Documentation for long term forcing data for wet seasons in Darwin, Australia

Laura Davies, email: laura.davies@sci.monash.edu.au

March 3, 2010

1 Introduction

This documentation is written to explain the forcing data available for three wet seasons over the ARM site in Darwin, Australia. The data set has been created with ARM grant XXXX under PI Prof Chrisitan Jakob. This documentation will explain the input data used to create this forcing data set and the objective analysis methodologies used to create it. This data set is the culmination project to exploit the potential of variational analysis to provide a long timseries best estimate forcing data set suitable use with Single Column Models and Cloud-Resolving Models. Furthermore, in conjunction with another ARM funded project to investigate the formation of an ensemble forcing data set, a long term ensemble forcing data set is provided with the best estimate forcing data set. Details will be given on the methodology of how this ensemble was formed. Finally, the file structure of the available data will be outlined.

2 Creation of the dataset

In this section the data input used to create the forcing data will be explained and outline of the variational analysis process will be given. There are several data inputs which are sourced from different locations. The domain of the data set is the region around Darwin which was used for TWP-ICE. Hence the TWP-ICE period can be considered a sub-period of this data set. Essentially the process here is similar to that for deriving the forcing data set of TWP-ICE which is also available from ARM. It will be highlighted here where the the method used here differs from the standard process.

2.1 Data input

The variational analysis requires **surface meteorology data, surface sensible and latent heat fluxes, surface radiation and top of the atmosphere radiation**. The data constrains the energy balance at the top and bottom of the atmosphere. This data is extracted from ECMWF analysis data of the four grid points encompassing the domain and is averaged over the domain.

Vertical profiles of dynamic and thermodynamic variables at the edge of the domain are also required for the variational analysis. The profiles are used to constrain the advection of temperature and moisture into and out of the domain. At Darwin there are five locations for radiosonde measurements around the ARM observation facility and one in the centre. Vertical profiles are normally derived from radiosonde data, for example as gathered during a field campaign such as TWP-ICE. However, long term radiosonde data sets are not available at sufficient temporal or spatial resolution to create a long term forcing data set. Investigations have shown the pusedo radiosonde data, which are derived from ECMWF vertical profiles, can be used in place of real radiosonde data with little impact on the resulting forcing data. ECMWF data on the four grid points encompassing the domain are interpolated to the radiosonde locations from TWP-ICE. Hence they simply replace radiosonde data in the variational analysis. Whilst Barnes interpolation is used, it has been found that the the pusedo radiosonde data is not sensitive to the method of interpolation used, Bi-linear interpolation has also been investigated.

The variational analysis uses **timeseries of liquid water path and total column water** to constrain the moisture equation. These are provided by a line-of-sight microwave radiometer product from ARM. The data is quality controlled to avoid contamination of the instrument by precipitation. The quality control involves using surface meteorology measurements of precipitation to determine if the instrument would be too wet to give useful results. The more intense the precipitation the longer required for the instrument to dry and hence a longer period of microwave data is screened out. Where too much data is missing the gap in filled by a mean value. Using mean values provides acceptable results as the column water terms have small weight in the variational analysis.

One of the key inputs to the variational analysis is a timeseries of **domain averaged rainfall**. It is the rainfall input that forms the basis for the ensemble forcing which will be discussed further in Section 3. The rainfall timeseries is derived from 10 min radar data over the domain. The radar data is temporally and spatially averaged to give rainfall values at 6 hr intervals, the frequency of ECMWF data.

2.2 Methodology

The above data is used to derive the best estimate forcing data set using variational analysis (Zhang and Lin, 1997). This process provides the best estimate of the time evolution of the atmosphere through constraining all the input data. The variational analysis process minimises the errors between between all the input data the observations and is constrained by mass, moisture, heat and momentum. Variational analysis provides a best estimate of the state of the atmosphere given the input data, however, if this data has errors, there will be errors in the derived best estimate. It has been shown (Zhang *et al.*, 2001) that input surface precipitation has a large effect on the derived forcing data set. For example, the analysed vertical velocity is very sensitive to surface precipitation.

3 Formation of ensemble data

The derivation of surface precipitation from radar data is highly complex and liable to large errors. It may be that these errors will have a large effect on the derived forcing data set. A possible method of addressing these errors is to use an ensemble technique. The methodology used here is identical to that used to create and ensemble forcing data set for the TWP-ICE period.

In a previous ARM funded project Christian Jakob and collaborators (publication to follow) have estimated the error in the radar-derived rainfall estimates from a comparison with rain gauge data. They used the distribution of these errors to calculate N rainfall scenarios (currently N=100) for the TWP-ICE period. Each of these scenarios was a percentile of a distribution that encompasses the full range of the errors in deriving the radar-derived rainfall and each scenario is possible given the uncertainties in the radar-derived rainfall. These new rainfall scenarios were used to create an N member ensemble for the TWP-ICE period.

Similarly, these TWP-ICE based rainfall error estimates have been used to construct 100 rainfall scenarios for the three complete wet seasons. This assumes that the error distribution in radar-derived rainfall will be the same in all wet seasons as it is in TWP-ICE. Whilst this cannot explicitly be proved it is likely and the ensemble itself could be argued to account for errors arising from this assumption. Variational analysis is then performed using each of the 100 rainfall scenarios with all other input data unchanged. Attempts to modify other variables in step with the changing rainfall did not have a large impact. For example, if an ensemble member has higher rainfall, it is likely top of the atmosphere outgoing longwave radiation (OLR) will also increase. Modifying OLR in parallel with rainfall did not have a significant effect on the derived forcing data.

4 Available data

The data that is provided through the ARM archive is available for three complete wet seasons - 0405, 0506 and 0607, however, as radar data is not sufficiently available for the duration of all three seasons the three seasons have been subdivided. There are six sub-periods in 0405, two sub-periods in 0506 and four sub-periods in 0607. The zipped and tarred data bundle ForcingData030310.tar.gz will be unpacked to ForcingData030310 which contains sub-folders for each sub-period of the wet seasons. Table 1 shows the sub-periods for which data is available. The sub-folders are labelled for each sub-period in Table 1.

Within each sub-period there is directories for the best estimate forcing data set (called best_est) and the 100 ensemble member forcing data sets (numbered p00-p99). The forcing data within these directories is called forcing.nc

All data is in netcdf format and have the same format as all ARM best estimate forcing data sets with the exception that ARSCL cloud data are not part of the files. All time variables are days since 1 October of the first year in the season (regardless of what any variable long name descriptor state).

References

Zhang, M. H. and Lin, J. L. (1997). Constrained variational analysis of sounding data based on column integrated budgets of mass, heat, moisture, and momentum: Approach and application to ARM measurements. *J. Atmos. Sci.*, **54**, 1503–1524.

Zhang, M. H., Lin, J. L., Cederwall, R. T., Yio, J. J., and Xie, S. C. (2001). Objective Analysis of ARM IOP Data: Method and Sensitivity. *Mon. Wea. Rev.*, **129**, 295–311.

	Start date and time			End date and time		
	Month	Day	Time	Month	Day	Time
Season 0405						
0405_p1	11	3	12Z	11	8	00Z
0405_p2	11	11	12Z	12	26	06Z
0405_p3	01	06	12Z	01	10	00Z
0405_p4	01	17	12Z	01	23	00Z
0405_p5	01	27	12Z	02	25	06Z
0405_p6	02	28	12Z	04	06	18Z
Season 0506						
0506_p1	11	10	06Z	04	15	18Z
0506_p2	04	18	12Z	04	30	06Z
Season 0607						
0607_p1	10	11	12Z	11	20	18Z
0607_p2	11	24	12Z	01	06	18Z
0607_p3	01	09	06Z	01	17	12Z
0607_p4	01	21	12Z	04	18	18Z

Table 1: Summary of data periods for which forcing data is available. XXXX_pY refers to the Y^{th} period of season XXXX. The seasons had to be subdivided as radar rainfall data was not continually available. All times are UTC.