

Assimilation of Satellite Observations
into
Coastal Biogeochemical Models

Matthew Robert John Turner

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

July 2006

Department of Civil and Environmental Engineering
The University of Melbourne

Abstract

This thesis has investigated the improvement of forecasting temperature in a coastal embayment through the assimilation of sea surface temperature (SST) observations. The research was prompted by the increasing pressures on the coastal marine environment. To better manage the environment, an improved understanding of its future state is necessary. Improving the forecasting of temperature advances our knowledge in this direction. Whilst assimilation of SST is routinely carried out for oceans, its use has been minimal in coastal regions, which is more complicated because of anisotropic covariances and a breakdown of geostrophy in the coastal region. Improvements in computing power, and the introduction of ensemble-based assimilation techniques have made the approach followed in this thesis possible.

Previous coastal data assimilation had focused on sea level and tidal prediction. More recently, data assimilation has been undertaken with simple ecological models, but temperature has rarely been the subject of research. Furthermore, most coastal assimilation studies have assimilated point scale in situ observations, rather than satellite derived spatial observations, which are the focus of this thesis.

The thesis was conducted using a case study of Port Phillip Bay (PPB), a large embayment in south eastern Australia, where in situ temperature measurements gave an accurate indication of the true state of the temperature of water against which to compare forecasts. Over the long term, the SST observations were found to have negligible bias, however a strong diurnal bias was apparent. The model of PPB replicated the main features of PPB well, although the temperature prediction was warm biased. Existing methods to initialise ensembles and to incorporate forecast error were deemed inappropriate and so a new method for initialising the ensemble was developed based on the singular value decomposition (SVD) of a long model simulation. An appropriate ensemble size was determined by using the system variance explained by the singular values. Forecast error was introduced through the development of a rigorous approach to the generation of perturbed forcing data, which aimed to reduce the introduction of bias into the perturbed forcing data set.

Initially, various configurations of ensemble assimilation were tested in a synthetic setting by means of twin experiments. The low correlation between temperature and other variables meant that the multivariate analysis gave poor results, relative to a univariate (temperature) analysis. The use of a heatflux biased model was found to make the analysis suboptimal and to distort the forecast error so that the estimate of the forecast uncertainty was inaccurate, however the assimilation into a biased model gave large improvements over an unassimilated forecast. A comparison between the EnKF and EnSRF exploring the impact of perturbed observations found no significant difference between the methods, although EnSRF maintains the shape of the ensemble anomalies better than the EnKF.

The actual assimilation of SST data was contrasted against a climatology forecast of PPB temperature. The assimilation of SST without any specific accounting for the diurnal bias improved the forecast, although errors due to observation bias were noted. Attempts to remove this bias using diurnal correction algorithms failed, owing to a larger than expected cool skin. Conditional merging, which combines spatial and in situ observations, was applied to the SST observations and improved the forecast accuracy by reducing the observation bias. An examination of the assimilation innovations indicated where the forecast accuracy could be improved further.

By demonstrating the improvement that the assimilation of satellite derived observations can have on forecasting models, this thesis forms a step towards the development of an operational coastal marine forecasting system. In doing so the work undertaken in this thesis and possible extensions to other biogeochemical processes will become a useful tool for managers to protect the coastal marine environment from the multitude of pressures being placed upon it.

This is to certify that

- (i) the thesis comprises only my original work towards the PhD,
- (ii) due acknowledgement has been made in the text to all other material used,
- (iii) the thesis is less than 100,000 words in length, exclusive of table, maps, bibliographies, appendices.

Signature_____

Date_____

Acknowledgments

While the production of a thesis is an individual effort, it could not be successful without the help and support of others, whose assistance I gratefully acknowledge here. To my supervisors Dr. Jeffrey Walker, Dr. Peter Oke, and Prof. Rodger Grayson, my heartfelt thanks for your assistance, encouragement, and suggestions as the thesis has unfurled.

The numerical studies that comprise the bulk of this work require data for operation and evaluation. Many agencies and people have kindly made data available for this research and I happily acknowledge their assistance. For tidal and bathymetry data, I thank the Port of Melbourne Corporation, and especially Paul Downie and Greg London. For in situ temperature and salinity data, I thank MAFRI and especially Andy Longmore. The use of the MAFRI data are on the understanding that: the West Channel Pile data are recognised as having Andy Longmore as the custodian and Fisheries Victoria as the funding source. The other data are recognised as having Andy Longmore as custodian and the Port Phillip Bay Environmental Management Plan as the funding source. For riverine data and WTP data, I thank Melbourne Water. For atmospheric data, I thank the Bureau of Meteorology and the EPA Victoria. For the PPBES data, I thank Rob Molloy of CSIRO, and for the AVHRR SST observations, I thank CSIRO and especially Peter Turner and Glenelg Smith. Furthermore, I was given invaluable support in the configuration, operation and trouble-shooting of MECO by Dr. Mike Herzfeld and Mr Jason Waring, thank you.

A CSIRO-Melbourne University Collaborative Research Grant supported this research and enables collaboration with the researchers at CSIRO Marine and Atmospheric Research. I truly enjoyed my visits to Hobart and the opportunity to collaborate with a world class group of researchers. Notably, I would like to thank

Dr John Parslow for supporting this project, and Drs. Pavel Sakov and Stuart Corney, for their support, encouragement, constructive criticisms and time. This thesis was also financially supported by a Melbourne Research Scholarship (MRS).

Finally, although undertaking scientific research is an impersonal pursuit, the emotional support of friends, family and loved ones eases the burden. To this end I would thank my colleagues in the Department of Civil and Environmental Engineering, and especially Adam Smith, Christoph Rüdiger, and Stephen Wealands for sharing so many good times in the dungeon over the past years. To my family and parents, thank you for all you have done. Y para mi preciosa, gracias por todo, mi bella, como tú sabes.

Contents

Acknowledgments	v
List of Figures	x
List of Tables	xiv
Glossary	xvi
List of Symbols	xviii
1 Introduction	1
1.1 Research Rationale and Motivation	2
1.2 Thesis Objectives and Scope	5
1.3 Approach	7
1.4 Thesis Outline	8
2 Data Assimilation	11
2.1 Introduction	11
2.2 Sequential Filtering	13
2.3 Kalman Filter Simplifications	16
2.4 Ensemble Filters	20
2.5 Divergence	31
2.6 Applications of Data Assimilation	34
2.7 Review Findings	39
2.8 Chapter Summary	40

3	Data: Port Phillip Bay	41
3.1	Port Phillip Bay	42
3.2	Data	45
3.3	Data Analysis	50
3.4	SST Satellite Observations	74
3.5	Chapter Summary	89
4	Model Testing and Ensemble Generation	93
4.1	Hydrodynamic Modelling	94
4.2	Model Testing	98
4.3	Sensitivity Analysis	107
4.4	Model Linearity and Data Assimilation	112
4.5	Ensemble Generation	117
4.6	Ensemble Size	122
4.7	Forecast Error	126
4.8	Chapter Summary	144
5	Twin Experiments	147
5.1	Introduction	147
5.2	Method	149
5.3	Univariate Multivariate Comparison	152
5.4	Model Bias Experimentation	159
5.5	Analysis Method	166
5.6	Dynamic Profile Covariances	176
5.7	Chapter Summary	179
6	SST Assimilation for PPB	181
6.1	Introduction	181
6.2	Method	182
6.3	Unadjusted SST Assimilation	188
6.4	SST Error Correction	194
6.5	Conditional Merging Assimilation	197
6.6	In Situ Assimilation	202
6.7	Innovations	206

6.8 Chapter Summary	213
7 Conclusions and Future Directions	215
7.1 Conclusions	216
7.2 Future Directions	223
Bibliography	227
A MECO Parameters Utilised	243
B Choice of ξ and χ values in semi and restricted data types	245
C Derivation of Y and y	247
D Forecast Error Covariance	249
E Assimilation Images	253
F Anomaly shape: EnKF vs EnSRF	257

List of Figures

2.1	Schematic of Kalman filter forecast and analysis steps	17
2.2	EnKF analysis and forecast steps.	22
3.1	Location of PPB in south eastern Australia.	42
3.2	Weather and tide stations in the vicinity of PPB.	46
3.3	Water monitoring sites within PPB.	51
3.4	Spatial variation of atmospheric data.	53
3.5	Wind speed exceedence at various weather stations.	56
3.6	Wind direction variation.	57
3.7	Daily evaporation at Melbourne and spatial variance.	58
3.8	Spectral analysis of air temperature variation at Point Wilson.	59
3.9	Precipitation variability around PPB.	61
3.10	Cloud cover probability density functions.	62
3.11	Relative humidity exceedence around PPB.	64
3.12	Comparison of Frankston and Point Wilson relative humidity.	65
3.13	Average annual flow in the Yarra from 1978 to 2003.	66
3.14	Variation in daily solar radiation.	67
3.15	Sea level variation around PPB.	69
3.16	Time series of salinity at West Channel Pile between 1990 and 2001.	70
3.17	Correlation between salinity and temperature.	71
3.18	Temperature variation around PPB.	73
3.19	Spatial temperature variation with distance in PPB based on under- way transect data.	74
3.20	Illustration of skin cooling and diurnal warming.	77
3.21	Diurnal warming predicted by Gentemann <i>et al.</i> (2003) model.	79

3.22	Diurnal warming predicted by Kawai and Kawamura (2002) model.	81
3.23	Satellite pixel average spatial variability.	82
3.24	AVHRR satellite data processing.	83
3.25	Satellite observation, shallow in situ value comparison.	85
3.26	Satellite observation, deep in situ value comparison.	86
3.27	Spatial correlation of temperature observation error.	87
3.28	Temporal variation of temperature observation error.	88
4.1	PPB numerical grid and bathymetry	96
4.2	Modelled residual surface currents for January 2000 to June 2003.	99
4.3	Modelled sea level from January 2000 to June 2003.	100
4.4	Observed and modelled sea level at various tide gauge sites.	102
4.5	Modelled sea level at Hovel Pile and predicted error.	103
4.6	Modelled temperature from January 2000 to June 2003.	104
4.7	Observed and modelled temperature.	105
4.8	Linearity test of PPB model.	116
4.9	Univariate temperature singular vectors used to initiate the ensemble.	122
4.10	Multivariate singular vectors used to initiate the ensemble.	123
4.11	Cumulative variance explained by singular vectors.	125
4.12	Examples of ensembles of perturbed forcing data.	137
4.13	Effects of forcing data ensembles on model predictions.	138
4.14	Comparison of a stochastic model against perturbed forcing data.	141
4.15	Covariance of temperature with other model fields.	143
5.1	Examples of synthetic truth, control, and observations.	151
5.2	Spatial RMSE of temperature, salinity and sea level.	154
5.3	Salinity near the centre of PPB and near the open boundary.	155
5.4	Surface salinity difference between the ensemble mean and the truth.	156
5.5	Temperature error and ensemble spread near the centre of PPB.	160
5.6	Histogram of innovations for both biased and unbiased runs.	162
5.7	EnSRF vs. EnKF RMSE.	168
5.8	Surface snapshots of assimilation results.	169
5.9	Ensemble filter internal mechanism comparison.	173
5.10	Spatial correlation between analysed and forecast anomalies.	174

5.11	Perturbation vector shape changes due to filter type.	175
5.12	Profile time series of conditions in the centre of PPB.	177
6.1	Examples of satellite SST observations.	184
6.2	Atmospheric conditions over the assimilation period.	187
6.3	Climatological temperature in PPB from 1990 to 2001.	188
6.4	Temperature forecasts with the assimilation of unadjusted SST data .	190
6.5	Spatial comparison of temperature forecasts.	192
6.6	Contrast of error and RMS spread for unadjusted SST assimilation. .	193
6.7	Comparison of actual and predicted observation error.	196
6.8	An example of ordinary kriging.	199
6.9	Assimilation of conditional merged and in situ observations	201
6.10	Contrast of error and RMS spread for Conditional Merging.	203
6.11	Contrast of PPB states predicted by CM, control and In situ.	205
6.12	Innovation mean of selected assimilation runs.	206
6.13	Innovation standard deviation of selected assimilation runs.	208
6.14	Innovation distributions for conditionally merged forecasts.	210
6.15	Spatial distribution of conditionally merged forecast innovations. . . .	211
D.1	Comparison of a stochastic model against perturbed forcing data. . .	250
D.2	Covariance of temperature with other model fields.	251
D.3	Covariance of temperature with other model fields.	252
E.1	Uncorrected NOAA-16 AVHRR SST images used for assimilation. . .	253
E.2	Uncorrected NOAA-16 AVHRR SST images ... cont.	254
E.3	Uncorrected NOAA-16 AVHRR SST images ... cont.	255
F.1	Anomaly shape changes due to filter type — ensemble member 1. . .	257
F.2	Anomaly shape changes due to filter type — ensemble member 2. . .	258
F.3	Anomaly shape changes due to filter type — ensemble member 3. . .	258
F.4	Anomaly shape changes due to filter type — ensemble member 4. . .	259
F.5	Anomaly shape changes due to filter type — ensemble member 5. . .	259
F.6	Anomaly shape changes due to filter type — ensemble member 6. . .	260
F.7	Anomaly shape changes due to filter type — ensemble member 7. . .	260

F.8	Anomaly shape changes due to filter type — ensemble member 8.	. . . 261
F.9	Anomaly shape changes due to filter type — ensemble member 9.	. . . 261
F.10	Anomaly shape changes due to filter type — ensemble member 10.	. . . 262

List of Tables

3.1	Annual fresh water exchange within PPB.	43
3.2	Bureau of Meteorology data obtained for years 2000–2003.	47
3.3	Riverine data of waters entering PPB.	48
3.4	Water monitoring data recorded in PPB.	50
3.5	Correlation coefficients of 3-hourly data at Melbourne	54
3.6	Correlation coefficients of daily data at Melbourne	54
3.7	Average and standard deviation of pressure.	55
3.8	Pressure anomaly correlation coefficients.	55
3.9	Rank of exceedence temperatures for 10% and 90% levels.	60
3.10	Annual precipitation across PPB.	61
3.11	Cloud cover observed at Laverton given a Melbourne observation. . .	63
3.12	Summary of in situ recorded temperature data.	72
3.13	Coefficients for Kawai and Kawamura (2002) model	80
3.14	Difference between in situ and satellite data	87
3.15	Summary of meteorological stations used for modelling.	90
4.1	RMSE between observed and modelled sea level.	101
4.2	RMSE between observed and modelled temperature.	106
4.3	Model Parameters Sensitivity.	109
4.4	Sensitivity of model to mixing scheme.	110
4.5	Influence of forcing on model predictions.	111
4.6	Sensitivity of model to temperature module and forcing data.	111
4.7	Mean std. dev. values from decomposition of long time series.	124
4.8	Values of ξ and χ for various unrestricted data types.	131
4.9	Values of ξ and χ for various semi-restricted data types.	134

4.10	Values of ξ and χ for various restricted data types.	135
5.1	Average RMSE for multi-univariate comparison.	153
5.2	Average RMSE for biased-unbiased model comparison.	161
5.3	Normality of the biased and unbiased innovations.	163
5.4	Significance of mean and standard deviation of unbiased innovations .	165
5.5	RMSE at a range of ensemble sizes.	170
6.1	Forecast RMSE for SST assimilation.	188
6.2	Forecast RMSE for skin-corrected SST assimilation.	195
6.3	Forecast RMSE for Conditionally Merged Assimilation.	200
6.4	Forecast RMSE for In Situ SST assimilation.	204
6.5	Variation of innovation statistics over simulation period.	207
A.1	Model timing parameters.	243
A.2	Wind drag law parameters.	243
A.3	Heat flux and solar radiation parameters	243
A.4	Bottom friction parameters.	244
A.5	Csanady vertical mixing parameters.	244
A.6	Horizontal mixing parameters.	244
A.7	Computational Settings.	244
B.1	ξ probability distribution: semi-restricted domain	246
B.2	ξ probability distribution: restricted domain	246

Glossary

ASL	— Above Sea Level
AVHRR	— Advanced Very High Resolution Radiometer
AATSR	— Advanced Along Track Scanning Radiometer
ATSR	— Along Track Scanning Radiometer
BLUE	— Best Linear Unbiased Estimator
CPSST	— Cross Product SST
EnKF	— Ensemble Kalman Filter
EnSRF	— Ensemble Square Root Filter
EPA	— Environment Protection Authority
EOF	— Empirical Orthogonal Function
ECMRWF	— European Centre for Medium Range Weather Forecasting
EnOI	— Ensemble Optimal Interpolation
MAFRI	— Marine and Fisheries Research Institute
MCSST	— Multi-Channel SST
MECO	— Model for Estuaries and Coastal Oceans
MERIS	— MEdition Resolution Imaging Spectrometer
MODIS	— Moderate Resolution Imaging Spectroradiometer
MOM	— Modular Ocean Model
NCEP	— National Centers for Environmental Prediction
NLSST	— Non-Linear SST
NOAA	— National Oceanic and Atmospheric Administration

Continued on next page

POM	— Princeton Ocean Model
PPB	— Port Phillip Bay
PPBES	— Port Phillip Bay Environmental Study
PROMISE	— PRe-Operational Modelling In the Seas of Europe
PSU	— Practical Salinity Units
RMSE	— Root Mean Squared Error
RMS	— Root Mean Squared
ROMS	— Regional Ocean Modelling System
SeaWiFS	— Sea-viewing Wide Field-of-view Sensor
SEEK	— Singular Evolutive Extended Kalman filter
SST	— Sea Surface Temperature
SVD	— Singular Value Decomposition
SWR	— Short Wave Radiation
UTC	— Coordinated Universal Time
WTP	— Western Treatment Plant

List of Symbols

Symbol	Definition	Units
a	calibration coefficient	[-]
b	calibration coefficient	[-]
c	calibration coefficient	[-]
d	calibration coefficient	[-]
d	semi variance correlation length scale	[-]
e	base of the natural logarithm	[-]
f	nonlinear function operator	[-]
g	gravity	[ms ⁻²]
h	forcing data record	[-]
h	semi variance distance value	[-]
k	time subscript	[-]
l	number of sources of system noise	[-]
m	number of observations	[-]
n	number of states	[-]
n_e	number of ensemble members	[-]
n_s	number of samples used in standard error	[-]
p	number of time points	[-]
s_n	number of successive correction iterations	[-]
s	successive correction counter	[-]
t	time	[s]

Continued on next page

Symbol	Definition	Units
t	Students- t statistic	[-]
u_1	flow in i direction	[ms ⁻¹]
u_2	flow in j direction	[ms ⁻¹]
w	flow in z direction	[ms ⁻¹]
w	weighting factor	[-]
w^*	best estimator of weighting factor	[-]
x	singular state	[-]
x^*	best linear estimator of the singular state	[-]
z	directional vertical axis	[-]
z_i	normally distributed random number	[-]
α_i	gaussian distributed random process $N(0,1)$	[-]
β	time-independent forcing data offset term	[-]
γ	semivariance term	[-]
γ	standard error	[-]
δ	linearity test perturbation vector	$(n \times 1)$
ε	observation perturbation vector	$(m \times 1)$
ϵ_k	forcing data error vector at time t_k	$(n \times 1)$
ζ_i	time-dependent forcing data error term	[-]
η	sea level	[m]
θ	linearity test statistic	[-]
μ	sample mean	[-]
ν_k	model error vector at time t_k	$(n \times 1)$
ξ	perturbation error standard deviation control term	[-]
π	ratio of a circle's circumference to its diameter	[-]
ρ	correlation coefficient	[-]
ρ	density	[kg m ⁻³]
ρ_0	reference density	[kg m ⁻³]
σ	standard deviation	[-]
σ^2	variance	[-]

Continued on next page

Symbol	Definition	Units
σ^{2*}	best linear estimate of variance	[-]
χ	perturbation offset standard deviation control term	[-]
ω	frequency	[s ⁻¹]
A	conditionally merged surface	[-]
D	Kolmogorov-Smirnov test statistic	[Wm ⁻²]
E_1	RMSE of ensemble mean	[-]
E_2	average RMSE of each ensemble member	[-]
F	χ^2 standard deviation test statistic	[-]
L	spatial observations surface	[-]
M	spatially interpolated surface based on spatial observations at in situ sites	[-]
$N(x, y)$	normal distribution with a mean of x , and a standard deviation of y	[-]
N	Brunt - Väisälä (stratification) frequency	[s ⁻¹]
P	spatially interpolated surface using in situ observations	[-]
Q_0	shortwave heat flux threshold	[Wm ⁻²]
$Q_{I_{\max}}$	peak shortwave heat flux	[Wm ⁻²]
$\overline{Q_I}$	mean daily shortwave heat flux at top of atmosphere	[Wm ⁻²]
Q_{K-S}	Kolmogorov-Smirnov test statistic probability function	[-]
Q_N	Net radiation	[Wm ⁻²]
Q_S	incoming solar radiation	[Wm ⁻²]
Q_L	long wave radiation	[Wm ⁻²]
Q_H	sensible heat flux	[Wm ⁻²]
Q_E	latent heat flux	[Wm ⁻²]
R	sample correlation coefficient	[-]
S	salinity	[PSU]
T	temperature	[°C]
$\overline{T}(t)$	spatially averaged temperature	[°C]
W	wind speed	[ms ⁻²]

Continued on next page

Symbol	Definition	Units
$W(s)$	Sucessive correction weights	$[-]$
\bar{W}	mean daily wind speed	$[ms^{-2}]$
\mathbf{b}_k	state bias estimate vector at time t_k	$(n \times 1)$
\mathbf{d}_k	observation vector at time t_k	$(m \times 1)$
\mathbf{h}_k	forcing data vector at time t_k	$(p \times 1)$
\mathbf{i}	horizontal cartesian coordinate vector	(3×1)
\mathbf{j}	horizontal cartesian coordinate vector	(3×1)
\mathbf{j}_{n_e}	averaging vector where each cell has a value of $\frac{1}{n_e}$	$(n_e \times 1)$
\mathbf{n}_k	observation error vector at time t_k	$(m \times 1)$
\mathbf{u}_k	system noise vector at time t_k	$(l \times 1)$
\mathbf{v}_k	normally distributed perturbation vector at time t_k	$(n \times 1)$
\mathbf{x}_k	state vector at time t_k	$(n \times 1)$
\mathbf{x}_k^a	analysis state vector at time t_k	$(n \times 1)$
\mathbf{x}_k^f	forecast state vector at time t_k	$(n \times 1)$
$\bar{\mathbf{x}}_k$	ensemble mean state vector at time t_k	$(n \times 1)$
\mathbf{x}'_k	state vector anomaly at time t_k	$(n \times 1)$
$\hat{\mathbf{x}}$	true state vector at time t_k	$(n \times 1)$
\mathbf{y}	Ensemble Kalman filter ensemble mean operator	$(n_e \times 1)$
\mathbf{z}	Ensemble square root filter ensemble mean operator	$(n_e \times 1)$
\mathbf{z}	vertical cartesian coordinate vector	(3×1)
\mathbf{D}	observation ensmeble matrix	$(m \times n_e)$
\mathbf{F}	matrix of physically realistic state anomalies	$(n \times p)$
\mathbf{G}_k	matrix relating system noise to states at time t_k	$(n \times l)$
\mathbf{H}_k	forward observation operator matrix at time t_k	$(m \times n)$
\mathbf{I}	identity matrix	$(-)$
\mathbf{J}_{n_e}	averaging matrix with each cell has value of $\frac{1}{n_e}$	$(n_e \times n_e)$
\mathbf{K}_k	Kalman gain matrix at time t_k	$(n \times m)$
$\tilde{\mathbf{K}}$	time-constant gain matrix	$(n \times m)$
\mathbf{K}'_k	anomaly gain matrix at time t_k	$(n \times m)$

Continued on next page

Symbol	Definition	Units
\mathbf{L}_k	state bias estimate gain matrix at time t_k	$(n \times m)$
\mathbf{M}	singular vector matrix	$(m \times m)$
\mathbf{P}_k	forecast error covariance matrix at time t_k	$(n \times n)$
\mathbf{P}_k^a	forecast error covariance matrix analysis at time t_k	$(n \times n)$
\mathbf{P}_k^f	forecast error covariance matrix forecast at time t_k	$(n \times n)$
\mathbf{P}_e	ensemble forecast covariance matrix at time t_k	$(n \times n)$
\mathbf{P}_k^b	state bias estimate error covariance matrix at time t_k	$n \times n$
$\tilde{\mathbf{P}}$	time-constant forecast error covariance matrix	$(n \times n)$
\mathbf{Q}_k	covariance matrix of system noise at time t_k	$(l \times l)$
\mathbf{R}_k	observation error covariance matrix at time t_k	$(m \times m)$
\mathbf{R}_e	ensemble observation error covariance matrix	$(m \times m)$
\mathbf{U}	singular vector matrix	$(n \times n)$
\mathbf{V}	singular vector matrix	$(p \times p)$
\mathbf{W}	an orthogonal matrix	$(n_e \times n_e)$
\mathbf{X}_k	ensemble state matrix at time t_k	$(n \times n_e)$
\mathbf{X}_k^f	ensemble state matrix forecast at time t_k	$(n \times n_e)$
\mathbf{X}_k^a	ensemble state matrix analysis at time t_k	$(n \times n_e)$
$\bar{\mathbf{X}}$	ensemble state mean	$(n \times n_e)$
\mathbf{X}'	ensemble state anomalies	$(n \times n_e)$
\mathbf{Y}	ensemble Kalman filter ensemble anomaly operator	$(n_e \times n_e)$
\mathbf{Z}	ensemble square root filter ensemble anomaly operator	$(n_e \times n_e)$
Λ	anomaly weighting matrix	$(n_e \times n_e)$
Σ	singular value matrix	$(m \times m)$
Υ	observation perturbation ensemble matrix	$(m \times n_e)$
Φ_k	linear system operator matrix at time t_k	$(n \times n)$

Chapter 1

Introduction

The potential improvements to coastal marine forecasting through the incorporation of accurate satellite-observed data into a forecast model are investigated in this thesis. Its focus is the application of ensemble data assimilation techniques to the coastal environment. While the potential scope of this application is enormous, this thesis limits itself to assimilation of satellite sea surface temperature (SST) observations into a hydrodynamic model. The research is undertaken through a case study approach, whereby a series of numerical studies are made to explore the data assimilation techniques. The testing of various data assimilation configurations is made in a synthetic setting before moving to actual SST observations.

This thesis represents a step towards having an operational biogeochemical forecasting system for the coastal marine environment. Such a system, akin to a daily weather report, would predict the biological, chemical and physical conditions of coastal marine waters, and would be of benefit to a wide range of users. Currently to the knowledge of the author no such system exists which forecasts a full range of conditions, however many systems predicting various physical conditions (e.g. sea level and currents) exist and undoubtedly efforts are being made in this direction.

This thesis improves on previous state of knowledge in a number of ways. While data assimilation techniques are widely applied in the atmospheric and oceanographic fields they have rarely been applied to high resolution coastal and estuarine situations with the exception of water level forecasting. The modelling of coastal waters is more complex with greater nonlinear effects. Another different aspect of this

research is the focus on the assimilation of satellite observations into the forecast. Existing coastal marine forecasting systems focus on in situ observed data. The limited spatial resolution of such data inhibits accuracy of the forecast. By using satellite derived observations, improvements in forecasting skill should be possible.

Developed over the last decade, ensemble data assimilation tools have rarely been applied in coastal applications. The use of ensemble techniques enables a more accurate assimilation than would otherwise be possible. Furthermore, as ensemble data assimilation techniques are a recent development, there has been rapid development of the field and a bewildering array of opinions and possibilities. In this regard, the thesis also seeks to find some structure or guidance to the ensemble data assimilation configuration used. In doing so perceived inadequacies of existing methods for initiating ensembles and incorporating forecast error prompt further development of these methods. New insight into ensemble filter performance is also found, with the choice of filter affecting the resulting analysis anomaly shape.

1.1 Research Rationale and Motivation

The motivation and rationale of this thesis — forming a step towards the development of an operational coastal marine forecasting system — is approached from two angles: i) the level of demand for such a system: the importance of, and pressures on, coastal environments are relevant here, and ii) the supply of such a system: does the capability and capacity to develop a coastal marine forecasting system exist? The following sections demonstrate that there is a definite need for such a system which will only increase in the future, and further demonstrate that recent advances in technology make the development of such a system realistic. This overview justifies the research conducted in this thesis and places it within the context of a wider research effort.

1.1.1 Demand for a coastal forecasting system

Demand for coastal biogeochemical forecasting systems stems from the increasing pressures being placed on the coastal marine environment. Shipping, fisheries, recreation and a multitude of other uses and activities are pursued in the marine environ-

ment. Also coastal waters are generally the terminal point for terrestrial pollution from both point and diffuse sources, which brings further stress onto the coastal environment. Coastal marine waters are thus important economically and socially (see for example Cracknell, 1999; Siefert and Plattner, 2004). Increased awareness of the environment coupled with increased interaction between human and natural concerns necessitates a better understanding and monitoring of the marine environment to ensure human and natural priorities can both be accommodated. A forecasting system giving knowledge of the future states of the coastal marine environment would assist in its management. Demand for such systems is also driven by military or defense concerns: a recent Australian study into developing forecasting capabilities — BLUElink (<http://www.marine.csiro.au/bluelink/>) — was jointly funded by the Australian navy.

Modelling of coastal marine processes already assists in providing an understanding of how the system functions and predicting what will occur under different scenarios. The development of a forecasting system builds on the potential of numerical models, their ability to predict future events, and their use in operational management and decision making. Because of their potential, biogeochemical models are increasingly being used to assist in the management of estuaries and coasts (e.g., Harris *et al.*, 1996; Webster *et al.*, 2001).

Two examples of the potential for coastal marine forecasting are for suspended sediment and algae concentrations. During off-shore construction or dredging works suspended sediment is resuspended. At high concentrations the sediment can smother and damage plant communities on the sea bed. A forecast system would predict the suspended sediment concentrations and enable managers to adapt work patterns to reduce environmental damage. In the second example, increased nutrient loads to coastal receiving waters has resulted in more occurrences of toxic algal blooms (Vorosmarty *et al.*, 2004). A forecasting system would predict the onset of a bloom as well as predict its direction of movement once a bloom has occurred. Such knowledge would allow authorities to lessen the impact of the bloom by implementing action plans, closing beaches and fisheries etc.

Improved forecasting of temperature also improves coastal management. Wind driven coastal upwelling can deliver cold nutrient laden waters to the surface leading to high biological production. Such events are discernable by marked changes in

temperature observable by satellites (Stegmann *et al.*, 2006). Furthermore, ecological growth is sensitive to temperature (Solidoro *et al.*, 2003). Therefore improving temperature forecasts will allow for improved coastal management.

Existing operational systems and studies reflect the coastal protection and shipping concerns that have driven coastal forecasting to date (Flather, 2000). Although suspended particulate matter was investigated in the Pre-Operational Models In the Seas of Europe (PROMISE) study (Prandle, 2000a,b), existing operational forecasting systems (Flather, 2000, and also <http://ocean.dmi.dk>) more typically concentrate on water movement, currents, sea levels and occasionally temperature.

1.1.2 Supply of a coastal forecasting system

Recent improvements in model development, satellite observation systems and data assimilation techniques have improved the capacity for constructing an operational forecasting system. Corresponding advances in computing power have supported these improvements and made the construction of coastal marine forecasting systems feasible. However, an operational coastal biogeochemical system does not exist yet.

The difficulty in developing operational forecasting for biogeochemical systems is that the processes are much more complicated than say, sea level forecasting. The capability to model highly complex biogeochemical marine systems is rapidly developing but these models are only useful if they can be rigorously tested against appropriate observations, and until recently the availability of necessary observations has not kept pace with the model developments. Existing forecasting systems are based on the assimilation of in situ observations into numerical models, and “the single most important factor limiting the progress towards operational water quality [biogeochemical] models is a lack of data” (James, 2002).

New data sets available from satellites alleviate this significant problem in the advancement of forecasting system development. As Cracknell (1999) writes “... we regard the coastal zone as the last remaining important frontier [...] for the application of remote sensing [satellite] techniques.” Until recently satellite data was of little use for coastal work because it was too infrequent, was of too coarse a resolution and without significant multispectral capacity. The launch of a series of satellite platforms over recent years — ATSR, MERIS, MODIS, and SeaWiFS amongst oth-

ers — have changed this situation and the regular observation of marine data in the coastal region is now realistic. The possibilities for operational forecasting using satellite derived observations are described by Johannessen *et al.* (2000).

Another motivation for the use of satellite observations is that “in many practical applications, a correct representation of [spatial] concentration patterns is more important than the representation of local concentrations” (Vos and ten Brummelhuis, 1997). That is, the correct spatial pattern is more important than the exact value at a point. Satellite observations are useful because they contain the spatial pattern implicitly. Furthermore, higher complexity phenomena, transported by advection diffusion processes (James, 2002), move at a much slower rate than wave energy. Local effects acting in these models can therefore become much more important than local effects acting in hydrodynamic models. Because of larger local differences, spatial information becomes much more important for verification, evaluation and assimilation. The construction and maintenance of an in situ network fine enough to resolve local detail is usually prohibitively expensive, while satellite observations can provide such spatial coverage on a regular basis.

Complementary developments in the field of data assimilation, particularly with the development of ensemble techniques (Evensen, 1994) has made forecasting more feasible for higher resolution models. Previously assimilation was restricted to variational approaches — requiring adjoint models or minimisation techniques that are time consuming for large state space models — or simplified sequential techniques such as Optimal Interpolation and Nudging with the associated reduction of accuracy. The rapid increase of computational power has also made the development of coastal marine forecasting systems more feasible. Therefore, this thesis by linking these improvements from disparate fields makes a step towards an operational biogeochemical forecasting system.

1.2 Thesis Objectives and Scope

While the broad context of the thesis — the development of a forecasting system for coastal marine biological, chemical and physical processes utilising satellite observations — has been outlined, this undertaking would be unrealistic for a single thesis and would be the focus of teams of researchers. Therefore, the scope of this thesis

is limited to the assimilation of satellite sea surface temperature (SST) observations into a high resolution hydrodynamic model, resulting in a more achievable thesis objective.

As individually both the modelling and satellite observation of temperature is well advanced and the algorithms are reasonably well understood, the assimilation of temperature is more likely to be successful. The modelling and observations of other biogeochemical phenomena are less developed and contain larger uncertainties, the forecasting of these less understood phenomena can be undertaken at a later stage after demonstrating the success of assimilating well-known phenomena (temperature). Data assimilation of SST observation is undertaken in oceanographic studies, however, it does not appear to have been undertaken in high resolution coastal hydrodynamic models. A challenge for this thesis is to apply the data assimilation techniques to an application to a physically more complex setting. While data assimilation improves the states of the model, analysis of the assimilation provides feedback on the adequacy of the model structure and the data quality. This prompts a secondary goal to discover if model (or observation) deficiencies can be revealed through the data assimilation process.

As well as seeking to improve temperature forecasting, this thesis also attempts to discover an appropriate ensemble data assimilation configuration, and where deficiencies are encountered to develop the techniques further. A wide array of ensemble data assimilation techniques have been introduced, but little guidance is available regarding the suitability of the different options. While not a primary objective, an attempt is made to bring a degree of order to the ensemble data assimilation field

The specific objectives of the thesis are to:

- discover whether high resolution satellite (SST) observations can successfully be assimilated into a high resolution hydrodynamic model to improve the forecasting skill of the model,
- develop a data assimilation configuration appropriate for the case study undertaken,
- where appropriate to develop improvements to existing data assimilation techniques,

- discover if analysis of the data assimilation can reveal deficiencies in the model or observations, and
- synthesise the array of ensemble data assimilation techniques to develop a degree of order.

1.3 Approach

The objectives of the thesis are met through a case study of a coastal embayment in south eastern Australia. Port Phillip Bay (PPB) was chosen because it was the subject of a major scientific study in the mid 1990s — The Port Phillip Bay Environmental Study (PPBES) (Harris *et al.*, 1996) — the data from the study is available for this project and provides a solid basis of knowledge about how it operates. Furthermore, PPB is the subject of ongoing monitoring due to its being adjacent to the city of Melbourne.

The case study is performed in both synthetic and real settings. The synthetic setting allows various data assimilation configurations to be tested. This enables the discovery of the configuration appropriate for the PPB case study. In the real setting using the data assimilation configuration determined, actual SST observations are assimilated to test for improved forecasting skill.

Forecasting is the process of using a tested model, together with expected or estimated forcing conditions, to predict what will occur in the future. Other analogous terms are nowcasting and hindcasting (Prandle, 2000a). Formally, all the experiments in this thesis are hindcasts: hindcasting is the process of providing the best estimate of a historical event. Throughout this thesis, however, the more familiar term forecast will be used in a relative sense, and because improved forecasting is the motivation of this thesis. The use of forecasting techniques to improve predictions will be judged testing the techniques with a series of hindcasts, comparing the results against accurately observed data.

1.4 Thesis Outline

The research embodied in this thesis is divided into 7 chapters. Chapter 2 gives a literature review of the data assimilation field focusing on ensemble data assimilation techniques. This is followed by a chapter describing the case study site and the data collected. Chapter 4 covers initial testing of the model as well as developing a method to incorporate forecast error into the ensemble. The final two chapters describe the data assimilation experimentation done in both a synthetic and actual setting. The findings of the thesis are summaries in the conclusions, together with recommendations for future research.

The literature review begins with a detailed review of the general techniques and theoretical bases for data assimilation. This thesis concentrates on ensemble techniques and these are extensively explored, in particular the generation of forecast covariance error and the addition of forecast error into the ensembles. The review covers ensemble generation, propagation, analysis and introduces the well-known Ensemble Kalman Filter (EnKF) as well as the range of Ensemble Square Root Filters (EnSRF) and the practical issues related to ensemble data assimilation. Other issues are ensemble size and divergence. The remainder of the chapter covers applications of data assimilation in the context of the coastal marine environment, although touching on oceanographic applications.

The case study site is described in chapter 3, which also summarises and analyses all the data collected for research in this thesis. The analysis focuses on the uncertainty of the data, particularly the spatial variability of the meteorological data recorded at sites around PPB. This analysis enabled the maritime climate sites to be selected and separated from the more continental dominated sites. The satellite SST observations are also analysed, and contrasted with in situ temperature observations to discern any bias.

Chapter 4 describes the initial model testing and ensemble generation made prior to assimilation. The numerical model used throughout this thesis is described, together with its configuration for the PPB conditions. Testing and sensitivity experimentation of the model is also undertaken.

The data assimilation forecast equation is stochastic, while the PPB model is deterministic. Thus a method is needed to incorporate forecast error into the en-

semble members. This is achieved by developing a method to generate perturbed forcing data. The method adds error to the forcing data based on the uncertainty characteristic of the data. The perturbed forcing method is evaluated by comparing its correlation and standard deviation fields against those generated from an actual stochastic version of the PPB model.

In chapter 5, the assimilation of SST observations into the PPB model is first tested in a twin experiment environment. Synthetic observations are extracted from a “truth” model run, and assimilated into another model with degraded initial conditions. The ability of the assimilated run to replicate the truth confirms the potential of the assimilation system. The issues that are investigated are ensemble size, multivariate versus univariate assimilation, model bias, and the analysis method.

The assimilation configuration determined previously is applied in chapter 6 to the assimilation of actual SST observations. The assimilation gives improvements over a control and climatology forecast. However, evidence of a bias in the SST observations was observed, due to the skin-bulk temperature difference. Attempts to remove this observational bias are made using diurnal warming algorithms from the literature, and conditional merging, a technique that combines spatial observations with point observations.

An analysis of the innovations and the error predictions pointed to deficiencies in both the model and the observations. This illustrates the power of data assimilation: not only improves forecasts, but to suggest improvements in biogeochemical models as well.

The conclusions are made in chapter 7. These are presented in three parts, the first being a brief summary of the thesis. This is followed by a findings and discussion section covering, in turn, data assimilation, numerical modelling, and observations. Finally recommendations for further research are made.

Chapter 2

Data Assimilation

Data assimilation is a collection of statistical techniques that can be used to improve the model predictions by merging the model estimates with information from observations. Numerical models can be used as a tool to predict future events, and make predictions with a good spatial and temporal resolution, but are subject to errors from a range of sources: understanding of the physical processes, errors in the model development and errors with the initial and boundary conditions. Observations of physical states are generally on a coarser temporal and spatial resolution than coastal models, yet their values are often more reliable when available. Assimilating the observations into a model is expected to give the best estimate of the current state of the system, which should then allow for the best forecast of the future conditions.

This chapter presents a review of data assimilation literature. The first part introduces data assimilation techniques with an emphasis given to sequential ensemble techniques. The second part details the current status of coastal marine data assimilation.

2.1 Introduction

The need for data assimilation is derived from the fundamental understanding that model predictions are imperfect and are only accurate within an uncertainty range. The model errors are due to i) model development errors (see for instance Dee, 1995),

ii) errors in the initial and boundary conditions, and iii) errors in the calibration or testing of the model. Because of these errors, the model prediction of a state will be in disagreement with the true value of the state and also with an independent observation of the state.

The purpose of data assimilation is to adjust the model prediction closer towards the true state. There are three levels of adjustment that can be made to improve the model prediction. Firstly, if the independent observation of the state is assumed to be more accurate than the model prediction, then model prediction will be closer to the truth if the disagreement between the model and the observations is added to the model. Essentially, this involves replacing the model states with the observed values.

The second level is to recognise that the independent observation has a degree of uncertainty also, due to measurement error. Therefore, instead of adding the disagreement between the model and the observation to the model, a proportion of the disagreement is added, determined by the ratio of the observation uncertainty to the model uncertainty. The resulting model state is then closer to the most certain estimate of the true state (whether the model prediction or the observation).

The final level to data assimilation is to consider that not all model states are observed and therefore cannot be adjusted directly. In this case the error covariance between the unobserved state and an observed state are calculated. If the errors of these states are highly correlated, the misfit at the unobserved states will vary in similar manner to those at the observed state. Therefore, an unobserved state can be improved by the addition of the disagreement between the model and an observation at an observed state, adjusted by their relative uncertainties and by the covariance between the observed state and the unobserved state.

Data assimilation, therefore improves the model prediction towards the truth, but requires a number of elements: model predictions, observations, knowledge of the uncertainties of both and their error covariances. The determination of the uncertainties and error covariances confound the application of data assimilation. This chapter explores data assimilation techniques, its application in the coastal marine environment, and advises on the techniques most suitable for this thesis. Finally, this thesis confines itself to sequential data assimilation. Another data assimilation approach is variational assimilation, but this will not be covered. The interested

reader is referred to the papers by Talagrand and Courtier (1987); Courtier and Talagrand (1987); Thacker and Long (1988) for details about variational assimilation.

2.2 Sequential Filtering

Sequential data assimilation improves a forecast by using independent estimates of the model state and observations to adjust the model closer to the true system state. The updating is based upon the choice of data assimilation technique and knowledge of the relative uncertainty of the model prediction and observation. Without updating the states as new information becomes available, forecasting would continuously compound model errors leading to poor predictions. By using data assimilation to update the states, the predictive capabilities of a model are maximised.

The concept of sequential data assimilation analysis, combining observations and a model to improve the estimate of a state, can be simply understood by investigating the best linear unbiased estimator (BLUE). Consider two independent estimates of a state x , x_1 and x_2 , which have variances σ_1^2 and σ_2^2 respectively. If w is a weighting factor such that,

$$x = (1 - w)x_1 + wx_2, \quad (2.1)$$

then following Barham and Humphries (1970), the best (in a least squares sense) linear estimator, x^* , of the true state of x is given by

$$x^* = x_1 - w^*(x_1 - x_2), \quad (2.2)$$

where w^* is given by

$$w^* = \frac{\sigma_1^2}{\sigma_1^2 + \sigma_2^2}, \quad (2.3)$$

and the best estimator has a variance of

$$\sigma^{*2} = \sigma_1^{*2}(1 - w^*). \quad (2.4)$$

Equation (2.2) shows that the best estimate of x is given by the average of the two estimates, x_1 and x_2 , weighted by their relative variances. This ensures that the more certain estimate is more heavily weighted than the less certain estimate.

Equation (2.4) shows that the variance of the resulting estimator is reduced; the best estimator is more certain than either of the original independent estimates. In the same manner data assimilation seeks the best estimate of a state by combining the two independent state estimates derived from the model and the observation. The issue for data assimilation is calculating the relative weighting.

Sequential data assimilation is introduced through the Kalman filter, a technique first proposed in the 1960's (Kalman, 1960). The concept behind the Kalman filter is that updating an uncertain forecast with uncertain observations improves the forecast. As a sequential approach, the Kalman filter process can be separated into two stages: a forecast stage and an analysis (correction) stage. The interested reader can find a derivation of the Kalman filter equations presented in the following two sections in Barham and Humphries (1970).

2.2.1 Forecast

The Kalman filter assumes that the underlying process being modelled can be described as a Markov process

$$\hat{\mathbf{x}}_{k+1} = \Phi_k \hat{\mathbf{x}}_k + \mathbf{G}_k \mathbf{u}_k, \quad (2.5)$$

where $\hat{\mathbf{x}}_k$, an $n \times 1$ vector, is the true state at time t_k . Φ_k , an $n \times n$ matrix, describes the known equations of the process propagation from t_k to t_{k+1} . \mathbf{u}_k , a $l \times 1$ vector represents random noise entering the system through l sources of error with a mean of zero and variance given by \mathbf{Q}_k , a $l \times l$ matrix. \mathbf{G}_k is an $n \times l$ matrix which relates the random noise to the states. Equation (2.5) indicates that the true state of a system is made up of a known component and a random or unknown component. Observations of this system are described by

$$\mathbf{d}_{k+1} = \mathbf{H}_{k+1} \hat{\mathbf{x}}_{k+1} + \mathbf{n}_{k+1}, \quad (2.6)$$

where \mathbf{d} is an $m \times 1$ vector representing m observations and \mathbf{n}_{k+1} (an $m \times 1$ vector) is measurement noise with a mean of zero and covariance of \mathbf{R}_{k+1} , an $m \times m$ matrix with a value of $E\langle \mathbf{n}_{k+1} \cdot \mathbf{n}_{k+1}^T \rangle$. This matrix is known as the observation error covariance. Generally, not all the model states are observed and so the observation estimates

are related to the model states through the forward observation operator \mathbf{H}_{k+1} , an $m \times n$ matrix that transforms the states to the observations. Equation (2.6) states that the observation is some function of the truth distorted by the measurement noise. Importantly, both the model and observations are assumed to be unbiased.

Starting from a previous analysis (best estimate) of the states, the best forecast of the states at a future time is given by

$$\mathbf{x}_{k+1}^f = \Phi_k \mathbf{x}_k^a, \quad (2.7)$$

where the superscripts a and f represent analysis and forecast respectively. The best estimate for the forecast covariance of the state errors is given by

$$\mathbf{P}_{k+1}^f = \Phi_k \mathbf{P}_k^a \Phi_k^T + \mathbf{G}_k \mathbf{Q}_k \mathbf{G}_k^T, \quad (2.8)$$

where the superscript T indicates the matrix transform.

The forecast (or background) error covariance, \mathbf{P}_k an $n \times n$ matrix, increases during the forecast with an additional component of variance due to the variance of the model noise. Importantly, this implies that the uncertainty associated with a noisy model prediction will increase over time, just as the analysis will act to reduce it.

2.2.2 Analysis

The Kalman filter analysis is simply an extension of BLUE from a scalar to a vector case. In this situation there are two independent estimates; an estimate derived from a model, \mathbf{x} , and an estimate derived from an observation, \mathbf{d} with covariances given by \mathbf{P} and \mathbf{R} for the model and observation respectively. For multiple dimensions equations (2.2) and (2.4) become

$$\mathbf{x}^a = \mathbf{x}^f + \mathbf{K} [\mathbf{d} - \mathbf{H}\mathbf{x}^f], \quad (2.9)$$

$$\mathbf{P}^a = \mathbf{P}^f - \mathbf{K}\mathbf{H}\mathbf{P}^f = [\mathbf{I} - \mathbf{K}\mathbf{H}] \mathbf{P}^f, \quad (2.10)$$

with

$$\mathbf{K} = \mathbf{P}^f \mathbf{H}^T [\mathbf{H}\mathbf{P}^f \mathbf{H} + \mathbf{R}]^{-1}, \quad (2.11)$$

where the subscripts denoting time have been removed in equations (2.9–2.11) for clarity and \mathbf{I} is the identity matrix. \mathbf{K} , an $n \times m$ matrix, is known as the Kalman gain and, by contrasting equation (2.11) against equation (2.3), can be seen as the multi-dimensional equivalent of w^* . These equations give the best linear (least squares) estimate of the true states of a system assuming an unbiased model and unbiased observations. A benefit of the Kalman filter is that as well as providing the best estimate of a system state, the forecast covariance matrix indicates the uncertainty associated with the estimate. This allows for confidence or otherwise in the estimate.

2.2.3 Kalman Filter

Thus using a linear model Φ_k , its noise characteristics \mathbf{Q}_k , a set of system observations \mathbf{d}_k , and its noise characteristics \mathbf{R}_k , the Kalman filter predicts the best most likely (in a squared root sense) future states \mathbf{x} , given an initial estimate of the states \mathbf{x}_0 and their uncertainty \mathbf{P}_0 . This is done by making repeated forecasts which are corrected through the analysis of observations as they become available. The repeated forecast and analysis steps of the sequential approach is shown in figure 2.1. A forecast of a state, \mathbf{x} , is made from a previous analysis or best guess estimate. During the forecast the uncertainty associated with the forecast, \mathbf{P} , increases. When an observation becomes available, an analysis is performed. The state is updated to incorporate information from the observation and the uncertainty of the state estimate reduces accordingly. This process is then repeated.

2.3 Kalman Filter Simplifications

While the Kalman filter gives the best linear estimate of a future state, its application in numerical models has practical limitations. Systems with few states can apply the Kalman filter, but for large systems, with many states, its application becomes very burdensome. This was particularly true in previous decades when computing power was very limited, and is relevant today where large distributed models are used. Two calculations are of particular concern, one is the matrix inverse in equation (2.11) and the other is the storage and propagation of \mathbf{P} in equation (2.8). Furthermore, it

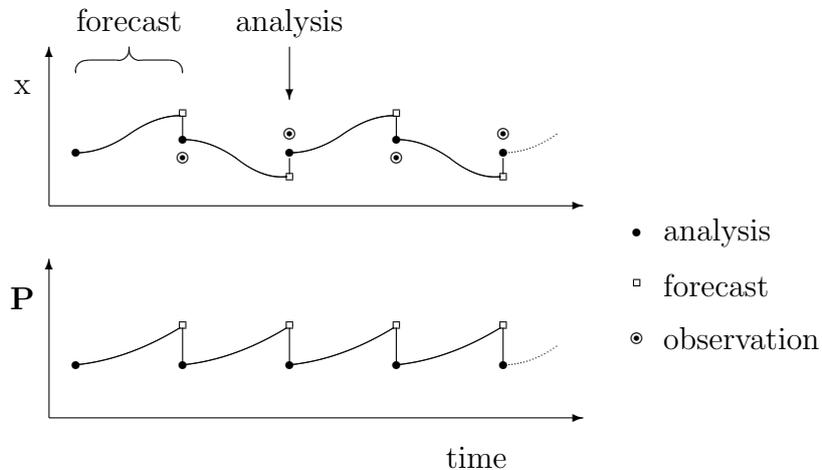


Figure 2.1: Schematic of the Kalman filter showing repeated forecast and analysis steps.

is often difficult to estimate \mathbf{Q} and the need to linearise a nonlinear model becomes a challenge as well. Because of the practical difficulties in applying the full Kalman filter, simplifications have been devised for operational use.

The essence of sequential approximations is to simplify the Kalman gain replacing the dynamic \mathbf{K}_k with a static version $\tilde{\mathbf{K}}$. A selection of these techniques will be briefly described to illustrate the developments in data assimilation. Further discussion on these simplifications to the Kalman filter is available from Dee (1991), De Mey (1997) and Walker and Houser (2005).

Direct Insertion

This method assigns all certainty to the observations, replacing the forecast model states, \mathbf{x}^f , with the observations, \mathbf{d} . This is essentially setting the Kalman gain to unity. While this has been successfully used — an example is the assimilation of sea surface height data (Haines, 1991) — a disadvantage of the method is that by changing the state values only at observations the dynamics of the system are changed, possibly inducing instability. Furthermore, the information contained in the observations is not optimally spread to other (unobserved) parts of the model domain.

Successive Correction

Slightly more advanced, Successive Correction assumes that the observations are more accurate than the model and iteratively nudges the model towards the data. Weights, $W(s)$, are chosen to smooth the model states towards the observations over s_n iterations. A problem is that, for some choices of $W(s)$ and s_n , the solution may not converge to the truth and for large values of s_n the analysis may not be consistent with the dynamics of the system. For more information see Daley (1991, chapter 3).

Successive correction has been used for the assimilation of sea surface height, combining in situ (buoy) and remotely sensed (altimetric) data to yield improvements over the assimilation of altimetric data alone (Ishikawa *et al.*, 1996). Temperature and velocity data were assimilated by successive correction in a simulated study of the Indian ocean (Moore *et al.*, 1987), although there were problems with resolving the horizontal temperature structure which were attributed to the assimilation scheme.

Analysis Correction

Analysis Correction is similar to successive correction with an iterative approach being employed. The improvement is that at each iteration the observations as well as the model states are adjusted. For certain choices of weights, analysis correction converges to optimal interpolation. Analysis correction was introduced into the British Meteorological Office for weather prediction in the early 1990's (Lorenc *et al.*, 1991), being faster and more simple than a previous optimal interpolation system. Although, now the Meteorological Office uses a 3-dimensional variational assimilation scheme (Lorenc *et al.*, 2000).

Optimal Interpolation

Optimal Interpolation (OI) provides an approximate solution to Kalman filtering by assuming a time-constant forecast error covariance, $\tilde{\mathbf{P}}$, which leads to a gain approximation,

$$\mathbf{K} = \tilde{\mathbf{P}}\mathbf{H}^T \left[\mathbf{H}\tilde{\mathbf{P}}\mathbf{H}^T + \mathbf{R} \right]^{-1}. \quad (2.12)$$

There are two approaches to determining $\tilde{\mathbf{P}}$. If calculating the inverse of equation (2.12) is not too difficult, $\tilde{\mathbf{P}}$ can be determined empirically from the asymptotic limit of a time-varying estimate of the forecast error covariance matrix, \mathbf{P}_k (Fu *et al.*, 1993). The Kalman filter is run off-line calculating values of \mathbf{P} , and the asymptotic limit of these are used in a real-time application. By employing a static forecast error covariance matrix, it is no longer necessary to propagate \mathbf{P} in time, substantially reducing computational cost.

If the cost of the matrix inversion is too great, $\tilde{\mathbf{P}}$ can be determined by a correlation function, dependent upon the distance between the grid points and observations, where correlation length scales are adjusted empirically. With this approach, only observations in a small region around grid points are used for analysis. This leads to a reduced block structure for \mathbf{K} , which reduces computation (Bouttier and Courtier, 1999). This is a form of localisation which will be discussed later. Optimal interpolation has been widely applied in the literature and further discussion can be found in Daley (1991, chapter 4).

State-space Reduction

In attempting to make the Kalman filter more practically applicable, many authors have developed techniques to reduce the state space of the model (Fiegunth *et al.*, 2003; Verlaan and Heemink, 1997) or its parameters (Navon, 1997). With a smaller state space, the Kalman filter equations are easier to handle. Specifying the error covariances is also easier (Dowd and Thompson, 1997; Hoang *et al.*, 1997). A popular approach is to reduce the state by selecting only the most significant modes of the Empirical Orthogonal Functions (EOFs). EOFs are the eigenvectors and eigenvalues of a data set (for more information see Preisendorfer, 1988). This approach has been followed by Cane *et al.* (1996), Braconnot and Frankignoul (1993) and Webster and Narayanan (1990). While the resulting reduced state-space is more computationally efficient, this is achieved at the loss of detail in the relationships between the states (or parameters) that may be important.

2.3.1 Sequential Filter Choice

There is a clear choice in the selection of a data assimilation technique between efficiency and accuracy. Simpler methods are much faster but suffer a lack of accuracy as a result. Houser *et al.* (1998), albeit an investigation into soil moisture assimilation, suggests using simpler methods because of data complexity trade-off, whereas Ballabrera-Poy *et al.* (2001) preferred more sophisticated methods, such as the Kalman filter, which are more likely to produce optimal estimates and provide quantitative error estimates.

The use of such simpler methods has often been preferable. Fu *et al.* (1993) compared an optimal interpolation forecast with the Kalman filter and found no statistical difference between the two, but with greatly reduced computational cost for the optimal interpolation. In this case it seemed that the noise inherent in the system hindered the more sensitive Kalman filter.

These simplification of the linear Kalman filter will not be explored in this thesis because of the static simplifications they employ. While these might be suitable for oceanic assimilation, where a geostrophic balance generally exists, in dynamic coastal environments such a steady state simplification is not appropriate. Furthermore, forecast covariance relationships in coastal regions are likely to be anisotropic, with correlation structures following the coastline. This diminishes the potential to use correlation length-scales to simplify the forecast covariance relationships. For these reasons the simplified Kalman Filter approaches will not be considered further in this thesis and instead ensemble filtering techniques are explored.

2.4 Ensemble Filters

In addition to the computational expense, a significant problem for the Kalman filter is that it is only applicable to linear systems. Unfortunately, most environmental models are highly nonlinear and linearisation of these models may lead to significant degradation of the state and covariance forecasts. Moreover, the variance distribution is assumed Gaussian; if this distribution is subjected to nonlinear propagation, the Gaussian distribution will be distorted invalidating the assumptions made to derive the filter. Linearising a nonlinear relationship introduces errors that will make

the assimilation no longer optimal.

An approach to solving this problem is to approximate nonlinear operators by their tangent linear approximations in the vicinity of \mathbf{x}^a (Bouttier and Courtier, 1999; Evensen, 1992). This approach is known as the Extended Kalman Filter, and is analogous to taking a first order Taylor series of the Kalman filter, where higher order terms have been discarded. Unfortunately, when these higher order terms become important the utility of the extended Kalman filter reduces.

The ensemble Kalman filter (EnKF) was introduced to account for nonlinear effects and failings of the extended Kalman filter (Evensen, 1994). Since its introduction other ensemble filters have been proposed with the differences in the method the analysis is calculated. These will be covered in the following sections. The improvement of ensemble filters over other sequential filters was that the states and forecast error covariance are propagated using the nonlinear model equations as an ensemble of forecasts. As the nonlinear model equations are used to propagate the states and covariance, no loss of higher order nonlinear terms is involved. Furthermore, using clever linear algebra methods the large \mathbf{P} matrix need never be explicitly calculated, saving computer memory. The formulation of the EnKF enables further improvements to computational speed through parallel implementation. For instance, each ensemble member run could be processed by a separate processor. The implementation of a parallel EnKF has been described by van Hees *et al.* (2003) and Keppenne (2000).

It has long been established that the average of an ensemble forecast is more accurate than a single forecast, however, the use of an ensemble in a Kalman filter was novel. Ensemble forecasting has been implemented in operational weather forecasting since 1992 (Toth and Kalnay, 1997; Molteni *et al.*, 1996), but although an ensemble forecast was made, operational analysis was still performed by optimal interpolation or a variational approach. Since its introduction the EnKF has been used in a wide variety of applications summarised by Evensen (2003).

As a sequential filter, the EnKF shares the repeated forecast and analysis steps of the Kalman filter; see figure 2.2. The difference is that multiple model runs are made giving an ensemble of forecasts. The spread of these forecasts replaces the explicit propagation of the forecast error covariance: contrast figure 2.2 with figure 2.1. While the replacement of \mathbf{P} eases the computational burden, the introduction of an

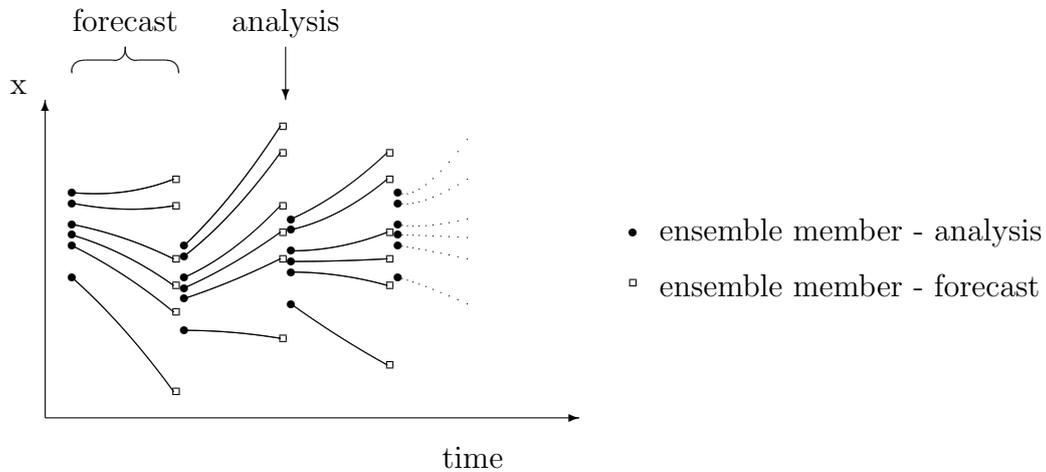


Figure 2.2: A graphical representation of the EnKF, with repeated analysis and forecast steps.

ensemble raises other issues such as having a sufficient number of ensemble members to characterise the state covariances, and ensuring the forecast runs spread the ensemble members appropriately.

2.4.1 Analysis

Since the introduction of the EnKF, its form has modified as improvements were suggested (e.g. Burgers *et al.*, 1998; Evensen, 2003, 2004). The 2003 paper by Evensen will be used as a starting point for discussion. The ensemble of model forecasts is created by replacing the state vector \mathbf{x} , with an ensemble of n_e state vectors, \mathbf{X} , an $n \times n_e$ matrix given by

$$\mathbf{X} = [\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_{n_e}]. \quad (2.13)$$

The ensemble can be separated into a mean and a perturbation matrix with the ensemble mean given by

$$\bar{\mathbf{X}} = \mathbf{X} \mathbf{J}_{n_e}, \quad (2.14)$$

where \mathbf{J}_{n_e} is an $n_e \times n_e$ matrix in which every element is $\frac{1}{n_e}$. The ensemble anomaly matrix is the difference between the ensemble and the ensemble mean,

$$\mathbf{X}' = \mathbf{X} - \bar{\mathbf{X}}. \quad (2.15)$$

The ensemble anomaly matrix represents the variation in the ensemble forecast about the mean. This matrix will be referred to as the ensemble spread as it describes the spread of the ensemble members about the mean. This allows the forecast error covariance, \mathbf{P} , to be approximated by

$$\mathbf{P}_e = \frac{\mathbf{X}'\mathbf{X}'^T}{n_e - 1} \quad (2.16)$$

where the subscript e indicates an ensemble approximation. This expression is the characterising feature of ensemble filtering techniques. The use of an ensemble approximation of the forecast error covariance matrix, \mathbf{P}_e , is based on the assumption that in the limit of an infinite number of ensembles members

$$\lim_{n_e \rightarrow \infty} \mathbf{P}_e = \mathbf{P}. \quad (2.17)$$

This approach finds resonance with the error subspace statistical estimation (ESSE) formulation of Lermusiaux and Robinson (1999), where a subspace of \mathbf{P} — equivalent to \mathbf{X}' — was estimated from the dominant n_e eigenmodes of a Singular Value Decomposition (SVD) (see Heath, 1997, p. 134) of \mathbf{P} . However, the ultimate source of \mathbf{P} is unclear, as reliable estimates of the initial error covariance are rarely available (Miller and Ehret, 2002). Similar to equation (2.17), the use of more eigenmodes leads to a more accurate estimate of the error subspace of \mathbf{P} . To maintain the correct forecast error covariance post-analysis, the appropriate spread of the ensemble members is required. This is achieved through the use of an ensemble of perturbed observations (Burgers *et al.*, 1998) that is given by \mathbf{D} an $m \times n_e$ matrix with

$$\mathbf{D} = [\mathbf{d} + \varepsilon_1, \mathbf{d} + \varepsilon_2, \dots, \mathbf{d} + \varepsilon_{n_e}], \quad (2.18)$$

where ε_i is a $m \times 1$ vector of Gaussian random perturbations with zero mean and standard deviation given by the uncertainty of the observation itself. Without the

use of perturbed observations the ensemble approximate analysis forecast error covariance matrix, \mathbf{P}_e^f will be deficient by an amount of $\mathbf{K}\mathbf{R}\mathbf{K}^T$ (Tippett *et al.*, 2003).

In the Evensen's 2003 paper, an approximation to the observation error covariance matrix is generated

$$\mathbf{R}_e = \frac{\Upsilon\Upsilon^T}{n_e - 1}, \quad (2.19)$$

where Υ is an $m \times n_e$ matrix of the perturbations used to generate the perturbed observations, $[\varepsilon_1, \varepsilon_2, \dots, \varepsilon_{n_e}]$. The use of this estimation as the observation error covariance matrix has been discredited by Kepert (2004). Thus, observations error covariance, \mathbf{R} , rather than its ensemble approximation, \mathbf{R}_e , will be used throughout this thesis. As the error associated with an observation is frequently known, there is no need to incorporate an ensemble estimate.

In ensemble form the Kalman filter analysis step, equations (2.9) and (2.11), becomes

$$\mathbf{X}^a = \mathbf{X}^f + \mathbf{P}_e \mathbf{H}^T [\mathbf{H}\mathbf{P}_e \mathbf{H}^T + \mathbf{R}]^{-1} [\mathbf{D} - \mathbf{H}\mathbf{X}^f]. \quad (2.20)$$

As \mathbf{X} contains information on both the mean state and its variance, the forecast error covariance, equation (2.10), is updated implicitly by equation (2.20) through the analysis of individual ensembles. This is achieved practically by substituting equation (2.16) into equation (2.20) giving

$$\mathbf{X}^a = \mathbf{X}^f + \mathbf{X}'\mathbf{X}'^T \mathbf{H}^T [\mathbf{H}\mathbf{X}'\mathbf{X}'^T \mathbf{H}^T + (n_e - 1)\mathbf{R}]^{-1} [\mathbf{D} - \mathbf{H}\mathbf{X}^f]. \quad (2.21)$$

A method for efficiently solving this equation is presented by Evensen (2003).

2.4.2 Forecast

An ensemble forecast is made by propagating each of the ensemble members in time using a numerical model. The forecast mean and variance are implicitly calculated as the mean and spread of the ensemble members. A challenge for all sequential assimilation techniques is correctly incorporating forecast error, the variance, \mathbf{Q} , which is generally difficult to assess. Failure to correctly account for forecast error leads to errors in the calculation of the forecast error covariance via the second right-hand term of equation (2.8). While ensemble methods do not calculate equation (2.8)

explicitly, without a mechanism to continually introduce realistic forecast error into the ensemble propagation the ensemble spread will collapse because at each analysis step the ensemble spread is reduced according to equation (2.10). This will reduce the optimality of the ensemble forecast.

For the purpose of this thesis, it is useful to distinguish two sources of forecast error. The first is model error, which covers errors associated with the structure of the model — for instance, errors introduced during the model development process outlined by Dee (1995) — as well as errors with incorrect parameter selection due to insufficient calibration. Issues with model errors include whether calibration can assign unique values to parameters (Janssen and Heuberger, 1995) and what are the limits to model testing and model validation (Power, 1993; Oreskes *et al.*, 1994; Rykiel, 1996; Balci, 1997). The second source of error is termed forcing error. These are errors due to both the initial conditions the model is subjected too, but also the forcing (or boundary) data that drives the model. Errors in the forcing data will necessarily impact on the accuracy of a model forecast.

While in theory the issue of forecast error is straight forward, the reality is less so. The challenge is determining the correct level of error to be introduced, as well as introducing it so as not to disturb the dynamics of the model.

Model Error

In his review paper Hamill (2002) lists three approaches to incorporating model error. These are i) using stochastic equations, ii) adding noise to the ensemble at the analysis time and iii) using multi-model ensembles.

The first approach involves reforming deterministic model equations into stochastic equivalents by introducing stochastic error terms into the model equations directly. According to Hamill (2002), little attention has been given to this approach, although Buizza *et al.* (1999) used it for their study and it was also discussed by Lermusiaux and Robinson (1999). A possible reason for lack of attention to this approach is that introducing stochastic equations defeats one of the implicit advantages of the EnKF: that the analysis is independent of the model used and that no model re-coding is necessary. Poorly known model parameters can also be made stochastic (Evensen, 2003).

The second approach of adding noise at the analysis time-step has been the most actively pursued. This formulation was used in the original EnKF (Evensen, 1994), and has subsequently been widely used (e.g. Keppenne, 2000; Mitchell and Houtekamer, 2000).

This is the standard EnKF approach of Evensen (2003); the ensemble forecast is made by a deterministic model subject to stochastic noise,

$$\mathbf{x}_{k+1} = f(\mathbf{x}_k) + \sqrt{\Delta t} \sigma \rho \mathbf{v}_k, \quad (2.22)$$

where Δt is the time step, σ is the standard deviation of the model error, \mathbf{Q} . \mathbf{v}_k is a $n \times 1$ vector of random perturbations with a mean of zero and variance of unity, ρ is a factor which controls the covariance of \mathbf{v}_k in time. The difference between equation (2.22) and equation (2.5) is that while Φ is linear, $f(\mathbf{x})$ need not be.

Generally, fast Fourier transform methods are used to generate spatially correlated smooth random fields with zero mean and specified standard deviation. Although, details as to the method of appropriately specifying the standard deviation, σ , remain vague. A further deficiency with this approach is that this method can generate physically unrealistic ensemble members. This deficiency is investigated further in this thesis and prompts the development of a method for incorporating forecast error described in chapter 4.

The third approach of using different models to produce the various ensemble members has been often discussed but not widely explored. Different models have different assumptions and equations often leading to a wide range of model predictions, which can be attributed to model error. Although initial attempts with this method have been disappointing (Hamill, 2002).

Forcing error

The impact of forcing data as a source of forecast error is not usually considered explicitly. This oversight is probably due to the mind set of the oceanic and atmospheric data assimilation community, where work is predominantly on large scale chaotic models. Although forcing data may be included in these models, it is generally of secondary importance and restricted to the initial state. Model error due to nonlinear equations dominates the forecast error. This mind set is starting to

change. For instance, in a recent paper by Borovikov *et al.* (2005) all the ensemble variability of an oceanic data assimilation experiment was derived from perturbed atmospheric forcing data.

For coastal biogeochemical models the influence of forcing data is much more important. Many of the systems modelled in the coastal marine environment are dispersive. This means that without external influences the systems evolve to a steady state spatially mean conditions. For instance without open boundary or wind forcing, the currents of a hydrodynamic model diminish as friction losses accumulate. Similarly without current and wave movement to keep sediment particles suspended and in the absence of new sources all the sediment in a water body eventually deposits on the bed. Due to the importance of forcing data in coastal biogeochemical models, errors in the forcing data will be an important source of forecast errors. Thus introducing error through forcing data is another mechanism to introduce forecast error into the ensemble.

Errors in forcing data are associated with the measurement of forcing data and its spatial representation. While discussion of forcing error is rare in the literature, recently it has been given more attention (Reichle *et al.*, 2002b; Robert and Alves, 2003; Brusdal *et al.*, 2003; Natvik and Evensen, 2003a). These papers generally include forcing error by adding Gaussian random noise to the forcing fields with a specified standard deviation, although the treatment of perturbed forcing appears to have been made in a simplistic manner. For instance, Reichle *et al.* (2002b) selected the size of the perturbations they added based on “simple order-of-magnitude considerations”. As such there appears to be considerable scope to deal with perturbed forcing data more rigorously.

2.4.3 Ensemble Generation

The uncertainty of the initial state estimates is represented by the initial spread of the ensemble members. In the method outlined by Evensen (2003) ensemble members are generated by taking an initial best-guess of the states, and then adding perturbations in the form of random correlated fields to each ensemble member. Importantly, this approach includes a recommendation to “integrate the ensemble over a time interval covering a few characteristic time scales of the dynamical system”

to ensure dynamic stability and correct multivariate correlations before commencing the assimilation. This approach is the basis of several papers (Houtekamer and Mitchell, 1998; Keppenne, 2000; Reichle *et al.*, 2002a).

An improved sampling scheme has been proposed by Evensen (2004), based upon the work of Pham (2001). This method uses an ensemble of randomly generated, spatially correlated fields. Perturbation independence is sought by performing a Singular Value Decomposition (SVD). The first n_e singular vectors are then combined with another random orthogonal matrix and the singular values are adjusted appropriately. Zupanski *et al.* (2006) attempts to address the initialisation problem more explicitly, extending on the previous methods. However a disadvantage of all the methods mentioned above is that they are applied prior to the assimilation period and require a spin up for dynamic stability at which time the prescribed error distribution may have been altered by the model equations.

Two additional methods are available from operational ensemble forecasting: the breeder method and optimal perturbations. The purpose of these two methods is to generate a set of the fastest growing errors. The two methods have been investigated in a paper by Miller and Ehret (2002) which studied the forecasting of multimodal systems with small ensemble sizes. They found that the optimal perturbations (also termed singular vectors) method performed well, although less well for systems with a large initial variance. In such cases the breeder method performed well, although it may fail to observe the bimodal evolution.

The breeder method introduced by Toth and Kalnay (1993), and utilised by the National Centers for Environmental Prediction (NCEP) for atmospheric modelling for operational weather forecasting (Toth and Kalnay, 1997), is based on the concept that errors within an analysis can be classified as either growing or non-growing. By focusing on the growing states through breeding, the leading local Lyapunov vectors are estimated. Local Lyapunov vectors are the vectors which grow asymptotically fastest in the period up to the analysis, and thus are likely to dominate the growth of analysis errors (Toth and Kalnay, 1997). These vectors are treated as perturbations and are added to the initial best guess of the states to generate the initial spread. A disadvantage of the method is that independence is not guaranteed and that the vectors may tend to cluster towards a preferred, fastest-growth direction. The failure to observe bimodal evolution reported by Miller and Ehret (2002) may

have been caused by this behaviour. Also experiments using the breeder method for ensemble forecasting have reported that it underestimates the forecast ensemble variance (Miller and Ehret, 2002; Wang and Bishop, 2003). Further discussion of Lyapunov vectors and growth of perturbations is given by Farrell and Ioannou (1999).

Optimal perturbations are utilised by the European Centre of Medium-Range Weather Forecasting (ECMWF) in the reduced resolution T63L19 version of the T213L31 forecast model. The basis of the method is that within a time period where the dynamics of perturbations are assumed to be linear, the directions of maximum instability are given by the singular vectors of the linear propagator multiplied by its adjoint. Growth by this method can be much faster than Lyapunov vectors (Molteni *et al.*, 1996). A disadvantage of this method is that it requires an adjoint model, which is not always available and may be difficult to construct.

2.4.4 Ensemble Size

Determination of the number of ensemble members required is an unresolved issue in ensemble filtering. Indeed many studies have focused on the sensitivity of an ensemble forecast system to ensemble size (e.g. Houtekamer and Mitchell, 2001). There appears two groups of opinion; those who use smaller ensemble sizes (e.g. Mitchell *et al.*, 2002; Whitaker and Hamill, 2002) and those who advance the use of larger ensemble sizes (e.g. van Leeuwen, 1999). In studying atmospheric ensemble forecasting with the T62/18 level version of the NCEP medium range forecast model, Toth and Kalnay (1997) found that in terms of forecast skill the greatest improvement in an ensemble mean forecast was in increasing from 4–10 ensemble members, and that little improvement was found over 20 ensemble members. This finding, they claimed, was in agreement with previous studies. However, for higher order moments certainly more than 40 ensemble members are useful.

Determining the appropriate ensemble size involves a trade-off: The accuracy of the covariance estimate increases with the number of ensemble members, however additional ensemble members increases the computational expense propagating them. This is a matter of efficiency; users seek those methods that allow for small ensemble sizes with the best performance (Miller and Ehret, 2002). Both ensemble

generation (e.g. Evensen, 2004) and the type of ensemble sequential method used (e.g. Whitaker and Hamill, 2002) impact on the efficiency. The preceding discussion assumes that the system noise is being accurately modelled. If this is not the case, increasing the ensemble size will not necessarily improve the assimilation.

Model and Observation Bias

The Kalman filter and subsequent EnKF equations are derived assuming both unbiased model and observations. Unbiased signifies that there is no systematic error associated with the forecast or observations

$$\mathbf{E}\langle \mathbf{u}_k \rangle = \mathbf{E}\langle \mathbf{n}_k \rangle = 0. \quad (2.23)$$

In reality, however, bias may exist in both models and observations due to incorrect parameters, numerical dispersion and faulty boundary conditions (Dee, 1995). As observations are collected independent of the data assimilation system, pre-processing should be targeted to identify and remove any observational bias.

A method to explicitly account for and remove forecast (or model) bias from a sequential filter was developed theoretically by Dee and da Silva (1998) and subsequently refined by Dee and Todling (2000). This is achieved by replacing equation (2.9) with

$$\mathbf{b}_k = \mathbf{b}_{k-1} + \mathbf{L}_k [\mathbf{d}_k - \mathbf{H}_k(\mathbf{x}^f - \mathbf{b}_{k-1})], \quad (2.24)$$

$$\mathbf{x}_k^a = \mathbf{x}_k^f + \mathbf{K}_k [\mathbf{d}_k - \mathbf{H}_k(\mathbf{x}_k^f - \mathbf{b}_{k-1})], \quad (2.25)$$

where \mathbf{b}_k , an $n \times 1$ vector, is the estimated forecast bias. The bias gain matrix ($n \times m$) is given by

$$\mathbf{L} = \mathbf{P}_e^b \mathbf{H}^T [\mathbf{H} \mathbf{P}_e^b \mathbf{H}^T + \mathbf{H} \mathbf{P}^f \mathbf{H}^T + \mathbf{R}]^{-1}, \quad (2.26)$$

with \mathbf{P}_e^b the error covariances of the bias estimates. \mathbf{P}_e^b is assumed to be proportional to \mathbf{P}_e^f , with a recommended proportionality constant between 0 and 2.

This bias correction method has been successfully applied to the EnKF by Kepenne *et al.* (2005). In this study bias correction improved the prediction of sea

level and SST in a 27 layer version of the Poseidon global ocean general circulation model.

2.5 Divergence

Continuous underestimation of forecast error leads to divergence (Hamill, 2002). When the forecast error covariance is underestimated, the Kalman gain weights the model forecast more strongly than the observation. As this continues the correction of the model trajectory is reduced and the forecast diverges from the truth. The prediction progressively worsens becoming of little use. Divergence can also be caused by insufficient ensembles to represent the error subspace of \mathbf{P} properly (Hamill, 2002; Houtekamer and Mitchell, 1998). Overestimation of \mathbf{R} can have a similar effect.

Theoretically, (Anderson, 2001) the ratio of the root mean squared error (RMSE) of the ensemble mean to the average RMSE of each ensemble member is given by

$$\frac{E_1}{E_2} = \sqrt{\frac{n_e + 1}{2n_e}}, \quad (2.27)$$

where E_1 is the RMSE of the ensemble mean

$$E_1 = \sqrt{\frac{1}{n} \sum_{i=1}^n \left(\frac{1}{n_e} \sum_{j=1}^{n_e} \mathbf{X}_i^j - \hat{\mathbf{X}}_i \right)^2}, \quad (2.28)$$

with $\hat{\mathbf{X}}_i$ the true system state. E_2 is the average RMSE of each ensemble member

$$E_2 = \frac{1}{n_e} \sum_{j=1}^{n_e} \sqrt{\frac{1}{n} \sum_{i=1}^n \left(\mathbf{X}_i^j - \hat{\mathbf{X}}_i \right)^2}. \quad (2.29)$$

Using this ratio, the onset of divergence in assimilation experiments has been detected (Whitaker and Hamill, 2002). In practice, the truth is unknown and so this ratio cannot be estimated in real time. A similar concept is that the root mean squared (RMS) spread estimates the RMSE (Keppenne, 2000), which has also been used to test for divergence (Houtekamer and Mitchell, 1998; Mitchell and

Houtekamer, 2000).

Other than improving the forecast error specification, divergence can be addressed by using an expansion factor (Anderson and Anderson, 1999), whereby the ensemble perturbations are multiplied by a factor slightly more than unity to spread them further. A dual EnKF proposed by Houtekamer and Mitchell (1998) reduced divergence by using each state ensemble to calculate the forecast error covariance for the other. Furthermore, covariance localisation can reduce divergence.

2.5.1 Localisation

During analysis, the states are updated relative to their covariance with the state associated, via \mathbf{H} , with the observation. It is expected that this covariance decreases with distance and at some distance the two states will be uncorrelated. In this case updating distant states is meaningless, and only wastes computing resources. Furthermore, because of the stochastic nature of the ensemble the ratio of noise to signal increases with distance. Therefore, at some distance a covariance may be a statistical artifact, rather than the truth. This is the basis for localisation (Hamill *et al.*, 2001; Ott *et al.*, 2004).

With localisation, analysis is only performed in a moving sub-domain. Beyond this sub-domain the correlations are taken as zero and all covariances ignored, and therefore no analysis occurs outside the moving sub-domain. The benefits of this are improved forecasts, as spurious correlations at distance are avoided, and faster filtering, as a smaller domain is analysed. A larger ensemble size reduces the need for localisation, as noise and signal can be more easily discriminated.

2.5.2 Ensemble Square Root Filters

A criticism of the EnKF is the use of perturbed observations: see equation (2.18). With small ensemble sizes, perturbed observations introduce sampling errors which affect the optimality of the analysis (Whitaker and Hamill, 2002). In response, a number of ensemble schemes have been proposed that use unperturbed observations. These include an Ensemble Adjustment Kalman Filter (Anderson, 2001), an Ensemble Transform Filter (Bishop *et al.*, 2001) and an Ensemble Square Root

Filter (Whitaker and Hamill, 2002). In fact, as Tippett *et al.* (2003) show, these methods and the EnKF itself are all ensemble forms of square root filters.

The EnKF analysis equation can be rewritten following Whitaker and Hamill (2002) as an ensemble mean equation,

$$\bar{\mathbf{x}}^a = \bar{\mathbf{x}}^f + \mathbf{K} [\mathbf{d} - \mathbf{H}\bar{\mathbf{x}}^f], \quad (2.30)$$

where the overbar represents an ensemble mean, and as an ensemble perturbation equation,

$$\mathbf{x}_i'^a = \mathbf{x}_i'^f + \mathbf{K}' [\varepsilon_i - \mathbf{H}\mathbf{x}_i'^f], \quad (2.31)$$

where the prime represents the perturbation from the mean for an individual ensemble member, \mathbf{x}_i . \mathbf{K}' is the gain used to update the perturbations and for the EnKF is equal to \mathbf{K} . If an analysis is performed with unperturbed observations ($\varepsilon_i = 0$), and with the same gain for both equations ($\mathbf{K}' = \mathbf{K}$), then the ensemble spread is reduced excessively. Tippett *et al.* (2003) show that this is equivalent to updating the perturbations via

$$\mathbf{x}'^a = [\mathbf{I} - \mathbf{K}\mathbf{H}] \mathbf{x}'^f, \quad (2.32)$$

in which case the ensemble spread is underestimated and \mathbf{P}_e^a is underestimated by the amount $\mathbf{K}\mathbf{R}\mathbf{K}^T$.

This difficulty is avoided by either i) finding a new value of \mathbf{K}' , the approach of Whitaker and Hamill (2002), or ii) solving equation (2.10) directly to ensure the correct spread is maintained, the approach of Bishop *et al.* (2001); Evensen (2004).

An ensemble square root filter that uses unperturbed observations can be derived following the second approach. The analysis of the ensemble mean is updated via equation (2.30), while the analysed ensemble perturbations are derived from equation (2.10). Substituting equation (2.16) into equation (2.10) gives

$$\Rightarrow \mathbf{X}'^a \mathbf{X}'^{aT} = \mathbf{X}'^f \mathbf{X}'^{fT} - \mathbf{X}'^f \mathbf{X}'^{fT} \mathbf{H}^T [\mathbf{H}\mathbf{X}'^f \mathbf{X}'^{fT} \mathbf{H}^T + (n_e - 1)\mathbf{R}]^{-1} \mathbf{H}\mathbf{X}'^f \mathbf{X}'^{fT},$$

which through rearrangement can be expressed as

$$\Rightarrow \mathbf{X}'^a = \mathbf{X}'^f \sqrt{\mathbf{I} - [\mathbf{H}\mathbf{X}'^f]^T [\mathbf{M}\Sigma^{-1}\mathbf{M}^T] \mathbf{H}\mathbf{X}'^f}, \quad (2.33)$$

where $\mathbf{M}\Sigma\mathbf{M}^T$ is the SVD of the $m \times m$ matrix $[\mathbf{H}\mathbf{X}'^f\mathbf{X}'^{fT}\mathbf{H}^T + (n_e - 1)\mathbf{R}]$. Since a square root of a matrix does not have a unique solution (Tippett *et al.*, 2003), equation (2.33) is also written as

$$\mathbf{X}'^a = \mathbf{X}'^f\mathbf{Z}\mathbf{W}, \quad (2.34)$$

with \mathbf{Z} an $n_e \times n_e$ matrix given by

$$\mathbf{Z} = \sqrt{\mathbf{I} - [\mathbf{H}\mathbf{X}'^f]^\text{T} [\mathbf{M}\Sigma^{-1}\mathbf{M}^T] \mathbf{H}\mathbf{X}'^f}, \quad (2.35)$$

and \mathbf{W} is any $n_e \times n_e$ orthogonal matrix. Evensen (2004) comments that the impact of \mathbf{W} is unknown. Throughout this thesis, \mathbf{W} will be taken as \mathbf{I} and subsequently ignored. This form of the equation gives rise to the name square root filter. Using equation (2.34) the ensemble perturbations can be analysed directly. As this formulation of ensemble square root filter, hereafter denoted EnSRF, does not require perturbed observations it is expected to perform better than the EnKF, especially where the ensemble size is small.

2.6 Applications of Data Assimilation

Data assimilation applications began in the atmospheric and oceanic fields. These remain the major areas of data assimilation research. Details on atmospheric and oceanographic assimilation are available from the review paper of Ghil and Malanotte-Rizzoli (1991).

In oceanic data assimilation, the focus of attention shifted in the 1990's (Malanotte-Rizzoli and Young, 1997). The original objective of oceanic data assimilation was to produce 4D realisations of the oceans states consistent with observations and model dynamics. This had the purposes of model improvement, to correct deficiencies in model parameterisation and boundary conditions, as well as producing realistic simulations of the oceans state. Now due to improvements in modelling and a greater emphasis on global weather/climate prediction, forecasting the future state of the oceans is becoming important.

As well as weather and climate change concerns, the main users of oceanic fore-

casts are shipping and naval concerns that are interested in currents and water movement. Most of the ocean assimilation systems developed so far that assimilate spatial satellite data, have concentrated on assimilation of altimetry (Ballabrera-Poy *et al.*, 2001; Evensen and van Leeuwen, 1996), the combined assimilation of altimetry and SST (Brusdal *et al.*, 2003; Testut *et al.*, 2003) and more recently combined altimetry, SST and salinity assimilation (Oke *et al.*, 2005).

However the direct observation of all variables is not crucial. In a recent multivariate assimilation paper by Keppenne and Rienecker (2003), the cross-correlations between temperature and other variables were used to update salinity and current states. While a relatively small ensemble size of 40 was used, the geostrophic balance between the variables contributed to the strong multivariate covariances that enabled the success of this study. In a coastal study where the geostrophic balance does not hold multivariate assimilation may not be as successful.

In oceanic assimilation no one technique seems dominant. Reduced rank filters are popular because of the large dimensionality of the oceanic domain. When assimilating water level anomalies and SST (Brusdal *et al.*, 2003) found that the three sequential filters tested (EnKF, ensemble Kalman Smoother, and SEEK — the Singular Evolutive Extended Kalman filter) gave a similar performance. For very large models the use of even a few ensembles are prohibitively costly and optimal interpolation, or the ensemble equivalent an ensemble optimal interpolation (EnOI) (Oke *et al.*, 2005) is used. EnOI is achieved by calculating a static forecast error covariance off-line from a series of ensemble runs.

The importance of using an advanced data assimilation technique, like ensemble filtering, rather than a simplified method is illustrated by the study of Kelley *et al.* (2002) that assimilated SST into a coastal ocean forecasting system for the USA east coast. The observations were assimilated horizontally using statistical interpolation and vertically using an interpolation method that assumed a good correlation between the surface temperature and the mixed layer. As might be expected the assimilation scheme had difficulty reproducing the ocean thermal structure in dynamically complex regions. The authors suggest the forecast would benefit from the assimilation of observed temperature profiles, but an improved data assimilation scheme would probably be of benefit also.

While water levels, currents and temperature are the focus for oceanic data

assimilation, biogeochemical assimilation has also been attempted. In a group of papers, Natvik and Evensen (2003a,b) describe an assimilation experiment where SeaWiFS ocean colour data was assimilated into an ecological model of the North Atlantic. The relative success of these studies was most likely due to its focus on deep ocean (Case I) waters that are optically dominated by phytoplankton, rather than on coastal regions, where coloured dissolved organic matter and suspended sediment confound the optical properties of the water. Furthermore, the relatively coarse resolution would have improved the accuracy and uncertainty estimate of the satellite data.

2.6.1 Coastal Marine Assimilation

While much data assimilation effort has taken place in oceanography, much less work has been done in coastal regions. Much of the coastal data assimilation has been application driven, which historically has been hydrodynamic modelling of sea levels. More recently attention has been given to ecological forecasting as well as suspended sediment forecasting, but little research has concentrated on temperature.

The importance of storm surge prediction for coastal protection and wave prediction which affect shipping (Verlaan and Heemink, 1997; Flather, 2000) has generally driven the research of data assimilation in this area. Initially assimilation used simplifications to the Kalman filter — usually in a reduced rank form (see Verlaan and Heemink, 1997; Dowd and Thompson, 1997) — and more recently extended and ensemble Kalman filters (Madsen and Cañizares, 1999) have been applied. Water level assimilation has also been demonstrated using local models that use chaos theory to perform the analysis (Babovic and Fuhrman, 2002).

The success of sea level–current assimilation is due to the high degree of spatial correlation within hydrodynamic models that enables the state reduction employed in the reduced rank filters: the dynamics of the system can often be expressed by a few key EOFs. Another advantage in forecasting sea level is that the observations are obtained from relatively few tidal gauge stations; a corollary of the simple dynamics. The small number of observations greatly simplifies the matrix inversion needed to perform the analysis, equation (2.20) or (2.11), saving computational time.

Generally this modelling assimilates point scale rather than spatial data as the

resolution of altimeters is generally too coarse for coastal applications, and the system is small enough for the use of buoys and tidal gauges (Flather, 2000). The data assimilation of water level into hydrodynamic models is now well established, and evidenced by Bertino *et al.* (2002) who used an EnKF to assimilate water levels into an estuary model. The assimilation of surface currents derived from coastal radar into coastal ocean models (Shulman *et al.*, 2002; Oke *et al.*, 2002; Hallock *et al.*, 2003) is a new direction of research for hydrodynamic forecasting.

In terms of ecological assimilation, state estimation is in its early stages (Eknes and Evensen, 2002) with only a handful of papers in the literature. Most of the papers discussing ecological assimilation are in the context of parameter estimation, where uncertain parameters are estimated by assimilating observed data generally using a variational assimilation approach. These papers seek to improve the underlying numerical model (see for example Fennel *et al.*, 2001; Navon, 1997; Solidoro *et al.*, 2003; Garcia-Gorriz *et al.*, 2003).

Papers investigating parameter estimation suggest a reason for the lack of state estimation papers: accurate ecological modelling is very difficult. Solidoro *et al.* (2003) were able to accurately calibrate 5 out of 42 model parameters owing to a high degree of parameter correlation. Interestingly, states were highly sensitive to parameters related to temperature, suggesting that improvements to predicting temperature would aid in ecological forecasting. Both Fennel *et al.* (2001) and Faugeras *et al.* (2003) report poor or unsatisfactory results in their parameter estimation studies. Usually these studies rely on data observed at a single location, although Garcia-Gorriz *et al.* (2003) sought parameter estimation of a coupled physical-biological model in the Adriatic Sea using SeaWiFS satellite data. A recent review of aquatic biogeochemical modelling by Arhonditsis and Brett (2004) supports the conclusion that there is still much work to be done, although the paper notes that recent advances in data assimilation technology are promising for future work.

The state estimation papers for coastal ecological assimilation generally focus on 1-dimensional models (Eknes and Evensen, 2002; Allen *et al.*, 2002). Allen *et al.* (2002) reported a predictability window of 2 days for their ecological forecasts, while Eknes and Evensen (2002) indicated that the assimilation could realistically estimate the error associated with the forecast state. Both papers highlight the future potential of ecological assimilation but recognise that further research is required.

Forecasting of suspended sediment is also at a fairly early stage. Efforts towards the operational forecasting of suspended sediment were made under the guise of the PROMISE project (Gerritsen *et al.*, 2000; Prandle *et al.*, 2000; Vos *et al.*, 2000; Baumert *et al.*, 2000). Although operational forecasting is still limited by knowledge of the physical processes and the availability and accuracy of suspended sediment observations. Suspended sediments are most prevalent during turbulent stormy conditions when clouds inhibit satellite observation and in coastal regions where retrieval is complicated (Ruddick *et al.*, 2000).

While assimilation of temperature into oceanographic models is routinely undertaken, surprisingly little assimilation of temperature into coastal models has been published. Annan and Hargreaves (1999) improved the predictive ability of a baroclinic model of the North Sea by assimilating SST with a highly simplified Kalman filter scheme. The use of a simplified Kalman filter was justified in that superior methods were not (in 1999) currently available. The simplifications were i) to ignore horizontal correlations and thus assimilate a series of one-dimensional vertical models, ii) to make temperature adjustments independent of turbulent kinetic energy, and iii) to assume that the waters above and below thermocline were well mixed, which leads to an assumption of uniform heating error in the upper water column.

While these simplifications were appropriate for a relatively coarse (20 km resolution) shelf-sea model, they will not be appropriate for a model with a smaller resolution, where horizontal transport becomes more important. Furthermore, as computational power has increased there is no longer a need to restrict to simplified methods and improving the forecasting of temperature with more advanced methods appears possible.

This review of the applications of data assimilation to biogeochemical modelling in the coastal marine environment shows that forecasting sea level is well established. Attempts at ecological and suspended sediment forecasting are being made, although these are at an early stage and at the limits of current technology and data availability. Assimilating SST into the coastal marine models has not been undertaken recently and work is needed to provide an update of the simplified method currently in the literature.

2.7 Review Findings

The literature review has found that data assimilation is well advanced in the oceanographic field, but less work has been performed in coastal regions. Most of the coastal data assimilation focus has been on the hydrodynamics, especially for coastal protection against storm surges. Preliminary work on developing suspended sediment forecasting systems has been undertaken, but many modelling issues need attention and there are difficulties with the satellite observation. Work is developing on forecasting ecological systems, but this is hindered by the complexity of the systems. Many researchers using data assimilation for parameter estimation report poor results.

Surprisingly little work has been done on temperature forecasting in the coastal environment. Only one paper was found which concentrated on its assimilation, and this used a very simplified version of the Kalman filter. There is clear scope for improving the knowledge of temperature prediction itself. Furthermore, temperature assimilation appears as a bridge between hydrodynamic assimilation which is well established and higher process assimilation, such as ecological and suspended sediment, that currently may be considered infeasible.

Coastal data assimilation exercises are more complex than oceanic studies because the forecast error covariances are likely to be anisotropic and the variables may not be in geostrophic balance. These conditions limit the sequential data assimilation techniques available for conducting the research. The Kalman filter was discounted for computational expense and because it is only applicable for linear models. Simplifications to the Kalman filter were discarded also as they apply varying complexities of steady state filters, whereas the variable conditions expected necessitate a dynamic filter. Ensemble based sequential filters appear suitable for this research, being dynamic, computationally efficient, and applicable to nonlinear systems.

The exploration of the ensemble data assimilation literature found that while the use of ensembles is well established, there is a bewildering variety of possibilities in regard to the configuration of an assimilation system. For example, there is a choice of a deterministic or stochastic filter, the initialisation of the ensemble, and how forecast error should be incorporated. The review also found that the incorpo-

ration of forecast error through perturbed forcing data has not been approached in a rigorous manner and there is scope for improvement.

In light of these findings this thesis will proceed by exploring the forecasting of temperature in a coastal case study, assimilating SST observations. Ensemble data assimilation techniques will be used and the incorporation of forecast error through perturbed forcing data will be further developed. As there is no clear guidance on the optimal ensemble data assimilation configuration, different configurations will be explored in a synthetic setting before proceeding to the actual assimilation.

2.8 Chapter Summary

This chapter presents a review of data assimilation techniques and their application to biogeochemical modelling of coastal regions. Ensemble data assimilation techniques appear most appropriate for data assimilation in a coastal marine setting because they adapt to dynamic conditions, are flexible in their computational cost depending on resources available, and allow for nonlinear systems. Within the field of ensemble data assimilation there is a wide range of choice, in terms of forecast error, ensemble generation, ensemble size, and filter type. In particular, the inclusion of forecasting error through forcing data has scope for further attention.

This review sets the theoretical base for the PPB SST assimilation case study which comprises the remainder of the thesis, and shows there is scope for improving the forecasting of temperature in coastal regions. Furthermore, the review shows that providing some guidance on the application of data assimilation in a coastal setting is of benefit.

Chapter 3

Data: Port Phillip Bay

The focus of the coastal marine data assimilation of this thesis is Port Phillip Bay (PPB), located in south east Australia. PPB was chosen because it is one of the most intensely monitored water bodies in this area. As the gateway to the city of Melbourne, home to more than three million people, its continued health is of importance, and for this reason government agencies conduct ongoing monitoring of PPB.

This chapter introduces PPB, describing the current state of knowledge in relation to its physical aspects. Then the PPB data collected for this thesis is described. The atmospheric, riverine, tidal, and water quality data was provided by various agencies and is summarised by its spatial location, length and frequency of record. A preliminary analysis of these data is presented, predominately to verify the quality of the data, but also to ascertain the uncertainty associated with the data. As the majority of the data was collected as point scale data, a determination of the spatial uncertainty of the data is important before the data are used for modelling. The satellite observations used for assimilation are also reviewed for data quality. By contrasting them with point measurements, the accuracy and bias is assessed. A description of the satellite processing to remove suspect pixels is detailed.

3.1 Port Phillip Bay

Port Phillip Bay is a large, almost completely enclosed, body of water located in south eastern Australia, south of the City of Melbourne (see figure 3.1). PPB is shallow with a maximum depth of about 23 m and with at least half of its area less than 8 m deep. Its surface area is 1,930 km² and it has a volume of about 26 km³. PPB is connected to the open sea at “The Heads”, a narrow entrance to the south. This relatively small opening leads to locally large tidal currents, but also limits exchange between PPB and Bass Strait.

PPB can be conceptualised as two parts; a northern basin, and a southern section known as the “Great Sands”. The northern basin is deeper in the centre, flat, and

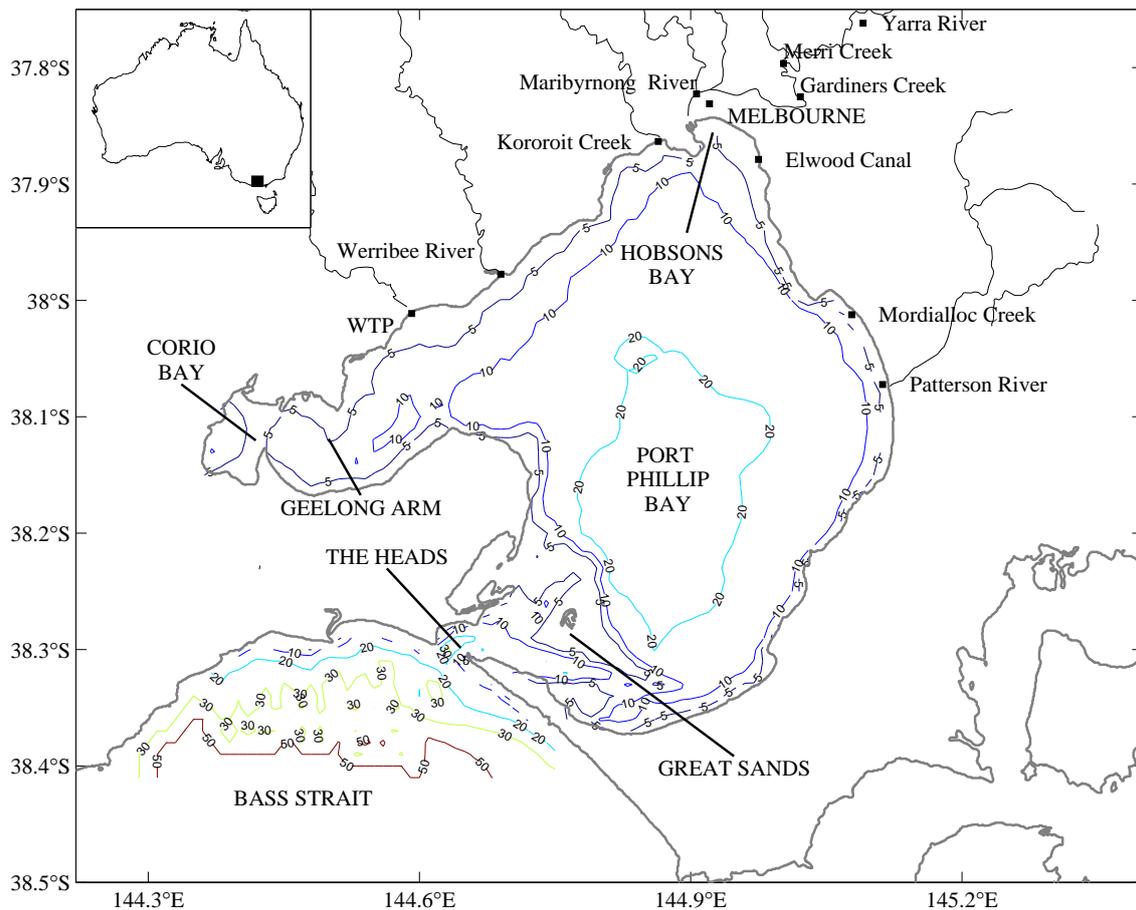


Figure 3.1: Location of PPB in south eastern Australia. Bathymetry contours indicated are in metres.

gradually shallows towards the edges. The Great Sands is relatively shallow, with a few deeper channels connecting the northern basin to Bass Strait. The shallow nature of the Great Sands, together with the narrow entrance at The Heads, acts as a resistance to flow and denotes the limit of significant tidal exchange.

A major source of scientific knowledge of PPB is the Port Phillip Bay Environmental Study (PPBES) conducted from 1992 to 1996 (Harris *et al.*, 1996). This study investigated the environmental condition of PPB and focused on physical aspects, nutrients, toxicants, and ecology. The findings of the study are briefly summarised below.

Physical Aspects In general, the wind characteristics are complex with few trends, but there is a tendency for southerly winds in summer and northerly winds in winter. Strong winds associated with storms are more frequent in the summer. The annual freshwater input to PPB are summarised in table 3.1. Direct rainfall is an important source of fresh water contributing about half the annual total. The other half comes from river flows with most of this from the Yarra River to the north-east. Rainfall varies across PPB, with more towards the south east and less towards the north west. This spatial variation in precipitation influences the riverine flows and so the predominant source of riverine inflow is from the eastern catchments. Annual evaporation is slightly less than the total freshwater inflow.

Water movement in PPB is driven by tides, wind and density differences. The strong tidal signal at the entrance to PPB is attenuated by the Great Sands. Maximum tidal velocity reduces from 1 ms^{-1} at the Heads to 0.05 ms^{-1} in the centre of PPB and to 0.02 ms^{-1} in the north. Wind generated currents are of the order of 0.05 ms^{-1} , and it is believed these currents follow the wind direction in the shallow

Table 3.1: Annual fresh water exchange within PPB. Source: PPBES (Harris *et al.*, 1996).

Source	Volume [km^3]
Rivers and Western Treatment Plant	1.6
Rainfall	1.3
Evaporation	-2.3

edges, and return as a counter current in the deeper centre. Fresh water inputs, principally from the Yarra River, generate density currents. In Hobsons Bay during rain events, such currents can reach 0.5 ms^{-1} . This reduces as the freshwater plume travels south, with mixing reducing the salinity difference. This plume generally hugs the eastern coastline under the Coriolis effect. It is thought that some density driven flow exists between PPB and Bass Strait, induced by a thermal difference, but its magnitude is unknown.

As PPB is shallow, it is generally well mixed vertically. Thermal stratification occurs in summer due to surface heating, however this is usually less than 3°C . Occasionally, though, it is enough to inhibit the transfer of oxygen to deeper levels. Horizontal thermal and salinity gradients are common, generated by different heating and cooling rates with water depth. With less mass the temperature and salinity of shallower waters vary more readily than those of deeper waters. The persistence of the horizontal gradients suggests that mixing is slow. The residence time of PPB is estimated to be of the order of 12 to 16 months: long because of restricted mixing over the Great Sands.

Temperature variation in PPB is due to solar radiation input, long wave radiation, sensible heat fluxes, evaporation, inputs of water from precipitation and riverine sources, and water exchange with Bass Strait. Depending on location, the temperature varies from 11°C to 21°C over the year. Salinity varies with fresh water inputs from precipitation and riverine sources, loss of water through evaporation, and exchange with Bass Strait. In Corio Bay and the Geelong Arm, evaporation may exceed fresh water inputs leading to locally high salinity levels.

Nutrients Nutrient inputs to PPB are dominated by inputs from the Yarra River and the Western Treatment Plant (WTP). The nutrient loads from the WTP vary seasonally with the largest loads in late winter. The Yarra flow varies considerably, which influences the nutrient loads. Nitrogen appears to accumulate in the catchment and be flushed out during storms, whereas phosphorus has a more constant input rate. Atmospheric sources of nutrients are also significant. Inorganic nitrogen is the limiting plant nutrient in PPB, and is predominantly removed through the conversion of soluble nitrate into insoluble nitrogen gas by denitrifying bacteria. This gas is then lost to the atmosphere. These biological processes, together with

relatively low levels of nutrient input, keep PPB in a low trophic state.

Toxicants Due to its proximity to a major city, PPB receives toxicants from agricultural, urban, and industrial sources. More than half the metal toxicants are derived from the Yarra River, as are many of the organic toxicants. The concentrations of toxicants are generally well within accepted water quality guidelines. Although, locally high concentrations of toxicants are found near industrial sources, especially in Corio Bay and Hobsons Bay.

Ecology In general, chlorophyll concentrations are low compared with similar bays world-wide. With low nutrient levels, PPB is classed as oligotrophic to mesotrophic. Its water quality is good by world standards, but high levels of algal growth are occasionally associated with proximity to nutrient sources. Still, the abundance of macroalgae and seagrass appears to be driven by physical factors rather than nutrient levels. Seagrass is usually found in shallower waters: less than five metres deep.

3.2 Data

The data needs of this thesis are meteorological data, riverine data, tidal data, bathymetry, water quality data, and satellite observations. Most of these data are collected by various government agencies and have been generously made available. These data are briefly summarised. Data were sought for the years 2000–2003, for assimilation studies and the years 1993–1995 for comparison with the PPBES findings.

3.2.1 Meteorological Data

Two organisations collect meteorological data in the vicinity of PPB. The Environment Protection Authority (EPA) collects wind and temperature data, while the Bureau of Meteorology collects most atmospheric variables at automatic weather stations. The locations of the stations nearest PPB are shown in figure 3.2. As the map indicates, most of the stations are centred to the north east of PPB in the

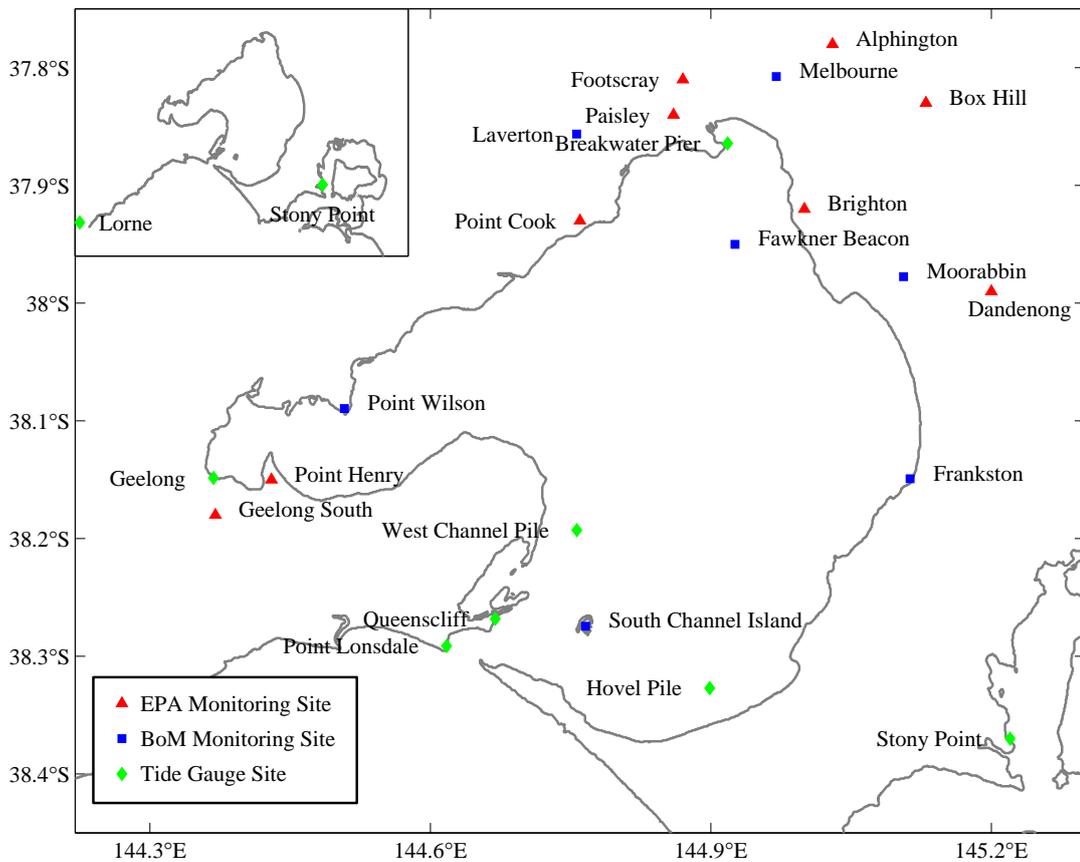


Figure 3.2: Location of weather and tide stations in the vicinity of PPB.

suburbs of Melbourne. There are, however, enough stations around PPB to give a reasonable spatial coverage.

The data obtained from the EPA is recorded hourly and covers the five year period from January 1998 through to December 2002. The data obtained from the Bureau of Meteorology covers the period from January 2000 through to November 2003, with the values reported at three hourly intervals. The Bureau of Meteorology collects all meteorological variables of interest — cloud cover, evaporation, precipitation, relative humidity, pressure, solar insolation, air temperature and wind — but not every station collects all variables. Table 3.2 summarises the variables collected at the Bureau of Meteorology stations.

While figure 3.2 shows that the Bureau of Meteorology monitoring stations are

Table 3.2: Summary of availability of meteorological data for years 2000–2003 collected at Bureau of Meteorology automatic weather stations. Data are available (✓) for cloud cover (CC), evaporation (E), precipitation (P), relative humidity (RH), atmospheric pressure (AP), solar insolation or short wave radiation (SWR), air temperature (AT), and wind vectors(W).

Station	CC	E	P	RH	AP	SWR	AT	W
Melbourne	✓	✓	✓	✓	✓	✓	✓	✓
Frankston	-	-	-	✓	-	✓	✓	✓
Laverton	-	-	✓	✓	✓	✓	✓	✓
Moorabbin	-	-	✓	✓	✓	✓	✓	✓
Point Wilson	-	-	-	✓	-	✓	✓	✓
Sth. Channel Is.	-	-	✓	-	-	✓	✓	✓
Fawkner Beacon	-	-	-	-	-	✓	-	✓

evenly distributed around PPB, table 3.2 shows that the majority of the variables are collected in the north. Only the station at Melbourne collects all variables; it is the only station to monitor cloud cover and evaporation. Atmospheric pressure is only monitored at three stations — Melbourne, Laverton and Moorabbin — all located to the north east of PPB. In the south, only precipitation, air temperature and wind speed are monitored at South Channel Island. Also, Fawkner Beacon, the only site other than South Channel Island actually in PPB, only collects wind data.

The incoming solar radiation data supplied is not actually measured at the stations, but instead is derived from the Japanese GMS-5 weather satellite. When this satellite data was tested in 1997, the findings were that on a typical clear day the average agreement between a pyrometer and the data was within 0.17%, and most values were within 6%. The satellite data overestimated solar radiation during wet, cloudy conditions, and underestimated it during dry conditions. Furthermore, the data are supplied with a warning that the performance of the satellite has degraded since June 2001, with all subsequent data considered less accurate relative to the data collected previously.

3.2.2 Riverine Inputs

The principal sources of fresh water to PPB are from the catchments to the north. The Yarra River dominates, although Kororoit Creek and the Werribee River to

Table 3.3: Summary of riverine data collected for waters entering PPB.

Station	Type	Start date	End date	Regularity	Source
Yarra River	Flow	Jan. 1978	Jun. 2003	daily	Melbourne Water
Yarra River	Water Quality	Jan. 1991	Oct. 1994	monthly	Data Warehouse
Yarra River	Water Quality	Jan. 1994	Jun. 2002	weekly - monthly	Melbourne Water
Maribyrnong River at Keilor	Flow	Dec. 1978	Apr. 2003	~ daily	Data Warehouse
Werribee River at Droomers	Flow	May 1979	Jul. 2001	~ daily	Data Warehouse
Werribee River at Werribee	Flow	Apr. 1982	Apr. 2003	~ daily	Data Warehouse

the west also contribute flow. Another source of fresh water is the WTP. The locations of the principal rivers and the outlets of the WTP are shown in figure 3.1. While data for the major rivers is generally available, the smaller rivers are ungauged. Yarra flow is composed of flow from the Yarra River, Merri Creek, Gardiners Creek, and the Maribyrnong River. Melbourne Water supplied daily flow data for the Yarra River, Merri Creek and Gardiners Creek for the period January 1978 through June 2003. Melbourne Water also recorded water quality data in these three rivers between 1994 and June 2002, which was downloaded from their website <http://www.melbournewater.com.au>. These data ranged in frequency from weekly to monthly. Melbourne Water also operates the WTP and supplied daily flow data for the period from January 2000 to December 2003.

Flow data were also downloaded from The Victorian Water Resources Data Warehouse, <http://www.vicwaterdata.net>, for the Maribyrnong River at Keilor and the Werribee River at both Droomers and Werribee. The Victorian Water Resources Data Warehouse also contained water quality data for the Yarra, which complemented the data obtained from Melbourne Water. The riverine data are summarised in table 3.3.

3.2.3 Tidal Data

Tidal data are collected inside PPB by the Port of Melbourne Corporation (formerly Victorian Channels Authority), and outside PPB by the National Tidal Authority. Both agencies report tidal data at six minute resolution. The Port of Melbourne Corporation supplied data recorded at Breakwater Pier, Geelong, Hovel Pile, Point Lonsdale, Queenscliff, and West Channel Pile for the years 2000–2002. The National Tidal Authority supplied data recorded at Lorne and Stony Point for the period January 2000–May 2003. The locations of the tidal gauges around PPB are shown in figure 3.2.

3.2.4 Bathymetry

High resolution bathymetry data were supplied by the Port of Melbourne Corporation. These data were used to generate the bathymetry contours of figure 3.1. The bathymetry was recorded at 25-metre resolution around The Heads, at 100-metre resolution over the Great Sands, and at 500-metre resolution at other locations.

3.2.5 Water Quality Data

Continuous in situ monitoring of PPB over a long period is quite rare, although a number of periods of intense monitoring exist. The period from August 2002 to June 2003 contains the highest concentration of continuously recorded data. The recorded data supplied by the Marine and Fisheries Research Institute (MAFRI) and the EPA are summarised in table 3.4. The locations of the monitoring sites are shown in figure 3.3.

MAFRI have a number of sites with long records of temperature and salinity. Data from four sites were supplied. At one site, West Channel Pile, the record covers ten years and includes both temperature and salinity data. The remaining three sites cover eight months with only temperature supplied.

The EPA also regularly monitors six sites within PPB, measuring water quality states such as concentrations of metals and nutrients, as well as physical states like temperature, dissolved oxygen and salinity. Unfortunately, the EPA data was recorded on average only once every month and a half.

Table 3.4: Summary of water monitoring data from stations in PPB recorded by the EPA and MAFRI.

Station	Start date	End date	Source	Regularity
West Channel Pile	August 1990	September 2001	MAFRI	hourly
Centre	August 2002 July 1984	June 2003 February 2003	MAFRI EPA	hourly ~45 days
Hobsons Bay	August 2002 July 1984	June 2003 February 2003	MAFRI EPA	hourly ~45 days
Longreef	August 2002 April 1990	June 2003 November 2001	MAFRI EPA	hourly ~45 days
Paterson River	April 1990	August 1996	EPA	~45 days
Dromana	April 1990	August 1996	EPA	~45 days
Corio Bay	July 1984	February 2003	EPA	~45 days

During the PPBES, twenty-five bay wide transects of underway data were collected approximately monthly. These data were made available by CSIRO for analysis. Each transect was collected over a 2–3 day period. An example of a transect path is shown on figure 3.3.

3.2.6 Satellite Data

CSIRO Marine supplied Advanced Very High Resolution Radiometer (AVHRR) data from the National Oceanic and Atmospheric Administration (NOAA) satellites for the years 2000 to 2003. Three types of errors were denoted by CSIRO Marine Research: erroneous or null valued data, the presence of land and the presence of clouds. The data were digitally encoded such that the minimum temperature resolution was 0.14°C.

3.3 Data Analysis

Before data can be used for modelling or assimilation, an assessment of its quality is required. Erroneous data must be noted and flagged. While most of the data are of good quality, it is necessary to determine which sites are most representative of the marine environment. For instance, as water has a large thermal mass its

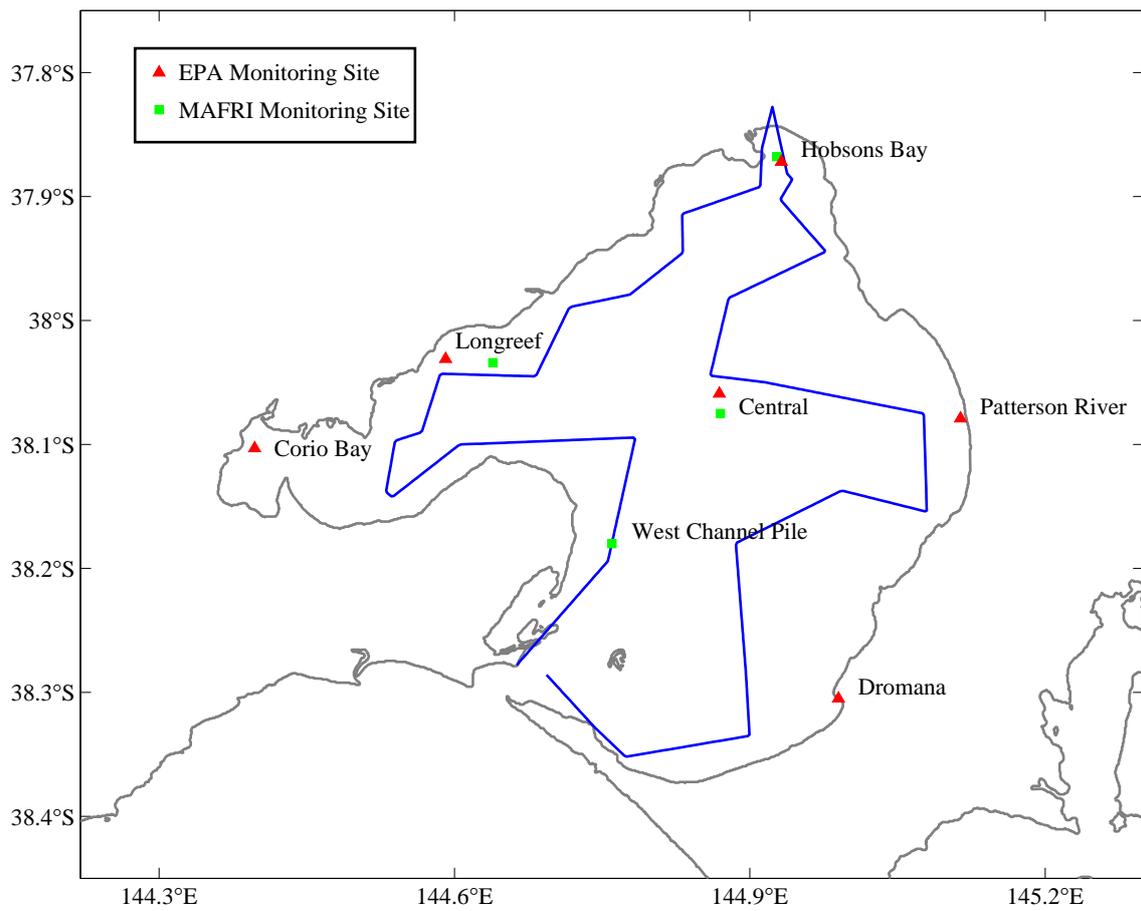


Figure 3.3: Location of water monitoring sites within PPB operated by the EPA and MAFRI. A typical transect path is shown as well.

temperature does not vary as readily as the land. This quality affects the local climate with locations close to large water bodies experiencing less extremes of temperature than locations further away. For this reason data collected at one site may be more suitable for modelling than data collected at another site.

As the literature review chapter highlights, an understanding of model and observation uncertainty is necessary for data assimilation, and thus an important component of this thesis. Therefore a thorough examination of the spatial uncertainty, or variation, associated with the data sets is made.

The spatial variation of most of the data sets is summarised by the scatterplots in figure 3.4, which plot the spatial mean value against the standard error of the spatial mean. The standard error is given by

$$\gamma = \frac{\sigma}{\sqrt{n_s}}, \quad (3.1)$$

where σ is the standard deviation and n_s is the number of samples used. Plots of this form indicate how the spatial variation of a data type changes. Not all data types can be summarised well in this form and the particular aspects of each data type are discussed individually in sections 3.3.1 through 3.3.9.

Pressure, figure 3.4a, is based on 3 stations and varies little spatially. The variation is generally constant with mean perturbation, although the spatial variability of pressure seems to be less at higher mean pressure perturbation values than at lower mean pressure perturbation values. The spatial variability of wind speed, figure 3.4b, is based on 8 stations. The panel indicates a slight tendency for increased variation with wind speed. Evaporation is only measured at two stations and so the difference is used to indicate spatial variation, rather than the standard error. Figure 3.4c shows a trend of increasing variation with magnitude. Air temperature spatial variation, shown in figure 3.4d, is based on 4 stations. Within the main range of air temperature values, 8–20°C, the spatial variation is constant. At higher mean air temperatures the spatial variation increases. The spatial variation of precipitation is based on 4 stations and is shown in figure 3.4e. This panel shows a tendency for increasing spatial variation of precipitation with magnitude. The spatial variation of river temperature, given in figure 3.4f, is based on 3 sites and is constant with the temperature magnitude. Finally, the spatial variation of daily average solar

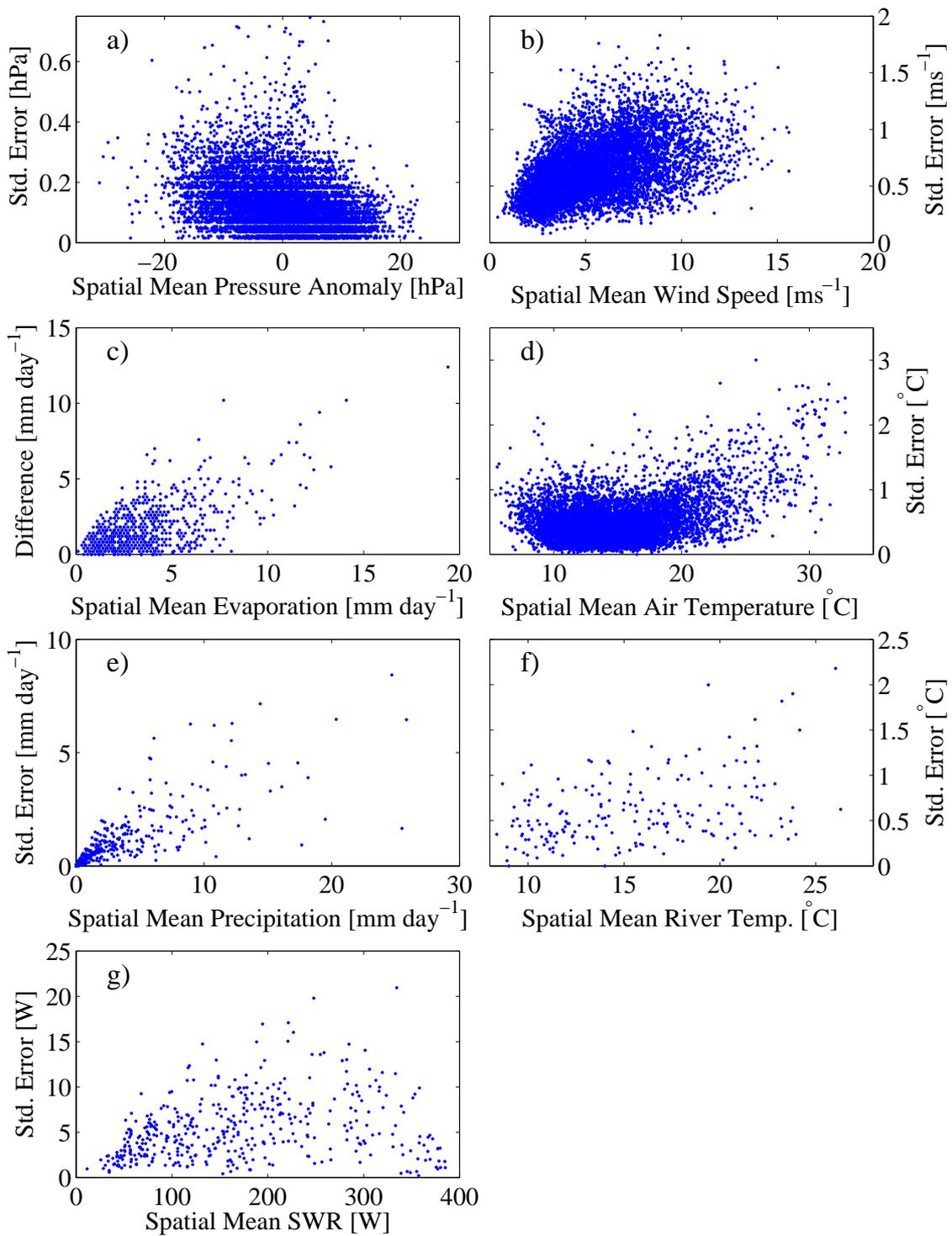


Figure 3.4: Scatter plots showing spatial variation of a) pressure, b) wind speed, c) evaporation, d) air temperature, e) precipitation, f) river temperature, g) daily solar radiation variation.

Table 3.5: Correlation coefficients between 3-hourly recorded data at Melbourne. Boldface values indicate coefficients with an absolute magnitude greater than 0.5.

	Cloud Cover	Relative Humidity	Air Temperature	Wind Speed
Atmospheric Pressure	-0.18	0.15	-0.30	-0.28
Cloud Cover	-	0.12	-0.04	0.11
Relative Humidity	-	-	-0.64	-0.43
Air Temperature	-	-	-	0.17

Table 3.6: Correlation coefficients between daily recorded data at Melbourne. Boldface values indicate coefficients with an absolute magnitude greater than 0.5.

	Evaporation	Solar Radiation
Precipitation	-0.14	-0.20
Evaporation	-	0.60

radiation variation is shown in figure 3.4g. Based on 7 stations, it shows that the maximum spatial variation occurs in the mid range of the mean value.

In general, the data does not vary considerably spatially, with the possible exception of relative humidity. This finding is important as it allows use of spatially uniform data, rather than spatially varying data, in the modelling. Furthermore, for some data types the spatial variation tends to vary with the spatial mean, while for other types the spatial variation is uniform. The division appears to be that those data types with a lower bound — e.g. evaporation and precipitation — have a spatial variation which depends on the mean value. This finding is the basis for the forecast error generation technique developed in the following chapter.

The relationship between the meteorological data collected is of interest. Any generation of ensembles of forcing data should take correlations between the data into account. Tables 3.5 and 3.6 list the correlation coefficients between the 3-hourly and daily data observed at Melbourne respectively. The tables show that generally there is little correlation amongst the data types. The exceptions are a negative correlation between relative humidity and air temperature and a positive correlation between evaporation and solar radiation. Whilst these correlations make sense physically, neither correlation is very strong. In general, it is safe to assume that the meteorological data are independent.

3.3.1 Atmospheric Pressure

Atmospheric pressure is collected at three stations around PPB: Laverton, Moorabbin and Melbourne. Physically, temporal variations in pressure are related to weather fronts moving eastwards across PPB. But in the vicinity of PPB pressure variations between stations should not be dramatic as the distance between the stations is small relative to the size of weather fronts.

An initial analysis removed the spatial mean from the data and considered the time series of residuals. A slight bias is evident at each station, which is understood to be due to the relative elevation of the stations. The Melbourne data appears more reliable with fewer recording errors. Table 3.7 indicates the average pressure at each site together with the standard deviation of the temporal variation for each station.

A second analysis considers the anomalies at each station about its temporal mean. Table 3.8 indicates these anomalies are highly correlated, which confirms the supposition that pressure does not vary significantly spatially. The scatterplot of spatial pressure variation with spatial mean pressure perturbation (figure 3.4a) indicates that the spatial variation is fairly constant, although it appears to slightly increase for negative perturbations. This is likely due to changes in pressure caused by the movement of cold fronts and unstable weather systems. An auto-correlation of the Laverton data indicates that the correlation in the pressure reduces significantly

Table 3.7: Station elevation, Average pressure and standard deviation of pressure at various stations around PPB.

Station	Elevation [m]	Average Pressure [hPa]	Standard Deviation Of Pressure [hPa]
Moorabbin	13	1015.4	0.19
Melbourne	32	1012.9	0.24
Laverton	20	1014.7	0.23

Table 3.8: Correlation coefficients indicating the relationships between the pressure anomalies at the various stations.

	Melbourne	Moorabbin
Laverton	0.99	0.99
Melbourne	-	0.99

after about 3 days. This is associated with the time it takes a weather front to pass.

3.3.2 Wind

Sixteen atmospheric stations record wind speed and direction. An exceedence plot of wind speed for each station is given in figure 3.5, indicating the proportion of time a given value is exceeded over its the entire record. For instance, the plot indicates that 50% of the time the wind speed at South Channel Island is greater than 7.5ms^{-1} . The stations appear to form three clusters of high, medium and lower wind speeds. The spatial location of the cluster members (figure 3.2) shows that the clustering is based on proximity to PPB. The three stations in the high wind speed cluster — Fawkner Beacon, Point Wilson and South Channel Island — are located either within PPB or at highly exposed locations on the edge of PPB. The five medium wind speed stations — Frankston, Laverton, Moorabbin, Point Cook, and Point Henry — are located in the middle of the bay. The remaining six stations — Geelong South, Paisley, Dandenong, Footscray, Brighton, Box Hill, Alphington, and Melbourne — are located in the inner bay or near the coast.

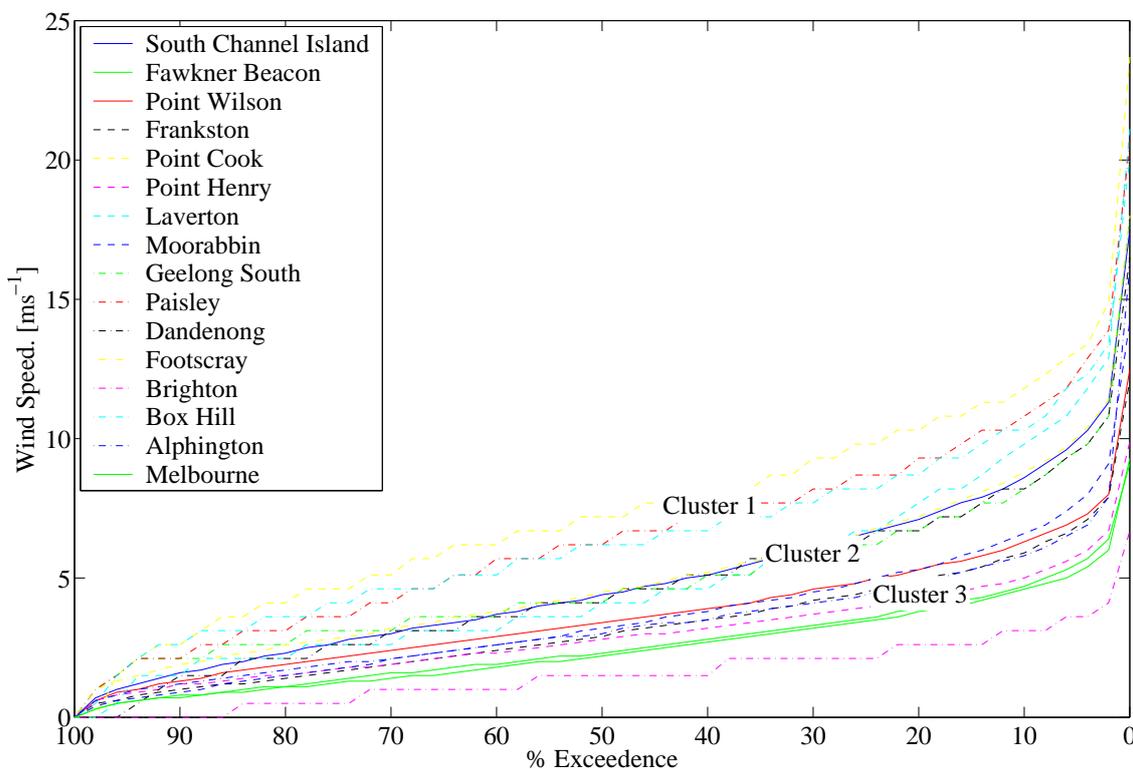


Figure 3.5: Exceedence plot of wind speed at various weather stations around PPB.

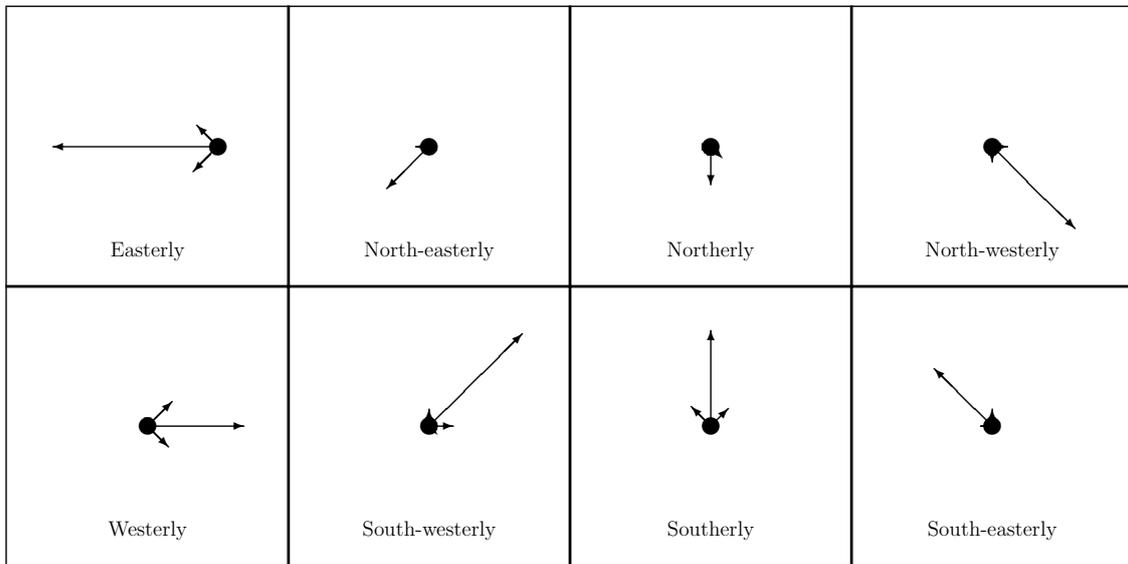


Figure 3.6: Spatial variation of wind direction about the mean wind direction calculated for each of the eight cardinal directions. The length of the arrow indicates the frequency of wind in the given direction.

and Point Henry — are located either on the edge of PPB or at air fields and so are relatively exposed to high winds. The lower wind speeds recorded at inland stations are due to the surface roughness over the land slowing the wind. This is most noticeable at Melbourne, which has a very low recorded wind speed. Logically, the data recorded by the high and, to a lesser extent, medium wind speed clusters are more representative of bay conditions.

To deduce the variation in wind speed for the eight high and medium wind speed stations, the spatial mean and standard deviation of wind speed are calculated at a given time. These data are shown as a scatterplot in figure 3.4b, indicating that while there is a wide variation, a slight trend of increasing variance with increasing wind speed is present. This is possibly an indication of gusting during stormy conditions.

The principal wind directions and their variation are also examined for the eight medium and high wind speed stations. For each record in the time series the wind direction at each station is assigned to one of the eight cardinal directions: north, north-east etc. The direction with the largest number of records is denoted the principal direction. For each principal direction a frequency distribution is made of the wind direction at the different stations given a principal direction. The

result is a matrix of the frequency of the wind direction given a principal direction that indicates the spatial variation associated with the wind blowing generally in a particular direction. Where a time level has two directions with equally high frequencies, both are declared principal directions and the frequency counts divided between them. The findings are plotted as wind roses in figure 3.6, where the length of the arrow indicates the relative frequency of that direction, and suggests that there are no clear trends. Furthermore, while there is evidence of some spatial variation, in general wind direction is spatially uniform.

3.3.3 Evaporation

For the period 2000 to 2003, daily evaporation data are only available at Melbourne. Evaporation has a strong seasonal signal with higher values recorded during summer and lower values recorded during winter, refer to figure 3.7a. Its values range from about 1 mm day^{-1} to about 7 mm day^{-1} , although peaks of up to 12 mm day^{-1} are observed.

An assessment of the variability in evaporation is possible as both Laverton and Melbourne recorded evaporation data between 1993 and 1995. Figure 3.7 plots

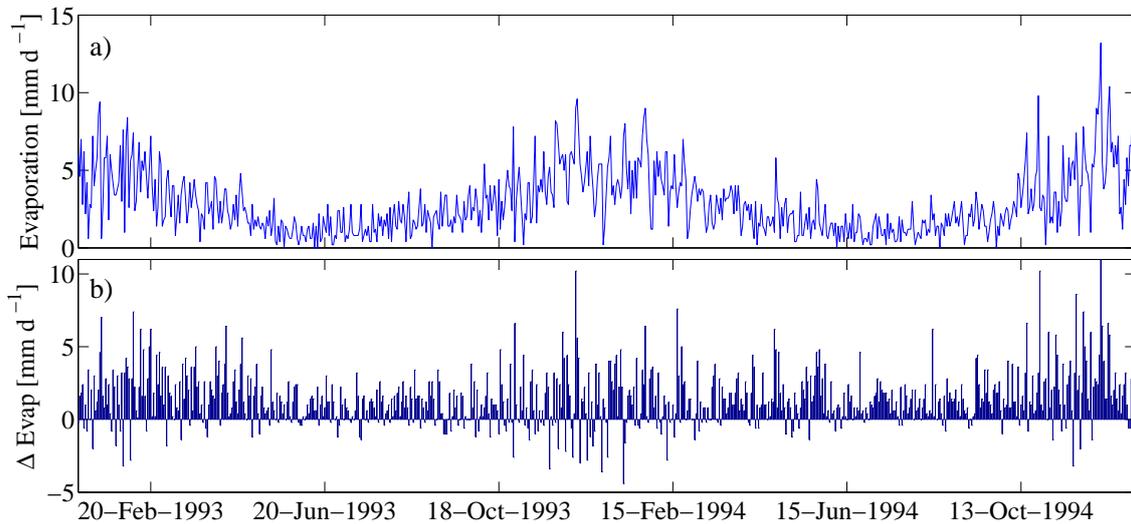


Figure 3.7: a) Daily evaporation at Melbourne, and b) difference in daily evaporation between Laverton and Melbourne over the period 1993 to 1995.

evaporation recorded at Laverton minus evaporation recorded at Melbourne, and shows that the difference is biased towards higher evaporation at Laverton with a mean difference of 1.41 mm day^{-1} . The standard deviation of the difference is 1.90 mm day^{-1} ; the bias is less than the variation of the data. The variation also appears greater in summer, which is confirmed by a scatterplot, figure 3.4c. This records the spatial mean evaporation for the period 1993–1995 against the spatial standard deviation. Variation in evaporation tends to increase proportionally to the data value.

3.3.4 Air Temperature

Like evaporation, air temperature has a strong seasonal pattern, but as figure 3.8 displays, it is recorded at a high enough resolution to indicate a strong diurnal signal. The proximity to the sea creates a maritime climate with the bulk mass of water reducing temperature extremes. A comparison of air temperature exceedence at the various weather stations easily distinguishes the maritime sites. Table 3.9 shows that the stations with the highest 10% exceedence temperature also have the lowest 90% exceedence temperature. This is characteristic of maritime conditions

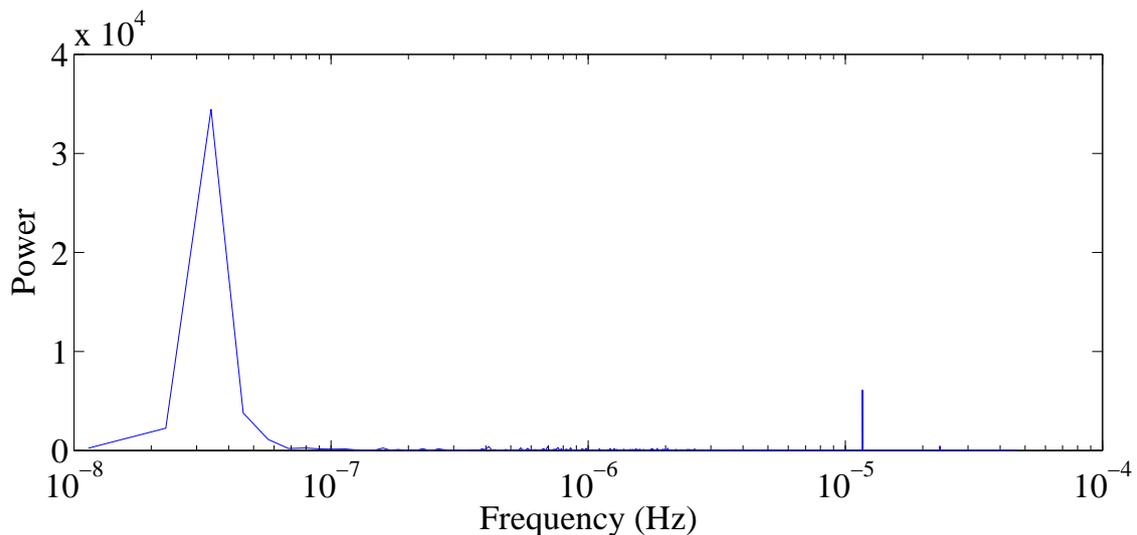


Figure 3.8: Spectral analysis of temperature variation at Point Wilson, evidencing the seasonal ($3.2 \times 10^{-8}\text{Hz}$) and the diurnal ($1.1 \times 10^{-5}\text{Hz}$) signal.

Table 3.9: Rank of exceedence air temperatures (AT) [$^{\circ}\text{C}$] for 10% and 90% levels. More maritime sites have a high rank for the 10% exceedence level and a low rank for the 90% exceedence level and are indicated in boldface.

10% exceedence			90% exceedence		
Rank	Station	AT	Rank	Station	AT
1	South Channel Island	10.3	15	South Channel Island	18.9
2	Brighton	9.6	14	Point Wilson	19.3
3	Frankston	9.5	13	Frankston	19.9
4	Point Wilson	9.3	12	Point Henry	20.1
5	Melbourne	9.2	11	Laverton	20.6
6	Point Henry	8.9	10	Geelong South	20.7
7	Paisley	7.2	9	Moorabbin	20.7
8	Dandenong	8.6	8	Point Cook	20.9
9	Footscray	8.3	7	Footscray	21.2
10	Geelong South	8.3	6	Box Hill	21.7
11	Moorabbin	7.9	5	Melbourne	21.9
12	Box Hill	7.6	4	Paisley	21.9
13	Alphington	7.5	3	Dandenong	22.0
14	Point Cook	7.2	2	Alphington	22.3
15	Laverton	7.0	1	Brighton	22.6

and indicates that South Channel Island, Frankston, Point Wilson and Point Henry are most representative of the air temperature in PPB. As with the wind analysis, these stations are situated on or by PPB.

A scatterplot of the spatial mean against the spatial standard deviation of the four maritime sites, figure 3.4d, suggests that the spatial variance is constant over the main air temperature range, 8–20 $^{\circ}\text{C}$. At higher temperatures this observation no longer holds and the standard deviation increases with temperature. The average spatial variation for mean temperatures in the range 8–20 $^{\circ}\text{C}$ is 1.13 $^{\circ}\text{C}$.

3.3.5 Precipitation

Precipitation is recorded at four stations around PPB. The time series at South Channel Island finishes in July 2002, while the remaining stations continue until November 2003. The average annual rainfall, for the period of record obtained, is presented in table 3.10. Similar to the findings of the PPBES, precipitation

Table 3.10: Annual precipitation for the period of record supplied indicating the significant variation in precipitation across PPB.

Station	Annual Precipitation [mm]
South Channel Island	412
Moorabbin	704
Melbourne	579
Laverton	526

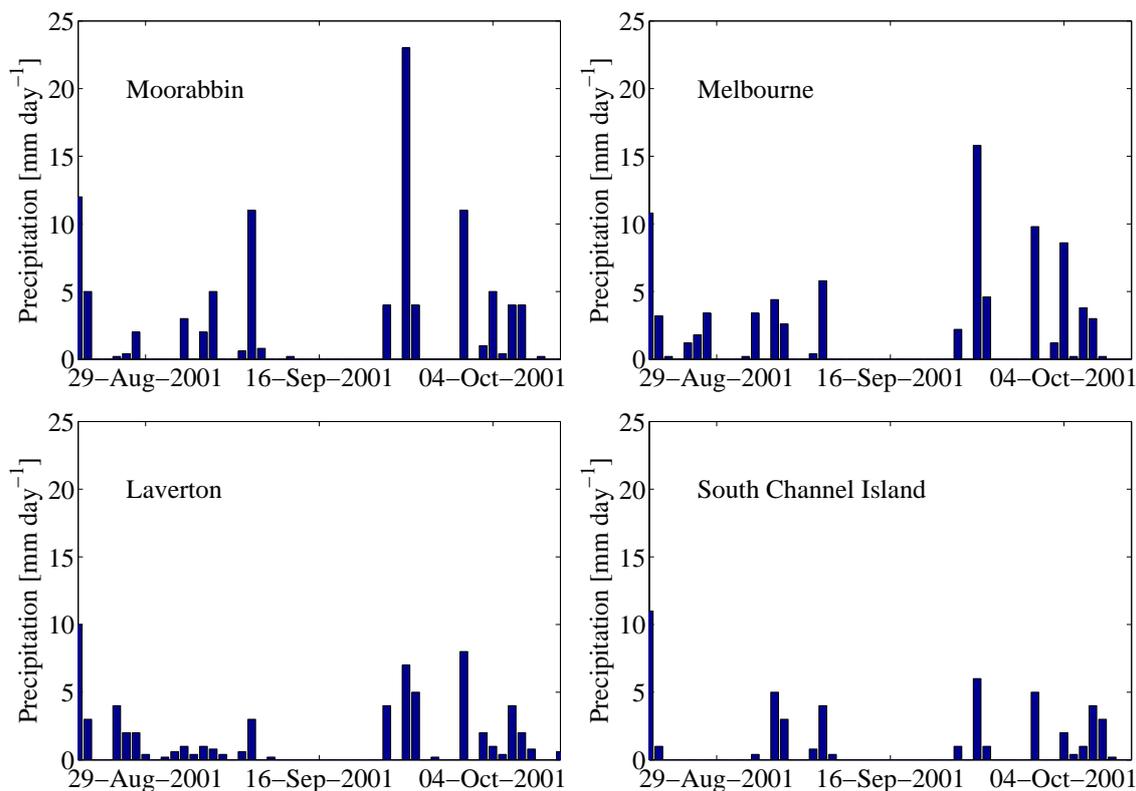


Figure 3.9: Variability of precipitation at four stations around PPB.

varies across PPB increasing from west to east. Although, notably South Channel Island receives significantly less rainfall than the other locations. This is because the open aspect of the site, situated within PPB, reduces the amount of precipitation recorded. As such the South Channel Island record should be treated cautiously.

The variability across the stations is indicated in figure 3.9, which displays the precipitation recorded at the four stations over a short period. The figure indicates

that generally the same pattern of rainfall is present across all stations, but the magnitude of the precipitation varies. This spatial variability is indicated by a scatterplot of mean spatial precipitation against spatial standard deviation, figure 3.4e, showing that the variability tends to be proportional to the magnitude of precipitation. The temporal correlation in precipitation is limited and dissipates after about 3–4 days, similar to atmospheric pressure.

3.3.6 Cloud Cover

Cloud cover is measured in oktas and varies from zero (cloud free) to eight (a completely covered sky). Unfortunately, during the period 2000 to 2003 cloud cover is only recorded at Melbourne. However, between 1993 and 1995 cloud cover is recorded at both Laverton and Melbourne. The validity of the cloud cover data at Melbourne was ascertained by contrasting probability density functions of the different cloud cover levels, see figure 3.10. All the histograms have the same shape, which, assuming weather conditions have not altered significantly in the intervening ten years, and that Laverton has similar weather to Melbourne (both quite valid assumptions) indicates that the Melbourne data are consistent. The histograms also imply that on average the area around PPB is often fairly cloudy.

The variability of cloud cover data is determined by calculating the probability of a certain cloud cover level being observed at Laverton, given an observed cloud

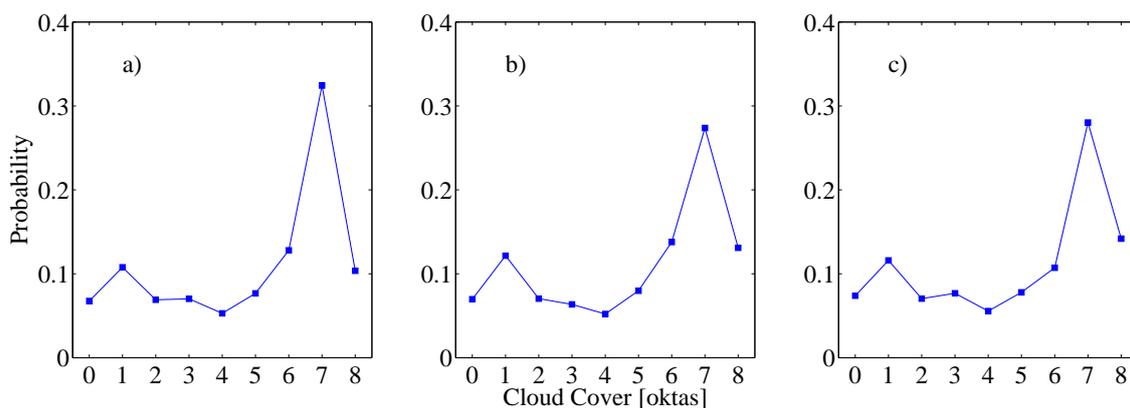


Figure 3.10: Probability density functions of cloud cover levels at a) Melbourne 2000–2003 b) Melbourne 1993–1995 and c) Laverton 1993–1995.

Table 3.11: Probability of cloud cover level being observed at Laverton (L) given a certain cloud cover level has been observed at Melbourne (M). Boldface values referred to in the text.

	M=0	M=1	M=2	M=3	M=4	M=5	M=6	M=7	M=8
Pr(L=0)	0.65	0.17	0.07	0.02	0.00	0.01	0.00	0.00	0.00
Pr(L=1)	0.21	0.51	0.24	0.19	0.08	0.05	0.01	0.00	0.00
Pr(L=2)	0.05	0.13	0.26	0.18	0.15	0.07	0.04	0.01	0.00
Pr(L=3)	0.03	0.07	0.19	0.27	0.19	0.17	0.07	0.01	0.00
Pr(L=4)	0.01	0.03	0.10	0.11	0.17	0.13	0.08	0.03	0.00
Pr(L=5)	0.02	0.03	0.09	0.09	0.16	0.21	0.15	0.06	0.01
Pr(L=6)	0.00	0.02	0.02	0.06	0.12	0.21	0.25	0.14	0.02
Pr(L=7)	0.01	0.02	0.03	0.04	0.09	0.12	0.35	0.60	0.36
Pr(L=8)	0.02	0.02	0.03	0.04	0.04	0.05	0.06	0.15	0.58

cover level at Melbourne. These probabilities are set out in table 3.11, and signify that variability is higher in the middle cloud cover levels (3–5 oktas) than at the edges (0–1 oktas or 7–8 oktas). For instance, if a cloud cover of 0 oktas is recorded at Melbourne, the probability of a cloud cover level of 2 oktas (a difference of 2 oktas) being recorded in Laverton is 5%; whereas if 5 oktas is recorded at Melbourne, the probability of 7 oktas (also a difference of 2 oktas) being recorded at Laverton is 12%. This indicates that at extremes of cloud cover, there is low spatial variability, while at intervening periods the spatial variability of cloud cover is higher.

3.3.7 Relative Humidity

An analysis of the relative humidity data pointed to a few errors in the Point Wilson record. Some minor errors were also noticed in the Moorabbin record. These records were discarded.

An exceedence plot of relative humidity, figure 3.11, allows the maritime stations to be identified. Frankston and Point Wilson have more mild conditions, characteristic of maritime locations: fewer low humidity periods. By contrast, Moorabbin and Laverton experience more extremes of humidity, with relatively more high and low humidity periods. The proximity to the sea reduces the number of low humidity periods, although it is not immediately clear why the occurrence of high humidity periods at the maritime sites is also reduced. Considering figure 3.11 the data

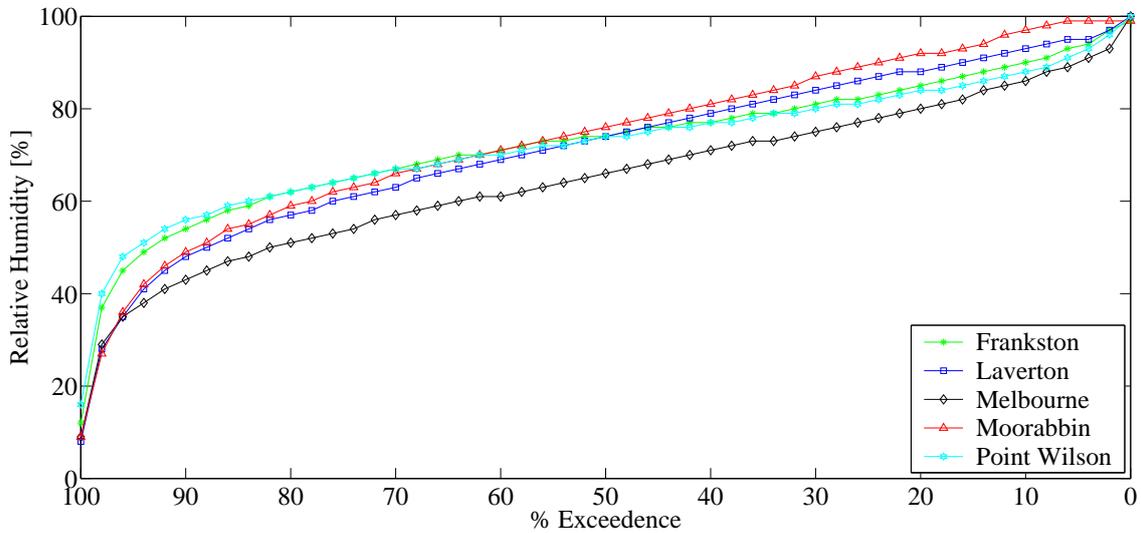


Figure 3.11: Exceedence plot of relative humidity differentiates the maritime stations, Frankston and Point Wilson with more moderate conditions, from the other stations.

measured at Melbourne does not appear representative of PPB.

Similar to the cloud cover analysis, spatial variability is assessed by assigning the relative humidity to 5%-width bins and, calculating the probability of a relative humidity level being observed at Frankston given a relative humidity level observed at Point Wilson. The results are not as clear cut as for cloud cover but tend to indicate that for lower relative humidities there is greater variation in the value. Figure 3.12 shows the probability of a Frankston value being within one bin ($\sim 7.5\%$) of the Point Wilson value. The lower probabilities for the low relative humidity values indicate greater variance. There appears to be slightly greater variance at the very high humidity values too.

The implication of this analysis is that humidity is spatially more varied, and is influenced more strongly by the local climate and conditions. This is especially true during low humidity conditions, when hot dry winds come from the north, then local conditions become more important. A site measuring relative humidity within PPB would be particularly useful here but none is available.

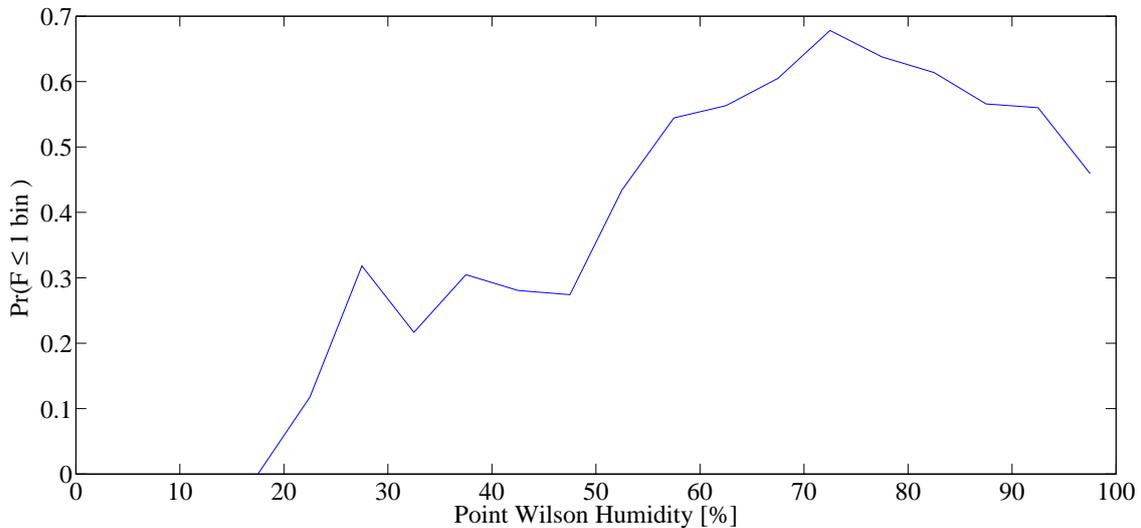


Figure 3.12: Probability of a Frankston relative humidity value being within distance of 7 percent of a Point Wilson value.

3.3.8 Riverine Inputs

While flow may be recorded a number of times along the length of a river, its uncertainty is difficult to assess. The method used to record stream flow correlates recorded water depth to flow and results in greater uncertainty for larger flows. Unfortunately, no rating curve information was available to quantify this. Similarly, spatial variability of stream flow is less meaningful as flow in one stream is quite different from flow in another depending on basin size. However, as stream flow is ultimately driven by precipitation, flow in all rivers around PPB will vary with the climatic regime. During dry conditions with less rain the flow in all rivers would reduce, but not in equal proportions, as only the major rivers with large baseflow components would contribute significant water to PPB.

A sense of the climatic conditions is given by figure 3.13, which plots the annual mean flow in the Yarra River from 1978 to 2003. Over this period the mean annual flow is about $14 \text{ m}^3\text{s}^{-1}$. During the period 2000–2003, and especially the later part, the PPB region suffered heavy drought conditions causing very low flows in the Yarra. If the flows in the Yarra are much lower than average, then the contributions from the smaller creeks, with much smaller catchments and smaller base flows, can be assumed negligible.

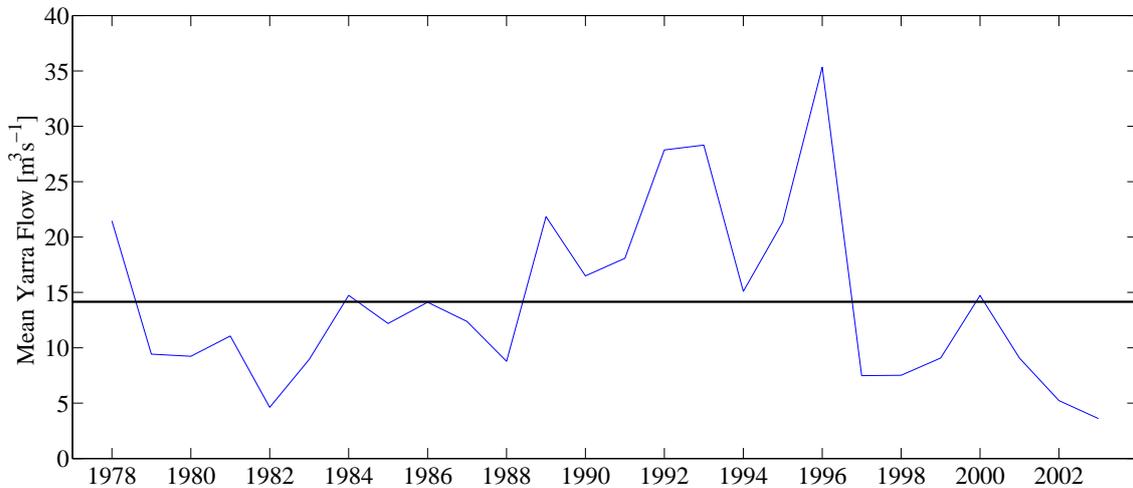


Figure 3.13: Average annual flow in the Yarra River from 1978 to 2003. The horizontal line denotes average conditions over this period.

During the 2000–2003 period average releases of water from the WTP were $4.3 \text{ m}^3\text{s}^{-1}$. Notably in 2003, the releases from the WTP are of the same order of magnitude as the flow in the Yarra, making their inclusion important. This is unlike the period of the PPBES when Yarra River flows are significantly larger. With the exception of Werribee River and Maribyrnong River, data for the other creeks and river shown in figure 3.1 are not available. Their inclusion or otherwise in the PPB modelling will be discussed in chapter 4.

The measurement of temperature in the Yarra River, Gardiners Creek and Merri Creek enables the temperature of waters entering PPB to be calculated. River temperature varies from $9\text{--}23^\circ\text{C}$, which is a larger range than the temperature variation of PPB quoted in the PPBES (Harris *et al.*, 1996). Determining the variation of temperature entering PPB is somewhat difficult as the temperatures at the three locations are not measured at the same time, although usually within a day or so of each other. However, allowing for this time difference the standard deviation of the temperature is 1.06°C . Similar to air temperature, the variation of river temperature with magnitude is constant except outside the main range, refer to figure 3.4f.

3.3.9 Solar Insolation

The GMS-5 satellite derived data reports daily solar energy per unit area. If these were to be used in a hydrodynamic model, disaggregation would be needed to resolve the diurnal cycle. An attempted disaggregation using solar data measured at Brighton (see figure 3.2) — the only data available — was disappointing with maximum insolation being significantly overestimated in the summer and significantly underestimated in the winter when compared with the expected values. The peak radiation should vary from 1100 Wm^{-2} in summer to about 550 Wm^{-2} in winter, with average solar radiation varying from 350 Wm^{-2} to 100 Wm^{-2} between summer and winter (Harris *et al.*, 1996).

The satellite derived daily solar data at each station is used to assess the spatial variation of solar radiation over time. This analysis was performed on data for the period 1993–1995 as the 2000–2003 data appeared unreliable, with significantly

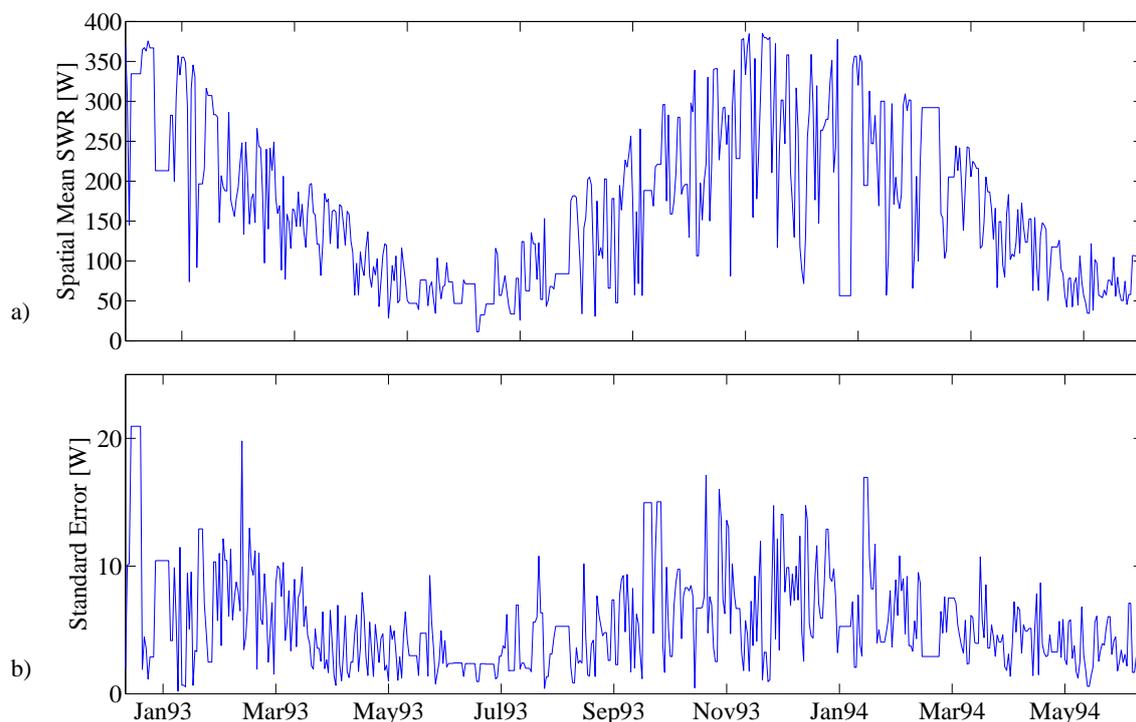


Figure 3.14: Variation in a) mean and b) standard deviation of daily solar radiation over a year and a half derived from GSM-5 data recorded over the period 1993–1995.

larger variance between stations. The results of the analysis indicate that, as expected, solar radiation is higher in summer than winter, figure 3.14a, but also that the variation in solar radiation was higher in summer than in winter, evidenced in figure 3.14b. Yet a scatterplot of spatial mean against spatial standard deviation, figure 3.4g, indicates that for both very low and very high mean solar radiation values, the variation in solar radiation is less. While these results appear to conflict, they can be explained. After seasonal effects, the main variation in solar radiation is due to clouds. At both high and low cloud levels, all areas will receive the same level of solar radiation, because the variation in cloud cover is low (refer to section 3.3.6).

3.3.10 Sea level

The quality of the supplied sea level data was assessed by plotting them all as a time series and checking for a consistent pattern amongst the sites. An example is shown in figure 3.15. Overall, the data was found to be of good quality, with occasional gaps of missing data and high frequency noise that contaminates the true tidal signal.

An investigation of the time series indicates the physical workings of PPB. The changes in sea level throughout PPB are controlled primarily by the sea level in Bass Strait. Higher sea levels in Bass Strait, as seen at Lorne, drive high sea levels in PPB, but the narrow and shallow entrance at The Heads significantly attenuates the tidal signal.

Just outside of The Heads, at Point Lonsdale, the amplitude of the tidal signal is reduced, compared with Lorne, indicating a loss of tidal energy as the tidal wave approaches PPB. Inside The Heads, at Queenscliff, the tidal signal is reduced further indicating the loss of energy across The Heads. Within PPB the tidal signals at all tidal gauges are more or less in-phase, with 3.5 hours lag relative to the Bass Strait signal. Most of the inner Bay sites, Hovel Pile, Breakwater Pier, and Geelong, have similar amplitudes, although the amplitude at West Channel Pile is slightly lower. The travel time across the Great Sands and The Heads leads to the phase difference between Lorne, Point Lonsdale, Queenscliff and the other Bay sites.

The similar tidal signals in the northern part of PPB confirms the assumption that the northern basin operates as a single unit, separate from the southern part

of PPB where the tidal signal varies. The implications for this thesis are that a model of PPB should reproduce these characteristics to be considered accurate. The tidal gauge at Stony Point is not used as the travel time to Western Port Bay (adjacent to PPB) distorts the Bass Strait signal, and so it is not as useful as Lorne in representing Bass Strait.

By analysing the mean sea level at Lorne using a moving average with a window of 1-week, the effect of frontal systems on the sea level can be assessed. The variation is generally within 20 cm of mean sea level, although it can reach 50 cm. Such storm surge events may have an impact on the flushing of PPB, but for this thesis they are ignored.

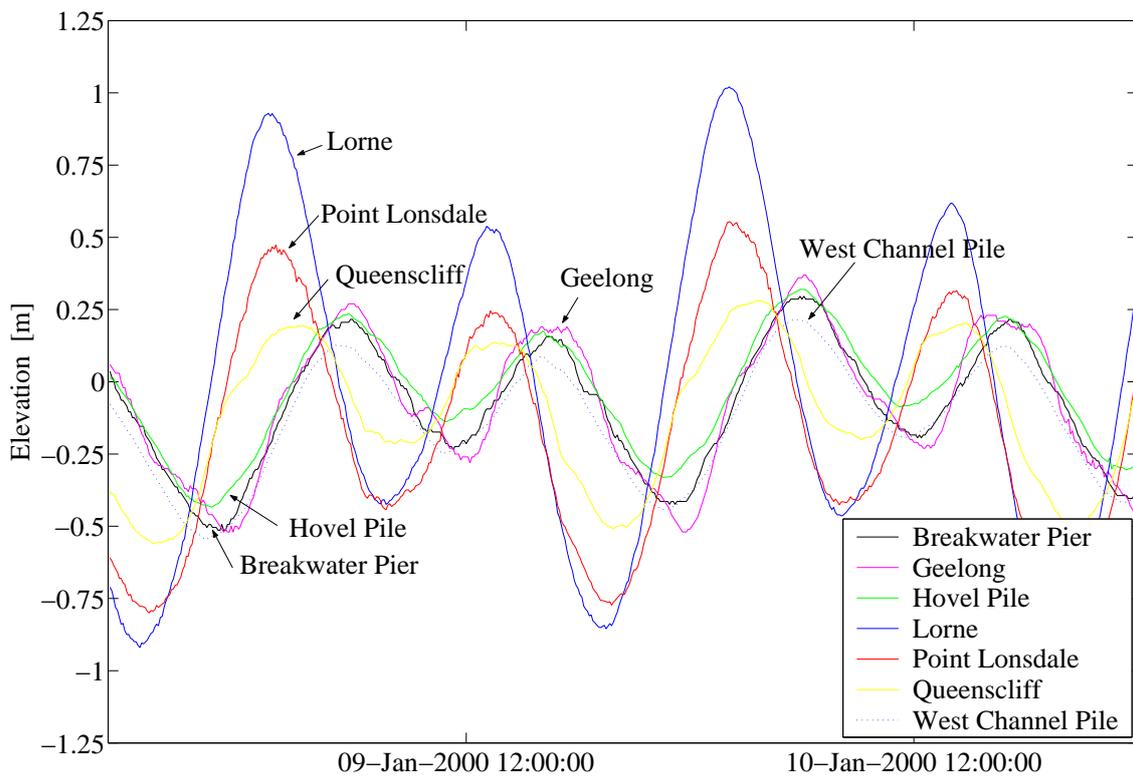


Figure 3.15: Variation of sea level phase and amplitude with location around PPB.

3.3.11 Salinity

Salinity is only recorded at West Channel Pile from 1990-2001. While there appears to be some seasonal variation, in general, salinity varies according to external oceanic forcing and over a long time-scale, refer to figure 3.16. In the ocean salinity is usually well correlated with temperature (see for example Knauss, 1997, figure 8.7), but as figure 3.17 shows this relationship does not exist in PPB. The data from the PPBES indicates that salinity is relatively uniform throughout PPB, although local variations are expected near fresh water sources.

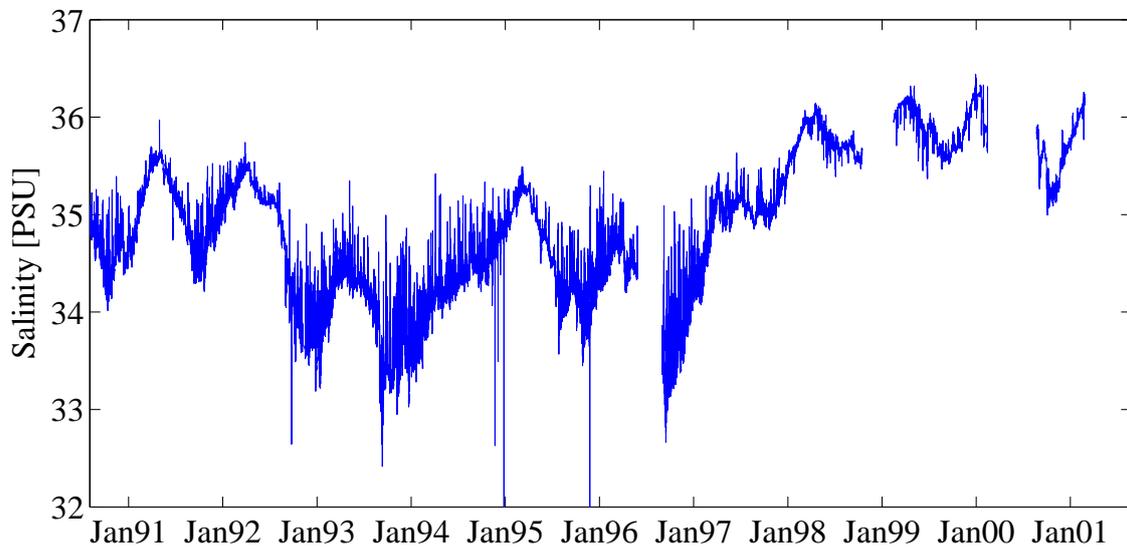


Figure 3.16: Time series of salinity at West Channel Pile between 1990 and 2001.

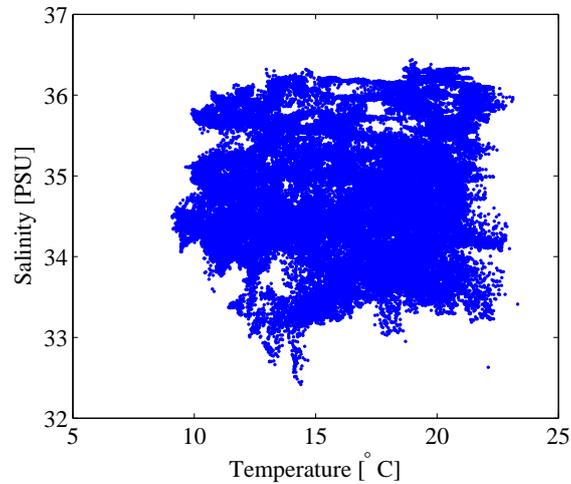


Figure 3.17: A plot of temperature against salinity of data collected at West Channel Pile between 1990 and 2001.

3.3.12 In-situ Temperature

The most useful set of in situ temperature data available for PPB are hourly data collected by MAFRI. While detailed water quality data from the EPA may be useful to detect longer term trends, because of the large interval between observations, the data are of little use for this thesis.

The MAFRI data are generally collected at depth, but towards the end of the record a second set of temperature probes were introduced to collect temperature near the surface. The availability of the data at the different depths is summarised in table 3.12, and they have been plotted for the period 2002–2003 in figure 3.18. The longest record is at West Channel Pile, although it terminates in 2001. The remaining three sites contain shorter records and at different depths.

The diurnal temperature variation decreases with depth. This is most obviously illustrated at the Central site (figure 3.18 right panel), where a clear diurnal signal can be observed in the surface record at about 3 metres depth, while no diurnal variation is present at 18 metres depth. When the surface temperature probe was lowered from the surface level to a deeper level on 30 March, the diurnal signal abated. The diurnal signal reaches to 6 metres as evidenced at Longreef, where a similar temperature signal is seen at both 3 and 6 metre depths. This allows the use of the (longer) Longreef Deep record to assess surface temperature variation, where

Table 3.12: Summary of in situ recorded temperature data.

Site	Depth Recorded	Start	End
Central Deep	18 m	08-Aug-2002	17-Jun-2003
Central Shallow	2.5 m	24-Mar-2003	16-Apr-2003
Hobsons Bay Deep	10 m	08-Aug-2002	17-Jun-2003
Hobsons Bay Shallow	3 m	24-Mar-2003	17-Jun-2003
Longreef Deep	4 m	08-Aug-2002	12-May-2003
Longreef Shallow	3.5 m	24-Mar-2003	16-Apr-2003
West Channel Pile	8 m	27-Aug-1990	18-Sep-2001

as the Central Deep record cannot be used. Furthermore, over most of its record the Longreef Deep record is located at a depth of 4 metres. While the record at the Hobsons Bay Deep site appears reasonable, the comparison between the surface and depth temperatures appears spurious because of the diverging observations, although, this might be on account of cooler fresh water from the Yarra overlying the warmer more saline PPB water.

The underway data, collected during the PPBES, are useful to investigate the spatial variation of temperature. The change in temperature is calculated for all points recorded within fifteen minutes of each other. These temperature differences were assigned to 100 m-width bins according to the distance between the two points. As the distribution of the temperature difference is highly skewed, each bin was analysed using percentiles. Figure 3.19 shows the changes in temperature expected as a function of distance, indicating that there is very little change in temperature over a distance of up to 1 km.

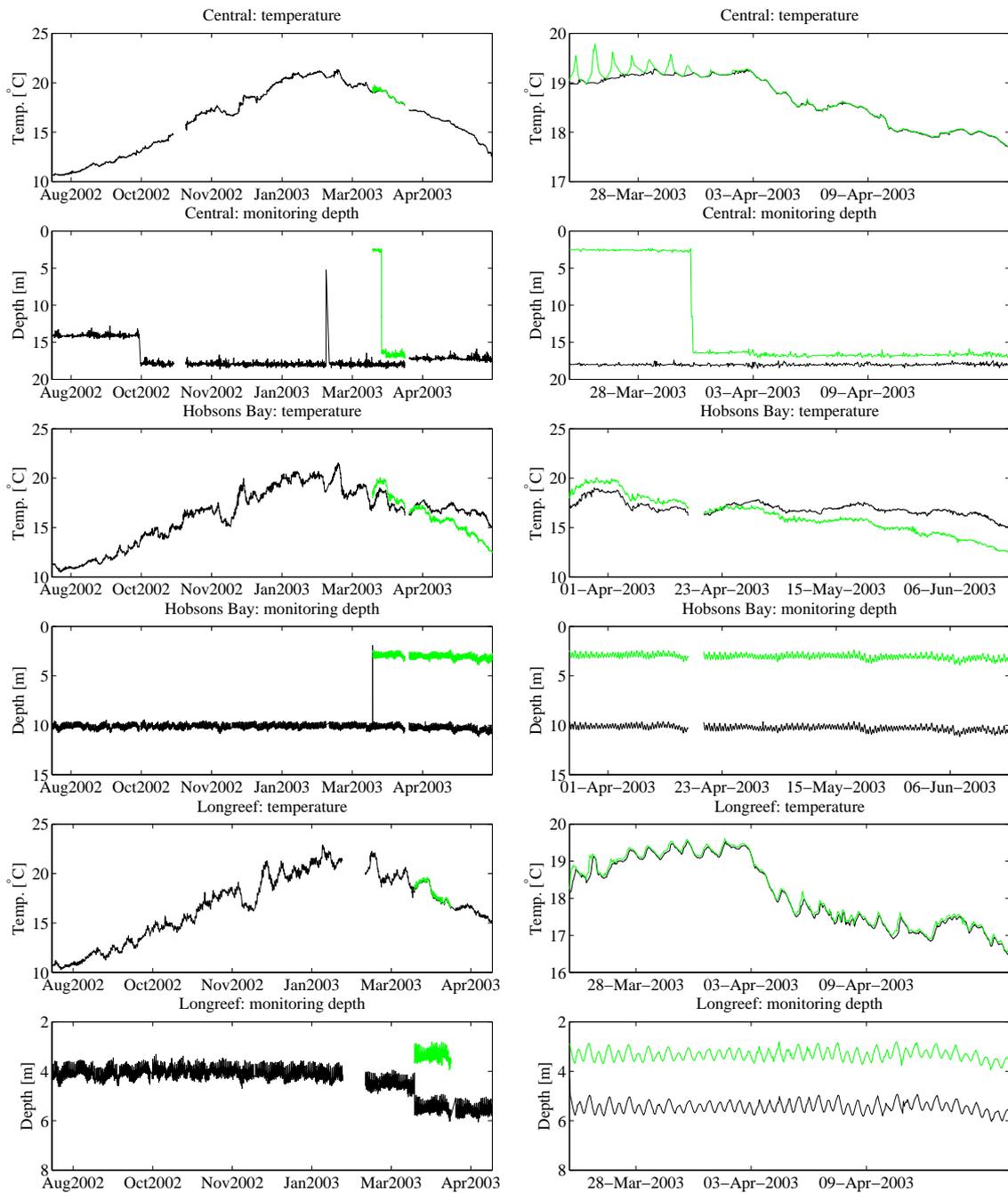


Figure 3.18: Variation in temperature with depth at three monitoring sites within PPB from August 2002 to June 2003. With reference to table 3.12, the Deep record lines are black and the Shallow record lines are green. The left column displays the entire record, while the right column focuses on the shallow record.

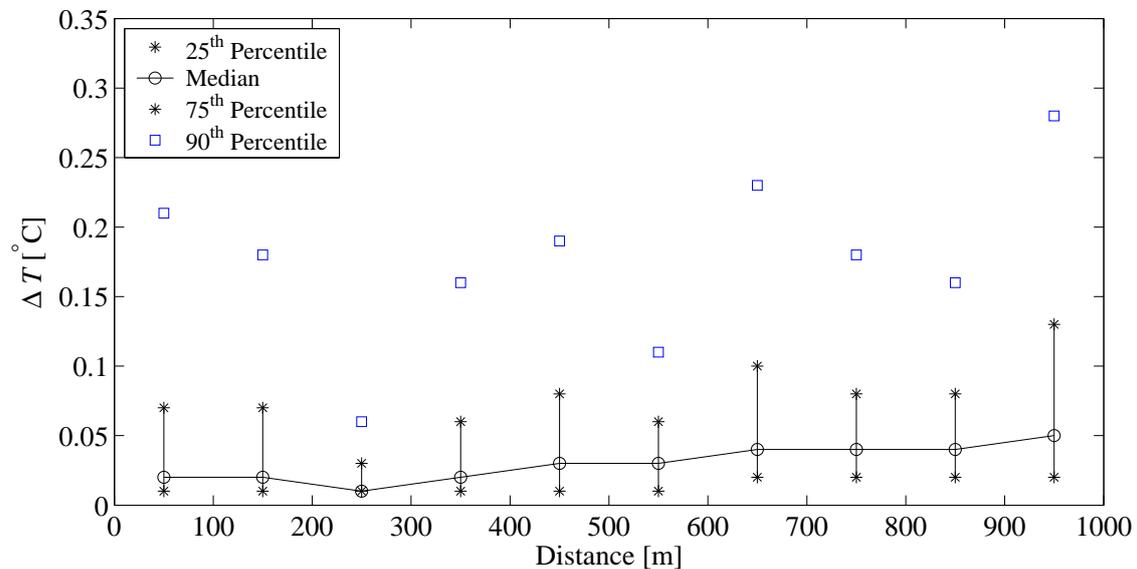


Figure 3.19: Spatial temperature variation with distance in PPB based on underway transect data.

3.4 SST Satellite Observations

The SST data are observed by the AVHRR aboard the NOAA series of satellites. For the period 2000–2003 data are available from NOAA-12, NOAA-14 and NOAA-16

SST is observed by measuring the intensity of the radiation emitted by the water at various wavelengths, which is related to the surface temperature. The AVHRR observes SST by measuring the intensity of radiation in the thermal infrared wavelengths, specifically in the 3–4 μm and 10–12 μm atmospheric windows (Liou, 2002). SST observations are also possible in the microwave bands but thermal IR is more accurate (Závody *et al.*, 1995). While the 3–4 μm window is more accurate its use is restricted to night time because of contamination from reflected solar insolation (Deschamps and Phulpin, 1980) making the images unusable (Brown and Minnett, 1999). Both day and night images are available for this thesis.

The radiation emitted by the water surface is subject to atmospheric interference. For accurate observations this must be removed. An early linear algorithm was outlined by Deschamps and Phulpin (1980). Due to assumptions made to derive the algorithm, this theoretical model departs from reality for high temperatures, high scan angles, and high humidity (Brown and Minnett, 1999). This prompted the

development of the linear Multi-Channel SST (MCSST) algorithm (McClain *et al.*, 1985), which accounts for high scan angles, and later the nonlinear Cross Product SST (CPSST) algorithm (Walton, 1988) of which the nonlinear SST (NLSST) algorithm is a simplification (Li *et al.*, 2001) and accounts for high temperatures and high scan angles. The NLSST algorithm is used operationally (Nalli and Smith, 1998).

Although the coefficients of the Deschamps and Phulpin (1980) algorithm can be determined theoretically, historically coefficients have been determined empirically by regressing the observed brightness temperatures against buoy and ship observations (Nalli and Smith, 1998). More recently, the coefficients of newer instruments (e.g. the Advanced Along Track Scanning Radiometer (AATSR)) have been calibrated against radiative transfer function models (Závody *et al.*, 1995). These models incorporate the physical understanding of the absorption and scattering processes in the atmosphere to predict the brightness temperature observed based on the SST and atmospheric conditions.

A key assumption in deriving the atmospheric correction algorithms is that the atmosphere is non-scattering at thermal infrared wavelengths, ignoring aerosols. In reality, stratospheric aerosols from volcanic eruptions can cause significant scattering, altering the retrieved surface temperatures by up to 1.5°C (Merchant *et al.*, 1999). Also, localised aerosols, such as dust from the Sahara blown off the north west African coast, remain a major source of residual error in current algorithms (Brown and Minnett, 1999). Furthermore, the absorption and scattering ability of marine aerosols, a function of the aerosol size, depends strongly on relative humidity (Závody *et al.*, 1995) as well as the wavelength in consideration, and thus becomes a further source of uncertainty.

The observation errors associated with AVHRR SST observation are a function of the error in the atmospheric correction and the noise characteristics of the instrument (Brown and Minnett, 1999). Furthermore, the errors are additive with the number of AVHRR channels used. For this reason the use of two channels is advised except for very low noise instruments (Deschamps and Phulpin, 1980). The standard error of the MCSST is 0.6–0.7°C, while the standard error of the NLSST is about 0.5°C (Nalli and Smith, 1998).

3.4.1 Skin vs Bulk Temperature

When using SST data, consideration of the difference between skin and bulk temperature is important. The skin surface temperature, which the satellite instrument observes may vary considerably from the temperature a few mm below the surface and certainly from the temperature up to a metre below. To clarify this the term *skin temperature* refers to the temperature of the water surface, whereas *bulk temperature* refers to the average temperature of the water body in, say, the top metre of the water column. Due to turbulence in the upper part of a water body, the bulk layer is considered well mixed and, as a result, has a more or less uniform temperature.

The importance of this difference is that while the skin temperature is observed, it is the bulk temperature that is desired. Numerical models discretise the water column into layers and typically, the vertical discretisation is of the order of 1 metre at the surface. At this resolution the model cannot resolve the skin bulk difference and only predicts the bulk temperature. The vertical resolution cannot be increased further without inducing errors through rapidly varying grid spacings. If SST data derived from skin temperatures are compared with the model prediction of bulk temperatures, an error may result on account of the skin-bulk effect. This is further complicated as the SST algorithm is generally calibrated on buoy observations: a bulk temperature measurement (Nalli and Smith, 1998).

The physical mechanism for the skin-bulk difference is energy transfer between the water and the atmosphere. Visible solar radiation penetrates the water column. While some of this radiation is scattered back and escapes to the atmosphere as water leaving radiance, the remaining radiation is mostly absorbed and transfers its energy to the water, heating it. As absorption is proportional to the distance travelled, more solar radiation is absorbed closer to the surface. Although some of this heat may be advected away by currents, to maintain thermodynamic stability, this heat must eventually be lost to the atmosphere at the surface. The different mechanisms of energy loss are discussed by Hasse (1971). The removal of energy at the surface cools it, creating a temperature gradient, which accounts for the surface skin being slightly cooler than the water directly beneath.

The processes of absorption of solar radiation and the loss of energy to the

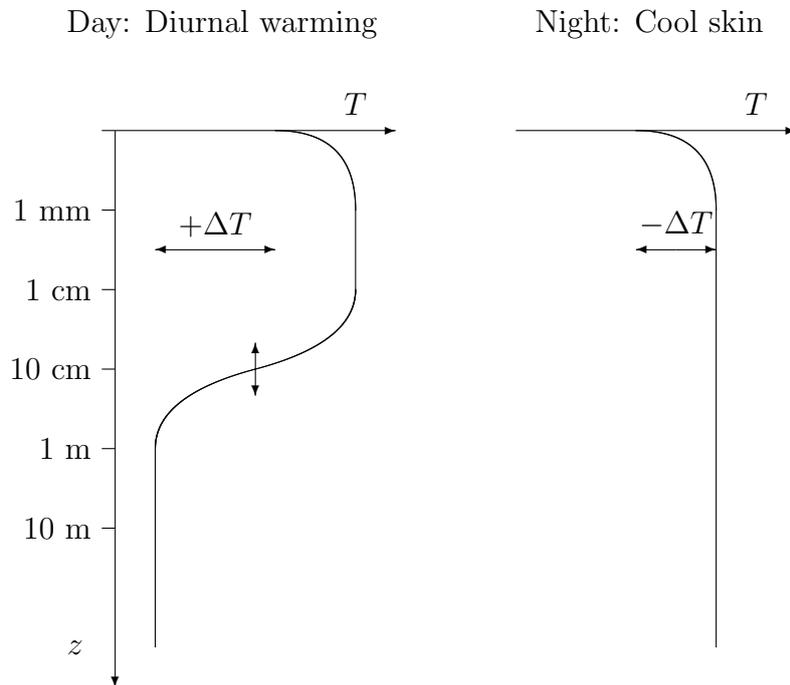


Figure 3.20: Schematic illustrating the skin cooling and diurnal warming during the day, based on illustration of Schluessel *et al.* (1990).

atmosphere combine to induce two effects (illustrated in figure 3.20): the cool skin effect and diurnal warming. The cool skin effect is where the loss of energy to the atmosphere results in the skin being slightly cooler than the water directly beneath. Diurnal warming occurs when the upper part of the bulk layer, including the cool skin, is warmed due to the absorption of solar radiation. This occurs during calm sunny conditions, when the mixing between the surface and deeper layers is suppressed and intense solar radiation heats the surface. As diurnal warming is driven by solar radiation, it only occurs during the day. During the night the heated surface layer gradually disappears and the cool skin effect remains. The magnitude of the diurnal warming can be of the order of degrees (Stuart-Menteth *et al.*, 2003), while the cool skin effect has a magnitude in tenths of degrees (Wick *et al.*, 1996).

Cool skin

The cool skin effect operates under different physical regimes which control the magnitude of the temperature difference (Emery *et al.*, 2001). However, when contrasting different models predicting the cool skin effect, Emery *et al.* (2001) found that although the models could accurately reproduce overall tendency, a large amount of unresolved variability remained.

For SST observations the cool skin effect is generally ignored. An attempt to explicitly account for the skin layer in deriving satellite estimates of bulk temperature found that no improvement in prediction resulted. Improvements in the model physics were offset by errors in parameters and forcing data (Emery *et al.*, 2001). The magnitude of the cool skin effect is within the uncertainty range of the SST data itself. As instrument design and atmospheric correction improves, the cool skin effect will become a more pressing problem. In this thesis the cool skin will not be dealt with. For more information the interested reader is directed to the papers by Saunders (1967), Grassl (1976), Schluessel *et al.* (1990), Wick *et al.* (1996), and Emery *et al.* (2001).

Diurnal warming

Diurnal warming has been observed in the ocean since the mid 1970's (Cornillon and Stramma, 1985), however theoretical relationships for diurnal warming appear less developed than those for the cool skin effect. Although physical models of diurnal warming are lacking, a range of semi-empirical equations have been proposed. As the magnitude of diurnal warming is larger than the uncertainties of the SST data its effect will be considered in this thesis. Two more recent approaches to the diurnal effect are presented. These methods differ significantly in their derivation, which is evidenced in their predictions.

The approach of Gentemann *et al.* (2003) used satellite (thermal and microwave) estimates of SST at a range of times and weather conditions. Their analysis suggested that the temperature difference between the skin and the bulk can commence as early as 8AM, then rises during the day to a peak at about 3PM. The temperature difference then reduces until about 11PM. They suggest the peak diurnal temperature difference can reach 2.8°C under optimal conditions. By regressing their data

against an empirical model they derived an algorithm based on wind speed, W , and the average daily solar flux at the top of the atmosphere, \overline{Q}_I .

The Gentemann *et al.* (2003) model is

$$\Delta T = \begin{cases} 0.344 \cdot f(t) e^{-0.29W} [(\overline{Q}_I - Q_0) - 1.444 \cdot 10^{-3} (\overline{Q}_I - Q_0)^2] & , \text{ for } \overline{Q}_I \geq Q_0 \\ 0 & , \text{ otherwise} \end{cases}, \quad (3.2)$$

with the diurnal variation accounted for by

$$\begin{aligned} f(t) = 0.001 \times [& 6.814 - 6.834 \cos(\omega t) - 8.427 \sin(\omega t) \\ & + 1.447 \cos(2\omega t) + 4.274 \sin(2\omega t) \\ & - 0.407 \cos(3\omega t) - 0.851 \sin(3\omega t) \\ & + 0.457 \cos(4\omega t) - 0.555 \sin(4\omega t) \\ & + 0.101 \cos(5\omega t) - 0.375 \sin(5\omega t)], \end{aligned} \quad (3.3)$$

where t is the time in hours, ω is the frequency equal to 0.2668 hr^{-1} and Q_0 is the minimum insolation level, 24 Wm^{-2} . Figure 3.21 illustrates equation (3.2) graphically. The left-hand plot describes the variation of the diurnal variation over a 24-hour period, while the right-hand plot shows the expected maximum skin-bulk difference at 3PM for a range of wind speeds and solar radiation levels. Greater di-

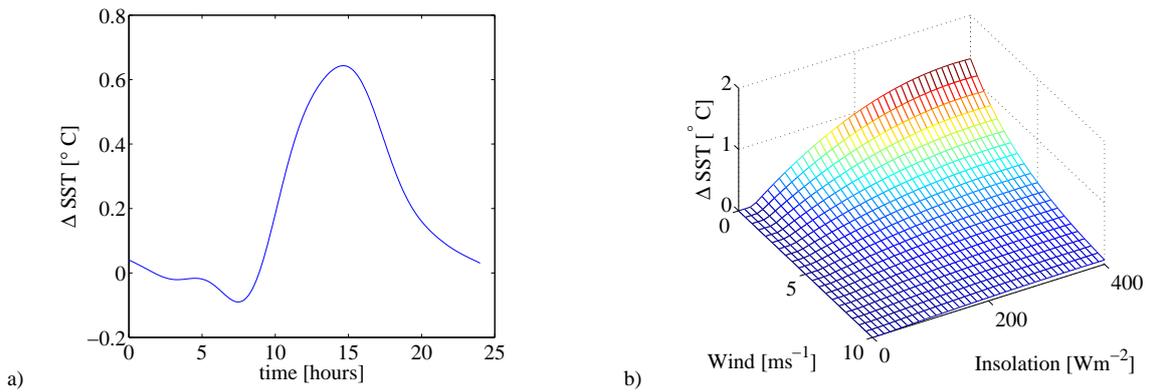


Figure 3.21: Plot indicating a) the temporal variation of diurnal warming predicted by Gentemann *et al.* (2003) based on a wind speed 2 ms^{-1} , and insolation 250 Wm^{-2} , and b) variation of maximum expected diurnal ($\sim 3 \text{ PM}$) warming.

urnal variation is expected at lower wind speeds and higher insolation. Lower wind speeds reduce the shear stress and therefore reduce mixing, allowing stratification to form. Stronger insolation provides more energy to heat the upper layers, magnifying the diurnal variation.

Another approach for calculating diurnal variance was presented by Kawai and Kawamura (2002). They validated a theoretical model of the skin-bulk difference with fixed and drifting buoy data. Skin and bulk temperature data for range of conditions was generated by this model and regressed into an empirical relationship of the maximum diurnal warming temperature difference. This approach was subsequently validated by Kawai and Kawamura (2003) and applied by Stuart-Menteth *et al.* (2003).

The Kawai and Kawamura (2002) model is

$$\Delta T = a(Q_{I_{\max}})^2 + b[\ln(\overline{W})] + c(Q_{I_{\max}})^2 \ln(\overline{W}) + d, \quad (3.4)$$

where $Q_{I_{\max}}$ is peak solar radiation, \overline{W} is daily mean wind speed, and the calibration coefficients, $a-d$ are given in table 3.13. Equation (3.4) is restricted that if ΔT is predicted as negative, ΔT is set to zero, i.e. no cool skin is allowed. The variation of maximum diurnal warming predicted is shown in figure 3.22. Similar to the Gentemann *et al.* (2003) model, diurnal warming increases with increasing solar radiation and with reducing wind speed. However, the Kawai and Kawamura (2003) model predicts a much higher diurnal warming.

Table 3.13: Coefficients for Kawai and Kawamura (2002) model presented in equation (3.4).

Coefficient	$W > 2.5\text{ms}^{-1}$	$W > 2.5\text{ms}^{-1}$
a	3.2708×10^{-6}	5.6814×10^{-6}
b	-7.9982×10^{-2}	4.0052×10^{-1}
c	-1.3329×10^{-6}	-3.9637×10^{-6}
d	7.3287×10^{-2}	-3.6700×10^{-1}

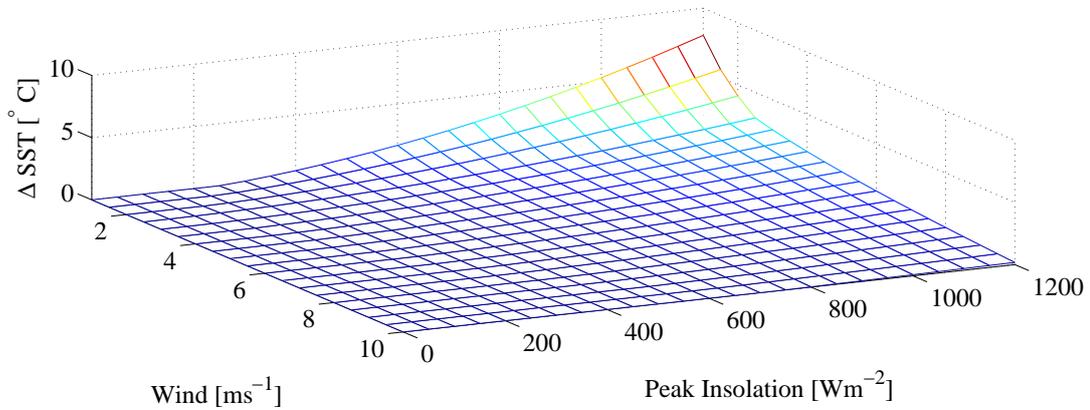


Figure 3.22: Plot indicating variation of expected maximum diurnal warming with wind speed and insolation based on Kawai and Kawamura (2002) model.

3.4.2 SST Pre-processing

The AVHRR data supplied by CSIRO Marine Research were originally in UTC, Coordinated Universal Time, which was converted into Australian Eastern Standard Time for use in the thesis by the addition of 10 hours. The data were supplied on a 0.01° grid, giving a pixel width of approximately 1 km. The satellite data was investigated by considering its spatial variability or spatial consistency. Two main errors contaminate the satellite images: land contamination and cloud contamination. The effect of land contamination is shown in figure 3.23, where the time-averaged spatial variability is plotted. The significantly higher values around the coast, while potentially indicative of a higher dynamic coastal activity, are actually a result of land contamination of pixels. Land contamination extends up to 2–3 kilometres from the shore.

Cloud contamination can also be a significant problem. A preliminary cloud mask had been applied by CSIRO when processing the data, but as figure 3.24b indicates this is not accurate enough and significant errors were observed in pixels surrounding cloud contaminated pixels.

The issues of land and cloud contamination are dealt with by applying further masks to the data: a land contamination mask 2 pixels wide, a cloud contamination mask 4 pixels wide. A mask 2-pixels wide means that if a cell is found to be contaminated, all cells within 2 pixel-widths of the contaminated cell are discarded

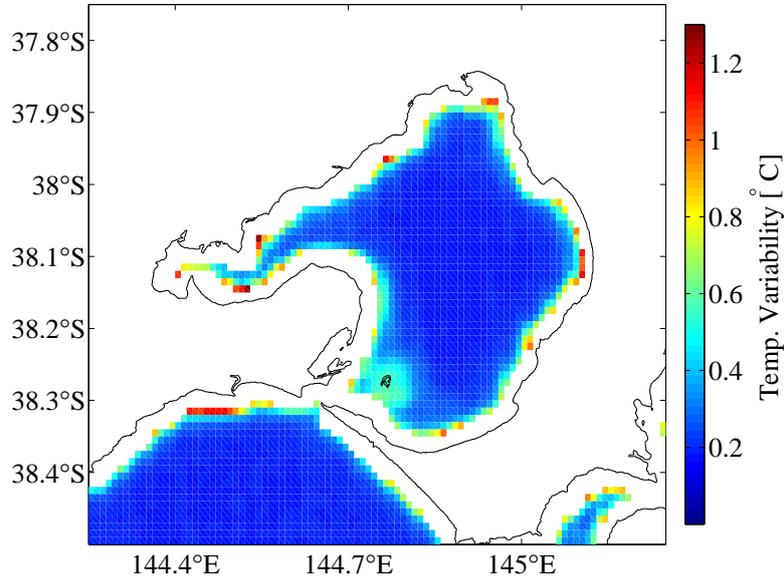


Figure 3.23: Average spatial variability of each satellite pixel, indicating that proximity to land contaminates the image.

as well as the contaminated cell. The cloud cover mask and the land mask are illustrated in figures 3.24b and 3.24c respectively. The post-masking images result is shown in figure 3.24d. This procedure greatly improves the quality of the satellite data. A further quality check of the data was performed by removing cells where the temperature standard deviation of the four surrounding cells was greater than 0.3°C . A sanity check was also made removing data outside the expected temperature range of $9\text{--}24^{\circ}\text{C}$

3.4.3 Comparison with In Situ Observations

While the spatial consistency of the satellite data is evaluated and improved by masking, the accuracy of the values observed by the satellite are gauged by contrasting the satellite data with the in situ observations. The geo-spatial location of each in situ station is used to assign the data to a corresponding satellite pixel. If a satellite image and an in situ record occur within half an hour of each other they are compared. The in situ datum is compared against the average of the nine satellite

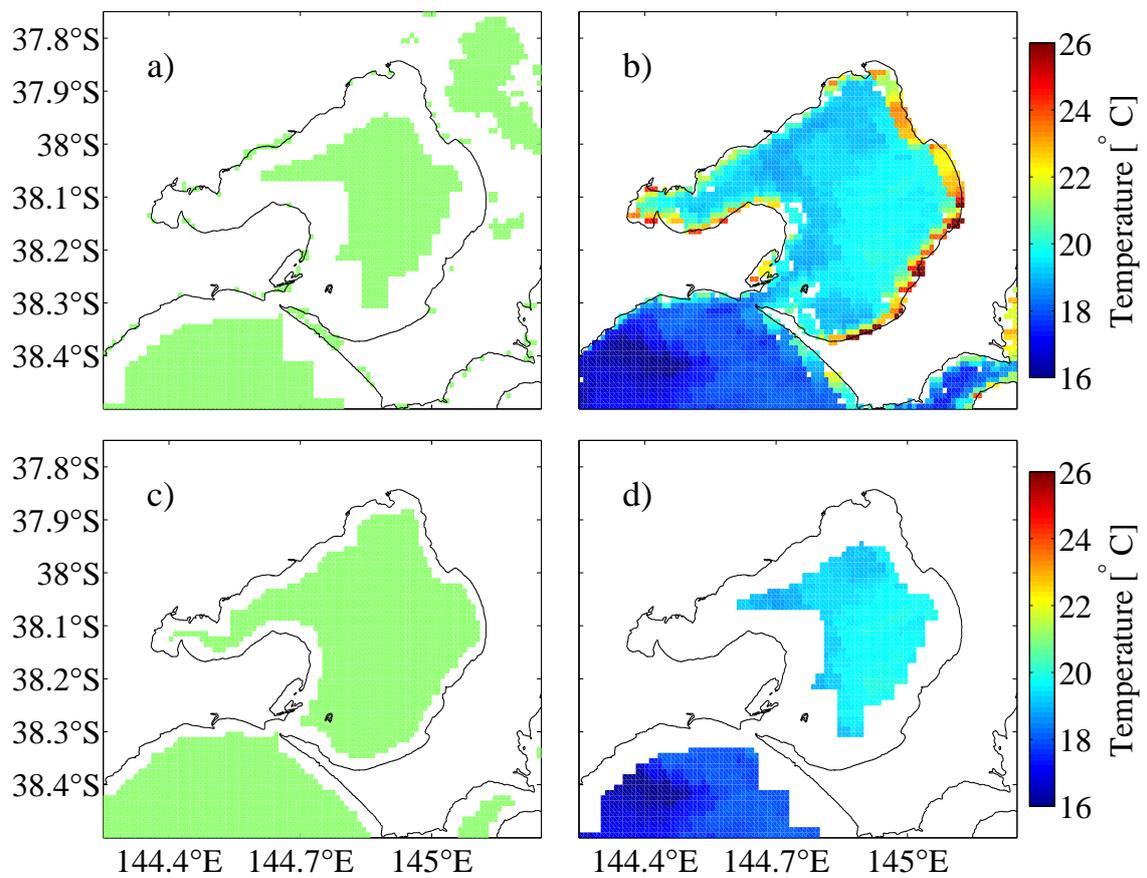


Figure 3.24: Images of AVHRR satellite data supplied for assimilation: a) the cloud cover mask, where coloured cells are uncontaminated, b) the original image, c) the land cover mask and d) the final processed image.

pixels about its geo-spatial location. The average is used to reduce any observation error. The in situ observations are measured with such accuracy that they can be taken as the true temperature. The observation error in the SST is then taken as the difference between the SST and the in situ observations.

The results are displayed in figures 3.25 and 3.26 for the shallow and deep sites respectively. The panels indicate that the satellite data mirrors the actual temperature well, capturing the seasonal variation. The different histograms show that except for Hobsons Bay Deep, the error distributions appear balanced. At Hobsons Bay Deep the satellite overestimates temperature during the summer, but underestimates temperature in autumn: towards the end of the record. As none of the other sites contain such error, the Hobsons Bay Deep in-situ record is thought to be responsible.

The comparison of satellite data against monitoring data should only be considered for sites which are shallow enough to pick up the diurnal signal; the deeper sites should match the seasonal trends but not the diurnal variations. This invalidates the comparison with the Central Deep site, the Hobsons Bay Deep site and West Channel Pile. The Central Shallow site should also be judged cautiously as during the latter part of its record the temperature probe was lowered to depth. For Longreef Shallow and the Hobsons Bay Shallow record the error distributions appear balanced, although the extensive Longreef deep record appears to indicate larger errors during summer. The statistics of the difference are given in table 3.14. The statistics indicate that the average mean error across all monitoring sites (excluding Hobsons Bay Deep) is -0.16°C , which is well within observational error. The standard deviation is on average 0.75°C excluding Hobsons Bay Depth). This value fits well with the standard error range of the NLSST (0.5°C) and MCSST ($0.6-0.7^{\circ}\text{C}$) algorithms. The analysis indicates that the satellite data have negligible bias temporally, which is important for the assimilation.

The spatial error correlation was assessed by constructing a scatterplot, where each axis represents the observation error at a different monitoring site. If the error is spatially correlated a linear trend should emerge from the data. The results are plotted in figure 3.27, and indicate that for the shallow sites there is a strong correlation of error between sites. The strength of the correlation is indicated by the sample correlation coefficient, R . The values of R show that the observation error

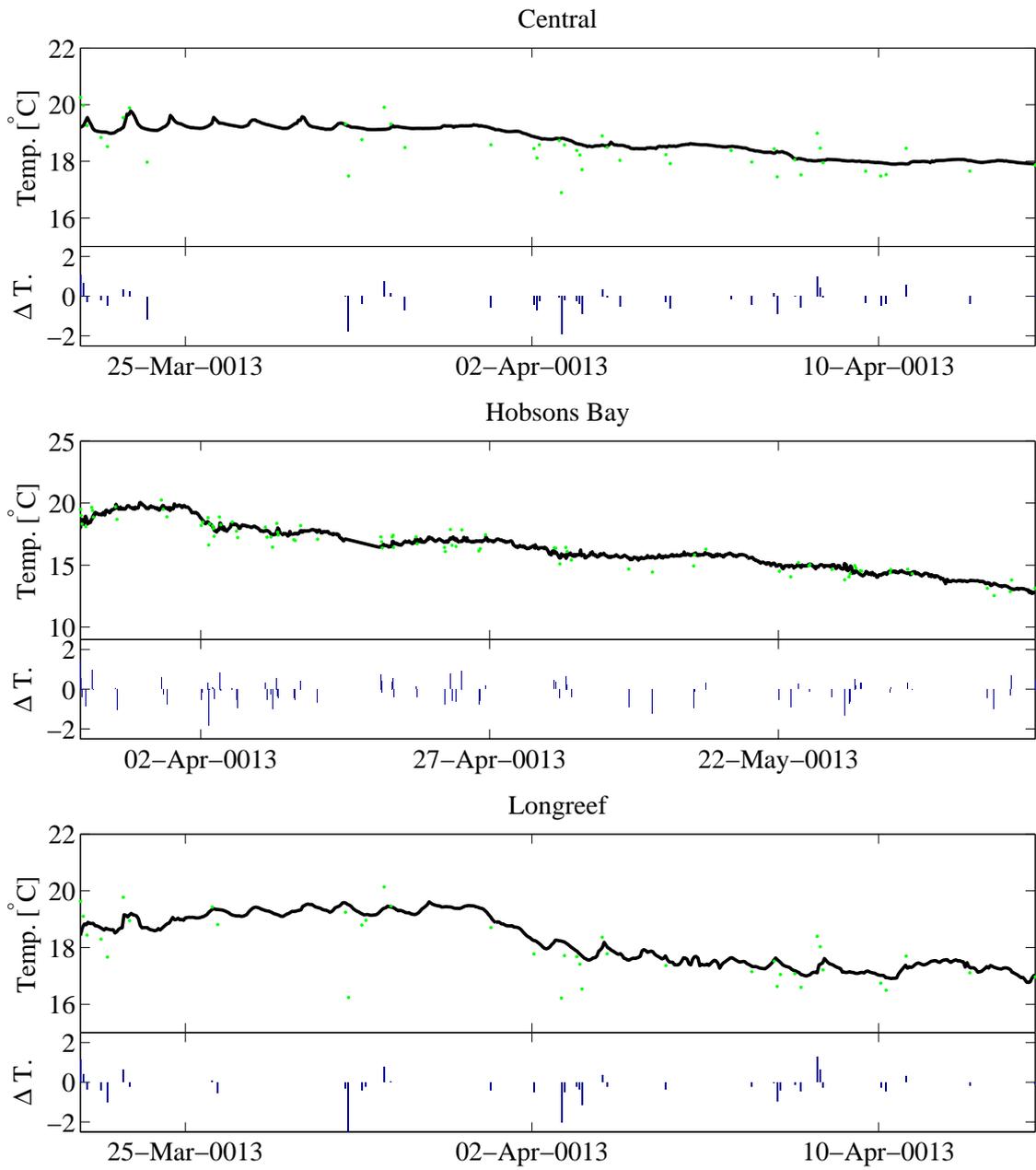


Figure 3.25: Comparison between satellite observations (dots) and in situ monitoring of temperature at shallow locations within PPB (thick black line). Lower panels indicate error in satellite data.

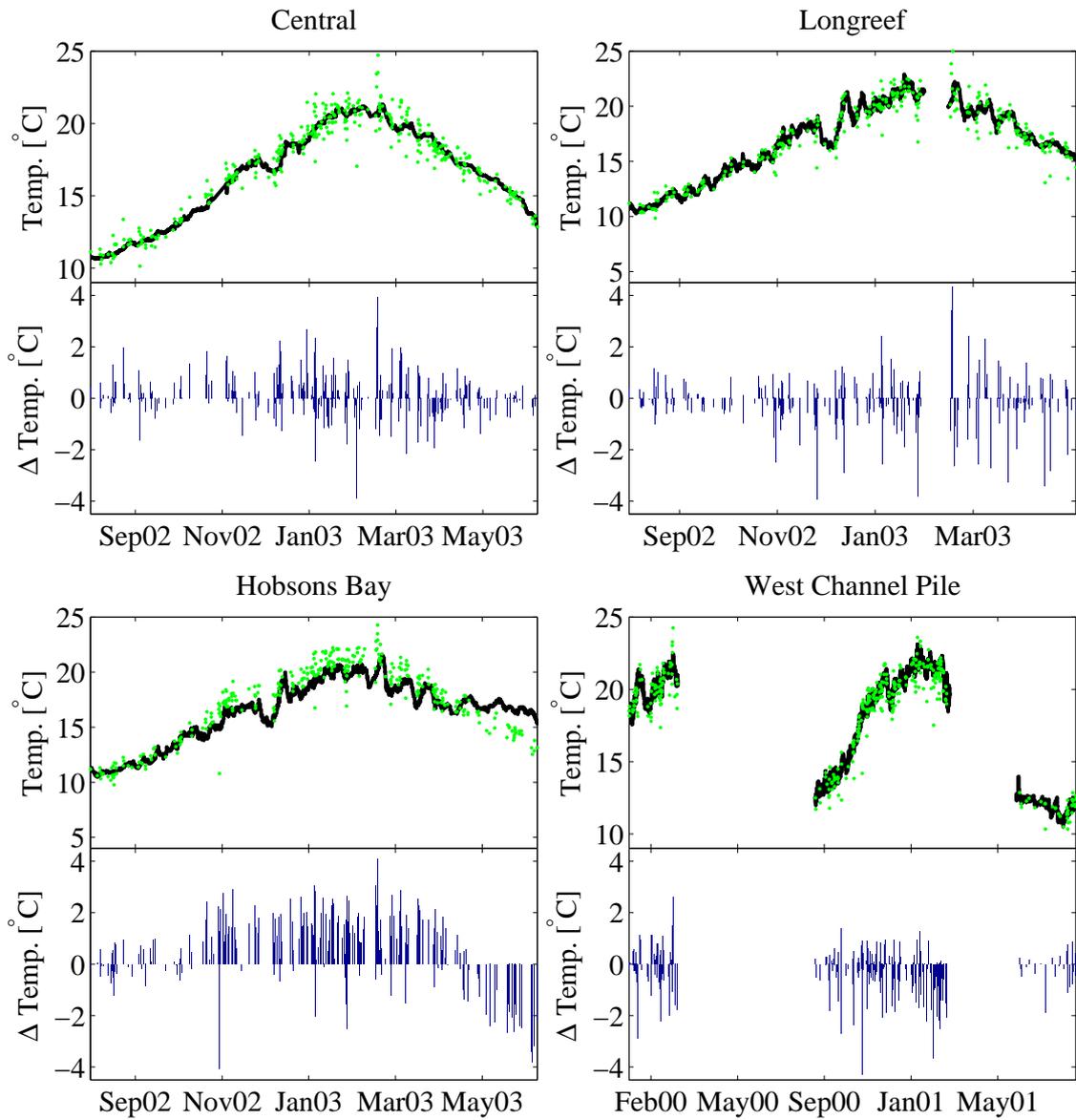


Figure 3.26: As for figure 3.25 but for locations at depth.

Table 3.14: Summary of mean difference and standard deviation between the monitored data and the satellite data.

Station	Mean Error	Standard Deviation
Longreef Deep	-0.18	0.97
Longreef Shallow	-0.26	0.77
Hobsons Bay Deep	0.60	1.26
Hobsons Bay Shallow	-0.13	0.60
Central Deep	0.11	0.82
Central Shallow	-0.23	0.60
West Channel Pile	-0.27	0.76

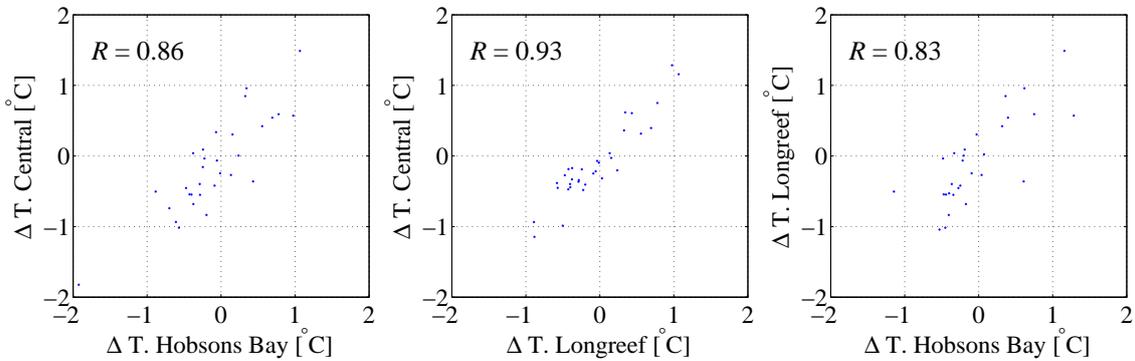


Figure 3.27: Scatter plots indicating spatial correlation of the difference between in situ water measurement and satellite observation at the three shallow locations within PPB.

is spatially well correlated.

This spatial correlation of the error appears connected to diurnal warming. Figure 3.28 plots the average of the Longreef deep errors according to the time of observation. During the early afternoon the average error is positive, overestimating temperature, whereas during the night more of the errors are negative. This pattern is indicative of a diurnal warming effect, although the night magnitude is significantly larger than expected. The reason for the larger than expected night error is unclear but could possibly be due to the global parameters used by AVHRR being unsuitable locally. Recognition of the spatial correlation of the error in the satellite data and removing its effects will need to be addressed for the successful implementation of an assimilation scheme and approaches to removing the error are

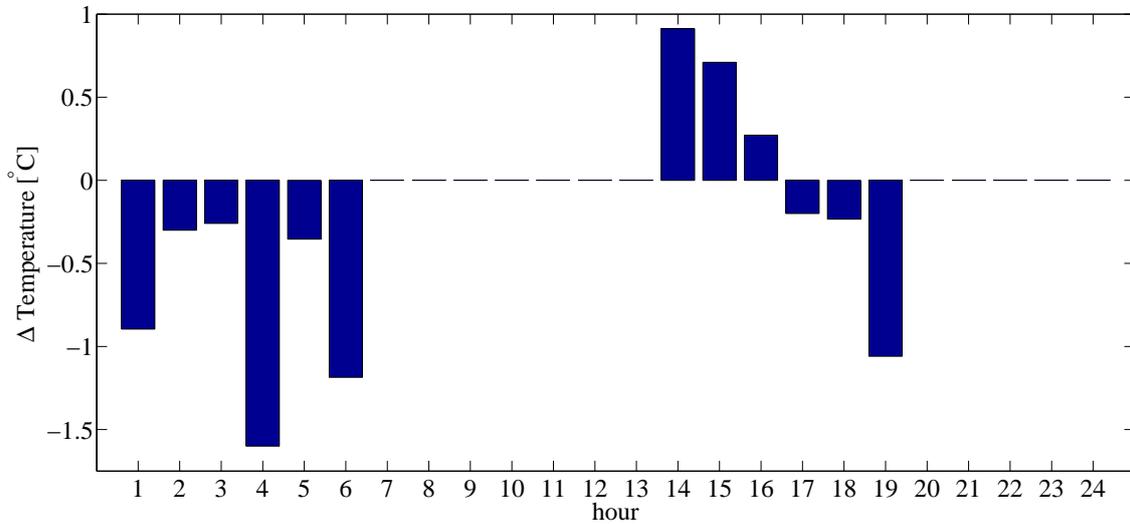


Figure 3.28: Bar plot of mean differences between in situ water measurement and satellite observation at the Longreef Depth site according to the hour of measurement.

applied in chapter 6.

3.4.4 Observation Error Covariance

The observation error covariance matrix, \mathbf{R} , indicates the relationship between the error of different observations. The diagonal elements indicate the error variance of the respective observations, and the off-diagonal elements indicate the covariance between the errors of two observations. If the error in the observations is well correlated, then the observations contain less information than if each were independent.

Three sources of error related to the satellite observation of SST are: i) atmospheric effects, ii) land contamination, and iii) errors in the sensor calibration.¹ While sensor errors may be assumed to be measurement errors and therefore independent, the other sources of error have a spatial component and will therefore introduce a covariance component to the observation error covariance matrix.

The SST retrieval algorithms, described in section 3.4, assume there is some level of atmospheric absorption and emittance of thermal infrared radiation. However,

¹The difference between the skin and bulk temperature is another source of observation error for assimilation, but this will be treated separately and dealt with in later chapters.

scattering by atmospheric particles is generally ignored and in the presence of large concentrations of water vapour (high humidity), the accuracy of the retrieval diminishes. Therefore, the magnitude of the observation error varies spatially also depending on the prevailing atmospheric conditions. As an air mass can extend over a wide area, so all pixels contained within this area should be affected by the algorithm in a similar manner. Clearly, the observation error covariance is not spatially white, but spatially correlated.

Similarly, proximity to land has an effect on the uncertainty of an observation because land-originating radiation may contaminate the result. This was observed in figure 3.23. The presence of nearby land, which generally has a significantly different temperature from the water, contaminates the observed signal and introduces a spatially correlated error.

The most serious atmospheric concerns for this thesis are with cloud and land contamination. These were addressed when land, cloud masks were applied to remove contaminated pixels. While it might be more realistic to assign error to cells based on atmospheric conditions and the proximity to land, and use correlation length scales to assign covariances, this has not been done because neither the variable magnitudes of the error nor the covariance structure is known. While this would be an interesting area of research, it is left for further study. Instead, it is assumed that error checking has removed the erroneous pixels and that all observations are independent.

3.5 Chapter Summary

PPB is the case study site of this thesis, and was chosen because of the extensive data available. The analysis of the atmospheric data has identified those weather stations that record data more representative of PPB conditions. These maritime stations were subjected to further analysis, which found that with the possible exception of relative humidity the spatial variation of the data is not excessive. This finding justifies the use of spatially uniform atmospheric forcing to drive the numerical model.

Based on the analysis of this chapter the following meteorological stations have been selected for use as atmospheric forcing data, and are summarised in table 3.15.

Table 3.15: Summary of meteorological stations used to represent the atmospheric inputs in the numerical modelling.

Atmospheric Input	Meteorological Station
Atmospheric Pressure	Moorabbin
Wind Vectors	South Channel Island
Precipitation	Melbourne
Evaporation	Melbourne
Cloud cover	Melbourne
Air temperature	Point Wilson
Relative Humidity	Frankston

Atmospheric pressure is taken from Moorabbin — the lowest elevation site — with missing data in-filled from the other stations and adjusted for the elevation difference. With the largest average windspeed, wind data are taken from South Channel Island, although Point Wilson or Fawkner Beacon would have been equally suitable. Precipitation data are taken from Melbourne because it has an annual average precipitation between those of Moorabbin and Laverton and so is the most likely station to represent spatially average conditions. Evaporation and cloud cover data are taken from Melbourne: the only available site. As the satellite derived radiation is corrupted, incoming solar radiation is derived theoretically, using a formulae proposed by Zillman (1972). Air temperature is taken from Point Wilson because the South Channel Island record terminates earlier. Finally, the Frankston record is used for relative humidity as it is more reliable than the Point Wilson record.

The analysis suggested a separation of the data types according to their spatial variation characteristics. Lower bounded types, like precipitation, tended to have increasing variance with mean value, while unbounded types like air temperature have constant variance throughout their data range. These findings are used in the development of a technique for generating perturbed forcing data in chapter 4.

An examination of the cloud and land masking supplied with the SST observations indicated that further processing was necessary. The land mask was extended by 2 pixels (~ 2 km) and the cloud mask by 4 pixels. Furthermore, a sanity check was made and where the variation of the surrounding pixel was greater than 0.3°C the pixel was discarded. By contrasting the SST observations with in situ temperature measurements, the observation bias and error were assessed. A negligible

(-0.16°C) temporal bias was detected, although a significant diurnal bias was found which will need addressing. The average error of the SST observations is 0.75°C , which is comparable with the quoted accuracy of standard AVHRR SST retrieval algorithms.

Chapter 4

Model Testing and Ensemble Generation

This chapter presents the numerical model and ensemble generation techniques used in this thesis, and forms the foundation for the later experimental chapters 5 and 6. The first part of the chapter presents and tests the numerical hydrodynamic model used in this thesis, while the second part presents the data assimilation framework. Moreover, a method for ensemble generation and perturbed forcing data generation is developed.

The CSIRO Model for Estuaries and Coastal Oceans (MECO) is used for the hydrodynamic modelling. The model is configured based on the data of chapter 3. The predictions of the model are also assessed qualitatively for realistic results and quantitatively with the recorded sea level and temperature data. The uncertainty of the numerical model is estimated through a sensitivity analysis of its parameters, mixing schemes, heatflux modules, and forcing data. Finally, the nonlinear behaviour of MECO is studied because ensemble data assimilation schemes allow nonlinear models to be used, but become suboptimal with overly nonlinear behaviour.

Findings in the literature study of chapter 2 showed a lack of consensus on methods to initialise the ensemble, and many of these techniques are not suitable for the PPB case study. Therefore, a new approach to ensemble initialisation is presented, including an ensemble size recommendation based on system complexity. Furthermore, a method for incorporating forecast error through perturbed forcing data is

developed. A comparison of the use of perturbed forcing data against a stochastic version of the hydrodynamic model demonstrates the validity of the approach.

4.1 Hydrodynamic Modelling

The hydrodynamic modelling in this thesis is undertaken using the CSIRO Model of Estuaries and Coastal Oceans (MECO). MECO is a finite difference model that solves the primitive equations using standard numerical techniques (Walker *et al.*, 2002; Herzfeld *et al.*, 2002). MECO is similar to freely available models such as Princeton Ocean Model (POM; Blumberg and Mellor, 1987), the Regional Oceanic Modelling System (ROMS; Shchepetkin and McWilliams, 2005), and the Modular Ocean Model (MOM; Pacanowski and Griffies, 1988).

Specifically, MECO solves the primitive equations of momentum, continuity and the conservation of heat and salt on z-levels with an orthogonal curvilinear Arakawa C grid. These equations are simplified by the hydrostatic and Boussinesq assumptions with turbulence closure met using the Csanady parameterisation. Mode splitting of the 2 and 3-dimensional modes allows the separation of fast moving (surface) gravity waves, from slower moving internal waves. The model uses explicit numerical schemes throughout, except for the vertical diffusion scheme which is calculated implicitly. The time stepping uses an Euler forward time scheme and a van Leer advection scheme.

The atmospheric heat flux is applied based on the bulk parameterisation formulae of Gill (1982). The bulk formulae account for longwave radiation, latent heat, and sensible heat. Incoming solar radiation is provided as a time series through the atmospheric forcing data, and is distributed through the water column according to the attenuation coefficient.¹ The net heatflux at the surface is introduced by adjusting the temperature of the surface layer. For more details on MECO or the numerical techniques used the interested reader is referred to the above-mentioned guides.

The first step in the hydrodynamic modelling was to develop a working model. The model of PPB used in this thesis is based on the CSIRO modelling of PPB for the

¹The attenuation coefficient denotes the rate at which light is absorbed by the water column. It has units of $[m^{-1}]$.

PPBES. Their original parameter files for the PPBES model were supplied and used as a basis. However, as MECO has been modified since the mid 1990's, alterations to the configuration of the model and the input files were made. This chapter describes the model configuration used throughout this thesis. The parameter set used for this thesis was based on the values used in the PPBES. The model parameters and their values adopted are listed in appendix A.

4.1.1 Numerical Grid and Bathymetry

The AVHRR satellite data are supplied on a 0.01 degree resolution grid. To simplify the assimilation process, PPB was modelled at the same horizontal resolution, avoiding the need to aggregate or disaggregate the satellite data, although (dis)aggregation is possible through data assimilation. Vertically, 14 layers were used with higher resolution in the upper layers. A 2-metre resolution was used from the surface to a depth of 20 metres, thence layer boundaries were located at 24, 32, 48, and 64 metre depths.

The decision to use a rectilinear, rather than a curvilinear grid, represents a tradeoff between simplifying the assimilation and retaining an accurate model, which represents the physical water movement well. Using a rectilinear grid reduces the accuracy around the entrance of PPB as the complex bathymetry and coastline of this region cannot be adequately resolved, but the water movement in the remainder of PPB should remain reasonably accurate. The grid adopted is illustrated in figure 4.1 with the bathymetry used also shown.

The bathymetry used is based on supplied data, but was fine tuned in the region around The Heads and the Great Sands. In this region deep narrow channels intersect shallow sandy extents. The numerical grid is too coarse to resolve these adequately, so wider, shallower channels than actually exist are included; a simplification attempting to balance the flow dynamics.

4.1.2 Open Boundary

The open boundary joins the southern extent of the model domain to Bass Strait. At this boundary sea level, temperature, and salinity are prescribed using clamped boundary conditions. The most obvious choice for prescribed sea level is the tide

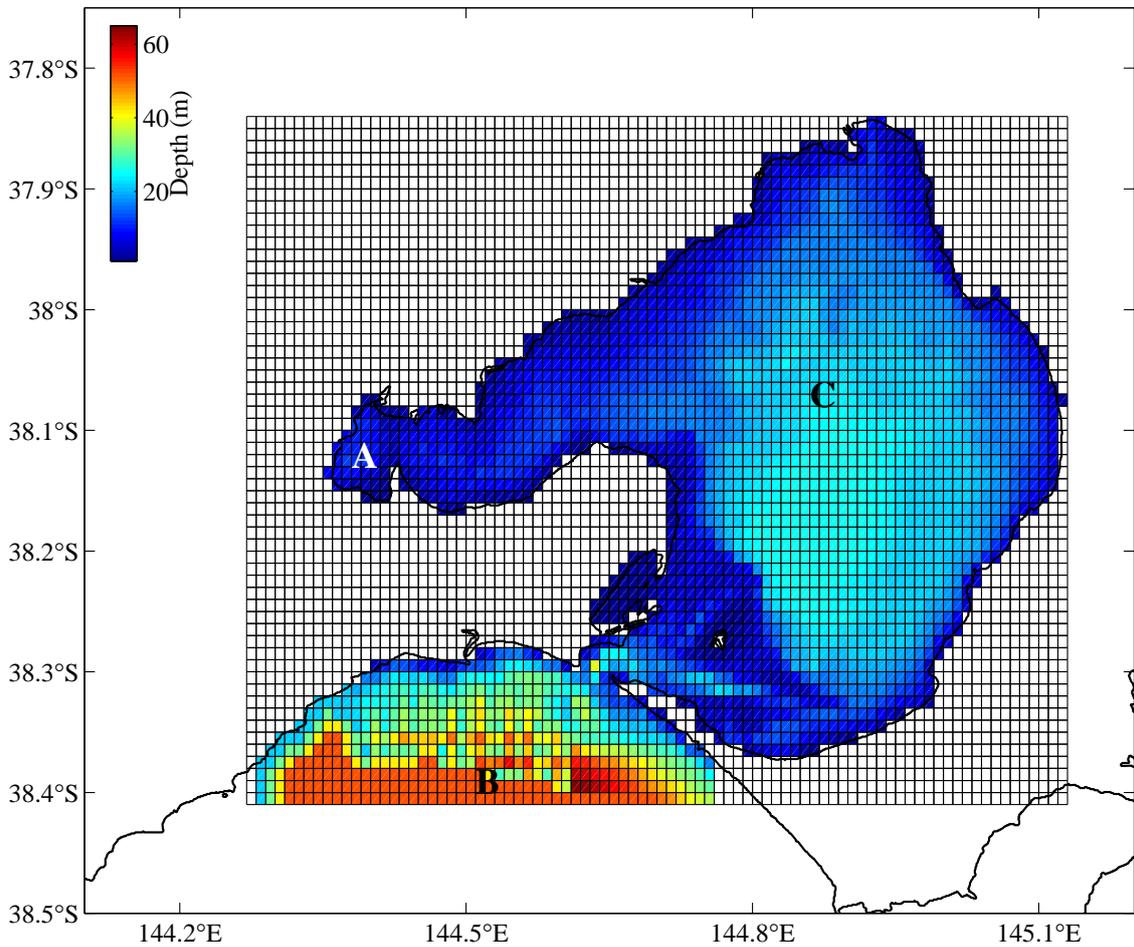


Figure 4.1: Outline of numerical grid, model domain, and bathymetry adopted. The open boundary is located along the southern edge of the domain. Letters A, B, and C indicate locations mentioned later in the text.

gauge record at Lorne (figure 3.2). This is complete except for November and December of 2001, where missing data was in-filled with tidal harmonics. The high frequency noise in the tidal signal could generate parasitic high frequency waves in the model, and was therefore removed by filtering the in-filled Lorne data through a low pass filter.

No measurement of Bass Strait temperature or salinity was available during the modelling period. Instead, following the approach taken in the PPBES, data recorded near The Heads were used to generate a climatology that was used for the temperature boundary condition. This open boundary climatology approach was

also followed by Kelley *et al.* (2002) in oceanic data assimilation. Temperature at the open boundary ranges from just less than 13°C to 18.5°C. A constant salinity of 35.4 PSU was used in the PPBES modelling and has also been adopted here.

4.1.3 Riverine Flows

Where rivers enter PPB they add additional mass (water) to the model. For the continuity of heat and salt, the temperature and salinity of the incoming water needs to be specified as well as the flow.

The PPBES included input from the Yarra River, Mordialloc Creek, Patterson River, Werribee River, Kororoit Creek and Elwood Canal (figure 3.1). Of these, flow data was only available for the Yarra River and Werribee River. Missing data in Kororoit Creek and Elwood Canal had been in-filled in the PPBES with values of 0.2 m³s⁻¹ and 0.18 m³s⁻¹ respectively. This data in-filling was continued for this thesis. The PPBES developed a regression relationship between flow in the Patterson River and flow in the Yarra River, and this relationship was also reused. In the PPBES, Mordialloc Creek flow data was in-filled with a constant value of 2.24 m³s⁻¹. This value was generated using data from the years 1988 and 1994, which are years of higher than average flow (figure 3.13). During 2003, the value of 2.24 m³s⁻¹ was at times about the same as the flow in the Yarra River, a much bigger river. Clearly this value overestimated the contribution of Mordialloc Creek and so it was removed. While the removal of Mordialloc Creek underestimated its contribution to the flow entering PPB, the in-filled values at Kororoit Creek, Elwood Canal and Patterson River probably overestimated their contribution and so a balance was achieved. The flows at the WTP outlets were also included in the model.

The variation in riverine temperature was modelled as

$$T = a \sin \left(\frac{2\pi t}{365} - b \right) + c, \quad (4.1)$$

where t is the Julian day. The coefficients a , b , c have values of 6.5, 74, and 15.3 respectively. These were derived by fitting the formula to the temperature observations in the Yarra River, Merri Creek and Gardiners Creek. This formula for

temperature was applied to all rivers and the WTP inputs. All sources of riverine water are assumed fresh, and so salinity was taken as zero.

4.1.4 Atmospheric Inputs

MECO allows atmospheric inputs to be specified as either spatially varying or spatially uniform data. Spatially varying data are important where influential variables are found to change considerably spatially. The analysis from chapter 3 suggests that, for the relatively small region of PPB, atmospheric conditions do not have significant spatial variation. An exception might be relative humidity, however for this variable a lack of sufficient data would prevent the preparation of a spatially varying data set. Therefore, throughout this thesis atmospheric forcing was applied in a spatially uniform manner. The meteorological stations used for the different atmospheric data inputs were determined in chapter 3. These were summarised in table 3.15. As the satellite derived radiation was corrupted, incoming solar radiation data were derived theoretically, using the formulae of Zillman (1972).

4.2 Model Testing

The hydrodynamic model of PPB was initially tested by comparing the results of a long model run, from January 2000 through to June 2003, against knowledge of the physical behaviour of PPB and observed data. The temperature was initialised uniformly to 18.16°C — the temperature at West Channel Pier on the 1st January 2000 — the salinity was initialised to 35 PSU, and components of water movement, u_1 , u_2 , and w , together with sea level are initialised to zero. As this model run is made over multiple years no specific model spin up is made. Currents and sea level should settle within a day or so, and temperature within a few days to weeks. Due to the low exchange through The Heads, salinity might take a year or more to settle. However, no data are available to confirm the salinity distribution and furthermore, it is not the focus of this thesis.

4.2.1 Currents

Two small residual currents were found in PPB, illustrated in figure 4.2. These currents flow south, generated by the net sea level gradient from the riverine inflows. The first flows along the eastern boundary of PPB, while the second flows from the north-western shore of PPB, south across the Geelong Arm and down to The Heads. The eastern current is well known, described in the PPBES as hugging the eastern coast under the influence of Coriolis. No mean current is observed either in the centre of PPB or in the Geelong arm, where tidal currents are equal and opposite for both the flood and ebb tides. At The Heads and the Great Sands, the residual currents are more confused owing to the complicated bathymetry and coarse resolution. For instance, the large westerly current apparent at The Heads is an artifact of the resultant direction as the water bends around the headland to enter and leave PPB.

A calculation of the geostrophic currents (not shown) indicates that the ageostrophic currents are of similar magnitude to the geostrophic currents. This implies

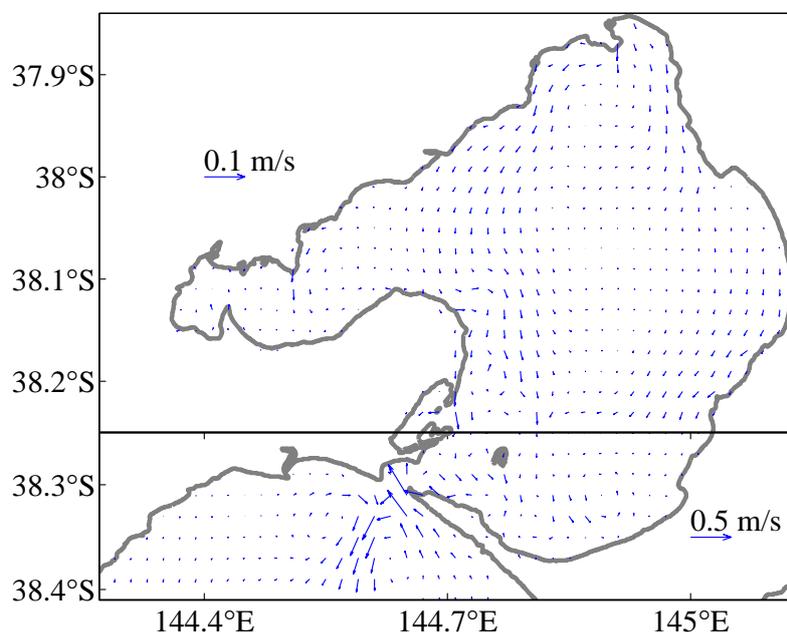


Figure 4.2: Plot of the residual surface currents based on modelled currents from January 2000 to June 2003. Note a different scale exists for the magnitude of the current vectors above and below 38.25°S also not all grid points have been plotted for clarity.

that the temperature is not in geostrophic balance with the currents and that there may be a low correlation between the variables.

4.2.2 Sea level

After removing the tidal variation any residual variation of the mean sea level is due to sources of water entering PPB, creating a sea level gradient. This is replicated by the model (figure 4.3a) with the highest mean sea level being found at the source of the largest fresh water inputs to PPB: the Yarra River and the Western Treatment Plant. Furthermore, the mean sea level is seen to be depressed at locations where sea level, locally constricted by the bathymetry and coastline, banks up on either side of the constriction to generate the elevation gradient necessary to overcome the constriction. This is most evident at The Heads, where the mean sea level is always lower than that in either Bass Strait or PPB but is also seen on a smaller scale at the entrance to Corio Bay (figure 3.1), where a shallow ridge across the Geelong Arm restricts the flow. The loss of energy as the tide passes through The Heads

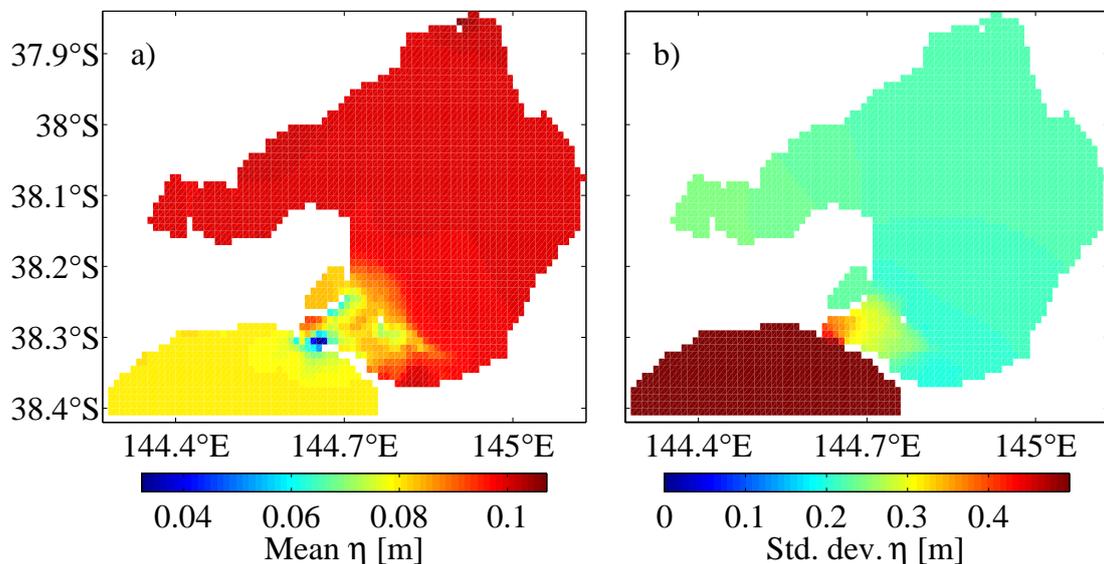


Figure 4.3: Plots of the temporal mean sea level (η) a), and the temporal standard deviation of sea level b) based on modelled sea level from January 2000 to June 2003.

Table 4.1: RMSE and mean error between observed and modelled sea level at various tidal gauges between 2000 and 2003.

Tidal Gauge	RMSE [m]	Mean Error [m]	Amplitude[†] [m]
West Channel Pile	0.114	0.107	-
Breakwater Pier	0.075	0.037	0.455
Geelong	0.093	0.073	0.501
Hovel Pile	0.043	-0.014	-
Point Lonsdale	0.184	0.138	0.818
Queenscliff	0.138	0.070	-

†Indicative amplitude based on summation of amplitudes of the M2, S2, K1, and O1 harmonic tidal components where available. Source: ANTT. Note as these components have different periods and phases, actual observed amplitudes are generally less than those indicated here.

and the Great Sands is reproduced in the model by the reduction of the sea level standard deviation moving from Bass Strait to PPB (figure 4.3b). Evidence for the northern basin operating as a single unit is found by the standard deviation of sea level in the north of PPB being spatially uniform.

A quantitative assessment of the sea level is presented through a series of comparisons with in situ observations. Due to the coarse resolution of the model, relative to the bathymetry and coastline at the entrance of PPB, it is unrealistic to expect an accurate prediction of sea level around The Heads. The RMSE and the mean error calculated at the tide gauge sites are given in table 4.1.

As expected, large errors in predicted sea level, up to 0.18 m RMSE, are found at Point Lonsdale and Queenscliff, while the prediction at the edge of the Great Sands is more accurate with 0.11 m and 0.04 m RMSE at West Channel Pile and Hovel Pile respectively. In the north of PPB the RMSE is 0.07 m at Breakwater Pier. The average mean error which gives an indication of the prediction bias is 0.07 m. This is reasonable considering the coarse resolution at The Heads, and that measurement error, based on the high frequency noise in the recorded data, is of the order of a few centimetres. Furthermore, the error is significantly less than the indicative tidal amplitudes.

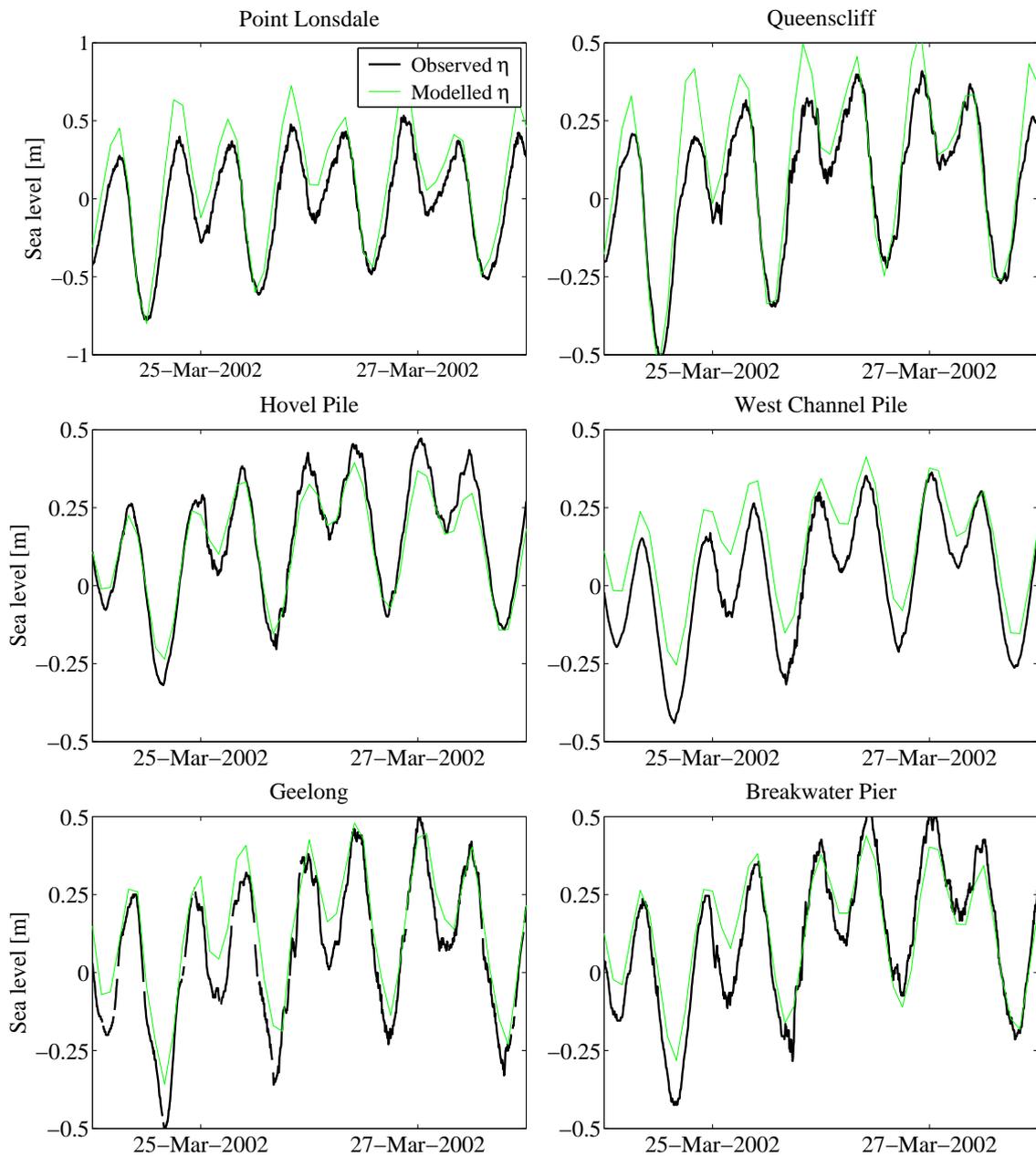


Figure 4.4: Observed and modelled sea level at various tidal gauge sites. Note the different y-axis scale for Point Lonsdale.

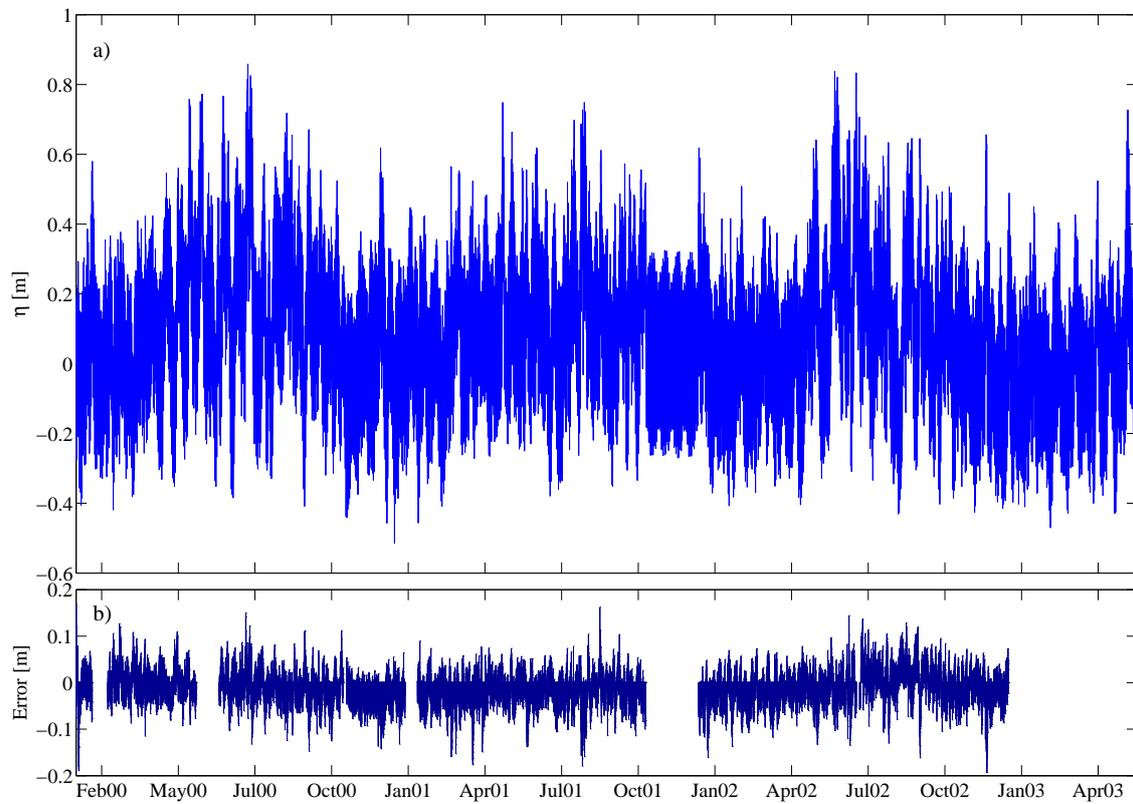


Figure 4.5: Plot of a) modelled sea level (η) at Hovel Pile, and b) prediction error.

A 4-day time series of the observed and modelled sea level are shown in figure 4.4. The phases of the tidal signal are well represented by the model; but for the period shown the amplitudes tend to be underestimated. A long-term prediction of sea level at Hovel Pile is shown in figure 4.5, and indicates that the long-term prediction error appears constant over the time period.

4.2.3 Temperature

Temperature inside PPB varies more temporally than the temperature in Bass Strait. The model replicates this with a larger temporal standard deviation of temperature inside PPB than outside (figure 4.6). Furthermore, the temperature of PPB varies more at the edges than in the centre, especially in Corio Bay. These results support the hypothesis that the temperature variation in PPB is dominated by atmospheric transfer rather than by advection of water from Bass Strait.

The predictions of temperature are compared with in situ observations in figure 4.7. Although, the model reproduced the trend and diurnal signal, temperature is overestimated. The West Channel Pile panel shows that the overestimation grows in time, indicating a prediction bias. This temperature bias is due to errors in the model, either the heatflux module or the parameters. While atmospheric data contains variations, they are assumed to be unbiased, implicating model bias as the source.

In spite of the overestimation, the diurnal temperature variation through the water column is consistent with predictions as Central (a deeper location) shows less variation than Longreef (a shallower location). Also, the seasonal variation of temperature predicted by the model seems to have the correct amplitude of about

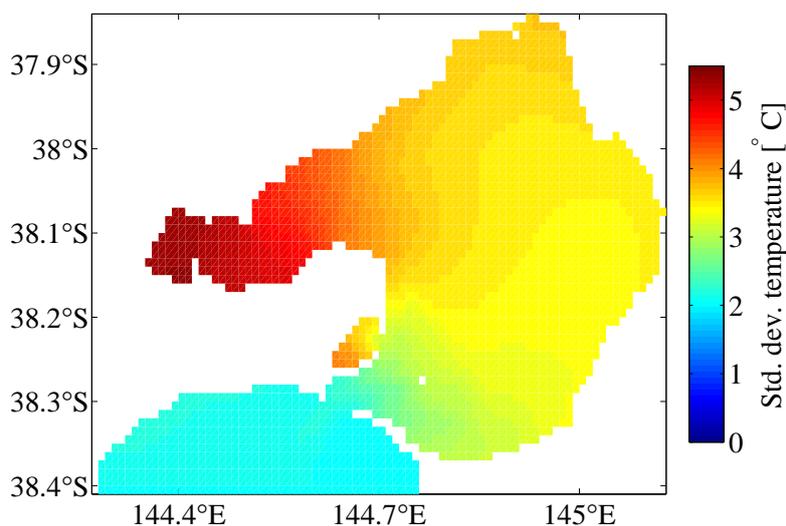


Figure 4.6: Temporal standard deviation of surface temperature based on modelled temperature from January 2000 to June 2003.

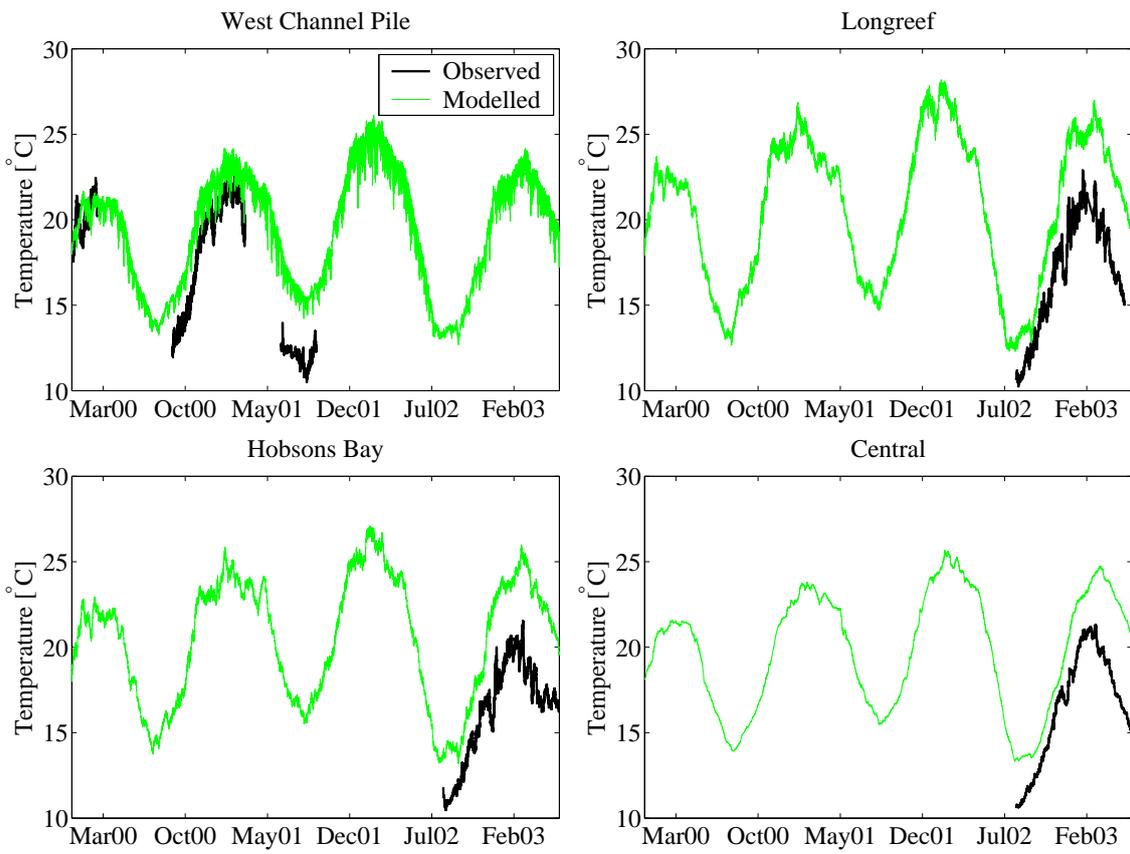


Figure 4.7: Observed and modelled temperature at the deep monitoring sites listed in table 3.12.

Table 4.2: The RMSE between observed and modelled temperature at various monitoring sites between 2000 and 2003.

Monitoring Site	RMSE [$^{\circ}\text{C}$]	Mean error [$^{\circ}\text{C}$]
West Channel Pile	2.72	2.22
Central	3.51	3.17
Hobsons Bay	4.33	4.08
Longreef	4.27	3.95

10 $^{\circ}\text{C}$, although it varies between 15 $^{\circ}\text{C}$ and 25 $^{\circ}\text{C}$ rather than 11 $^{\circ}\text{C}$ and 21 $^{\circ}\text{C}$.

The RMSE and mean error in temperature prediction is summarised in table 4.2. The RMSE at West Channel Pile is smaller because its observed record occurs during the first half of the model run and so the bias had less time to effect the prediction. Central is slightly more accurate than Longreef or Hobsons Bay, possibly because of the stability of the temperature at depth. The mean error is similar to the RMSE indicating the prediction error is mainly due to bias.

4.2.4 Model Testing Summary

Qualitatively, the model replicates the main features of the physical behaviour of PPB well. Specifically,

- greater temperature variations are observed nearer the coast and in PPB than in Bass Strait,
- residual currents flow south through PPB driven by riverine inputs,
- sea level energy is lost through the Heads and over the Great Sands, and
- the northern basin is found to operate as a single unit.

The sea level is well predicted relative to wave amplitude and observation error. The prediction is less accurate around The Heads due to the complex coastline and bathymetry, however improvements of sea level prediction are limited by the model resolution and forcing data.

Temperature prediction is biased toward warmer temperature with predictions and observations diverging over time, although the seasonal range and diurnal fea-

tures appear correct. While it might be possible to improve the temperature prediction through parameter tuning, the problem is understood to be connected to the bulk parameterisation in the heatflux module not adequately characterising the processes. This finding of model bias is significant, as one of the assumptions of the data assimilation techniques, equation (2.5), is that the model predictions are unbiased. The importance of this will be explored in chapters 5 and 6. Nevertheless, the error in the model prediction of temperature is large relative to the SST observation error (table 3.14) suggesting that the application of data assimilation techniques will yield significant improvements to the forecasting of temperature.

Another finding of the model testing is further confirmation of the domination of atmospheric heat transfer as the controlling temperature process and the de-coupling of temperature from the other variables. The evidence for this is

- the small residual currents in the northern basin, figure 4.2,
- the larger variation of temperature within PPB than outside, figure 4.6,
- the temporal mean sea level showing the local constriction of flow entering Corio Bay and The Heads, and
- the similar magnitudes of the geostrophic and ageostrophic currents.

The significance of the de-coupling of temperature is that it may allow temperature to be assimilated independently of currents, salinity, and sea level, which would allow for fewer ensemble members and more efficient assimilation. In general, this is not possible. In the open ocean temperature, currents, and salinity, are in a geostrophic balance and should be jointly adjusted to preserve this balance. The model testing has shown that the model adequately represents the main features of PPB, and that although temperature is biased, it is suitable for use in a data assimilation study.

4.3 Sensitivity Analysis

A sensitivity analysis determines those parts of a model that influence predictions to the greatest extent. These parts of the model are also likely to be most sensitive

to errors and uncertainty, as for a similar level of error they have the greatest leverage. Therefore, for a data assimilation study, those components of the model that dominate the sensitivity should be used to guide the incorporation of forecast error. The sensitivity analysis is undertaken here in three parts, considering in turn the influence of model parameters, mixing schemes, and temperature modules and forcing data.

Grid resolution and bathymetry influence the uncertainty of the model prediction but are not considered here. These model properties impact more significantly on currents and sea level than on temperature, and due to the apparent de-coupling of temperature from flow their impact is expected to be minor.

The sensitivity analysis was performed by altering the model configuration and calculating the difference between the sensitivity runs and the long model run used in section 4.2, here termed the control run. The comparison was made over the month of January 2003. The control run states on 1 January are used as the initial conditions, and so no model spin-up is required. The model states of temperature and sea level were used as the basis of comparison: temperature, because it is the variable of interest for the assimilation, and sea level, because it acts in unison with currents indicating changes in the hydrodynamics.

The sensitivity in temperature and sea level was assessed by calculating the RMSE between the control and the sensitivity run at the four MAFRI monitoring sites. As three of these stations have both deep and shallow records, seven values were used to calculate the average RMSE.

4.3.1 Parameter Settings

Six sets of model parameters were identified for sensitivity testing. The sensitivity tests were performed by increasing and decreasing the parameter values by 10%, refer to appendix A. The parameter sets were wind stress (wind drag coefficients), albedo, attenuation of incoming solar radiation, bottom friction (minimum drag coefficient, background friction velocity and bottom roughness length), the Csanady mixing parameters (background viscosity, background diffusivity and their representative constants), and horizontal mixing (horizontal diffusivity and horizontal viscosity). Where there were two or more individual parameters within a parameter set, they

Table 4.3: Average RMSE of temperature (T) and sea level (η) generated by a 10% increase or decrease in the model parameter values for different parameter sets. Values in boldface indicate the most significant parameter for that variable.

No.	Description	RMSE T [$^{\circ}\text{C}$]		RMSE η [m]	
		+10%	-10%	+10%	-10%
1	Wind stress	0.02	0.02	0.001	0.001
2	Albedo	0.22	0.22	0.001	0.001
3	Attenuation	0.02	0.03	0.001	0.001
4	Bottom friction	0.02	0.02	0.004	0.004
5	Csanady mixing	0.01	0.00	0.001	0.001
6	Horizontal mixing	0.02	0.02	0.001	0.001

were jointly adjusted so that the adjustment reinforces the sensitivity. For instance, when friction coefficients were tested, both were altered in the direction of increasing friction for one run, and in the direction of decreasing friction for the other run.

The results of the six experiments are presented in table 4.3. Of the parameters, only albedo significantly influenced temperature, with an average RMSE of 0.22°C . Albedo affects the amount of incoming solar radiation entering the water column, and so directly influences temperature. The other parameters only affected the distribution of the temperature within the water column either directly in the case of attenuation, or indirectly in the case of the others, and so had less effect on the temperature.

The most influential parameter set for sea level were the bottom friction parameters. These affect the dissipation of momentum, dampening velocity and consequently altering sea level. But the variation was inconsequential with an average RMSE of 0.004 m.

4.3.2 Mixing Schemes

The sensitivity of the model to the mixing scheme was made by re-compiling MECO with two alternative mixing schemes, while leaving the rest of the model unchanged. The alternative mixing schemes applied were the $k-\epsilon$ (Burchard *et al.*, 1998) and the Mellor-Yamada 2 (Mellor and Yamada, 1982) schemes.

The mixing scheme determines the value of the vertical viscosity and diffusivity terms in the hydrodynamic momentum equations. The Csanady scheme, used in

Table 4.4: Average RMSE of temperature (T) and sea level (η) generated by using an alternative mixing scheme.

No.	Mixing Scheme	RMSE T [$^{\circ}\text{C}$]	RMSE η [m]
7	k- ϵ	0.22	0.017
8	Mellor-Yamada 2	0.19	0.028

MECO, calculates values for these terms by estimating the change in them from an assumed background level. The Mellor-Yamada 2 and k- ϵ schemes predict viscosity and diffusivity more rigorously by relating them to the amount of turbulent kinetic energy and a turbulent length scale. In the Mellor-Yamada 2 scheme the length scale is parameterised analytically, while in the k- ϵ scheme the length scale is calculated from the dissipation rate of the turbulent kinetic energy. The k- ϵ scheme is the most advanced, although computationally expensive, whereas the Csanady scheme is faster but less accurate. The Mellor-Yamada 2 scheme lies between the two in terms of speed and accuracy. For further details see Herzfeld *et al.* (2002).

The results, presented in table 4.4, suggested that the choice of mixing scheme was generally more influential than the parameter settings (table 4.3). The k- ϵ scheme had the same impact on temperature as albedo, while Mellor-Yamada 2 had slightly less. Both mixing schemes were an order of magnitude more significant than the other parameters.

The choice of mixing scheme affects water column mixing which in turn affects the temperature distribution. Similarly mixing also dissipates the momentum and in this way reduces velocity. The effect on sea level was obvious; the model was an order of magnitude more sensitive to changes in mixing schemes than the parameter settings.

4.3.3 Heatflux Modules and Model Forcing

The final set of sensitivity experiments considered the influence of the heatflux module and the forcing data on the model's predictions. The heatflux module of the control was based on the bulk parameterisation of Gill (1982). To test its sensitivity, the heatflux module was replaced by the bulk parameterisations of Zillman (1972) and MECO was re-compiled. The key difference between the two heatflux modules

Table 4.5: Changes to data sources used to assess the influence of forcing on model predictions.

Data Type	Control Run Station	Sensitivity Run Station
Air Temperature	Point Wilson	Frankston
Pressure	Moorabbin	Laverton
Rainfall	Melbourne	Moorabbin
Wind	South Channel Island	Fawkner Beacon
Relative Humidity	Frankston	Point Wilson

Table 4.6: Average RMSE of temperature (T) and sea level (η) generated by using an alternative heatflux module and by using different forcing data.

No.	Description	RMSE T [$^{\circ}\text{C}$]	RMSE η [m]
9	Altered heatflux module	0.71	0.006
10	Different forcing data	0.34	0.002

is that the Zillman module has a more explicit calculation of longwave radiation and the latent heat flux, and the net surface energy flux is introduced as a surface boundary condition, whereas in the Gill module the flux is introduced by changing the temperature of the surface layer directly. Again for further details see Herzfeld *et al.* (2002).

The sensitivity of the prediction to the forcing data was analysed by running the model with a different set of forcing data, taken from different locations. As both sets of data are collected around PPB they are equally valid as forcing data and the differences in results are attributable to errors (spatial and measured) in the data. The changes to the forcing data are summarised in table 4.5.

The results to changes in the heatflux module and the forcing data are presented in table 4.6. Temperature was especially sensitive to changes in the heatflux module. This was expected as changing the heatflux modules leads to different rates of heating or cooling, which impacts on the predictions. The model was also quite sensitive to changing forcing data. In fact, as neither cloud cover nor incoming solar radiation forcing data were altered, both influential inputs for temperature, the average RMSE value of 0.34°C probably underestimated the sensitivity of the model to changes in forcing data. Neither altering the heatflux module nor the forcing data had a significant influence on water level.

4.3.4 Implications for data assimilation

The sensitivity analysis found that in terms of sea level only changes to the mixing schemes had a significant impact on the model. With the exception of albedo, the temperature of the model was not particularly sensitive to the choice of model parameters. It was more sensitive to changes in mixing schemes, the heatflux modules, and forcing data. Furthermore, given the probable underestimation of the model sensitivity to changes in forcing data, the sensitivity of the model to forcing data changes was of the same order as the model sensitivity to changes in the model structure.

This means that for short term (~ 1 month) forecasting, forcing data errors are at least as significant as model errors. The implication of this for coastal data assimilation is that consideration of forcing data errors and their incorporation is important for accurate forecast error estimation. This is further emphasised by the finding in chapter 2 that forcing errors are generally not dealt with rigorously, and highlights the need to develop the generation of perturbed forcing data.

Finally, the sensitivity errors for temperature are of the order of 0.5°C and model testing found a prediction RMSE of $3\text{-}4^{\circ}\text{C}$ (larger than due to bias). In comparison the sensitivity for sea level was 0.005 m and model testing found a prediction RMSE of the order of 0.10 m. The sensitivity of sea level was much less than the temperature, relative to their respective prediction RMSEs. This suggests that prediction error in sea level is due to elements not tested in the sensitivity analysis, such as bathymetry or the open boundary conditions.

4.4 Model Linearity and Data Assimilation

Ensemble data assimilation techniques extend the linear assimilation techniques, like the Kalman filter, for use in nonlinear models. While the ensemble techniques allow the use of nonlinear models, the models cannot be highly nonlinear without violating the analysis assumptions. Nonlinear propagation of the model states and forecast error is possible, via equation (2.22), but the analysis, equation (2.20), is still based on an assumption of a Gaussian error distribution (Barham and Humphries, 1970). When a numerical model is highly nonlinear, the forecast error distribution may not

be Gaussian, and the resulting analysis will not be optimal.

An example of this is to consider what happens when a bifurcation occurs in a strongly nonlinear ensemble setting. The ensemble members will be divided, following diverging trajectories of the bifurcation. The resulting ensemble mean will exist between the two clusters in a physically unrealistic location. The ensemble spread would be excessive, and the ensemble distribution bi-modal, rather than uni-modal and normally distributed. With luck, accurate observations would direct the model to the correct mode but this could not be assured, nor could an optimal analysis be expected.

Nonlinear behaviour of a weakly nonlinear model can be seen as time scale dependent. Within a certain time frame the model may be quasi-linear, but after a longer time has elapsed the behaviour may be characterised as nonlinear. In practical terms, a study of the nonlinear behaviour of a weakly nonlinear model involves identifying the maximum time-window within which the behaviour of the model is still quasi-linear.

This section investigates the nonlinear behaviour of MECO, especially in regard to temperature with the aim of establishing its time scale of quasi-linearity. If the observations are generally less frequent than this time scale, the optimality of the data assimilation analysis should be questioned, and the suitability of ensemble filter methods discussed.

4.4.1 Testing Linearity

One test to determine linearity is to perturb the initial states of a model in equally opposite directions, and to run both the positive and negative perturbed model configurations as well as an unperturbed control. If the magnitude of the perturbations relative to the control remain equal over time the model is behaving linearly. If the magnitude of the perturbations vary, the model is behaving in a nonlinear fashion. The size of the initial perturbation also affects the linearity. With a small initial perturbation a model may remain quasi-linear longer than with a larger perturbation.

This test is the basis for a linearity statistic, which following Daescu and Carmichael

(2003) is given by

$$\theta(\delta, t) = \frac{\|\delta^+(t) + \delta^-(t)\|}{0.5(\|\delta^+(t)\| + \|\delta^-(t)\|)}, \quad (4.2)$$

where δ is the initial perturbation vector, t is the time elapsed after the start of the model run at t_0 , the change in perturbations overtime is given by $\delta^+(t)$ and $\delta^-(t)$ where the superscript signifies the addition or subtraction of the perturbations from the original state, also $\|\cdot\|$ signifies the vector norm. The changes in the perturbations are given by

$$\delta^+(t) = f(\mathbf{x} + \delta, t) - f(\mathbf{x}, t), \quad (4.3)$$

$$\delta^-(t) = f(\mathbf{x} - \delta, t) - f(\mathbf{x}, t). \quad (4.4)$$

When the model is quasi-linear the magnitudes of $\delta^+(t)$ and $\delta^-(t)$ will be similar, giving a θ value close to zero. When the difference between $\delta^+(t)$ and $\delta^-(t)$, is the same order as their magnitude, θ approaches one, indicating nonlinear behaviour.

The factors influencing the statistic are the length of time after the addition of the perturbation, and the magnitude of the perturbation. The time scale of the numerical model linearity is assessed by considering a time series of θ at a range of perturbation magnitudes.

4.4.2 Linearity of MECO

The timescale of quasi-linear behaviour of MECO is assessed by adding and removing a smooth random perturbation at a range of standard deviations to (and from) the temperature fields, and then calculating the timeseries of the θ statistic. Perturbations could be added to the other variables, but temperature is the variable of focus in this thesis.

Temperature is typically spatially correlated so correlated random fields are generated for the perturbations rather than uncorrelated random fields. The use of correlated random fields for adding perturbations to models was discussed in the literature review in sections 2.4.2 and 2.4.3. Thus, as well as testing the linearity of the model, this experiment tests the behaviour of the model to the addition of correlated random fields that are used for ensemble generation and incorporating model error into the forecast.

Initially, linearity was tested by varying the magnitude of the standard deviation of a 3-dimensional perturbation field. However, instead of the expected gradual transition from linear to nonlinear behaviour, nonlinear behaviour was observed immediately. The cause of this was the use of the 3-dimensional correlated random field.

As the field is generated randomly physically unrealistic fields can be formed. For instance, a physically unrealistic temperature field with warmer water underlying cooler water may be generated when larger perturbations are created below smaller perturbations. The resulting vertical density gradient is unstable and through buoyant forces causes overturning and mixing of the water, destroying the linearity of the model. Adding the opposite field to the states strengthens any stratification, opposing mixing which further enhances the nonlinearity. If the water were strongly stratified a slight negative vertical gradient in temperature perturbations might be tolerated. However, as the PPB water column is generally well mixed, any imposed negative gradient induces overturning.

The problem of using 3-dimensional correlated fields was remedied by repeating the linearity test experiments using a 2-dimensional correlated random field horizontally, assigning each cell vertically through the water column the same perturbed value. In this case, the results, shown in figure 4.8, were much more reasonable. The linearity of the model decreased with perturbation magnitude, but for small perturbations the quasi-linear behaviour of the model remained for at least 2–3 days.

Over the time scale of a few days the numerical model of PPB behaves in a quasi-linear fashion. The NOAA satellites collect SST data approximately every 12 hours so that there is confidence that the assimilation can successfully be made. Of course gaps between observations may be longer, due to cloudy conditions, and although the analysis may still give realistic results, in such cases optimality might be questionable.

The implications of this linearity testing for data assimilation are

1. that suboptimality due to excessive nonlinear behaviour will not impact the PPB case study, and
2. that using 3-dimensional correlated random fields induces unstable behaviour,

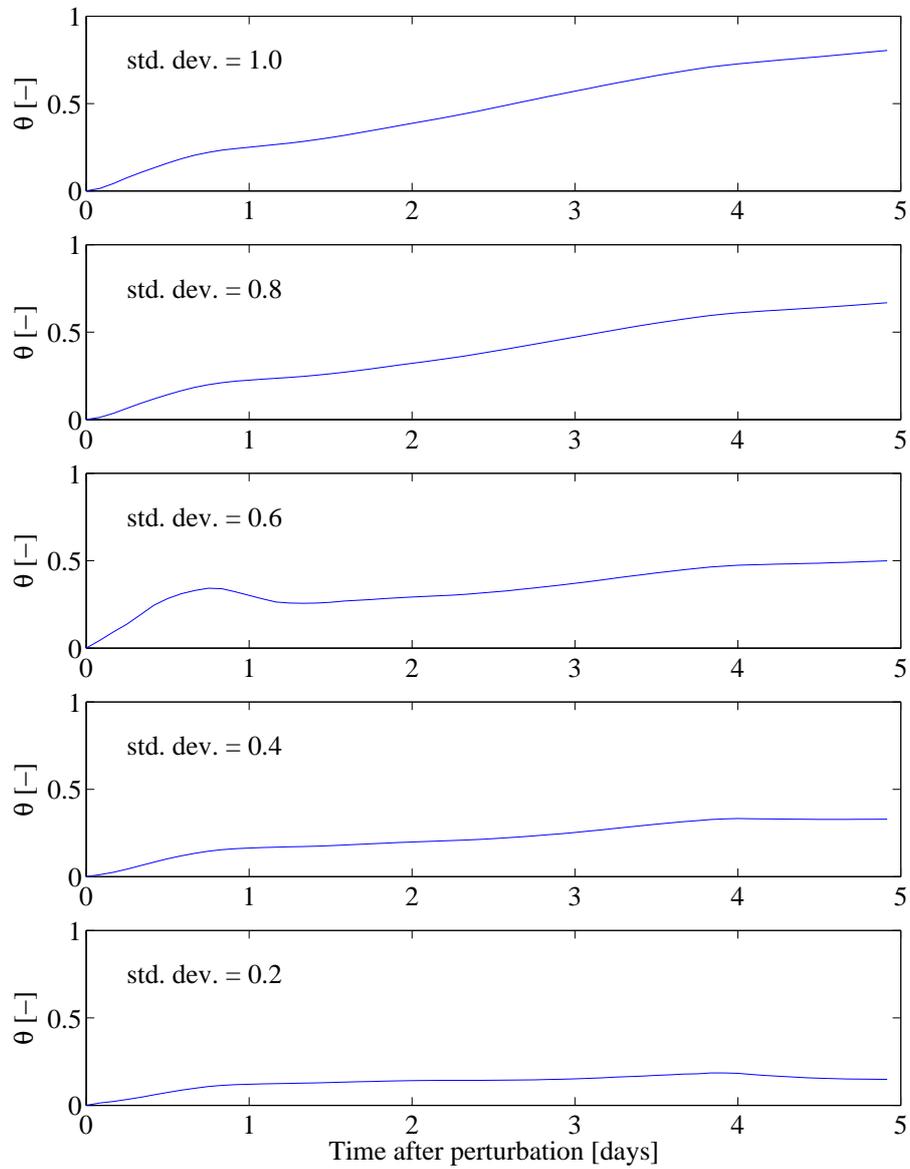


Figure 4.8: Linearity test using 2-dimensional perturbation field. The standard deviation value indicates the magnitude of the perturbation field [$^{\circ}\text{C}$]. Results are based on temperature states.

which impact on their use to incorporate forecast error into the ensemble.

Although, the use of 3-dimensional correlated random fields to incorporate model error is a popular approach, this finding suggests that alternative approaches for incorporating forecast error should be considered. This is discussed further in the following sections of this chapter.

4.5 Ensemble Generation

This section covers the generation of a set of perturbations used to initialise the data assimilation ensemble. The initial ensemble needs some variation that represents the *a priori* uncertainty of the initial value of the model state. The literature review identified four methods for initiating ensembles. However, none of these methods are deemed suitable for the PPB model, as outlined below, and a new approach to ensemble generation is proposed based on anomaly independence. Thus this section discusses

1. the importance of ensemble anomaly independence,
2. the reasons that existing ensemble initialisation methods are unsuitable, and
3. the method developed for this thesis.

Physical realism, as the previous section highlights, is important for retaining quasi-linear model behaviour. While the need for ensemble spread is clear, the importance of the rank of the ensemble is not so obvious.

4.5.1 Ensemble Independence

In addition to the ensemble spread, it is important that the ensemble anomalies have a high rank². This allows for smaller ensemble sizes, and makes ensemble techniques more efficient. This can be seen by considering the underlying EnKF equations. As Evensen (2003) has shown, the EnKF analysis equation (2.20) can be written as

$$\mathbf{X}^a = \mathbf{X}^f + \mathbf{X}'^f \mathbf{\Lambda}, \quad (4.5)$$

²The matrix rank indicates the number of linearly independent rows or columns of a matrix.

where

$$\Lambda = \mathbf{X}'^f \mathbf{H}^T [\mathbf{H} \mathbf{X}'^f \mathbf{X}'^f \mathbf{H}^T + (n_e - 1) \mathbf{R}]^{-1} [\mathbf{D} - \mathbf{H} \mathbf{X}^f] \quad (4.6)$$

is an $n_e \times n_e$ matrix. The EnKF analysis therefore involves the summation of a linear combination of the ensemble anomalies, \mathbf{X}' . To state this more explicitly

$$\mathbf{X}_{i,j}^a = \mathbf{X}_{i,j}^f + \sum_{k=1}^{n_e} \mathbf{X}_{k,j}'^f \Lambda_{i,k}, \quad (4.7)$$

where i and j denote the column and row indices respectively. As the analysis is a linear combination of the ensemble anomalies, analysis is more efficient if the ensemble anomalies are independent. Therefore, producing ensembles with anomalies that are linearly independent results in a more efficient assimilation filter.

A set of n_e initial state vectors $[\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_{n_e}]$ should thus be generated by adding a set of n_e independent perturbation vectors $[\mathbf{x}'_1, \mathbf{x}'_2, \dots, \mathbf{x}'_{n_e}]$ to the best-guess initial condition, \mathbf{x} . The initial state vectors \mathbf{x}_i become the column vectors in \mathbf{X} , while the independent perturbation vectors \mathbf{x}'_i become the column vectors of \mathbf{X}' .

4.5.2 Existing Methods

The literature review in section 2.4.3 identified four methods for the ensemble generation: the original method of Evensen (2003), the improved method of Evensen (2004), the breeder method (Toth and Kalnay, 1993) and the optimal perturbation method (Molteni *et al.*, 1996).

The optimal perturbation method requires an adjoint model to generate the fastest growing errors. An adjoint is unavailable in this thesis and is usually unavailable for coastal biogeochemical models, unless specifically developed. For this reason, in spite of the obvious benefits of the optimal perturbation method, its use is impractical for this thesis.

While the breeder method is simple to apply, as the literature review notes, there are questions as to its ability to accurately estimate forecast error variance. Furthermore, because it is a random method and relies on the model to generate perturbations in the direction of the largest growing error there is the possibility

that all perturbations generated will cluster towards one direction, thus reducing the ensemble rank. Also, as the method relies on the inherent nonlinearities of climate models (for which it is derived) to breed the perturbations, it will not be so effective for a diffusive hydrodynamic model which is only weakly nonlinear in the short term.

The original method, relying on correlated random fields, does not guarantee independence. Furthermore, dynamic consistency is achieved only by integrating the ensemble over a few characteristic time scales. For certain systems this could require a long spin up time, but it also makes specifying a particular level of ensemble spread difficult. With a long spin up, by the time the model has reached the start of the analysis period it has altered the initial spread imposed. The improved method uses the SVDs of correlated random fields instead of just the correlated random fields, but it is not clear how the initial distribution is sampled and the difficulty of a long spin up remains.

4.5.3 Proposed Method

As existing initialisation methods did not show promise for this case study, a new approach is proposed. Another related issue is the appropriate magnitude of the initial ensemble spread. This thesis selects a magnitude based on an analysis of the long model run, but the results of the sensitivity testing could also be used.

In general, the *a priori* temperature distribution is unknown so a starting point is to assume a uniform temperature throughout the model based on climatological data. This is considered appropriate because of the limited spatial extent of the model. A range of physically realistic anomalies can then be obtained by taking snapshots of temperature state values from a long model run and removing the spatial mean from each at that instant in time. This gives a $n \times p$ matrix, \mathbf{F} , where n is the number of temperature states and p is the number of snapshots extracted. Each column is a vector representing physically realistic temperature perturbations about a zero mean. This can be expressed as

$$\mathbf{F}_{t,x} = T(t, x) - \bar{T}(t), \quad (4.8)$$

where $T(t, x)$ is the temperature state value at position x and time t , and $\bar{T}(t)$ is

the spatially averaged temperature at time t .

By extracting snapshots at a time interval less than the smallest temporal scale and over a time period longer than the largest characteristic time scale, the full dynamic range of conditions of the temperature field will be covered, thus spanning a wide range of anomaly possibilities. A set of n_e ensemble anomalies is then taken from the first n_e spatial singular vectors of a SVD of \mathbf{F} as described below.

The matrix \mathbf{F} is decomposed using a SVD such that

$$\mathbf{F} = \mathbf{U}\mathbf{\Sigma}\mathbf{V}^T, \quad (4.9)$$

where \mathbf{U} and \mathbf{V} are square orthogonal column matrices of dimensions $n \times n$ and $p \times p$ respectively and $\mathbf{\Sigma}$ is a diagonal matrix with diagonal elements that are the singular values of \mathbf{F} arranged in non-increasing order. The singular values express the importance of their respective (spatial) singular vector, the columns of \mathbf{U} . As the columns of \mathbf{U} are orthogonal, perturbation independence is assured. By using the singular vectors contained in the first n_e columns of \mathbf{U} , the range of dynamic states is objectively, and concisely, represented as these vectors explain the most significant spatial variation in the model.

The singular vectors are then scaled so that their standard deviation is equal to an *a priori* assumption of the initial state uncertainty. If the initial variance of the ensemble is unknown, the average univariate vector standard deviation gives guidance for the initial spread of the ensemble members. Except for the last step of scaling the singular vectors to the *a priori* initial uncertainty, this method provides an objective means for initialising an ensemble, ensuring linear independence of each ensemble anomaly.

The technique can be extended to a multivariate case. For the PPB model this would include currents in the i and j directions u_1, u_2 ; sea level η ; and salinity S , in addition to temperature T . In this case \mathbf{F} is a composite matrix, where n is the total number of states. Before the decomposition is made the states are normalised: as well as removing the spatial mean, the vector of each variable at each snapshot in time is divided by its standard deviation to give a uniform standard deviation of

one. In this case \mathbf{F} is expressed as

$$\mathbf{F} = \begin{bmatrix} \frac{T(t,x) - \bar{T}(t)}{\sigma_T(t)} \\ \frac{S(t,x) - \bar{S}(t)}{\sigma_S(t)} \\ \frac{\eta(t,x) - \bar{\eta}(t)}{\sigma_\eta(t)} \\ \frac{u_1(t,x) - \bar{u}_1(t)}{\sigma_{u_1}(t)} \\ \frac{u_2(t,x) - \bar{u}_2(t)}{\sigma_{u_2}(t)} \end{bmatrix}. \quad (4.10)$$

The average standard deviation over the t snapshots is recorded for each variable, and after decomposition the vector of each variable is re-scaled such that its standard deviation equals this average standard deviation. This re-scaling of the singular vectors will disturb the orthogonality, but the perturbations will still fully span the state domain.

Figure 4.9 displays the first two singular vectors generated when only temperature is used, while figure 4.10 displays the first two singular vectors generated from a multivariate decomposition. Table 4.7 lists the average spatial standard deviation for each variable. The mean standard deviation for rescaling temperature was calculated as 2.1°C , but this was reduced to 1°C based on the temperature spread achieved in the sensitivity analysis. The value of 2.1°C is overestimated because the temperature bias overly warms PPB (and especially the Geelong Arm), while the Bass Strait temperature remains at more realistic levels because of the open boundary.

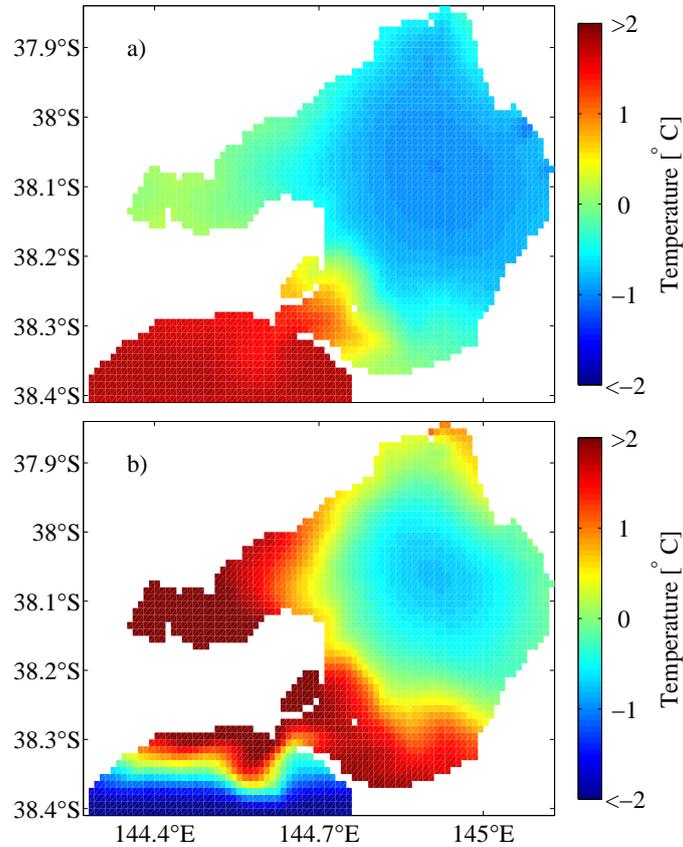


Figure 4.9: The first two univariate temperature singular vectors used as perturbations to initiate the ensemble.

4.6 Ensemble Size

As discussed in chapter 2, specification of an appropriate ensemble size is an unresolved issue for ensemble techniques. If the ensemble anomalies are independent, equation (4.7) sets an upper limit on the number of ensemble members required: the number of model state elements. A larger ensemble size would imply some level of perturbation dependency and the resulting analysis would be inefficient.

Setting the number of ensemble members equal to the number of model state elements is unrealistic for a large distributed model. For the PPB model, this would entail an ensemble with between 20,000–100,000 members, depending on whether univariate (temperature only) or multivariate (all variables) assimilation is pursued.

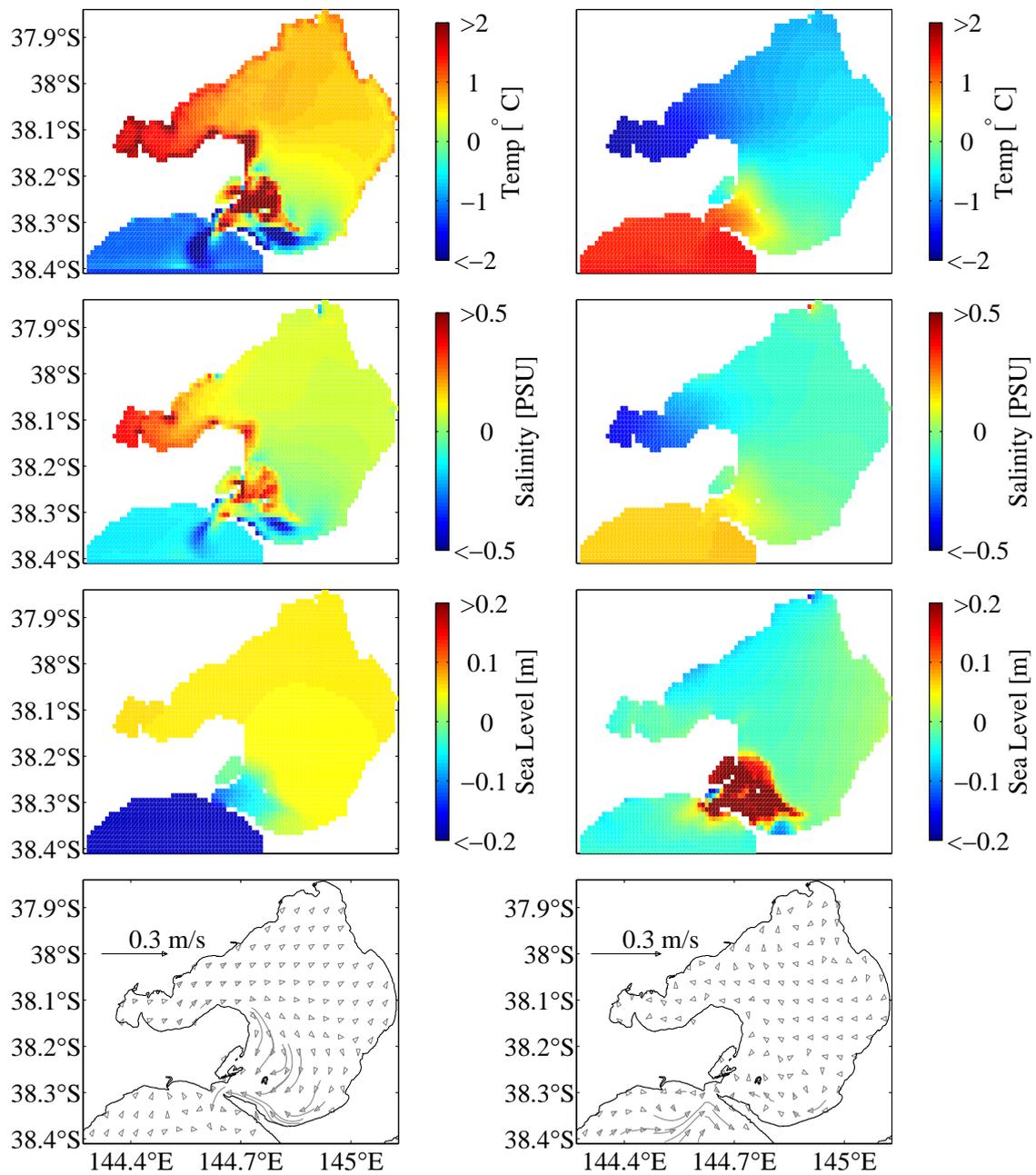


Figure 4.10: The first two multivariate singular vectors for different variables used as perturbations to initiate the ensemble.

Table 4.7: Mean standard deviation values from decomposition of long time series.

State Type	Units	Mean Standard Deviation
Temperature	[°C]	1
Sea level	[m]	0.063
Current in i direction	[ms ⁻¹]	0.071
Current in j direction	[ms ⁻¹]	0.046
Salinity	[PSU]	0.070

A thirty-day model run takes about one hour of real time to compute³, using 20,000 ensemble would take over two years to compute. Although, parallelism could reduce this time somewhat, it is clearly impractical.

Such an exercise is also unnecessary in a model that contains a high degree of state interdependence. Where model states are evolved by similar equations and forced by similar conditions, their values become highly correlated. As this is the case in hydrodynamic models, it is reasonable to assume that the errors are also highly correlated. When ensemble errors are highly correlated fewer independent vectors would be required to describe the range of ensemble perturbations and the ensemble size can be reduced accordingly.

The degree of state independence can be investigated through the significance of singular values obtained from the SVD performed in the previous section. For instance, if 95% of the variance in the system is explained by the first 50 singular values, using 500 ensemble members would be excessive.

Practically, the number of ensemble members to be used is also a function of the cost involved in propagating the ensemble through time, as the example above illustrates, taking years to complete a 30-day assimilation run. Nevertheless, the use of the SVD analysis to determine an optimal number of ensemble members allows meaningful evaluation of the trade-off made when reducing the ensemble size to account for cost constraints.

This approach considers an optimal case, which assumes that the ensemble anomalies are independent. In reality, subject to similar forcing data and model equations, some degree of ensemble anomaly dependence must develop over time, implying that more ensemble members are needed than the SVD suggests. For in-

³using a dedicated sunfire v60x server.

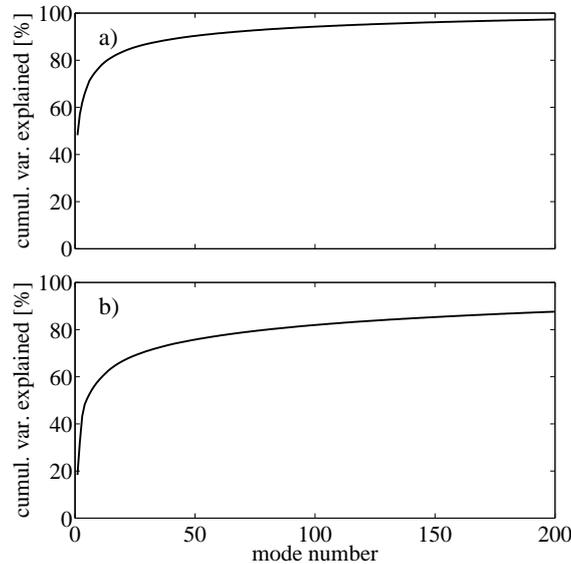


Figure 4.11: Cumulative variance of the system explained by singular vectors for a) the univariate (temperature), and b) multivariate cases.

stance, even with one state, a minimum number of ensemble members are needed to express the probability distribution of the error about the state. As such, ensemble assimilation methods benefit large-state models more than small-state models.

An indication of the number of ensemble members required for assimilation is obtained using the singular values of the previous section. Figure 4.11 plots cumulative percentage variance explained by the mode number⁴ of singular values for the a) univariate and b) multivariate cases. The plots show that most of the variance is characterised by the first few singular modes and relatively less by later singular modes.

The univariate plot (figure 4.11a) indicates a rapidly diminishing singular value with mode number, and suggests that around 20 well spread ensemble members will characterise 84% of the variance. The multivariate assimilation case is presented in figure 4.11b. Here the amount of variance explained by the first 20 ensemble members is much less than the univariate case, with only 67% of the variance explained and with 120 modes needed to explain 84% of variance. This is understandable,

⁴Mode number indicates the ranking of the singular values and vectors. The first singular vector and its associated singular value is assigned mode 1 and so on. By convention the singular values are listed in non-increasing order.

with multiple variables the combined dynamics are more complicated and require more ensemble members to explain the variance. A more complex system cannot be explained as the linear combination of the same number of independent vectors as a less complex system.

Based on this discussion assimilation experiments will be conducted with 20 ensemble members. This number should be sufficient for a univariate assimilation of temperature. Such a small ensemble size might be insufficient for a multivariate, but the computational cost of 100 plus ensemble members is too much for the computational facilities available for this thesis. The sufficiency of an ensemble size of 20 will be tested in chapter 5.

4.7 Forecast Error

Correct specification of forecast error is necessary to ensure the correct spread of ensembles, used to represent ensemble mean uncertainty. Issues of significance to the specification of forecast error are its i) magnitude, ii) spatial distribution, and iii) cross-correlation. While the literature review presented a range of methods, there was no clear guidance on how forecast error should be incorporated or what the forecast error magnitude should be. There are two sources of forecast error: i) model (equations and parameter) error, which is the subject of the majority of the forecast error methods, and forcing data, which the literature review found to be applied the in a simplistic manner.

This thesis will concentrate on forcing data as a source of forecast error because

1. none of the standard model error methods is deemed appropriate for the PPB case study,
2. for short term forecasting, forcing data errors are as significant as model errors and so to first order accuracy may replace model error, and
3. forcing data appears a more appropriate mechanism to determine the magnitude of forcing data error, and to effectively introduce a spatial distribution of the forecast error.

The first part of this section describes why the existing model error methods described in the literature review were rejected. The second part develops the perturbed forcing data method more rigorously.

4.7.1 Existing Model Error Methods

Three approaches to incorporating forecast error were discussed in chapter 2: i) use of multiple models, ii) the addition of correlated random fields at analysis, and iii) the use of stochastic models. Unfortunately, none of these methods is considered appropriate for the coastal data assimilation case study as this section briefly explains.

Attempts to incorporate model error through the use of multiple models were avoided for practical reasons. While other freely available hydrodynamic models exist it is impractical to install, configure, and calibrate each different model as well as translating results between the different file formats used by each model. Furthermore, as relatively few models exist, obtaining an uncertainty range based on more than a handful of ensemble members appears difficult, and as noted in the literature review, previous attempts using this approach have been disappointing (Hamill, 2002).

The utility of correlated random fields is placed in doubt by the linearity testing of section 4.4. The addition of such fields may generate physically unrealistic model states. The remedy applied in the linearity tests, applying a 2-D perturbation equally through the column, is inappropriate for the forecast error as it would automatically impose a perfectly correlated error structure through the water column. While only unstable temperature states were generated by the perturbed fields in the linearity testing, other states could also have been affected. For instance, certain combinations of sea level, and current perturbations could induce unstable shocks which would crash the system.

The use of random fields assumes that the magnitude of the forecast error is spatially uniform. There is no physical reason for this, rather, the error distribution in a model should vary spatially. Consider the distribution of incoming solar radiation in a model where the attenuation coefficient is uncertain. At the surface there is an error associated with the amount of light absorbed, based on the error distribution

of the coefficient. At some depth, however, no light will penetrate regardless of the coefficient value, and the error distribution will be zero. In this case the error varies spatially throughout the domain. Finally, there is no clear guidance on how to apply correlated random fields in a multivariate case, in terms of variable correlations or magnitudes. These arguments discount the use of correlated random fields for model error.

The remaining approach has been to use a stochastic model. Correlated random fields were introduced to avoid the use of stochastic models, and the use of stochastic models defeats an implicit benefit of ensemble techniques: the de-coupling of the model propagation and the analysis. More explicitly, there is no need to adjust the code of an existing deterministic model to attempt ensemble assimilation: a working, well-calibrated model can be applied ‘as is’. Furthermore, specification of the magnitude of the stochastic terms may not be obvious. Nevertheless, the advantages of stochastic models are that the resulting states are physically realistic and the error will be spatially distributed and appropriate to the model rather than spatially uniform. Both these advantages are on account of the stochastic terms in the model equations generating and spreading the error. If a stochastic model is already available, this option becomes realistic. In this thesis the numerical model was already developed, and while some minor tinkering is possible, developing a fully stochastic model was not attempted.

4.7.2 Perturbed Forcing Data

As the literature review shows, perturbed forcing data has already been used to introduce forecast error. The approach outlined in this thesis is more rigorous than previous work. A theoretical framework for incorporating forcing data error is developed and then applied to the data types used by the PPB model.

The nonlinear forecast model can be explicitly represented as

$$\mathbf{x}_{k+1} = f(\mathbf{x}_k, \mathbf{h}_k + \epsilon_k) + \nu_k, \quad (4.11)$$

where \mathbf{h}_k is the forcing data at time k , and ϵ_k and ν_k are random processes with means of zero, representing forcing data error and model error respectively. Errors in forcing data are associated with the measurement errors and spatial errors. An

advantage of using perturbed forcing data is that establishing the uncertainty associated with forcing data is simpler than establishing model uncertainty. This is because the uncertainty of recording instruments is known, and as data are collected at various locations the spatial uncertainty can be estimated. This was performed in section 3.3, especially in figure 3.4, which plots the spatial uncertainty associated with the various forcing data sets used in the model.

The remainder of this section develops an approach to generating perturbed forcing data. The aim of the approach is to avoid the addition of bias to the forcing data while adding perturbations that represent the forcing data uncertainty. Throughout the discussion the data are assumed to be point time series. This is appropriate for the PPB case study as spatially uniform data are used, although, there would appear to be little difficulty in extending the techniques described to spatially varying fields.

A framework for generating perturbed forcing data for typical data types is developed by considering an observed time series of forcing data with p records in time

$$\hat{\mathbf{h}} = [\hat{h}_1, \hat{h}_2 \dots, \hat{h}_p]^T \quad (4.12)$$

used to force a model with n_e ensemble members. If the forcing data are to be unbiased ($E\langle\epsilon_k\rangle = 0$), then generation of an ensemble of n_e forcing data sets $\mathbf{h}^1, \mathbf{h}^2 \dots, \mathbf{h}^{n_e}$, is required such that $E\langle\mathbf{h}_i^j\rangle = \hat{h}_i$. This condition ensures that the ensemble of forcing data is unbiased relative to the original forcing data.

While various forms are possible, an error and offset form was adopted to reflect that the data may suffer from calibration (offset) as well as sampling errors. The adopted form is

$$\mathbf{h}_i^j = \hat{h}_i + \zeta_i + \beta, \quad (4.13)$$

where ζ_i is a time dependent error term of $N(0, \sigma_1)$, being a normally distributed random number with zero mean and a standard deviation of σ_1 , and β is an $N(0, \sigma_2)$ offset applied to the entire time series. Applying an offset β , in addition to the error term ζ_i , provides an additional mechanism for spreading the members while also retaining some of the structure or temporal correlation in the original time series. This is useful for data to which the model is highly noise sensitive and for data that has a high degree of structure in its time series. Without β and relying entirely

on ζ_i may lead to excessive imposed noise that could disturb the model, as well as unrealistic data values.

The form of equation (4.13) was selected after considering a range of possible formulations including temporally correlated noise, and natural (cubic) splines. An advantage of the adopted formulation is that it is simple and easily calculated in real time, and as will be discussed, the parameters controlling spread the standard deviation of β and ζ can be assigned a physical meaning.

The form of equation (4.13) ignores autocorrelation of the error and any correlations between the variables. However, with the exception of sea level, the data used in this thesis is collected at 3-hourly and daily intervals, reducing the autocorrelation of the forcing data errors. Furthermore, as tables 3.5 and 3.6 show the correlations between atmospheric data types are low, generally less than 0.3 and have a maximum of 0.64.

The magnitude and form of the perturbation added to generate the ensemble forcing fields is controlled by two standard deviation terms σ_1 and σ_2 . Realistic values for these parameters can be obtained by analysing the error in observed data and this allows control over the introduction of forcing error.

Based on the form of equation 4.13 the generation of three types of forcing data are considered: i) unrestricted, ii) semi-restricted, and iii) restricted. The notion of a restricted, or otherwise, data type relates to whether the data type has a fixed boundary outside of which values are not physically allowable. For example cloud cover is a restricted data type because its values must fall between zero and eight, whereas precipitation is semi-restricted with a lower bound of zero, and air temperature is unrestricted. The different data types are considered because the spatial error analysis of section 3.3, suggests that the spatial error distribution varies with data type. The specification of error according to data type aids the generation of unbiased physically realistic data sets.

Sea level data are treated separately to avoid introducing high frequency noise to the model. This is because the elevation data are recorded more frequently than other data sets: every six minutes rather than 3-hourly or daily. For the sea level data normally distributed random numbers with a standard deviation of 0.05 m are generated at 12 hour intervals and a cubic spline is fitted through them. This produces a temporal correlation in the perturbations. This perturbation series is

Table 4.8: Values of ξ and χ for various unrestricted data types. The final column indicates the number of weather stations used to derive the ξ and χ values.

Variable	units	ξ	χ	No. of data sets
Air Temperature	°C	1.4	0.6	4
River Temperature	°C	0.5	0.5	3
Open Boundary Temperature	°C	0.25	0.25	none
Air Pressure	Pa	204	204	3
Open Boundary Salinity	PSU	0.0	0.5	none
Wind Vector	m s ⁻¹	2.5	0.7	8

then added to the original sea level time series.

4.7.3 Unrestricted value fields

An unrestricted data type is one whose value is not physically constrained over its normal range. An example of an unrestricted data field is air temperature. As the value of the data can range freely throughout the domain, it follows that the data error is independent of the data value. Therefore, the instrument error in measuring air temperature is assumed constant irrespective of the actual temperature. Unrestricted value fields therefore have the standard deviation of the error term specified as

$$\sigma_1 = \xi, \quad (4.14)$$

where ξ is a constant. The standard deviation of the offset is given by

$$\sigma_2 = \chi, \quad (4.15)$$

where χ is also constant.

Table 4.8 lists the unrestricted data inputs to the model. For unrestricted data, ξ and χ have the same units as the data and represent the standard deviation of the error. The values in table 4.8 are obtained based on consideration of the data analysis in figure 3.4, but they are not directly comparable because the tabulated values are using standard deviation rather than the standard error used in the figures. Where no data sets are used to derive the ξ and χ values they are based on best guess estimates.

The temperature scatter plots, figures 3.4d and 3.4f, show that variation over their main range is constant. The scatterplot of pressure, figure 3.4a, suggests that its variation is more complicated than equations (4.14) and (4.15) imply, but to a first order approximation this is sufficient. According to figure 3.4, the variation in wind is seen to increase with magnitude, suggesting a semi-restricted data type. Nevertheless, wind is incorporated into the model as a vector with i and j components, making the use of an unrestricted form appropriate; negative wind vectors merely indicate the wind blows in the opposite direction. While the error distribution of wind may not be optimal, the wind data will be unbiased.

No ensemble members are generated for river salinity as riverine inputs are assumed to be fresh: zero PSU. This is not always the case and for large saline estuarine systems a variable salinity boundary may be needed. For the open boundary salinity, the prescribed temporal error applied for salinity is zero. This sets a constant salinity boundary with the offset providing the variation between ensemble members.

4.7.4 Semi-restricted value fields

A semi-restricted data type is physically constrained by an upper or lower limit. For the lower limited case the domain is (h_{\min}, ∞) and for the upper limited case the domain is $(-\infty, h_{\max})$. Examples of semi-restricted data fields are precipitation and river flow: both are lower bounded by the value of zero. In the semi-restricted case the standard deviation of the error σ_1 is generally proportional to the magnitude of the data. For example, the uncertainty associated with determining a flow value for a river in flood from a stage measurement is higher than for a low flow event contained within the river banks, and the uncertainty associated with the flow value becomes zero as the river dries up.

In addition to increased error with magnitude, there is a chance that events occur that are not measured. This is especially true for precipitation. In this case an observation of zero cannot be assumed to have an uncertainty of zero. Although this case is not dealt with here, such events may be added when a rare value is chosen from a random sample.

As a first approximation, the standard deviation of the error term for semi-

restricted value fields can be specified as

$$\sigma_1 = (\hat{h}_i - h_{\min})\xi \quad (4.16)$$

for the lower limited case and

$$\sigma_1 = (h_{\max} - \hat{h}_i)\xi \quad (4.17)$$

for the upper limited case. σ_1 is linearly dependent upon the difference between the value \hat{h}_i and the data limit h_{\min} or h_{\max} with the proportionality constant ξ . The offset is similarly formed as

$$\sigma_2 = (\hat{h}_i - h_{\min})\chi \quad \text{or} \quad \sigma_2 = (h_{\max} - \hat{h}_i)\chi \quad (4.18)$$

for the lower and upper limited cases respectively.

Applying a variational error to semi-restricted value data significantly reduces the bias associated with out-of-range values. Data such as precipitation have a lower bound of zero and a significant proportion of zero-valued data. If the unrestricted perturbation approach were applied, on average, half of the ensemble values that were originally zero would be perturbed outside the boundary, requiring truncation to zero to bring them back within the boundary and thus introducing bias. Using a variational error avoids this situation, because the applied error reduces as the boundary is approached, reducing (but not eliminating) the possibility for perturbed values to exceed the boundary. If out of range values are produced they are set to the boundary value.

Using a normally distributed error allows bias to be minimised through judicious choice of ξ and χ . The value of ξ and χ needed to reduce the chance of the perturbed data leaving the boundary is guided by the relationship

$$\xi, \chi \preceq \frac{-1}{z_i}, \quad (4.19)$$

where z_i is a $N(0, 1)$ random number and \preceq is taken to indicate ‘generally’ less than or equal to. This equation is exact only if either ξ or χ is zero. As probabilities can be associated with the chance of a certain value of z_i being exceeded, the probability

Table 4.9: Values of ξ and χ for various semi-restricted data types. The final column indicates the number of weather stations used to derive the ξ and χ values.

Variable	units	ξ	χ	No. of data sets
Precipitation	mm d ⁻¹	0.25	0.25	4
Evaporation	mm d ⁻¹	0.2	0.2	2
Short Wave Radiation	W m ⁻²	0.05	0.05	6
River Discharge	m ³ d ⁻¹	0.05	0.2	none

of domain exceeding values occurring can be estimated. For example, to reduce the chance of a domain violating error being introduced to less than one in one thousand, $\Pr(z_i \leq -3)$, ξ should be set to less than about 0.32. A similar argument can be constructed for the choice of χ . Further discussion on this and the derivation of equation (4.19) is given in appendix B.

The ξ and χ values adopted for semi-restricted data types are listed in table 4.9. Again, they are based on the analysis of figure 3.4 or best guess estimates. For semi-restricted data types the ξ and χ values are non-dimensional and can be viewed as scaling factors. For instance, if the standard deviation of the perturbation added to precipitation data is 0.25 times the data value, this would mean that the perturbation added to the data is $\pm 25\%$ of the precipitation value. The ξ and χ values for river discharge are best guess estimates; the higher offset (χ) component reflects the larger lag correlations found within river flow data.

4.7.5 Restricted value fields

A restricted data type is physically constrained by an upper and lower bound (h_{\min}, h_{\max}). Applying a perturbation term to this type of data requires somewhat more thought to the error distribution. A constant error such as equation (4.15) could be used, with any bias due to the truncation of domain exceeding values accepted. A better approach is to assume the maximum standard deviation occurs at the mid point of the domain and reduces linearly to zero at the domain boundaries, giving a triangular shaped distribution by

$$\sigma_1 = \begin{cases} \frac{\hat{h}_i - h_{\min}}{h_{\text{mid}} - h_{\min}} \xi, & h_{\min} \leq \hat{h}_i \leq h_{\text{mid}}, \\ \frac{h_{\max} - \hat{h}_i}{h_{\max} - h_{\text{mid}}} \xi, & h_{\text{mid}} < \hat{h}_i \leq h_{\max}, \end{cases} \quad (4.20)$$

Table 4.10: Values of ξ and χ for various restricted data types. The final column indicates the number of weather stations used to derive the ξ and χ values.

Variable	units	ξ	χ	No. of data sets
Relative Humidity	%	5.0	5.0	5
Cloud Cover	oktas	0.4	0.3	2

where h_{mid} is $\frac{h_{\text{max}}+h_{\text{min}}}{2}$. An example of a restricted data type is cloud cover. Cloud cover data refers to the proportion of the sky covered by clouds with zero signifying clear skies and eight indicating completely cloudy skies. It is reasonable to associate an error distribution following equation (4.20) with cloud cover data, as it is easy to decide if the sky is completely covered or is completely free from clouds, but to determine whether cloud cover is four, five or six oktas is more difficult and subjective. The analysis of chapter 3 showed that cloud cover is more uncertain for midrange values. The applied error distribution takes this into account.

The offset is formed in a similar fashion with

$$\sigma_2 = \begin{cases} \frac{\hat{h}_i - h_{\text{min}}}{h_{\text{mid}} - h_{\text{min}}} \chi, & h_{\text{min}} \leq \hat{h}_i \leq h_{\text{mid}}, \\ \frac{h_{\text{max}} - \hat{h}_i}{h_{\text{max}} - h_{\text{mid}}} \chi, & h_{\text{mid}} < \hat{h}_i \leq h_{\text{max}}. \end{cases} \quad (4.21)$$

As with the semi-restricted case, the choice of ξ and χ values affects the probability that perturbed data leaves the domain. For the restricted case the guiding relationship is

$$\xi, \chi \preceq \frac{h_{\text{min}} - h_{\text{max}}}{2z_i}. \quad (4.22)$$

The derivation of equation (4.22) is given in appendix B. As with equation (4.19), probabilities can be assigned to different values of z_i . Thus equation (4.22) indicates that a ξ (or χ) value of less than one sixth of the data range yields a probability of generating a domain leaving perturbed data value of less than one in one thousand.

For the restricted value fields, ξ and χ have the same units as the data type and the values adopted are listed in table 4.10, based on the analysis of sections 3.3.6 and 3.3.7. The standard deviation of the perturbation added to relative humidity data, at the mid point of its range (50%) is 5%. For cloud cover this distribution compares well with the data analysis of section 3.3.6. Relative humidity has a complicated error distribution, and while not ideal, applying this distribution to

relative humidity data avoids generating physically unrealistic and biased ensembles.

4.7.6 Application of Perturbed Forcing Data

A sample of perturbed forcing data generated using the above method is presented in figure 4.12, based on the forcing data used in the PPB model. These figures represent only a small time window of a longer time series and are presented to illustrate some of the details discussed above.

The advantage of using variable perturbation terms to avoid introducing bias at the lower boundary of the semi-restricted data is well displayed by precipitation (panel b) and the triangular error distribution of a restricted data type (relative humidity) is observed in panel d, with larger variation about the midrange (40%–60%) values than the boundary (80%+) values. The advantage of adding both offset and error terms to the data rather than just an error term is well illustrated by panel c, which shows river flow. The ensemble members all have more point to point variation than the original time series, but maintain the same time series trend. This retention of the trend is attributable to the offset; if the same variation for each point was attempted by the application of only the error term the structure of the time series would be lost because of the the magnitude of the error needed.

To demonstrate the ability of perturbed forcing data to spread the forecast, an ensemble of 20 members was run, with each member using a different realisation of the forcing data. All ensemble members start with the same initial conditions, taken from the long model run of section 4.2 and therefore avoiding the need for a model spin up. The evolution of the ensemble members at the surface locations A, B, and C (indicated in figure 4.1) is shown in figure 4.13 for temperature. Any spread of ensemble members is due to the perturbed forcing data, and demonstrates that perturbed forcing data is a suitable mechanism to introduce forecast error into an ensemble. At the end of the simulation period, the error at the 3 locations A, B, and C is 2.48°C, 0.28°C, and 1.15°C respectively. Although these values are less than the average 3–4°C error found during the model testing, the error at location B compares well with the RMSE of 1.1°C found at West Channel Pile over the first 67 days of the long model run.

Initially, the largest ensemble spread is observed at the open boundary (panel b),

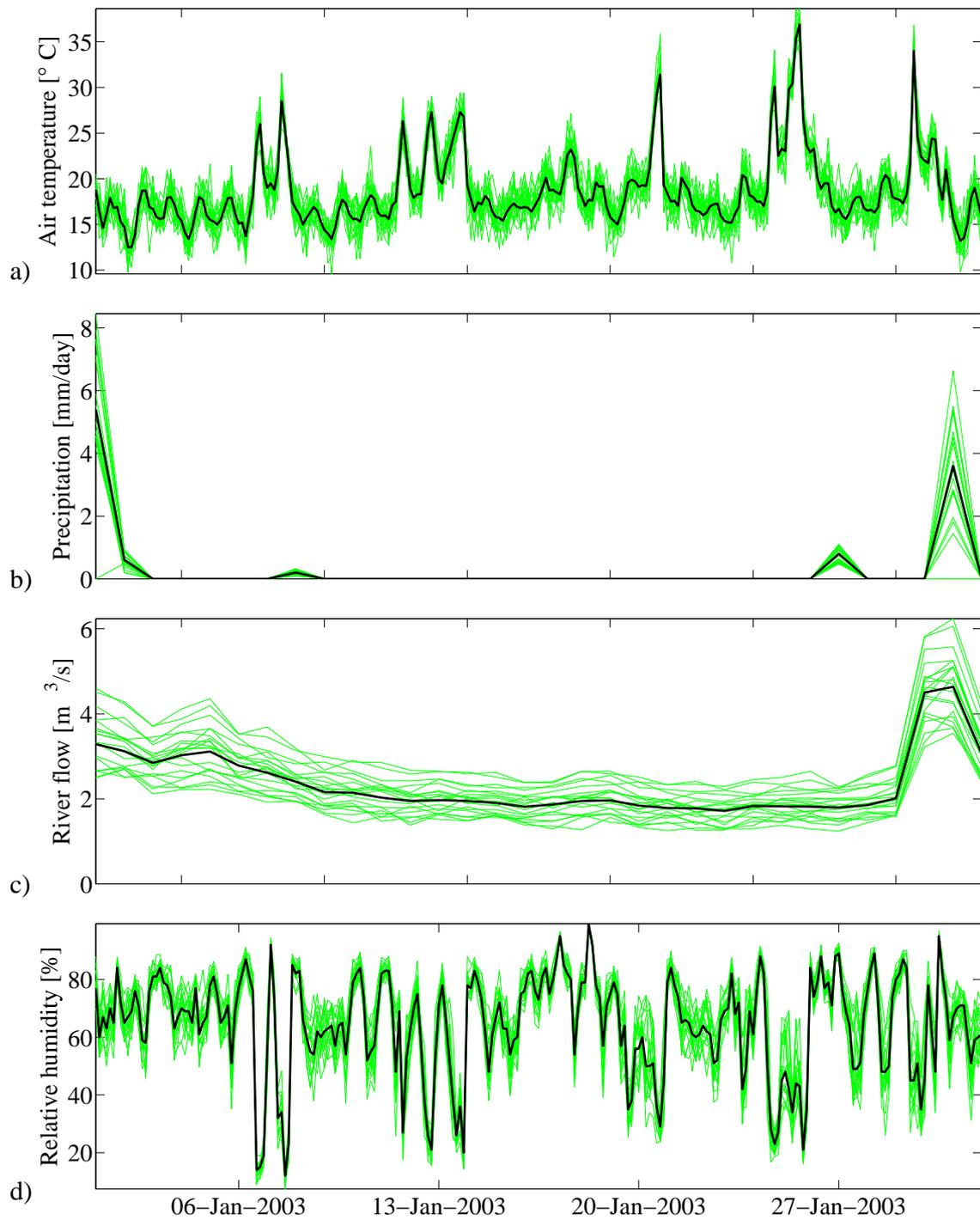


Figure 4.12: Examples of perturbed forcing data ensembles. Thin green lines represent the ensemble members and thick black lines the original data of a) air temperature, b) precipitation, c) river flow, and d) relative humidity.

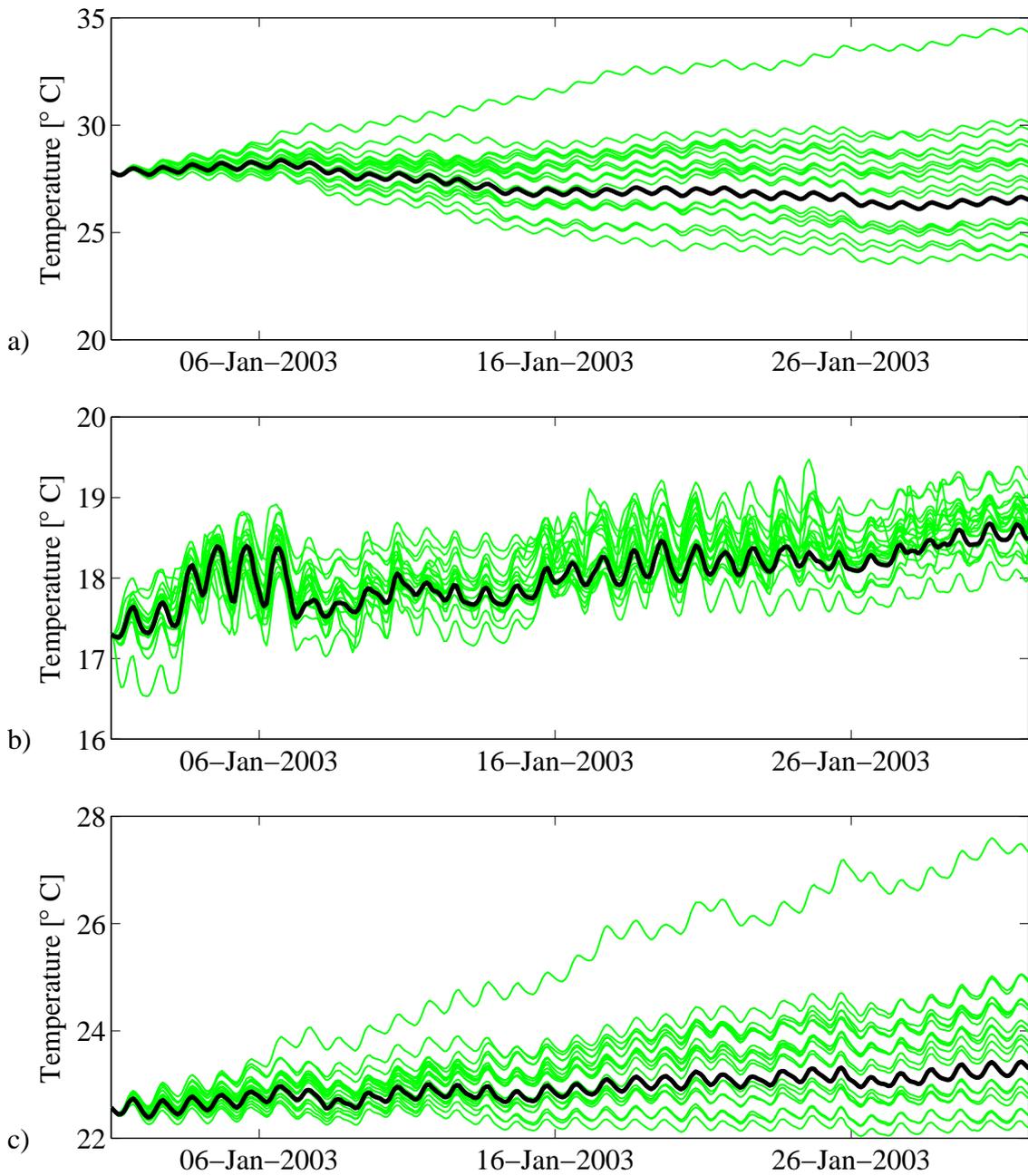


Figure 4.13: Time series illustrating the effects of forcing data ensembles on model predictions of temperature at surface sites a) A, western arm, b) B, open boundary and c) C, centre of PPB. The locations are indicated in figure 4.1. Thin green lines represent the ensemble and thick black lines the truth.

driven by the perturbed temperature forcing at the open boundary. This spread is constant in time, constrained by the open boundary, and so does not grow further. Over time a gradual increase of variation in temperature is observed within PPB (panel c). The effects of the perturbed atmospheric forcing data act more slowly than the open boundary. The spread increases according to the perturbations in the various atmospheric forcing data, finding a limit based on the size of the perturbations. The spread due to atmospheric forcing was most pronounced at the edges of PPB and especially in the western arm (panel a). A similar ensemble spread results for other variables (sea level, currents etc.) as well.

With a different ensemble spread found at each location, figure 4.13 also shows that the resulting forecast error is not uniform, but varies spatially with location. This is also seen in figure 4.14a, which plots the surface ensemble standard deviation at the end of the simulation. The spread also depends on the variable considered and the magnitude of the perturbations of the forcing data that dominates that particular variable: the distribution may be significantly different for salinity and sea level.

A brief sensitivity analysis of the perturbed forcing data (not included) indicated that the spread of temperature is most sensitive to changes in open boundary temperature, air temperature and short wave radiation values. However, the forecasts are fairly insensitive to changes of up to 50% in the χ and ξ values.

This section has shown that

1. perturbed forcing data can introduce unbiased error into the ensemble to represent forecast error,
2. the forecast error introduced is not spatially uniform, but varies locally, depending on the model and forcing data interaction, and
3. the forecast error predicted is in the range of that estimated by the model testing.

This exercise gives confidence that the perturbed forcing data method developed is behaving well and is appropriate for use in the data assimilation case study of chapters 5 and 6.

4.7.7 Forecast Error Covariance

This section extends the findings of the previous sections by exploring the forecast error covariance structure formed by the ensembles. To validate the forecast error covariance predicted by the perturbed forcing data for temperature, the forecast covariance values are compared against those obtained independently using a version of MECO adjusted to have a stochastic heatflux module. Then the variable covariances are investigated to explore the likely effects of multivariate assimilation.

The propagation of the ensembles, driven by perturbed forcing, allows a first analysis of the forecast error covariance. The covariance estimate of two states, x_i and x_j is given by

$$\mathbf{P}_{i,j} = \rho_{i,j}\sigma_i\sigma_j, \quad (4.23)$$

where $\rho_{i,j}$ is the correlation between x_i and x_j , and σ_i and σ_j are the standard deviations of x_i and x_j respectively. An ensemble estimate of the covariance is given by equation (2.16). During analysis, the updating of a cell is based on the covariance between the cell in question and the cell corresponding to an observation. Comparing the strength of the covariance between the different states gives an indication of the effect the assimilation of the satellite observations will have on model states.

A stochastic version of the MECO heatflux module was built by multiplying each component of the net heatflux — Q_N : incoming solar radiation Q_S , longwave radiation Q_L , latent heat Q_E , and sensible heat Q_H — by a spatially uniform stochastic component. Thus the net heatflux equation becomes

$$Q_N = \alpha_1 Q_S - \alpha_2 Q_L - \alpha_3 Q_H - \alpha_4 Q_E, \quad (4.24)$$

where α_i are Gaussian distributed random processes with a mean of one and a standard deviation of one. The standard deviation values were chosen arbitrarily.

The stochastic model was initialised with the same initial conditions as used previously for the perturbed forcing data (values taken from the long model run, and so no spin up is required). 20 stochastic 30-day runs were made with each ensemble member using the same set of forcing data (rather than perturbed forcing data), as specified in section 4.1. The results are presented in figure 4.14. As it is difficult to represent a covariance matrix visually, the surface standard deviation

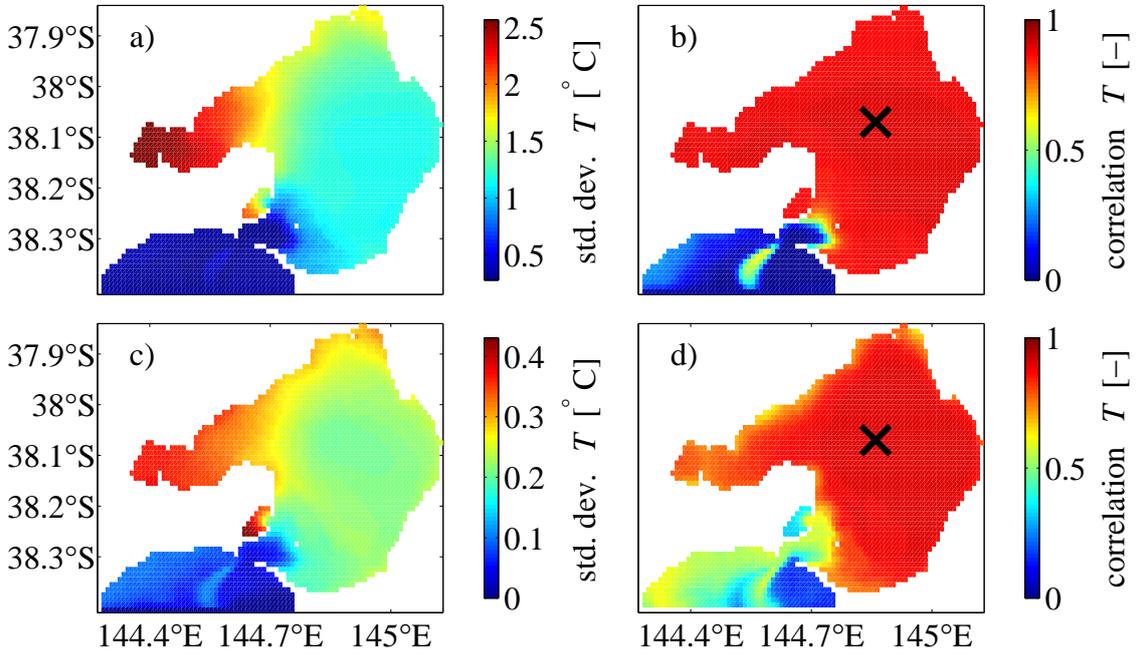


Figure 4.14: Comparison of the forecast error covariance of surface temperature calculated by perturbed forcing data (top) and by a stochastic model (bottom). at the end of a 30 day simulation. Panels a and b display the surface standard deviation, while correlation between a cell in the centre of PPB (indicated by the cross) and other surface cells is shown in panels c and d.

and correlation components of covariance are displayed separately. The correlation is shown as the correlation between a selected cell in the centre of the bay (indicated by a cross) and all other surface cells. Similar correlation plots based on other locations are presented in appendix D. As the standard deviation values of equation (4.24) were chosen arbitrarily the comparison between the standard deviations of the two forecast error covariance predictions should be based on the relative pattern rather than the actual values.

Around the centre of PPB the patterns of the correlation (panels b and d) and the standard deviation (panels a and c) are similar. In Bass Strait the correlations and standard deviations are less similar. This is because the stochastic model only varies with the heatflux module. At the open boundary it predicts zero variation, as the temperature at the open boundary is constant for all stochastic model runs. The perturbed forcing predicts a variation in Bass Strait equal to the imposed variation

of the open boundary data, (figure 4.13b). Similar correlation patterns are found based on other locations (appendix D)

The conclusion of this comparison is that the forecast error covariances predicted by the perturbed forcing data are consistent with other methods. This gives confidence that the values are accurate and enables an investigation of the covariances between temperature and other variables.

The covariance between variables is made by plotting the surface covariance between a surface temperature state in the centre of the PPB (indicated by a cross) and the surface cells of another variable. Again the covariance is displayed by splitting it into components of standard deviation and correlation. This is presented in figure 4.15, with correlation plots for other locations of the temperature cell in appendix D.

An analysis of the covariance between variables will indicate the effect that the assimilation of temperature will have on the other variables. If low covariances are found then there is little connection between the variables, and the assimilation of SST will have little effect on the other variables. The independence of temperature and salinity (figure 3.17) and lack of geostrophic balance in PPB (section 4.2) point to a de-coupling of temperature and other variables. This is supported by the plots of figure 4.15, where the small correlations between temperature and the other variables evidence the lack of relation between the variables.

The salinity uncertainty, measured by the standard deviation, is higher near the open boundary, sources of fresh water and in the Geelong Arm, where changing evaporation rates have a larger impact on salinity. Sea level uncertainty is larger towards the edges of PPB and lower in the Great Sands. The uncertainty is of the order of 0.03 m. This seems reasonable given the sea level amplitude in PPB of about 0.4 m. The largest uncertainty with the currents is in Bass Strait and the Great Sands, with values ranging from 0.01 to 0.07 ms^{-1} .

All the variables have low correlation with temperature. There is virtually no correlation between the temperature and salinity. There appears to be some minor correlation with the sea level in Bass Strait, although this may be an artifact of the tides. Also, there appears to be some apparently random correlations between the currents and temperature, although still relatively minor. While this analysis is not comprehensive, it strengthens the supposition of de-coupled temperature and

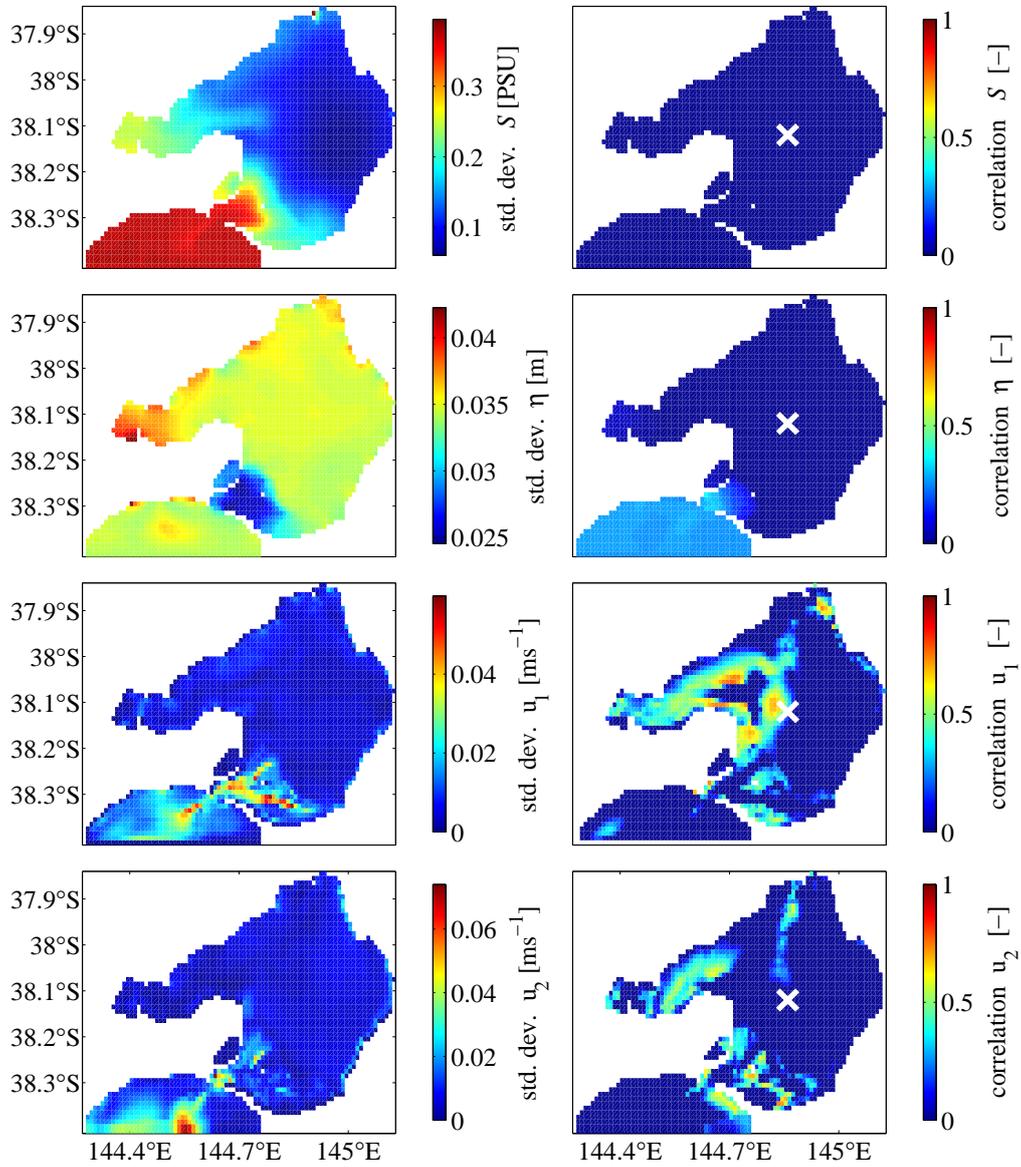


Figure 4.15: Comparison of the forecast error covariance of the temperature with other model fields. The cross in the correlation plots is the location of the temperature cell that all the correlations are made against.

the other variables. This is important because if there is no covariance between temperature and the other variables (they are independent), then assimilation SST observations would have no direct impact on the other variables. In this case a univariate assimilation of temperature could be performed rather than a multivariate assimilation. This possibility will be tested in the experimentation of chapter 5.

The investigation of this section has shown that

1. the forecast error covariances generated by the perturbed forcing data method are similar to those generated by other methods, and
2. that low covariances exist between temperature and other variables.

The implication of these findings for the data assimilation are that the perturbed forcing data method is further validated and that a univariate rather than a multivariate assimilation analysis might be possible.

4.8 Chapter Summary

Confidence in the numerical model and the ensemble assimilation approach are fundamental to a data assimilation study. This chapter lays the foundations of the later data assimilation experiments through the testing of the numerical model and the development of the ensemble assimilation techniques used in this thesis.

The numerical modelling uses the Model for Estuaries and Coastal Oceans (MECO) and is based on the modelling performed for the PPBES. The model is reconfigured for the study period and is found to replicate the main features of PPB: currents, sea levels and seasonal and diurnal temperature changes. Sea level prediction accuracy in the vicinity of The Heads was hindered by the relatively coarse grid resolution. Nevertheless, away from The Heads RMSE errors were between 0.05 and 0.10 m. Temperature was less accurately predicted, with the model predictions biased towards high values. The offset in the temperature prediction is of the order of 3–4°C. This is much larger than the error in the satellite observations and indicates that assimilation should significantly improve the model forecasts.

A sensitivity analysis of the model indicated that the errors in the model structure are of the same order of magnitude as errors in the forcing data. Therefore, to

a first order approximation, perturbed forcing data can act as a surrogate for model errors when incorporating forecast error into the assimilation ensemble.

The generation of independent ensemble perturbations was pursued by using the singular vectors from the decomposition of a long model run. The singular values from this decomposition indicated that for a univariate assimilation an ensemble of 20 members should suffice. The incorporation of forecast error into the ensemble is made using perturbed forcing data. By assigning the forcing data to one of three data types: restricted, semi-restricted, and unrestricted, the introduction of bias into the forcing error is minimised. The forcing data errors produced are the same order as actual prediction errors found in the model testing.

Chapter 5

Twin Experiments

This chapter identifies the best assimilation configuration for a coastal hydrodynamic model to assimilate SST observations using the PPB case study. This is achieved by conducting a series of synthetic experiments and enables the improved assimilation of actual SST observations that will be undertaken in the following chapter. These experiments use the data of chapter 3 with the model and data assimilation approach outlined in chapter 4 to address the issues raised in chapter 2.

The ensemble assimilation configuration is explored in an twin experiment environment, using the ensemble generation and forecast error propagation techniques that were developed in the previous chapter, together with the PPB model. Four elements of the assimilation configuration are investigated: i) the importance of multivariate versus univariate assimilation, ii) the effect of model bias on the analysis and subsequent forecasting skill, iii) the efficiency of deterministic versus stochastic ensemble filters, and, iv) the need for ensemble filters.

5.1 Introduction

While the form of the ensemble generation and forecast error techniques were decided in the previous chapter, there remains a bewildering array of choices as to the configuration of an ensemble assimilation filter. In this chapter various ensemble assimilation configurations are tested in the PPB case study setting. By testing various configurations this chapter aims to give advice on the best configuration for

use in a coastal hydrodynamic model assimilating SST and also to advise where effort to improve the forecast should be best focused. The testing is enacted by means of identical twin-experiments also known as synthetic studies.

The essence of a twin experiment is to use data assimilation to reproduce the true system state from the prediction of an initially degraded model. The general form of a twin experiment is that an initial model simulation is made, which is designated the truth. Snapshots of the states of this truth simulation are extracted to form synthetic observations with random noise added to simulate measurement error. The effectiveness of the assimilation scheme is determined by degrading the model (by altering any or all of the initial conditions, model parameter settings, and forcing data), and then assimilating observations into the model forecasts as they become available. Ideally, the assimilation scheme returns the degraded simulation states to the true model states, minimising the RMSE between the ensemble mean and the truth. The improvements that the assimilation brings are calculated by comparison with a control run, made by running the model from the degraded initial condition without assimilation.

A related, but different approach to testing assimilation schemes in a synthetic environment is through observation system simulation experiments (OSSEs) (Houtekamer *et al.*, 1996). The difference between OSSEs and twin experiments is that OSSEs typically use different models to produce the truth and the forecasts. An implication of using twin experiments rather than OSSEs is that because the underlying models equations are essentially the same (or similar) the twin experiments may give an overly optimistic indication of the potential performance of assimilation scheme tested. This is not of concern for this thesis because in chapter 6 the assimilation scheme is applied to real observations where its true performance is assessed.

The first experiment compares a multivariate assimilation against a univariate assimilation. This comparison considers whether the analysis should update all model states or be restricted solely to temperature. A multivariate assimilation should be more accurate, because it updates all the states dynamically, rather than relying on model equations to spread the analysis correction in subsequent forecasts. However, when there is a low covariance between temperature and the other variables, and especially with small ensemble sizes, a univariate assimilation may be more efficient, as it is less affected by sampling error in the calculation of the ensemble forecast

error covariance matrix.

Model bias is an issue for virtually all model numerical models and particularly for the PPB model which is warm biased. While the assimilation equations assume an unbiased model, investigating model bias gives an indication of the severity of this assumption. By comparing a biased assimilation against an unbiased assimilation, the impact of model bias can be assessed.

As discussed in the literature review, ensemble filters can be classed as either stochastic (e.g. the EnKF), which achieve the correct analysis spread by using perturbed observations, or as deterministic (e.g. the EnSRF), which spread the ensemble anomalies using deterministic equations. At small ensemble sizes it is claimed that deterministic filters will outperform stochastic filters because sampling error in the perturbed observations reduces their performance. This is investigated at a range of ensemble sizes by comparing the EnKF and the EnSRF. Furthermore, an analysis of both filter equations gives insight into the shape of the resulting ensemble anomalies, which helps the selection of an appropriate ensemble filter.

Ensemble methods are appropriate in dynamic environments where the error covariance structure of the system varies with time. If the covariance structure of the system does not vary significantly in time, the extra computational expense of an ensemble filter may be unnecessary. Justification of the use of ensemble methods is sought by displaying the conditions through the water column over time.

5.2 Method

The same synthetic study setting is used for all the experiments, with the simulation period being the month of January, 2003. The experiments conducted in this chapter use the PPB model configuration described in section 4.1. Specifically, a rectangular 0.01 degree resolution model grid is used with 14 vertical layers (see figure 4.1), which is open to Bass Strait along its southern edge. Open boundary conditions are a climatological temperature, constant salinity and observed sea level. Atmospheric forcing is spatially uniform and the data are taken from stations listed in table 3.15. Incoming solar radiation is derived using the formula of Zillman (1972). River flow (and temperature and salinity) enter the PPB model at various locations listed in section 4.1.3 and shown in figure 3.1.

The various assimilation ensembles are initialised with singular vectors derived from the SVD of a long model run developed in section 4.5. In univariate assimilations only the temperature states of the ensemble are initialised (based on equation 4.8), while in the multivariate assimilations all variables of the ensemble are initialised (based on equation 4.10).

Forecast error is incorporated into the ensemble using the perturbed forcing data method developed in chapter 4. Based on the study of the proportion of system variance explained by singular vectors in section 4.6, 20 ensemble members are used except where explicitly mentioned. The analysis will be made using the EnSRF formulation described in section 2.5.2 — specifically equations (2.30) and (2.34) — except where otherwise noted.

5.2.1 Synthetic Truth and Observations

Two truth simulations were generated. One used the same model and forcing data as the assimilation simulation and became an unbiased truth. In the second truth simulation the Gill heatflux module used in MECO was replaced by the Zillman module. This created a biased truth. The biased truth was therefore taken from the long model run of section 4.2, while the unbiased truth was taken from experiment 9 of the sensitivity study of section 4.3.2. The truth simulations values were recorded at 2-hourly intervals and subsequent RMSE calculations are based on this 2-hour record frequency.

Synthetic observations were created from both truth runs, by extracting snapshots of the surface temperature state of the truth runs at midnight once every two days, observation error was added based on the observation error covariance. An independent error with a standard deviation of 0.5°C has been adopted for these experiments, based on the standard error quoted for the NLSST algorithm (Nalli and Smith, 1998). A discussion on the observation error covariance choice was made in section 3.4.4. This the observation error covariance adopted for the experiments was

$$\mathbf{R} = 0.5^2\mathbf{I}. \quad (5.1)$$

Thus normally distributed error with a mean of zero and a standard deviation of 0.5 is added to the extracted truth snap-shots to generate the synthetic observations.

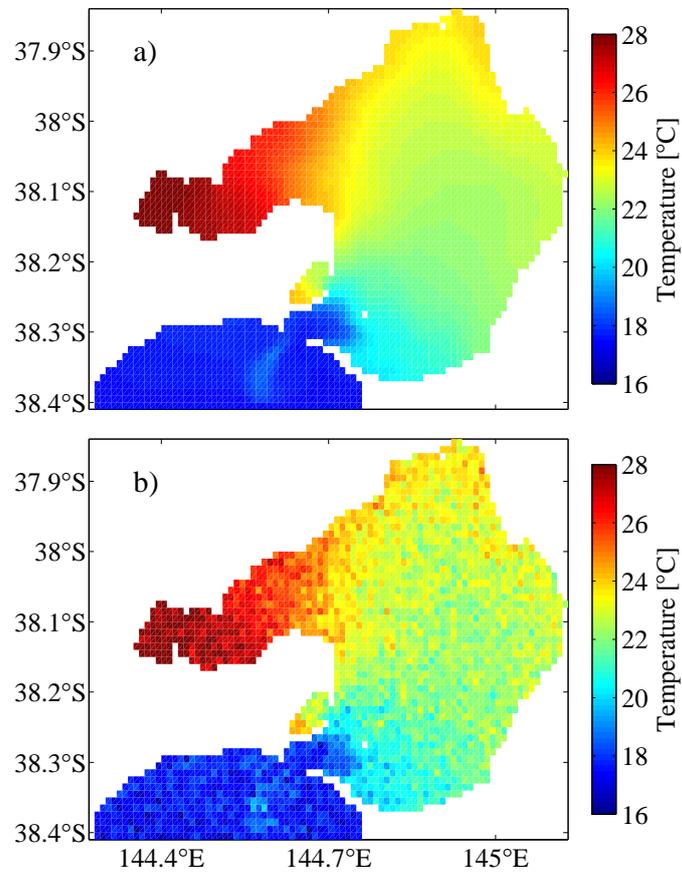


Figure 5.1: Surface snapshots of observations for 19 January 2003 showing the a) synthetic truth, and b) synthetic observation.

The difference between skin and bulk temperature is not an issue in this chapter because the observations are taken directly from simulated bulk temperature.

An example of the biased truth and a synthetic observation are shown in figure 5.1. The speckled appearance of the observation is a result of the uncorrelated observation error assumption. When the biased synthetic observations are assimilated, the forecast model is biased, conversely when the unbiased synthetic observations are used the forecast model is unbiased. Unless otherwise specified, the biased truth and observations will be used in the experimentation.

5.2.2 Initial Conditions and Control Runs

The initial conditions of the assimilation runs are degraded such that temperature was uniformly 18°C throughout the model, salinity was 35 PSU, and sea level and currents were initially zero. With these initial conditions a 30-day control run was made with no assimilation. This control run was the basis against which subsequent assimilation experiments were judged.

5.3 Univariate Multivariate Comparison

Formally, ensemble assimilation can update any unobserved state based on an innovation (the difference between an observation of a state and a forecast of that observed state) and the covariance between the errors of the observed and unobserved state. If the errors are uncorrelated, the covariance is zero and increments are also zero. A disadvantage of the Monte Carlo approach, used in ensemble filtering, is that, due to sampling errors, states with uncorrelated errors can appear to have correlated errors by chance. This is especially pronounced for small ensemble sizes, and analysis using such erroneous correlations would lead to less accurate forecasts.

The covariance investigation of section 4.7.7 suggests that the covariance between temperature errors and the other variables is quite low. In this case multivariate assimilation might be susceptible to sampling errors when only 20 ensemble members are used, and limiting the analysis to univariate temperature assimilation would give a more accurate result.

This section contrasts univariate assimilation with multivariate assimilation for ensembles of 20 and 50 members. While section 4.6 suggested that 120 members might be necessary for multivariate assimilation, this is unfeasible based on the computing resources available for this thesis. In the multivariate assimilation, the ensemble state vector \mathbf{X} is a composite of temperature, salinity, sea level and current fields, while for the univariate assimilation it only contains the temperature states. Also in the multivariate assimilation each of the fields uses perturbed initial conditions, while for the univariate assimilation only the temperature ensemble members are perturbed. The other fields — salinity, sea level etc. — are initially constant,

and are separated by the perturbed forcing data.

5.3.1 Univariate Multivariate: Results

An initial evaluation of performance is made by comparing the temporally averaged RMSE of the different runs for different variables, presented in table 5.1. The results of the control run are also presented for comparison. The results for currents are not listed, with sea level used as a surrogate for currents: if sea level is accurate, currents are assumed accurate also.

When compared with the control run, the assimilation should reduce the error. However, this only occurs for temperature. In the univariate assimilation, the other variables perform equally as well as in the control run, while in the multivariate assimilation the other variables perform significantly worse. Except for the case of temperature with an ensemble size of 50, the univariate assimilation performs better than the multivariate assimilation, with a lower RMSE for all variables.

Figure 5.2 displays a time series of the RMSE for the different variables. The RMSE of the control run is also plotted for reference. In the univariate assimilation, salinity and sea level RMSE follow the control RMSE, while in the multivariate assimilation large increases in RMSE occur at analysis times when observations are assimilated. The analysis is introducing the error through erroneous covariances.

Increasing the ensemble size from 20 to 50 reduces both the error in the multivariate and univariate assimilations, but except for temperature, the univariate assimilation remains more accurate. This indicates that sampling error is a problem

Table 5.1: Average RMSE between the ensemble mean and the truth of various model configurations for Salinity S , Temperature T , and Sea Level η . Boldface values indicate lowest RMSE for the respective variable.

Description	RMSE S [PSU]	RMSE η [m.]	RMSE T [°C]
Control	1.20	0.0041	3.14
Univariate, ensemble size 20	1.21	0.0080	0.69
Multivariate, ensemble size 20	1.66	0.0848	0.75
Univariate, ensemble size 50	1.20	0.0066	0.68
Multivariate, ensemble size 50	1.59	0.0276	0.65

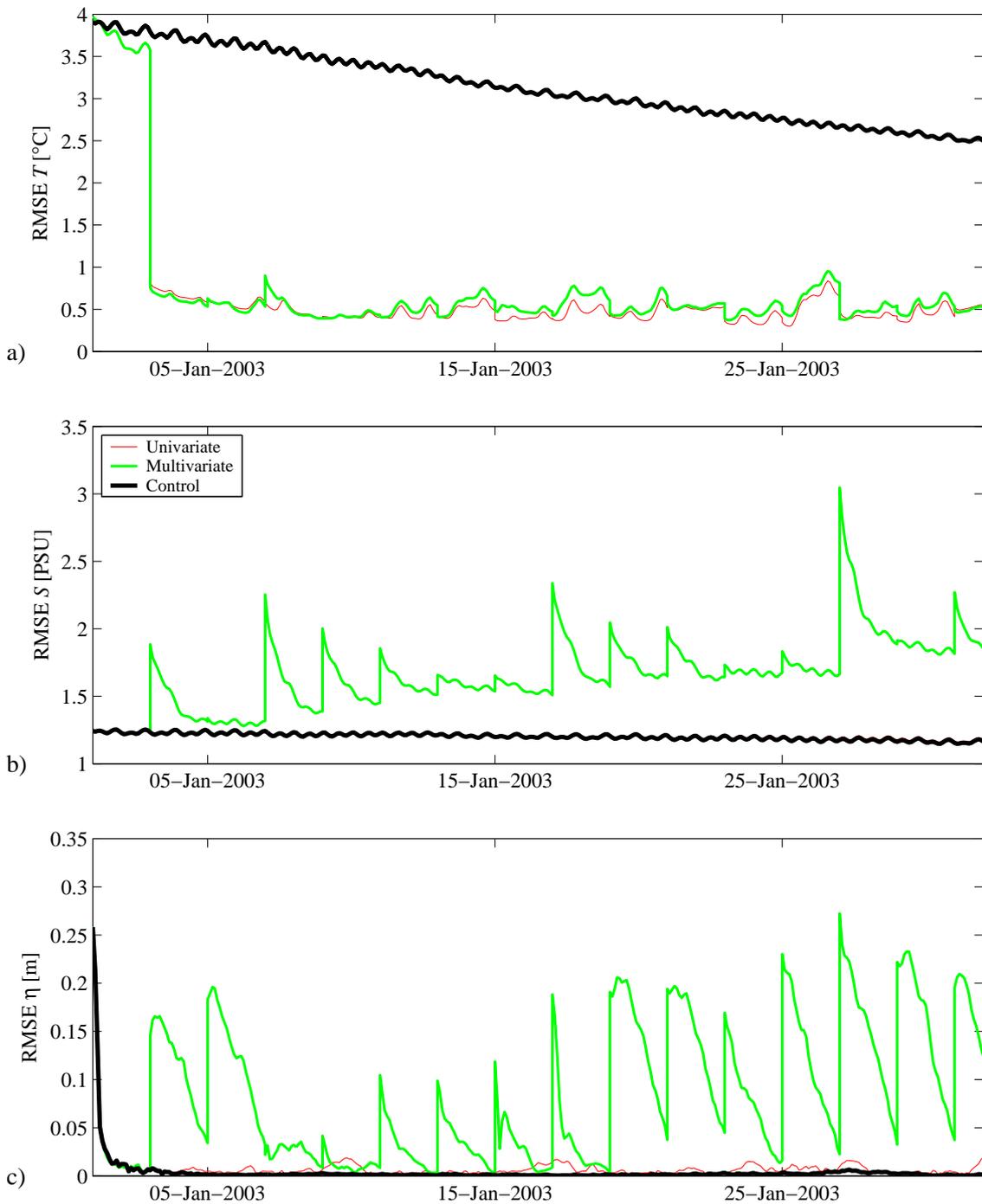


Figure 5.2: Time series of spatial RMSE of a) temperature T , b) salinity S , and c) sea level η . Present on the graphs are the multivariate, univariate runs for 20 ensemble members together with the control run.

for the multivariate assimilation. The results are explored in more detail for each variable.

Salinity

In all cases the salinity fields are poorly predicted. The univariate assimilations and control run produced an average RMSE of around 1.20 PSU, while the multivariate assimilation produced an average RMSE of 1.66 and 1.59 PSU for 20 and 50 ensemble members respectively.

The reason for the multivariate case performing worse than the univariate case is shown in figure 5.3. This figure plots the ensemble mean salinity for both the univariate and multivariate runs at two locations shown in figure 4.1: site C in

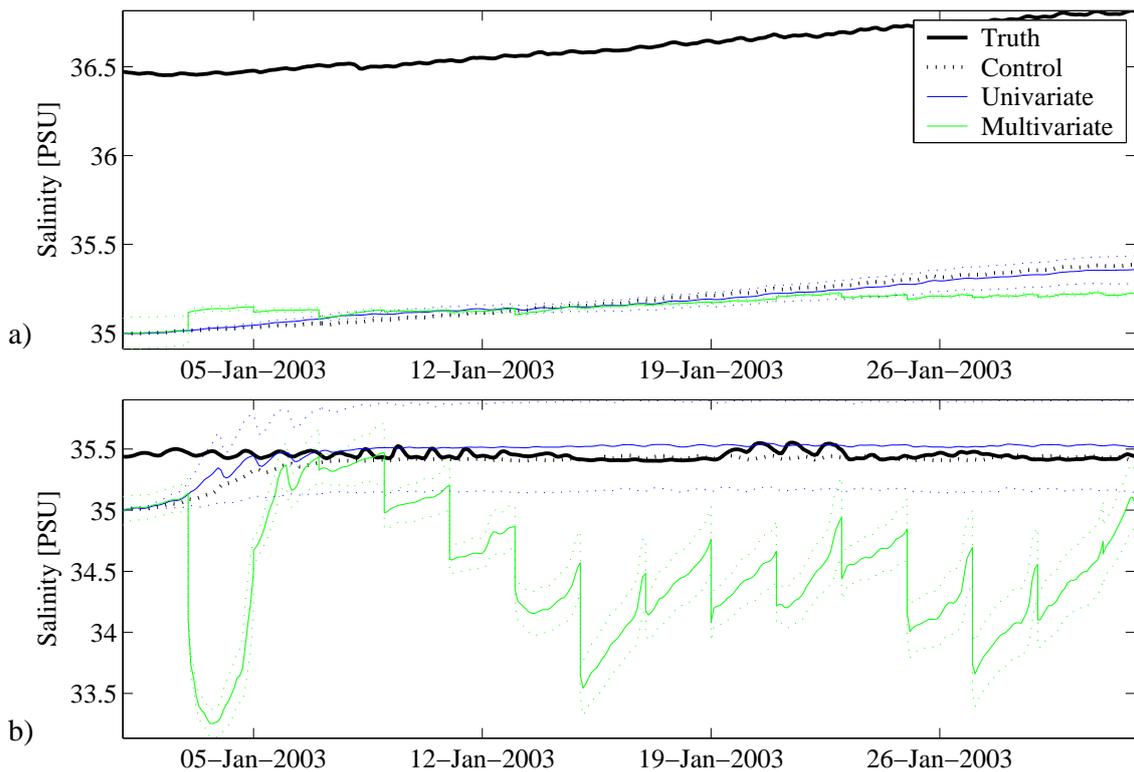


Figure 5.3: Time series plots of salinity at a) site C in the centre of PPB and b) site B near the open boundary (refer to figure 4.1) for the 20 ensemble member results. The coloured dotted lines either side of the ensemble means indicate the standard deviation of the ensemble spread.

the centre of PPB and site B near the open boundary in Bass Strait. At both locations the univariate ensemble mean mimics the control, explaining their similar RMSE values, but in the multivariate simulations the ensemble mean is adjusted away from the truth at analysis times, explaining the larger RMSE values. This is exacerbated near the open boundary where the larger standard deviation, due to the introduction of forcing error at the open boundary. This allows larger magnitudes of analysis steps in Bass Strait than in the centre of PPB, where the salinity uncertainty is less and the ensemble spread is correspondingly lower.

While salinity varies locally in response to evaporation, precipitation and fresh-water inputs, the poor agreement is largely due to the fact that background salinity is controlled by the salinity imposed at the open boundary. Since mass exchange between PPB is a much slower process than energy exchange, it takes considerably longer for the salinity forcing data at the open boundary to impact on the model when compared with the sea level forcing data: months rather than hours.

This process of the open boundary resetting the background salinity can be seen in figure 5.4 which maps the difference between the true salinity and the ensemble mean salinity at the end of the experiment, 31 January 2003. Within PPB both panels have a similar pattern with the smallest differences towards the south of PPB and largest differences in the Geelong Arm. In Bass Strait there is a negligible difference from the truth.

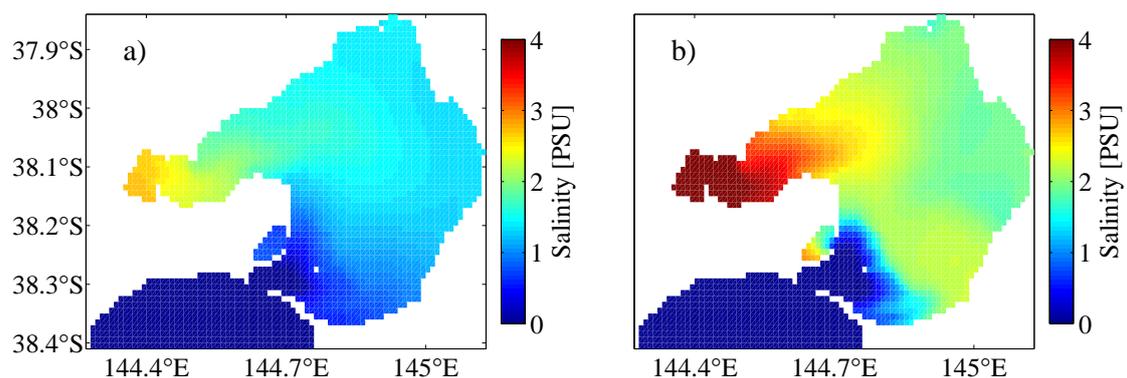


Figure 5.4: Plots of surface salinity difference on 31 January between the ensemble mean of the 20 member assimilation runs and the truth for a) univariate, and b) multivariate assimilation runs.

Sea level

In the univariate assimilation, sea level is not updated and the average RMSE (table 5.1) is less than 0.01 m. As figure 5.2c indicates, all runs start with a RMSE of about 0.25 m. This rapidly reduces as the forcing data at the open boundary establishes the true hydrodynamic regime, irrespective of the initial conditions. While each ensemble member is driven by a different realisation of the perturbed forcing data, the expected values of the forcing data are the true forcing data, and thus the ensemble mean approaches the truth. After about one day the ensemble mean sea level has stabilised to the truth and the RMSE remains low through to the end of the assimilation for the control and the univariate assimilation runs. However, similar to salinity, the sea level RMSE increases during the analysis of the multivariate assimilation, indicating that the states are adjusted away from the truth. This is due to the analysis using erroneous covariances between the sea level and the temperature, an artefact of sampling error in the ensembles, compounded by the small correlation between the variables. Increasing the ensemble size reduces the sampling error, but for the computationally realistic ensemble sizes shown the univariate assimilation remains more accurate.

A difficulty was that the multivariate assimilation analysis occasionally induced numerical instability. This was due to the analysis adjustment of the sea level, which in Bass Strait led to an elevation discontinuity at the open boundary. The subsequent propagation of this discontinuity within the model produced the numerical instability. This was another consequence of the multivariate analysis using erroneous covariances. The numerical instability was remedied by adjusting the boundary forcing data post-analysis, adding an amount equal to the difference at the analysis time between the boundary value and the model states. This adjustment was reduced linearly over 2-day period so that the applied open boundary sea level forcing gradually relaxed to the actual open boundary sea level forcing values.

Temperature

Assimilation significantly improves the temperature forecast. This is obvious from the reduction of RMSE for all runs relative to the control indicated in table 5.1. In the absence of assimilation the RMSE of the control run gradually reduces, indicating

that the control approaches the truth. This is because both use the same forcing data set. However, this occurs over a long time frame (of the order of months), akin to the open boundary resetting the salinity, and would not occur in reality where forcing error exists.

The improvement of the univariate assimilation over the multivariate assimilation for the 20 member ensemble is contrary to expectation. The analysis scheme is state independent: for the same innovations and ensemble members the same analysed temperature fields should result, irrespective of whether the unobserved variables are updated or not. At the first analysis (figure 5.2) the multivariate assimilation has a lower RMSE, but subsequently performs worse than the univariate assimilation. This suggests that feedback from the erroneous variables is inducing errors in the temperature forecast. For an ensemble size of 50 both the univariate assimilation and multivariate assimilation runs give a similar result.

5.3.2 Univariate Multivariate Comparison: Implications

Updating multiple variables of the PPB model through the assimilation of SST observations was not found to improve the forecasting skill of the model. On the contrary, the forecasting skill was reduced. Clearly, the multivariate covariances generated by the model are poor. Either there is some covariance between temperature and the other variables that is being swamped by sampling noise, or there is no covariance but sampling error is generating one artificially. Increasing the ensemble size reduces the error slightly. The consequence of this is that for the remainder of the thesis univariate assimilation will be pursued with an ensemble size of 20. There is little benefit in attempting univariate assimilation with an ensemble size of 50 to justify the additional computational expense. As univariate assimilation is pursued, subsequent discussion will focus on temperature.

If information of the state of another variable was required, other than increasing the ensemble size, one possibility would be to determine a minimum correlation, or covariance, threshold below which all correlations are statistically equivalent to zero. No updating would be performed on any states found to have a correlation lower than the threshold, the value of which would vary with the uncertainty of the states in question and with the ensemble size. This idea finds resonance with the

distance-dependent filtering proposed by Hamill *et al.* (2001) where updating was not performed beyond a certain distance where the state correlations had reduced and become indistinguishable from noise. Alternatively, the direct observation and assimilation of another state such as salinity would bring large improvements to its forecast accuracy. For instance, the literature review indicated that many studies have shown that assimilating sea level observations improves its forecast accuracy.

5.4 Model Bias Experimentation

The ensemble equations assume an unbiased forecast model, yet analysis in section 4.2 has shown that the PPB model is warm biased. This discrepancy is explored by comparing the assimilation results using a biased model against those using an unbiased model. The biased assimilation simulation is made by assimilating the observations, derived with the Zillman heatflux module, into the model forecast, which uses the Gill heat flux module. The unbiased assimilation simulation is made by assimilating observations, derived with the Gill heatflux module, into the model forecast, which also uses the Gill heatflux module. The differences between the two heatflux modules were described in section 4.3.3. Although in reality only one control simulation is made (with a model using the Gill heatflux module), as there are two synthetic truths (biased and unbiased) against which the control is compared, there are two average RMSE control simulation values.

5.4.1 Model Bias Experimentation: Results

Table 5.2 lists the results for both the biased assimilation, unbiased assimilation and control simulations. While it is logical that assimilation with an unbiased model should give a better result than assimilating with a biased model, the table shows that both give significantly improved results relative to the control simulation.

The performance of both runs is assessed further by considering the time series of temperature forecast at site C (refer figure 4.1). This is shown in figure 5.5. Initially, the ensemble mean of both the biased and unbiased assimilation simulations follow the control (panel a). At the first analysis step, temperature is adjusted towards the truth, while the control simulation remains on its original trajectory. Both the

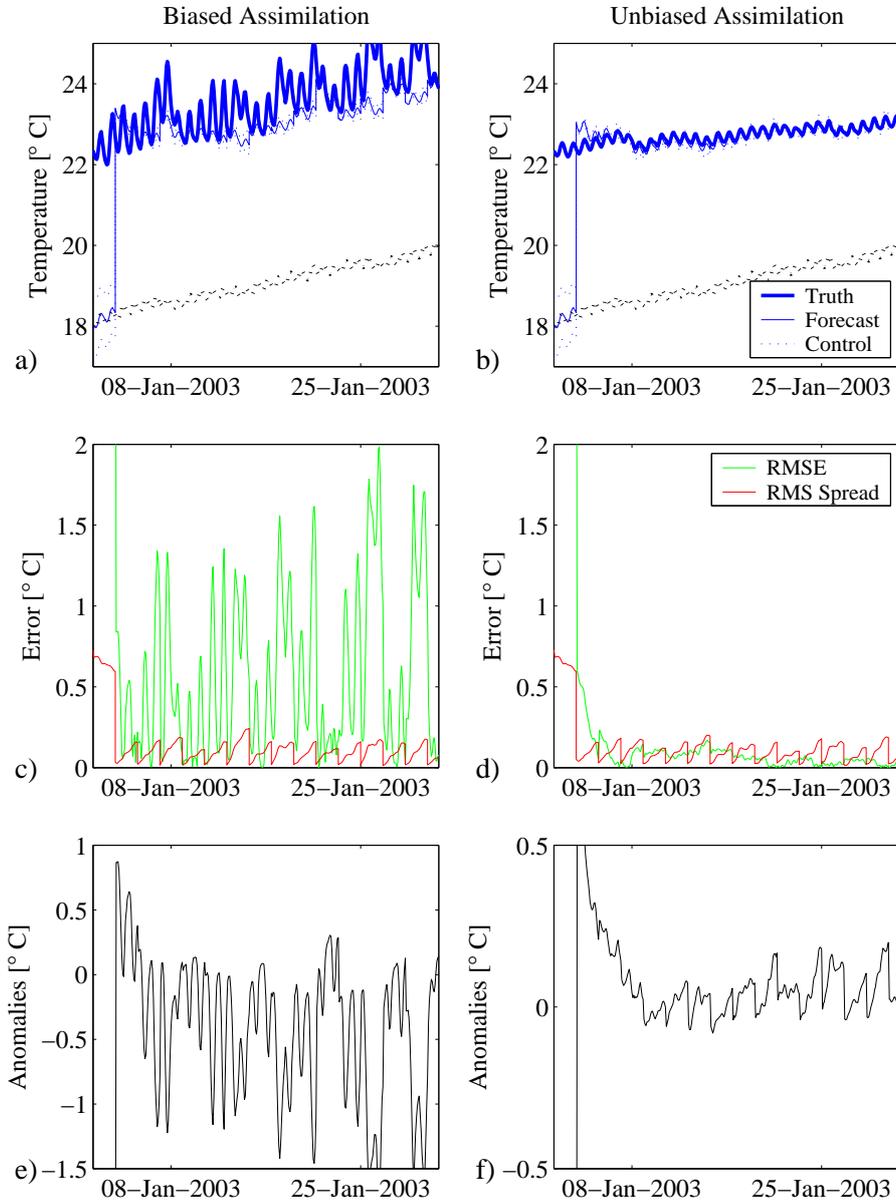


Figure 5.5: Plot of temperature error and ensemble spread at site C (refer figure 4.1) cell near the centre of PPB for biased and unbiased assimilations. The upper panels present the truth, ensemble mean, and control. The middle panels indicate the RMS spread and the difference between the ensemble mean and the truth, the lower panels presents the difference between the RMSE and the RMS spread. Note different scale for panels e and f.

Table 5.2: Average RMSE between the ensemble mean and the truth for various model configurations.

Description	Temperature [$^{\circ}\text{C}$]
Control Unbiased	3.21
Unbiased Assimilation	0.55
Control Biased	3.14
Biased Assimilation	0.69

biased and unbiased simulations then follow the truth, although the biased model follows a slightly different trajectory and requires ongoing correction. The unbiased case, driven by the same model, follows the truth more closely and requires only minor correction. For this reason the RMSE of the unbiased assimilation is lower.

An assimilation scheme is performing well if the RMSE is of the same order as the RMS spread (Keppenne, 2000). If this holds the RMS spread then provides an accurate measure of the uncertainty of the ensemble mean. This occurs in the unbiased model panel d (of figure 5.5), where the RMS spread matches the RMSE well, but not in the biased run (panel c). This is more obvious in panels e and f, where the difference between the RMS spread and the RMSE is plotted. Towards the end of the unbiased run, the RMS spread over estimates the RMSE: the truth and the assimilation are driven by the same model and forcing data set and so converge. In reality this would not occur as model error and forcing error would conspire to separate the forecast from the truth.

The RMS spread of biased assimilation underestimates the RMSE. This is because the ensemble mean of the biased model deviates from the truth, driven by different model equations. Figure 5.5c shows that the RMSE varies with a diurnal cycle due to the different diurnal heating and cooling rates of the different heatflux modules, while the RMS spread varies according to the two-day analysis cycle. As the RMS spread is underestimated, the biased analysis cannot be optimal.

5.4.2 Normality of Innovations

The performance of the assimilation is further analysed by investigating the innovations. The requirement that model forecasts and observations are unbiased and

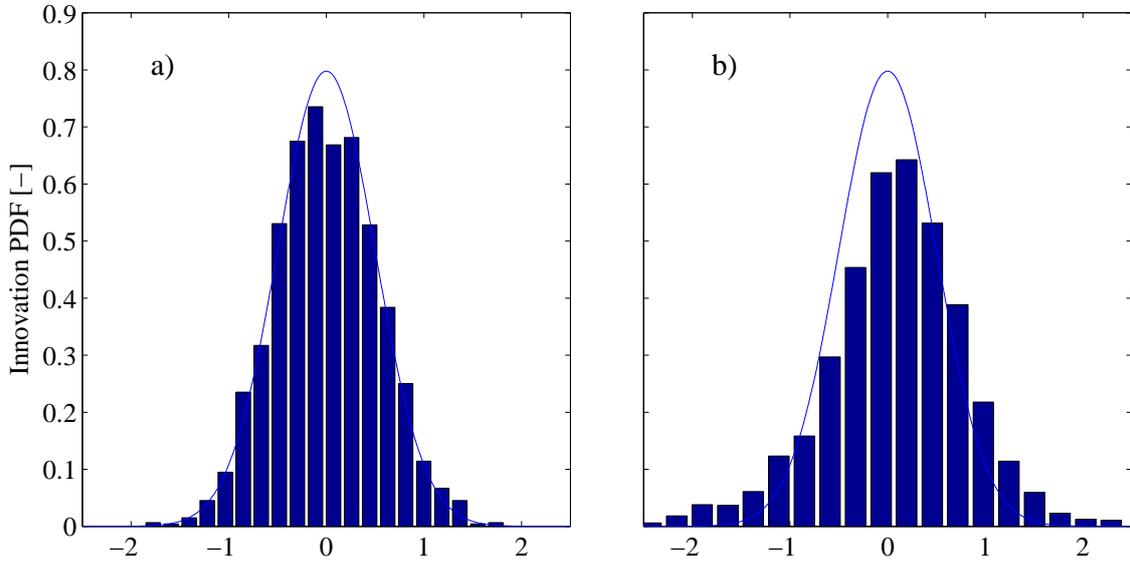


Figure 5.6: Histogram of innovations on 17 January 2003 of a) an unbiased assimilation, and b) a biased assimilation. The plots are overlaid by a normal distribution with a mean of zero and standard deviation of 0.5°C .

that the error is normally distributed leads to the result that the innovations

$$\mathbf{d} - \mathbf{H}\bar{\mathbf{x}}$$

should be normally distributed (Oke *et al.*, 2002) with a mean of zero and a standard deviation equal to the square root of the observation variance plus the forecast variance. In the case of an unbiased perfect model, the forecast error disappears and the standard deviation of the innovations approaches 0.5°C , the standard deviation of the observation error. A brief inspection of histograms of the innovations, examples are displayed in figure 5.6, indicates that both the biased and unbiased innovation distributions are unimodal, and centered about a mean of zero suggesting that the distribution of innovations are normal.

The normality of the innovations is formally assessed using a Kolmogorov-Smirnov test. This test concerns a statistic D , which is the maximum absolute difference between the cumulative distribution functions of the two distributions being compared. In this case, the distribution of the innovations and a normal distribution with a mean of zero and standard deviation of 0.5°C . The significance of D is given by the

Table 5.3: Results of the normality of the biased and unbiased innovation testing using the Kolmogorov-Smirnov method. D is the maximum difference between the innovation cumulative distribution function and a normal $N(0,0.5)$ cumulative distribution function. Q_{K-S} is the probability that both distributions are realisations of the same process. The significance indicates whether they are likely to be from the same distribution, using a 97.5% cutoff threshold.

Date	Unbiased Assimilation			Biased Assimilation		
	D	Q_{K-S}	Signif.	D	Q_{K-S}	Signif.
03 Jan	0.789	0	No	0.78	0	No
05 Jan	0.230	0	No	0.11	0	No
07 Jan	0.146	0	No	0.50	0	No
09 Jan	0.143	0	No	0.06	0	No
11 Jan	0.077	0	No	0.09	0	No
13 Jan	0.029	0.026	No	0.25	0	No
15 Jan	0.058	0	No	0.34	0	No
17 Jan	0.023	0.15	No	0.10	0	No
19 Jan	0.011	0.90	No	0.22	0	No
21 Jan	0.034	0.01	No	0.53	0	No
23 Jan	0.019	0.35	No	0.27	0	No
25 Jan	0.050	0	No	0.28	0	No
27 Jan	0.037	0.00	No	0.54	0	No
29 Jan	0.033	0.01	No	0.24	0	No
31 Jan	0.030	0.02	No	0.33	0	No

function

$$Q_{K-S}(\lambda) = 2 \sum_{j=1}^{\infty} (-1)^{j-1} \exp^{-2j^2\lambda^2}, \quad (5.2)$$

where

$$\lambda = (\sqrt{m} + 0.12 + 0.11/\sqrt{m})D, \quad (5.3)$$

and m is the number of observations. Q_{K-S} returns the probability that the distributions are two realisations of the same underlying distribution and statistically the same. Further details on the Kolmogorov-Smirnov test are available from Press *et al.* (1992, pages 623-627). The test was applied to the biased and unbiased assimilations and the results are displayed in table 5.3. Neither the unbiased or biased assimilation innovations are statistically identical to the normal distribution.

The Kolmogorov-Smirnov test may be too severe and so the innovations are tested again. This time the difference between the innovation mean and zero, and the difference between the innovation standard deviation and 0.5°C are considered to determine if the differences are statistically significant or not.

Student's t test is used for investigating the significance of the mean. The statistic t is given by

$$t = \frac{\bar{x} - \mu}{\frac{s}{\sqrt{m}}}, \quad (5.4)$$

where μ and s are the sample mean and sample standard deviation respectively and \bar{x} is the population mean, in this case zero. The t -distribution (with $m - 1$ degrees of freedom) gives the probability that a t value observed could have been as large by chance, i.e. that the difference between the means is not significant statistically (Spiegel and Stephens, 1998, pages 219 and 242-243).

A χ^2 test is used to investigate whether the standard deviation is significantly different from an expected value (Snedecor and Cochran, 1980, page 75), in this case 0.5°C . The statistic F is given by

$$F = (m - 1) \left(\frac{s}{\sigma} \right)^2, \quad (5.5)$$

where σ is the population standard deviation, in this case 0.5°C . F follows a χ^2 -distribution with $N - 1$ degrees of freedom. The cumulative χ^2 -distribution gives the probability that the observed F is greater than would be expected by chance. Therefore cumulative probability values greater than 0.975 and less than 0.025 indicate that it is unlikely that the distribution of the innovation is statistically equivalent to 0.5°C .

The results of these tests are given in tables 5.4. For the unbiased assimilation innovations, four of the innovation means are not statistically different from zero and five of the standard deviation are not statistically different from 0.5°C . During the later dates, the mean of the innovations remains statistically different from zero even though the innovation mean is very close to zero (less than 0.04°C difference). The large number of observations (> 2500) makes the t -test very sensitive. The failure is most likely because the ensemble mean component of the innovations is an average of only 20 ensemble members and so slight variations, due to sampling

Table 5.4: The sample mean μ and sample standard deviation s of the biased and unbiased innovations are assessed using the Student's t and χ^2 variance test. For the t -test a 'No' signifies that there is no statistically significant difference between the sample mean and zero at a 97.5% level, and for the χ^2 -test a 'No' signifies that there is no statistically significant difference between the sample standard deviation and 0.5 at a 95% level.

Date	Unbiased Model Assimilation				Biased Model Assimilation			
	Student's t		χ^2		Student's t		χ^2	
	μ [$^{\circ}$ C]	Signif.	s [$^{\circ}$ C]	Signif.	μ [$^{\circ}$ C]	Signif.	s [$^{\circ}$ C]	Signif.
03 Jan	3.66	Yes	2.58	Yes	3.68	Yes	2.18	Yes
05 Jan	0.29	Yes	0.56	Yes	0.10	Yes	0.66	Yes
07 Jan	0.18	Yes	0.54	Yes	0.77	Yes	0.73	Yes
09 Jan	0.19	Yes	0.56	Yes	-0.03	Yes	0.65	Yes
11 Jan	0.08	Yes	0.55	Yes	-0.08	Yes	0.69	Yes
13 Jan	0.01	No	0.55	Yes	0.29	Yes	0.70	Yes
15 Jan	0.05	Yes	0.53	Yes	0.43	Yes	0.65	Yes
17 Jan	0.01	No	0.53	Yes	0.06	Yes	0.72	Yes
19 Jan	-0.00	No	0.51	No	0.27	Yes	0.70	Yes
21 Jan	-0.03	Yes	0.53	Yes	0.83	Yes	0.78	Yes
23 Jan	-0.01	No	0.51	No	-0.46	Yes	0.76	Yes
25 Jan	-0.06	Yes	0.51	No	0.35	Yes	0.67	Yes
27 Jan	-0.04	Yes	0.52	Yes	0.85	Yes	0.71	Yes
29 Jan	-0.03	Yes	0.50	No	-0.35	Yes	0.73	Yes
31 Jan	-0.03	Yes	0.50	No	0.47	Yes	0.71	Yes

errors, will cause the ensemble mean to deviate from the truth, and the failure of the test.

The results for the biased univariate assimilation innovations are also presented in table 5.4. Neither the means nor the standard deviations of the innovations are statistically similar to the expected values. The average mean of the innovations after 3 January is 0.25° C, while the average standard deviation of the innovations is 0.7° C.

5.4.3 Model Bias Experimentation: Implications

The findings of the investigation of the effect of model bias on the assimilation results are that

1. both biased and unbiased assimilation forecasts give improvement over an unassimilated (or control) forecast,
2. as expected, using an unbiased assimilation model is more accurate than using a biased assimilation model,
3. in the case of a biased model assimilation, the RMSE is not well predicted and is underestimated by the RMS spread: the ensemble spread gives a poor estimate of the uncertainty of the ensemble mean forecast,
4. the failure of the unbiased ensemble innovations to approach expected values was due to sampling error in the ensemble forecast on account of the relatively small ensemble size, and
5. the failure of the biased ensemble innovations to approach expected values was due to the model bias indicating the suboptimal performance of the analysis.

The implication of these findings for coastal data assimilation and the PPB case study are that because the RMSE is underestimated in the biased assimilation, its analysis is not as responsive as for an unbiased model. In a very dynamic system this may lead to problems, however, the temperature of PPB does not change drastically and so long as the assimilation is every two to three days the forecasting of average conditions no difficulties are anticipated. These findings also suggest that while an unbiased model is preferable, the PPB system is fairly tolerant of a biased model as long as regular assimilation is performed: the forecast will be accurate relative to an unassimilated forecast. This gives confidence that the extension for assimilation of actual SST data and the replication of real conditions using the warm biased PPB model should be possible.

5.5 Analysis Method

As ensemble size is constrained by cost considerations, the most efficient filter is sought for the analysis. The literature review (section 2.5.2) noted that most ensemble filters are forms of square root filters and can be classified as stochastic or

deterministic. A stochastic filter contains random features such as perturbed observations to maintain correct ensemble spread, whereas a deterministic filter maintains the correct ensemble spread using additional (deterministic) equations. The experiments conducted so far in this thesis have used an EnSRF: a deterministic filter.

The promoted benefit of deterministic ensemble filters is that they are not affected by additional sampling error from perturbed observations; all ensemble filters are susceptible to sampling errors in the ensemble forecast, leading to errors in the forecast error covariances¹. As deterministic filters do not use perturbed observations they are expected to outperform stochastic filters, especially for small ensembles where sampling error of the perturbed observations is expected to have a larger impact as fewer random numbers are added to the observations, the chances of abnormal distributions being realised increases. With more random numbers the sample distribution approaches the population distribution and sampling error reduces.

The efficiency of deterministic and stochastic filters are contrasted in this section by comparing the results of assimilation simulations using the EnKF and the EnSRF. This is achieved by testing both filters at a range of ensemble sizes. The experimental analysis compares the filter performance in terms of RMSE and RMS spread. Another comparison is made by investigating the analysis weighting matrices of each filter. These give insight into the behaviour of the ensemble filters. The biased model forms the basis of these experiments. While the analysis will not be as optimal as using an unbiased model, considering the known model temperature bias, the results should be more realistic.

5.5.1 Analysis Method: Results

With an ensemble size of 20 both the EnKF and the EnSRF give similar results. Figure 5.7 illustrates this with a time series of the RMSE for both filters. Both RMSE time series mirror each other with very little to separate them. Initially, the EnSRF has the lower RMSE, while afterwards the EnKF has the lower RMSE.

A spatial comparison of the two filters (figure 5.8) shows the temperature in

¹This is a separate issue and is related to ensemble size. It was discussed previously when comparing the univariate to multivariate assimilation in section 5.3.

the surface layer at the first analysis (3 January) for the pre- and post-analysis forecasts together with their anomalies. On account of the degraded (lowered) initial temperature, the temperature in PPB is greatly underestimated by between 2°C and 5°C. The temperature in Bass Strait is closer to the truth, because its temperature is controlled by the open boundary forcing, which is the same as for the synthetic truth. This does not occur in PPB because i) the atmospheric temperature changes occur more slowly than across an open boundary, ii) bias in the forecast heatflux module drives the temperature in a different direction from the synthetic truth, and iii) low mass exchange through The Heads limits the impact of Bass Strait temperature on PPB temperature. The post-analysis plots (panels e–h) illustrate the improvement the assimilation makes to the forecast, the post-analysis anomalies are spread about zero. While there are small differences between the results, both forecasts and anomalies for the EnKF and EnSRF give a similar spatial pattern.

Table 5.5 lists the RMSE of assimilation runs for both the EnKF and EnSRF with ensemble sizes of 10, 20 and 50. Relative to the ensemble size of 20, the performance of both filters decreases when the ensemble size is reduced to 10 ensemble members with the EnKF performing worse than the EnSRF with a significantly larger RMSE. This supports the hypothesis that when using a smaller ensemble size the EnKF is subject to sampling error. However, at such a low ensemble size

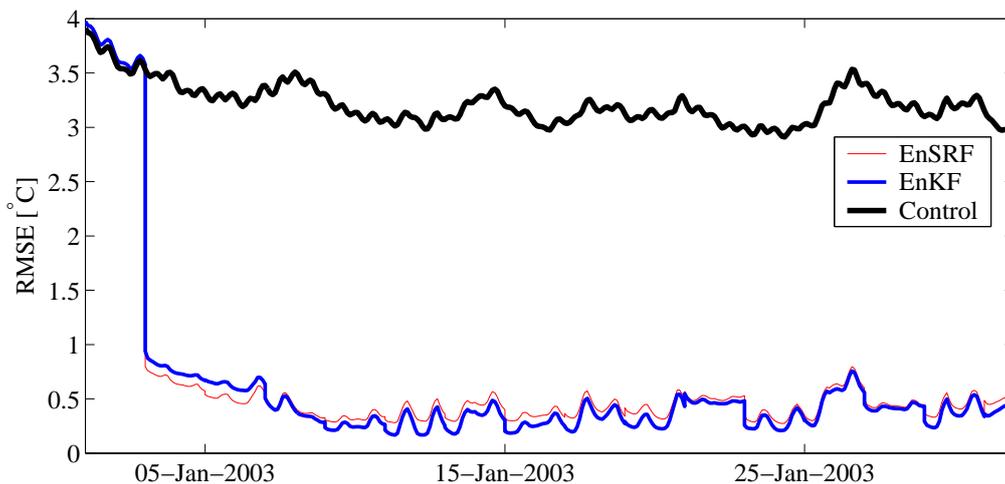


Figure 5.7: Plot of RMSE for the EnSRF, EnKF, and the control for the biased univariate case using 20 ensemble members.

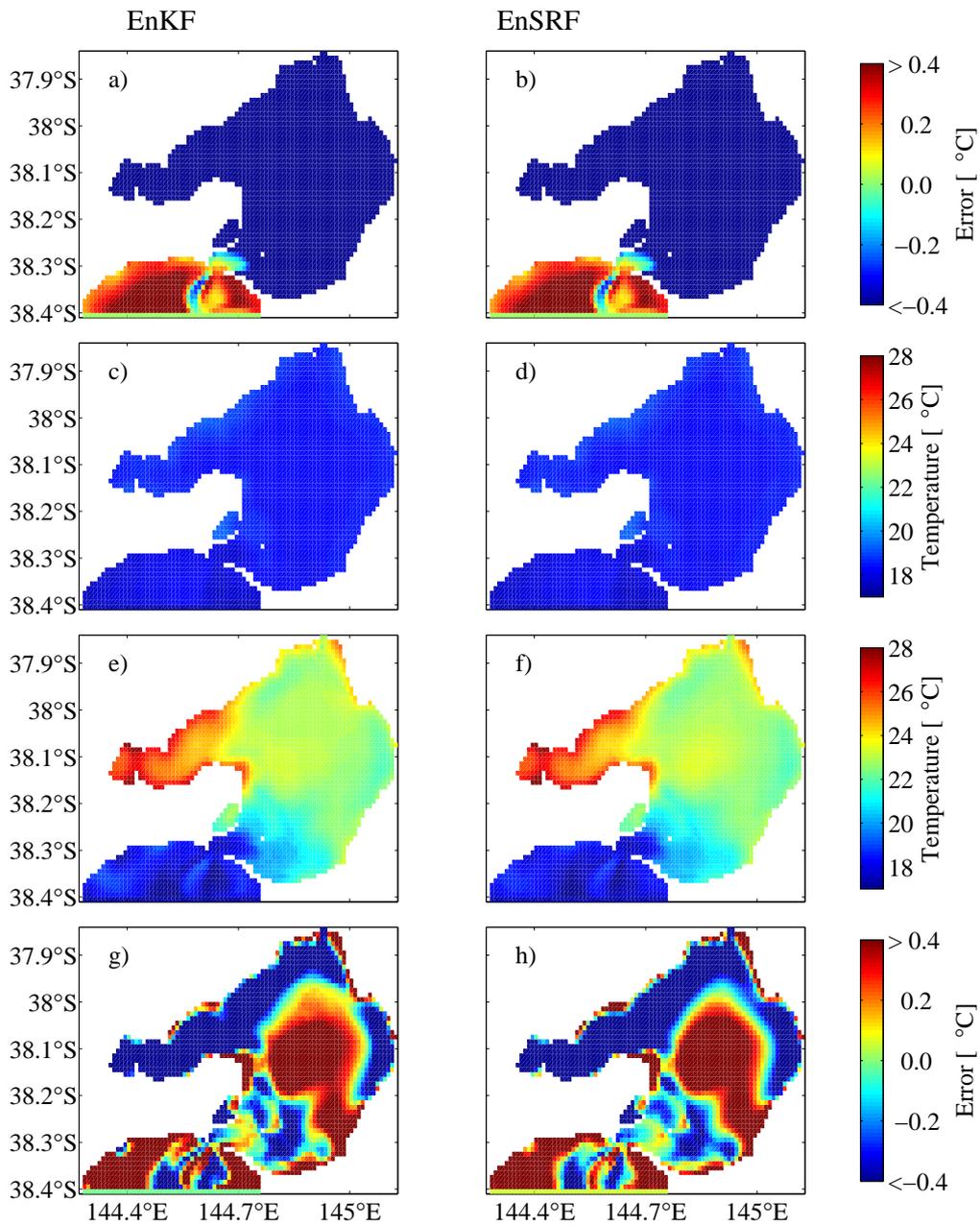


Figure 5.8: Surface snapshot of assimilation results on 3 January 2003 at the first analysis. The left-hand column are the results for the EnKF and the right-hand column for the EnSRF. Panels a) and b) show the ensemble mean forecast error, the ensemble mean forecasts minus the truth; panels c) and d) show the ensemble mean forecasts; panels e) and f) show the ensemble mean analyses; and panels g) and h) show the ensemble mean analysis error.

Table 5.5: Spatially and temporally averaged RMSE temperature for the EnSRF and the EnKF at a range of ensemble sizes.

Ensemble Size	EnSRF [°C]	EnKF [°C]
10 members	0.91	1.00
20 members	0.69	0.63
50 members	0.68	0.77

neither filter performs optimally. An ensemble size of 10 is too small to adequately characterise the temperature variation within PPB. At an ensemble size of 50, the RMSE of the EnSRF reduces slightly, whereas the RMSE of the EnKF unexpectedly increases. The reason for this was unclear but was thought to be related to the perturbations added to the ensembles, which would make the EnKF less stable. This possibility was discounted by repeatedly analysing the same set of forecast ensemble members with the EnKF, using a different realisation of perturbed observations each time. The results (not shown) were that the perturbed observations have very little effect on the resulting analysis states. This is possibly due to the large number of observations (> 2500) contained in each synthetic image, which would reduce the possibility for sampling error in the perturbations.

The discrepancy between the filter results in table 5.5 was explored by repeating the 20 ensemble member simulations, but instead using a different set of perturbed forcing data (and different perturbed observations for the EnKF). For each ensemble size tested (10, 20 and 30) in the experiments shown in table 5.5, the same forcing data set was used for both the EnKF and EnSRF assimilations. By using a different set of perturbed forcing data the effect of external factors on the filters can be assessed. When a different set of forcing data was used, the results were that the RMSE of the EnSRF was 0.95°C (previously 0.69°C), while the RMSE of the EnKF increased to 0.65°C (previously 0.63°C). Evidently the variation in the forecast error, through the perturbed forcing (and model nonlinearity), which affects the forecast error covariance matrix has a greater impact on the resulting analysis than the type of analysis method used.

While these variations would be reduced by increasing the ensemble size, it is calculated that there would be no substantive reduction in RMSE. The differences between the 10 and 20 ensemble member runs in table 5.5 are greater than the

differences between the 20 and 50 ensemble member runs, pointing to a convergence of results with increasing ensemble size. Increasing the ensemble size would merely reduce the variance of different RMSE values resulting from the use of a different perturbed forcing set.

However, there is clearly a difference between these two filters, as the differing results, when applied to the same conditions indicates (table 5.5). These differences are explored further by considering the shape of the ensemble filters.

5.5.2 Ensemble Filter Analysis

The experiments have indicated that there is a slight difference between the EnKF and the EnSRF, either due to sampling errors or their intrinsic structure. These differences can be investigated by rearranging the filter equations and considering some of the resulting matrices.

The EnSRF analysis, equation (2.30), can be rearranged in the form

$$\bar{\mathbf{x}}^a = \bar{\mathbf{x}}^f + \mathbf{X}'^f \mathbf{z}, \quad (5.6)$$

where \mathbf{z} is a $n_e \times 1$ vector given by

$$\mathbf{z} = \mathbf{X}'^{fT} \mathbf{H}^T [\mathbf{H} \mathbf{X}'^f \mathbf{X}'^{fT} \mathbf{H}^T + (n_e - 1) \mathbf{R}]^{-1} [\mathbf{d} - \mathbf{H} \bar{\mathbf{x}}^f]. \quad (5.7)$$

The analysis of the ensemble mean can be seen as the addition of a linear combination of the forecast anomaly vectors to the ensemble mean forecast, with \mathbf{z} as the weighting vector. Similarly for the ensemble spread equation (2.34),

$$\mathbf{X}'^a = \mathbf{X}'^f \mathbf{Z} \quad (5.8)$$

expresses the reduction in ensemble spread through a weighting matrix, \mathbf{Z} . The ensemble spread is analysed as a linear combination of the forecast anomaly vectors. The ensemble analysis performed by the square root filter can therefore be characterised by two weighting factors: a vector \mathbf{z} , which adjusts the ensemble mean through equation (5.6), and a matrix \mathbf{Z} , which adjusts the ensemble spread via equation (5.8).

In the EnKF both these operations — mean adjustment and ensemble spread reduction — are performed by a single weighting matrix $\mathbf{\Lambda}$, given in equation 4.6, with the EnKF equation (2.20) becoming

$$\mathbf{X}^a = \mathbf{X}^f + \mathbf{X}'^f \mathbf{\Lambda}. \quad (5.9)$$

Again the analysis is performed by adding a linear combination of the forecast anomalies to the forecast, this time with $\mathbf{\Lambda}$ as the weighting matrix.

Equation (5.9) can be rearranged to form the EnKF equivalents to equations (5.6) and (5.8). When this is done the EnKF equivalent to \mathbf{z} is given by

$$\mathbf{y} = \mathbf{\Lambda} \mathbf{j}_{n_e}, \quad (5.10)$$

where \mathbf{j}_{n_e} is a $n_e \times 1$ vector with each element of value $\frac{1}{n_e}$, and the EnKF equivalent to \mathbf{Z} is given by

$$\mathbf{Y} = \mathbf{I} + \mathbf{\Lambda} - \mathbf{\Lambda} \mathbf{J}_{n_e}, \quad (5.11)$$

where \mathbf{J}_{n_e} is an $n_e \times n_e$ matrix with each element of value $\frac{1}{n_e}$. The derivation of these relationships is given in appendix C.

Differences between the EnKF and EnSRF will present themselves in the differences between \mathbf{y} and \mathbf{z} for the ensemble mean and between \mathbf{Y} and \mathbf{Z} for the ensemble spread. By calculating these weighting vectors and matrices for the same forecast ensemble, sampling error, due to the perturbed observations, should be more apparent.

The weighting vectors and matrices are presented using the forecast on 3 January. The results are presented in figure 5.9 for the 10 ensemble members experiments. The values from the 10 ensemble member experiments are displayed because the matrices are smaller and it is easier to visualise differences. Also the effects of any perturbed observation sampling error will be more obvious.

The effect of perturbed observations on the ensemble mean can be seen in figure 5.9a where the weighting for some elements of \mathbf{y} (4 and 5) has increased and for others (8, 9 and 10) it has decreased relative to the EnSRF weights. Moreover, for the ensemble spread \mathbf{Y} is fairly unstructured when compared to \mathbf{Z} , which has a dominant diagonal feature. This diagonal feature is due to the subtraction from the

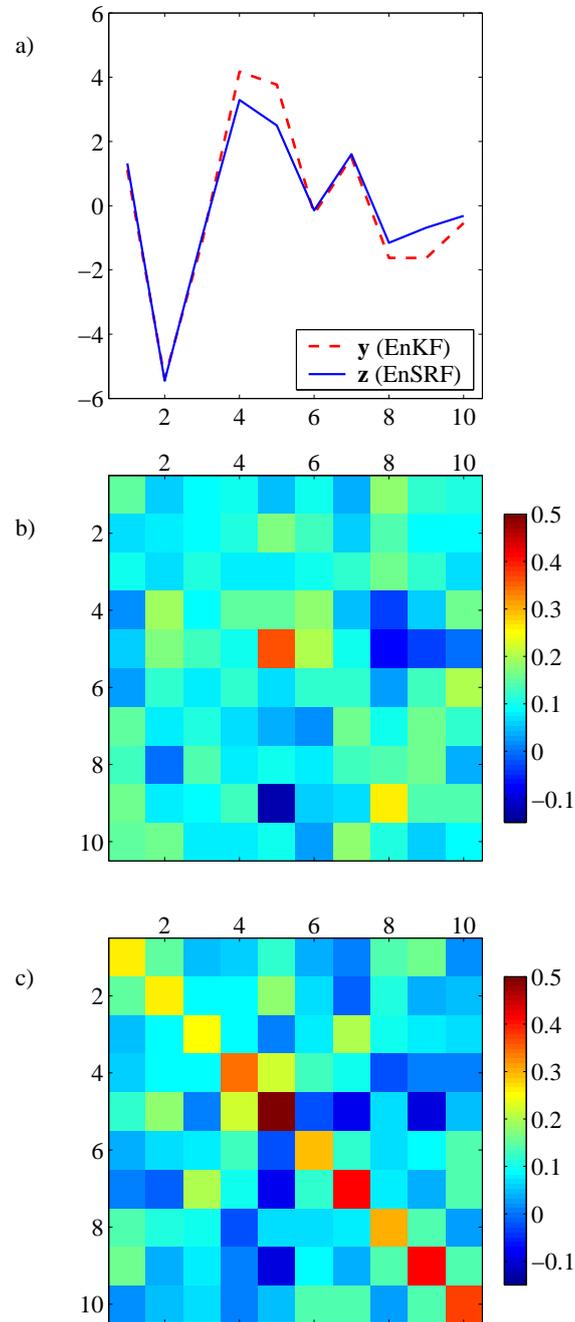


Figure 5.9: Plots showing a) EnKF ensemble mean weighting vector \mathbf{y} (dashed line) with the EnSRF equivalent \mathbf{z} (solid line), b) the EnKF weighting matrix \mathbf{Y} , and c) the EnSRF ensemble spread reduction matrix \mathbf{Z} .

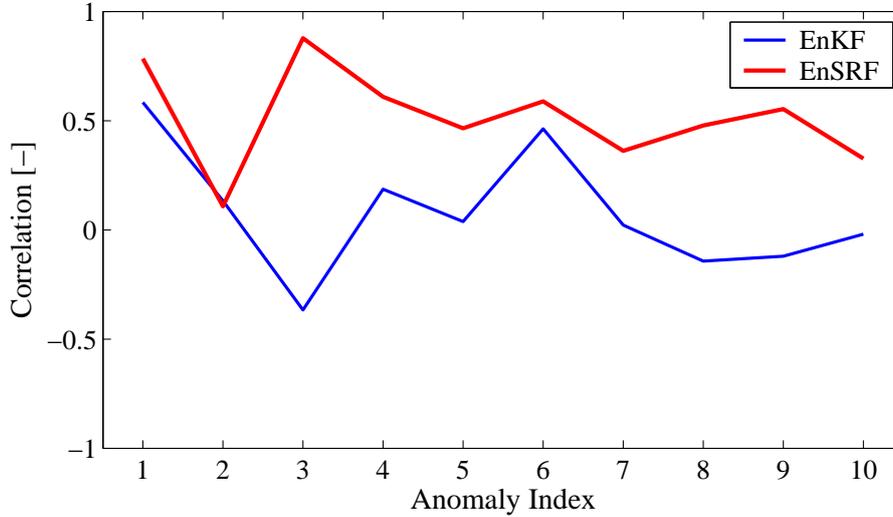


Figure 5.10: Spatial correlation between analysed and forecast anomalies for EnSRF and EnKF.

identity matrix in equation (2.35).

These observations have implications on the shape of the resulting analysed anomaly vectors. In the case of the EnKF, the resulting analysis anomaly vectors should consist of a somewhat random combination of the forecast anomalies reduced by their respective weights. However, in the case of the EnSRF, the shape of the i^{th} analysed anomaly is based on the i^{th} forecast anomaly vector reduced by the i^{th} diagonal value of \mathbf{Z} , with minor adjustments from the other forecast anomalies weighted by their corresponding off-diagonal values. The result of this is that the analysis anomalies generated by the EnSRF will be more closely aligned to the forecast anomalies than those generated by the EnKF.

This finding is of importance if anomaly independence is sought. While neither filter preserves the anomaly shape perfectly — the off-diagonals of \mathbf{Y} and \mathbf{Z} are non-zero — the EnSRF does not distort the analysed anomalies as much as the EnKF. This is evidenced in figure 5.10, which plots the anomaly spatial correlation between the i^{th} forecast and analysis anomaly for all 10 anomalies. The higher correlations for the EnSRF analysis anomalies indicates that the EnSRF filter maintains the anomaly shape better. Figure 5.11 displays this by plotting the third forecast anomaly against the corresponding analysis anomalies of the EnKF and the En-

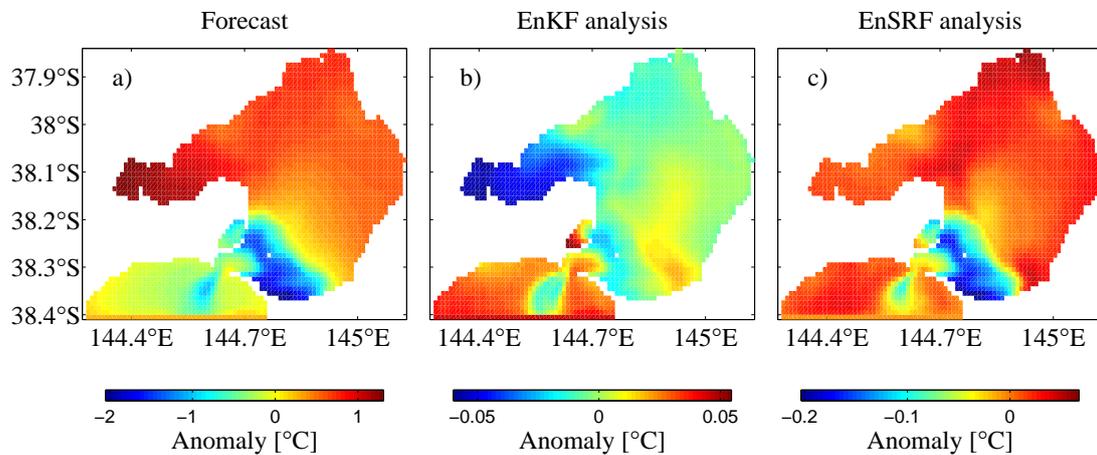


Figure 5.11: Changes in vector shape of third ensemble member through the application of the EnKF and the EnSRF. Note different scale for each image.

SRF. The third anomaly is shown because it displays this difference most vividly (see figure 5.10).

5.5.3 Analysis Method: Implications

The findings of the investigation of the effect of the analysis method on the assimilation results are that

1. there is little difference between the analysis methods in terms of forecast accuracy,
2. the internal structure of the filters is different, which impacts on the resulting ensemble anomalies, with the EnSRF analysis anomalies resembling the forecast anomalies closer than the EnKF ones,
3. sampling error in the forecast error covariance matrix has a larger impact on the resulting accuracy of the forecast than the filter choice, and
4. increasing ensemble size would not appear to improve the forecast accuracy greatly as there are diminishing improvements as the ensemble size increases.

The implication of these findings for coastal data assimilation and the PPB case study are that neither ensemble filter is preferred, although if the maintenance of anomaly independence is important, the EnSRF is a better choice. Furthermore, improvements to forecasting accuracy and reliability would probably be improved by focusing on the forecasting error rather than the filter choice, as the forecasting error appears to have a larger bearing on the result.

5.6 Dynamic Profile Covariances

The ensemble assimilation techniques applied in this thesis are constrained by a large computation cost. This cost, due to the propagation of each ensemble member, is proportional to the ensemble size. While the ensemble techniques are effective, a simpler technique such as Optimal Interpolation, which requires only a single model run, would be much more cost effective. The utility of ensemble techniques is in dynamic forecasting environments, where the error covariance structure changes with time. If the covariance structure does not vary in time, ensemble techniques are unnecessary and a steady-state covariance structure would suffice, reducing the computational burden. This section tests the need for a sequential filter using a dynamic forecast error covariance matrix, by displaying the changes in covariance and conditions through the water column over an assimilation simulation.

A justification for the use of ensemble techniques is given by figure 5.12. This plot presents a time sequence of the conditions through the water column at site C (refer to figure 4.1) in the centre of PPB using the EnSRF with 20 ensemble members and assimilating biased observations. The values used are the horizontal average values of 21 cells.

The upper panel, fig. 5.12a, indicates the average wind conditions, which suggest that the January 2003 was generally quite blustery and that strong vertical mixing through the water column is likely. The assimilation period is during summer so if the winds were mild, strong stratification might occur, due to the high levels of insolation. The second panel, fig. 5.12b, indicates the standard deviation of the ensemble, which decreases with depth. This is expected as uncertainty enters the model through the boundary conditions at the surface, and therefore a larger variation is expected. The regular sudden decrease in standard deviation is due to

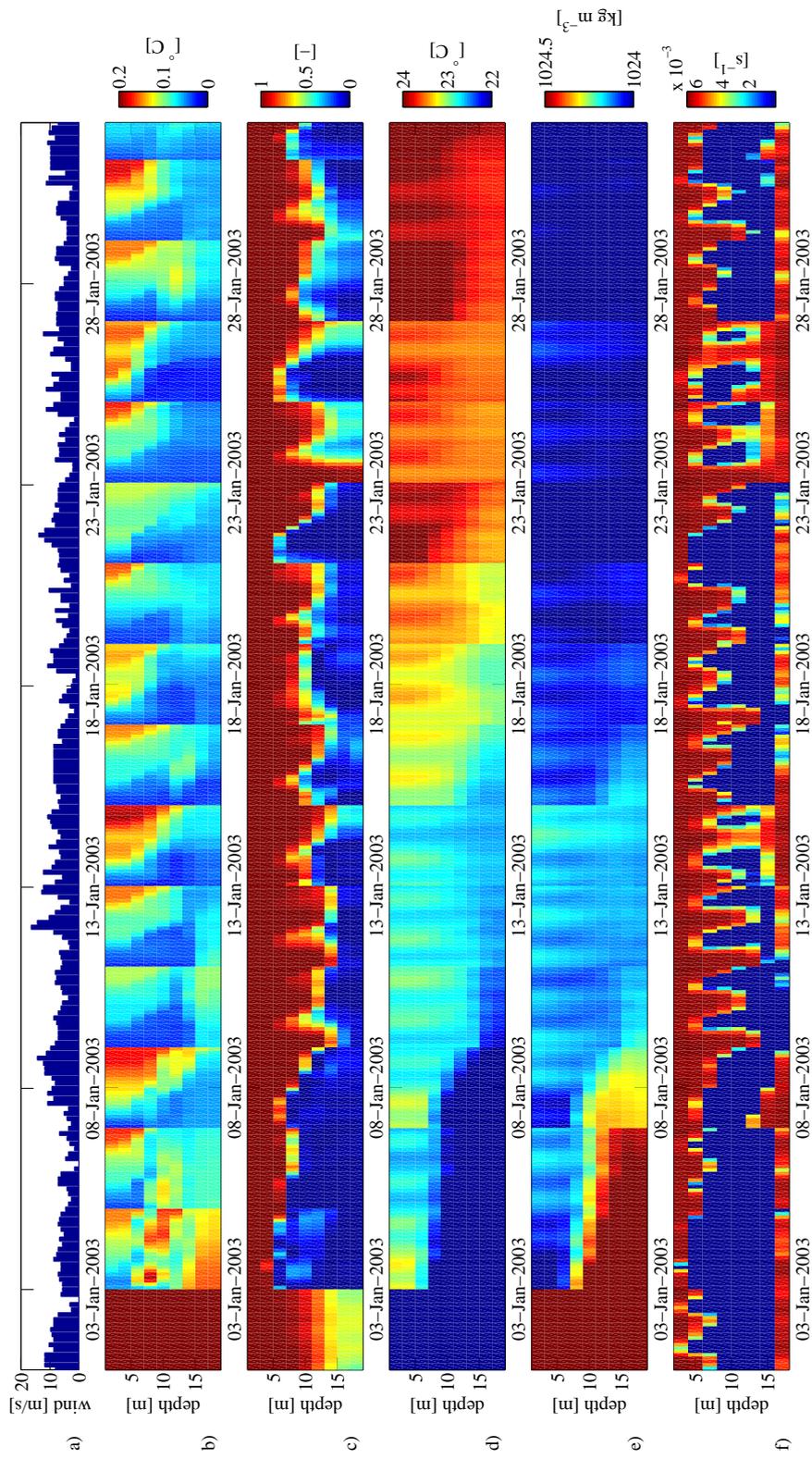


Figure 5.12: Profile time series of horizontally averaged conditions of nine cells located at site C near the centre of PPB (refer figure 4.1) indicating a) wind speed, b) ensemble standard deviation, c) correlation between the surface and depth, d) temperature, e) density, and f) stratification frequency. The 2-day assimilation cycle can easily be seen in the regular variation of the ensemble standard deviation (b).

the analysis reducing the ensemble spread.

The third panel, fig. 5.12c, shows the correlation coefficient, R^2 , between the surface and the depth. When the correlation is high the cells will be updated more during analysis, whereas when the correlation is low the cells remain untouched. In general, there is a constant high correlation in the upper layers, but at depth the correlations vary depending on conditions.

The fourth panel, fig. 5.12d, displays the temperature. The diurnal cycle can be made out as the regular temperature variation at the surface, also the gradual heating of the water column from above the surface with heat moving down the water column is seen. The period from 15–18 January shows this clearly. The density distribution, based on temperature and salinity is shown in panel fig. 5.12e. These values are used to calculate the Brunt-Väisälä (or stratification) frequency N , which is given by

$$N^2 = -\frac{g}{\rho_0} \frac{d\rho}{dz}, \quad (5.12)$$

where ρ_0 is a reference density and higher values of N indicate the strength of any stratification. Further details on this formula are available from Cushman-Roisin (1994, pp. 123-129). Where N is undefined (the frequency values are imaginary) the value is set to zero and indicates that the water is well mixed. Stratification induced at the surface by incoming solar radiation is seen as the daily rise and fall of the well mixed layer.

An example of the dynamic nature of the PPB system is the period 25–28 January. On the 25th the water column is somewhat stratified indicated by the column of well mixed water being narrower and higher up the water column than average. Correspondingly, the correlations between the surface and the deeper layers are quite low. With mixing this stratification breaks down by the 28th and well mixed conditions reestablish themselves. This variable nature would not have been picked up by an assimilation technique using a steady-state covariance structure, and confirms the need for a dynamic forecast error covariance provided by ensemble assimilation techniques.

5.7 Chapter Summary

This chapter explored various ensemble assimilation configurations, with the aim of determining a suitable configuration for the assimilation of SST into a coastal hydrodynamic model. The configurations were tested through OSSEs of the PPB case study in which the assimilation of synthetic observations was used to recover the synthetic truth from an initially degraded model.

Computational cost considerations limit the ensemble size possible. Because of this and due to the low correlation between temperature and other variables, univariate assimilation of temperature was found to outperform multivariate assimilation. Even with an ensemble size of 50 the covariances between the variables cannot be accurately established, although only 20 members are needed to characterise the temperature variation.

Some degree of model bias is regrettable, but often unavoidable. The comparison of an assimilation of a biased with an unbiased model confirmed that an unbiased model is preferable as it conforms to the ensemble assimilation assumptions. However, both assimilations (biased and unbiased) produced a clear improvement to the forecast over an unassimilated control. The main detraction of a biased model is that the assimilation scheme no longer accurately predicts the error of the forecast relative to the truth: the RMS spread no longer approximates the RMSE.

Surprisingly, given the number of ensemble filters proposed in the literature, the ensemble filter choice had far less impact on the assimilation than the other configuration options. At a small ensemble size (10), the EnSRF slightly outperformed the EnKF, but both results were degraded relative to an ensemble size of 20. With larger ensemble sizes, the differences between the forecast results were overwhelmed by external influences: differences in the perturbed forcing data set used. Differences were discovered between the shape of the ensemble anomalies, with the EnSRF anomalies resembling the forecast anomalies more closely than the EnKF anomalies.

These findings describe the best configuration of a coastal hydrodynamic model data assimilation system for incorporating SST observations into the PPB case study using the resources available for this thesis. Based on these findings the thesis can proceed to assimilate actual SST observations in the following chapter confident in the ensemble assimilation configuration adopted.

Chapter 6

SST Assimilation for PPB

The investigations of the preceding chapters are applied through the assimilation of actual SST observations from the NOAA–16 AVHRR into the PPB model. By doing so this chapter demonstrates that the assimilation techniques developed and selected in the preceding chapters function in an actual as opposed to synthetic setting, and that short-term forecasting of temperature can be improved through the assimilation of satellite observations of SST.

The accuracy of an assimilation forecast is diminished by known bias in the model and diurnal error in the observations. While the removal of model bias is important, this chapter focuses on observation error, and specifically the diurnal error induced by the differences between the difference between the skin and bulk temperatures. Two error correction approaches are tested: i) diurnal warming algorithms, and ii) conditional merging.

6.1 Introduction

This final chapter marks the verification of the thesis objectives proposing to improve temperature prediction in coastal models through the assimilation of SST. This is demonstrated through the PPB case study. The work of the preceding chapters is combined to assimilate actual SST observations into a model of PPB using ensemble assimilation techniques. Two issues that have been identified in the thesis that might lessen the results are: i) the warm bias in the model, and ii) a diurnal error in the SST

observations. While model bias is a significant issue for assimilation, this chapter will focus on observation error and its removal.

The assimilation equations assume that the model forecast is unbiased. However, the experiments of section 5.4 have shown that while an unbiased model is preferable, regular assimilation of unbiased observations into a biased model improves the forecast over an unadjusted forecast, although the analysis will not be optimal. This finding will enable the warm biased PPB model to be applied.

A further assumption of data assimilation is that observations are unbiased. If biased observations are used, then they must be dealt with in an appropriate manner. The analysis of section 3.4.3 has shown that over a long time series the SST observations are unbiased. However, a diurnal error does exist with an observation overestimating or underestimating the temperature depending on the time of day. This diurnal error appears to be due to the temperature difference between the skin and bulk temperatures of the water, and the inability of the satellite retrieval algorithm to resolve this difference.

A series of approaches to remove the observational error are attempted in this chapter. These are

1. to apply correction formulae proposed in the literature as described in section 3.4.1 following the algorithms of
 - (a) Gentemann *et al.* (2003) and
 - (b) Kawai and Kawamura (2002), and
2. to apply Conditional Merging, which adjusts the satellite data using in situ observations to improve the accuracy.

The unadjusted SST is also assimilated to gauge the improvement of the error correction approaches and the utility of assimilating spatial data is validated by contrast with an assimilation of only point scale in situ observations into the model.

6.2 Method

Following the investigations of chapter 5, a univariate assimilation with an ensemble size of 20 is used throughout this chapter. While the investigation of the previous

chapter found that the EnKF and EnSRF give similar forecast accuracy for an ensemble size of 20, the EnSRF has been chosen for performing the analysis here. For all runs the same perturbed forcing data set will be used, ensuring that any forecast differences are due to the observations. The ensemble generation, use of perturbed forcing, model configuration, and forcing data set detailed in chapter 5 will also be used in this chapter.

The simulation is conducted over the period from 18 March 2003 to 22 April 2003, which is the only period in which in situ temperature data are collected near the surface as well as at depth (see table 3.12). At other periods temperature is only monitored at depth. This is important because the surface temperature measurements allow a determination of SST error to independently test the effectiveness of the diurnal correction algorithms and a measure of observations error is necessary to perform conditional merging. Furthermore, the in situ data are accurate enough to indicate the true temperature, and as figure 3.19 shows the surface temperature varies little over at least a cell width of the model. A greater number of in situ locations allows for a more rigorous testing of the assimilation.

The temperature in the model is initially set to 19.27 °C, which is the climatology temperature value for the 18 March based on 11 years of record at West Channel Pile. Salinity is initially set to 35 PSU and the currents and sea level are set to zero. The experimentation of chapter 5 found that current and sea level values rapidly established themselves (within a few hours). Due to the file outputting procedure of MECO there is a difference of at most an hour between the model time and the observation time, the impact of this difference on the results is considered negligible.

6.2.1 Satellite Observations

Over the assimilation period AVHRR SST satellite data are available from both the NOAA-12 and the NOAA-16 platforms. As the NOAA-16 platform has a marginally more complete record, 39 available images over the period as compared to 35, it is used for the assimilation and the NOAA-12 images are kept aside to give an independent data set for validation of the resulting spatial patterns.

While the satellite data are already processed — buddy checks, sanity checks, and masking for clouds and land (refer to section 3.4) — a final inspection of the

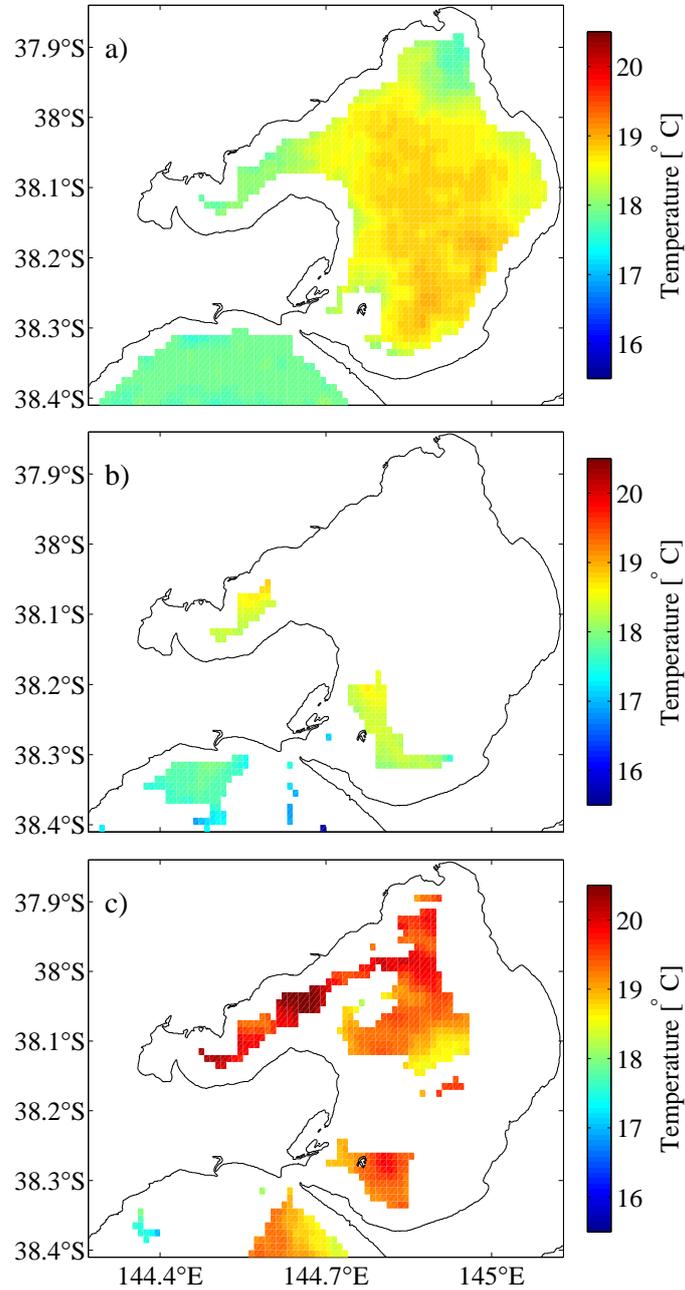


Figure 6.1: Examples of satellite SST observations: panel a) taken on 24 March 2003 01:27 is accepted, while panel b) observed on 30 March 2003 02:00 is rejected for too few pixels and panel c) observed on 18 March 2003 00:58 is rejected because the image is too fragmented.

data is used to remove images where there are too few pixels or where the image is fragmented or patchy. Also, if none of the pixels collocates with an in situ observation site then the image is rejected as well; at least one in situ observation is needed for the conditional merging and to assess observation error. Figure 6.1 presents some examples of images that are accepted and rejected. The final set of 21 observations used for assimilation over the 36 day period are presented in appendix E. This small number is on account of the high frequency of cloud cover; 21 NOAA–16 images are used for the assimilation experiments, whereas 72 overpasses were made over the simulation period. This is consistent with the high levels of cloud cover experienced in the PPB area (see section 3.3.6).

Throughout the chapter the satellite observations are assumed independent, that is

$$\mathbf{R} = \sigma^2 \mathbf{I}, \quad (6.1)$$

where σ is the standard deviation of observation error, which is taken as 0.5°C , based on quoted AVHRR accuracy. The basis for the independent observations assumption has been discussed in chapter 3.

6.2.2 Atmospheric Conditions

A brief indication of the weather conditions over the assimilation period is given by considering wind speed, cloud cover, and air temperature. Wind and cloud cover influence the skin-bulk relationship, while air temperature indicates the general state of the weather.

The wind speed over the assimilation period is plotted in figure 6.2 and indicates that wind speeds vary between 4 and 10 ms^{-1} , with a calmer period between 26–32 March. The amount of cloud cover during the period is given in figure 6.2. The data indicates that while there are periods of low cloud cover, high levels of cloud cover predominate. A cloud cover level of 7 oktas occurred 35% of the time. As a result fewer images were available for assimilation than would be present during more sunny conditions. The effect of clouds to reduce the number of observations is illustrated in figure 6.2, where the adopted observations are indicated (panel b) clustered in the periods where the cloud cover is low. The air temperature data are used as a secondary validation of the atmospheric conditions, observed temperature

and climatology. The air temperature at Point Wilson shows a general cooling of temperatures over the assimilation period in figure 6.2. However, during 24 to 30 March there appears to be an extended period of warm weather (panel c). There are also relatively lower wind speeds (panel a) during this period.

6.2.3 Temperature Climatology

The effectiveness of the assimilation performance has, so far, been evaluated by comparing the improvement of the assimilation forecast in predicting the truth with that of a control simulation, a model simulation starting at the same initial conditions but without any assimilation. Unfortunately, this procedure is not impartial. If the initial conditions are severely degraded from the truth, improvement by assimilation is easy. The assimilation performance is improved merely by selecting initial conditions far from the truth, in which case, the comparison proves little other than that the assimilation scheme is working. A more expedient test is to compare the assimilation forecast against a climatological forecast.

A climatological forecast is the statistical average conditions that would be expected at a given time of the year: the conditions expected based purely on historical observations, and ignoring any recent information. A climatological forecast of the daily temperature in PPB is calculated from the long term record of temperature at West Channel Pile. The West Channel Pile record is sorted by Julian day, and the climatological temperature is calculated as the average of all the temperature observations for a particular day, irrespective of the time of day or year. The climatology for the assimilation period is displayed in figure 6.3. It indicates the gradual cooling of the temperature from 19.3°C to 16.5°C , as late summer turns into autumn.

Each of the assimilation simulations is assessed by comparing the forecast temperature at the 6 in situ sites with the actual recorded observations. The performance of the assimilation is judged by comparing the RMSE against the RMSE for the site calculated using the climatology temperature. The assimilation forecast is only of use if it improves on the climatology temperature, otherwise the assimilation effort is of little benefit.

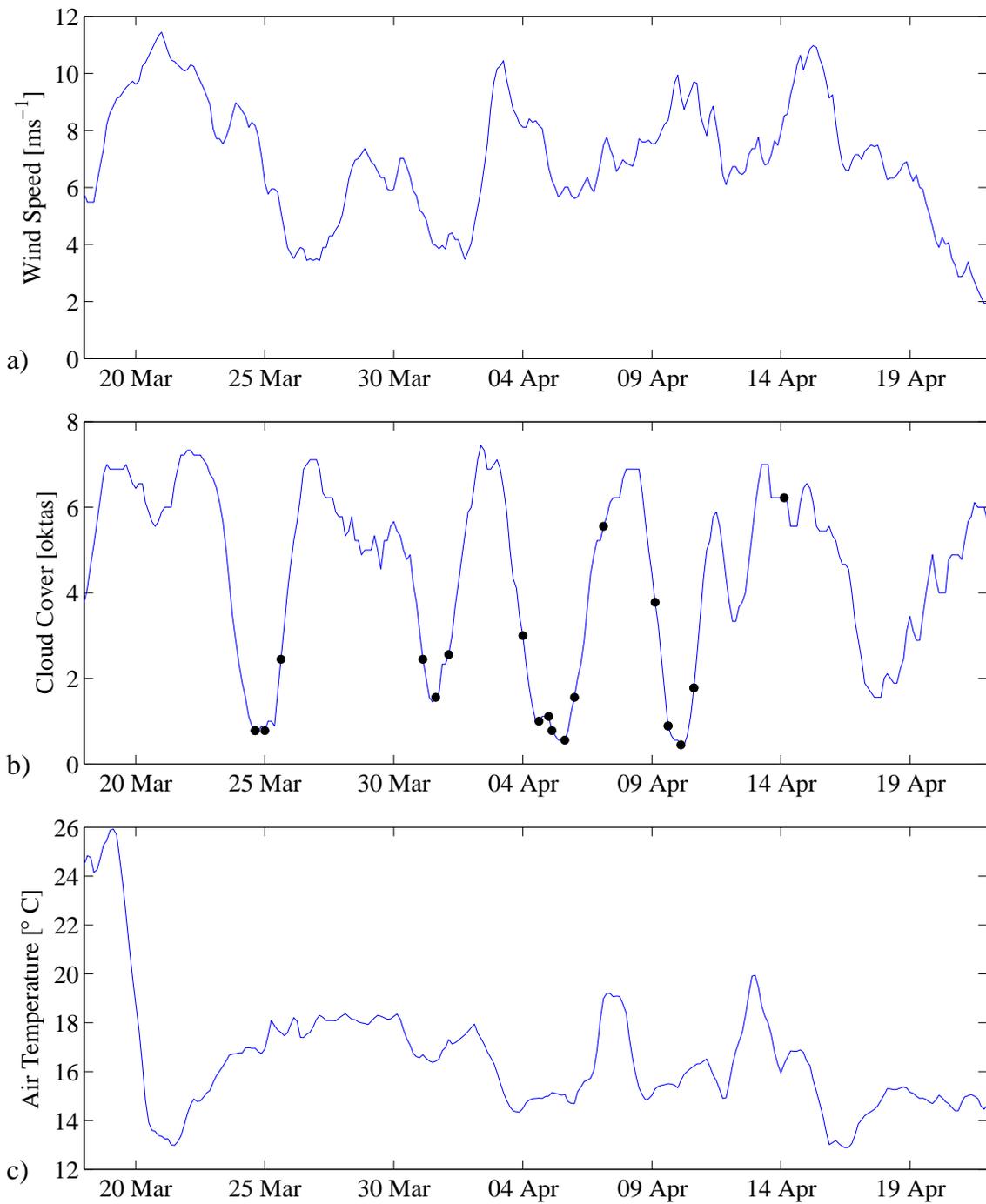


Figure 6.2: Observed a) wind speed (South Channel Island station), b) cloud cover (Melbourne station), and c) air temperature (Point Wilson station) over the assimilation period. The figures displayed were filtered with an 24 hour moving average to reduce the noise. The black dots indicate the times when adopted SST observations were recorded.

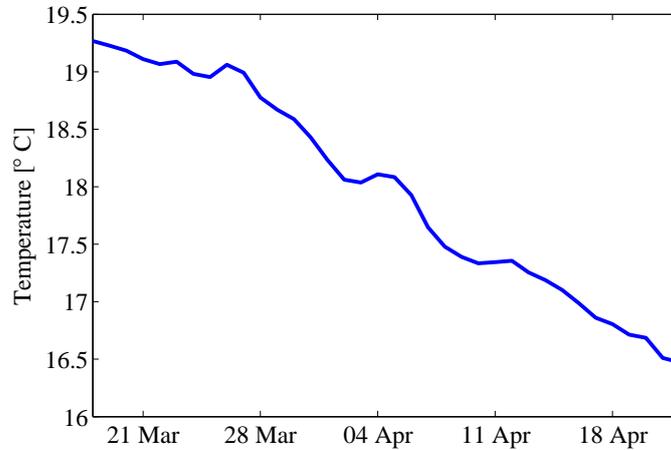


Figure 6.3: Climatological temperature in PPB based on the observation record at West Channel Pile from 1990 to 2001.

6.3 Unadjusted SST Assimilation

The first set of experiments assimilated the SST observations directly into the model (without any explicit correction of the skin-bulk difference) and is termed unadjusted SST. The accuracy of the forecast was measured against the observed in situ observations. The summary of this assimilation simulation is given in table 6.1. The assimilation gave clear improvement over the control simulation, and for three of the six monitoring stations the assimilation predicted more accurately than climatology. An average RMSE for the 6 sites indicated both the climatology and unadjusted SST assimilation give similar performance with a value of 0.63°C . But if the Hobsons Bay

Table 6.1: RMSE calculated for ensemble forecast temperature using unadjusted SST, control and climatology forecast temperatures relative to in situ observations. Boldface values indicate lowest RMSE at the respective site.

Monitoring Site	SST [°C]	Control [°C]	Climatology [°C]
Central Deep	0.38	0.96	0.66
Hobsons Bay Deep	1.14	2.10	0.85
Longreef Deep	0.70	1.70	0.58
Central Shallow	0.36	0.15	0.46
Hobsons Bay Shallow	0.58	1.28	0.72
Longreef Shallow	0.53	1.61	0.53

Deep record (which analysis of section 3.4.3 shows as being suspect) is ignored, the assimilated forecast shows improvement with an average RMSE of 0.51°C compared with 0.59°C for the Climatology. The control simulation did quite poorly with an average prediction error of 1.2°C as it fails to predict the cooling after 30 March. It only predicted the Central Shallow record accurately by luck, because this record was so short: only 6 days long. At other locations the control simulation greatly overestimated the temperature.

The improvement assimilation brings to a forecast is more clearly displayed in figure 6.4, with each panel representing a recorded in situ time series. The climatology prediction had mixed results, predicting the latter half of the simulation more accurately than the earlier half. The Deep records were not so well predicted, but Hobsons Bay and Longreef were more accurately predicted than Central. The rising temperature found in the in situ observations of figures 6.4d, e, and f around 26–32 March is due to an unseasonal warming period evidenced in the atmospheric data by warmer air temperatures and lower wind speeds (figure 6.2). Information on such weather events is not available in the climatology and so it failed to predict this event. This short-term warming event was not observed in the Central Deep in situ record, figure 6.4a, because the greater thermal mass provides inertia against short term fluctuations in temperature, and the Central Shallow record was too short to distinguish this event.

With the exception of the Hobsons Bay Deep result (and the early parts of Central Deep and Hobsons Bay Deep due to initialisation error), the control simulation predicted the temperature at all locations reasonably well over the first third of the assimilation period. Thereafter, the water cools rapidly, while the control continued to predict a warmer temperature and the forecast diverges. This is particularly apparent at Longreef and Hobsons Bay.

The assimilation of SST initially tracked the control — and the in situ observations — but as the temperature cools, the assimilation of satellite observations corrected the trajectory and insured that the model assimilation simulation continued to track the in situ record. However, the assimilation did not function perfectly and bias in the observations (and model) degraded the assimilated forecast. An example of observation error is the two vertical lines around 31 March, indicated on figures 6.4d, e, and f by a black circle. These lines are due to the assimilation

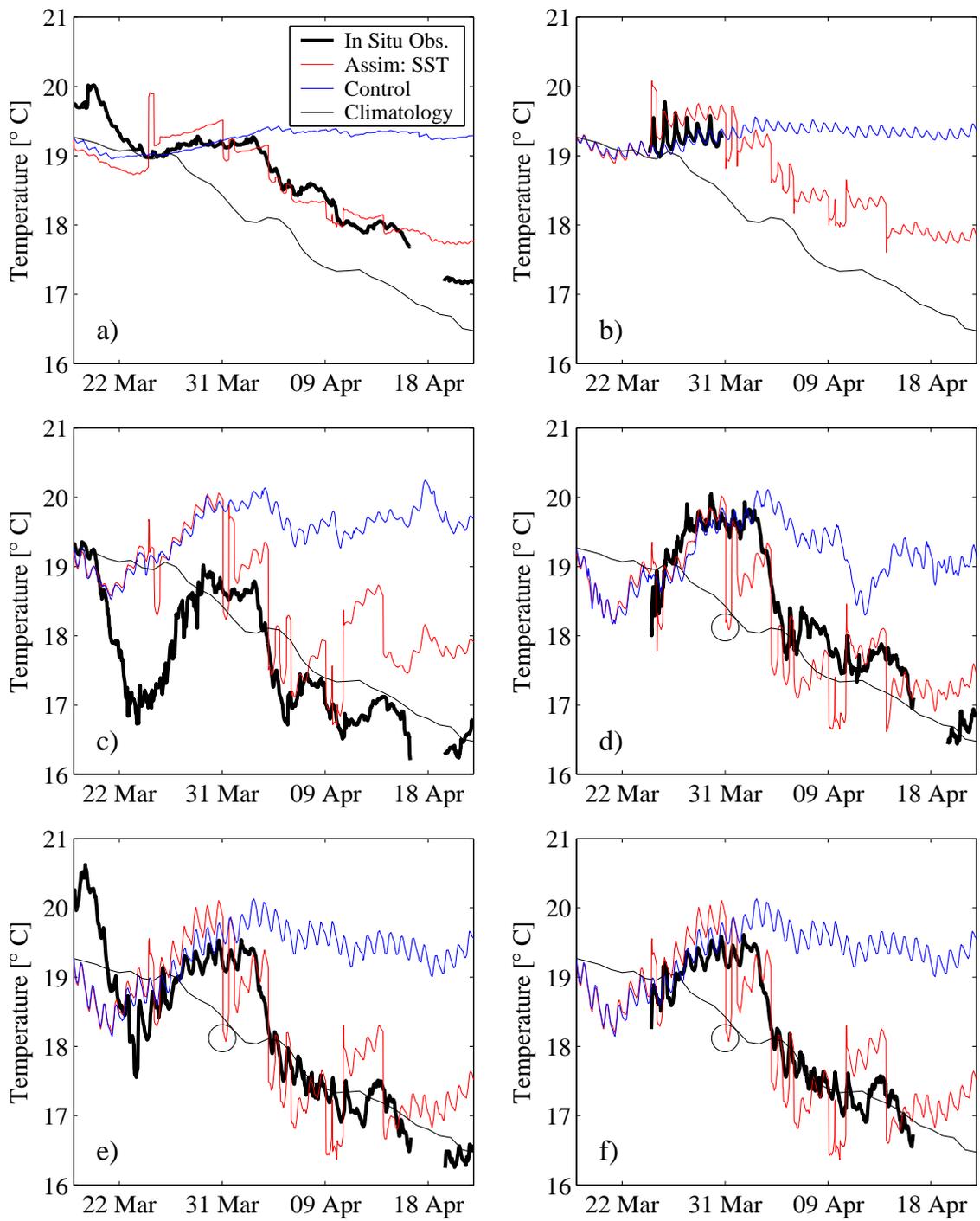


Figure 6.4: Plots of temperature forecasts with the assimilation of unadjusted SST data at a) Central Deep and b) Shallow, c) Hobsons Bay Deep and d) Shallow, and e) Longreef Deep and f) Shallow. Circles indicate presence of observation bias

analysing two images — images 5 and 6 in appendix E — 12 hours apart. Image 5, taken during the night (1:50AM), was cold by about 1.5°C , and reduced the temperature across the model accordingly, while image 6, taken during the day (2:11PM), was without error and returned the model to the in situ temperature.

As well as improving the timeseries of forecasts, the assimilation of SST observations improved the spatial prediction. Figure 6.5 compares the spatial predictions of the assimilation scheme with independent spatial observations from the NOAA-12 satellite. The comparisons are made at three times where the NOAA-12 images have a good spatial coverage of PPB. These images are evenly spread through the simulation period: 24 March, 4 April and 10 April.

For a spatial comparison, the attribute that is of importance is that the relative values of the images are correct, rather than the absolute values being correct (Vos and ten Brummelhuis, 1997). This means that the areas of warmer water in the images correspond to areas of warmer water in the forecast, and conversely with areas of lower temperatures.

The assimilated forecasts were more spatially varied than the control simulation and, from a visual inspection, match the NOAA-12 image more closely. In the NOAA-12 image, the warmest part of PPB was generally in the centre, with cooler parts in Bass Strait and the Geelong Arm. In the forecast (excluding the extreme edges) this pattern is replicated, but this is not found in the control which has a warmer temperature in the Geelong Arm on 4 April and 10 April.

An indication of the prediction skill of the forecast is given by comparing the absolute error of the temperature forecast relative to the in situ observations against the spread of the ensemble members. This comparison is done in figure 6.6 for each of the in situ sites. The results were unsatisfactory and indicate that the RMS spread is underestimating the absolute error of the forecast.

The large variation between the forecast and the in situ observation is likely due to the warm bias in the heatflux scheme, as the variation in error tends to follow a diurnal pattern similar to figure 5.5b. If the model does not predict the extent of the diurnal cycle accurately a diurnal variation in absolute error will result. However, it is promising that in the period after the first analysis — 24–30 March — in panels b, e, and f of figure 6.6 both the absolute error and the RMS spread increase with a similar gradient. This indicates that while the ensemble filter had difficulty

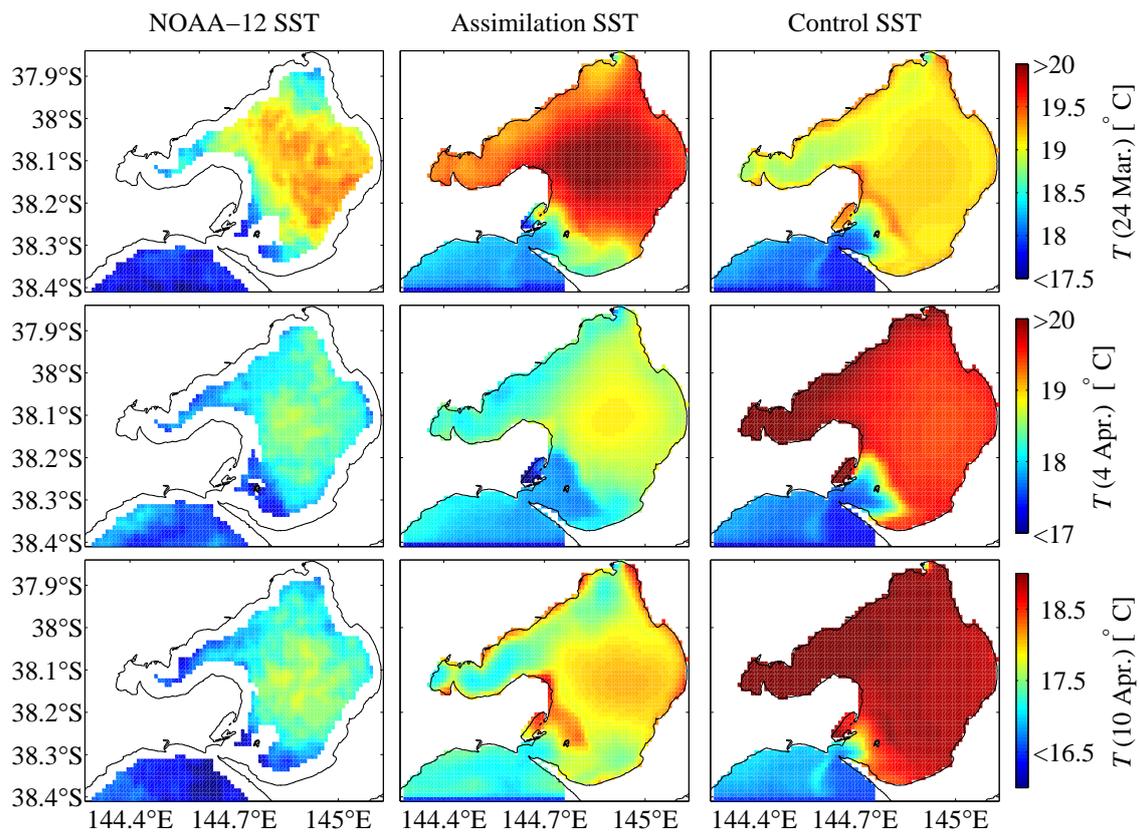


Figure 6.5: Spatial comparison of temperature forecasts of the SST assimilation after analysis and the control run against independent NOAA-12 images. The top row corresponds to the 24 March, the middle row to the 4 April and the lower row to the 10 April.

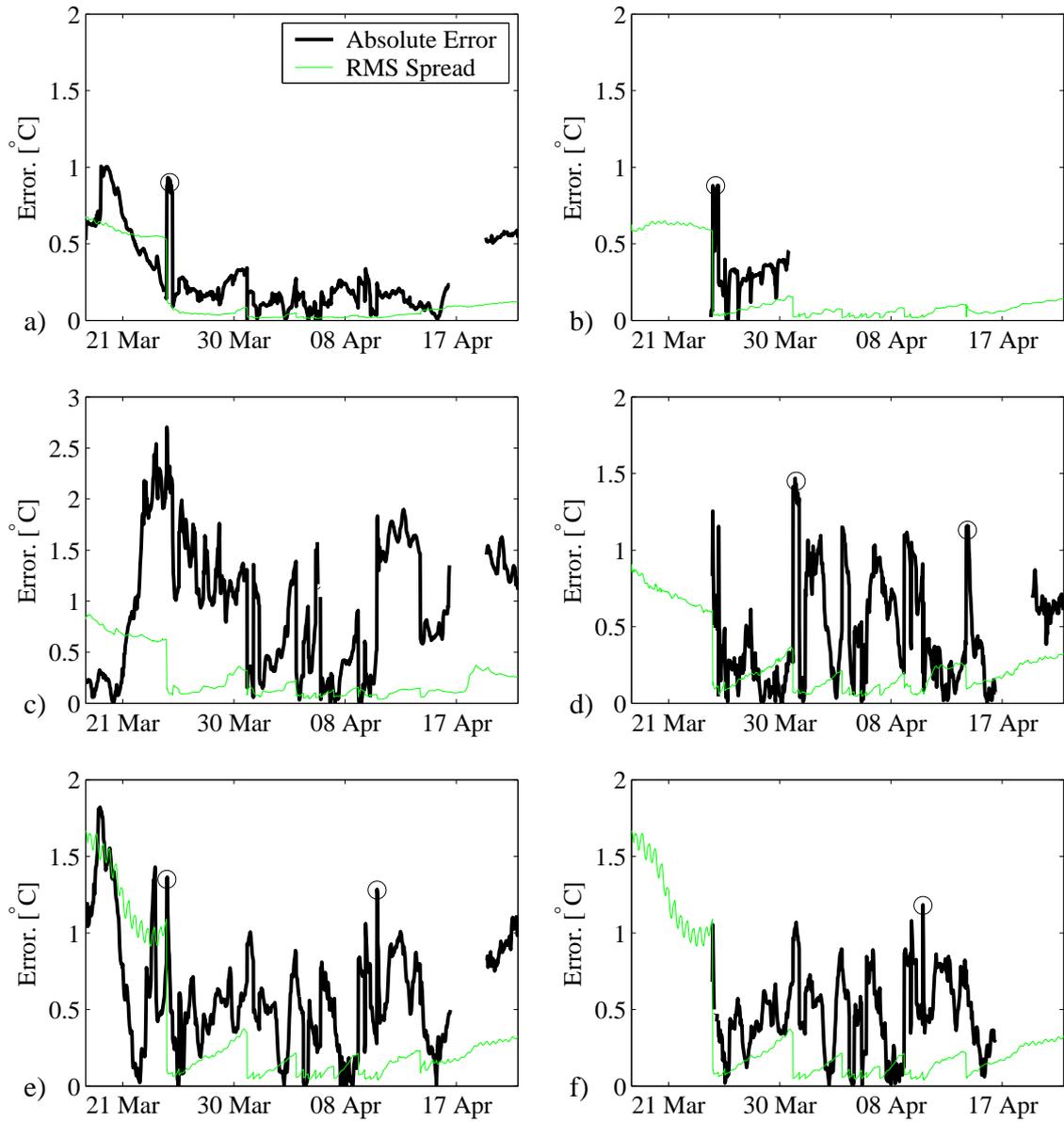


Figure 6.6: Plots contrasting absolute error between the unadjusted SST forecast and the in situ observation with the RMS spread at a) Central Deep and b) Shallow, c) Hobsons Bay Deep and d) Shallow, and e) Longreef Deep and f) Shallow. Circles indicate error attributable to observation error.

predicting its actual value, it appeared to predict the rate of absolute error growth well.

While the model bias produced a growth of absolute error, some of the larger spikes indicated by circles in figure 6.6 were due to observation error. The removal of observation error would lessen the impact of model bias, reducing the RMSE error. This will be pursued in the following sections.

In summary, the assimilation of the unadjusted SST observations significantly improved the prediction skill of the model, even without any observation error correction. Improvement is gained over a control simulation, and the forecast was better than climatology for three out of six in situ sites. However, during some analyses, large movements away from the truth were observed, indicating error in the SST observations. This was compounded by the underestimation of the absolute error by the RMS spread: due to both observation errors and model bias.

Assimilation of SST observations also improved the spatial pattern predicted by the model when contrasted against a control simulation. The comparison was performed visually. A more objective spatial comparison could be performed by applying a specific spatial comparison technique such as those discussed by Wealands *et al.* (2005), but this is not undertaken here.

6.4 SST Error Correction

The SST analysis of section 3.3.12 and assimilation in the previous section have shown that the SST observations contain an error relative to the bulk temperature. A likely cause of this error is diurnal warming which induces a temperature difference between the surface skin of the sea, observed by the satellite, and the bulk surface temperature, which is monitored in situ and modelled. This was explained in section 3.4.1.

The diurnal warming temperature difference is predicted by the formulae of Gentemann *et al.* (2003) (Gentemann) and Kawai and Kawamura (2002) (KK). By removing the temperature difference predicted by these formulae from the satellite observations, the diurnal error should be removed and the corresponding assimilation forecast improved. As spatially uniform forcing data are used, the error corrections are also spatially uniform.

The summary results of the assimilation using the Gentemann and KK corrections are present in table 6.2. These show that applying the diurnal warming corrections to the SST images improved the forecasting skill of the assimilation relative to the climatology at a few sites. However when contrasted against the unadjusted SST results of table 6.1, the only RMSE improvement was for KK at the Hobsons Bay Deep site, which (as noted previously) appears to have suspect data. In all other locations the RMSE increased relative to the unadjusted SST assimilation results or in the case of Gentemann remained unchanged.

The reason for the poor results can be seen in figure 6.7, which plots the actual observation error recorded at Longreef, Hobsons Bay and Central against that predicted by the diurnal warming algorithms. The Gentemann algorithm gave similar forecasts to the unadjusted SST observations, because its estimates of observation error were close to zero. The KK algorithm overestimated the observation error, often predicting a positive error when the actual error was negative. This is because the KK algorithm predicts the maximum diurnal warming temperature difference, and does not account for the actual time of day when the difference may be less.

For about half of the observations, the error at all the in situ sites is negative: the SST observations are colder than the in situ temperature. The theory of diurnal warming predicts that the skin temperature can be much higher than the bulk temperature but that it does not allow for the opposite. The cool skin effect theory (section 3.4.1) accounts for skin temperature being cooler than the bulk, but it

Table 6.2: RMSE calculated for ensemble forecast temperature using algorithm-corrected SST data relative to in situ observations. Boldface values indicate lowest RMSE at the respective site. For sites without a boldface value lowest RMSE is found in table 6.1.

Site Location	Gentemann algorithm [°C]	KK algorithm [°C]	Climatology [°C]
Central Deep	0.38	0.52	0.66
Hobsons Deep	1.14	0.91	0.85
Longreef Deep	0.73	0.74	0.58
Central Shallow	0.37	0.45	0.46
Hobsons Shallow	0.65	0.88	0.72
Longreef Shallow	0.58	0.65	0.53

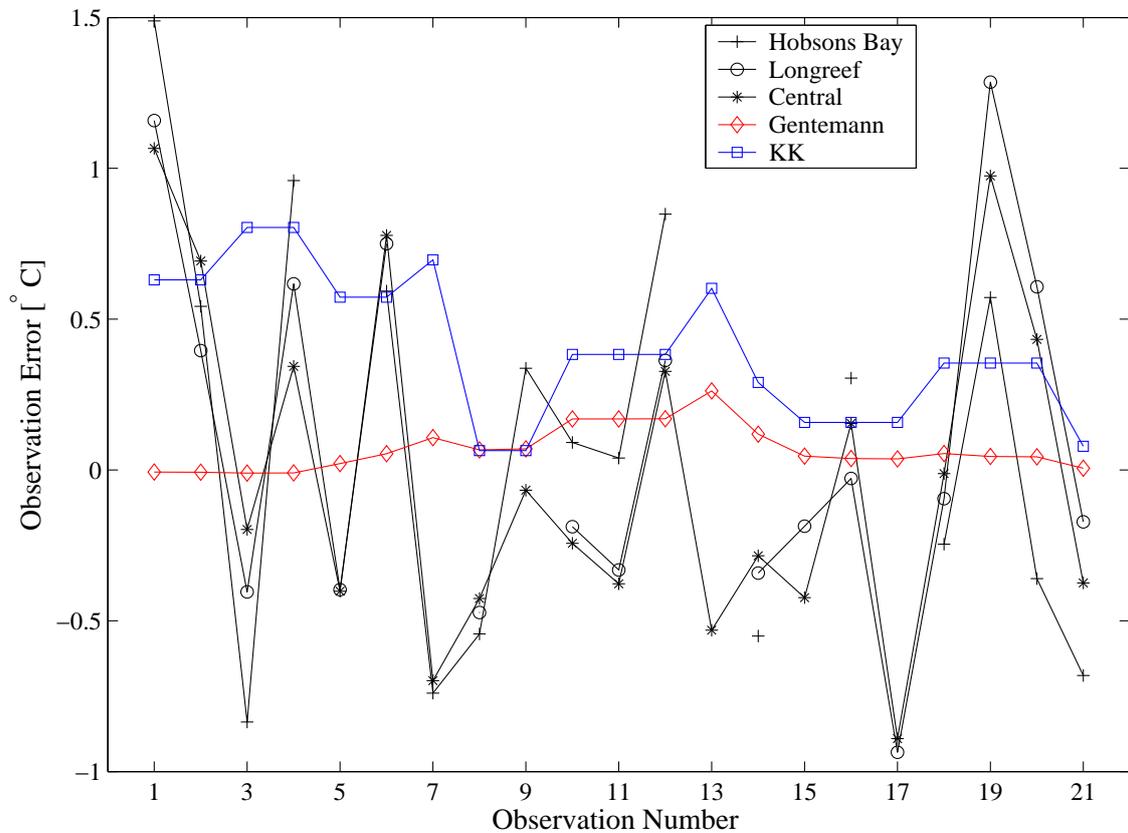


Figure 6.7: Comparison of actual observation error at the 3 monitoring sites and predicted observation error given by the Gentemann and KK algorithms. As the extent of the satellite image varies with cloud cover, the actual observation error is not always present for all sites.

predicts a magnitude of the order of a few tenths of a degree, rather than the -0.5 and -0.8°C range observed. This means that some other effect is inducing this *enhanced cool skin*.

The failure of the algorithms to predict the skin-bulk temperature difference is probably because both algorithms are derived for global applications to open ocean conditions. While the formulae may be globally applicable, they may not necessarily be accurate at a given location and may need to be tuned to local conditions. Furthermore, in PPB, which is a coastal environment, local effects such as bathymetry and riverine inputs may all effect the skin-bulk relationship in ways not intended by the diurnal warming formulae, which only account for wind, incoming solar radia-

tion, and for the Gentemann algorithm time-of-day. It may be possible to improve the forecast by developing a new diurnal model specific to PPB conditions, however this will not be pursued. Instead, improvement will be sought by merging the SST observation with additional observations of temperature.

6.5 Conditional Merging Assimilation

The assimilation of SST has improved the forecasting of temperature within PPB, but it has been hindered by diurnal observation error, which the diurnal warming algorithms were unable to remove. If the spatial pattern in the satellite observations containing diurnal error were combined with accurate, but spatially limited in situ observations, an improved observation set would result. Assimilating such a data set should correspondingly improve the forecast.

This concept is realised through conditional merging (Sinclair and Pegram, 2004), a technique originally developed to improve rainfall radar data. The underlying assumption of conditional merging is that spatial data give a good indication of the relative pattern but are not so accurate, while in situ data are accurate but contain little or no spatial information. By merging spatial and in situ data the resultant data set retains the spatial pattern, but is more accurate especially in the vicinity of the in situ sites.

A conditionally merged surface, A , is given by

$$A = (M - P) + L, \tag{6.2}$$

where L is a surface given by the spatial data, M is a surface generated by spatially interpolating in situ observations, and P is a surface generated by spatially interpolating values of L taken from the in situ observation locations. The spatial interpolation is performed by ordinary kriging, which gives the optimal spatial interpolation based on an assumption that the underlying process is stationary (Journel and Huijbregts, 1978, page 304). The combined $M - P$ field can be seen as a spatial interpolation of the observation error at the in situ locations that is added as a correction to the SST observations, L . Applying this technique assumes that the observation error is spatially correlated, which the discussion of section 3.4.3 and

especially figure 3.27 supports.

The basis to kriging is that the value of a process at a particular location can be optimally estimated as the weighted average of the known value of the process at other nearby locations. The weightings are based upon the spatial covariances, which are presented as a semivariogram, indicating the semivariance between two locations as a function of their separating distance. For more details refer to Journel and Huijbregts (1978, pages 303–313).

The semivariogram is generated based on observations. However, when this was attempted for the satellite data, a considerable amount of noise precluded the determination of the semivariogram and so an assumed semivariogram based on a spherical form was adopted and used for the kriging. The adopted semivariogram was taken as

$$\gamma(h) = a \left(\frac{3h}{2 \times d} - \frac{1}{2} \left(\frac{h}{d} \right)^3 \right), \quad (6.3)$$

where h is the distance between the two locations, a is variance at zero distance with a value of 0.01°C^2 , and d is the correlation length-scale taken as 56 km, being a best guess estimate based on high correlation of observation error shown in figure 3.27. Using this form of equation the covariance decays from 0.01°C^2 at a distance of zero, to zero at a distance of 56 km. Further information on this form of covariance and kriging is available from Pebesma (2001). As an assumed semivariogram is used, the kriged surface cannot be considered optimal.

An example of a kriged temperature surface is given in figure 6.8. The figure demonstrates how kriging is used to spread the observation error known at a few in situ locations to the rest of the satellite image. The accuracy of the resulting conditionally merged image is limited by the number of in situ observations and decorrelation scales. Near the in situ observations the correction is more accurate, while further away it becomes less accurate.

The conditional merging requires in situ observations, however if the observation image does not have a pixel that corresponds to the in situ observation (due to cloud contamination), then that in situ observation cannot be used. Therefore, depending on the image (and the amount of cloud), between one and three points were used to generate the kriged field. When one point is used the kriged surface becomes a uniform offset akin to the diurnal warming corrections of section 6.4. While the

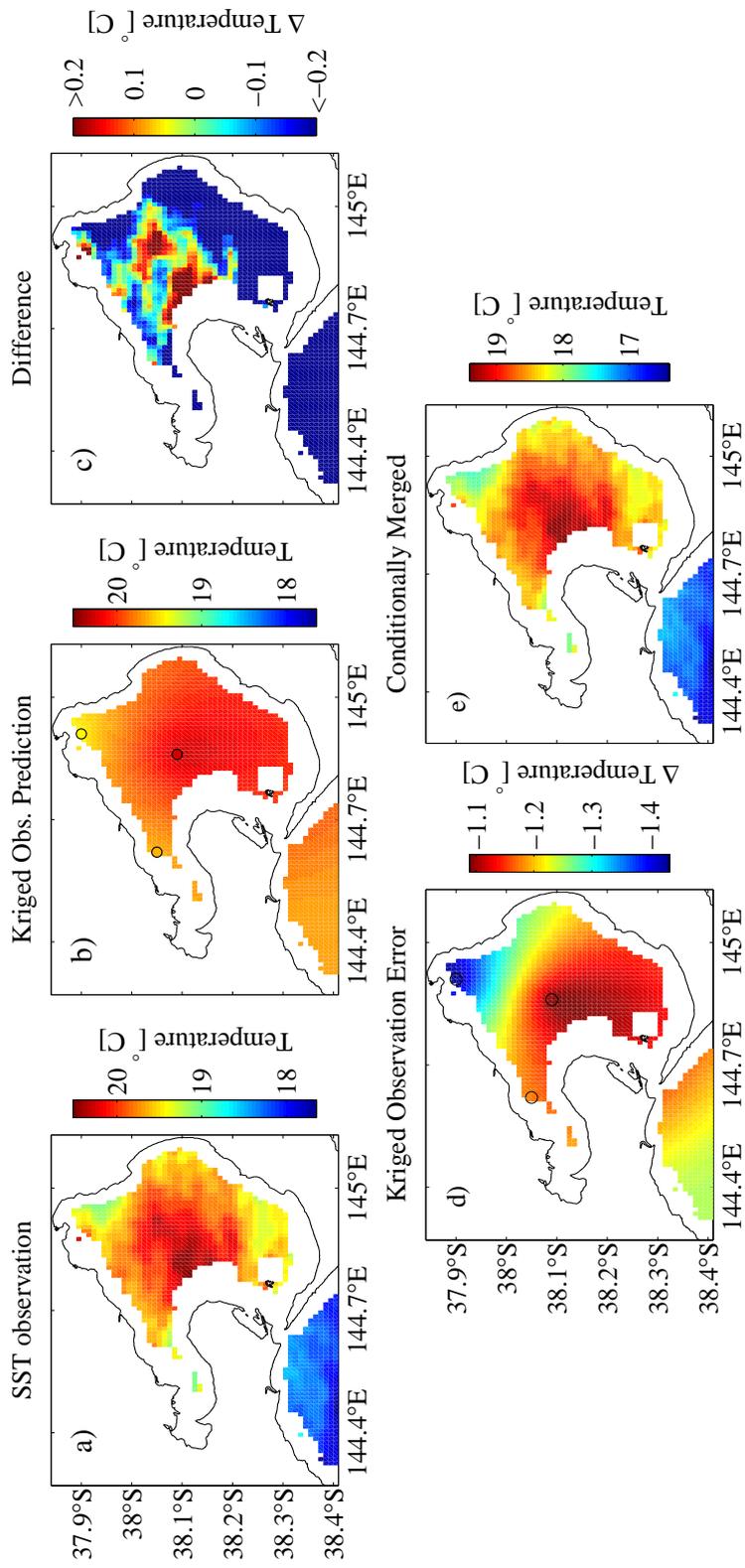


Figure 6.8: An example of the generation of a conditionally merged field. Panel a) displays the satellite observations, panel b) displays the kriged field based on the satellite observations at three in situ sites, indicated by the black circles, panel c) displays the SST prediction error (the difference between panels a and b), panel d) displays a kriged observation error based on observation error at the three in situ sites indicated, and panel e) displays the resulting Conditionally Merged field when panel d) is added to panel a).

standard deviation of the in situ observations is less than 0.01°C and at a distance of 500m is usually within $\pm 0.02^{\circ}\text{C}$, the observation error standard deviation of the corrected observations is left as 0.5°C , refer equation 6.1. The conditional merging reduces the variance in the image. However, as the semivariogram is assumed, a reduction of the variance has not been applied to the observation error covariance matrix.

The results from assimilating conditionally merged satellite observations into the PPB model are summarised in table 6.3. The conditionally merged assimilation forecast produced improvements with an average RMSE excluding Hobsons Bay Deep of 0.44°C , this contrasts well with 0.59°C for the climatology and 0.51°C for the unadjusted SST assimilation. The conditionally merged forecast is more accurate than climatology at four out of the six sites are illustrated in table 6.3.

The time series of forecasts at the in situ sites are displayed in figure 6.9. The improvement of the conditionally merged forecast over the unadjusted SST forecast is observed by contrasting the assimilation jumps on 30 March. In the unadjusted assimilation, large jumps were observed pushing the forecast away from the in situ observations (black circles in figure 6.4). The conditional merging has reduced this bias significantly and correspondingly was more accurate. If a greater number of shallow in situ sites were available the resulting forecasts should be even more accurate, but installing more in situ sites is costly and so a trade-off between cost and accuracy is made.

Table 6.3 indicates that the climatology forecast was more accurate than con-

Table 6.3: RMSE calculated for ensemble forecast temperature using conditionally merged (CM) SST observations and climatology relative to in situ observations. Boldface values indicate lowest RMSE at the respective site. For sites without a boldface value the lowest RMSE is found in tale 6.1

Site	CM [$^{\circ}\text{C}$]	Climatology [$^{\circ}\text{C}$]
Central Deep	0.38	0.66
Hobsons Deep	1.15	0.85
Longreef Deep	0.66	0.58
Central Shallow	0.18	0.46
Hobsons Shallow	0.58	0.72
Longreef Shallow	0.42	0.53

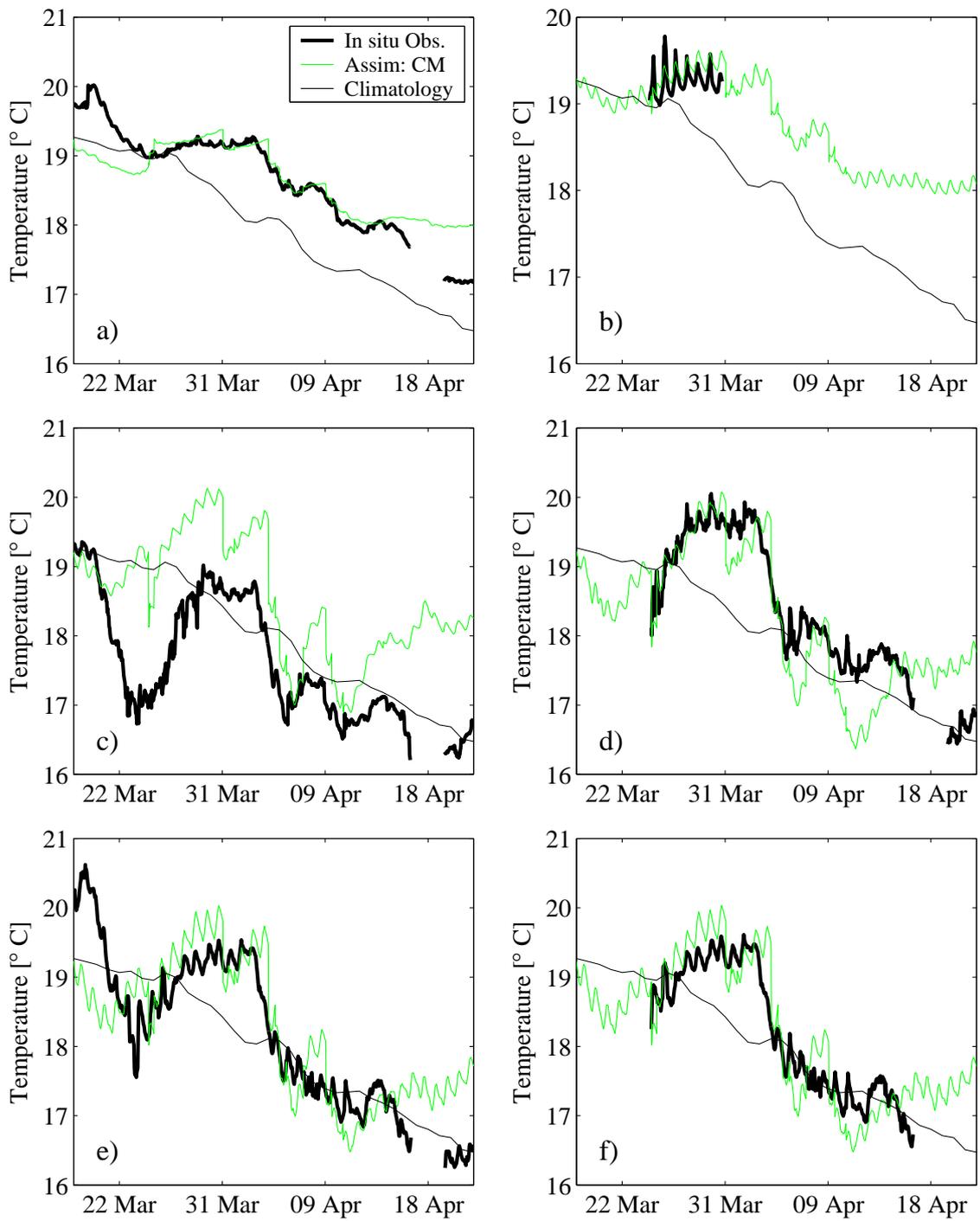


Figure 6.9: Plots of temperature forecasts with the assimilation of conditionally merged observations at a) Central Deep and b) Shallow, c) Hobsons Bay Deep and d) Shallow, and e) Longreef Deep and f) Shallow.

ditional merging at the Longreef Deep site, even though at the Longreef Shallow site conditional merging is more accurate than the climatology. The reason for this is that within the simulation period the SST observations were only available from 24 March to 14 April. Within this period the conditional merging forecast follows the in situ observations more accurately than the climatology (figure 6.9, panel e). However, before and after this period the climatology is more accurate as initialisation errors and model bias, respectively, reduce the accuracy of the conditionally merged assimilation forecast.

The effect of the conditional merging on the prediction of the absolute error (being the difference between the ensemble mean forecast and the in situ observations) is displayed in figure 6.10 and contrasts with unadjusted SST absolute error in figure 6.6. The removal of the diurnal observation error has correspondingly lessened the RMSE spikes identified in figure 6.6. Over the first third of the assimilation period the RMS spread is of the same order as the absolute error, indicating that the ensemble is characterising the absolute error well. The high absolute error at the end of the simulation occurs because no SST observations were available after 14 April and with no correction the model overestimated the temperature because of the warm bias.

The assimilation of conditionally merged SST observations gave an more accurate forecast than either the unadjusted SST assimilation forecast or the climatology forecast. This was due to the conditional merging removing the diurnal observation error through the incorporation of accurate in situ measurements.

6.6 In Situ Assimilation

To investigate an assumption that the assimilation of spatial data is more useful than the assimilation of point scale data, a subsequent run is made assimilating the in situ data observed at the Longreef surface site directly into the model. While the satellite must wait for an overpass, this is unnecessary for in situ data as it is collected continuously and so the data are assimilated every six hours. Assimilation at a higher frequency is possible, but the marginal benefit relative to the change in temperature over this period is negligible. The standard deviation of the uncertainty associated with the in situ data is taken as 0.01°C .

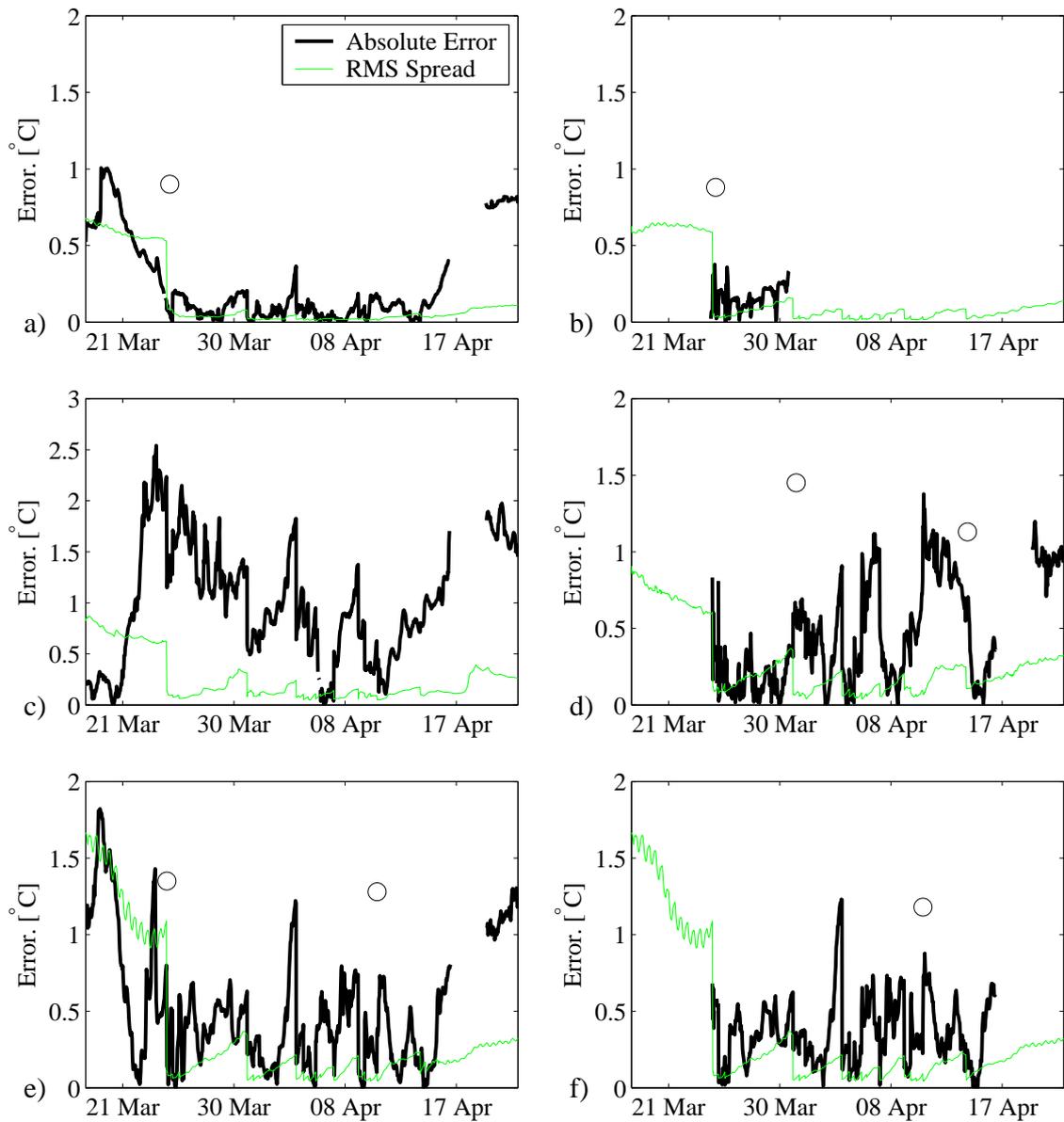


Figure 6.10: Plots contrasting conditionally merged absolute error between the forecast and the in situ observation with the RMS spread at a) Central Deep and b) Shallow, c) Hobsons Bay Deep and d) Shallow, and e) Longreef Deep and f) Shallow. Circles show where conditional merging has reduced the observation error (contrast with figure 6.6).

Table 6.4: RMSE calculated for ensemble forecast temperature using in situ observations at Longreef Shallow and climatology relative to in situ observations. Boldface values indicate the lowest RMSE at each respective site. For sites without boldface values the lowest RMSE is found in tables 6.1 or 6.3.

Site	In Situ [$^{\circ}\text{C}$]	Climatology [$^{\circ}\text{C}$]
Central Deep	0.92	0.66
Hobsons Deep	1.71	0.85
Longreef Deep	0.51	0.58
Central Shallow	1.16	0.46
Hobsons Shallow	0.64	0.72
Longreef Shallow	0.12	0.53

As expected, the assimilation of the Longreef in situ observations into the model significantly improved the accuracy of the forecast at Longreef (see table 6.4). However, this improvement was spatially confined and at the other sites the accuracy of the forecast deteriorated, with an average RMSE, excluding Hobsons Bay Deep, of 0.67°C . This is much higher than the average climatology value of 0.59°C and the conditionally merged value of 0.44°C .

While the three sites are a similar distance from each other, in situ forecasts at Hobsons Bay were more accurate than at Central, probably because Hobsons Bay and Longreef share similar characteristics: both are of similar depth (around 10 and 6 metres respectively) and have a similar proximity to the coast. Therefore they operate in a similar fashion, subject to similar errors and are expected to have a high error covariance. By contrast the Central site functions differently and so Central and Longreef are less correlated. (See figure D.1g of appendix D).

These findings are confirmed by the plots of figure 6.11 that contrast the spatial temperature distribution of the in situ forecast with the conditionally merged assimilation and the control run. The area of influence of the assimilated in situ data appears to extend 10-15km around Longreef. Within the area of influence the in situ forecast gives similar results as the conditionally merged assimilation. At more distant locations the temperature predicted by the in situ forecast approaches that of the control run, indicating that the assimilation is not improving the forecast at these locations. The improvement of the temperature prediction by the conditionally merged assimilation and in situ assimilation, within its area of influence,

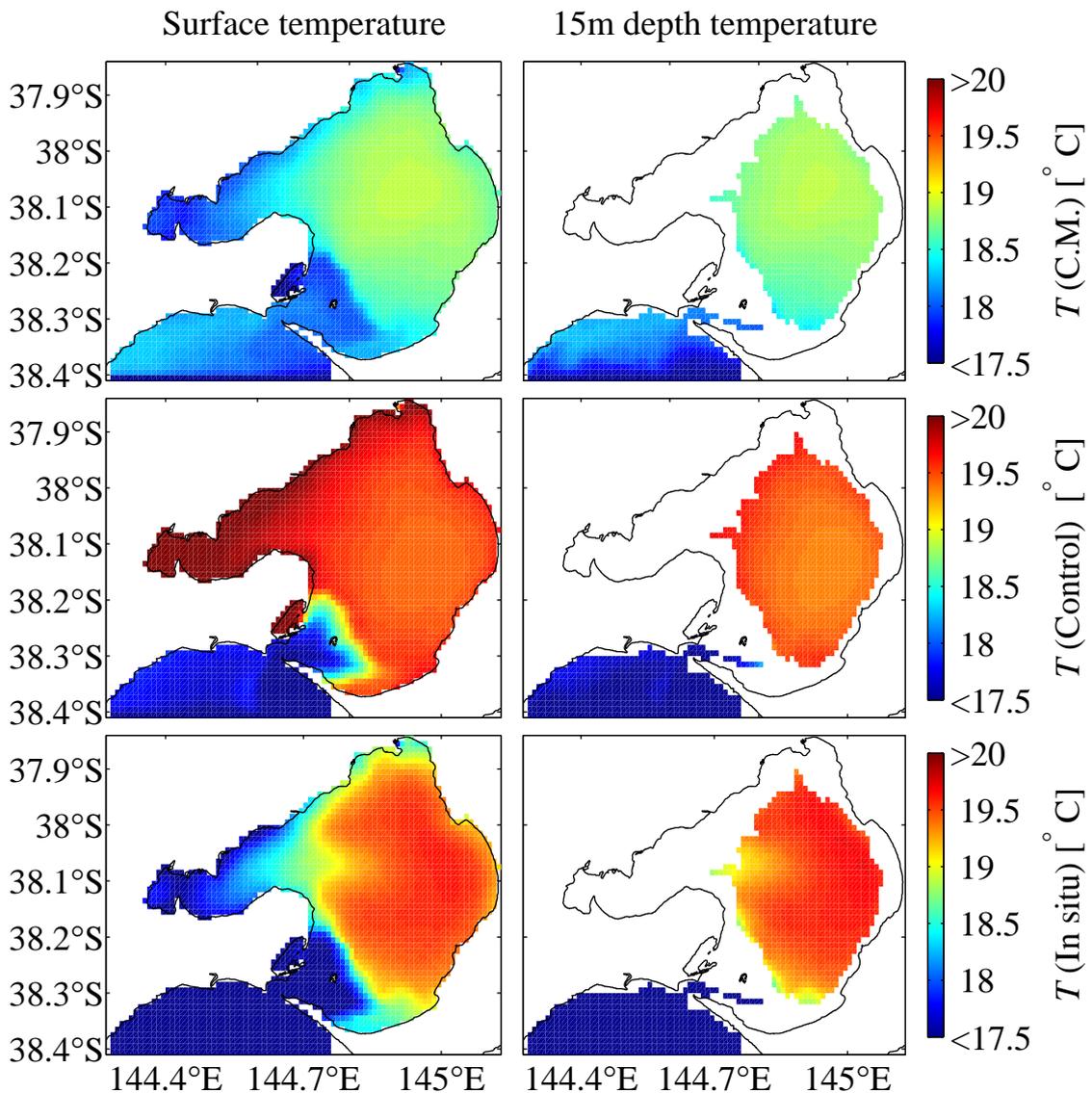


Figure 6.11: A comparison of the temperature states of PPB on 4 April predicted by the assimilation of conditionally merged observations (top row), the control run (middle row), and the assimilation of in situ observations at Longreef (bottom row). Predictions are shown for the water surface and at 15m. depth.

at surface and at depth illustrates the ability of data assimilation to update model states at both observed and unobserved locations.

The results of the assimilation of an in situ point illustrate the importance of assimilating spatial data. Even though the in situ observation is more accurate the

spatial variation in temperature is better represented by less accurate (but unbiased) spatial observations. The improvement of spatial observations over in situ observations will depend upon the spatial covariances of the variable being studied. Those which are greatly influenced by local conditions will react more strongly to the assimilation of spatial rather than point scale observations.

6.7 Innovations

An analysis of the innovations gives further insight into the performance of the assimilation scheme and suggests where further improvement is possible. Three aspects of the innovations are investigated: i) time series of the mean and standard deviation of the innovations, which is suggestive of bias and the observations uncertainty, ii) histograms of the innovations for unimodal distribution, and iii) spatial distribution of the innovations for any indication of spatial error.

The variation of the mean of the innovation over time is plotted in figure 6.12 for both the unadjusted SST and conditionally merged forecasts. The diurnal error in the unadjusted SST values is obvious as the day-time, between 10AM and

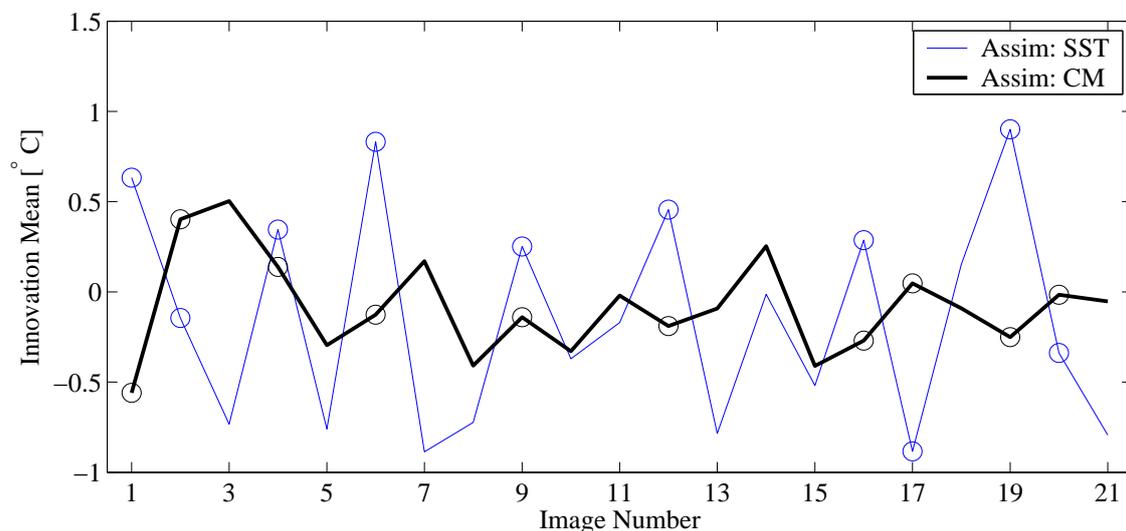


Figure 6.12: Plot of innovation mean of the unadjusted SST and conditionally merged observations against observation image number. Day-time images, between 10AM and 6PM, are circled.

Table 6.5: Temporal average of innovation spatial mean and standard deviation over the simulation period for the unadjusted SST assimilation (SST) and the conditional merging assimilation (CM).

	Average Innovation Mean [°C]	Average Innovation Standard Deviation [°C]
SST innovations	-0.16	0.26
CM innovations	-0.08	0.26

6 PM (circled), observations are predominately positive. The impact of the conditional merging is significant, reducing the temporal variation of the innovation mean. However, the presence of some jagged spikes indicates that a degree of diurnal error remains.

The average innovation statistics are presented in table 6.5. The average conditionally merged innovation mean is half that of the average uncorrelated SST innovation mean, but both are negative. As the diurnal observation error has mostly been removed, it can be concluded that this is predominately due to model bias: the observations are continuously cooling the model as the model is warm biased. Further improvement in forecast performance could therefore be gained by reducing the model bias, either by improving the thermodynamic module used in the assimilation, or by applying a bias aware assimilation system (e.g. Dee and Todling, 2000).

The temporal variation in innovation standard deviation is plotted in figure 6.13 for both the unadjusted SST and conditionally merged observations, with the time-averaged values given in table 6.5. Conditional merging has little effect on the standard deviation of the innovations. This is because the diurnal observation error is generally spatially uniform. As a result applying the conditional merging reduces the error but the spatial distribution remains the same. Moreover, figure 6.13 indicates that the standard deviations are distributed into two groups. In the earlier innovations the standard deviation averages around 0.35–0.4°C, while for the later innovations the variation has reduced to about 0.15–0.2°C. This will be explored further through the innovation histograms and spatial distribution.

A final point for the innovation standard deviations is that in all cases the standard deviation is less than the 0.5°C and in some cases quite significantly. As noted

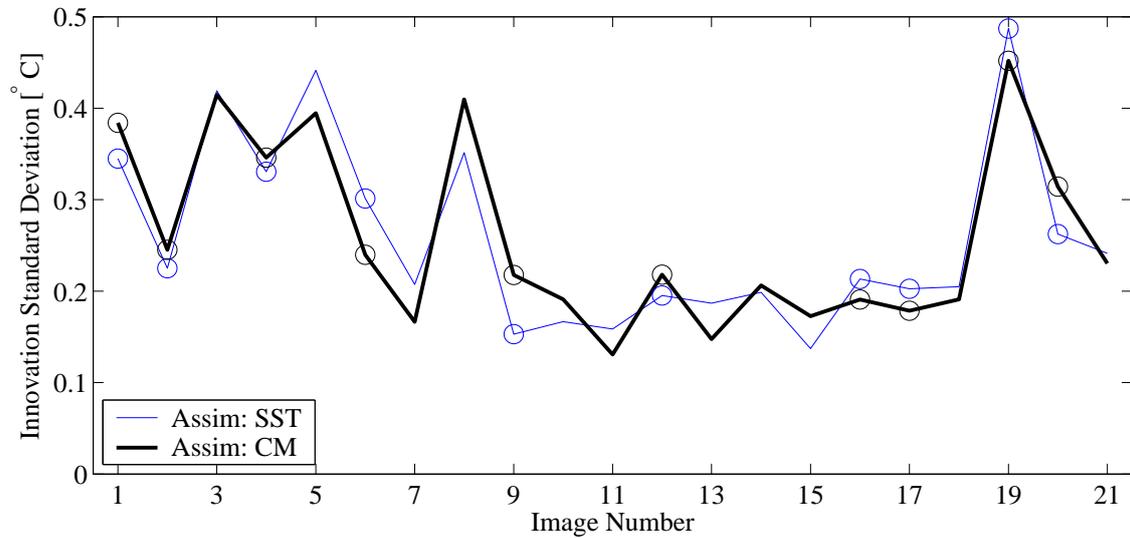


Figure 6.13: Plot of innovation standard deviation of the unadjusted SST and conditionally merged (CM) observations against observation image number. Day-time images, between 10AM and 6PM, are circled.

in section 5.4.2 the variation of the innovations is given by the summation of the model variation plus the observation error variation. Whatever the true level of model error, the standard deviation of the observation error is quoted as 0.5°C . As the innovation standard deviation is less than this value, there is evidence that the observation error contains spatial correlation and that the observations are not independent as assumed, refer to discussion in section 3.4.4. The absolute error of a temperature observation relative to the actual temperature is 0.5°C , but the variation of a temperature observation relative to its neighbour is much less. While this does not necessarily diminish the results, it suggests that further research into the observation error covariance structure is warranted.

A first step would be to assume an isotropic bias correlation length scale and use that to construct the observation error covariance. This approach is similar to one of the two approaches used to generate the forecast error covariance matrix in optimal interpolation (refer to section 2.3). The cross correlation of observation error values at the in situ stations (figure 3.27) together with the separation distance could give advice for an appropriate correlation length scale. The innovation statistics can be used to judge if the observation error covariances used in subsequent simulations are

more appropriate. Furthermore, the effect of conditional merging on the observations error covariance structure also needs to be considered as well. Incorporating further information into the observations reduces the observation error. This can be calculated statistically using kriging techniques (refer to Journel and Huijbregts, 1978; Pebesma, 2001).

The innovation distributions are plotted as histograms to check for unimodal behaviour in figure 6.14. In contrast to the synthetic experiments of chapter 5 (figure 5.6), the distribution of the innovations is far less regular. Images 3, 5 and 8 are definitely bimodal and image 19 appears so also. Interestingly, these images correspond to the innovations, with the largest standard deviations indicated in figure 6.13.

This bimodal nature is explored further by considering the spatial distribution of the innovations. This is displayed in figure 6.15. If the observations are independent, as assumed, then the distribution of the innovations should be independent also. A minor degree of spatial correlation will give structure to the image, but striking variations are not expected. In most of the images there is no distinct innovation pattern, but for images 3, 5, and 8, the innovations in Bass Strait are clearly larger than in PPB and it is this difference which is creating the bimodal innovation distribution. The likely reason is the use of a climatology forcing for the open boundary in Bass Strait, which was used as no temperature observations were available. The bimodal events occur on 25, 31 March and 4 April. This period coincides with the extended warming seen in figure 6.2c, which is not evident in the temperature climatology (figure 6.3). As a climatology is used to force the temperature in Bass Strait, the temperature over this period is underestimated, leading to the higher innovation values observed.

Another possible reason for the bimodal innovation structure is that a difference exists in the skin-bulk relationship between Bass Strait and PPB, which (because there are no in situ records in Bass Strait) the conditional merging does not take into account. As images 3, 5, and 8 are all night images, a skin-bulk issue may be the cause, but this is unlikely because no other records have such a clearly demarked Bass Strait–PPB innovation difference and it only occurs during the early part of the simulation period.

Image 19, recorded at 1:50PM, also has a bimodal distribution, which is definitely

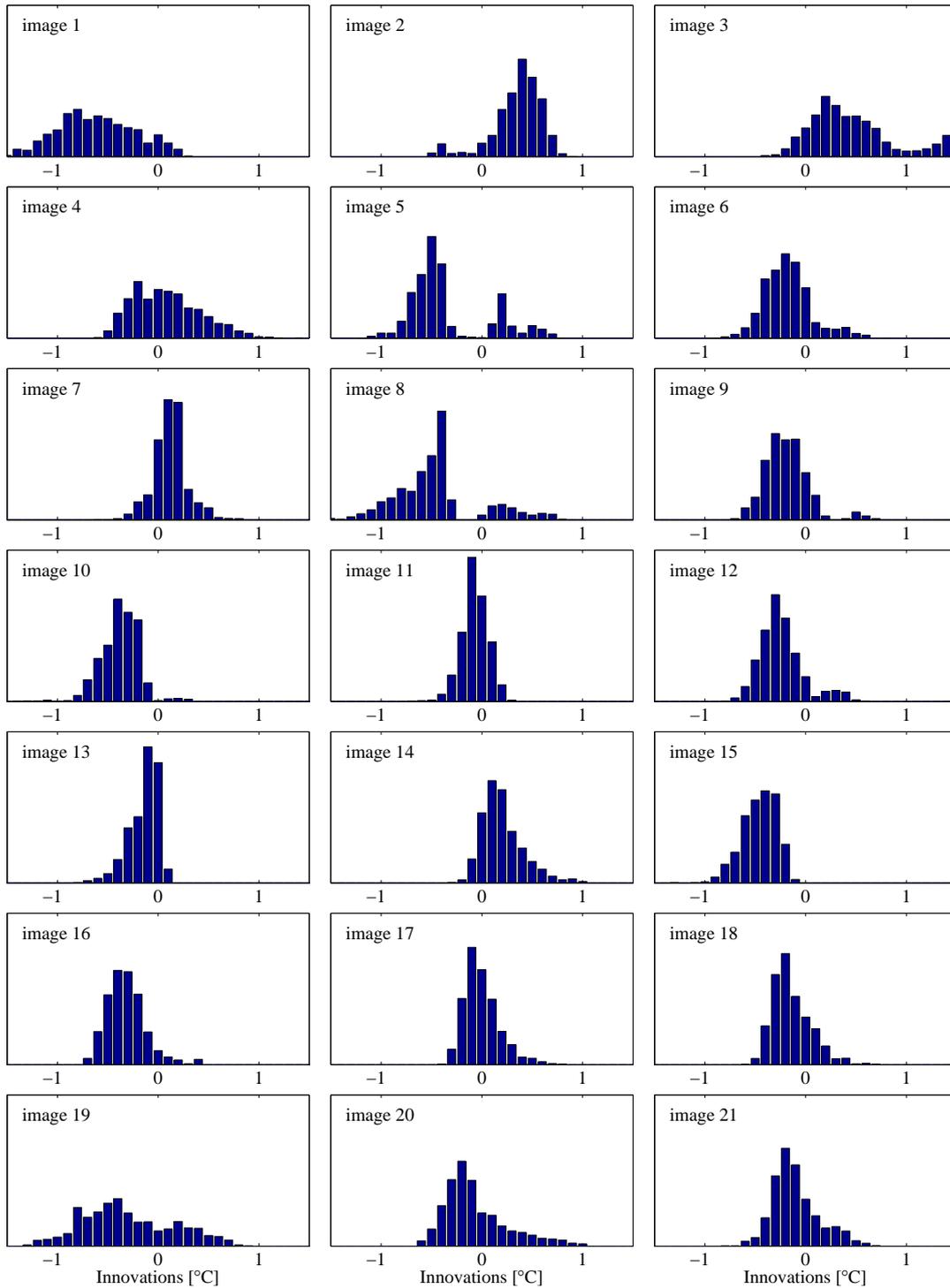


Figure 6.14: Innovation distributions for conditionally merged forecasts. Each histogram has been normalised by dividing by the total frequency.

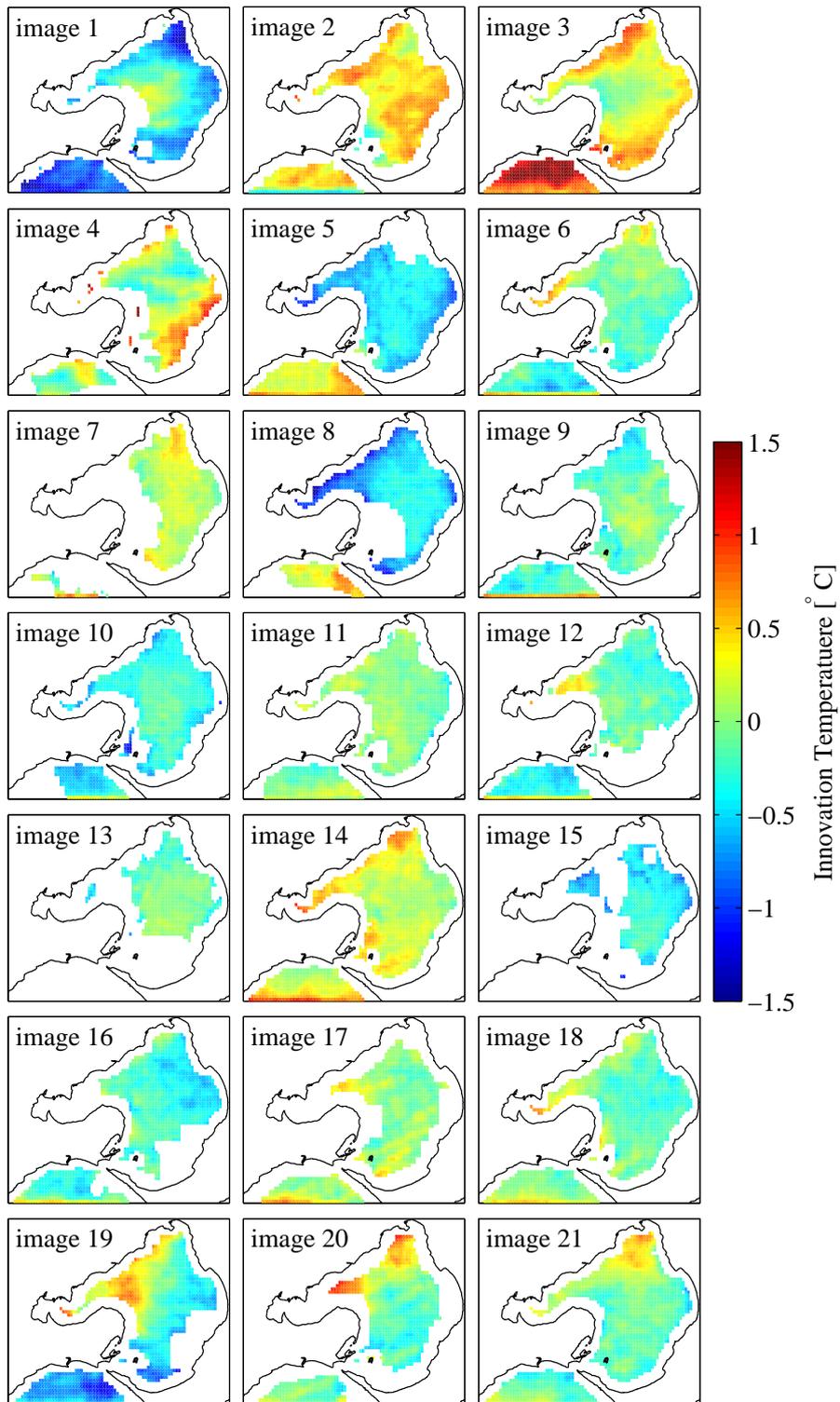


Figure 6.15: Spatial distribution of innovations for conditionally merged forecasts.

spatially distributed (figure 6.14). One mode is located in the western side of PPB and in the Geelong Arm, and the other mode comprises the eastern side of PPB and Bass Strait. Considering the image 18 (Appendix E) and the ensemble forecast, the bimodal innovation distribution suggests an observation error in image 19. The warmest part of the image is to the western edge of PPB, while in other images the warmest area is generally near the centre of PPB (see for instance images 18 and 20 of appendix E). Evidently some form of contamination, possibly due to fog, is distorting image 19. This prompts the need for more careful preprocessing

The spatial distribution and histogram of the innovations has enabled errors in both the model forecast and contamination in the observations to be recognised. This highlights the power of using innovations as a diagnostic tool. Based on the innovation analysis, the forecast can be improved by changing the open boundary conditions. Presently the open boundary uses a clamped forcing condition. However, if the assimilation brings improvements perhaps a combination boundary which is initially a zeroflux (or no gradient) boundary after an analysis, but which relaxes towards a clamped climatological boundary over time is more appropriate.

As the error in the observation image (19) was not detected either more detailed preprocessing is needed or an online check is required whereby the innovations are tested in the manner described in this section prior to assimilation. Those images that fail the innovation tests can be either manually processed further (for example, in the case of image 19 by removing the half of the image thought to be contaminated), or removed from the assimilation and the forecast continued until the next viable observation becomes available.

This section has found that

1. conditional merging reduces the observation bias, with the average innovation mean being halved,
2. the assumption of independent observation error is incorrect as the standard deviation of the innovations is continuously less than the assumed observation error, and
3. errors in the model structure and errors in the observations can be detected by exploring the innovation distribution.

The implication of these findings for data assimilation is that the information contained in the innovations can be used to validate the assimilation and direct measures to further improve forecasting accuracy.

6.8 Chapter Summary

This chapter applies ensemble data assimilation techniques to an actual event in PPB. It has been shown in chapter 5 that data assimilation improves the short term forecast in a synthetic setting; the purpose of this chapter is to discover how the assimilation works in a real environment and to illustrate improvement over a climatology forecast.

This chapter focused on observation error and its removal to improve the forecast. Observations containing diurnal errors resulted in adjustments to the forecasts away from the truth thus reducing forecast accuracy. Attempts to remove diurnal observation error were made using diurnal warming algorithms and conditional merging. The diurnal warming formulae failed to accurately predict the skin-bulk temperature difference; as a result these techniques gave poorer results than the assimilation of unadjusted SST. The use of conditional merging led to improved forecasting skill, halving the average innovation mean from -0.16°C to -0.08°C , relative to an unadjusted SST assimilation forecast. This improvement was on account of the removal of the diurnal observation error. The persistence of the negative average innovation mean is due to warm bias in the heatflux module of the model. Overall, the tests were successful with the average RMSE of the conditionally merged assimilation forecast (excluding the Hobsons Bay Deep site) of 0.44°C , a large improvement over the climatology average RMSE of 0.59°C .

An analysis of the innovations of the unbiased and conditionally merged assimilation runs gave further insight into the assimilation performance and prompted suggestions for improving the forecasting performance. Through the investigation of bimodal innovation distributions and their spatial distribution, improvements to the open boundary setting of the model are suggested and suspicious satellite observations detected.

Chapter 7

Conclusions and Future Directions

This thesis has investigated the improvement of forecasting temperature in a coastal embayment through the assimilation of SST observations. The research was prompted by the increasing pressures on the coastal marine environment. To better manage the environment an improved understanding of its future state is necessary. Improving the forecasting of temperature advances our knowledge in this direction. Whilst assimilation of SST is routinely carried out for oceans, its use has been minimal in coastal regions. Advances in computing power and the introduction of ensemble-based assimilation techniques have made the approach pursued in this thesis possible.

Previous coastal data assimilation had focused on sea level and tidal prediction. More recently, data assimilation has been undertaken with simple ecological models, but temperature has rarely been the subject of research in its own right. Furthermore, most coastal assimilation studies have assimilated point scale in situ observations, rather than satellite derived spatial observations, which were the focus of this thesis.

The thesis was conducted using a case study site, Port Phillip Bay (PPB): a large embayment in south eastern Australia. Various configurations of ensemble assimilation were tested in a synthetic setting by means of a twin experiment. Theoretically, a multivariate analysis should outperform a univariate analysis. However, in the PPB case, the low correlation between temperature and other variables meant that the multivariate analysis gave poor results due to sampling error in the ensem-

ble compounded by the small ensemble size. The effect of model bias on the analysis was explored by contrasting a biased assimilation run with an unbiased one. Model bias was found to remove the optimality of the analysis and distort the forecast error such that the assimilation estimate of the model forecast error relative to the truth was no longer accurate. But the assimilation of a biased model, still gave large improvements over an unassimilated forecast. The debate over stochastic and deterministic ensemble filters was approached by contrasting the EnKF and a form of EnSRF. No clear preference was established, however it was found that EnSRF maintained the shape of the ensemble anomalies better than the EnKF. This would be of importance if ensemble member independence was required.

The actual assimilation of SST data was contrasted against a climatology forecast of PPB temperature. The assimilation of SST without any specific accounting for the diurnal error improved the forecast, although errors due to observation error were noted. Attempts to remove this error using diurnal correction algorithms failed, owing to the larger than expected cool skin. Conditional Merging, which combines spatial and in situ observations was applied to the SST observations and improved the forecast accuracy by reducing the diurnal observation error. An examination of the assimilation innovations indicated how the forecast accuracy could be improved further.

7.1 Conclusions

Data assimilation research is a multidisciplinary field. Successful data assimilation requires, not only competence with the data assimilation equations and techniques, but also a detailed knowledge of numerical models, data collection, and a solid understanding of the errors associated with each. An understanding of the biogeochemical processes being modelled is also necessary. As such the findings of this thesis will be synthesised according to these three disparate aspects: i) data assimilation methods, ii) modelling and iii) observations.

7.1.1 Data Assimilation

One of the objectives of this thesis was to develop some guidance in the application of data assimilation to coastal applications. This was proposed because of the bewildering array of choice and limited explanation of the conditions under which each method is most appropriate in the application of ensemble techniques. Within this thesis the elements of the assimilation system that were explored were: i) the ensemble size, ii) ensemble initialisation, iii) forecast error incorporation, iv) multivariate or univariate analysis, and v) the analysis method. All of these features must be considered for a successful assimilation. However, the conclusions of this thesis are that for application to coastal modelling some elements have a larger bearing on the resulting forecast accuracy than others.

Specifically, ensemble size and the use of multivariate or univariate analysis are considered of greater impact to the performance of a coastal data assimilation study, while forecast error inclusion, analysis method, and ensemble initialisation are, in turn, considered of lesser importance.

The choice of an appropriate ensemble size is an aspect of the data assimilation that is of greater importance. This is because with increasing ensemble size the ensemble forecast error covariance matrix approaches the actual forecast error covariance matrix, assuming the system noise is accurately modelled. An indication of an appropriate ensemble size was based on the complexity of the system, and determined using the variance explained by the singular values of the system. Subsequent analysis confirmed that the recommended values were appropriate. Therefore the thesis does not recommend an optimal ensemble size, but recommends one based on the complexity of the coastal system being forecast. While there is not a distinct threshold, there appears to be an ensemble size range, below which the accuracy of forecast rapidly diminishes. Above this range increasing the ensemble size will bring improvement, but the marginal benefits tends to decrease and computational costs increase.

The choice of univariate or multivariate assimilation is also of high importance. This is interrelated with ensemble size, as a multivariate assimilation has a larger state space it therefore needs a larger ensemble size to explain an equivalent variance. The research found that including variables with low correlation can have a

detrimental impact on the analysis, not just on the unobserved variable, but also on the observed variable. This is significant for coastal applications because the collapse of geostrophic balance that generally holds in oceanic waters means that high correlations between variables may not exist, which is the case for PPB.

The impact of deteriorating forecast due to erroneous low correlations would be reduced by exploring the variable correlation structure prior to an assimilation study and discarding the analysis of those variables which have a low correlation with observed variables. An alternative approach would be to calculate correlations at analysis time. Those states where correlations were less than a threshold would be discarded from the analysis as they would be considered statistically equivalent to zero, or in other words uncorrelated. If analysis were desired for unobserved variables that appear to have a low correlation with the observed variable, a more satisfactory solution might be to seek direct observations of the desired variable instead.

The last three aspects of the data assimilation system are considered of secondary importance because they seek greater efficiency of the assimilation system once ensemble size and analysis variables have been determined. Forecast error incorporation and ensemble initialisation is concerned with spreading ensemble members and ensemble anomaly independence, whereas the analysis method seeks the most efficient way of performing the analysis.

The addition of forecast error is necessary to represent the increasing uncertainty of a model forecast as model errors and forcing data errors are compounded. The accuracy of the forecast error impacts on the optimality of the analysis as the observation error is usually determined a priori: the ratio of the observation error to the forecast error is the basis for the analysis. An important finding of this thesis was that the forecast errors due to forcing data errors were at least as significant as model errors. This prompted the development of a method to generate perturbed forcing data that reduced the possibility of introducing model bias. Forcing data errors were pursued because their magnitude could be easily calculated by analysing spatial uncertainty in the forcing data. Testing found that perturbed forcing error could replicate the actual forecast error. Furthermore, it was shown that by using perturbed forcing data the forecast error was distributed unevenly through the domain. This finding is valuable because the forecast error distribution is more

realistic than the uniform model error added using the predominant existing method to incorporate model error. This finding is salient because in coastal models forcing data impacts more on the results than perhaps atmospheric climate models, where initial conditions and model nonlinearities are more significant.

A review of the ensemble data assimilation literature seems to indicate that the correct specification of forecast error does not appear to have been given the same degree of attention as have other elements, such as the analysis method. This is unfortunate as the forecast error was found to be more influential on the results than the analysis method. Whereas in this thesis forcing data error alone was used to improve realism and the spatial distribution of the forecast error. In future coastal forecasting studies a combination of model and forcing error would probably be best. Based on the conclusions of the thesis forecast error methods that introduce error more realistically should be applied. In addition to perturbed forcing data, using stochastic model equations and or perturbing the model parameters would appear to introduce error best in this regard.

The findings of this thesis are that the choice of ensemble filter for the analysis is not vitally important, although the choice influences the resulting shape and independence of the ensemble anomalies. An issue for further research is the conclusion that the observations error is not independent. For the EnKF, this would result in a more complicated generation of observation perturbations to reflect differing covariances. For the EnSRF calculating the square root of a more complicated matrix might also be more challenging. As knowledge of the complexity of the observation error covariance improves (e.g. Eresmaa and Järvinen, 2005), whichever filter can better accommodate a more complex \mathbf{R} will be preferable.

The importance of the ensemble initialisation was not studied in this thesis, however in one sense it could be argued that how an ensemble is initialised is not actually important for coastal hydrodynamic models because they are dispersive: over time the influence of errors/perturbations in the initial conditions diminishes and the influence of forcing data increases. This contrasts with other highly non-linear models that are particularly sensitive to changes in initial conditions. The importance of ensemble initialisation could be tested by exploring the variation in ensemble anomaly rank (or the linear dependency of the anomaly vectors) as the ensemble evolves.

7.1.2 Numerical Modelling

An understanding of numerical modelling is fundamental to data assimilation. The key finding of this thesis in regards to modelling relates to model bias. While the assimilation equations are based on assumptions that model forecasts are unbiased, the forecasting skill using a model with a biased heatflux module was greatly improved using assimilation. Nevertheless, an unbiased model gives better results. A serious failing of the assimilation system when a biased model is used is the underestimation of the error margin associated with the forecast.

At the conclusion of the thesis, the main source of residual error in the forecast was due to model bias, thought to be related to the heatflux module. Two possible ways of removing this bias are through improvements to the model structure or through applying a bias aware assimilation filter. This thesis considers that addressing errors in the model structure would be more beneficial than attempting a bias correction filter. The view (of removing bias via improving the model structure) stems from an understanding that data assimilation should not be used to redeem a poor model (Walker *et al.*, 2001), although data assimilation can be used to reveal faults in a model. Initial attempts at applying a bias-aware assimilation filter (not presented in this thesis) were disappointing, probably because the bias varies diurnally rather than with a constant offset. Moreover, anecdotally other heatflux modules are known to be more accurate than the Gill formulation used in MECO (Pers. Comm. M. Herzfeld). Therefore, a conclusion of this thesis is to reduce any model bias, wherever possible. If model bias is significant the effort made to improve data assimilation techniques, such as improving the accuracy of the forecast covariance error structure, described above, would be better spent improving the model structure.

From an investigation of the innovations a specific deficiency in the model was found. A climatology open boundary forcing for temperature in Bass Strait was used because no temperature record in Bass Strait was available. Due to the low flow exchange between Bass Strait and PPB, this did not affect the temperature forecast in PPB (at least where measurements were recorded in the northern half), however this did induce problems with a systematic underestimation of Bass Strait temperature during an unseasonably warm period when temperature diverged from

the climatology. This was discovered through histogram and spatial plots of the innovations. After the analysis adjusted the temperature towards the correct SST, the open boundary rapidly re-established the incorrect climatology temperature. The proposed solution is to change the open boundary conditions to a zero-flux boundary relaxing to clamped boundary with time after an analysis. Such a boundary condition would ensure that if no other information (observations) were available the open boundary would be driven by the climatology forcing. But when more recent information from observations became available, it should be used in preference, then as the relevance of this (new) information decreases with time, more credence would correspondingly be given to the climatology.

7.1.3 Observations

The final aspect of the data assimilation system is the observations that are assimilated into the model. Issues raised by this thesis are i) remotely sensed observation bias, ii) the observations error covariance structure, and iii) the use of innovation to discern errors in observations.

The examination of the satellite observations revealed that although the observations had negligible bias over the long term, a diurnal error remained. This indicates that the satellite retrieval algorithms do not account for all processes and additional correction of satellite observations using a technique like conditional merging, which combines a less accurate spatial image with accurate point scale measurements to improve the accuracy of the entire image, may be appropriate.

Conditional merging was found to be a useful tool to remove error in the satellite observations. This conclusion is of generic benefit because improving the accuracy of and removing error from satellite images is an issue for all satellite observations and not merely SST imagery. Conditional merging is appropriate where the collection of a few in situ observations to augment the satellite observation is possible. It is of particular benefit where the accuracy of the satellite observations is known to be suspect, but the relative pattern is trusted.

The conditional merging removed diurnal error in SST observations that could not be removed by retrieval algorithms. Furthermore, it was successful because the observation errors had large correlation length scales. This allowed observations

from relatively few in situ sites to correct the entire satellite image. However, while the conditional merging improved the forecast skill, it should be used judiciously.

In situ observations may not always be present at the top of the water column, and using in situ observations from deeper in the water column would introduce other uncertainties. Furthermore, the spatial interpolation of the observation error has been made assuming that the process is stationary and that the observation error is isotropic. While the correlation of the observation error between the in situ sites indicates that the relationship holds over long distances, it may not merely be a function of distance, and other factors such as proximity to land, water depth and currents might influence the observations bias in some way. In this case a spatial interpolation method which takes into account these other factors would give a more accurate conditionally merged observation (e.g co-kriging or external drift kriging). Alternatively, using the spatial structure of the model itself might be beneficial.

If an in situ network was to be set up to apply conditional merging to satellite observations, the optimal spacing would be the correlation length scale of the observation bias. For SST this is the correlation length scale of the diurnal warming as this is the mechanism which is inducing the observations bias; for other variables where the atmospheric retrieval is less accurate the correlation length scale might be the length scale of weather systems, with a correspondingly larger spacing of the in situ sites. For PPB the use of 3 in situ sites appears reasonable, but they could probably be more efficiently spaced out, for instance by placing a site in the Geelong Arm or in the south east of PPB.

The SST observations were taken as having an error standard deviation of 0.5°C , based on the quoted instrument accuracy. Analysis of the observation error showed that this value was appropriate. The thesis also assumed that this error was independent, however the standard deviation of the innovations underestimated the expected value, which indicated that the observation error was significantly correlated.

Correlation in the observation errors reduces the value of each individual observation, and strengthens the need for a technique like conditional merging to remove observation bias. Because of the conditional merging the correlation of the observation errors is not thought to be problematic, however the assumption of independent observations is not ideal and a more accurate observation covariance structure should

be sought. Furthermore, the introduction of in situ information through conditional merging increases the observation accuracy near in situ locations. This information should also be reflected in the an observation error covariance matrix.

The investigation of the innovations led to the discovery of an erroneous image, which had not previously been determined by the cloud detection algorithms or the pre-analysis checks. The bimodal nature of the innovations enabled the detection of the image. Operational forecasting systems should therefore perform real time checking of the innovations as observations are being assimilated. Suspect images could then be further processed prior to analysis, or discarded and the forecast continued.

7.2 Future Directions

A number of recommendations for future research is made based on the findings of this thesis. These recommendations apply both to the data assimilation techniques and to applications of forecasting in coastal waters.

To further improve the ensemble data assimilation techniques used in this thesis, attention should focus on improving the accuracy of the error covariance matrices: both for the observations and forecast. The conclusions indicated that the forecast error covariance is most easily improved by increasing the ensemble size, however this is externally limited by computing power. Instead this thesis recommends that greater attention be given to the addition of forecast error to the ensemble forecast.

The techniques that should be pursued are those that introduce error into the ensemble realistically, either through the model equations (and or parameter values), or through the forcing data. This allows the forecast error to be targeted with greater accuracy. The resulting forecast error should vary spatially according to uncertainty and be correlated (between states and variables) according to the physical understanding of the processes, expressed through the numerical model.

The procedure developed in this thesis for introducing perturbed forcing error could be further improved by extending it spatially and incorporating temporal correlations to the error, although temporal correlation is partially introduced via the offset term. These aspects were deemed unnecessary for the PPB case study because of the limited domain, but would be necessary in other applications that

extend over larger domains. If a time varying surface was used for forcing instead of a time series, at a given time the perturbed error term would be the multiplication of a (smooth) random surface (rather than a random number) with a standard deviation, which would vary depending on the data type: restricted, semi-restricted, or unrestricted. The covariances of the surface would depend on the characteristics of the data. Furthermore, time correlation of perturbed forcing data can be included by adding a suitable autocorrelation term to the random number generator producing the timeseries. The error-offset form of the perturbation, added to the forcing data appears to operate well, but other forms of perturbation are possible and should be explored further.

For the observation error covariance, a better understanding of the covariances for satellite derived observations should be gained. Furthermore, as conditional merging reduces the uncertainty of the observation, by incorporating more accurate in situ observations, its affect on the observation covariance matrix should be explored. Necessary actions are to improve the accuracy of the semivariogram, and test if the temperature error correlations are isotropic or influenced by other factors which would impact on the spatial interpolation employed.

The ensemble efficiency could be explored by investigating the shape of the ensemble anomalies over time and whether there is a tendency for the anomalies to converge. This would indicate that the rank of the ensemble was reducing, with a corresponding reduction in efficiency. If this was found a mechanism to maintain the ensemble anomaly rank should be sought.

The success of this thesis in assimilating satellite derived SST observations into a hydrodynamic model to improve the temperature forecasting prompts the recommendation that the application be extended to the forecasting of other coastal biogeochemical processes that are observed by satellite. Suspended sediment and ecological applications are the most obvious direct applications. In both of these applications the observation accuracy is an issue so that conditional merging is likely to be beneficial.

Finally, for the specific case study explored in this thesis, the observations assimilated were derived from the NOAA AVHRR instrument. While this is still the current operational SST observation platform, more recent, new-generation instruments such as ASTER and AATSR are available. They have greater accuracy,

although their coverage is on-demand rather than continuous. Their use in future studies is recommended.

The motivation of this thesis was to form a step towards the development of an operational coastal marine forecasting system. Such a system would present managers with the information they need to protect and maintain the coastal marine environment. Undoubtedly, efforts are being made in this direction with prototype systems under development and within a few years such systems may become routine. By improving the accuracy of the forecast error covariance through the development of perturbed forcing data, removing bias from observations, and extending the assimilation of SST to high resolution coastal models this thesis has made practical developments in this direction.

Bibliography

- Allen, J. I., M. Eknes, and G. Evensen. An ensemble Kalman filter with a complex marine ecosystem model: hindcasting phytoplankton in the Cretan Sea. *Annales Geophysicae*, 20:1–13, 2002.
- Anderson, J. L. An ensemble adjustment Kalman filter for data assimilation. *Monthly Weather Review*, 129(12):2884–2903, 2001.
- Anderson, J. L. and S. L. Anderson. A monte carle implementation of the nonlinear filtering problem to produce ensemble assimilations and forecasts. *Monthly Weather Review*, 127(12):2741–2758, 1999.
- Annan, J. D. and J. C. Hargreaves. Sea surface temperature assimilation for a three-dimensional baroclinic model of shelf seas. *Continental Shelf Research*, 19(11):1507–1520, 1999.
- ANTT. *Australian national tide tables: Australia and Papua New Guinea*. Australian Hydrographic Service, The Royal Australian Navy, Canberra, 2006.
- Arhonditsis, G. B. and M. T. Brett. Evaluation of the current state of mechanistic aquatic biogeochemical modeling. *Marine Ecology Progress Series*, 271:13–26, 2004.
- Babovic, V. and D. R. Fuhrman. Data assimilation of local model error forecasts in a deterministic model. *International Journal for Numerical Methods in Fluids*, 39:887–918, 2002. doi:10.1002/fld.350.
- Balci, O. Principles of simulation model validation, verification, and testing. *Transactions of the Society for Computer Simulation*, 14(1):3–12, 1997.

- Ballabrera-Poy, J., P. Brasseur, and J. Verron. Dynamical evolution of the error statistics with the SEEK filter to assimilate altimetric data in eddy-resolving ocean models. *Quarterly Journal of the Royal Meteorological Society*, 127(571):233–253, 2001.
- Barham, P. M. and D. E. Humphries. Derivation of the Kalman filtering equations from elementary statistical principles. In C. T. Leondes, editor, *Theory and Applications of Kalman Filtering*, AGARDograph No. 139, pp. 45–49. NATO, Advisory Group for Aerospace Research and Development, 1970.
- Baumert, H., G. Chapalain, H. Smaoui, J. P. McManus, H. Yagi, M. Regener, J. Sündermann, and B. Szilagyi. Modelling and numerical simulation of turbulence, waves and suspended sediments for pre-operational use in coastal seas. *Coastal Engineering*, 41(1-3):63–93, 2000.
- Bertino, L., G. Evensen, and H. Wackernagel. Combining geostatistics and Kalman filtering for data assimilation in an estuarine system. *Inverse Problems*, 18:1–23, 2002.
- Bishop, C. H., B. J. Etherton, and S. J. Majumdar. Adaptive sampling with the ensemble transform Kalman filter. Part I: Theoretical aspects. *Monthly Weather Review*, 129(3):420–436, 2001.
- Blumberg, A. F. and G. L. Mellor. A description of a three-dimensional coastal ocean circulation model. In N. S. Heaps, editor, *Three-Dimensional Coastal Ocean Models*, pp. 1–16. American Geophysical Union, 1987.
- Borovikov, A., M. M. Rienecker, C. L. Keppenne, and G. C. Johnson. Multivariate error covariance estimates by monte carlo simulation for assimilation studies in the pacific ocean. *Monthly Weather Review*, 133(8):2310–2334, 2005.
- Bouttier, F. and P. Courtier. Data assimilation concepts and methods, 1999. Training course notes, ECMWF.
- Braconnot, P. and C. Frankignoul. Testing model simulations of the thermocline depth variability in the tropical Atlantic from 1982 through 1984. *Journal of Physical Oceanography*, 23(4):626–647, 1993.

- Brown, O. B. and P. J. Minnett. MODIS infrared sea surface temperature algorithm. Algorithm Theoretical Basis Document ATBD-MOD-25, NASA, 1999. Version 2.0.
- Brusdal, K., J. M. Brankart, G. Halberstadt, G. Evensen, P. Brasseur, P. J. van Leeuwen, E. Dombrowsky, and J. Verron. A demonstration of ensemble-based assimilation methods with a layered OGCM from the perspective of operational ocean forecasting systems. *Journal of Marine Systems*, 40-41:253–289, 2003.
- Buizza, R., M. Miller, and T. N. Palmer. Stochastic representation of model uncertainties in the ECMWF ensemble prediction system. *Quarterly Journal of the Royal Meteorological Society*, 125(560):2887–2908, 1999.
- Burchard, H., O. Peterson, and T. P. Rippeth. Comparing the performance of the Mellor Yamada and the k- ϵ two-equation turbulence models. *Journal of Geophysical Research*, 103(C5):10,543–10,554, 1998.
- Burgers, G., P. J. van Leeuwen, and G. Evensen. On the analysis scheme in the ensemble Kalman filter. *Monthly Weather Review*, 126:1719–1724, 1998.
- Cane, M. A., A. Kaplan, R. N. Miller, B. Y. Tang, E. C. Hackert, and A. J. Busalacchi. Mapping tropical pacific sea level: Data assimilation via a reduced state space Kalman filter. *Journal of Geophysical Research*, 101(C10):22,599–22,617, 1996.
- Cornillon, P. and L. Stramma. The distribution of diurnal sea surface warming events in the western Sargasso Sea. *Journal of Geophysical Research*, 90(C6):11,811–11,815, 1985.
- Courtier, P. and O. Talagrand. Variational assimilation of meteorological observations with the adjoint vorticity equation II: numerical results. *Quarterly Journal of the Royal Meteorological Society*, 113(478):1329–1347, 1987.
- Cracknell, A. P. Remote sensing techniques in estuaries and coastal zones—an update. *International Journal of Remote Sensing*, 19(3):485–496, 1999.
- Cushman-Roisin, B. *Introduction to Geophysical Fluid Dynamics*. Prentice Hall, New Jersey, 1994. Pp. 320.

- Daescu, D. N. and G. R. Carmichael. An adjoint sensitivity method for the adaptive location of the observations in air quality modeling. *Journal of the Atmospheric Sciences*, 60(2):434–450, 2003.
- Daley, R. *Atmospheric data analysis*. Cambridge University Press, Melbourne, 1991. Pp. 457.
- De Mey, P. Data assimilation at the oceanic mesoscale: A review. *Journal of the Meteorological Society of Japan*, 75(1B):415–427, 1997.
- Dee, D. P. Simplification of the Kalman filter for meteorological data assimilation. *Quarterly Journal of the Royal Meteorological Society*, 117(498):365–384, 1991.
- Dee, D. P. A pragmatic approach to model validation. In D. R. Lynch and A. M. Davies, editors, *Quantitative Skill Assessment for Coastal Ocean Models*, no. 47 in Coastal Estuarine Studies, chap. 1. American Geophysical Union, 1995.
- Dee, D. P. and A. M. da Silva. Data assimilation in the presence of forecast bias. *Quarterly Journal of the Royal Meteorological Society*, 124:269–295, 1998.
- Dee, D. P. and R. Todling. Data assimilation in the presence of forecast bias: The GEOS moisture analysis. *Monthly Weather Review*, 128(9):3268–3282, 2000.
- Deschamps, P. Y. and T. Phulpin. Atmospheric correction of infrared measurements of sea surface temperature using channels at 3.7, 11 and 12 μm . *Boundary-Layer Meteorology*, 18(2):131–143, 1980.
- Dowd, M. and K. R. Thompson. Forecasting coastal circulation using an approximate Kalman filter based on dynamical modes. *Continental Shelf Research*, 17(14):1715–1735, 1997.
- Eknes, M. and G. Evensen. An ensemble Kalman filter with a 1-D marine ecosystem model. *Journal of Marine Systems*, 36:75–100, 2002.
- Emery, W. J., S. Castro, G. A. Wick, P. Schluessel, and C. Donlon. Estimating sea surface temperature from infrared satellite and in situ temperature data. *Bulletin of the American Meteorological Society*, 82(12):2773–2785, 2001.

- Eresmaa, R. and H. Järvinen. Estimation of spatial global positioning system zenith delay observation error covariance. *Tellus A*, 57(2):194–203, 2005.
- Evensen, G. Using the extended Kalman filter with a multilayer quasi-geostrophic ocean model. *Journal of Geophysical Research*, 97(C11):17,905–17,924, 1992.
- Evensen, G. Sequential data assimilation with a nonlinear quasi-geostrophic model using Monte Carlo methods to forecast error statistics. *Journal of Geophysical Research*, 99(C5):10,143–10,162, 1994.
- Evensen, G. The ensemble Kalman filter: theoretical formulation and practical implementation. *Ocean Dynamics*, 53:343–367, 2003.
- Evensen, G. Sampling strategies and square root analysis schemes for the EnKF. *Ocean Dynamics*, 54:539–560, 2004.
- Evensen, G. and P. J. van Leeuwen. Assimilation of Geosat altimeter data for the Agulhas Current using the ensemble Kalman filter with a quasigeostrophic model. *Monthly Weather Review*, 124(1):85–96, 1996.
- Farrell, B. F. and P. J. Ioannou. Perturbation growth and structure in time-dependent flows. *Journal of the Atmospheric Sciences*, 56:3622–3639, 1999.
- Faugeras, B., M. Levy, L. Memery, J. Verron, J. Blum, and I. Charpentier. Can biogeochemical fluxes be recovered from nitrate and chlorophyll data? A case study assimilating data in the northwestern Mediterranean Sea at the JGOFS-DYFAMED station. *Journal of Marine Systems*, 40-41:99–125, 2003.
- Fennel, K., M. Losch, J. Schröter, and M. Wenzel. Testing a marine ecosystem model: sensitivity analysis and parameter optimization. *Journal of Marine Systems*, 28(1-2):45–63, 2001.
- Fieguth, P. W., D. Menemenlis, and I. Fukumori. Mapping and pseudoinverse algorithms for ocean data assimilation. *IEEE Transactions of Geoscience and Remote Sensing*, 41(1):43–51, 2003.
- Flather, R. A. Existing operational oceanography. *Coastal Engineering*, 41(1-3):13–40, 2000.

- Fu, L. L., I. Fukumori, and R. N. Miller. Fitting dynamic-models to the Geosat sea-level observations in the tropical Pacific-Ocean. part II: A linear, wind-driven model. *Journal of Physical Oceanography*, 23(10):2162–2181, 1993.
- Garcia-Gorriz, E., N. Hoepffner, and M. Ouberdous. Assimilation of SeaWiFS data in a coupled physical-biological model of the Adriatic Sea. *Journal of Marine Systems*, 40-41:233–252, 2003.
- Gentemann, C. L., C. J. Donlon, A. Stuart-Menteth, and F. J. Wentz. Diurnal signals in satellite sea surface temperature measurements. *Geophysical Research Letters*, 30(3), 2003. doi:10.1029/2002GL016291.
- Gerritsen, H., R. J. Vos, T. van der Kaaij, A. Lane, and J. G. Boon. Suspended sediment modelling in a shelf sea (North Sea). *Coastal Engineering*, 41:317–352, 2000.
- Ghil, M. and P. Malanotte-Rizzoli. Data assimilation in meteorology and oceanography. *Advances in Geophysics*, 33:141–266, 1991.
- Gill, A. *Atmosphere-Ocean Dynamics*. Academic Press, London, 1982. Pp. 662.
- Grassl, H. The dependence of the measured cool skin of the ocean on wind stress and total heat flux. *Boundary-Layer Meteorology*, 10:465–474, 1976.
- Haines, K. A direct method for assimilating sea surface height data into ocean models with adjustments to the deep circulation. *Journal of Physical Oceanography*, 21(6):843–868, 1991.
- Hallock, Z. R., P. Pistek, J. W. Book, J. L. Miller, and H. T. Shay, L. K. and Perkins. A description of tides near the Chesapeake Bay entrance using in situ data with an adjoint model. *Journal of Geophysical Research*, 108(C3), 2003. doi:10.1029/2001JC000820.
- Hamill, T. Ensemble-based data assimilation: A review, 2002. University of Colorado and NOAA-CIRES Climate Diagnostics Centre, Unpublished manuscript.

- Hamill, T. M., J. S. Whitaker, and C. Snyder. Distance-dependent filtering of background error covariance estimates in an ensemble Kalman filter. *Monthly Weather Review*, 129(11):2776–2790, 2001.
- Harris, G., G. Batley, D. Fox, D. Hall, P. Jernakoff, R. Molloy, A. Murray, B. Newell, J. Parslow, G. Skyring, and S. Walker. *Port Phillip Bay Environmental Study Final Report*. CSIRO, Canberra, Australia, 1996.
- Hasse, L. The sea surface temperature deviation and the heat flow at the sea-air interface. *Boundary-Layer Meteorology*, 1:368–379, 1971.
- Heath, M. T. *Scientific Computing: An Introductory Survey*. McGraw-Hill, Boston, 1997.
- van Hees, F., A. J. van der Steen, and P. J. van Leeuwen. A parallel data assimilation model for oceanographic observations. *Concurrency and Computing: Practice and Experience*, 15:1191–1204, 2003.
- Herzfeld, M., J. Waring, J. Parslow, N. Margvelashvili, P. Sakov, and J. Andrewartha. *MECO- Model for Estuaries and Coastal Oceans V4.0 Scientific Manual*. CSIRO Marine Research, 1.0 ed., 2002.
- Hoang, S., R. Baraille, O. Talagrand, X. Carton, and P. De Mey. Adaptive filtering: application to satellite data assimilation in oceanography. *Dynamics of Atmospheres and Oceans*, 27(1-4):257–281, 1997.
- Houser, P. R., W. J. Shuttleworth, J. S. Famiglietti, H. V. Gupta, K. H. Syed, and D. C. Goodrich. Integration of soil moisture remote sensing and hydrologic modeling using data assimilation. *Water Resources Research*, 34(12):3405–3420, 1998.
- Houtekamer, P. L., L. Lefaire, J. Derome, H. Ritchie, and H. L. Mitchell. A system simulation approach to ensemble prediction. *Monthly Weather Review*, 124(6):1225–1242, 1996.
- Houtekamer, P. L. and H. L. Mitchell. Data assimilation using an ensemble Kalman filter technique. *Monthly Weather Review*, 126(3):796–811, 1998.

- Houtekamer, P. L. and H. L. Mitchell. A sequential ensemble Kalman filter for atmospheric data assimilation. *Monthly Weather Review*, 129(1):123–137, 2001.
- Ishikawa, Y., T. Awaji, and K. Akitomo. Successive correction of the mean sea surface height by the simultaneous assimilation of drifting buoy and altimetric data. *Journal of Physical Oceanography*, 26(11):2381–2397, 1996.
- James, I. D. Modelling pollution dispersion, the ecosystem and water quality in coastal waters: a review. *Environmental Modelling and Software*, 17(4):363–385, 2002.
- Janssen, P. H. M. and P. S. C. Heuberger. Calibration of process-oriented models. *Ecological Modelling*, 83(1-2):55–66, 1995.
- Johannessen, O. M., S. Sandven, A. D. Jenkins, D. Durand, L. H. Pettersson, H. Espedal, G. Evensen, and T. Hamre. Satellite earth observation in operational oceanography. *Coastal Engineering*, 41(1-3):155–176, 2000.
- Journel, A. G. and C. J. Huijbregts. *Mining Geostatistics*. Academic Press, New York, 1978.
- Kalman, R. E. A new approach to linear filtering and prediction problems. *Transactions of the ASME-Journal of Basic Engineering*, 82(Series D):35–45, 1960.
- Kawai, Y. and H. Kawamura. Evaluation of the diurnal warming of sea surface temperature using satellite-derived marine meteorological data. *Journal of Oceanography*, 58:805–814, 2002.
- Kawai, Y. and H. Kawamura. Validation of daily amplitude of sea surface temperature evaluated with a parametric model using satellite data. *Journal of Oceanography*, 59:637–644, 2003.
- Kelley, J. W., D. W. Behringer, H. J. Thiebaut, and B. Balasubramaniyan. Assimilation of SST into a real-time coastal ocean forecasting system for the U.S. East Coast. *Weather and Forecasting*, 17(4):670–691, 2002.
- Kepert, J. D. On ensemble representation of the observation-error covariance in the ensemble Kalman filter. *Ocean Dynamics*, 54:561–569, 2004.

- Keppenne, C. L. Data assimilation into a primitive-equation model with a parallel ensemble Kalman filter. *Monthly Weather Review*, 128(6):1971–1981, 2000.
- Keppenne, C. L. and M. M. Rienecker. Assimilation of temperature into an isopycnal ocean general circulation model using a parallel ensemble Kalman filter. *Journal of Marine Systems*, 40-41:363–380, 2003.
- Keppenne, C. L., M. M. Rienecker, N. P. Kurkowski, and D. A. Adamec. Ensemble Kalman filter assimilation of temperature data with bias correction and application to seasonal prediction. *Nonlinear Processes in Geophysics*, 12:491–503, 2005.
- Knauss, J. A. *Introduction to Physical Oceanography*. Prentice Hall, New Jersey, 1997.
- van Leeuwen, P. J. Comment on “Data assimilation using an ensemble Kalman filter technique”. *Monthly Weather Review*, 127(6):1374–1377, 1999.
- Lermusiaux, P. F. J. and A. R. Robinson. Data assimilation via error subspace statistical estimation. Part I: Theory and schemes. *Monthly Weather Review*, 127(7):1385–1407, 1999.
- Li, X., W. Pichel, P. Clemente-Colón, V. Krasnopolsky, and J. Sapper. Validation of coastal sea and lake surface temperature measurements derived from NOAA/AVHRR data. *International Journal of Remote Sensing*, 22(7):1285–1303, 2001.
- Liou, K. N. *An introduction to atmospheric radiation*, vol. 84 of *International Geophysical Series*. Academic Press, London, 2 ed., 2002.
- Lorenc, A. C., S. Ballard, R. Bell, N. Ingleby, P. Andrews, D. Barker, J. Bray, A. Clayton, T. Dalby, D. Li, T. Payne, and F. Saunders. The Met. Office global three-dimensional variational data assimilation scheme. *Quarterly Journal of the Royal Meteorological Society*, 126(570):2991–3012, 2000.
- Lorenc, A. C., R. S. Bell, and B. MacPherson. The meteorological office analysis correction data assimilation scheme. *Quarterly Journal of the Royal Meteorological Society*, 117(497):59–89, 1991.

- Madsen, H. and R. Cañizares. Comparison of extended and ensemble Kalman filters for data assimilation in coastal area modelling. *International Journal for Numerical Methods in Fluids*, 31:961–981, 1999.
- Malanotte-Rizzoli, P. M. and R. E. Young. Gulf Stream system assimilation experiments: A sensitivity study. *Journal of Atmospheric and Oceanic Technology*, 14(6):1392–1408, 1997.
- McClain, E. P., W. G. Pichel, and C. C. Walton. Comparative performance of AVHRR-based multi-channel sea surface temperatures. *Journal of Geophysical Research*, 90(C14):11,587–11,601, 1985.
- Mellor, G. L. and T. Yamada. Development of a turbulence closure model for geophysical fluid problems. *Reviews of Geophysics*, 20:851–875, 1982.
- Merchant, C. J., A. R. Harris, M. J. Murray, and A. M. Závody. Toward the elimination of bias in satellite retrievals of sea surface temperature 1. Theory, modeling and interalgorithm comparison. *Journal of Geophysical Research*, 104(C10):23,565–23,578, 1999.
- Miller, R. N. and L. L. Ehret. Ensemble generation for models of multimodal systems. *Monthly Weather Review*, 130(9):2313–2333, 2002.
- Mitchell, H. L. and P. L. Houtekamer. An adaptive ensemble Kalman filter. *Monthly Weather Review*, 128(2):416–433, 2000.
- Mitchell, H. L., P. L. Houtekamer, and G. Pellerin. Ensemble size, balance and model-error representation in an ensemble Kalman filter. *Monthly Weather Review*, 130(11):2791–2808, 2002.
- Molteni, F., R. Buizza, T. N. Palmer, and T. Petroliagis. The ECMWF ensemble prediction system: Methodology and validation. *Quarterly Journal of the Royal Meteorological Society*, 122(529):73–119, 1996.
- Moore, A. M., N. S. Cooper, and D. L. T. Anderson. Initialization and data assimilation in models of the Indian Ocean. *Journal of Physical Oceanography*, 17(11):1965–1977, 1987.

- Nalli, N. R. and W. L. Smith. Improved remote sensing of sea surface skin temperature using a physical retrieval method. *Journal of Geophysical Research*, 103(C5):10,527–10,542, 1998.
- Natvik, L. J. and G. Evensen. Assimilation of ocean colour data into a biochemical model of the North Atlantic - Part 1. Data assimilation experiments. *Journal of Marine Systems*, 40-41:127–153, 2003a.
- Natvik, L. J. and G. Evensen. Assimilation of ocean colour data into a biochemical model of the North Atlantic - Part 2. Statistical analysis. *Journal of Marine Systems*, 40-41:155–169, 2003b.
- Navon, I. M. Practical and theoretical aspects of adjoint parameter estimation and identifiability in meteorology and oceanography. *Dynamics of Atmospheres and Oceans*, 27(1-4):55–79, 1997.
- Oke, P. R., J. S. Allen, R. N. Miller, G. D. Egbert, and P. M. Kosro. Assimilation of surface velocity data into a primitive equation coastal ocean model. *Journal of Geophysical Research-Oceans*, 107(C9), 2002. doi:1029/2000JC000511.
- Oke, P. R., A. Schiller, D. A. Griffin, and G. B. Brassington. Ensemble data assimilation for an eddy-resolving ocean model. *submitted to: Quarterly Journal of the Royal Meteorological Society*, 2005.
- Oreskes, N., K. Shrader-Frechette, and K. Belitz. Verification, validation, and confirmation of numerical models in the earth sciences. *Science*, 263(5147):641–646, 1994.
- Ott, E., B. R. Hunt, I. Szunyogh, A. V. Zimin, E. J. Kostelich, M. Corazza, E. Kalnay, D. J. Patil, and J. A. Yorke. A local ensemble Kalman filter for atmospheric data assimilation. *Tellus*, 56A:415–428, 2004.
- Pacanowski, R. C. and S. M. Griffies. The MOM 3 manual. Tech. rep., Geophysical Fluid Dynamics Laboratory, Princeton, NJ, USA, 1988.
- Pebesma, E. J. *Gstat User's Manual*, 2001.

- Pham, D. T. Stochastic methods for sequential data assimilation in strongly non-linear systems. *Monthly Weather Review*, 129(5):1194–1207, 2001.
- Power, M. The predictive validation of ecological and environmental models. *Ecological Modelling*, 68(1-2):33–50, 1993.
- Prandle, D. Introduction. operational oceanography in coastal waters. *Coastal Engineering*, 41(1-3):3–12, 2000a.
- Prandle, D. Operational oceanography - a view ahead. *Coastal Engineering*, 41(1-3):353–359, 2000b.
- Prandle, D., J. C. Hargreaves, J. P. McManus, A. R. Campbell, K. Duwe, A. Lane, P. Mahnke, S. Shimwell, and J. Wolf. Tide, wave and suspended sediment modelling on an open coast – Holderness. *Coastal Engineering*, 41(1-3):237–267, 2000.
- Preisendorfer, R. W. *Principal component analysis in meteorology and oceanography*, vol. 17 of *Developments in atmospheric science*. Elsevier Science Pub. Co., Amsterdam, 1988. Posthumously compiled and edited by Curtis D. Mobley.
- Press, W. H., S. A. Teukolsky, W. T. Vetterling, and B. P. Flannery. *Numerical Recipes in C: The Art of Scientific Computing*. Cambridge University Press, Cambridge, second ed., 1992.
- Reichle, R. H., D. B. McLaughlin, and D. Entekhabi. Hydrologic data assimilation with the ensemble Kalman filter. *Monthly Weather Review*, 130(1):103–114, 2002a.
- Reichle, R. H., J. P. Walker, R. D. Koster, and P. R. Houser. Extended versus ensemble Kalman filtering for land data assimilation. *Journal of Hydrometeorology*, 3(6):728–740, 2002b.
- Robert, C. and O. Alves. Tropical pacific ocean model error covariance from Monte-Carlo simulations. Tech. rep., Bureau of Meteorology Research Centre, Melbourne, Australia, 2003. To be submitted.
- Ruddick, K. G., F. Ovidio, and M. Rijkeboer. Atmospheric correction of SeaWiFS imagery for turbid coastal and inland waters. *Applied Optics*, 39(6):897–912, 2000.

- Rykiel, E. J. Testing ecological models: the meaning of validation. *Ecological Modelling*, 90(3):229–244, 1996.
- Saunders, P. The temperature at the ocean-air interface. *Journal of Atmospheric Science*, 24(3):269–273, 1967.
- Schluessel, P., W. J. Emery, H. Grassl, and T. Mammen. On the bulk-skin temperature difference and its impact on satellite remote sensing of sea surface temperature. *Journal of Geophysical Research*, 95(C8):13,341–13,356, 1990.
- Shchepetkin, A. F. and J. C. McWilliams. The regional oceanic modeling system (ROMS): a split-explicit, free-surface, topography-following-coordinate oceanic model. *Ocean Modelling*, 9:347–404, 2005.
- Shulman, I., C.-R. Wu, J. K. Lewis, J. D. Paduan, L. K. Rosenfeld, J. C. Kindle, S. R. Ramp, and C. A. Collins. High resolution modeling and data assimilation in the Monterey Bay area. *Continental Shelf Research*, 22:1129–1151, 2002.
- Siefert, R. L. and G.-K. Plattner. The role of coastal zones in global biogeochemical cycles. *EOS, Transactions, American Geophysical Union*, 85(45):469, 2004.
- Sinclair, S. and G. Pegram. Combining radar and rain gauge rainfall estimates for flood forecasting in South Africa. In *Sixth International Symposium on Hydrological Applications of Weather Radar*. Melbourne, Australia, 2004.
- Snedecor, G. W. and W. G. Cochran. *Statistical Methods*. Iowa State University Press, Ames, Iowa, seventh edition ed., 1980. 507 pages.
- Solidoro, C., A. Crise, G. Crispi, and R. Pastres. An a priori approach to assimilation of ecological data in marine ecosystem models. *Journal of Marine Systems*, 40-41:79–97, 2003.
- Spiegel, M. R. and L. J. Stephens. *Theory and Problems of Statistics*. McGraw-Hill, New York, 3rd ed., 1998. 538 pages - Schaum's outline of.
- Stegmann, P. M., D. G. Foley, C. King, F. B. Schwing, H. Price, S. J. Bograd, and D. M. Palacios. Integrating enhanced satellite data maps into coastal management. *EOS, Transactions*, 87(14):137,142, 2006.

- Stuart-Menteth, A. C., I. S. Robinson, and P. G. Challenor. A global study of diurnal warming using satellite-derived sea surface temperature. *Journal of Geophysical Research*, 108(C5), 2003. doi:10.1029/2002JC001534.
- Talagrand, O. and P. Courtier. Variational assimilation of meteorological observations with the adjoint vorticity equation I: theory. *Quarterly Journal of the Royal Meteorological Society*, 113(478):1311–1328, 1987.
- Testut, C.-E., P. Brasseur, J.-M. Brankart, and J. Verron. Assimilation of sea-surface temperature and altimetric observations during 1992–1993 into an eddy permitting primitive equation model of the North Atlantic Ocean. *Journal of Marine Systems*, 40-41:291–316, 2003.
- Thacker, W. C. and R. B. Long. Fitting dynamics to data. *Journal of Geophysical Research*, 93(C2):1227–1240, 1988.
- Tippett, M. K., J. L. Anderson, C. H. Bishop, T. M. Hamill, and J. S. Whitaker. Ensemble square root filters. *Monthly Weather Review*, 131(7):1485–1490, 2003.
- Toth, Z. and E. Kalnay. Ensemble forecasting at NMC: The generation of perturbations. *Bulletin of the American Meteorological Society*, 74(12):2317–2330, 1993.
- Toth, Z. and E. Kalnay. Ensemble forecasting at NCEP and the breeding method. *Monthly Weather Review*, 125(12):3297–3319, 1997.
- Verlaan, M. and A. W. Heemink. Tidal flow forecasting using reduced rank square roots filters. *Stochastic Hydrology and Hydraulics*, 11(5):349–368, 1997.
- Vorosmarty, C., D. Lettenmaier, C. Leveque, M. Meybeck, C. Pahl-Wostl, J. Alcamo, W. Cosgrove, H. Grassl, H. Hoff, P. Kabat, F. Lansigan, R. Lawford, and R. Naiman. Humans transforming the global water system. *EOS, Transactions, American Geophysical Union*, 85(48):509,513–514, 2004.
- Vos, R. J. and P. G. J. ten Brummelhuis. Goodness-of-fit criteria for the simultaneous assimilation of remote sensing and in-situ data in SPM transport models. Tech. rep., Delft Hydraulics, 1997. Report prepared for CEC-DGXII_MAST 3 and Netherlands Remote Sensing Board (BCRS).

- Vos, R. J., P. G. J. ten Brummelhuis, and H. Gerritsen. Integrated data-modelling approach for suspended sediment transport on a regional scale. *Coastal Engineering*, 41(1-3):177–200, 2000.
- Walker, J. P. and P. R. Houser. Hydrologic data assimilation. In A. Aswatharayan, editor, *Advances in Water Science Methodologies*, p. 230. A.A. Balkema, The Netherlands, 2005.
- Walker, J. P., G. R. Willgoose, and J. D. Kalma. One-dimensional soil moisture profile retrieval by assimilation of near-surface measurements: A simplified soil moisture model and field application. *Journal of Hydrometeorology*, 2(4):356–373, 2001.
- Walker, S. J., J. R. Waring, M. Herzfeld, and P. Sakov. *MECO User Manual*. CSIRO Marine Research, Hobart, Australia, 4.01 ed., 2002.
- Walton, C. C. Nonlinear multichannel algorithms for estimating sea surface temperature with AVHRR satellite data. *Journal of Applied Meteorology*, 27:115–124, 1988.
- Wang, X. and C. H. Bishop. A comparison of breeding and ensemble transform Kalman filter ensemble forecast schemes. *Journal of the Atmospheric Sciences*, 60(9):1140–1158, 2003.
- Wealands, S. R., R. B. Grayson, and J. P. Walker. Quantitative comparison of spatial fields for hydrological model assessment—some promising approaches. *Advances in Water Resources*, 28(1):15–32, 2005.
- Webster, I. T. and S. Narayanan. On the use of EOF analysis in dynamic model testing. *Dynamics of Atmospheres and Oceans*, 14(6):529–542, 1990.
- Webster, I. T., P. J. S., R. B. Grayson, R. P. Molloy, J. Andrewanatha, P. Sakov, K. S. Tan, S. J. Walker, and B. B. Wallace. *Gippsland Lakes Environmental Study - Assessing options for improving water quality and ecological function*. CSIRO for Gippsland Coastal Board, 2001. Pp.84.
- Whitaker, J. S. and T. M. Hamill. Ensemble data assimilation without perturbed observations. *Monthly Weather Review*, 130(7):1913–1924, 2002.

- Wick, G. A., W. J. Emery, L. H. Kantha, and P. Schlüssel. The behavior of the bulk-skin sea surface temperature difference under varying wind speed and heat flux. *Journal of Physical Oceanography*, 26(10):1969–1988, 1996.
- Závody, A. M., C. T. Mutlow, and D. T. Llewellyn-Jones. A radiative transfer model for sea surface temperature retrieval for the along track scanning radiometer. *Journal of Geophysical Research*, 100(C1):937–952, 1995.
- Zillman, J. A study of some aspects of the radiation and heat budgets of the southern hemisphere oceans. Tech. Rep. Meteorological Study No 26, Bureau of Meteorology, Department of the Interior, 1972.
- Zupanski, M., S. J. Fletcher, I. M. Navon, B. Uzunoglu, R. P. Heikes, D. A. Randall, T. D. Ringler, and D. Daescu. Initiation of ensemble data assimilation. *Tellus*, 58A:159–170, 2006.

Appendix A

MECO Parameters Utilised

Table A.1: Model timing parameters.

parameter	value	units	description
DT	360	[s]	Model time step
IRATIO	60	[-]	Mode splitting ratio

Table A.2: Wind drag law parameters.

parameter	value	units	description
V0	10.0	[ms ⁻¹]	lower wind speed threshold
V1	26.0	[ms ⁻¹]	upper wind speed threshold
CD0	0.00114	[-]	lower drag coefficient
CD1	0.00218	[-]	upper drag coefficient

Table A.3: Heat flux and solar radiation parameters

parameter	value	units	description
SWR- ATTENUATION	0.2	[-]	short wave energy attenuation through water column
HEATFLUX- REFH	10	[m]	heat flux reference height
ALBEDO	0.2	[-]	albedo

Table A.4: Bottom friction parameters.

parameter	standard	units	description
QBFC	0.003	[-]	minimum bottom drag coefficient
UF	0.01	[ms ⁻¹]	background friction velocity
Z0	0.0002	[m]	bottom roughness length

Table A.5: Csanady vertical mixing parameters.

parameter	value	units	description
VZ0	0.0001	[ms ⁻¹]	background viscosity
VZALPHA	0.0625	[-]	viscosity coefficient
KZ0	0.00001	[m ² /s ⁻¹]	background diffusivity
KZALPHA	0.03	[-]	diffusivity coefficient

Table A.6: Horizontal mixing parameters.

parameter	value	units	description
U1VH	1.0	[m/s ⁻¹]	horizontal viscosity in u1 direction
U2VHV	1.0	[m/s ⁻¹]	horizontal viscosity in u1 direction
U1KHK	30	[m ² /s ⁻¹]	horizontal diffusivity in u1 direction
U1VHK	30	[m ² /s ⁻¹]	horizontal diffusivity in u1 direction

Table A.7: Computational Settings.

parameter	value	units	description
NONLINEAR	YES	[-]	include nonlinear model equations
CALCDENS	YES	[-]	calculate density
HMIN	0.14	[m]	minimum cell depth
CFL	YES	[-]	apply CFL criteria
TRA SCHEME	VANLEER	[-]	horizontal advection scheme
MOM SCHEME	ORDER1	[-]	horizontal momentum scheme

Appendix B

Choice of ξ and χ values in semi and restricted data types

The semi-restricted data type required that the perturbed data point remain above a lower bound (or below an upper bound). Equation (4.19) is derived by substituting equations (4.16) and (4.18) into equation (4.13). For the lower-bounded case,

$$h_{\min} \leq \hat{h}_i + (\hat{h}_i - h_{\min})\xi z_i + (\hat{h}_i - h_{\min})\chi z \quad (\text{B.1})$$

$$\Rightarrow \frac{h_{\min} - \hat{h}_i - (\hat{h}_i - h_{\min})\chi z}{\hat{h}_i - h_{\min}} \leq \xi z_i \quad (\text{B.2})$$

$$\Rightarrow \frac{-1 - \chi z}{z_i} \geq \xi \quad (\text{B.3})$$

where $z_i < 0$. For a given time series χz is constant and has an expected value of zero. In which case equation (B.3) becomes

$$\xi \preceq \frac{-1}{z_i}, \quad (\text{B.4})$$

the \preceq is used rather than \geq because the relationship generally holds, but occasionally may not hold depending on the value of χz . This value being random is only known when the equation is applied. A similar derivation can be constructed for the upper-bounded case. The same relationship can be derived for χ by assuming that ξz_i is zero.

As z_i is a normally distributed Gaussian random number, probabilities can be assigned to the possibility of z_i being less than a given value (Table B.1). Using the

values in this table, it can be calculated that to reduce the probability of a domain exceeding value being generated to one in a thousand, ξ should be less than 0.32. A

Table B.1: The probability that a normally distributed random number z_i is less than a particular value as a function of ξ for the case of a semi-restricted variable.

ξ	z_i	exceedence probability
1.0	-1	0.1587
0.5	-2	0.0228
0.33	-3	0.0014

similar derivation can be constructed for the restricted data type, which limits the occurrence of perturbed data exceeding the restricted boundary. To simplify the derivation the offset term has been removed, although as with the semi-restricted case it will have an effect. Equation (4.22) is derived by substituting equation (4.20) into equation (4.13).

$$h_{\min} \preceq \hat{h}_i + \frac{\hat{h}_i - h_{\min}}{\frac{h_{\max} + h_{\min}}{2} - h_{\min}} \xi z_i \quad (\text{B.5})$$

$$\Rightarrow h_{\min} - \hat{h}_i \preceq \frac{2(\hat{h}_i - h_{\min})}{h_{\min} - h_{\max}} \xi z_i \quad (\text{B.6})$$

$$\Rightarrow \frac{(h_{\min} - \hat{h}_i)(h_{\min} - h_{\max})}{2(\hat{h}_i - h_{\min})} \succeq \xi z_i \quad (\text{B.7})$$

$$\Rightarrow \frac{(h_{\min} - h_{\max})}{2z_i} \succeq \xi. \quad (\text{B.8})$$

Values of ξ can be associated with probability of exceedence values as indicated in table B.2. The same relationship can be derived for χ by ignoring ξ .

Table B.2: The probability that a normally distributed random number z_i is less than a particular value as a function of ξ for the case of a fully restricted variable.

ξ	z_i	exceedence probability
$\frac{(h_{\min} - h_{\max})}{2}$	-1	0.1587
$\frac{(h_{\min} - h_{\max})}{4}$	-2	0.0228
$\frac{(h_{\min} - h_{\max})}{6}$	-3	0.0014

Appendix C

Derivation of \mathbf{Y} and \mathbf{y}

The derivation begin with equation (5.9),

$$\mathbf{X}^a = \mathbf{X}^f + \mathbf{X}'^f \mathbf{\Lambda}. \quad (\text{C.1})$$

This equation can be split in the ensemble mean and anomaly components

$$\bar{\mathbf{X}}^a + \mathbf{X}'^a = \bar{\mathbf{X}}^f + \mathbf{X}'^f + \mathbf{X}'^f \mathbf{\Lambda}. \quad (\text{C.2})$$

However, the ensemble mean can be formed by multiplying equation (C.1), by an averaging matrix \mathbf{J}_{n_e} , which is a $n_e \times n_e$ matrix with each element of value $\frac{1}{n_e}$. This gives

$$\bar{\mathbf{X}}^a = \bar{\mathbf{X}}^f + \mathbf{X}'^f \mathbf{\Lambda} \mathbf{J}_{n_e}. \quad (\text{C.3})$$

If equation (C.3) is subtracted from equation (C.2), the resulting equation will calculate the ensemble anomalies

$$\mathbf{X}'^a = \mathbf{X}'^f + \mathbf{X}'^f \mathbf{\Lambda} - \mathbf{X}'^f \mathbf{\Lambda} \mathbf{J}_{n_e}. \quad (\text{C.4})$$

This equation can be rearranged as

$$\mathbf{X}'^a = \mathbf{X}'^f [I + \mathbf{\Lambda} - \mathbf{\Lambda} \mathbf{J}_{n_e}], \quad (\text{C.5})$$

and therefore

$$\mathbf{Y} = \mathbf{I} + \mathbf{\Lambda} - \mathbf{\Lambda}\mathbf{J}_{n_e}, \quad (\text{C.6})$$

Since all columns of the ensemble means are the same \mathbf{y} is simply the first column vector of $\mathbf{\Lambda}\mathbf{J}_{n_e}$ from equation (C.3).

Appendix D

Forecast Error Covariance

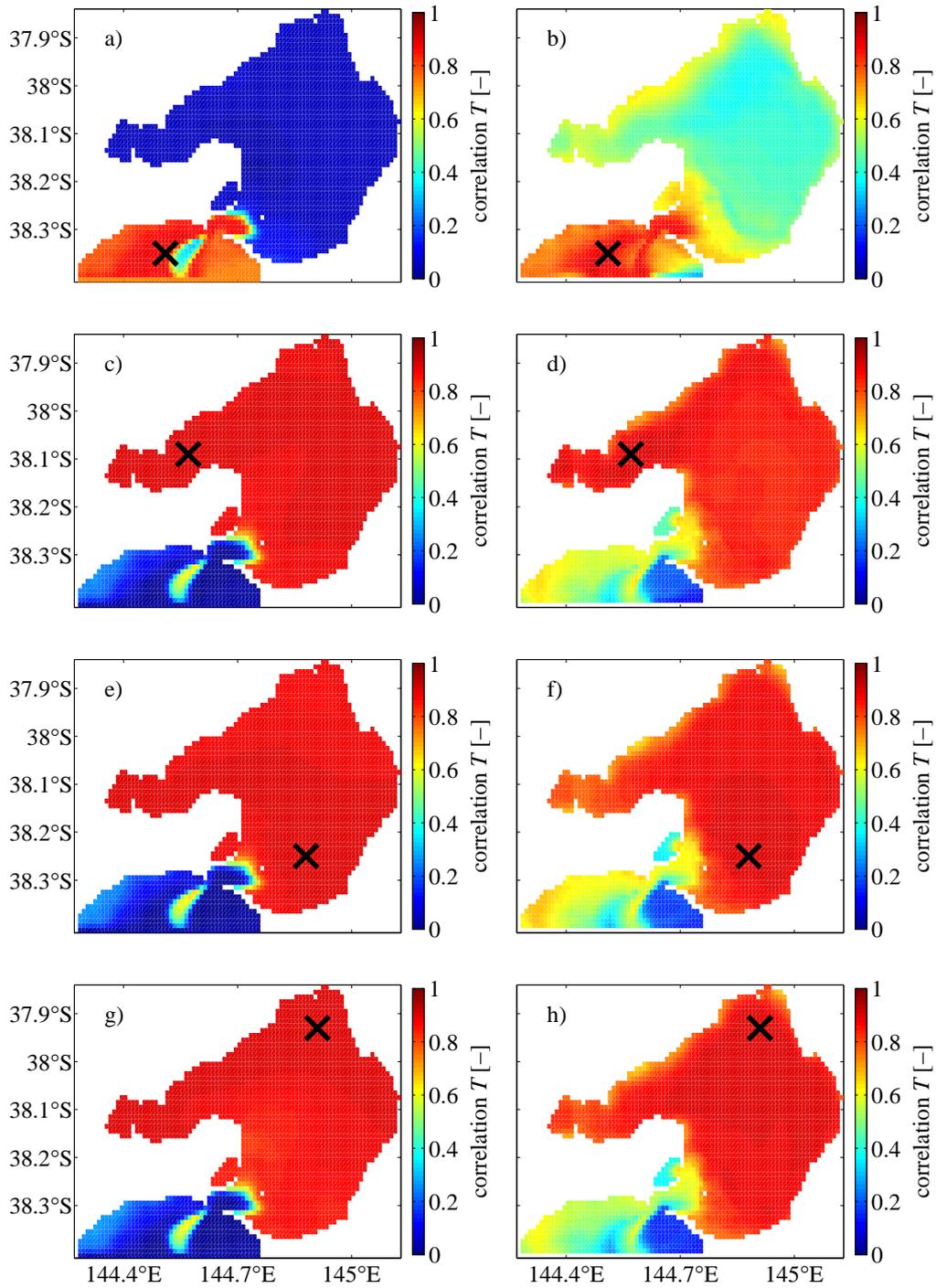


Figure D.1: Comparison of the forecast error correlation of temperature calculated by perturbed forcing (left) and that calculated by a stochastic model (right).

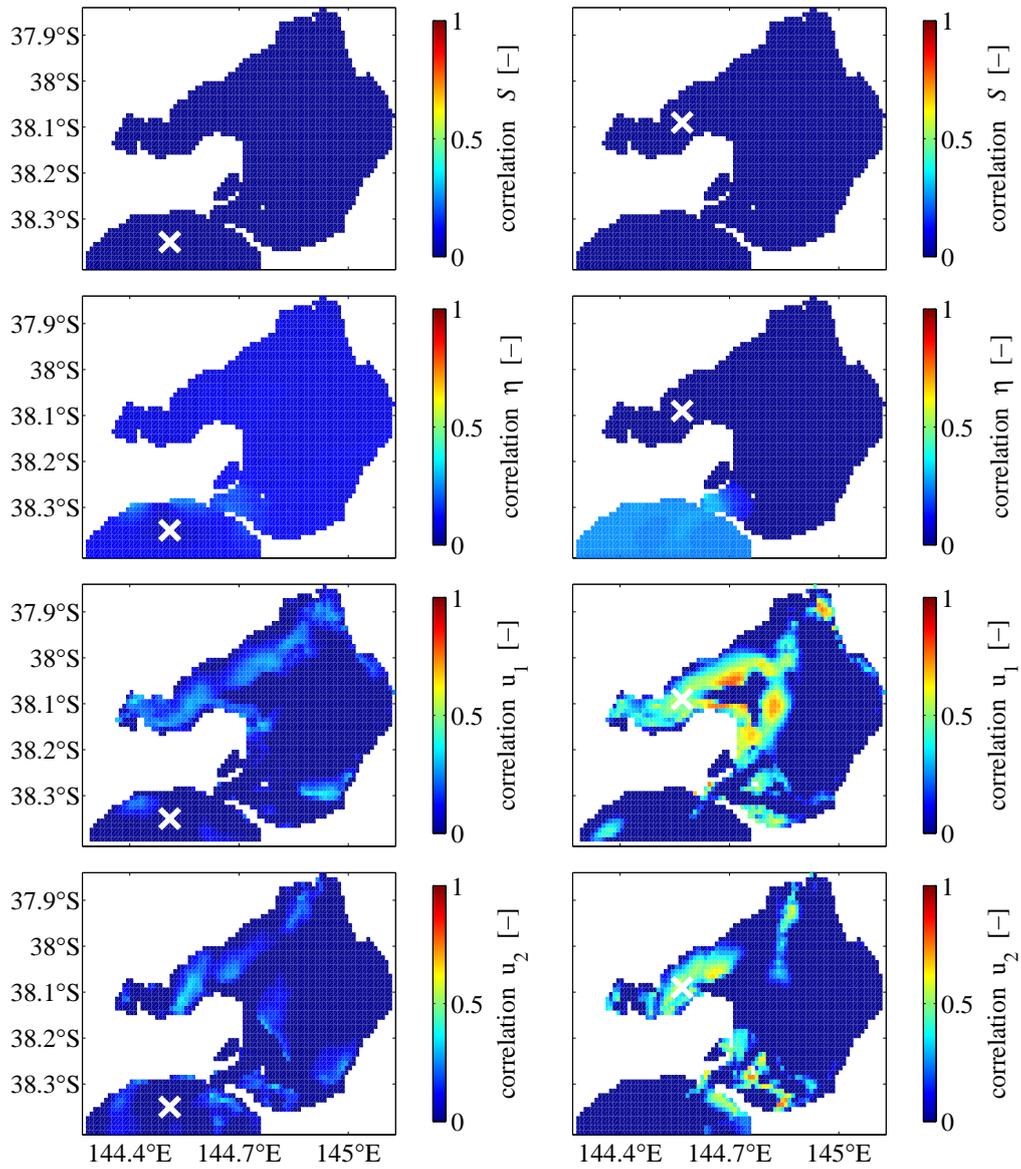


Figure D.2: Comparison of the forecast error covariance of the temperature with other model fields

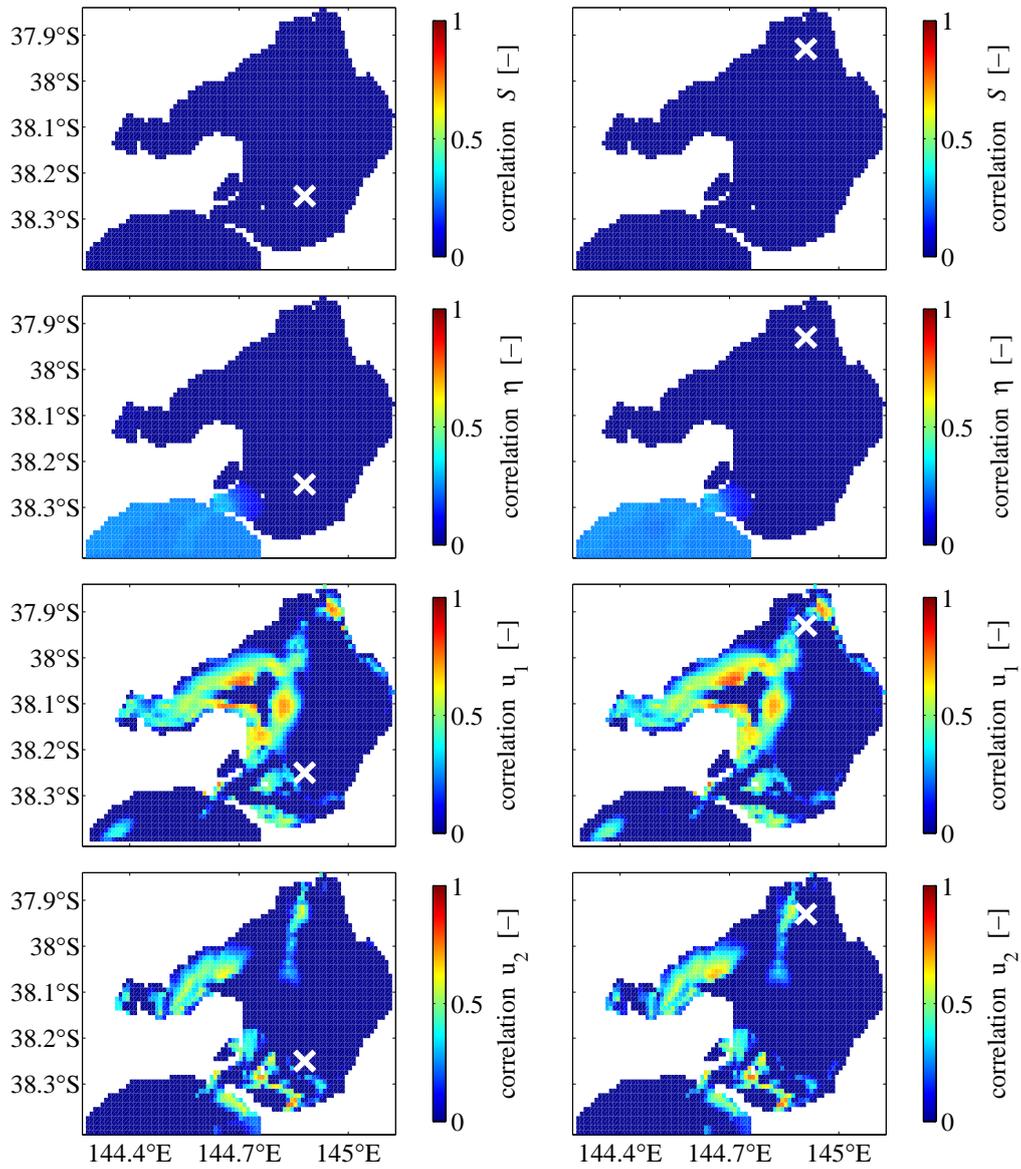


Figure D.3: Comparison of the forecast error covariance of the temperature with other model fields

Appendix E

Assimilation Images

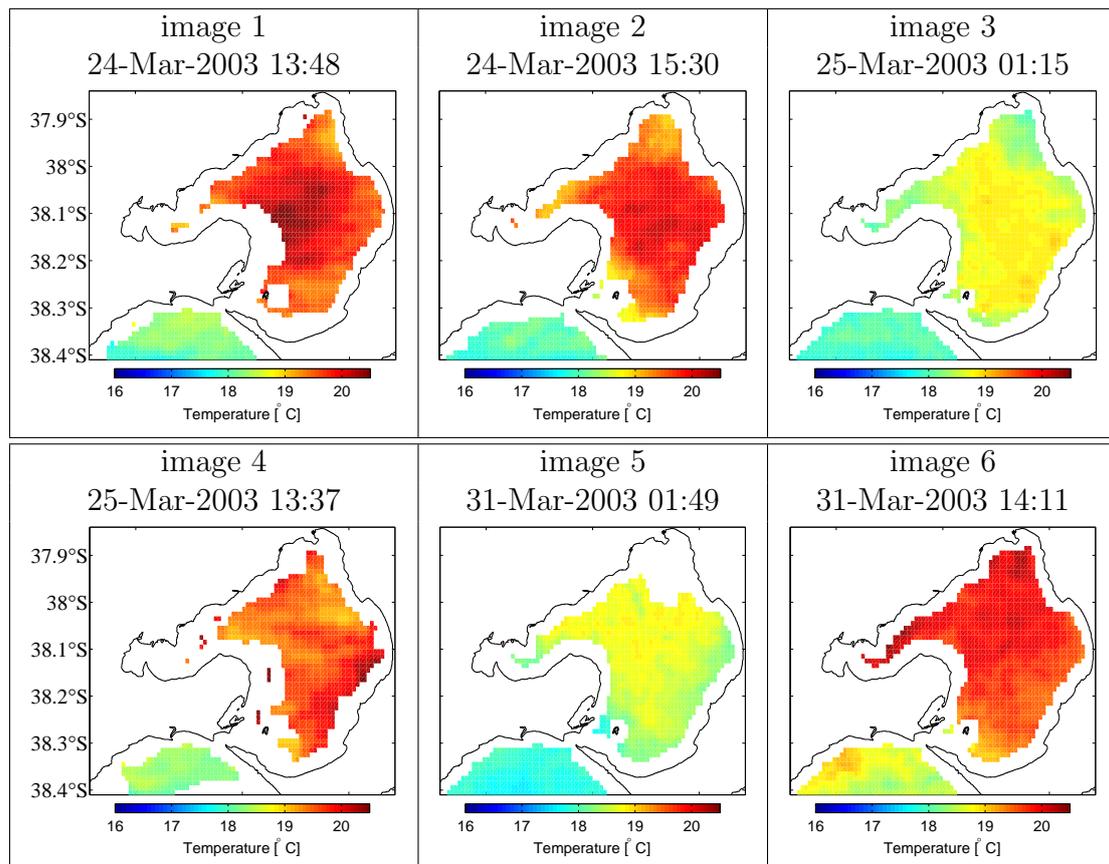


Figure E.1: Uncorrected NOAA-16 AVHRR SST images 1–6 used for assimilation.

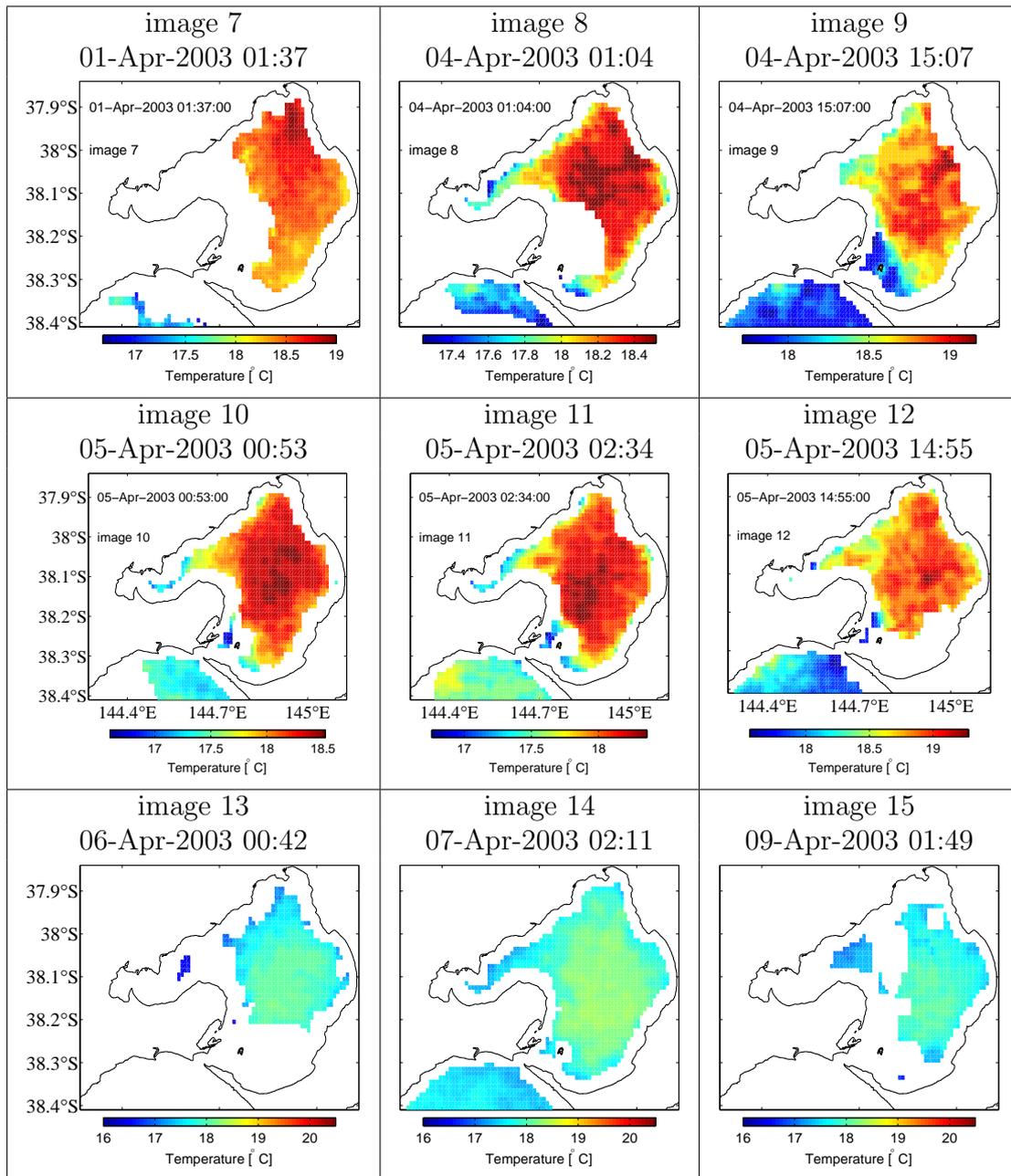


Figure E.2: Uncorrected NOAA-16 AVHRR SST images 7–15 used for assimilation.

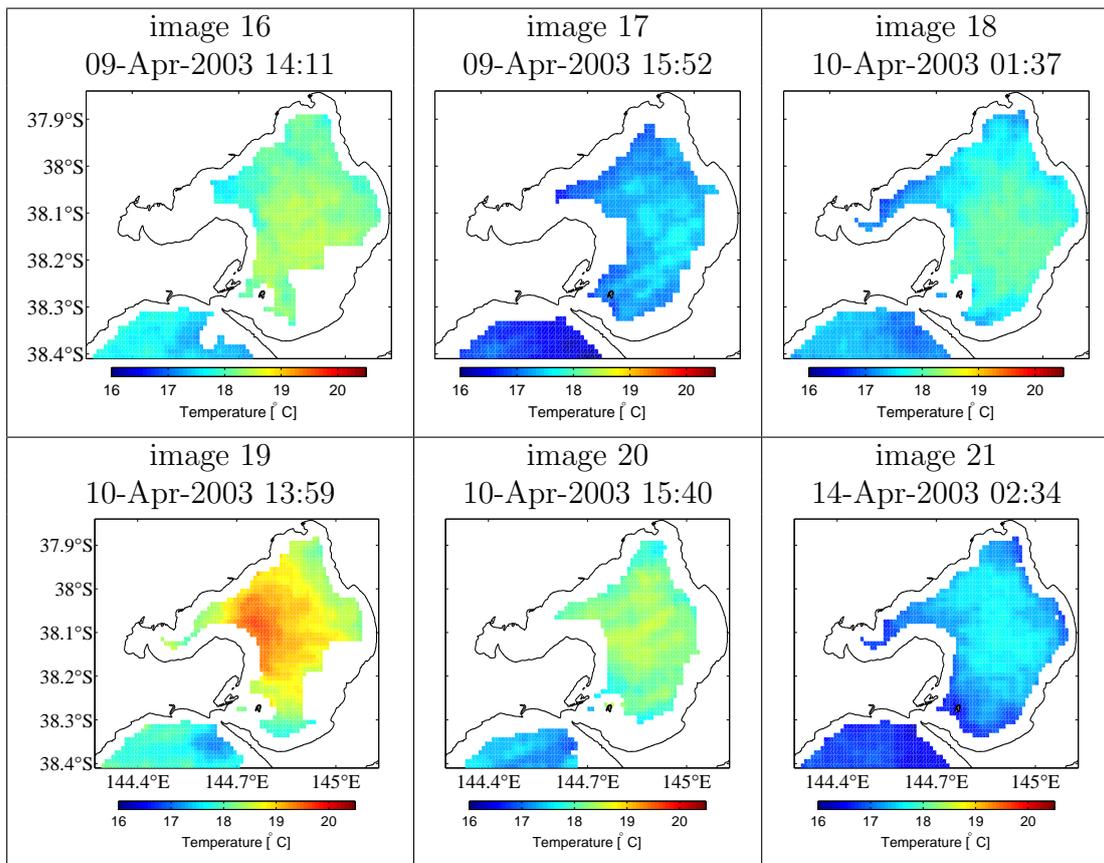


Figure E.3: Uncorrected NOAA-16 AVHRR SST images 16–21 used for assimilation.

Appendix F

Anomaly shape: EnKF vs EnSRF

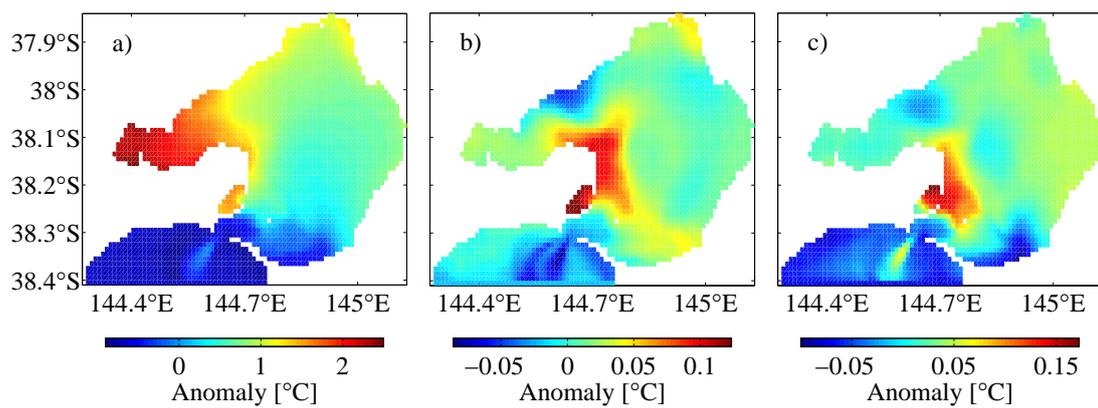


Figure F.1: Changes in vector shape of first ensemble member through the application of the EnKF and the EnSRF; a) forecast, b), EnKF analysis, and c) EnSRF analysis.

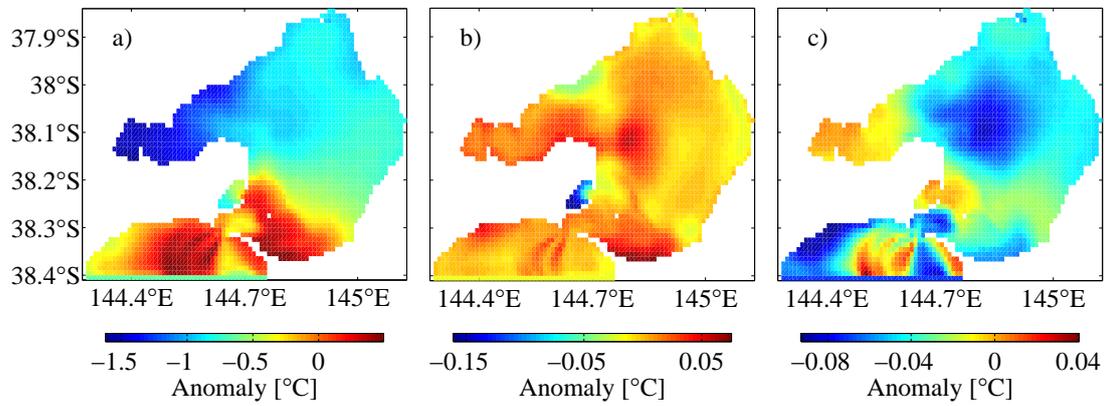


Figure F.2: Changes in vector shape of second ensemble member through the application of the EnKF and the EnSRF; a) forecast, b), EnKF analysis, and c) EnSRF analysis.

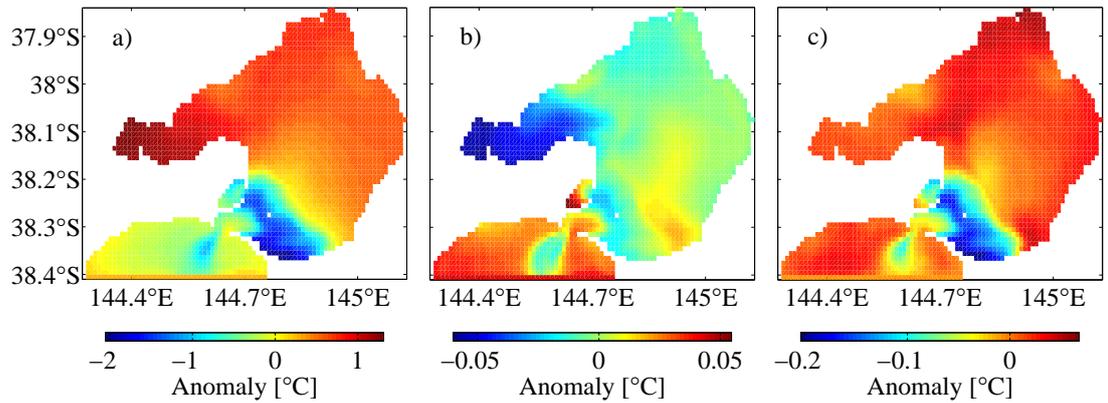


Figure F.3: Changes in vector shape of third ensemble member through the application of the EnKF and the EnSRF; a) forecast, b), EnKF analysis, and c) EnSRF analysis.

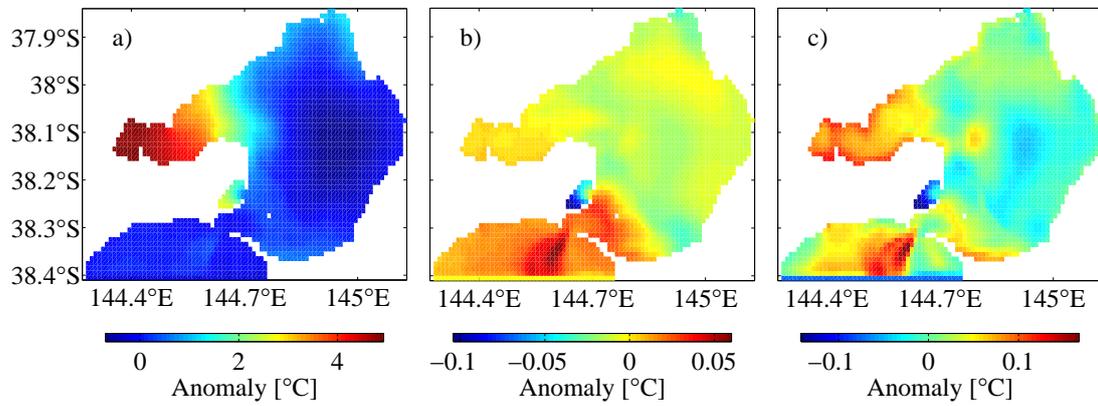


Figure F.4: Changes in vector shape of fourth ensemble member through the application of the EnKF and the EnSRF; a) forecast, b), EnKF analysis, and c) EnSRF analysis.

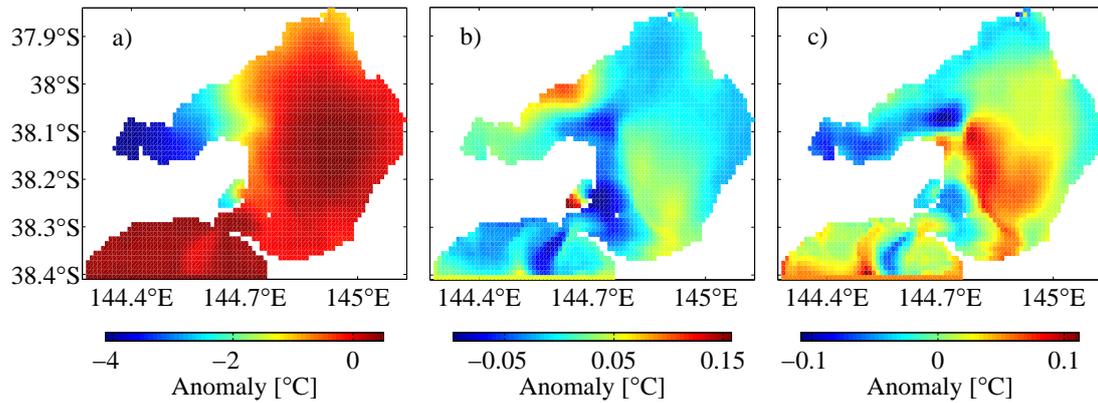


Figure F.5: Changes in vector shape of fifth ensemble member through the application of the EnKF and the EnSRF; a) forecast, b), EnKF analysis, and c) EnSRF analysis.

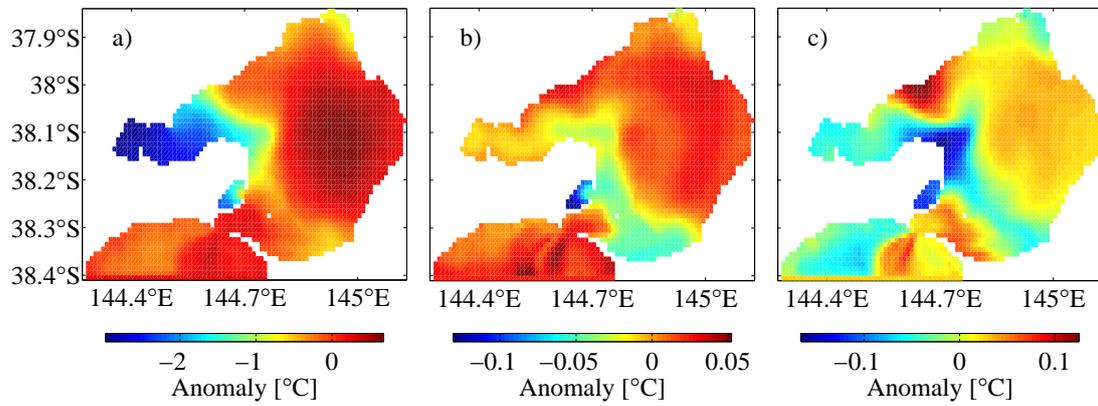


Figure F.6: Changes in vector shape of sixth ensemble member through the application of the EnKF and the EnSRF; a) forecast, b), EnKF analysis, and c) EnSRF analysis.

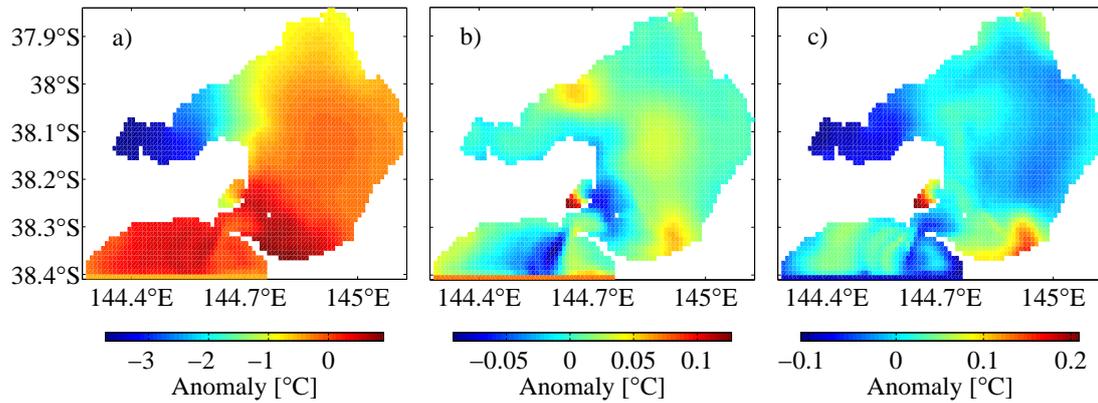


Figure F.7: Changes in vector shape of seventh ensemble member through the application of the EnKF and the EnSRF; a) forecast, b), EnKF analysis, and c) EnSRF analysis.

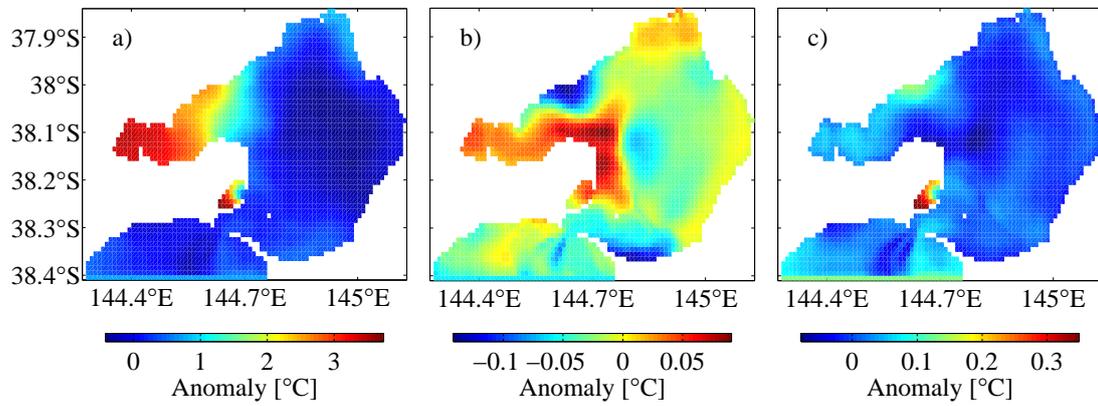


Figure F.8: Changes in vector shape of eighth ensemble member through the application of the EnKF and the EnSRF; a) forecast, b), EnKF analysis, and c) EnSRF analysis.

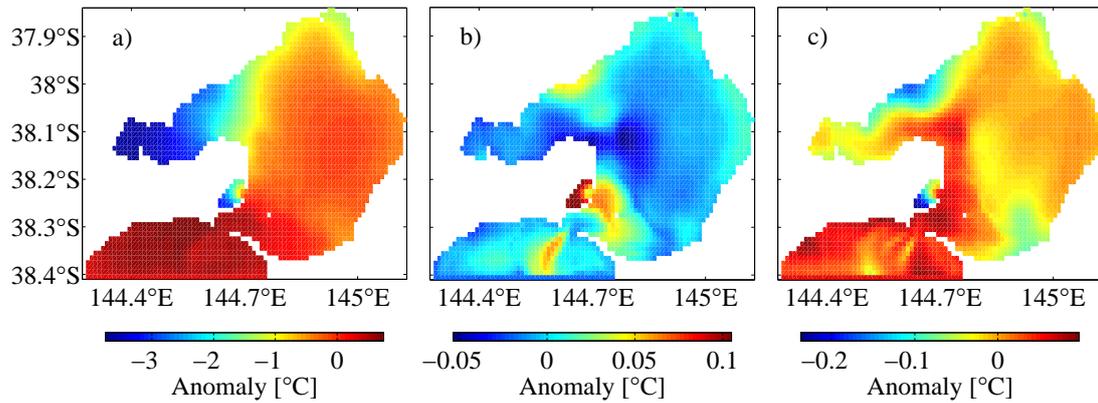


Figure F.9: Changes in vector shape of ninth ensemble member through the application of the EnKF and the EnSRF; a) forecast, b), EnKF analysis, and c) EnSRF analysis.

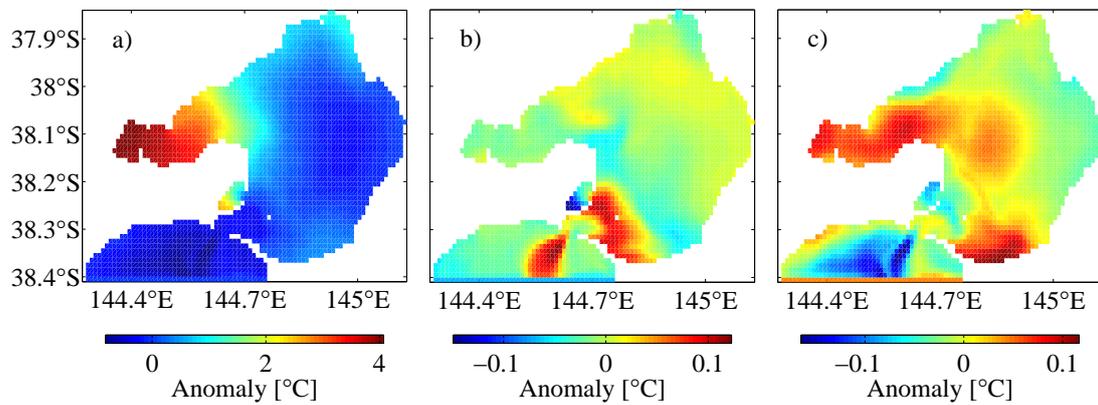


Figure F.10: Changes in vector shape of tenth ensemble member through the application of the EnKF and the EnSRF; a) forecast, b), EnKF analysis, and c) EnSRF analysis.