference index, *IEEE Transactions on Geoscience and Remote Sensing*, **39**, 1643 – 1654.
- Peel, M., Finlayson, B. and McMahon, T. (2007). Updated world map of the Köppen-Geiger climate classification, *Hydrology and Earth System Sciences*, **11**, 1633–1644.
- Pescod, N. (1994). A four-parameter, three-layer model of soil moisture based on hydraulic properties of the soil in the absence of vegetation, in J. Jasper and P. Meighen (eds), *Bureau of Meteorology Research Report, No 46. Parametrisation of Physical Processes: Papers Presented at the 5th BMRC Modelling Workshop*, pp. 101–106.
- Pielke, R. (2001). Influence of the spatial distribution of vegetation and soils on the prediction of cumulus convective rainfall, *Reviews of Geophysics*, **39**, 151–177.
- Pietroniro, A. and Leconte, R. (2005). A review of Canadian remote sensing and hydrology, 1999-2003, *Hydrological Processes*, **19**, 285–301.
- Puri, K. (2006). Overview of ACCESS, in A. Hollis and A. Kariko (eds), *The Australian Community Climate and Earth System Simulator (ACCESS) - Challenges and Opportunities: Extended Abstracts of Presentations at the Eighteenth Annual BMRC Modelling Workshop*, pp. 1–4.

- Puri, K., Dietachmayer, G., Mills, G., Davidson, N., Bowen, R. and Logan, L. (1998). The new BMRC Limited Area Prediction System, LAPS, *Australian Meteorological Magazine*, **47**, 203–223.
- Quintana-Seguí, P., Le Moigne, P., Durand, Y., Martin, E., Habets, F., Bailon, M., Canellas, C., Franchisteguy, L. and Morel, S. (2008). Analysis of near-surface atmospheric variables: Validation of the SAFRAN analysis over France, *Journal of Applied Meteorology and Climatology*, **47**, 92–107.
- Rawlins, F., Ballard, S., Bovis, K., Clayton, A., Li, D., Inverarity, G., Lorenc, A. and Payne, T. (2007). The Met Office global four-dimensional variational data assimilation scheme, *Quarterly Journal of the Royal Meteorological Society*, **133**, 347–362.
- Reichle, R. and Koster, R. (2003). Assessing the impact of horizontal error correlations in background fields on soil moisture estimation, *Journal of Hydrometeorology*, **4**, 1229–1242.
- Reichle, R. and Koster, R. (2004). Bias reduction in short records of satellite soil moisture, *Geophysical Research Letters*, **31**, L19501.
- Reichle, R. and Koster, R. (2005). Global assimilation of satellite surface soil moisture retrievals into the NASA Catchment Land Model, *Geophysical Research Letters*, **32**, L02404.
- Reichle, R., Crow, W. and Keppenne, C. (2008a). An adaptive ensemble Kalman filter for soil moisture data assimilation, *Water Resources Research*, **44**, W03423.
- Reichle, R., Crow, W., Koster, R., Sharif, H. and Mahanama, S. (2008b). Contribution of soil moisture retrievals to land data assimilation products, *Geophysical Research Letters*, **35**, L01404.
- Reichle, R., Koster, R. and Mahanama, S. (2004a). Bias correction of satellite soil moisture and assimilation into the NASA catchment land surface model, in A. Teuling, H. Leihnse, P. Troch, J. Sheffields and E. Wood (eds), *Proceedings of the 2nd International CAHMDA Workshop on The Terrestrial Water Cycle: Modelling and Data Assimilation Across Catchment Scales*, pp. 149–1153.

- Reichle, R., Koster, R., Dong, J. and Berg, A. (2004b). Global soil moisture from satellite observations, land surface models, and ground data: Implications for data assimilation, *Journal of Hydrometeorology*, **5**, 430 – 442.
- Reichle, R., Koster, R., Liu, P., Mahanama, S., Njoku, E. and Owe, M. (2007). Comparison and assimilation of global soil moisture retrievals from the Advanced Microwave Scanning Radiometer for the Earth Observing System (AMSR-E) and the Scanning Multichannel Microwave Radiometer (SMMR), *Journal of Geophysical Research*, **112**, D09108.
- Reichle, R., McLaughlin, D. and Entekhabi, D. (2001). Variational data assimilation of microwave radiobrightness observations for land surface hydrology applications, *IEEE Transactions on Geoscience and Remote Sensing*, **39**, 1708 – 1718.
- Reichle, R., McLaughlin, D. and Entekhabi, D. (2002a). Hydrologic data assimilation with ensemble Kalman filter, *Monthly Weather Review*, **130**, 103–114.
- Reichle, R., Walker, J., Koster, R. and Houser, P. (2002b). Extended versus Ensemble Kalman Filtering for land data assimilation, *Journal of Hydrometeorology*, **3**, 728–740.
- Richards, L. (1931). Capillary conduction of liquids in porous medium, *Physics*, **1**, 318–333.
- Robock, A., Vinnikov, K., Srinivasan, G., Entin, J., Hollinger, S., Speranskaya, N., Liu, S. and Namkhai, A. (2000). The Global Soil Moisture Data Bank, *Bulletin of the American Meteorological Society*, **81**, 1281 – 1299.
- Rodriguez, E., Navascues, J., Ayuso, S. and Järvenoja (2003). Analysis of surface variables and parameterisation of surface processes in HIRLAM. Part 1: Approach and verification by parallel runs, HIRLAM Technical Report No. 59, 52 pp., Swedish Meteorological and Hydrological Institute, Norrköping, Sweden.
- Rooney, G. and Claxton, B. (2006). Comparison of the Met Office’s Surface Exchange Scheme, MOSES, against field observations, *Quarterly Journal of the Royal Meteorological Society*, **136**, 425–446.

- Rüdiger, C., Calvet, J.-C., Gruhier, C., Holmes, T., de Jeu, R. and Wagner, W. (2009). An intercomparison of ERS-Scat and AMSR-E soil moisture observations with model simulations over France, *Journal of Hydrometeorology*, **10**, 431–447.
- Rüdiger, C., Hancock, G., H.Hemakumara, Jacob, B., Kalma, J., Martinez, C., Thyer, M., Walker, J., Wells, T. and Willgoose, G. (2007). Goulburn River experimental catchment data set, *Water Resources Research*, **43**, W10403.
- Schaake, J., Duan, Q., Koren, V., Mitchell, K., Houser, P., Wood, E., Robock, A., Lettenmaier, D., Lohmann, D., Cosgrove, B., Sheffield, J., Luo, L., Higgins, R., Pinker, R. and Tarpley, D. (2004). An intercomparison of soil moisture fields in the North American Land Data Assimilation System (NLDAS), *Journal of Geophysical Research*, **109**, D01S90.
- Schmugge, T. and Jackson, T. (1996). Soil moisture variability, in *Scaling Up in Hydrology Using Remote Sensing*, Stewart, J., Engman, E., Feddes, R., and Kerr, Y. (eds), pp. 183–192, John Wiley and Sons, Chichester, UK.
- Scipal, K., Drusch, M. and Wagner, W. (2008). Assimilation of a ERS scatterometer derived soil moisture index in the ECMWF numerical weather prediction system, *Advances in Water Resources*, **31**(8), 1101–1112.
- Scipal, K., Holmes, T., de Jeu, R., Naeimi, V. and Wagner, W. (2008). A possible solution for the problem of estimating the error structure of global soil moisture data sets, *Geophysical Research Letters*, **35** L22403.
- Seth, A. and Giorgi, F. (1998). The effects of domain choice on summer precipitation simulation and sensitivity in a regional climate model, *Journal of Climate*, **11**, 2698–2712.
- Seuffert, G., Wilker, H., Viterbo, P., Drusch, M. and Mahfouf, J.-F. (2004). The usage of screen-level parameters and microwave brightness temperature for soil moisture analysis, *Journal of Hydrometeorology*, **5**, 516–531.
- Sheffield, J., Goteti, G., Wen, F. and Wood, E. (2004). A simulated soil moisture based drought analysis for the United States, *Journal of Geophysical Research*, **109**, D24108.

- Smith, R., Blyth, E., Finch, J., Goodchild, S., Hall, R. and Madry, S. (2006). Soil state and surface hydrology diagnosis based on MOSES in the Met Office Nimrod nowcasting system, *Meteorological Applications*, **13**, 89–109.
- Stoffelen, A. (1998). Towards the true near-surface wind speed: Error modeling and calibration using triple collocation, *Journal of Geophysical Research*, **103**, 7755–7766.
- Taillefer, F. (2002). CANARI (Code for the Analysis Necessary for Arpege for its Rejects and its Initialization) Technical Documentation, Internal CNRM/GMAP Report, CNRM, 55 pp., Météo-France, Toulouse, France.
- Tapley, B., Bettadpur, S., Ries, J., Thompson, P. and Watkins, M. (2004). GRACE measurements of mass variability in the Earth system, *Science*, **305**, 503–505.
- Todisco, F., Vergni, L. and Mannocchi, F. (2009). Operative approach to agricultural drought risk management, *Journal of Irrigation and Drainage Engineering-ASCE*, **135**, 654–664.
- Trenberth, K. and Guillemot, C. (1996). Physical processes involved in the 1988 drought and 1993 floods in North America, *Journal of Climate*, **9**, 1288–1298.
- Vachaud, G., Passerat De Silans, A., Balabanis, P., and Vauclinm M. (1985). Temporal Stability of Spatially Measured Soil Water Probability Density Function, *Soil Science Society of America Journal*, **9**, 822–828.
- van den Hurk, B., Ettema, J. and Viterbo, P. (2008). Analysis of soil moisture changes in Europe during a single growing season in a new ECMWF soil moisture assimilation system, *Journal of Hydrometeorology*, **9**, 116–131.
- Vinnikov, K. and Yeserkepova, I. (1991). Soil moisture: Empirical data and model results, *Journal of Climate*, **4**, 66–79.
- Vinnikov, K., Robock, A., Qiu, S. and Entin, J. (1999). Optimal design of surface networks for observation of soil moisture, *Journal of Geophysical Research*, **104**, 19743–19749.
- Viterbo, P. P. C. (1995). The importance of soil water for medium-range weather forecasting: Implications for data assimilation, *Workshop on Imbalance of*

- Slowly Varying Components of Predictable Atmospheric Motions*, abstract, WMO, Beijing, China, March 1995.
- Wagner, W., Lemoine, G. and Rott, H. (1999). A method for estimating soil moisture from ERS scatterometer and soil data, *Remote Sensing of Environment*, **70**, 191–207.
- Wagner, W., Naeimi, V., Scipal, K., de Jeu, R. and Martínez-Fernández, J. (2007). Soil moisture from operational meteorological satellites, *Hydrogeology Journal*, **15**, 121–131.
- Wagner, W., Scipal, K., Boogard, H., van Diepen, K., Beck, R., Nobbe, E. and Savin, I. (2000). Assessing water-limited crop production with a scatterometer based crop growth monitoring system, *Geoscience and Remote Sensing Symposium 2000 Proceedings, IGARSS 2000*, **4**, 1696 – 1698.
- Wagner, W., Scipal, K., Pathe, C., Gerten, D., Lucht, W. and Rudolf, B. (2003). Evaluation of the agreement between the first global remotely sensed soil moisture data with model and precipitation data, *Journal of Geophysical Research*, **108**, 4611.
- Walker, J. and Houser, P. (2001). A methodology for initializing soil moisture in a global climate model: Assimilation of near-surface soil moisture observations, *Journal of Geophysical Research*, **106**, 11761 – 11774.
- Walker, J., Ursino, N., Grayson, R. and Houser, P. (2003). Australian root-zone soil moisture: Assimilation of remote sensing observations, *MODSIM 2003 International Congress on Modelling and Simulation*, pp. 380–386.
- Walker, J., Willgoose, G. and Kalma, J. (2001a). One-dimensional soil moisture profile retrieval by assimilation of near-surface measurements: A simplified soil moisture model and field application, *Journal of Hydrometeorology*, **2**, 356–373.
- Walker, J., Willgoose, G. and Kalma, J. (2001b). One-dimensional soil moisture profile retrieval by assimilation of near-surface observations: A comparison of retrieval algorithms, *Advances in Water Resources*, **24**, 631–650.
- Weymouth, G., Mills, G., Jones, D., Ebert, E. and Manton, M. (1999). A continental-scale daily rainfall analysis system, *Australian Meteorological Magazine*, **48**, 169–179.

- Wilker, H., Drusch, M., Seuffert, G. and Simmer, C. (2006). Effects of near-surface soil moisture profile on the assimilation of L-band microwave brightness temperature, *Journal of Hydrometeorology*, **7**, 433–442.
- Yevdjovich, V. (1964). Statistical and probability analysis of hydrologic data, part III: analysis of variance, covariance and time series. in *Handbook of Applied Hydrology, A Compendium of Water-Resources Technology*, pp. 8.43–8.67, McGraw-Hill Book Company, New York.
- Young, R., Walker, J., Yeoh, N., Smith, A., Ellett, K., Merlin, O. and Western, A. (2008). Soil moisture and meteorological observations from the Murrumbidgee catchment, Technical Report, 54 pp., Department of Civil and Environmental Engineering, The University of Melbourne, Melbourne, Australia.
- Zhang, H. and Frederiksen, C. S. (2003). Local and nonlocal impacts of soil moisture initialization on AGCM seasonal forecasts: A model sensitivity study, *Journal of Climate*, **16**, 2117 – 2137.
- Zhou, Y., McLaughlin, D. and Entekhabi, D. (2006). Assessing the performance of the ensemble Kalman filter for land surface data assimilation, *Monthly Weather Review*, **134**, 2128–2142.

Appendix A

Related assimilation methods

The EKF used in this thesis (described in Section 3.2) is closely related to a number of existing land surface analysis techniques, and the relationship to each is briefly reviewed in this Appendix. In summary, the EKF is derived from a variational method, initially suggested by Mahfouf (1991) for the assimilation of screen-level observations. Hess (2001) then simplified this method by linearising the forecast model over the assimilation window, enabling the calculation of a direct solution to the minimisation of the cost function (rather than the iterative approach taken by Mahfouf (1991)). With this linearisation the method resembles an EKF, and it is referred to as both a “simplified 2-D Var” (Balsamo et al., 2004) and a “simplified EKF” (Mahfouf et al., 2009). The similarity between the simplified variational approach and the EKF that is used here is discussed below. Additionally, the EKF as it is applied here, is also compared to the EKF assimilation of soil moisture in which the forward model is not used in the observation operator, since the latter is more commonly used for soil moisture assimilation (e.g., Walker and Houser, 2001; Reichle et al., 2002b; Muñoz Sabater et al., 2007)

A.1 Comparison to simplified 2-D Var

To examine the relationship between the simplified 2-D Var of Balsamo et al. (2004) and the EKF implemented here, consider the example of the ISBA model from Chapter 5, in which a two-layer soil moisture state is updated from observations of the upper-most layer. For the simplified 2-D Var of Balsamo et al. (2004) the model update is made at the beginning of an assimilation window, at t_i , based on the observations during the assimilation window, in this case

at the end of the assimilation window at time $t_i + 6$.¹ As outlined by Mahfouf (2010b) the simplified 2-D Var update is written:

$$\begin{aligned}\Delta\mathbf{x}(t_i) &= \mathbf{x}^a(t_i) - \mathbf{x}^b(t_i) \\ &= \mathbf{P}_0\mathbf{H}_i^T(\mathbf{H}_i\mathbf{P}_0\mathbf{H}_i^T + \mathbf{R}_i)^{-1}(\mathbf{y}^o(t_i + 6) - \mathcal{H}_i[\mathbf{x}^b(t_i)])\end{aligned}\quad (\text{A.1})$$

where \mathbf{P}_0 indicates a static \mathbf{P} matrix, describing the model errors at the beginning of each assimilation window. The observation operator in equation 3.6 can be written as $\mathcal{H}_i = \hat{\mathcal{H}}\mathcal{M}_{t_i \rightarrow t_i+6}$, where $\hat{\mathcal{H}} = \begin{pmatrix} 1 & 0 \end{pmatrix}$. For a 6 hour assimilation cycle, $\mathcal{M}_{t_i \rightarrow t_i+6}$ is \mathcal{M}_i . Substituting $\mathcal{H}_i = \hat{\mathcal{H}}\mathcal{M}_i$ into equation A.1 gives:

$$\begin{aligned}\Delta\mathbf{x}(t_i) &= \mathbf{P}_0(\hat{\mathbf{H}}\mathbf{M}_i)^T[(\hat{\mathbf{H}}\mathbf{M}_i)\mathbf{P}_0(\hat{\mathbf{H}}\mathbf{M}_i)^T + \mathbf{R}]^{-1} \\ &\quad \times (\mathbf{y}^o(t_i + 6) - \hat{\mathcal{H}}\mathcal{M}_i[\mathbf{x}^b(t_i)])\end{aligned}\quad (\text{A.2})$$

Applying the forecast model to equation A.2 carries the update forward to the end of the assimilation window:

$$\begin{aligned}\Delta\mathbf{x}(t_i + 6) &\approx \mathbf{M}_i\Delta\mathbf{x}(t_i) \\ &= (\mathbf{M}_i\mathbf{P}_0\mathbf{M}_i^T)\hat{\mathbf{H}}^T[\hat{\mathbf{H}}(\mathbf{M}_i\mathbf{P}_0\mathbf{M}_i^T)\hat{\mathbf{H}}^T + \mathbf{R}]^{-1} \\ &\quad \times (\mathbf{y}^o(t_i + 6) - \hat{\mathcal{H}}[\mathbf{x}^b(t_i + 6)])\end{aligned}\quad (\text{A.3})$$

The use of the forecast model in the observation operator can be interpreted as evolving \mathbf{P}_0 to the time of the observations, as is done by equation 3.4 for an EKF, giving the similarity between the simplified 2-D Var and the EKF. Hence, while Balsamo et al. (2004) refers to this method as a simplified 2-D Var, where ‘‘simplified’’ refers to the use of a linearisation of the forecast model by finite-differencing, in place of a model adjoint, Mahfouf et al. (2009) refers to it as a simplified EKF, where ‘‘simplified’’ refers to the use of static \mathbf{P}_0 at each update time. In this thesis, the full EKF is applied, and equations 3.4 and 3.5 are used in place of \mathbf{P}_0 in the above, and so the EKF terminology is adopted here.

¹Note that if observations were prior to the end of the assimilation window, the two techniques would differ, since a variational smoother ingests all of the observations in one analysis, while a Kalman filter assimilates them sequentially.

A.2 Comparison to standard EKF for assimilating near-surface soil moisture data

For the assimilation of soil moisture observations, equation A.3 raises another approach to implementing an EKF, which is to make the model update at the time of the observations and use $\hat{\mathcal{H}} = \begin{pmatrix} 1 & 0 \end{pmatrix}$ as the observation operator. This is a much more common method for the EKF assimilation of soil moisture (e.g., Walker and Houser, 2001; Reichle et al., 2002b; Muñoz Sabater et al., 2007) than that used here, and it has the advantage of avoiding the additional integrations required to linearise the observation operator. However it relies on the observations being included in the state update vector, and the screen-level variables cannot be sensibly included in the update vector of the off-line land surface models used in these experiments since they are not prognostic (and are diagnosed by interpolating between the surface humidity and temperature, and the forcing fields for the first atmospheric model layer). Hence, this approach cannot assimilate screen-level observations into the off-line land surface models, and so has not been used here. However, for the assimilation of soil moisture data, examining the analysis equations indicates that these two approaches are very similar. To demonstrate this, consider the update equation at time $t_i + 6$ for the assimilation with an observation operator of $\hat{\mathcal{H}} = \begin{pmatrix} 1 & 0 \end{pmatrix}$:

$$\begin{aligned} \Delta \mathbf{x}(t_i + 6) &= \mathbf{P}^f(t_i + 6) \hat{\mathbf{H}}^T [\hat{\mathbf{H}} \mathbf{P}^f(t_i + 6) \hat{\mathbf{H}}^T + \mathbf{R}]^{-1} \\ &\quad \times (\mathbf{y}^o(t_i + 6) - \hat{\mathcal{H}}[\mathbf{x}^b(t_i + 6)]) \end{aligned} \quad (\text{A.4})$$

For the EKF used here, the analysis update can be obtained by substituting $\mathbf{P}^f(t_i)$ from equation 3.4 in place of \mathbf{P}_0 in equation A.3. This gives the EKF analysis update at time t_i carried forward 6 hours to $t_i + 6$:

$$\begin{aligned} \Delta \mathbf{x}(t_i + 6) &\approx (\mathbf{M}_i \mathbf{P}^f(t_i) \mathbf{M}_i^T) \hat{\mathbf{H}}^T [\hat{\mathbf{H}} (\mathbf{M}_i \mathbf{P}^f(t_i) \mathbf{M}_i^T) \hat{\mathbf{H}}^T + \mathbf{R}]^{-1} \\ &\quad \times (\mathbf{y}^o(t_i + 6) - \hat{\mathcal{H}}[\mathbf{x}^b(t_i + 6)]) \end{aligned} \quad (\text{A.5})$$

The equations describing the analysis update from the observations at time $t_i + 6$ are very similar. In equation A.4 the update is made at the time of the observations, and \mathbf{P} is the forecast value at the time of the observations,

obtained from equation 3.4 (i.e., $\mathbf{P} = \mathbf{M}_i \mathbf{P}^a(t_i) \mathbf{M}_i^T + \mathbf{Q}(t_i)$) In contrast, for equation A.5 the update is made 6 hours before the observations, and \mathbf{P} is the forecast value at time t_i (from equation 3.4) carried forward another 6 hours to $t_i + 6$ (i.e., $\mathbf{P} = \mathbf{M}_i (\mathbf{M}_{i-1} \mathbf{P}^a(t_{i-1}) \mathbf{M}_{i-1}^T + \mathbf{Q}(t_{i-1})) \mathbf{M}_i$). However, once equation A.5 has been carried forward 6 hours, the analysis updates differ only in the value of \mathbf{P} used. The similarity between the two approaches has been confirmed by comparing the analysis results generated by the two methods for the experimental set-up used in Chapter 5. The differences are limited to the magnitude of the analysis increments, with the increments being larger (yet showing the same spatial pattern) when the model is used as the observation operator.

Appendix B

Comparing coupled MOSES and off-line JULES simulations

JULES was launched in October 2006, with the same programming as was used in the UM MOSES model at that time, however the Met Office updates the UM code four times a year, so that the MOSES code in ACCESS now differs from that in JULES. This Appendix outlines the changes that have been made to the JULES model for use in this thesis so that it better mimics the coupled MOSES model. Additionally, it is shown that the off-line JULES model with these changes represents a reasonably accurate representation of the coupled MOSES model.

Two significant updates have been made to the MOSES code since the JULES model was launched (pers. comm. Imtiaz Dharssi):

- In June 2007 (PS15) the soil hydrology was updated so that excess soil moisture generated from a frozen or saturated soil layer flows into the next soil layer down, rather than being lost as runoff. This change had the greatest impact under snow, and after short periods of intense rain.
- In April 2008 (PS18) several changes were made to MOSES. A long-standing error in the calculation of soil hydraulic properties was fixed, and several other changes, including the introduction of soil temperature nudging and assimilation of SYNOP data, were introduced. These changes were a response to the poor performance of the UM over the European summer of 2006, and they had a positive impact on the atmospheric forecast skill, including a significant reduction of errors in the screen level temperature and humidity forecasts (Dharssi et al., 2009).

252 B. Comparing coupled MOSES and off-line JULES simulations

To take advantage of the improved model skill from the PS18 update, the experiments conducted in this thesis have been restricted to the period after April 2008, and the JULES code has been amended to include the PS15 and PS18 updates. Two additional model changes were also made in response to model short-comings accounted in this research:

- A bug in the calculation of bare soil evaporation was amended, to decrease the occurrence of strong nonlinearities in the model. In the operational MOSES model, bare soil evaporation occurs only when the nonsurface soil layers are above the wilting point. This has been amended so that bare soil evaporation depends only on the layer one soil moisture, and can continue when the nonsurface layers are below the wilting point. This change will be introduced to the operational MOSES code in the near future (pers. comm. Imtiaz Dharsasi).
- A lower bound of 10% of the wilting point has been placed on the layer one soil moisture, to prevent model instability when the layer one soil moisture is very dry. In the ACCESS suite the same bound is applied by the soil moisture analysis scheme (pers. comm. Jin Lee).

To check that the off-line MOSES model environment described above provides a reasonable representation of the MOSES model within ACCESS, the forecast soil moisture and temperature from each model have been compared over one diurnal cycle, from 7 UTC on 1 November 2008. There will be differences between the surface state forecast by the off-line and coupled version of MOSES, due to:

- differences between the atmospheric forcing applied to each due to the use of hourly fields from the atmospheric model (rather than generating an atmospheric forecast for every model time-step)
- differences between the model code, both due to issues associated with decoupling MOSES from the atmospheric model, or due to unaccounted for changes to MOSES after JULES was launched
- rounding off errors in writing out fields from the coupled model and reading them into off-line land surface model

The difference between the soil moisture forecast by the off-line and coupled MOSES models decreases rapidly with depth and only the surface layer is discussed here (the differences decrease by an order or magnitude for each of the

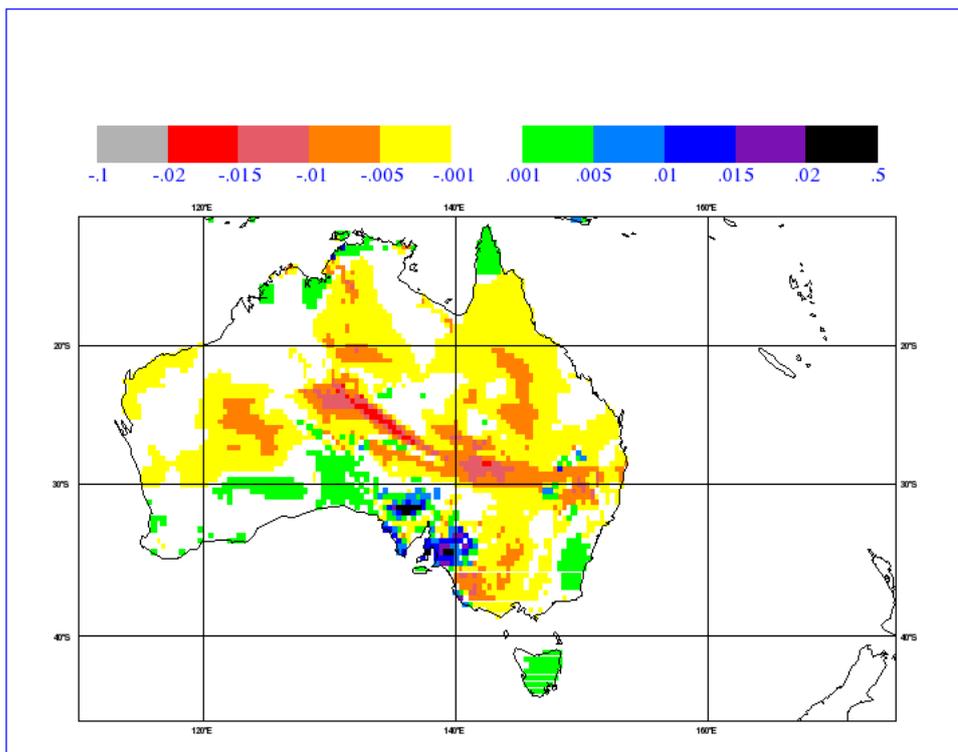


Figure B.1: Difference between the layer one soil moisture (as a fraction of saturation) from the off-line and coupled MOSES models (off-line minus coupled), for the 24-hour forecast to 07:00 UTC on 2 November, 2008.

under-lying soil layers). Figure B.1 shows the difference in the layer one soil moisture between the 24-hour forecasts (to 07:00 UTC on 2 November 2008) from the off-line and coupled MOSES forecasts. The difference in layer one ranges from -1.8% to 2.9% of the soil moisture at saturation, with a mean of -0.17%. Less than 4% of grids have an absolute difference greater than 1% of the soil moisture at saturation (equivalent to about $0.005 \text{ m}^3 \text{ m}^{-3}$). Figure B.2 show the large-scale and convective precipitation over the same time period. Comparison to Figure B.1 highlights that the greatest differences between the off-line and coupled soil moisture forecasts occur in precipitating regions, particularly with convective precipitation. These differences are due to the use of hourly mean precipitation in the off-line system, since infiltration of moisture into the soil is sensitive to the timing of the precipitation (although the separation of precipitation into large-scale and convective fractions has reduced this difference).

The soil temperature forecasts from the off-line and coupled models show greater differences than the soil moisture forecasts, and the difference in soil

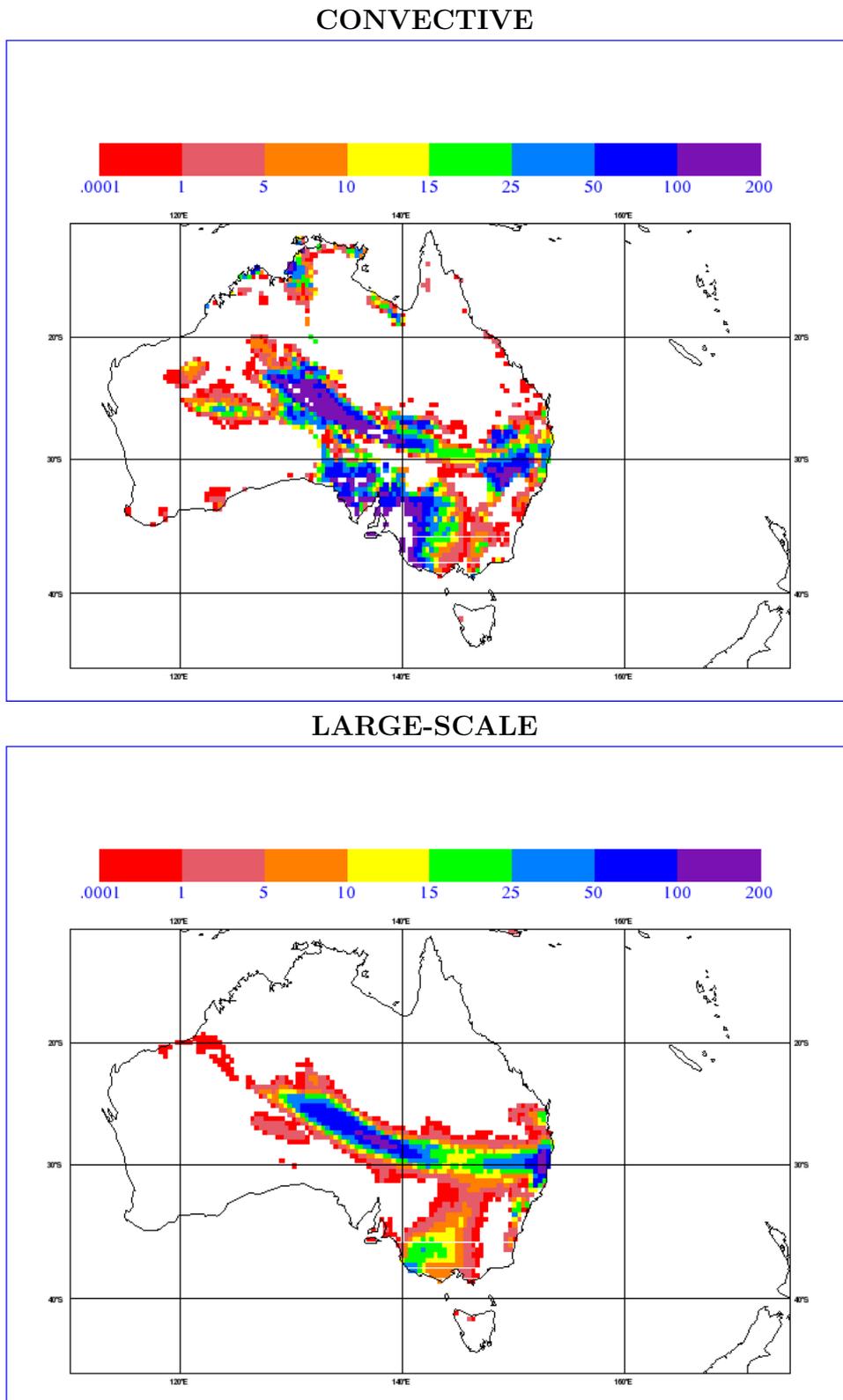


Figure B.2: ACCESS forecast convective (upper) and large-scale (lower) rain-fall (mm/day) for the 24-hours to 07:00 UTC on 2 November, 2008.

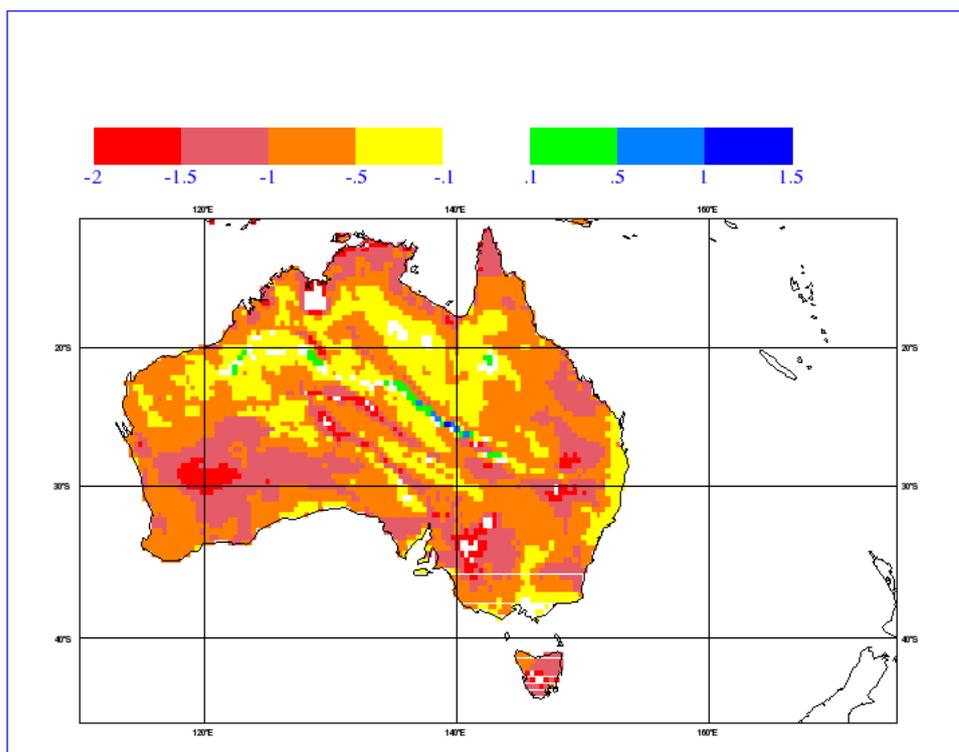


Figure B.3: Difference between the layer one soil temperature from the off-line and coupled MOSES models (off-line minus coupled), for the 24-hour forecast to 07:00 UTC on 2 November, 2008.

temperature ranges from -3.9 K to 1.0 K, with a mean of -0.8 K. As with the moisture, the temperature differences decrease with depth, with the mean reducing to -0.1 in the second layer (and less than 0.001 K in the third and fourth layers). Figure B.3 shows the difference in the layer one soil temperature differences in the 24-hour forecasts: there is a consistent tendency across Australia for the off-line forecast to be cooler than the coupled forecasts, usually by less than 1 K. Time series plots (not shown) indicate that these differences evolve gradually throughout the forecast, particularly during periods of heating. The reduction with depth suggests that the differences are driven by the forcing (although the width of the soil layers get deeper with distance from the surface, so that energy changes are dispersed across a greater volume), and the discrepancies may be due to use of hourly radiation means in the forcing data. While this discrepancy would need to be further investigated before the off-line land surface analysis is coupled to the ACCESS model, the off-line system is considered to be close enough to the coupled simulation for use in the (fully decoupled) experiments conducted here.

256 B. Comparing coupled MOSES and off-line JULES simulations

Appendix C

Statistics of fit between in situ data for non-surface layers at Kyeamba

Table C.1: RMSD, correlation, and anomaly correlation between layer 2 soil moisture from the individual monitoring stations in the Kyeamba-A pixel, for April '08 to April '09. For each statistic the minimum and maximum off-diagonal values are in bold.

	RMSD				
	K1	K2	K3	K5	K7
K1	0.000	-	-	-	-
K2	0.034	0.000	-	-	-
K3	0.034	0.036	0.000	-	-
K5	0.039	0.029	0.022	0.000	-
K7	0.032	0.025	0.030	0.022	0.000

	r_{abs}						r_{anm}				
	K1	K2	K3	K5	K7		K1	K2	K3	K5	K7
K1	1.00	-	-	-	-	K1	1.00	-	-	-	-
K2	0.84	1.00	-	-	-	K2	0.78	1.00	-	-	-
K3	0.84	0.98	1.00	-	-	K3	0.85	0.79	1.00	-	-
K5	0.85	0.97	0.98	1.00	-	K5	0.85	0.69	0.89	1.00	-
K7	0.89	0.93	0.93	0.95	1.00	K7	0.55	0.45	0.67	0.68	1.00

Table C.2: RMSD, correlation, and anomaly correlation between layer 3 soil moisture from the individual monitoring stations in the Kyeamba-A pixel, for April '08 to April '09. For each statistic the minimum and maximum off-diagonal values are in bold.

	RMSD				
	K1	K2	K3	K5	K7
K1	0.000	-	-	-	-
K2	0.015	0.000	-	-	-
K3	0.094	0.089	0.000	-	-
K5	0.115	0.111	0.025	0.000	-
K7	0.209	0.207	0.120	0.100	0.000

	r_{abs}						r_{anm}				
	K1	K2	K3	K5	K7		K1	K2	K3	K5	K7
K1	1.00	-	-	-	-	K1	1.00	-	-	-	-
K2	0.90	1.00	-	-	-	K2	0.89	1.00	-	-	-
K3	0.80	0.97	1.00	-	-	K3	0.63	0.82	1.00	-	-
K5	0.86	0.98	0.97	1.00	-	K5	0.87	0.83	0.60	1.00	-
K7	0.73	0.88	0.90	0.92	1.00	K7	-0.41	-0.35	-0.10	-0.36	1.00

Table C.3: RMSD, correlation, and anomaly correlation between layer 4 soil moisture from the individual monitoring stations in the Kyeamba-A pixel, for April '08 to April '09. For each statistic the minimum and maximum off-diagonal values are in bold.

	RMSD				
	K1	K2	K3	K5	K7
K1	0.000	-	-	-	-
K2	0.020	0.000	-	-	-
K3	0.075	0.062	0.000	-	-
K5	0.119	0.108	0.050	0.000	-
K7	0.201	0.193	0.136	0.087	0.000

	r_{abs}						r_{anm}				
	K1	K2	K3	K5	K7		K1	K2	K3	K5	K7
K1	1.00	-	-	-	-	K1	1.00	-	-	-	-
K2	0.88	1.00	-	-	-	K2	-0.11	1.00	-	-	-
K3	0.85	0.99	1.00	-	-	K3	-0.15	0.91	1.00	-	-
K5	0.87	0.99	0.98	1.00	-	K5	-0.14	0.97	0.89	1.00	-
K7	0.88	0.91	0.90	0.89	1.00	K7	0.88	-0.39	-0.35	-0.39	1.00

Table C.4: RMSD, correlation, and anomaly correlation between the root-zone soil moisture from the individual monitoring stations in the Kyeamba-A pixel, for April '08 to April '09. For each statistic the minimum and maximum off-diagonal values are in bold.

	RMSD				
	K1	K2	K3	K5	K7
K1	0.000	-	-	-	-
K2	0.017	0.000	-	-	-
K3	0.056	0.055	0.000	-	-
K5	0.069	0.070	0.017	0.000	-
K7	0.123	0.127	0.077	0.062	0.000

	r_{abs}						r_{anm}				
	K1	K2	K3	K5	K7		K1	K2	K3	K5	K7
K1	1.00	-	-	-	-	K1	1.00	-	-	-	-
K2	0.91	1.00	-	-	-	K2	0.88	1.00	-	-	-
K3	0.89	0.99	1.00	-	-	K3	0.92	0.91	1.00	-	-
K5	0.90	0.99	0.99	1.00	-	K5	0.88	0.92	0.93	1.00	-
K7	0.91	0.97	0.98	0.97	1.00	K7	0.58	0.39	0.57	0.48	1.00

Appendix D

Additional time series plots
comparing original AMSR-E
retrieval algorithms to in situ
data.

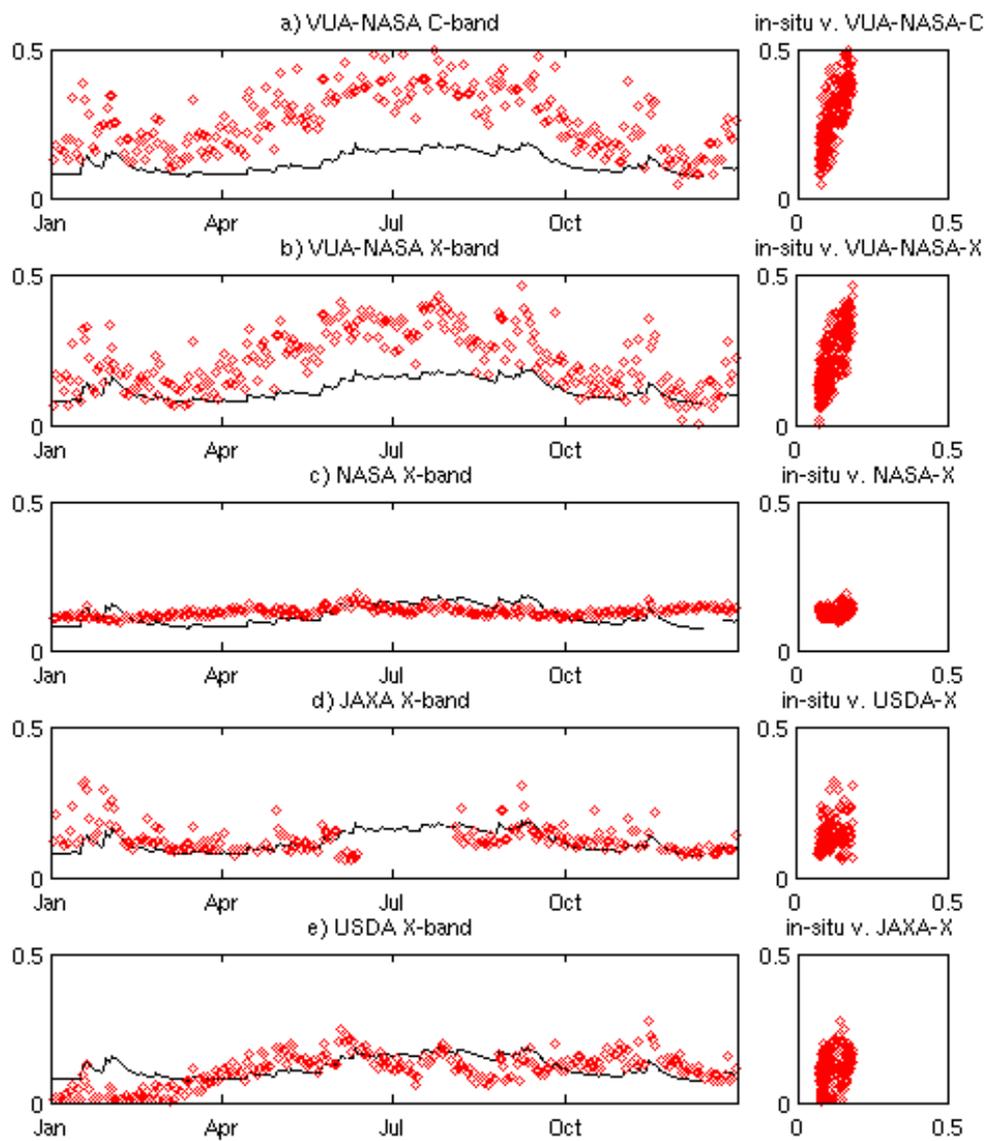


Figure D.1: Comparison of in-situ (solid lines) and original AMSR-E (red diamonds) near-surface soil moisture (m^3m^{-3}) for each retrieval algorithm, at Cooma Airfield (M1) over 2006.

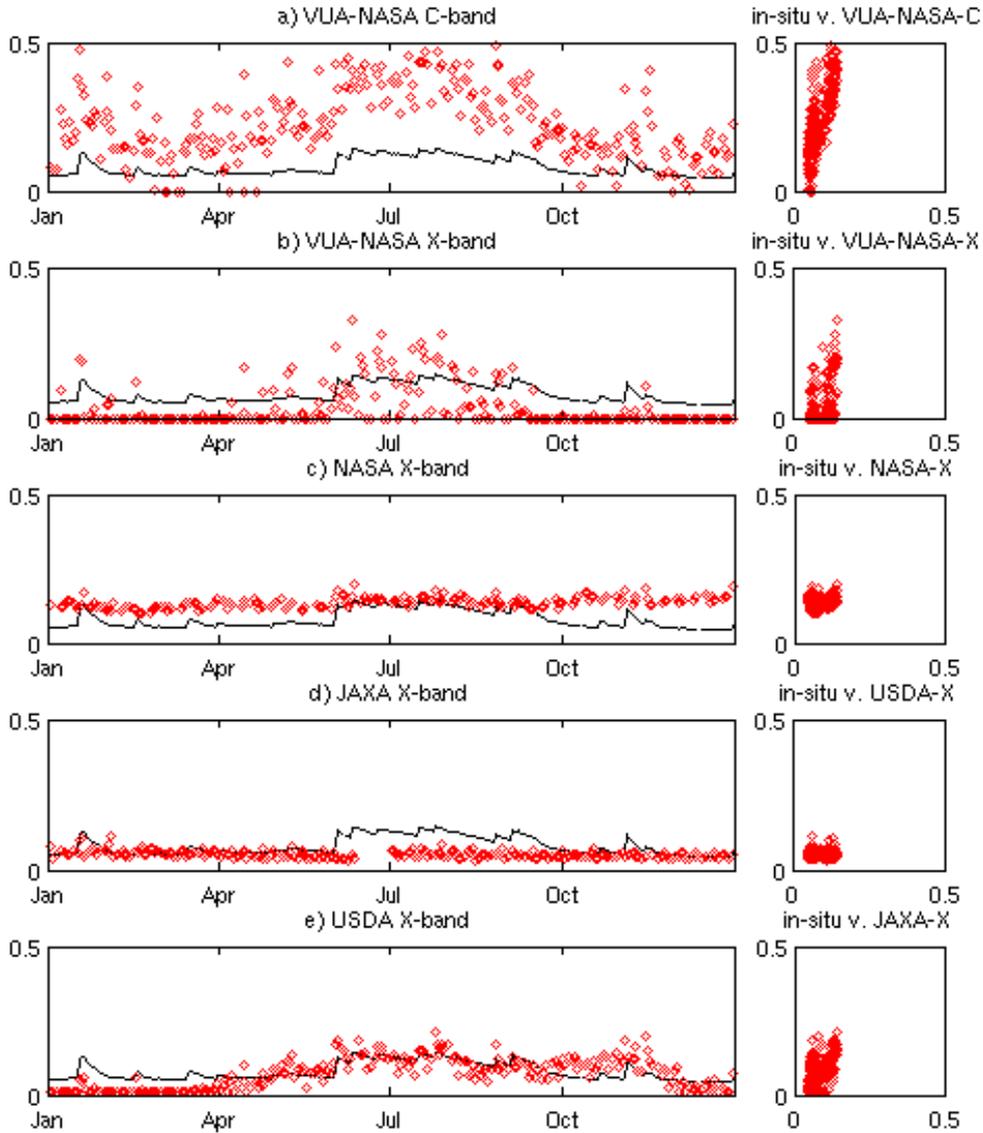


Figure D.2: Comparison of in-situ (solid lines) and original AMSR-E (red diamonds) near-surface soil moisture (m^3m^{-3}) for each retrieval algorithm, at Canberra Airport (M2) over 2006.

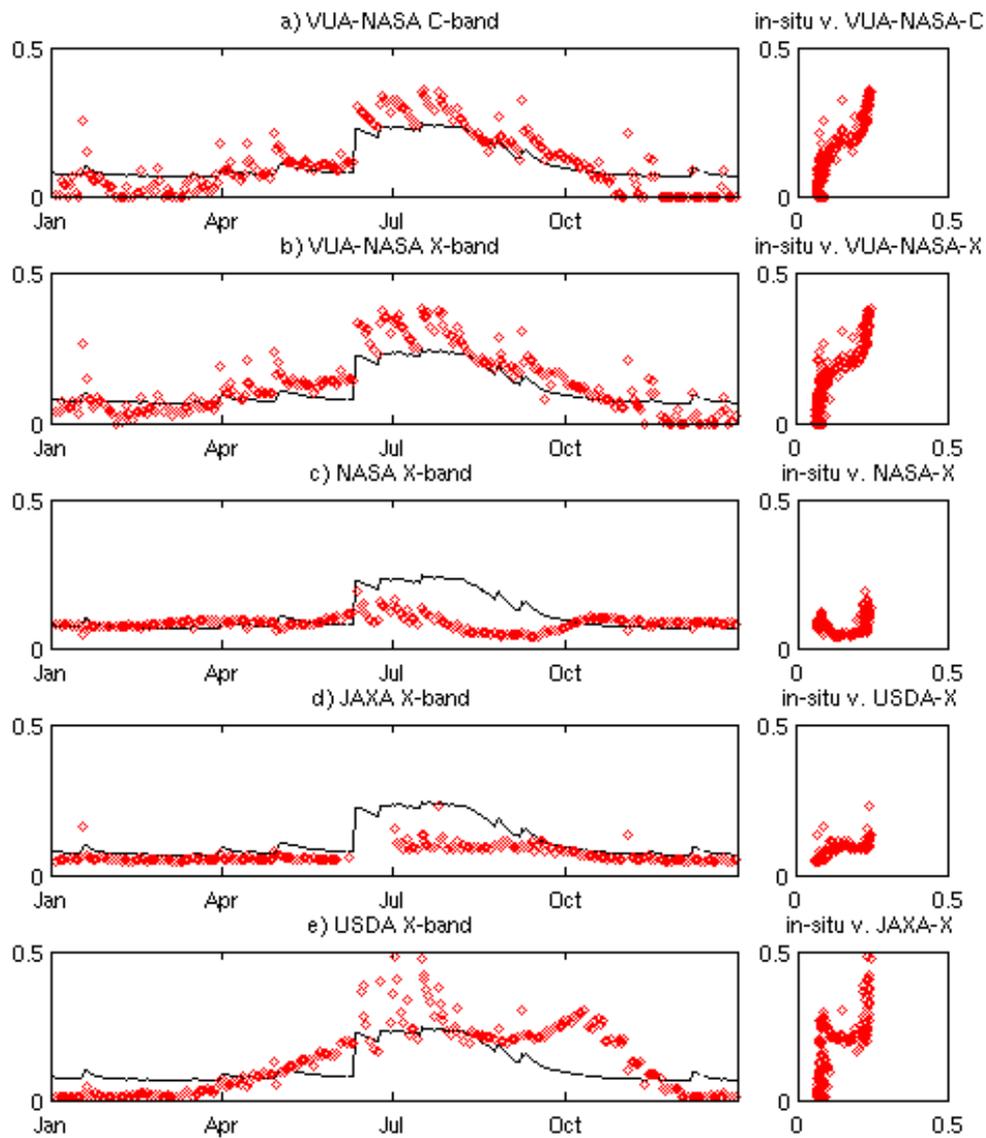


Figure D.3: Comparison of in-situ (solid lines) and original AMSR-E (red diamonds) near-surface soil moisture (m^3m^{-3}) for each retrieval algorithm, at West Wyalong Airfield (M4) over 2006.

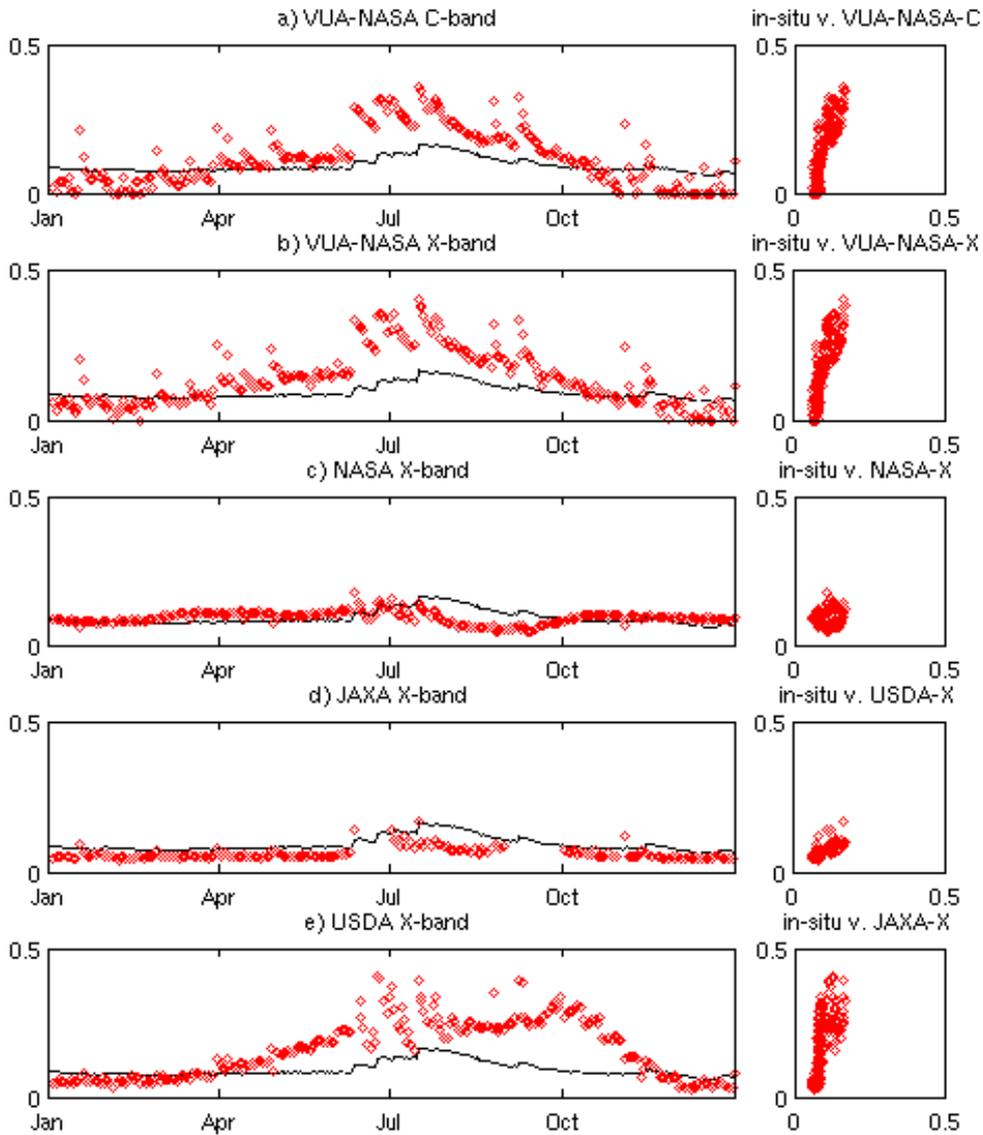


Figure D.4: Comparison of in-situ (solid lines) and original AMSR-E (red diamonds) near-surface soil moisture ($m^3 m^{-3}$) for each retrieval algorithm, at Balranald (M5) over 2006.

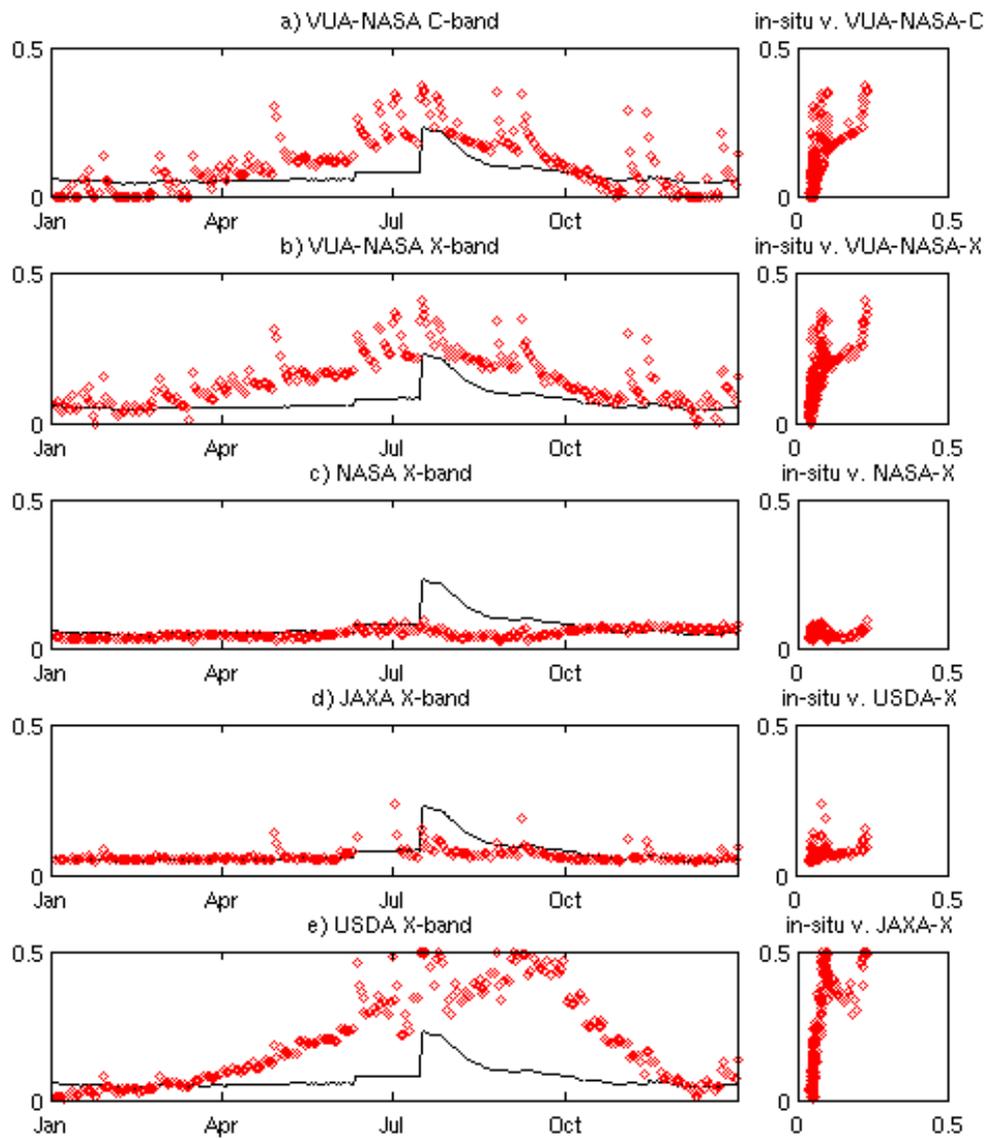


Figure D.5: Comparison of in-situ (solid lines) and original AMSR-E (red diamonds) near-surface soil moisture (m^3m^{-3}) for each retrieval algorithm, at Hay (M6) over 2006.

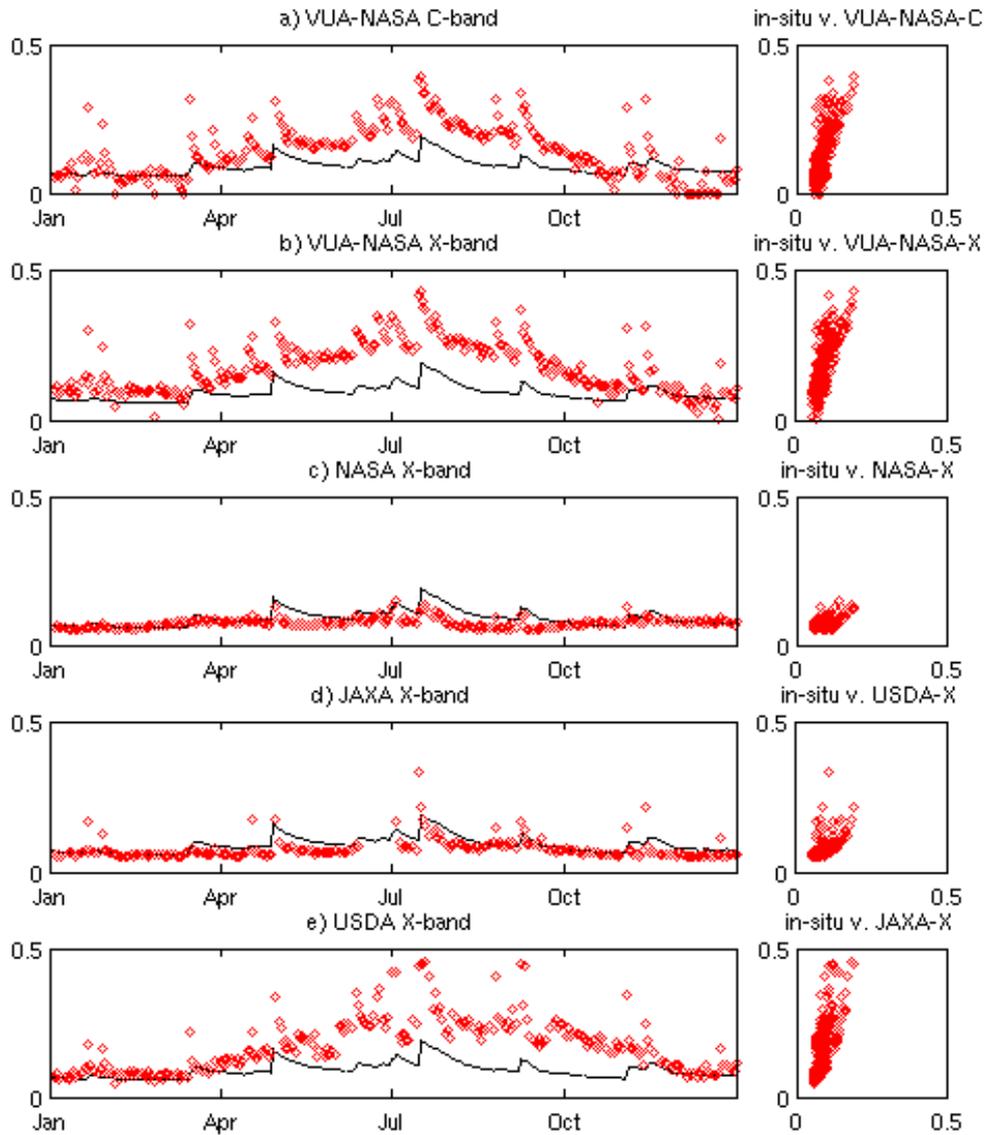


Figure D.6: Comparison of in-situ (solid lines) and original AMSR-E (red diamonds) near-surface soil moisture ($m^3 m^{-3}$) for each retrieval algorithm, at Griffith Aerodrome (M7) over 2006.

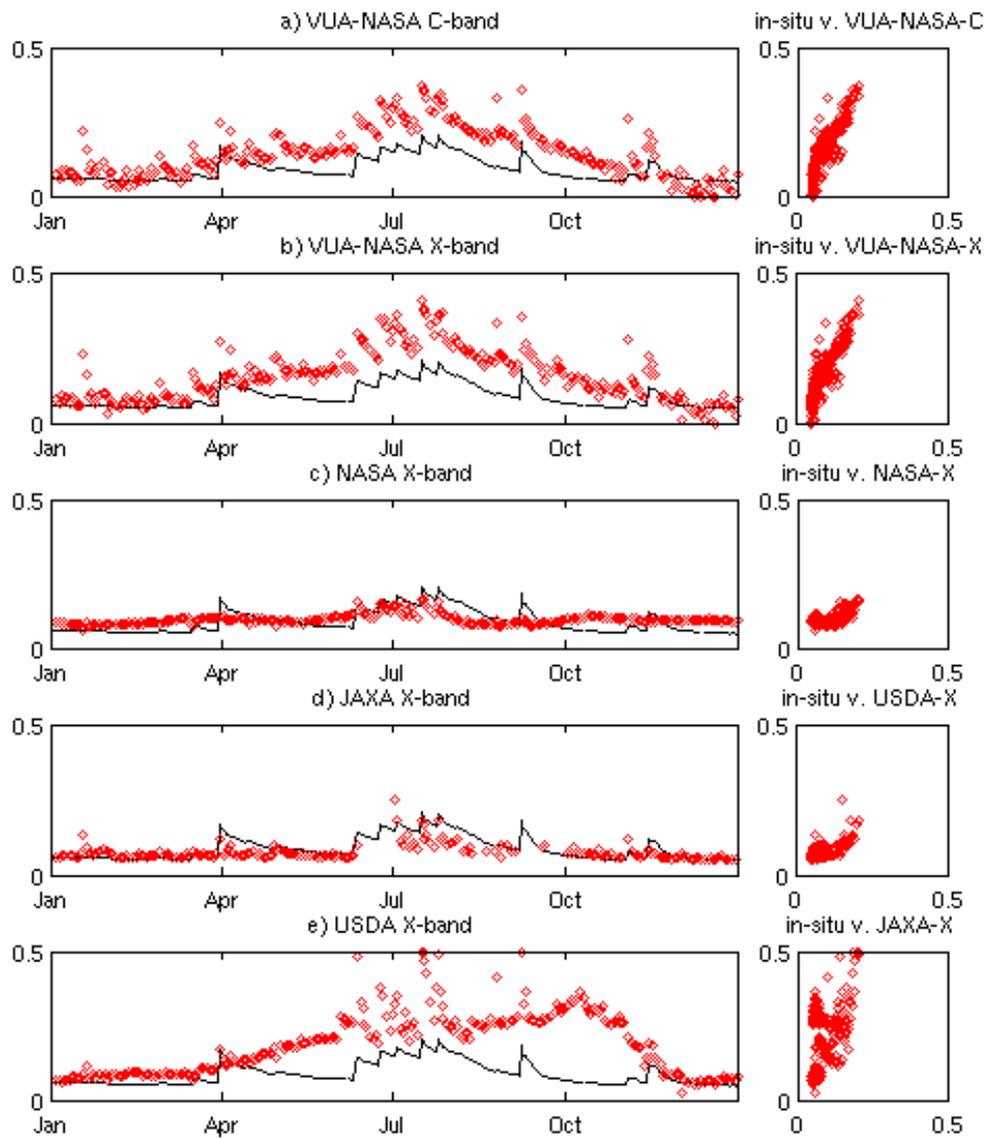


Figure D.7: Comparison of in-situ (solid lines) and original AMSR-E (red diamonds) near-surface soil moisture (m^3m^{-3}) for each retrieval algorithm, at Yanco (M8) over 2006.

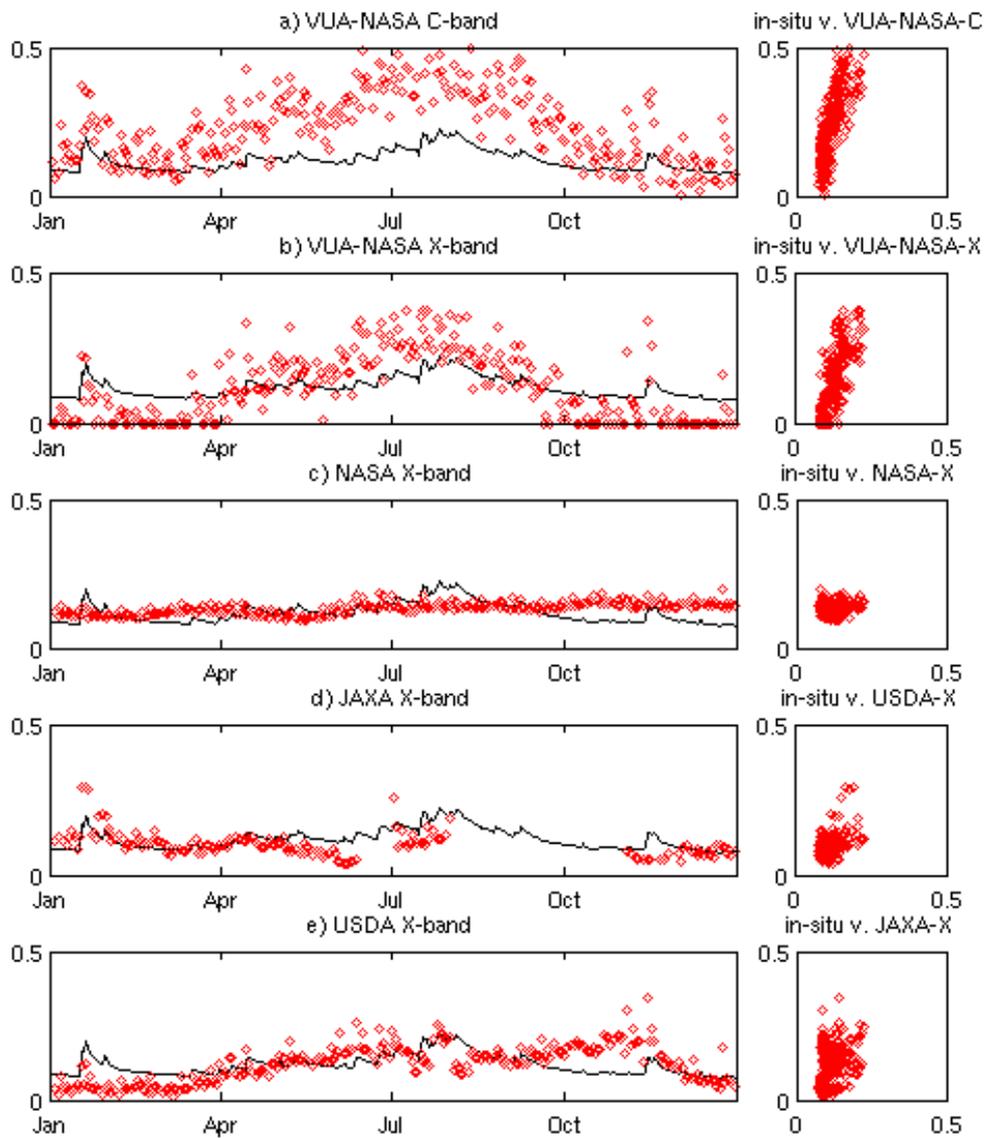


Figure D.8: Comparison of in-situ (solid lines) and original AMSR-E (red diamonds) near-surface soil moisture ($m^3 m^{-3}$) for each retrieval algorithm, at Adelong (M10) over 2006.

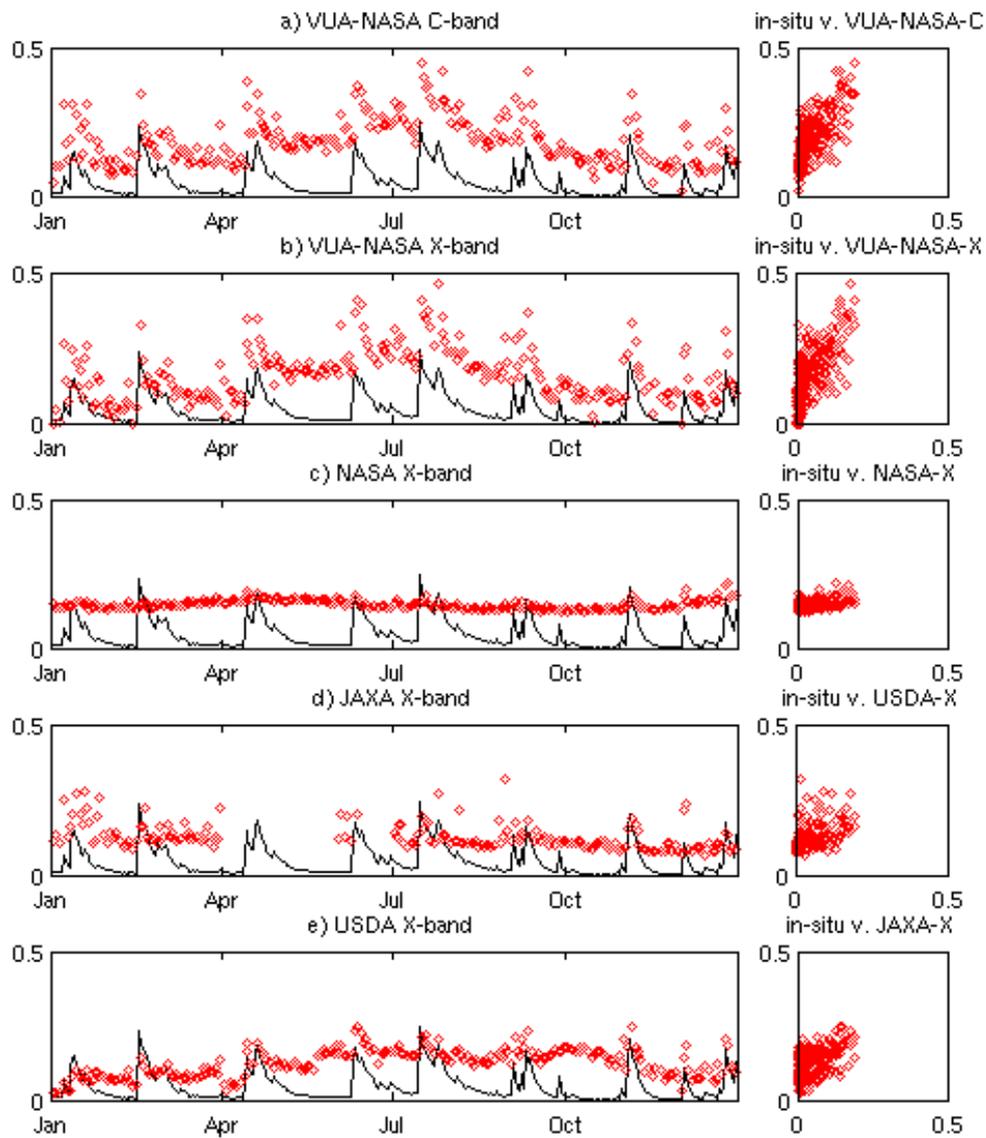


Figure D.9: Comparison of in-situ (solid lines) and original AMSR-E (red diamonds) near-surface soil moisture (m^3m^{-3}) for each retrieval algorithm, at Merriwa (G1) over 2006.

Appendix E

Additional time series plots
comparing normalised AMSR-E
retrieval algorithms to in situ
data.

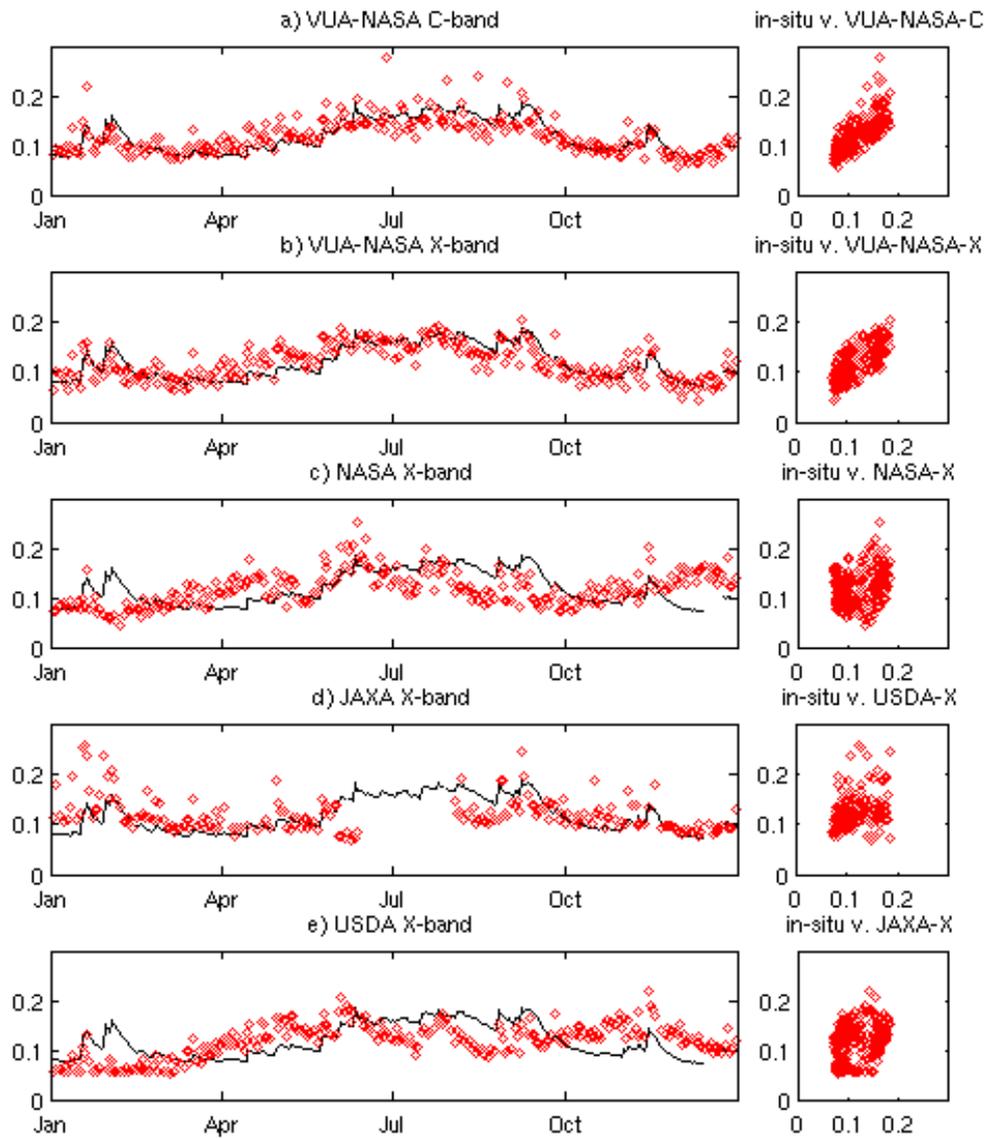


Figure E.1: Comparison of in-situ (solid lines) and normalised AMSR-E (red diamonds) near-surface soil moisture ($m^3 m^{-3}$) for each retrieval algorithm, at Cooma Airfield (M1) over 2006.

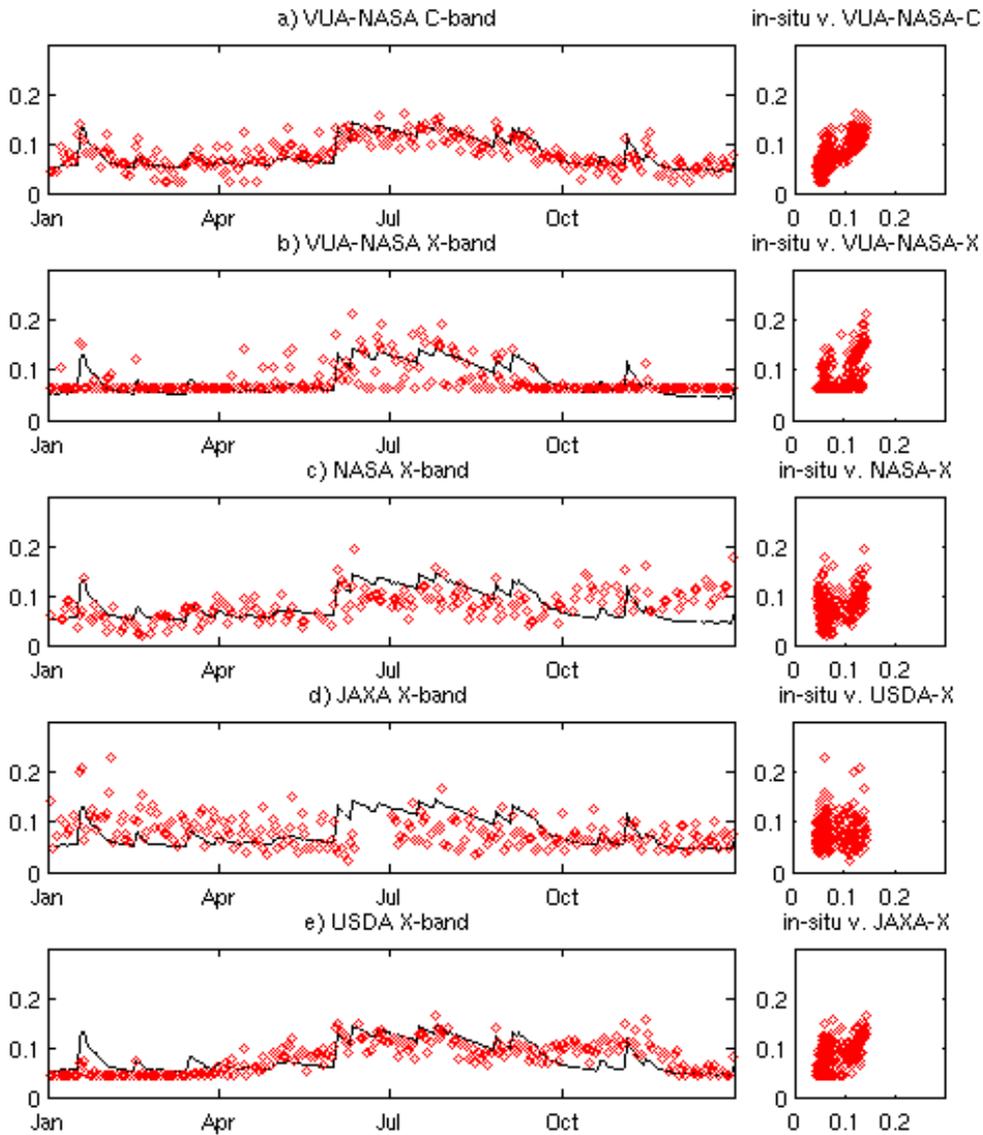


Figure E.2: Comparison of in-situ (solid lines) and normalised AMSR-E (red diamonds) near-surface soil moisture (m^3m^{-3}) for each retrieval algorithm, at Canberra Airport (M2).

274 E. Additional time series plots comparing normalised AMSR-E retrieval algorithms to in situ data.

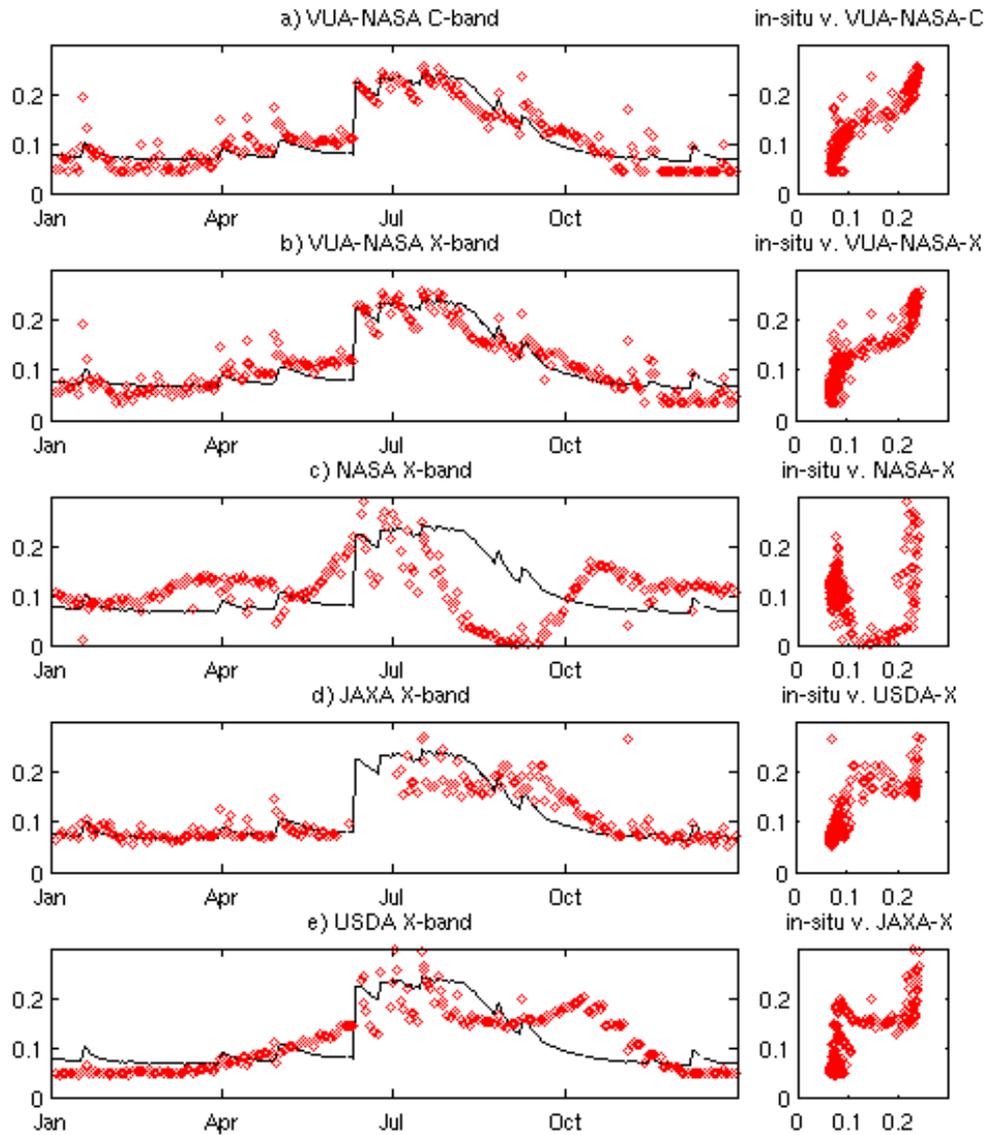


Figure E.3: Comparison of in-situ (solid lines) and normalised AMSR-E (red diamonds) near-surface soil moisture (m^3m^{-3}) for each retrieval algorithm, at West Wyalong Airfield (M4) over 2006.

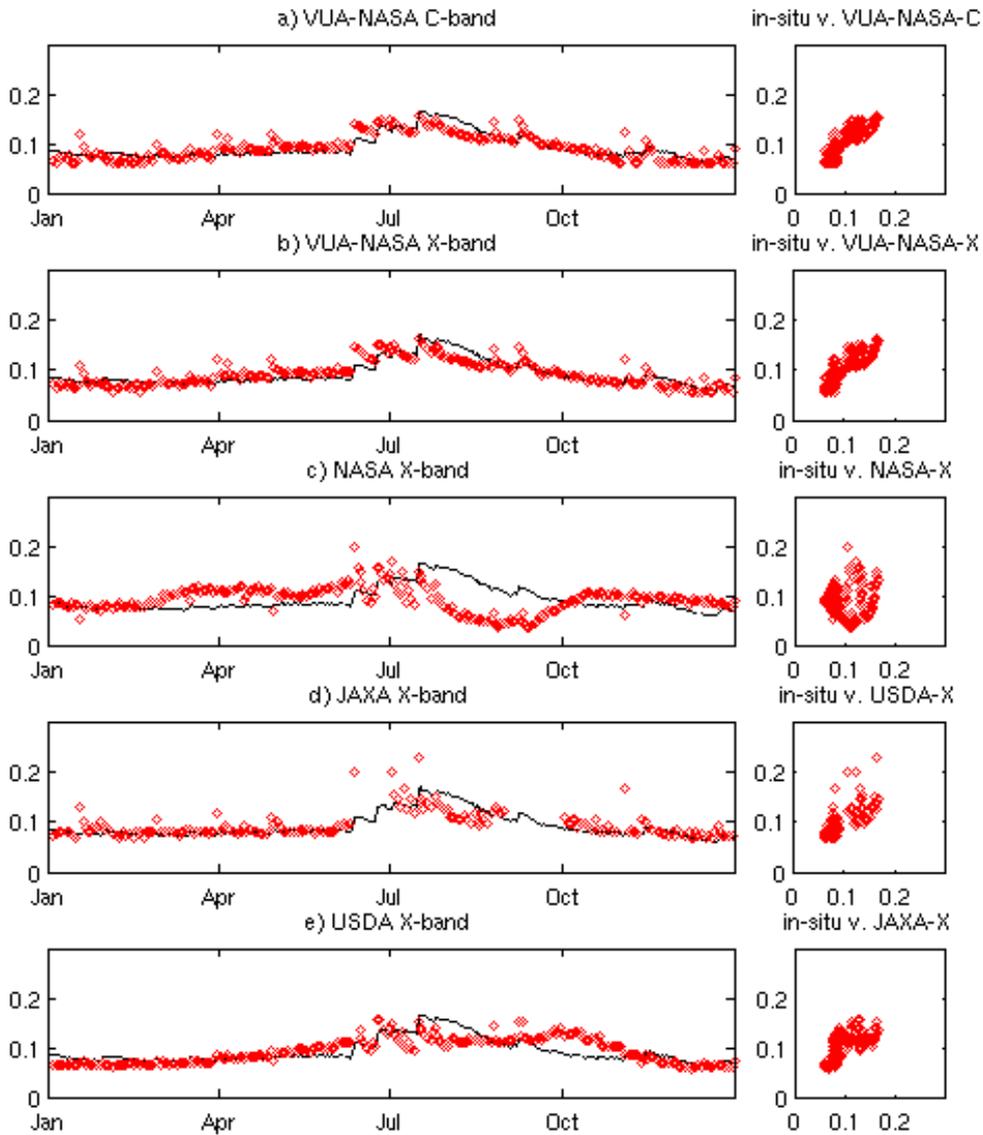


Figure E.4: Comparison of in-situ (solid lines) and normalised AMSR-E (red diamonds) near-surface soil moisture (m^3m^{-3}) for each retrieval algorithm, at Balranald (M5) over 2006.

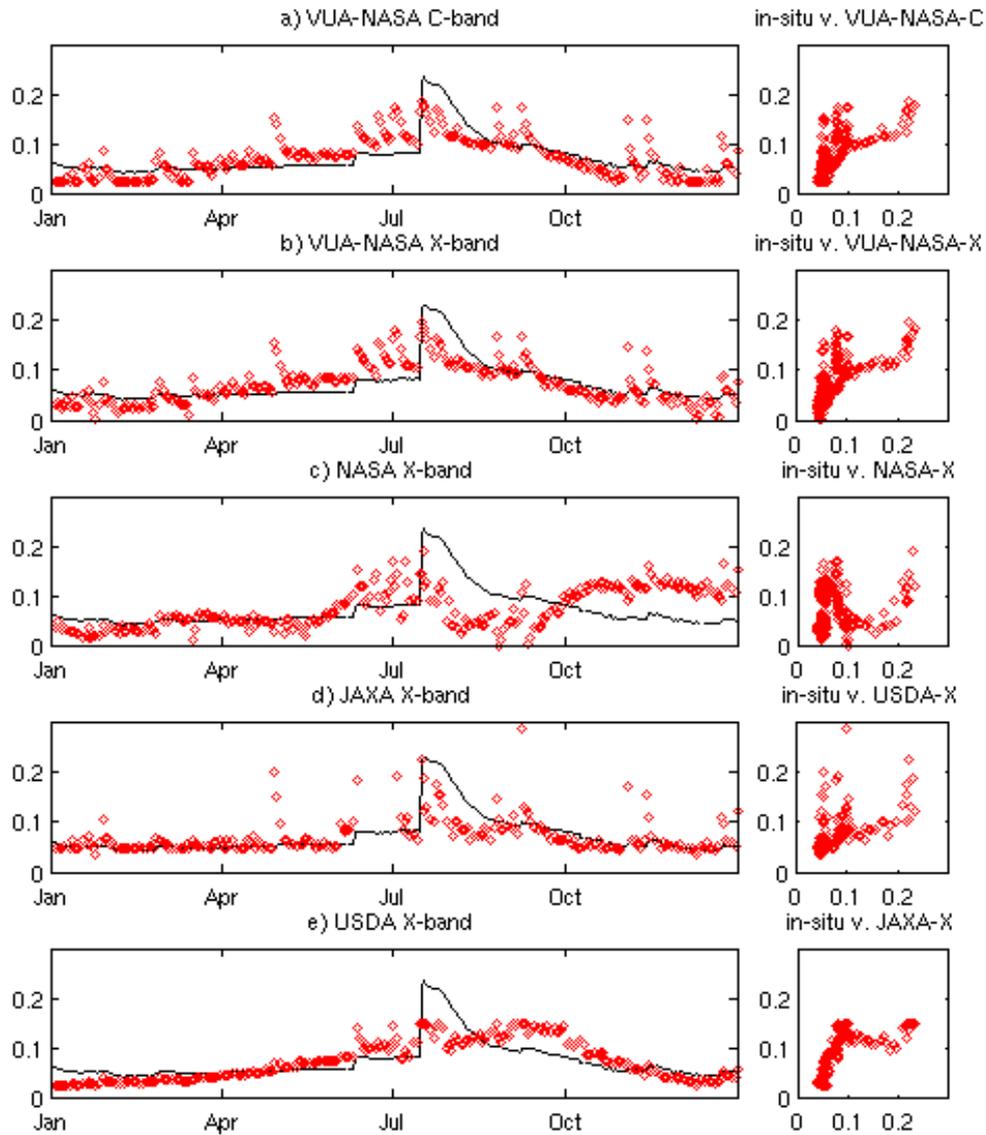


Figure E.5: Comparison of in-situ (solid lines) and normalised AMSR-E (red diamonds) near-surface soil moisture (m^3m^{-3}) for each retrieval algorithm, at Hay (M6) over 2006.

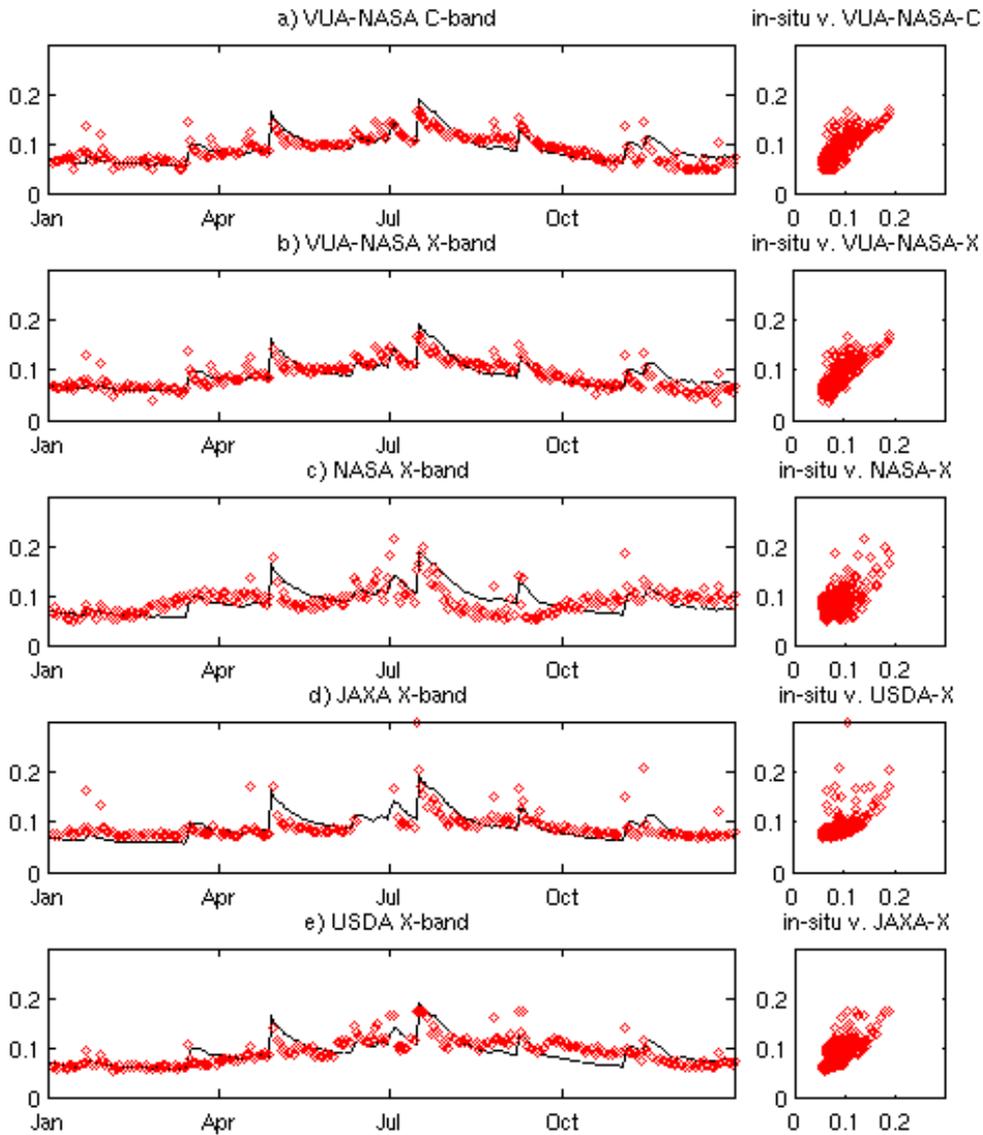


Figure E.6: Comparison of in-situ (solid lines) and normalised AMSR-E (red diamonds) near-surface soil moisture ($m^3 m^{-3}$) for each retrieval algorithm, at Griffith Aerodrome (M7) over 2006.

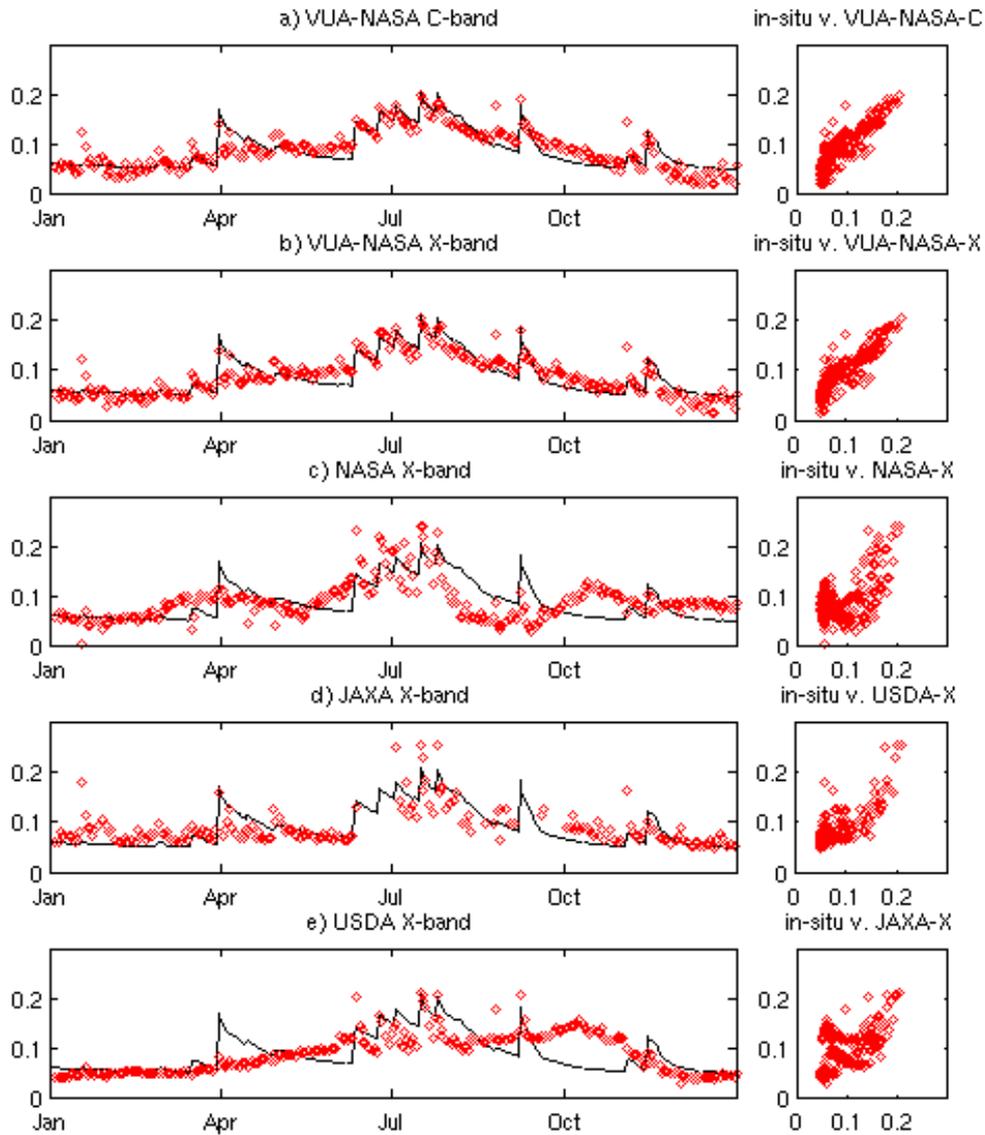


Figure E.7: Comparison of in-situ (solid lines) and normalised AMSR-E (red diamonds) near-surface soil moisture (m^3m^{-3}) for each retrieval algorithm, at Yanco (M8) over 2006.

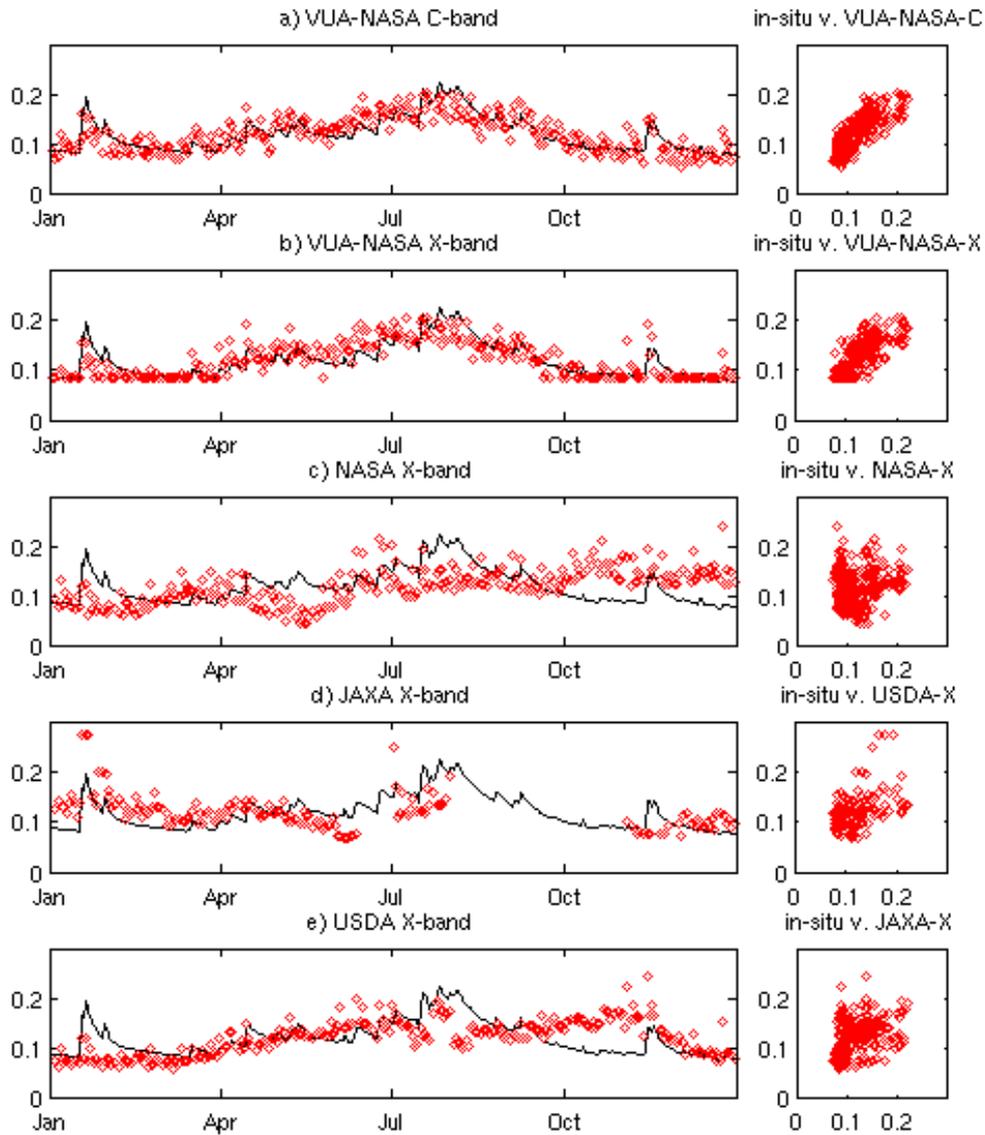


Figure E.8: Comparison of in-situ (solid lines) and normalised AMSR-E (red diamonds) near-surface soil moisture (m^3m^{-3}) for each retrieval algorithm, at Adelong (M10) over 2006.

280 **E. Additional time series plots comparing normalised AMSR-E retrieval algorithms to in situ data.**

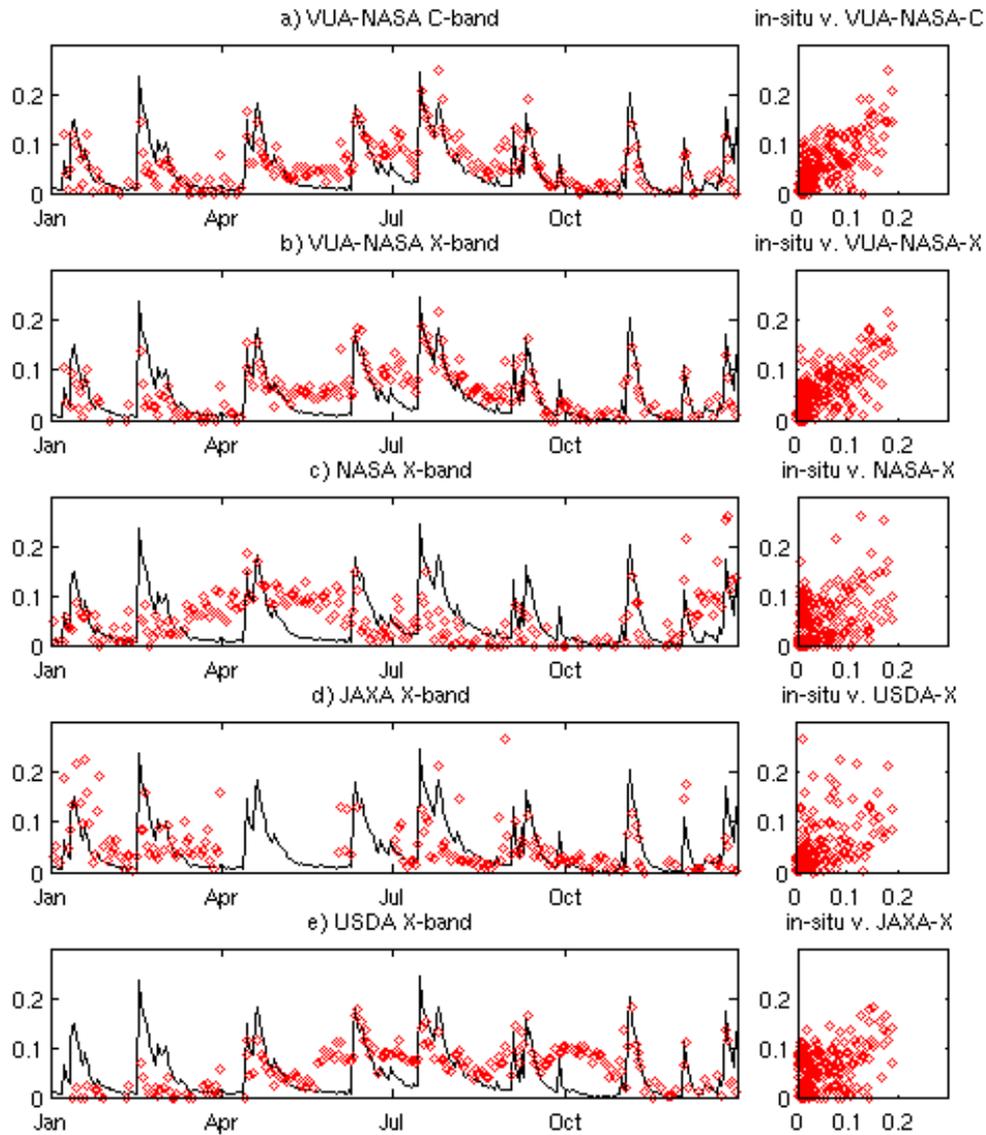


Figure E.9: Comparison of in-situ (solid lines) and normalised AMSR-E (red diamonds) near-surface soil moisture (m^3m^{-3}) for each retrieval algorithm, at Merriwa (G1) over 2006.

Appendix F

Additional time series plots
comparing original VUA-NASA
AMSR-E retrievals to in situ
data.

282 F. Additional time series plots comparing original VUA-NASA
AMSR-E retrievals to in situ data.

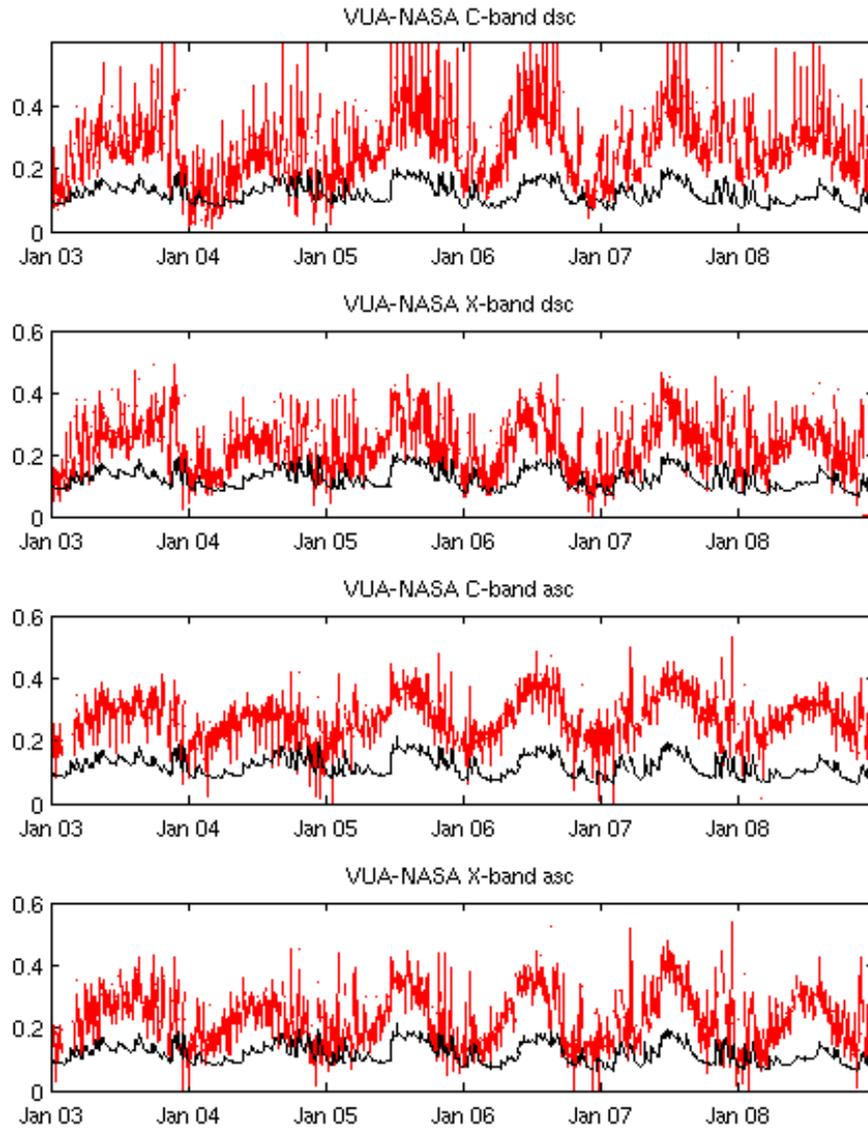


Figure F.1: Time series of in-situ (black) and original AMSR-E (red) near-surface soil moisture (m^3m^{-3}) from the VUA-NASA retrieval algorithm at Cooma Airfield (M1), from 2003-2008.

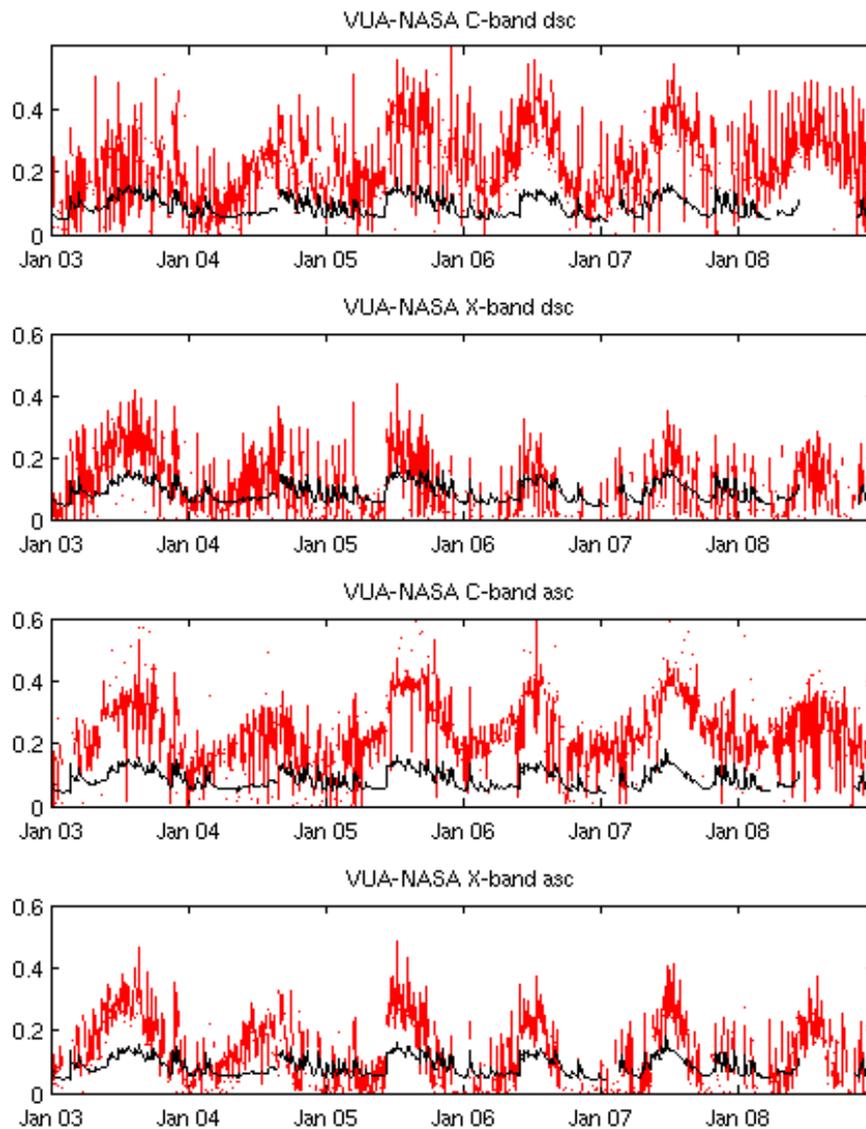


Figure F.2: Time series of in-situ (black) and original AMSR-E (red) near-surface soil moisture (m^3m^{-3}) from the VUA-NASA retrieval algorithm at Canberra Airport (M2), from 2003-2008.

284 F. Additional time series plots comparing original VUA-NASA
AMSR-E retrievals to in situ data.

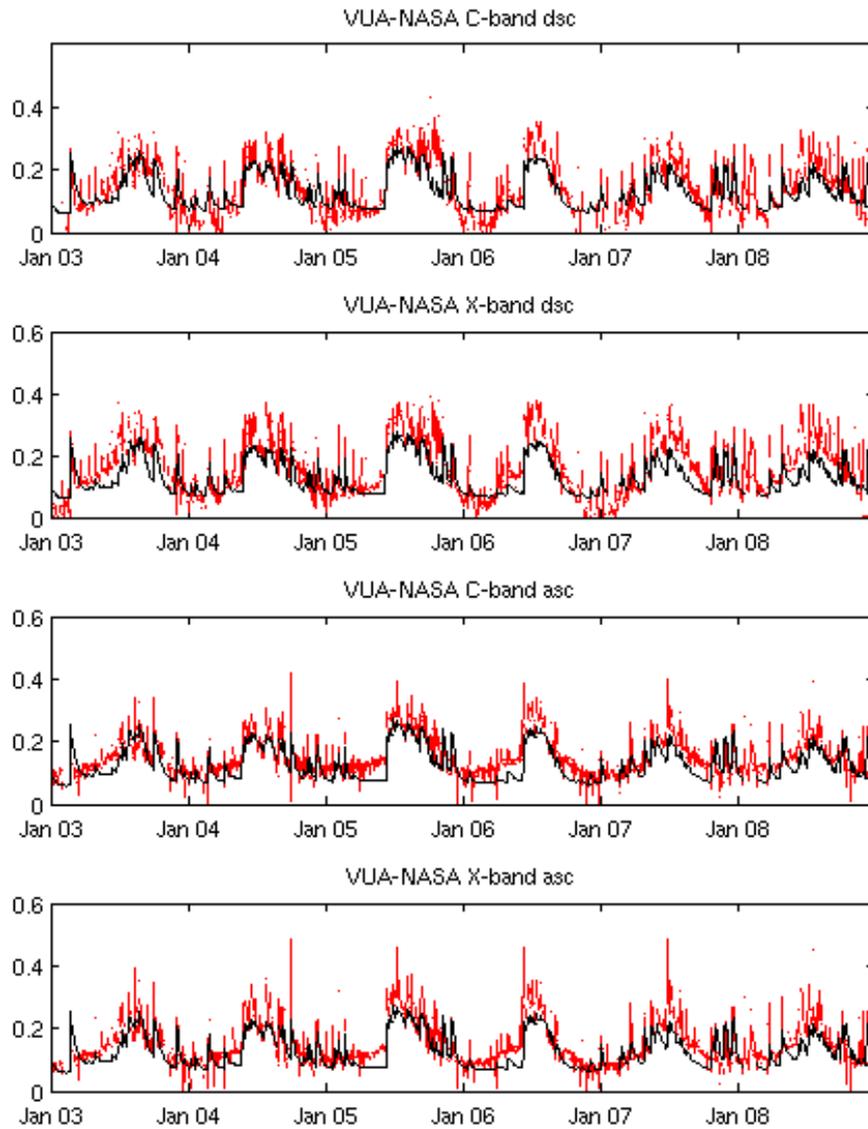


Figure F.3: Time series of in-situ (black) and original AMSR-E (red) near-surface soil moisture (m^3m^{-3}) from the VUA-NASA retrieval algorithm at West Wyalong Airfield (M4), from 2003-2008.

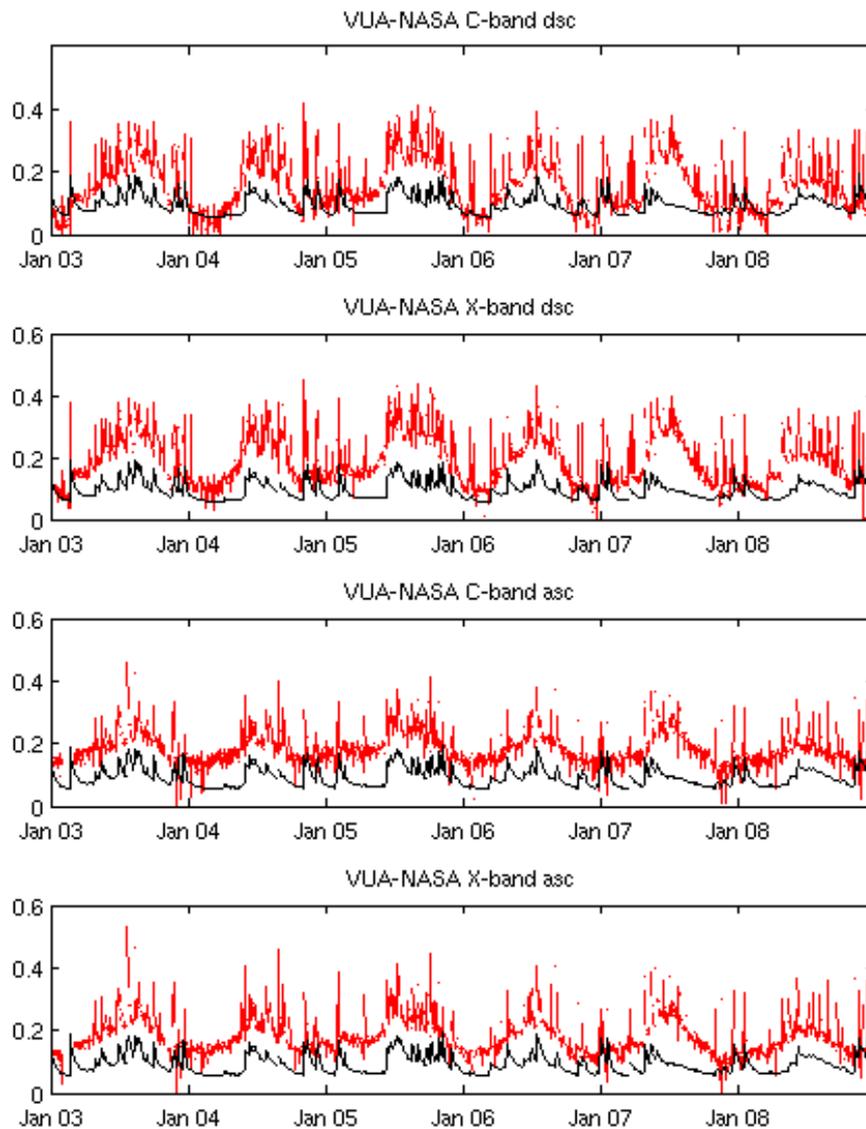


Figure F.4: Time series of in-situ (black) and original AMSR-E (red) near-surface soil moisture ($m^3 m^{-3}$) from the VUA-NASA retrieval algorithm at Balranald (M5), from 2003-2008.

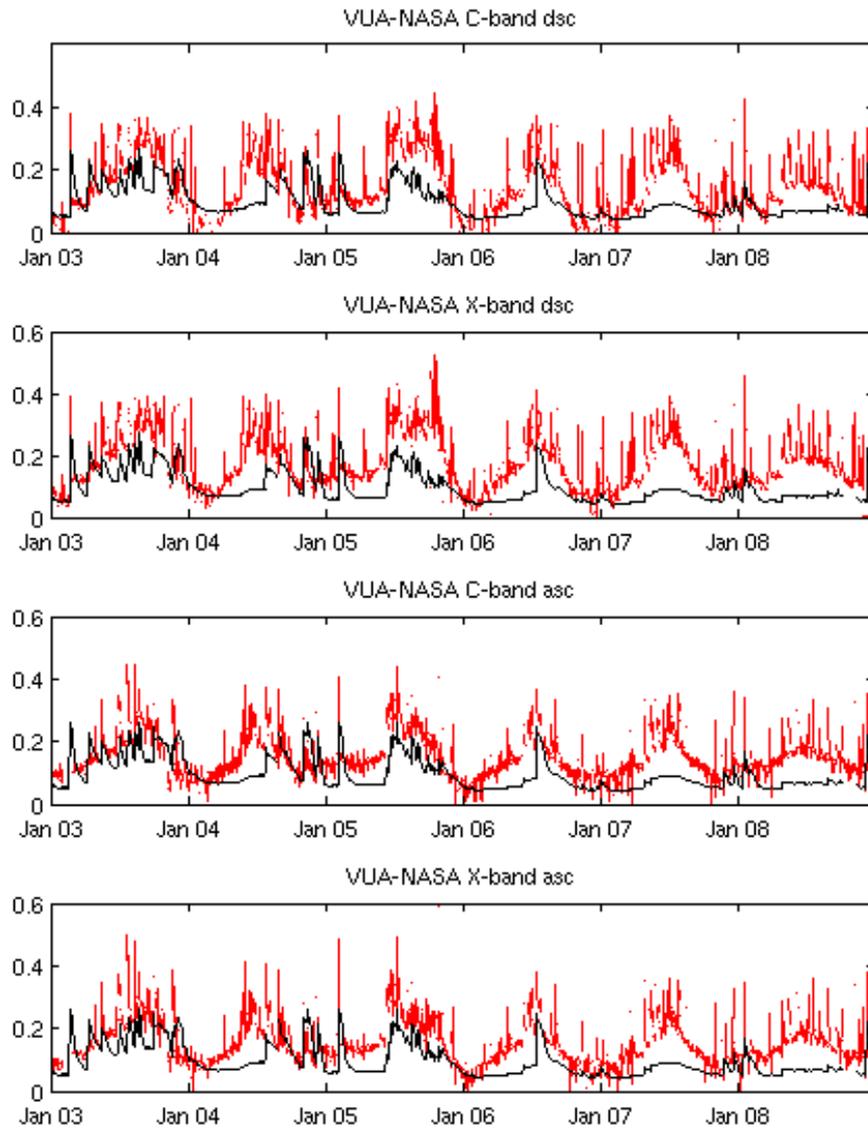


Figure F.5: Time series of in-situ (black) and original AMSR-E (red) near-surface soil moisture (m^3m^{-3}) from the VUA-NASA retrieval algorithm at Hay (M6), from 2003-2008.

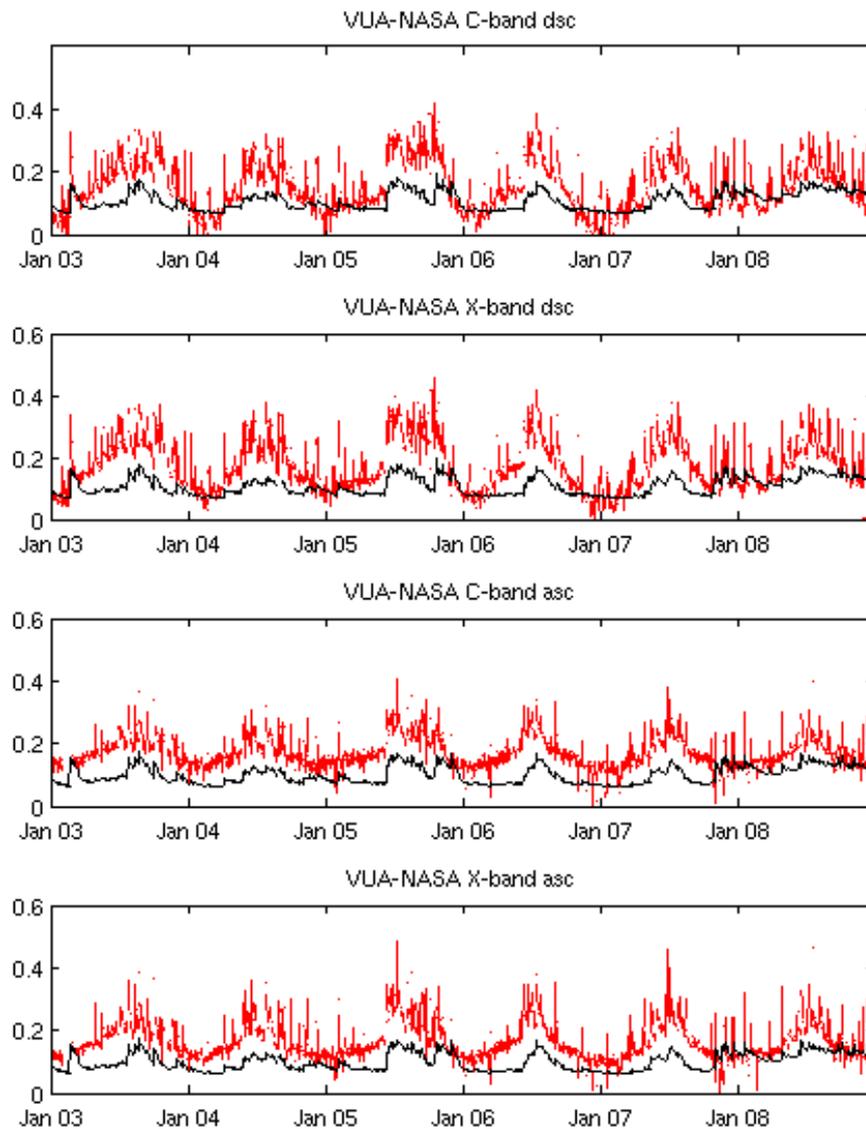


Figure F.6: Time series of in-situ (black) and original AMSR-E (red) near-surface soil moisture (m^3m^{-3}) from the VUA-NASA retrieval algorithm at Griffith Aerodrome (M7), from 2003-2008.

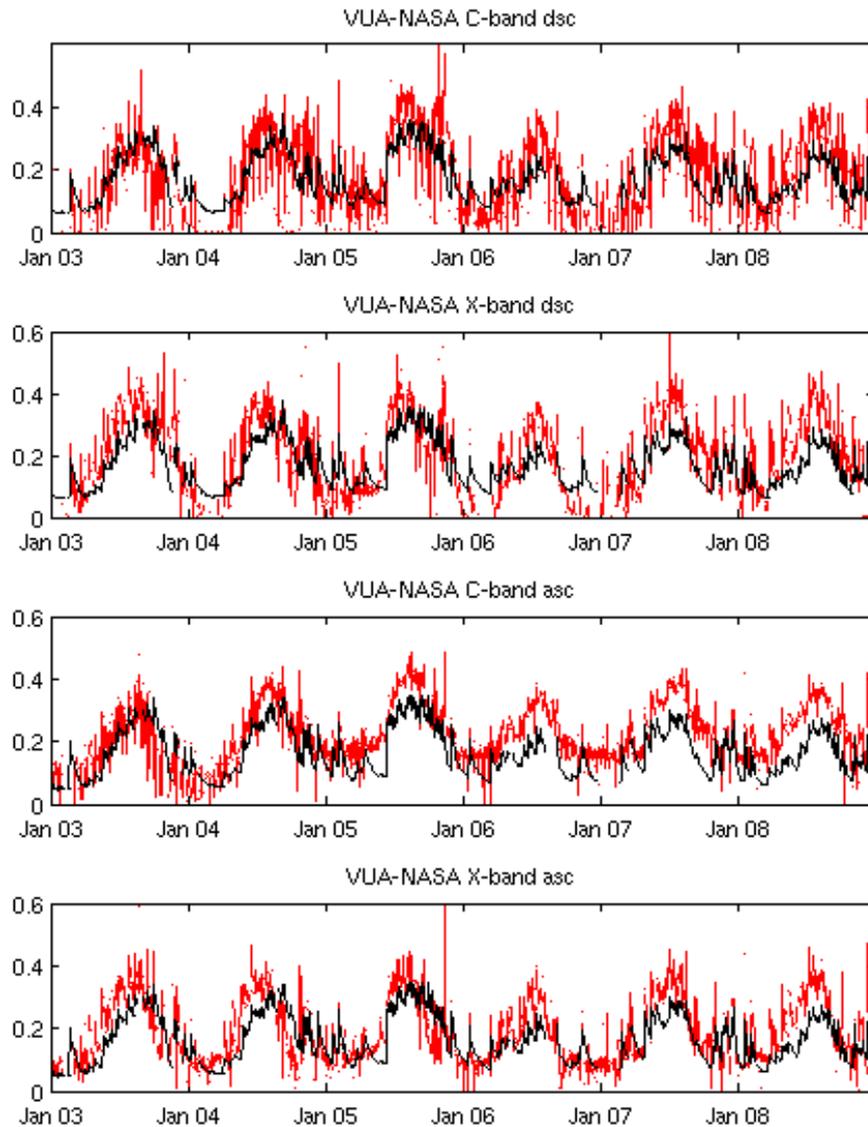


Figure F.7: Time series of in-situ (black) and original AMSR-E (red) near-surface soil moisture (m^3m^{-3}) from the VUA-NASA retrieval algorithm at Yanco (M8), from 2003-2008.

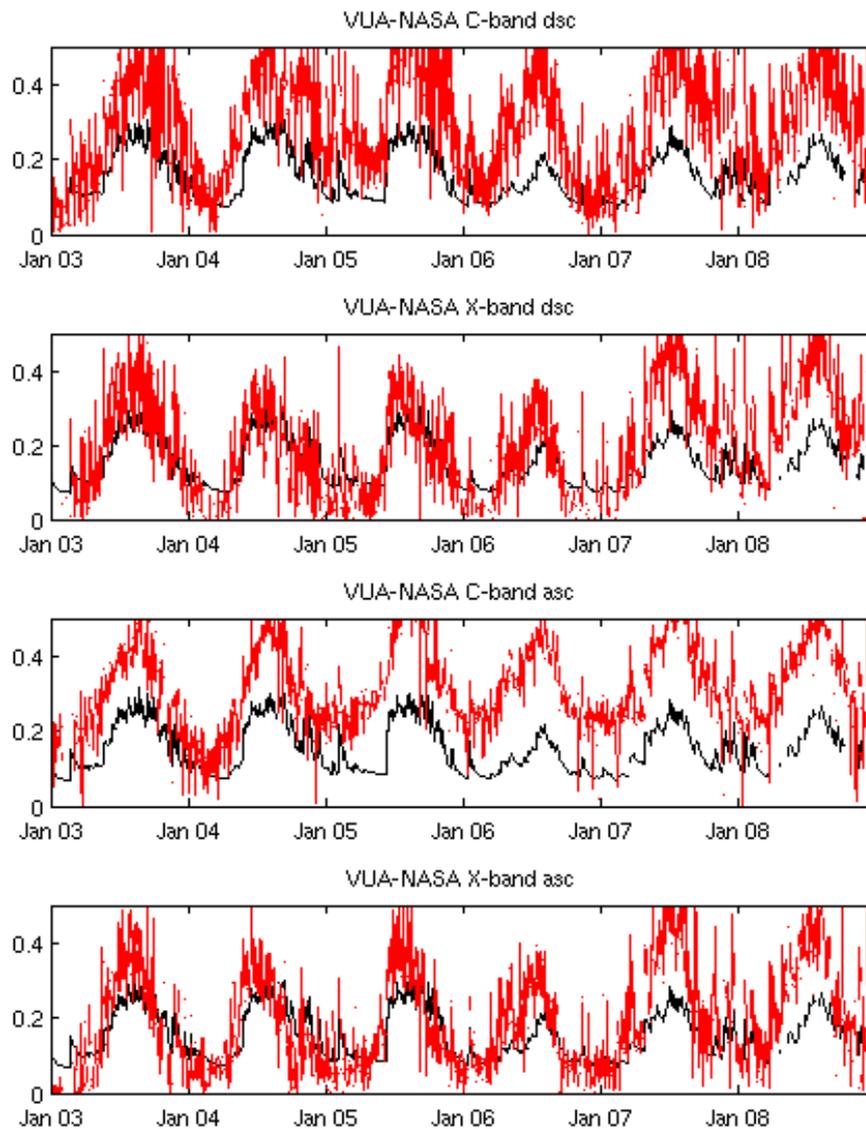


Figure F.8: Time series of in-situ (black) and original AMSR-E (red) near-surface soil moisture (m^3m^{-3}) from the VUA-NASA retrieval algorithm at Adelong (M10), from 2003-2008.

**F. Additional time series plots comparing original VUA-NASA
290 AMSR-E retrievals to in situ data.**

Appendix G

Additional time series plots
comparing normalised
VUA-NASA AMSR-E retrievals
to in situ data.

G. Additional time series plots comparing normalised VUA-NASA
292 AMSR-E retrievals to in situ data.

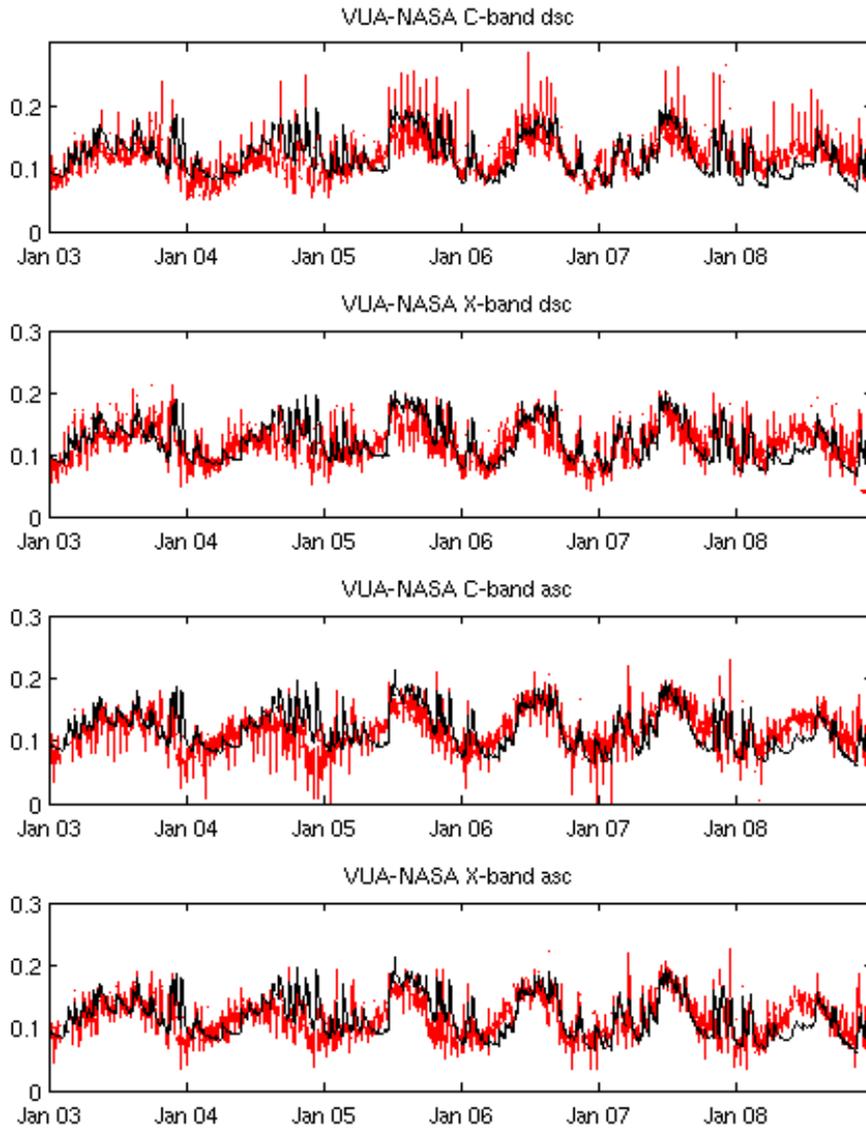


Figure G.1: Time series of in-situ (black) and normalised AMSR-E (red) near-surface soil moisture (m^3m^{-3}) from the VUA-NASA retrieval algorithm at Cooma Airfield (M1), from 2003-2008.

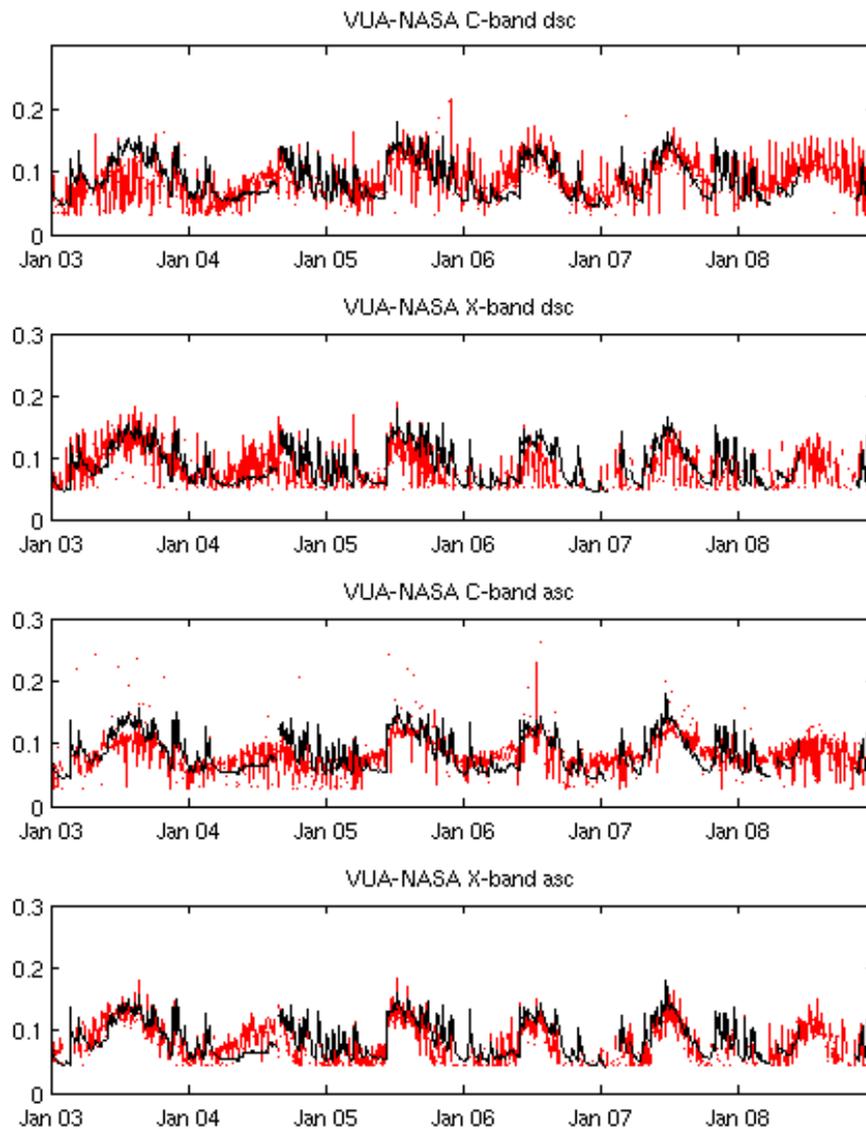


Figure G.2: Time series of in-situ (black) and normalised AMSR-E (red) near-surface soil moisture (m^3m^{-3}) from the VUA-NASA retrieval algorithm at Canberra Airport (M2), from 2003-2008.

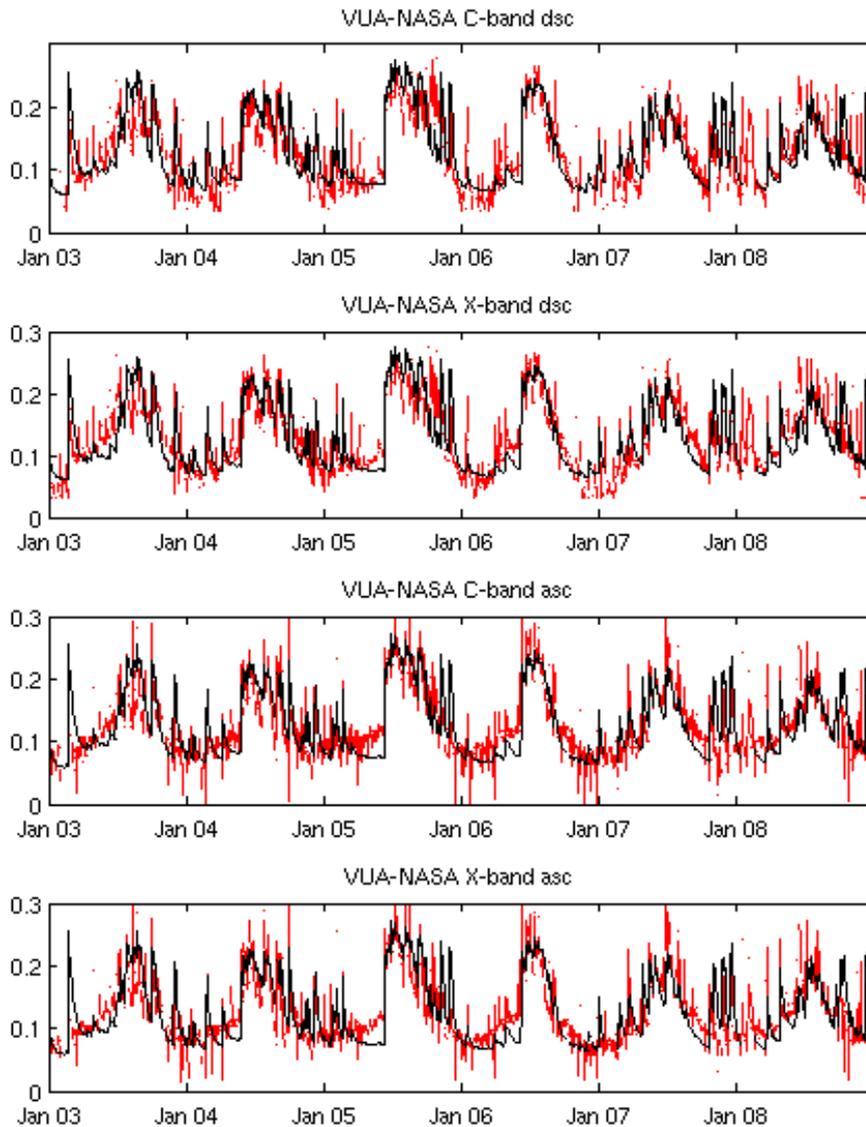


Figure G.3: Time series of in-situ (black) and normalised AMSR-E (red) near-surface soil moisture (m^3m^{-3}) from the VUA-NASA retrieval algorithm at West Wyalong Airfield (M4), from 2003-2008.

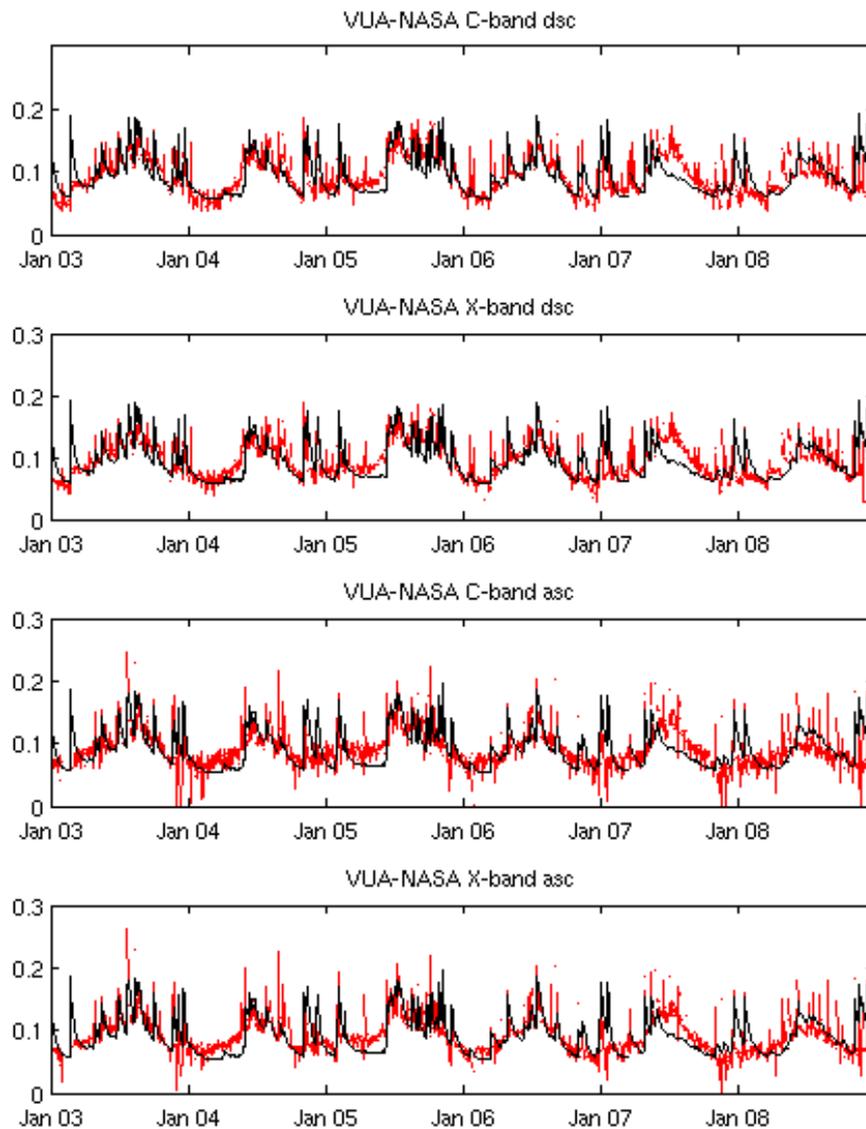


Figure G.4: Time series of in-situ (black) and normalised AMSR-E (red) near-surface soil moisture (m^3m^{-3}) from the VUA-NASA retrieval algorithm at Balranald (M5), from 2003-2008.

G. Additional time series plots comparing normalised VUA-NASA
296 AMSR-E retrievals to in situ data.

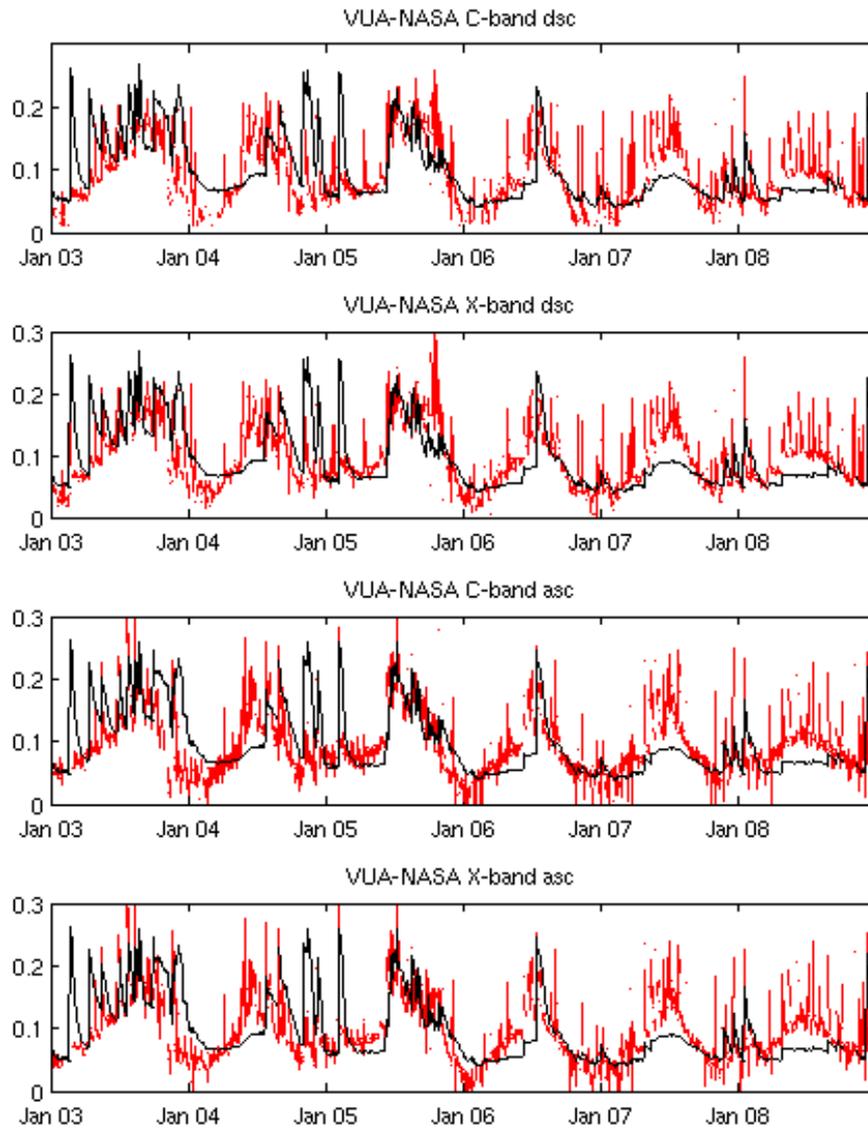


Figure G.5: Time series of in-situ (black) and normalised AMSR-E (red) near-surface soil moisture (m^3m^{-3}) from the VUA-NASA retrieval algorithm at Hay (M6), from 2003-2008.

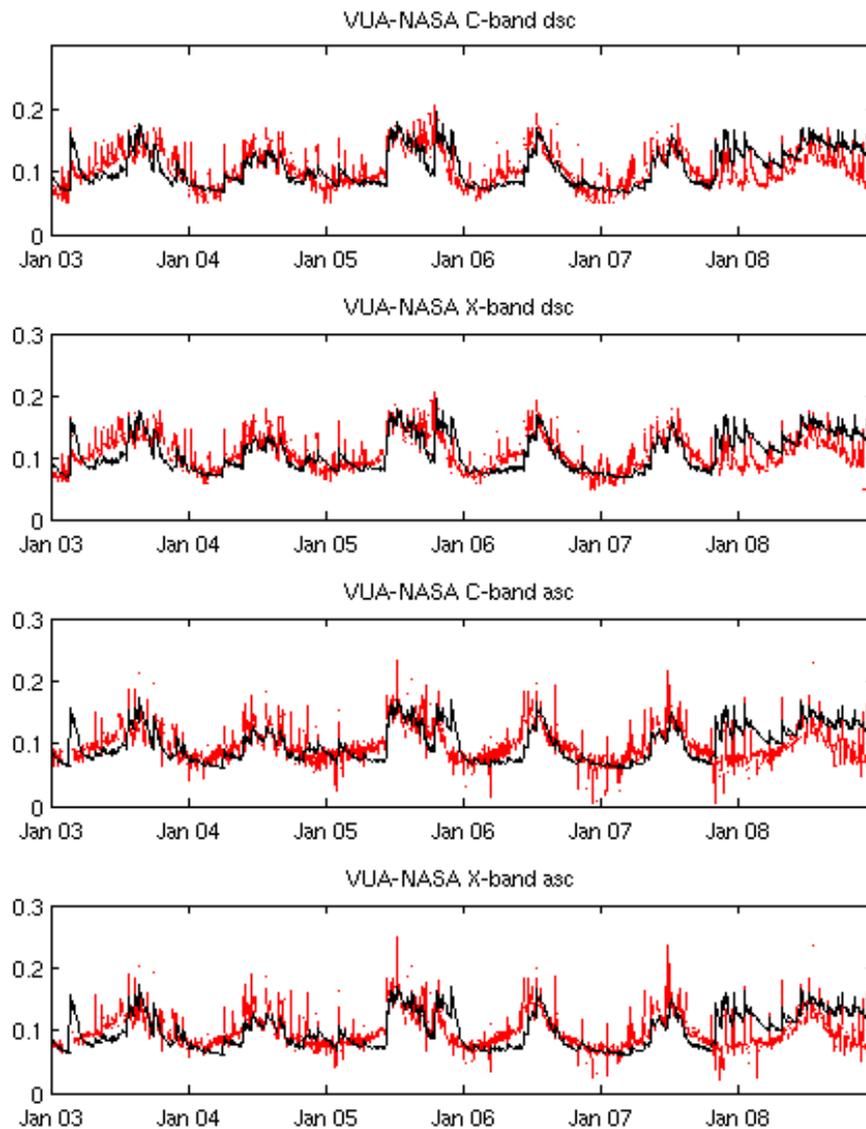


Figure G.6: Time series of in-situ (black) and normalised AMSR-E (red) near-surface soil moisture (m^3m^{-3}) from the VUA-NASA retrieval algorithm at Griffith Aerodrome (M7), from 2003-2008.

G. Additional time series plots comparing normalised VUA-NASA
298 AMSR-E retrievals to in situ data.

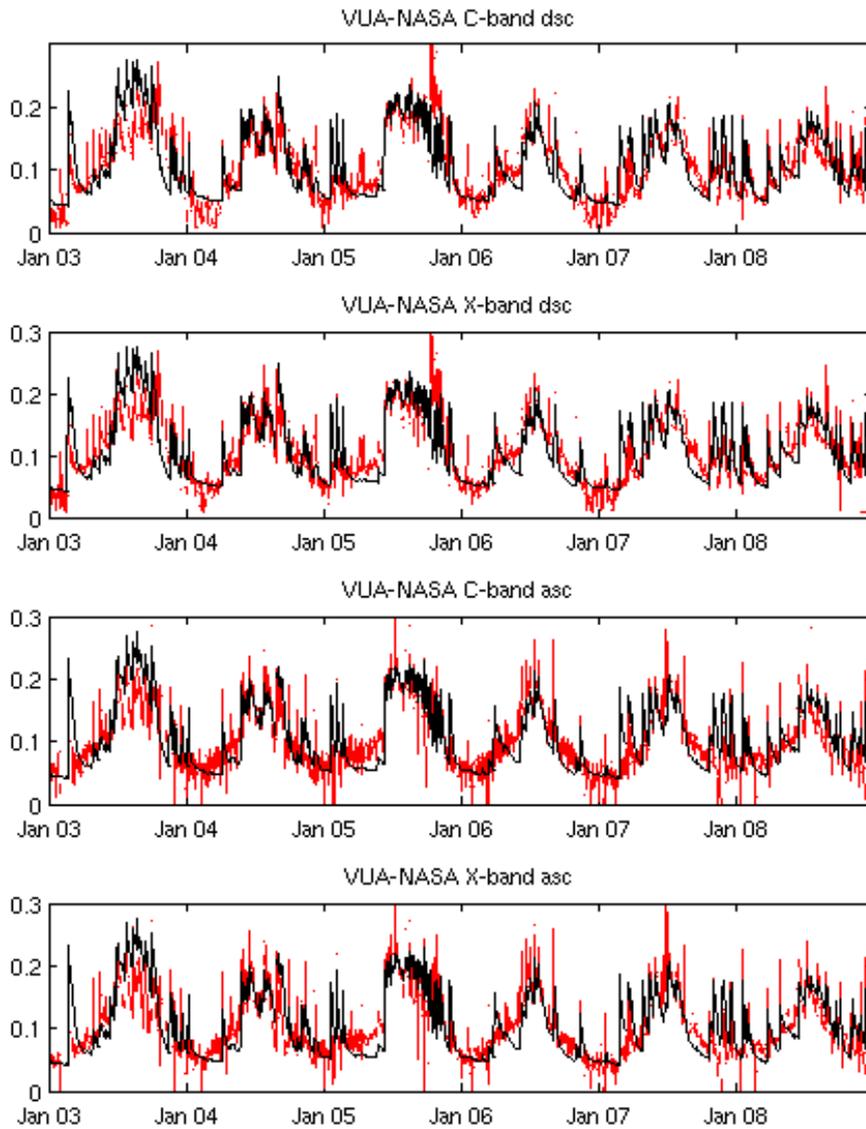


Figure G.7: Time series of in-situ (black) and normalised AMSR-E (red) near-surface soil moisture (m^3m^{-3}) from the VUA-NASA retrieval algorithm at Yanco (M8), from 2003-2008.

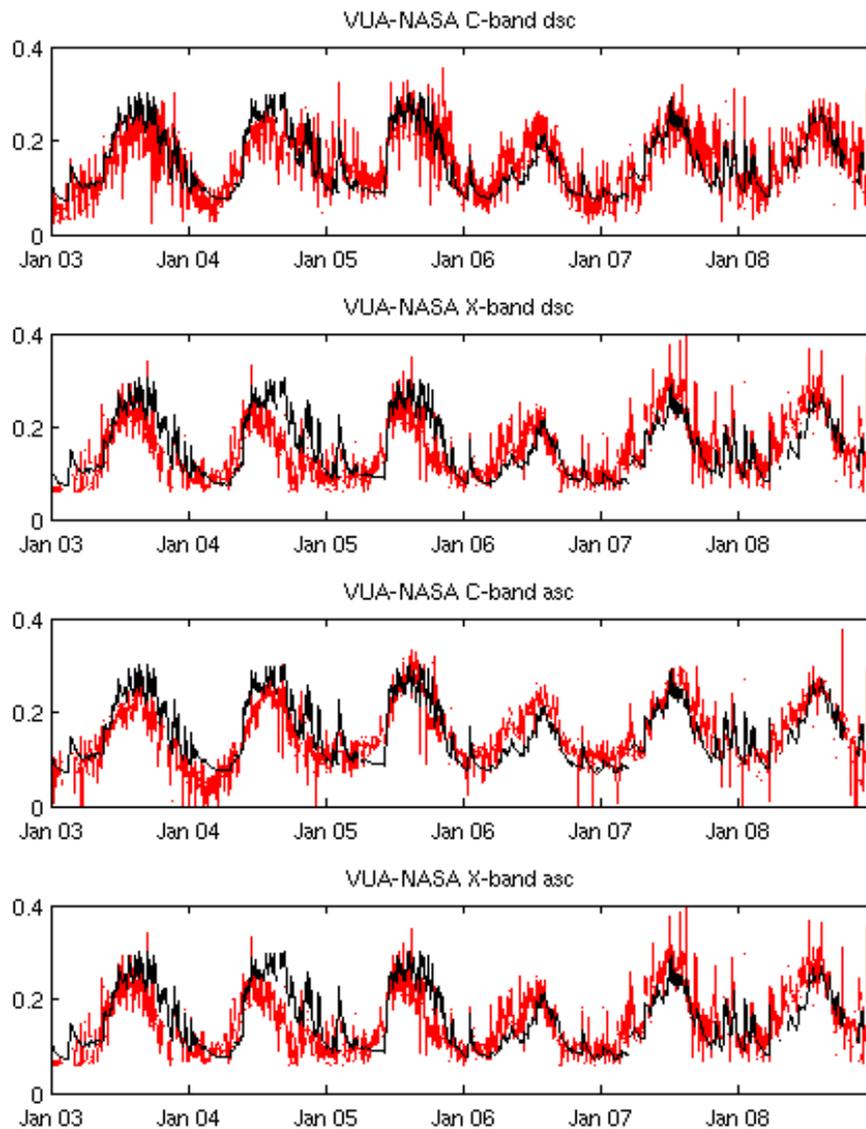


Figure G.8: Time series of in-situ (black) and normalised AMSR-E (red) near-surface soil moisture (m^3m^{-3}) from the VUA-NASA retrieval algorithm at Adelong (M10), from 2003-2008.