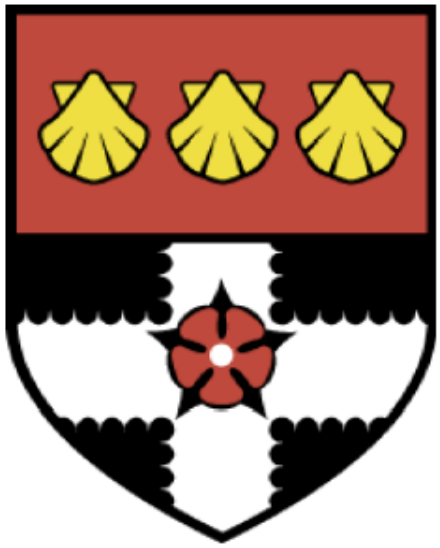
University of Reading
Department of Meteorology



The assimilation of misplaced boundary layer features

by

Alison M. Fowler

A thesis submitted for the degree of Doctor of Philosophy

February 2010

Abstract

An essential part of the data assimilation problem is a good knowledge of forecast error statistics. For the boundary layer (BL) it can be difficult to estimate these error statistics as they are often dependent upon the atmospheric state, in particular, they are shown to depend upon the presence of a capping inversion, and the strength and position of this inversion.

In traditional variational assimilation only small scale amplitude errors are accounted for, but it is shown that the inversion (when present) may also be associated with a significant positional error in the forecast which has non-Gaussian statistics. Without a description of the positional error statistics, the assimilation of observations can weaken the inversion strength; this is shown to give a poor analysis of the thermal and moisture structure of the BL and in turn a poor diagnosis of cloud.

In this thesis a new assimilation scheme is proposed which explicitly includes the non-Gaussian positional error of the inversion. This scheme, referred to as the 'floating BL scheme', is applied to the vertical variational assimilation of temperature and humidity. Within this context the floating BL scheme is tested with a series of idealised experiments and real data from radiosondes. It is shown that on average the floating BL scheme gives an analysis with an improved inversion height and is more capable of retaining the inversion structure. It is also shown that the floating BL scheme allows for a more realistic diagnosis of BL cloud, ensuring that the cloud remains capped by the inversion. It is speculated how this scheme may be generalised to 3D and 4DVar and its potential for use with other observations.

Declaration

I confirm that this is my own work and the use of all material from other sources has been properly and fully acknowledged.

Some of the work from this thesis has been published in the following journal article:

Fowler, A. M., R. N. Bannister, J. Eyre, 2010: Characterising the background errors for the boundary layer capping inversion. *Austral. J. Ocean. Meteorol.*, **59**, 17-24

Alison Fowler, Reading, 2010.

Acknowledgements

I'd like to acknowledge the help and guidance of my supervisors Ross Bannister and John Eyre. I'd also like to thank Stefano Migliorini and Alan O'Neill for their additional support and NERC and the Met Office for their sponsorship.
The University of Reading has been an ideal place to study especially due to their monitoring committee policy which has given the regular reassurance and advice needed. My monitoring committee members were Sarah Dance and Bob Plant. I'd also like to thank my office mates over the three years I've been studying. Thanks also goes to Chris, my partner, and Coco, my cat, for their support, patience and source of distraction. A final thanks goes to my parents and siblings for encouraging me to come this far.

Contents

List of Figures

List of Tables

Chapter 1

Introduction

The boundary layer (BL) forms the lowest kilometre or so of the Earth's atmosphere. When warm air overlies the BL, such as in a region of high pressure, the BL is often capped by a feature that has interesting thermal characteristics. A sharp positive gradient in the vertical temperature profile can form, this is known as a temperature inversion and acts as a lid on the cooler BL air below preventing it from mixing with the warmer free atmospheric (FA) air above. The structure and position of the inversion are therefore important for understanding such things as the dispersion of pollutants originating from within the BL. Furthermore if cloud is present within the BL then the height of the inversion acts as an upper bound on the cloud top height, which in turn impacts on surface properties such as surface temperature.
Accurate characterisation of the BL capping inversion in the initial conditions for numerical weather prediction (NWP) models is important for accurate forecasts of atmospheric features affected by the inversion. However, the inversion is a difficult feature in itself to model, as it is dependent upon both large scale forcings, such as the positioning of fronts and regions of subsidence, and small scale forcings such as turbulence. In order to give the best possible initial conditions for the forecast, observations of the BL should be combined with a previous forecast valid for the current time. This blending of data is referred to as *data assimilation* (DA) (see Kalnay, 2003; Daley, 1997 for an introduction) and is intended to keep the forecast close to reality.

The best way to perform the assimilation is an area of active research and there are a great number of challenges that need to be addressed (see e.g. Rabier, 2005; Errico et al., 2007; Dance, 2004 for a discussion of contemporary issues). One such challenge is how best to represent the errors associated with the forecast (e.g. Bannister, 2008a, 2008b). In this thesis particular thought will be given to the representation of the errors in the forecast for the BL when an inversion is present. Some important issues arise from the variability associated with the presence, position and structure of the inversion, and the presence of cloud.

This chapter aims to discuss the motivation of the assimilation problem for the BL. Firstly the importance of the BL structure will be discussed and the consequences for the forecast if the BL structure is not correctly assimilated. Secondly the data available that typically provide information about the BL structure will be presented.

1.1 The boundary layer (BL)

Transport processes at the Earth's surface can modify the lowest 100 to 3000m of the atmosphere. This region, which is strongly coupled to the surface, is known as the *boundary layer* (BL). The rest of the troposphere which lies above the BL is often referred to as the *free atmosphere* (FA). The Earth's surface forces the BL through frictional drag, evaporation and transpiration, heat transfer, pollutant emission, and terrain induced flow modification. It is the variability of these forcings both temporally and spatially that lead to large differences in the depth of the BL (see Stull, 1988, for an introduction to the BL). Over land for instance, the BL structure and depth can be seen to have a strong diurnal cycle. This cycle is due to the ground warming and cooling in response to solar radiation which in turn induces turbulence within the BL. Over the oceans a prominent diurnal cycle is not seen due to the sea surface temperature (SST) varying much more slowly than the land surface temperature. Changes to BL depth over the oceans are mainly due to synoptic and mesoscale vertical motion and advection of different air masses.
Figure 1.1 shows typical vertical profiles during an anticyclonic regime, for key atmospheric variables, for the BL over land, for day (top) and night (bottom). During the day mixing caused by surface heating causes potential temperature and humidity to have a near uniform value throughout the BL above the surface. At the surface during the day the profile is unstable as the ground is heated and surface water evaporates, giving a negative gradient in both temperature and humidity. During the night turbulence is not induced by surface heating and the air at the surface is cooled in response to radiative cooling at the ground, giving a more stable profile. It is also seen in Figure 1.1 that the BL and FA have very different properties in anticyclonic conditions. The BL air is much more humid and has a lower potential temperature than the FA, which is stably stratified.

1.1.1 The boundary layer capping inversion

An inversion refers to a sharp increase in temperature with increasing height. Inversions form at the top of the BL when warmer less dense air moves over colder air within the BL such as at a warm front. During anti-cyclonic events a well defined temperature inversion may also form at the top of the BL as the descending FA air warms by adiabatic compression. The temperature difference across such an inversion is typically between 2K and 15K. On average the inversion depth is 40% the depth of the BL (Gyr and Rys, 1995). When an inversion is present a rising parcel of air from the surface loses its positive buoyancy as it enters the inversion due to the ambient air being warmer than the air parcel (Stull, 1988). This prevents the BL air from escaping the BL and mixing with the FA. The inversion therefore acts to separate the BL from the FA, resulting in two distinct layers of air; the BL which is turbulent, moist and has a large concentration of aerosols originating from the surface, and the FA which is drier, warmer and more stable. If the BL air is very turbulent or the inversion is very weak, some of the rising thermals may overshoot the inversion before sinking back down into the BL. Through this process it is possible for some of the FA air to become mixed down into the BL air. This is known as *entrainment* and leads to the growth of the BL and the reduction of the inversion strength.

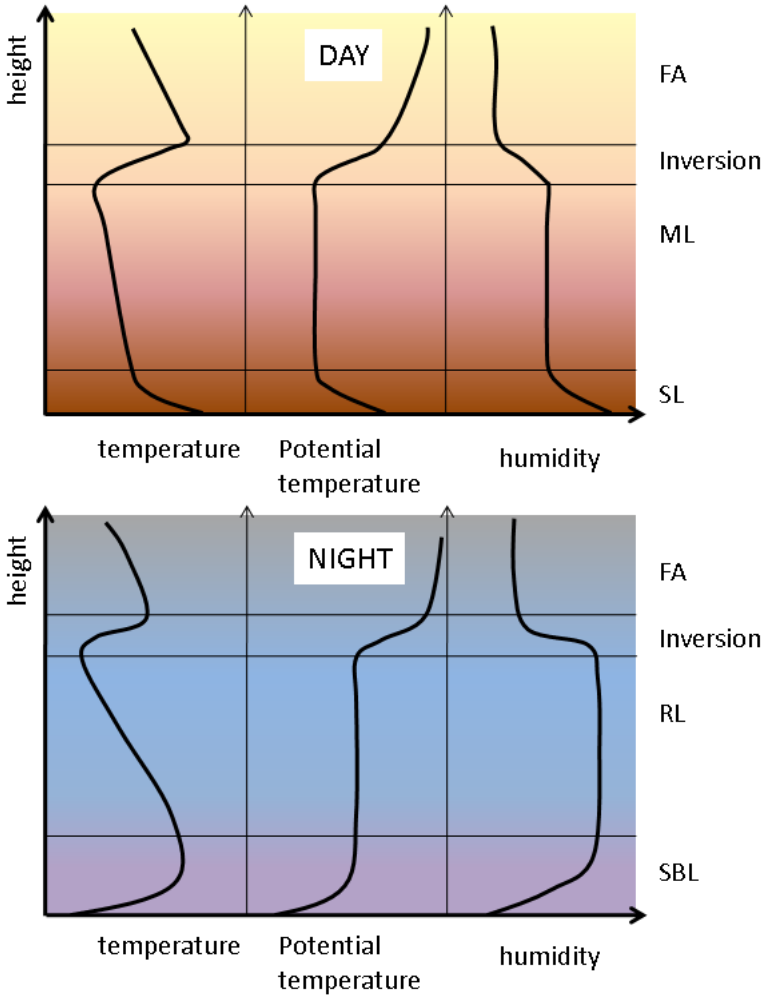


Figure 1.1: Typical profiles of temperature, potential temperature and specific humidity in the BL for anti-cyclonic conditions over land. During the day (top) a surface layer (SL) forms which has a higher temperature and humidity compared to the profile higher up. Above this is the mixed layer (ML) where turbulence causes the potential temperature and humidity to become approximately uniform. This is capped by an inversion. This inversion marks the transition between the BL and FA air which is drier and more stable. During the night (bottom), the BL is not forced by surface heating and a stable BL (SBL) forms underlying a residual layer (RL). *Image modified from Wallace and Hobbs, 1977.*

Figure 1.2 shows a schematic of the BL evolution throughout the day over land, over a region of high pressure. During the early afternoon the BL is often at its deepest due to the activity of rising thermals causing the FA to become *entrained* down into the BL. During the day, the region of the BL which is not in contact with the surface is known as the *mixed layer* (ML). As the name suggests the ML is often well mixed in atmospheric properties (potential temperature, specific humidity etc.) due to turbulence (seen in Figure 1.1). The *subsidence* of the overlying air causes a strong capping inversion to form which inhibits the entrainment and the growth of the ML. After sunset the ML begins to collapse and a shallow stable BL (SBL) forms, also referred to as the nocturnal BL, topped by a residual layer (RL) containing air from the previous day's ML. After sunrise a ML forms again and

begins to deepen through entrainment (during this time the inversion layer may also be known as the *entrainment zone*).

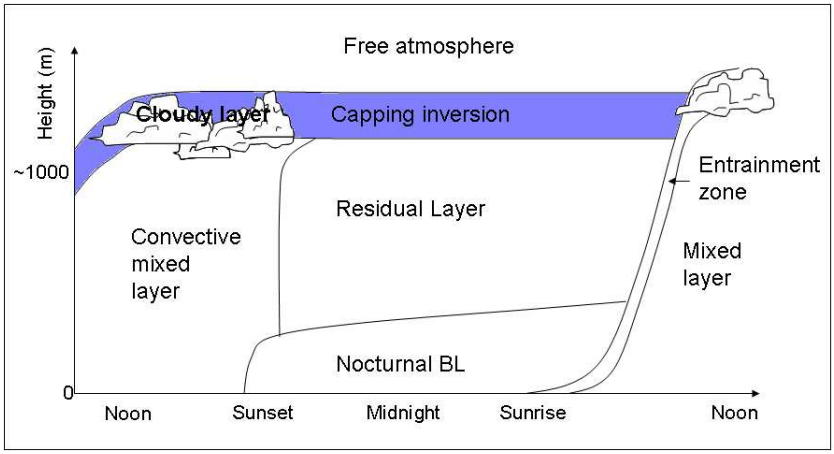


Figure 1.2: Schematic of the diurnal cycle of the BL over land during a region of high pressure. *Figure reproduced from Stull, 1988.*

This kind of behaviour is frequently observed, for example by lidar, which observes scattered laser light from aerosol and cloud particles. Figure 1.3 shows a Doppler Lidar time series from the Chilbolton Observatory in Southern England. The top panel shows the attenuated backscatter and the bottom panel gives the vertical Doppler velocity. In the top panel cloud is distinguished from the rest of the BL due to the liquid cloud particles being much larger than the aerosols and so the signal is much stronger (red[1]). In the bottom panel it is easy to pick out the convective mixed layer which is present during the day (approximately 8-22UTC) due to the strong up (red) and down (blue) drafts observed in the vertical Doppler velocity time series. This is seen to be replaced by a stable layer during the night when vertical velocities are close to zero (green).

The BL capping inversion is dependent on large scale weather patterns and forcing: such as subsidence and warm fronts; the diurnal cycle; and small scale features: such as turbulence and BL clouds. This means that it can be very variable spatially and temporally both in its height, structure and presence. This is seen in the middle panel of Figure 1.4 where the pressure level of the inversion (the lowest level where the temperature gradient changes from negative to positive) is plotted for the NAE (North Atlantic and Europe) domain for 25th January 2008. No inversion is represented by the absence of colour. We can compare the pattern in the presence of the inversion (middle panel) to the analysis of surface pressure given for 12hrs either side (top and bottom panels). The regions where an inversion is present roughly coincide with areas of high pressure and warm fronts as expected. The height is inhomogeneous in some of these regions, whilst in others there is little variation.

[1]Note the logarithmic scale.

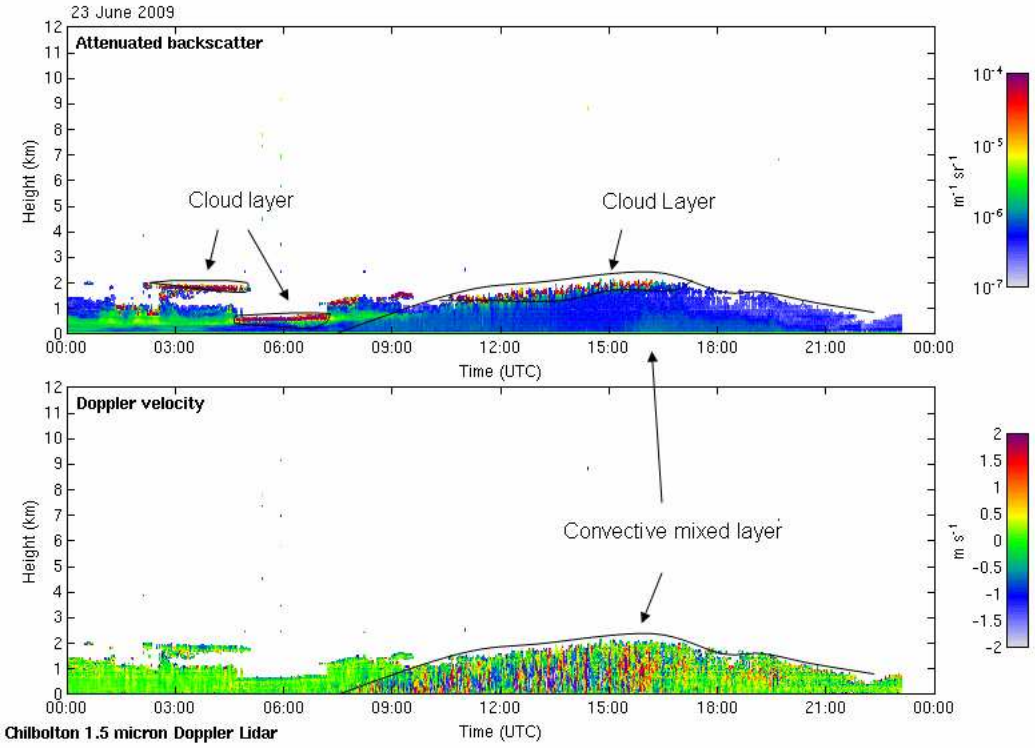


Figure 1.3: Observations from a ground based lidar at Chilbolton, Hampshire on 23rd June 2009 during a region of high pressure. Top panel shows time series of attenuated backscatter. The bottom panel shows vertical Doppler velocity (updrafts are red and downdrafts are blue). *Image from http://www.met.reading.ac.uk/radar/realtime/archive/doppler-lidar/.*

1.1.2 The importance of the boundary layer capping inversion

The presence of the BL capping inversion is important for understanding the re-distribution of atmospheric constituents originating from within the BL. These include water vapour and pollution which can become trapped within the BL if an inversion is present. The presence, location and structure of an inversion is therefore important for predicting the formation of clouds. Clouds form in regions where the air is saturated with water vapour: this region is bounded between the *lifting condensation level* (LCL) and the inversion, where the latter often becomes synonymous with the cloud top height (CTH). The LCL is the level at which a parcel of moist air lifted dry-adiabatically would become saturated (see e.g. Barry and Chorley, 1968). The height of the LCL is dependent upon the relative humidity profile of the atmosphere. Over the oceans higher humidity leads to a lower LCL and so there is often a higher fraction of BL cloud than over land.

As well as the height of the inversion restricting the CTH, the strength of the inversion (the gradient of temperature with respect to height) may also regulate the depth and persistence of the cloud layer. A weak inversion (a small rate of increase in temperature with height) allows for more entrainment of dry FA air. Entrainment of drier air can cause the LCL to rise and so the cloud layer thins.

Two types of cloud are often present within the BL over mid-latitudes. The first is *fair weather cumulus* cloud which form due to rising thermals. The second is *stratocumulus* cloud (Sc) which can form a thick blanket of cloud. Sc is the most common type of cloud

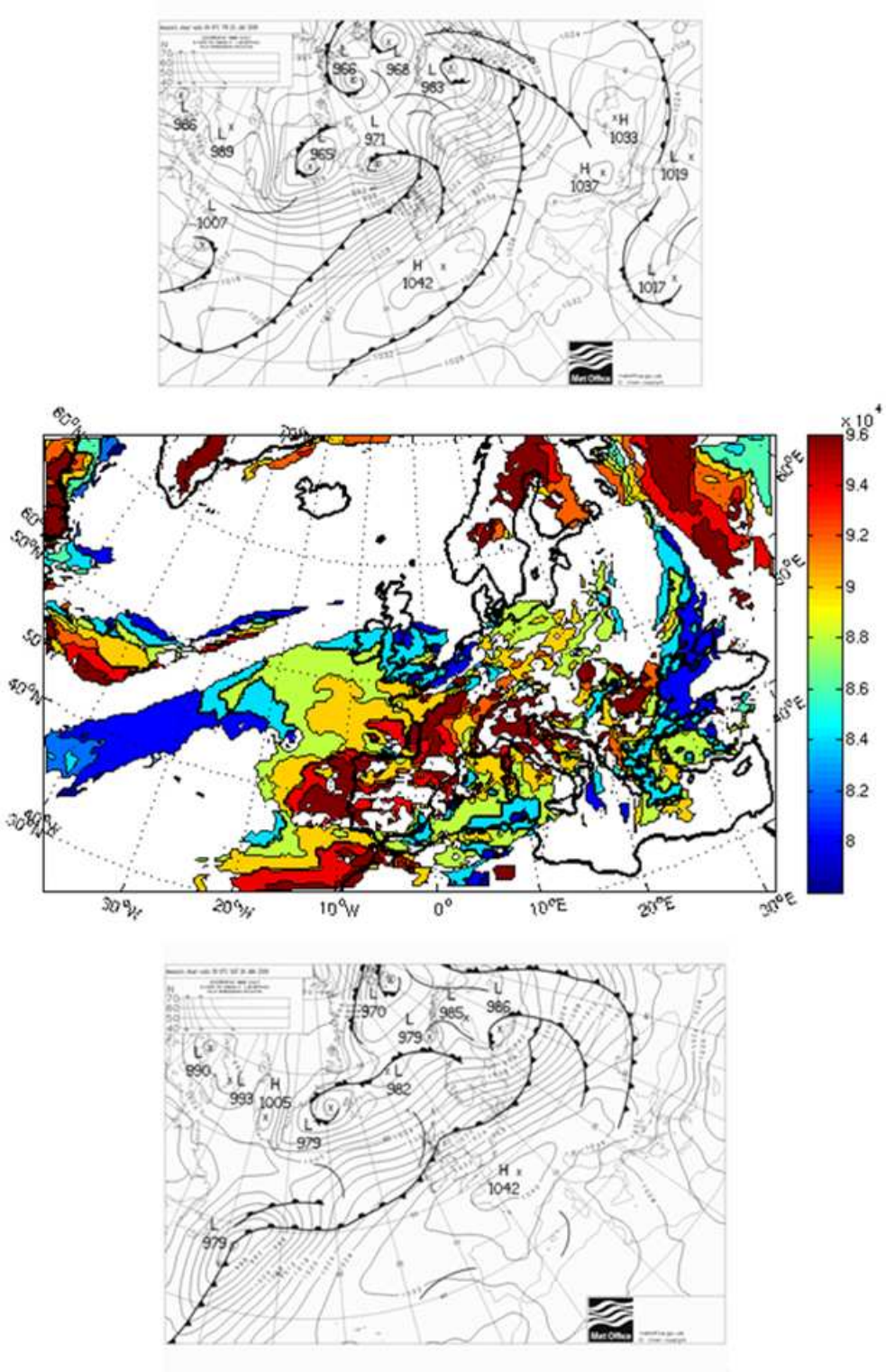


Figure 1.4: Top panel: The surface pressure analysis for the 25th January 2008, 0Z (*Crown Copyright*). Middle panel: The pressure (Pa) of the inversion (if present) mapped for the NAE domain for 25th January 2008, 12Z. Red indicates a low inversion, blue a high inversion. Bottom panel: The surface pressure analysis for 26th January 2008, 0Z (*Crown Copyright*).

globally (Warren et al., 1986) and in anticyclonic conditions can form persistent layers. Sc has a large impact on the distribution of the short and long-wave radiative flux divergence in the BL. These radiative effects, along with latent heating, regulate the BL

dynamics, turbulence generation and evolution. At the surface, the effect of Sc is to weaken the strength of the diurnal cycle of surface temperature by reflecting incoming short wave (SW) radiation back to space during the day and absorbing and re-emitting long wave (LW) radiation from the surface during the night.
The effect Sc has on surface temperature is important in the formation of many atmospheric features, such as fog, frost and convective storms, which affect the way people live their lives. Fog and frost can lead to large disruptions as safety is compromised. An example of such is the chaos at Heathrow, Christmas 2006. Here poor forecasts of the persistent fog made it impossible for the airport to plan and many people were left stranded at the airport due to flight cancellations. The poor forecasts of the fog, in this case, were contributed to by a difficulty in correctly representing the BL capping inversion (Lorenc, 2007b).

1.2 Sources of information about the boundary layer

The inversion height and BL structure, have been shown to be strongly linked to surface conditions and other important meteorological phenomena. A meteorological forecast such as that made by the Met Office's Unified Model (Davies et al., 2005), attempts to predict the BL structure by solving a (largely) initial value problem. Therefore for an accurate forecast of these features, it is important that we make use of all the information available about the BL structure to characterise the initial conditions of the model. Sources of information about the current state of the BL can be split into two main categories.

- *Observations* which have been made close to the present time and can provide information about a variety of variables at different locations.
- *A priori data* which is based on theoretical reasoning such as information about the previous state of the atmosphere and known physical and dynamical models. This source gives an estimate of the model variables described on the model grid.

The information provided by these two sources can be combined using data assimilation techniques to give an estimate of the true structure of the BL. This estimate should be a complete, physically consistent four dimensional representation of the true state of the BL (Daley, 1997). The observations alone are not enough to give this. This is primarily due to their inhomogeneity in space and time which means they cannot provide all the information necessary for each model grid point. Observations also contain errors due to the instrument. However, the observations are essential as they sample reality. On the other hand a priori data available from a previous forecast does give a representation of the model variables on the model grid points (as needed) but, due to model and initial condition errors will deviate from the truth. The information that these two sources provide about the BL is now discussed.

1.2.1 Observations

A vast amount of global data measuring atmospheric properties are available. These data include in-situ measurements such as from drifting buoys, aircraft, and vertical soundings

from radiosondes, and remote measurements from satellite radiometers and scatterometers and ground or space based radars and lidars. To capture the structure of the BL we require:

- High vertical resolution profiles of temperature and humidity in order to represent the inversion- at least 100m resolution.
- A high temporal resolution to capture the small scale processes which influence the height of the inversion and lead to the formation of cloud.

In this thesis we shall conduct a one dimensional (vertical) study and so at this point it is not necessary to obtain data with a high temporal resolution nor data with a high horizontal resolution.

a) In-situ observations

In-situ observations can give direct, essentially point measurements. An example of in-situ data is the vertical soundings of radiosondes. Radiosondes generally give high resolution temperature and humidity vertical profiles. This allows for the diagnosis of regions of saturated air giving the vertical location of cloud layers. An example of a radiosonde sounding is shown in the Stuve diagram in Figure 1.5. This gives a profile for temperature (right profile) and dew point temperature (left profile). Within the BL the temperature profile follows a dry adiabat (green lines). At approximately 920hPa (900m) the temperature and dew point temperature values coincide and so it can be deduced that a layer of cloud may be found here. It is also evident that this cloud is capped by an inversion.

Radiosondes sample the atmosphere at fixed times and from fixed locations. Often they are released only once or twice a day so do not give a high temporal resolution. They also tend to be released only from populated regions over land (see Figure 1.6 for a map of the coverage of all sondes) and so do not give an evenly distributed description of the global atmosphere.

b) Remotely sensed observations

Many remote sensing instruments measure electromagnetic radiances from a distance. They can be either passive, detecting natural radiation that is emitted or reflected by the atmosphere, or active, emitting energy in order to then detect and measure the radiation that is reflected or backscattered back from the atmosphere. Passive instruments are often on board satellites orbiting the Earth. Such Earth observation satellites can either be in a near-polar or geostationary orbit. Polar orbiting satellites orbit at a typical altitude of 850 km which means they pass over the poles and are in a sun-synchronous orbit, viewing the same region twice a day in roughly the same light conditions. Geostationary orbiting satellites orbit the Earth above the equator at altitudes of 35,880 km. The altitude of this orbit means they appear stationary with respect to the rotating Earth. These two orbits both have their own advantages and disadvantages depending on the feature being observed. Satellites in polar orbits have a higher spatial resolution due to orbiting at a lower altitude, although can view only a smaller region at a time. Satellites in geostationary orbits have a higher temporal resolution as they are viewing the same area constantly. Satellites, unlike sondes, allow for near global coverage as seen in Figure

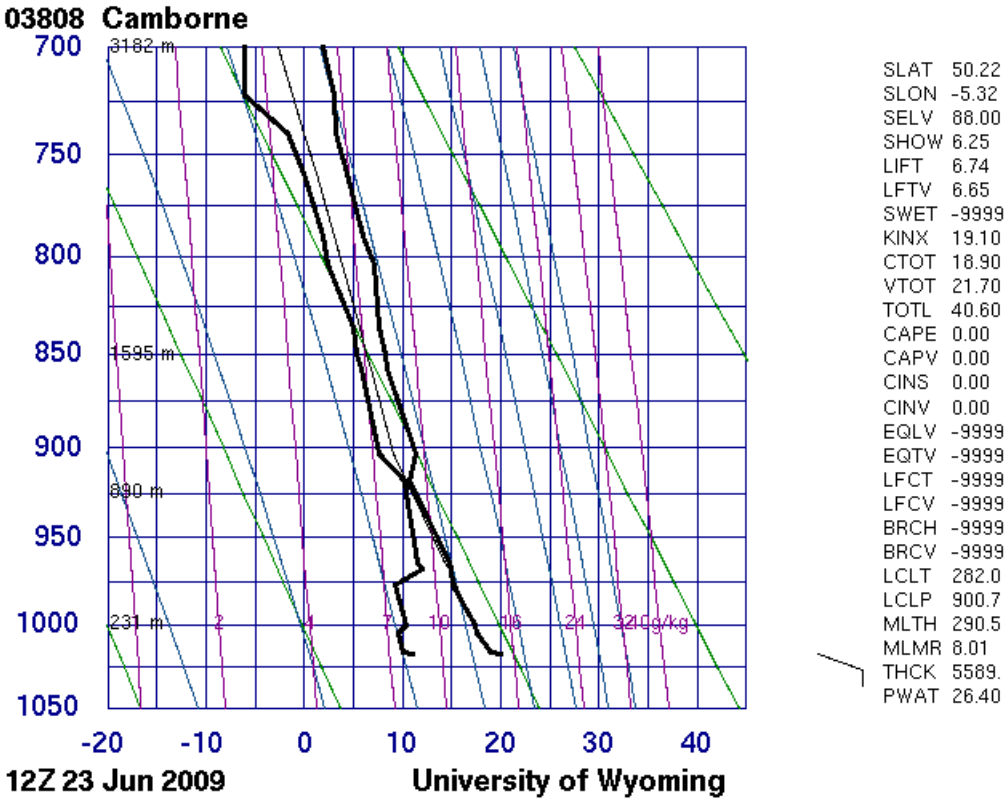


Figure 1.5: Radiosonde sounding at 12Z 23rd June 2009 over Cambourne, Cornwall. (The x-axis gives temperature (°C) and the y-axis gives pressure (hPa). The green diagonal lines give dry adiabats and the curved blue lines give moist adiabats. The diagonal pink line represents constant saturation mixing ratio (g/kg)). The profile on the right gives temperature and the profile on the left gives dew point temperature. This can be compared to Figure 1.3 of a lidar observation from the same day approximately 300km east of Cambourne. The BL here appears to be deeper at noon (2km compared to 1km). This highlights the spatial as well as temporal variation of the BL height. *Image from http://weather/uwyo/edu/cgi-bin/sounding.*

1.7. This Figure shows the coverage of ATOVS (Advanced TIROS Operational Vertical Sounder), which is in a polar orbit, in a 6hr period.

Many satellites measure top of the atmosphere (TOA) radiances, $L_\lambda(\text{TOA})$. The subscript λ represents the wavelength measured. From these measurements it is often possible to infer information about the whole atmospheric profile (Eyre, 1991). Contributions to TOA radiance for a particular wavelength are received from the whole depth of the atmosphere and $L_\lambda(\text{TOA})$ can be modelled by considering the thermal emission from and absorption by each layer of the atmosphere as the radiation propagates vertically towards the satellite overhead

$$L_\lambda(\text{TOA}) = L_\lambda(0)\tau_\lambda(0,\text{TOA}) + \int_0^{\text{TOA}} B_\lambda[T(s)]W_\lambda(s)ds. \tag{1.1}$$

Here τ_λ(0,TOA) is the transmittance of the atmosphere from the surface to TOA, and so the first term here gives the surface's contribution to the observed radiance. The second term gives the cumulative contribution to the observed radiance from each layer of the atmosphere, s, to TOA where $B_\lambda[T(s)]$ is the energy flux emitted by a black body at tem-

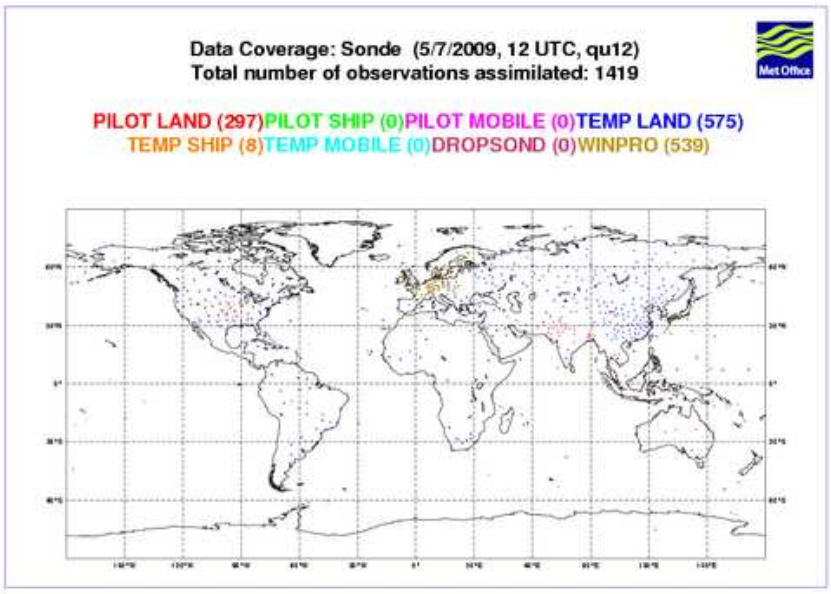


Figure 1.6: Coverage of sonde observations at 12UTC on 2nd July. *Image from http://www.metoffice.gov.uk* (*Crown Copyright*).

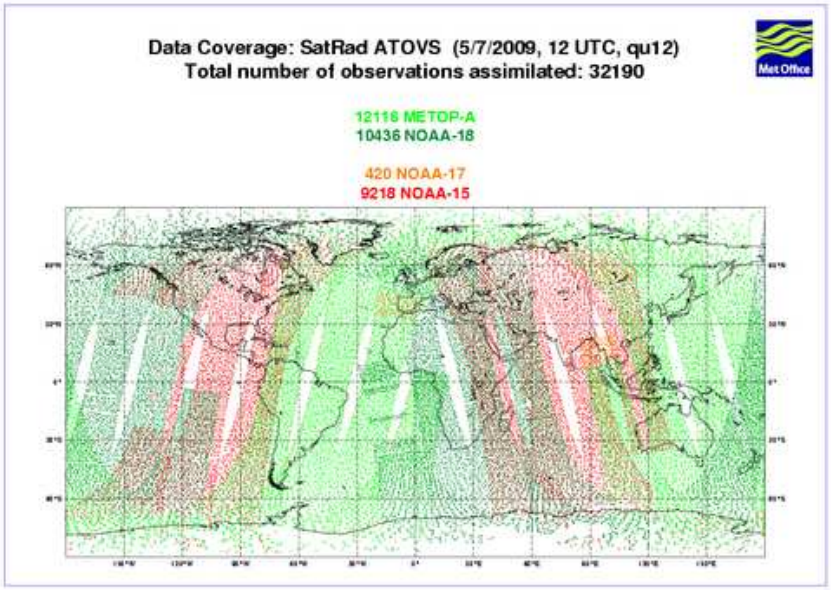


Figure 1.7: Coverage of ATOVS observations in a 6 hour period. *Image from http://www.metoffice.gov.uk* (*Crown Copyright*).

perature $T(s)$ (e.g. Salby, 1996) and $W_\lambda(s)$ is the weighting function $\approx \frac{d\tau_\lambda(s,\mathrm{TOA})}{ds}$ which specifies the layer from which the radiation emitted to space originates. The observed radiance is therefore potentially sensitive to temperature throughout the whole column of air and to different atmospheric constituents (such as carbon dioxide, water vapour, methane and ozone) which each absorb and emit thermal radiation to different degrees at different wavelengths. This causes the weighting function to peak at different heights for different wavelengths and so observations made at a variety of wavelengths can give information about different atmospheric constituents and the different regions of the atmosphere. This

means that observations of radiance at a large number of observed wavelengths are necessary to obtain height resolved information. The weighting functions however are broad and overlap. The later property means that the information given about the atmospheric profile by observations of TOA radiances at different wavelengths is not independent (see Eyre (1991) for examples of weighting functions).
The instrument with the highest spectral resolution that is currently used for meteorological data assimilation is the Infrared Atmospheric Sounding Interferometer (IASI), on board the Metop satellite which flies in a sun-synchronous orbit at an altitude of 817 km. IASI measures approximately 8000 different wavelengths (Matricardi and Saunders, 1999). These wavelengths are in the infrared band between wavelengths of 3.4 and 15.5 microns. This enables retrievals of temperature and humidity to be made at high vertical resolution in the troposphere and the lower stratosphere[2]. However it is still difficult to gain the necessary resolution to look at the thermal and moisture structure associated with the boundary layer inversion.
Clouds also have a large impact on the utility of satellite observations in the infra-red. The strong absorption of thermal radiation by clouds makes it difficult to obtain information from below cloud top when clouds are present (Errico et al., 2007). Due to the particular difficulties in interpreting satellite data in the presence of cloud only the clear sky radiances, as modelled by (1.1), have been used until recently within operational forecasts at the Met Office (MetO). However cloud effected regions are particularly important to observe (McNally, 2002) and so is an area of active research. Recently satellite observed radiances in cloudy regions which are not sensitive to cloud have been assimilated (Pavelin et al., 2008), by choosing to assimilate wavelengths whose weighting functions peak above the cloud level.
Important examples of active remote sensors are radars (RAdio Detection And Ranging) and lidars (LIght Detection And Ranging) (Figure 1.3 shows the information that can be derived from a lidar instrument). These can be either ground based or on board a satellite (e.g. CloudSat (Stephens et al., 2002) and CALIPSO (Winker et al., 2009)). From Figure 1.3 we can see that ground based lidars can give a lot of information about the height of an inversion, due to the sharp decrease in aerosols, although it is again difficult to infer information from these measurements about the temperature and humidity profiles.

Due to the poor vertical resolution of operational nadir (downward) viewing satellite observations and the difficulty with observing in regions of cloud, it is difficult to obtain enough information about the structure and height of the BL capping inversion and the associated temperature and humidity profiles. For this thesis it has therefore been decided to make use of radiosonde observations for a 1D study of the assimilation of the BL information. However, the benefits of satellite data providing a continuous observation source means that their future use for solving this problem should be considered. This will be discussed further in Chapter 6.

[2]Compared to other satellite instruments. A vertical resolution of approximately 1km is possible.

1.2.2 A priori data

A-priori data for the atmosphere are produced from numerical weather forecast models. These models represent the evolving atmosphere and use a variety of grid sizes. The MetO for instance currently run their NAE model at a resolution of 12km with a nested higher resolution of 1.5km over the UK. 50 vertical levels are available in the UM (Unified Model) which are spread across 65km of the atmosphere. Even though the vertical spacing of the levels is lowest at lower levels it is unlikely that the inversion is captured well in some cases.

As discussed already, the BL structure is a difficult feature to adequately model. This is due in part to the formation of the BL capping inversion being dependent on many different processes acting over a range of scales. Due to the limited spatial resolution of the model it is difficult to explicitly model all of these processes and therefore the representation of some of these processes is dependent upon parametrisation schemes. These parameterisation schemes relate sub-grid scale processes to known model parameters (see Stensrud, 2007). Turbulence and cloud, in particular, must be parametrised (e.g. Tiedtke 1993; Lock 1998). These inadequacies in the model-generated a-priori data are one of the reasons why data assimilation is needed in order to make use of observational data.

1.2.3 Data uncertainty

Both observations and the forecasts are subjects to errors. To make the most of both sets of data the statistical characteristics of these errors need to be understood, as will be seen in Chapter 2. Observational uncertainty comes from instrument noise and the way the instrument is used. Representativeness error is also often coupled with observational error. Representativeness error is due to the discrete sampling of a continuous medium (Daley, 1993; Liu and Rabier, 2002) and arises when the model is compared to the observations, via the *observation operator*. The observation operator itself may not be correctly specified, which is another source of error.

Errors in the forecast can come from a larger range of sources. There may be errors in the initial conditions and boundary conditions. There are also certainly errors in the model itself especially in the modelling of sub-grid scale processes which must be parameterised. However, even if the model was perfect, small initial condition errors growing over time would also be observed due to the non-linearity and instability of the dynamics of the atmosphere (Leith, 1978). These errors can be biased[3], spatially or temporally correlated and dependent upon the atmospheric state. The characteristics of these errors are summarised in table 1.1.

Traditionally only errors in amplitude are accounted for in data assimilation, which are usually assumed to be Gaussian. Amplitude errors are small scale errors which can be described on the model or observational level (see Figure 1.8, left panel). However, positional errors, which are associated with a feature and can be of a much larger scale than amplitude errors, can also be present. Positional errors are particularly important for sharp features such as the BL capping inversion assuming the inversion is captured by the model resolution. If the inversion is not captured by the model resolution then it is not

[3]In this thesis biases (or any systematic errors) are not considered. It is assumed that it is possible to remove these errors in a preprocessing step if they are indeed present.

Observations	SYSTEMATIC	• Bias due to incorrect use of instrument
	RANDOM	• Noise in measurement process
Observation operator	SYSTEMATIC	• Imperfections in formulation of operator • Unrepresented processes in input state that would affect the measurement
	RANDOM	• Unknown stochastic processes affecting measurement • Sub-grid scale fluctuations unaccounted for
A-priori	SYSTEMATIC	• Bias in initial conditions in model used to generate forecast • Model formulation imperfections • Unknown forcings
	RANDOM	• Unknown initial and boundary conditions • Random stochastic forcing of model

Table 1.1: Error Classification.

possible to define a positional error for this feature. A schematic of these two errors for a temperature inversion is given in Figure 1.8. The black line represents the mean temperature expected (identical in each case), the red bars represent the error standard deviation from this mean and so the grey region represents the region where, given these two types of errors, it is expected that the temperature values will lie (taking into account the effect of adiabatic motion for the positional error). If positional errors are not explicitly described within the assimilation scheme then they can appear as large systematic amplitude errors, due to the whole region between the estimated inversion and true inversion being warmer (if the inversion is too low) or cooler (if the inversion is too high). Neglecting even small positional errors can lead to a poor analysis (e.g. Hoffman, 1995; Ravela et al., 2007). For instance, later it will be shown that not accounting for positional errors in the BL capping inversion can weaken the inversion strength and erroneously affect the diagnosis of cloud. This could have a potentially large impact on the utility of high resolution forecasts

and lead to large errors in the forecasts of BL Sc and subsequent features affected by the presence of BL Sc.

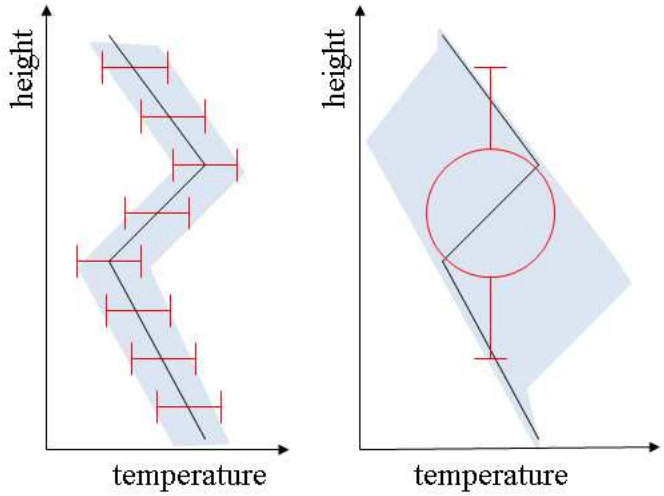


Figure 1.8: Schematic of amplitude (left) and positional (right) errors. The errors are represented by the bars and the grey region represents the typical values expected for temperature as a function of height given the two different errors (including adiabatic effects associated with the positional error).

1.3 This thesis

The aim of this thesis is to give an improved representation of the BL structure to be used as the initial conditions for a numerical weather forecast. This will be achieved through the assimilation of radiosonde data and a previous forecast valid for the current time.
For the assimilation to be optimal an accurate description of the errors associated with each source of data is necessary so that their influence can be weighted accordingly. The variability of the inversion complicates this, especially for the a-priori data, making a description of the error dependent upon the state of the atmosphere desirable.
The presence of the inversion in the temperature profile also means that an amplitude description of the errors alone is no longer adequate. If there is a positional error in the forecast of the inversion, it is important that this is included in the assimilation explicitly. This thesis examines this problem in the following way:

- In Chapter 2 the theory of Variational data assimilation is derived as it is conventionally formulated (by trying to reduce amplitude errors). From this we will see the importance of the forecast error statistics and give a flow-dependent estimate of their correlation structure for the BL. This will give a description only of the amplitude errors. If positional errors are also present this will be shown to lead to problems in the assimilation and the inversion structure may be lost in the final analysis.

- In Chapter 3 a new assimilation method will be formulated which explicitly allows for the inclusion of a positional error in a 1DVar framework (in the vertical). This is referred to as the 'floating boundary layer scheme'.

- In Chapter 4 some initial results from the univariate assimilation of temperature will be given. These will highlight some of the unique aspects of the floating BL scheme. In the final part of this chapter the positional error of the inversion in the forecast will be quantified using a large sample of radiosonde profiles. From this sample we will also compare the performance of the floating BL scheme to the standard scheme which only takes into account the amplitude errors.

- The new scheme will be extended in Chapter 5 to assimilate both temperature and humidity. A cloud scheme will also be incorporated to study how the the floating BL scheme affects the diagnosis of cloud.

- In Chapter 6 a summary of the conclusions will be given. Possible future work will also be discussed, including how the scheme could be incorporated in a full 3D or 4D variational data assimilation system.

Chapter 2

Variational Data Assimilation (Var)

There are several different ways a-priori and observed information can be combined. These range from very simple methods such as *successive correction methods* (SCM) (Bergthorsson and Döös, 1995; Cressman, 1959), *nudging* (Hoke and Anthes, 1976) and *optimal interpolation* (OI) (Gandin, 1963) to much more sophisticated and expensive methods based on the *Kalman filter* (KF) (e.g., Ghil et al., 1981). Most operational weather centres use *variational* data assimilation (Var) which in the simplest sense can be thought of as finding the least squares fit to the a-priori data and observations (Lewis et al., 2006) and can be shown to be equivalent to KF in some limits (Lorenc, 1986).
In a 6hr data assimilation (DA) cycle, a-priori information comes from the *background*, $\mathbf{x}^{\mathrm{b}}$, which is usually a previous 6hr forecast valid for the current time (the Ide et al. (1997) notation is used where possible). The background is taken to be the first guess of the true state of the atmosphere. This guess is then improved by taking into account the information provided by the observations, $\mathbf{y}$. In order to directly compare the model data with observations, the background is interpolated to the observation locations and if necessary also transformed to the observation variables. The transform, $\mathcal{H}$, which acts on $\mathbf{x}$ is known as the 'observation operator' or the 'forward model'. The difference between the observations and the background, $\mathbf{y} - \mathcal{H}(\mathbf{x}^{\mathrm{b}})$ is called the *innovation*. In order to obtain the *analysis*, $\mathbf{x}^{\mathrm{a}}$, the innovation is weighted and added to the background,

$$\mathbf{x}^{\mathrm{a}} = \mathbf{x}^{\mathrm{b}} + \mathbf{K}[\mathbf{y} - \mathcal{H}(\mathbf{x}^{\mathrm{b}})]. \tag{2.1}$$

The way the weight, $\mathbf{K}$, given to the innovation is chosen differs between SCM, OI, Var and KF. In the SCM, which is one of the earliest and simplest methods of DA, the weight is determined empirically based on the distance between the observations and the grid point. In OI the weight is chosen such that the analysis error is minimised and in Var it is chosen so that the probability of the analysis is maximised. In the next section the general Var approach will be formulated.

2.1 Underlying Theory

Var aims to robustly combine observations and a-priori data to obtain the most probable state of the atmosphere. Using Bayes' theorem we can express the conditional probability of the state of the atmosphere given the observations, $P(\mathbf{x}|\mathbf{y})$, known as the posterior probability, as:

$$P(\mathbf{x}|\mathbf{y}) = P(\mathbf{y}|\mathbf{x})P(\mathbf{x})/P(\mathbf{y}). \tag{2.2}$$

The normalising constant, $P(\mathbf{y})$ is often ignored because it is independent of $\mathbf{x}$, leaving $P(\mathbf{x}|\mathbf{y}) \propto P(\mathbf{y}|\mathbf{x})P(\mathbf{x})$.
$P(\mathbf{y}|\mathbf{x})$, the likelihood probability, peaks at the observed value of the true state of the atmosphere but, due to errors in the observations, will also have a spread. $P(\mathbf{x})$, the a-priori probability, is the information we know about the true state before the observations are made, i.e. a previous forecast, and is also known to have a spread due to errors in the forecast. Taking the natural logarithm and then the exponential of (2.2) we can give a new expression for the posterior probability,

$$P(\mathbf{x}|\mathbf{y}) \propto \exp\left(\ln P(\mathbf{y}|\mathbf{x}) + \ln P(\mathbf{x})\right). \tag{2.3}$$

The $\mathbf{x}$ for which the posterior is maximised is a useful estimate of the truth and is called the mode of $P(\mathbf{x}|\mathbf{y})$. Maximising this probability is equivalent to minimising the negative logarithm of the same probability:

$$\max\left(P(\mathbf{x}|\mathbf{y})\right) = \min\left(-\ln P(\mathbf{y}|\mathbf{x}) - \ln P(\mathbf{x})\right). \tag{2.4}$$

The function to be minimised is known as the *cost function*, $J(\mathbf{x})$:

$$J(\mathbf{x}) = -\ln P(\mathbf{y}|\mathbf{x}) - \ln P(\mathbf{x}). \tag{2.5}$$

The variables for which J is minimised with respect to are known in the meteorological data assimilation literature as the *control variables* (or when collected together into a vector they are known as the control vector). In this case the control vector is $\mathbf{x}$, although it is sometimes practical to minimise J with respect to a control vector that is different from, but related to, $\mathbf{x}$, an illustration of this is shown in Chapter 5. The value of $\mathbf{x}$ at the minimum of J is the *analysis*.

In order to evaluate the likelihood and a-priori probabilities the errors in the background and observations must be known. The background errors, $\boldsymbol{\eta}^{\mathrm{b}}$, can be defined as

$$\boldsymbol{\eta}^{\mathrm{b}} = \mathbf{x}^{\mathrm{b}} - \mathbf{x}^{\mathrm{t}}, \tag{2.6}$$

where $\mathbf{x}^{\mathrm{t}}$ is the true state of the atmosphere. If the errors are unbiased then the average of $\boldsymbol{\eta}^{\mathrm{b}}$, represented by $\langle\boldsymbol{\eta}^{\mathrm{b}}\rangle$, should be zero and so the background error covariances can be expressed as the outer product $\mathbf{B} = \langle\boldsymbol{\eta}^{\mathrm{b}}(\boldsymbol{\eta}^{\mathrm{b}})^{\mathrm{T}}\rangle$ (superscript T indicates the vector transpose). This is an $N \times N$ matrix, where N is the length of the state vector, $\mathbf{x}$. The diagonal elements of this matrix are the error variance for each element of $\mathbf{x}^{\mathrm{b}}$. The off diagonal elements are the error covariances between the different elements of $\mathbf{x}^{\mathrm{b}}$. These can include both the spatial correlation for one variable (autocorrelations) and multivariate

correlation. For example geostrophic balance within the forecast model implies that if background pressure values need to be corrected then so too should the wind values and the background errors of these two fields should be correlated.
Similarly the observation errors, $\boldsymbol{\eta}^{\mathrm{o}}$, can be defined as

$$\boldsymbol{\eta}^{\mathrm{o}} = \mathbf{y} - \mathcal{H}(\mathbf{x}^{\mathrm{t}}). \tag{2.7}$$

Again assuming the observations are unbiased the observation error covariances are $\mathbf{R} = \left\langle \boldsymbol{\eta}^{\mathrm{o}}(\boldsymbol{\eta}^{\mathrm{o}})^{\mathrm{T}} \right\rangle$ which are stored in a $P \times P$ matrix, where P is the length of the observation vector. It is often assumed this matrix is diagonal for simplicity as P may be large making it difficult to store and operate with a full $\mathbf{R}$. It can also be difficult to accurately define the observed error correlations and so instead often the diagonal elements are increased to allow for some effect of correlations present in the observations. This has an impact on the density of observations that can be used and it is often necessary to thin them (e.g., Liu and Rabier, 2002; 2003).

It is reasonable in many cases to assume that both the likelihood and a-priori probabilities can be modelled as Gaussian and non biased, so a description of the error (co)variances is sufficient to give a complete description of the probability distribution functions (PDFs). This assumption allows us to now express the likelihood probability as

$$P(\mathbf{y}|\mathbf{x}) = \frac{1}{((2\pi)^P|\mathbf{R}|)^{1/2}} \exp\left(-\frac{1}{2}(\mathbf{y} - \mathcal{H}(\mathbf{x}))^{\mathrm{T}}\mathbf{R}^{-1}(\mathbf{y} - \mathcal{H}(\mathbf{x}))\right), \tag{2.8}$$

where $|\mathbf{R}|$ denotes the determinant of the matrix $\mathbf{R}$. The priori probability can be expressed as

$$P(\mathbf{x}) = \frac{1}{((2\pi)^N|\mathbf{B}|)^{1/2}} \exp\left(-\frac{1}{2}(\mathbf{x} - \mathbf{x}^{\mathrm{b}})^{\mathrm{T}}\mathbf{B}^{-1}(\mathbf{x} - \mathbf{x}^{\mathrm{b}})\right). \tag{2.9}$$

Substituting (2.8) and (2.9) into (2.5) gives

$$J(\mathbf{x}) = \frac{1}{2}(\mathbf{x} - \mathbf{x}^{\mathrm{b}})^{\mathrm{T}}\mathbf{B}^{-1}(\mathbf{x} - \mathbf{x}^{\mathrm{b}}) + \frac{1}{2}(\mathbf{y} - \mathcal{H}(\mathbf{x}))^{\mathrm{T}}\mathbf{R}^{-1}(\mathbf{y} - \mathcal{H}(\mathbf{x})) + \mathrm{C}, \tag{2.10}$$

where $\mathrm{C} = \ln\left(\left((2\pi)^{(N+P)}|\mathbf{B}||\mathbf{R}|\right)^{1/2}\right)$. This constant may be ignored as it has no impact on finding the conditional mode, and so the cost function takes the more familiar form,

$$J(\mathbf{x}) = \frac{1}{2}(\mathbf{x} - \mathbf{x}^{\mathrm{b}})^{\mathrm{T}}\mathbf{B}^{-1}(\mathbf{x} - \mathbf{x}^{\mathrm{b}}) + \frac{1}{2}(\mathbf{y} - \mathcal{H}(\mathbf{x}))^{\mathrm{T}}\mathbf{R}^{-1}(\mathbf{y} - \mathcal{H}(\mathbf{x})). \tag{2.11}$$

The cost function evaluates the difference between the atmospheric state and the background (first term) and also the difference between the atmospheric state and the observations (second term), each of these terms are then weighted by the inverse of their errors. The assumption of Gaussian error statistics for the background and observation terms along with a linear observation operator means that Var becomes equivalent to the OI approach, which gives the analysis with the minimum analysis error variance (Lorenc, 1986). This is because for a Gaussian PDF the mode of the posterior PDF is the same as the mean, which is found in the OI approach.
If the errors of $\mathbf{x}^{\mathrm{b}}$ and $\mathbf{y}$ are truly Gaussian and non-biased, if $\mathbf{B}$ and $\mathbf{R}$ accurately describe

their error covariances, and if $\mathbf{H}$ is correct, then it can be shown that the expected vale of J evaluated at the analysis is equal to $P/2$ (half the number of observations). This is referred to as the Bennett-Talagrand theorem, a thorough proof of which is given in Appendix A. This result may be used to test the validity of the assumptions made about the prior and likelihood distributions.

In the case when the observation operator, $\mathcal{H}$, is linear, the errors are Gaussian and $\mathbf{B}$ and $\mathbf{R}$ are positive definite matrices, the cost function is quadratic and concave in $\mathbf{x}$ and its minimum can be calculated explicitly. First an expression for the gradient of the cost function with respect to the model state is needed,

$$\nabla_{\mathbf{x}} J = \mathbf{B}^{-1}\left(\mathbf{x}-\mathbf{x}^{\mathrm{b}}\right) - \mathbf{H}^{\mathrm{T}}\mathbf{R}^{-1}\left(\mathbf{y}-\mathbf{H}(\mathbf{x}-\mathbf{x}^{\mathrm{b}})-\mathcal{H}(\mathbf{x}^{\mathrm{b}})\right), \tag{2.12}$$

where $\mathcal{H}(\mathbf{x})$ has been approximated by $\mathbf{H}(\mathbf{x}-\mathbf{x}^{\mathrm{b}})-\mathcal{H}(\mathbf{x}^{\mathrm{b}})$. $\mathbf{H}$ is the Fréchet derivative of $\mathcal{H}(\mathbf{x})$ with respect to $\mathbf{x}$ evaluated at $\mathbf{x}=\mathbf{x}^{\mathrm{b}}$ (i.e. $\mathbf{H}=\frac{\partial\mathcal{H}(\mathbf{x})}{\partial\mathbf{x}}|_{\mathbf{x}^{\mathrm{b}}}$). $\mathbf{H}$ is a $P\times N$ matrix and is referred to as the linearised observation operator. The analysis is then given when (2.12) is equal to zero,

$$\mathbf{x}^{\mathrm{a}} = \mathbf{x}^{\mathrm{b}} + (\mathbf{B}^{-1}+\mathbf{H}^{\mathrm{T}}\mathbf{R}^{-1}\mathbf{H})^{-1}\mathbf{H}^{\mathrm{T}}\mathbf{R}^{-1}(\mathbf{y}-\mathcal{H}(\mathbf{x}^{\mathrm{b}})). \tag{2.13}$$

The weight given to the innovation (see (2.1)) can hence be expressed as

$$\mathbf{K} = (\mathbf{B}^{-1}+\mathbf{H}^{\mathrm{T}}\mathbf{R}^{-1}\mathbf{H})^{-1}\mathbf{H}^{\mathrm{T}}\mathbf{R}^{-1} \tag{2.14}$$

and is the so-called Kalman gain matrix. For the scalar case where $\mathbf{B}=\sigma_{\mathrm{b}}^2$ and $\mathbf{R}=\sigma_{\mathrm{o}}^2$, and $\mathbf{H}$ is the identity, (2.14) reduces to

$$K = \frac{\sigma_{\mathrm{o}}^{-2}}{\sigma_{\mathrm{b}}^{-2}+\sigma_{\mathrm{o}}^{-2}}. \tag{2.15}$$

Multiplying the numerator and denominator of (2.15) by $\sigma_{\mathrm{b}}^2\sigma_{\mathrm{o}}^2$ gives

$$K = \frac{\sigma_{\mathrm{b}}^2}{\sigma_{\mathrm{b}}^2+\sigma_{\mathrm{o}}^2}, \tag{2.16}$$

and (2.13) reduces to a linear combination of x^{b} and y:

$$x^{\mathrm{a}} = \left(\frac{\sigma_{\mathrm{o}}^2}{\sigma_{\mathrm{b}}^2+\sigma_{\mathrm{o}}^2}\right)x^{\mathrm{b}} + \left(\frac{\sigma_{\mathrm{b}}^2}{\sigma_{\mathrm{b}}^2+\sigma_{\mathrm{o}}^2}\right)y \tag{2.17}$$

and so the weight given to the observations is increased as the proportion of the error in the background to the error of the observation is increased (i.e. for $\sigma_b >> \sigma_o$, $x^a \to y$).

If the observation operator is non-linear then $\mathbf{H}$ is only an approximation valid at the point of linearisation and the minimum of J must be found iteratively. An estimate of the true state, $\mathbf{x}_{n+1}$, can be given from a previous guess, $\mathbf{x}_n$ (subscript n is the iteration number),

$$\mathbf{x}_{n+1} = \mathbf{x}_n + (\mathbf{B}^{-1}+\mathbf{H}_n^{\mathrm{T}}\mathbf{R}^{-1}\mathbf{H}_n)^{-1}\left(\mathbf{B}^{-1}(\mathbf{x}_n-\mathbf{x}^{\mathrm{b}})+\mathbf{H}_n^{\mathrm{T}}\mathbf{R}^{-1}(\mathbf{y}-\mathcal{H}(\mathbf{x}_n))\right). \tag{2.18}$$

At each iteration, n, $\mathbf{H}$ can be estimated as $\mathbf{H}_n = \frac{\partial \mathcal{H}(\mathbf{x})}{\partial \mathbf{x}}|_{\mathbf{x}_n}$. A non-linear observation operator gives a non quadratic cost function and so the first guess must be close to the truth to be confident the global minimum is found. A sensible first guess to use in this situation is the background.

The use of error statistics in the formulation of the analysis makes Var a more robust and theoretically valid method than more primitive methods such as SCM and nudging. It allows more weight to be given to good quality observations and also makes sure that poor observations have little impact on the analysis as seen for the scalar case (2.17). It is also possible to get a more realistic analysis by imposing physical relationships within the B-matrix so that the analysis is shifted away from the background in a way that is consistent in space and between different fields. This also ensures that the resulting analysis is realistically smooth. This shall be discussed further in Section 2.2.
The cost function (2.11) can be expanded to 4-dimensions allowing differences between the model state and observations made at different times, as well as locations to be taken into account. $\mathcal{H}$ then needs to also include the forecast model, $\mathcal{M}$, which needs to be linearised so that the gradient of the 4DVar cost function can be calculated. It is then often thought necessary to take into account the model error covariances (known as weak constraint 4DVar, e.g. Zupanski, 1997). For linear approximations weak constraint 4DVar is equivalent to a Kalman smoother at any time in the observation window (Fisher et al. 2005).

Assuming the analysis is unbiased then in a similar way to (2.6) and (2.7) the error in the analysis can be defined as

$$\boldsymbol{\eta}^{\mathrm{a}} = \mathbf{x}^{\mathrm{a}} - \mathbf{x}^{\mathrm{t}}. \tag{2.19}$$

Substituting in the value for $\mathbf{x}^{\mathrm{a}}$ given by (2.13) and linearising $\mathcal{H}$ about $\mathbf{x}^{\mathrm{t}}$ we have

$$\begin{aligned} \boldsymbol{\eta}^{\mathrm{a}} &= \mathbf{x}^{\mathrm{b}} - \mathbf{x}^{\mathrm{t}} + \mathbf{K}(\mathbf{y} - \mathcal{H}(\mathbf{x}^{\mathrm{t}}) - \mathbf{H}(\mathbf{x}^{\mathrm{b}} - \mathbf{x}^{\mathrm{t}})) \\ &= \boldsymbol{\eta}^{\mathrm{b}} + \mathbf{K}(\boldsymbol{\eta}^{\mathrm{o}} - \mathbf{H}(\boldsymbol{\eta}^{\mathrm{b}})), \end{aligned} \tag{2.20}$$

where $\boldsymbol{\eta}^{\mathrm{b}}$ and $\boldsymbol{\eta}^{\mathrm{o}}$ are defined by (2.6) and (2.7) respectively.
The analysis error covariance matrix $\mathbf{P}^{\mathrm{a}}$ can now be given as

$$\mathbf{P}^{\mathrm{a}} = \left\langle \boldsymbol{\eta}^{\mathrm{a}} \left(\boldsymbol{\eta}^{\mathrm{a}}\right)^{\mathrm{T}} \right\rangle = \left\langle \left(\boldsymbol{\eta}^{\mathrm{b}} + \mathbf{K}\left(\boldsymbol{\eta}^{\mathrm{o}} - \mathbf{H}(\boldsymbol{\eta}^{\mathrm{b}})\right)\right)\left(\boldsymbol{\eta}^{\mathrm{b}} + \mathbf{K}\left(\boldsymbol{\eta}^{\mathrm{o}} - \mathbf{H}(\boldsymbol{\eta}^{\mathrm{b}})\right)\right)^{\mathrm{T}} \right\rangle. \tag{2.21}$$

The angled brackets, $\langle . \rangle$, represent averaging over a suitable sample. By assuming background and observation errors are uncorrelated, (2.21) can be shown to simplify to

$$\mathbf{P}^{\mathrm{a}} = (\mathbf{I} - \mathbf{K}\mathbf{H})\mathbf{B}. \tag{2.22}$$

and so $\mathbf{P}^{\mathrm{a}}$ is the background error reduced by information given by observations. The analysis error is unbiased and Gaussian if so too are the background and observation errors. In the case of Gaussian statistics this is the minimum variance, assuming the errors for the background and observations are correctly given.

2.2 Background Error Covariance Matrix (B-matrix)

A detailed review of the background error covariances is given by Bannister (2008a, 2008b). The background error covariance matrix is central to the correct formulation of variational data assimilation. A poorly specified $\mathbf{B}$, in an otherwise reasonable DA system, will not give rise to an optimal estimate of $\mathbf{x}^{\mathrm{a}}$. Equation (2.14) which specifies the weight given to the innovation when calculating the analysis can be rewritten using the identity $\left(\mathbf{B}^{-1}+\mathbf{H}^{\mathrm{T}}\mathbf{R}^{-1}\mathbf{H}\right)\mathbf{B}\mathbf{H}^{\mathrm{T}}=\mathbf{H}^{\mathrm{T}}\mathbf{R}^{-1}\left(\mathbf{R}+\mathbf{H}\mathbf{B}\mathbf{H}^{\mathrm{T}}\right)$, giving:

$$\mathbf{K}=\mathbf{B}\mathbf{H}^{\mathrm{T}}(\mathbf{R}+\mathbf{H}\mathbf{B}\mathbf{H}^{\mathrm{T}})^{-1}. \tag{2.23}$$

$\mathbf{K}$ written in this way shows that in the calculation of the analysis increment, $\mathbf{B}$ is the last operator to act. Bannister (2008a) illustrates the importance of $\mathbf{B}$ for the case when only one observation is present. The increment added to the background is proportional to the column of $\mathbf{B}$ which corresponds to the observation (this is known as a *structure function*). In this way the covariances specified in the B-matrix are responsible for the spreading of information so that one observation is able to have an effect on the analysis at grid points away from the location of the observation. More generally the B-matrix controls:

1. The weighting of the observation and background terms, seen in (2.16). In Figure 2.1 (left hand panel) this point is illustrated for a single observation of $\mathbf{x_1}$ at model level k (=11). $\mathbf{x_1}$ is a vector described on model levels $l=1:20$. Three analyses are calculated using (2.13) with $B_{\mathrm{kk}}/\sigma_{\mathrm{o}}^2$ =1 (blue), 2 (red) and 0.5 (green). Changing this ratio changes the weight given to the observation in the analysis.

2. The way information from an observation is spread to nearby points, allowing for a smooth and realistic looking analysis. The analyses in the left hand panel of Figure 2.1 have been calculated using a B-matrix which has the structure function, $\mathbf{B}_{l\mathrm{k}}$, given in the right hand panel of Figure 2.1, resulting in the analyses shifting away from the background in the whole region where the background errors are correlated with level k

3. How an observation in one field produces an analysis increment in another field, for example known physical relations can be implemented in the multivariate background error covariances. This is illustrated in Figure 2.2, where now $\mathbf{x}=\begin{pmatrix}\mathbf{x_1}\\ \mathbf{x_2}\end{pmatrix}$, $\mathbf{x_1}$ again is a vector described on model levels $l=1:20$ and $\mathbf{x_2}$ is a vector described on model levels $m=1:20$. Only one observation is available and is of $\mathbf{x_1}$ at level k (left hand panel). However this observation of the $\mathbf{x_1}$ field can affect the analysis of $\mathbf{x_2}$ (middle panel) through correlations between errors in $\mathbf{x_1}$ at level k and errors in $\mathbf{x_2}$ as given by the structure function $\mathbf{B}_{m\mathrm{k}}$ (right hand panel).

4. The ability of one observation to reinforce another. In Figure 2.3 two observations of $\mathbf{x_1}$ are available at k=11 and also at k=8. Left hand panel: the analyses for the assimilation of the individual observations are given by the blue and green profiles. When both observations are assimilated the red analysis is given. Right hand panel: in the region were the the structure functions overlap the analysis has more trust in the observations when both are available.

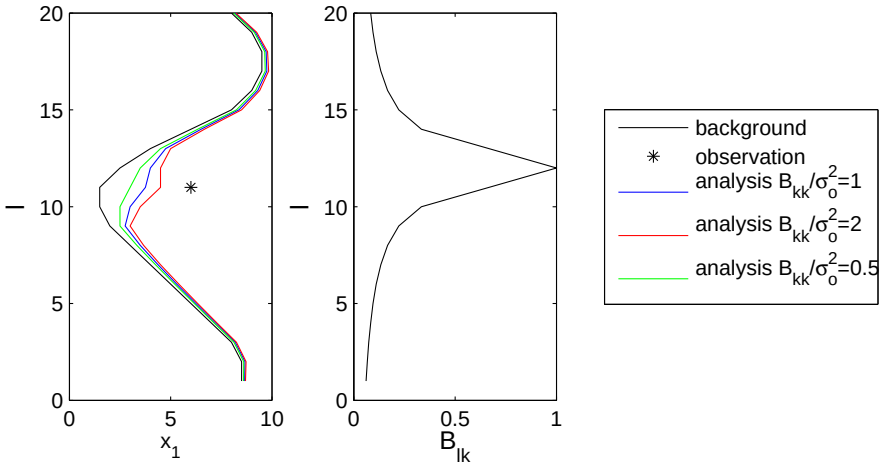


Figure 2.1: Left hand panel: The effect of varying the ratio of σ_o^2 to $B_{\rm kk}$ for an observation at level k (star) on the analysis (green, blue, red). Right hand panel: The structure function, $\mathbf{B}_{lk}$, controls the spread of the information from the single observation at level k locally to other levels: l=1:20.

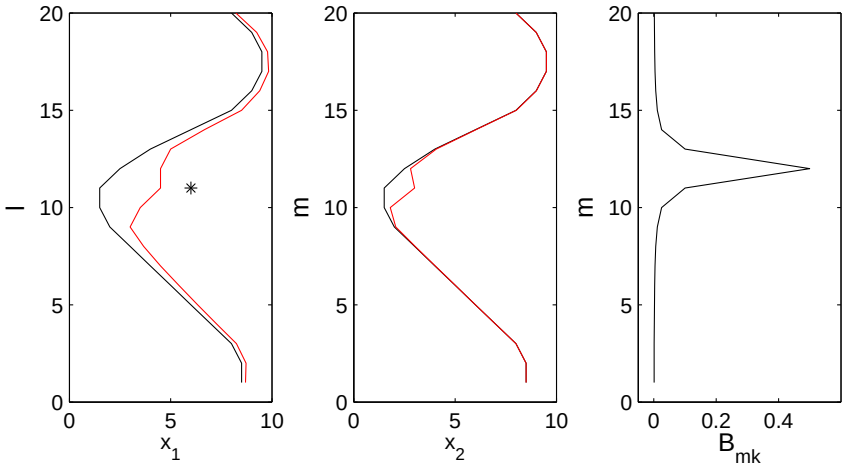


Figure 2.2: Correlations in $\mathbf{B}$ also allow for information from an observation to spread to other fields for example the middle panel gives the analysis calculated for field $\mathbf{x}_2$ (elements $m = 1 : 20$) when only an observation is available of $\mathbf{x}_1$ at level k (seen in the left hand panel). The multivariate part of the structure function $\mathbf{B}_{mk}$, is given in the right hand panel.

The third point here is important when producing an analysis of temperature and humidity especially when cloud is present. The temperature and humidity fields must remain consistent with one another as each is varied. For example when cloud is present in the BL and the cloud is capped by an inversion in the temperature profile there will also simultaneously be a sudden drying of the humidity profile above the cloud top. The inclusion of covariance between temperature and humidity fields within $\mathbf{B}$ is one way these two fields can be forced to remain consistent as their values are varied in the minimisation of J, so that diagnosed cloud is realistic.

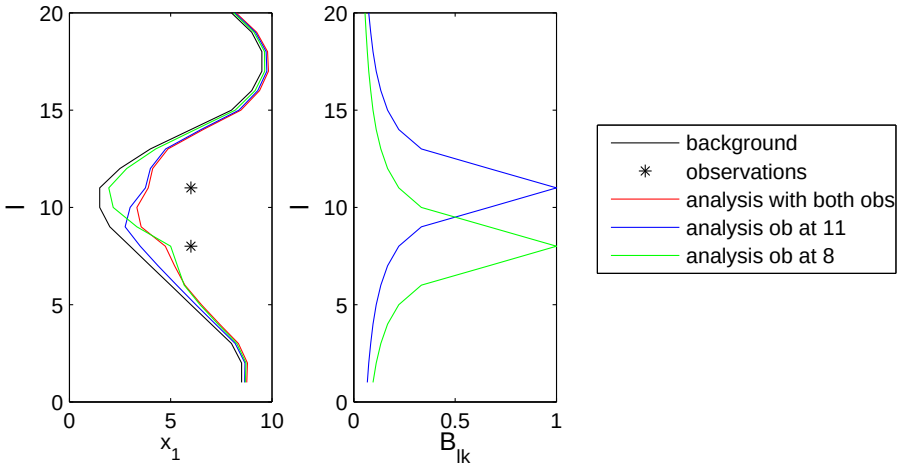


Figure 2.3: If two observations are available at levels which have overlapping structure functions then they can reinforce each other and have a larger impact on the analysis in the region where the structure functions overlap. Left hand panel: Shows the analysis when two observations are available (red) compared to the analysis of either one (green and blue profiles). Right hand panel: The structure functions can be seen to overlap.

As given in the previous section, assuming there are no biases in the background error, $\mathbf{B}$, can be defined as

$$\mathbf{B} = \left\langle \boldsymbol{\eta}^{\mathrm{b}} (\boldsymbol{\eta}^{\mathrm{b}})^{\mathrm{T}} \right\rangle \tag{2.24}$$

where $\boldsymbol{\eta}^{\mathrm{b}}$ is defined by equation (2.6). If biases are present and known these can be removed prior to calculating (2.24). Provided that background errors are sufficiently sampled by the population of $\boldsymbol{\eta}^{\mathrm{b}}$, (2.24) produces a full positive definite matrix (implying all eigenvalues are greater than zero, (Gaspari and Cohn, 1999)). This means that the background term in the cost function is convex in all directions in state space, and so a minimum exists.

In most data assimilation systems, errors are assumed to be Gaussian, which (strictly speaking) they rarely are. If the errors are not Gaussian then $\mathbf{B}$ does not give a complete description of the background errors. Gaussian errors can be a particularly poor assumption for humidity variables such as specific humidity which cannot have negative values or cloud variables such as cloud fraction which is bounded by zero and one. However incorporating higher order moments to give a more complete description of the background errors is usually infeasible due to the size of the atmospheric system.

The background state is derived from a 6hr forecast and so the background errors have characteristics in common with errors in the forecast. The errors are flow dependent and so should ideally be calculated as the forecast evolves. The Kalman filter (KF) allows for this by using the forecast model itself to compute the forecast error covariance:

$$\mathbf{P}^{\mathrm{f}} = \mathbf{M}\mathbf{P}^{\mathrm{a}}\mathbf{M}^{\mathrm{T}} + \mathbf{Q}, \tag{2.25}$$

where $\mathbf{P}^{\mathrm{a}}$ is the analysis error covariance matrix from the previous cycle, $\mathbf{M}$ is the tangent linearisation of the forecast model, that runs between the last analysis and the current

background, and $\mathbf{Q}$ is the model error covariance matrix. The KF is derived from estimation theory where as Var is derived from control theory. In (2.25) it is assumed that the model error is unbiased and not temporally correlated. This would give an optimal background error covariance matrix if the forecast model and observation operator are linear (i.e. $\mathbf{Mx} = \mathcal{M}(\mathbf{x})$ and $\mathbf{Hx} = \mathcal{H}(\mathbf{x})$). If $\mathcal{M}$ is not linear (as is the case for NWP) or $\mathcal{H}$ is not linear then $\mathbf{P}^{\mathrm{f}}$ should no longer be Gaussian. The non-linear form of KF is called the extended Kalman filter (EKF). This produces higher order statistical moments and so a moment closure scheme is necessary, which can cause the EKF to diverge from observations if they are too sparse. $\mathbf{P}^{\mathrm{f}}$ would be the natural choice for $\mathbf{B}$ in (2.11) unfortunately (2.25) is too computationally expensive due to the size of the matrices involved for NWP. Therefore a $\mathbf{B}$ is used that can only be an estimate of $\mathbf{P}^{\mathrm{f}}$. Different methods of estimating $\mathbf{B}$ are described in the next section.

2.2.1 Methods of estimating the B-matrix

In order to calculate the background errors using equation (2.24) the true state of the atmosphere must be known. Unfortunately this is never possible and so estimates are necessary. If observations available are reliable and plentiful then information about the background error may be gained by studying observed minus background values. This is the method of *analysis of innovations* (Rutherford, 1972). This method eliminates the 'truth' term from expressions for $\boldsymbol{\eta}^{\mathrm{b}}$ and $\boldsymbol{\eta}^{\mathrm{o}}$ by substituting (2.6) into (2.7), assuming there are no biases, and linearising $\mathcal{H}$ about $\mathbf{x}^{\mathrm{b}}$,

$$\mathbf{y} - \mathcal{H}(\mathbf{x}^{\mathrm{b}}) \approx \boldsymbol{\eta}^{\mathrm{o}} - \mathbf{H}\boldsymbol{\eta}^{\mathrm{b}}. \tag{2.26}$$

In this way the innovation gives information about both the observation and background errors in observation space. Taking the product of (2.26) at different elements of the observation space, i and j, and averaging gives

$$\left\langle \left(y_i - (\mathcal{H}(\mathbf{x}^{\mathrm{b}}))_i\right)\left(y_j - (\mathcal{H}(\mathbf{x}^{\mathrm{b}}))_j\right)\right\rangle = \left\langle \left(\eta_i^{\mathrm{o}} - (\mathbf{H}\boldsymbol{\eta}^{\mathrm{b}})_i\right)\left(\eta_j^{\mathrm{o}} - (\mathbf{H}\boldsymbol{\eta}^{\mathrm{b}})_j\right)\right\rangle. \tag{2.27}$$

Here the angled brackets represent averaging over space and time. Now assuming that observation errors and background errors are uncorrelated and observation errors are uncorrelated between themselves allows us to give an expression for $\mathbf{B}$ in observation space, for when $i \neq j$,

$$\left\langle \left(y_i - (\mathcal{H}(\mathbf{x}^{\mathrm{b}}))_i\right)\left(y_j - (\mathcal{H}(\mathbf{x}^{\mathrm{b}}))_j\right)\right\rangle \approx \left\langle \left(\mathbf{H}\boldsymbol{\eta}^{\mathrm{b}}\right)_i \left(\mathbf{H}\boldsymbol{\eta}^{\mathrm{b}}\right)_j \right\rangle = B_{ij}. \tag{2.28}$$

When $i = j$

$$\left\langle \left(y_i - \mathcal{H}(x^{\mathrm{b}})_i\right)\left(y_i - \mathcal{H}(x^{\mathrm{b}})_i\right)\right\rangle \approx R_{ii} + B_{ii}, \tag{2.29}$$

and so the observation error variances must be known to extract the background error variances. Assuming $\mathbf{B}$ is approximately static and homogeneous allows the errors to be averaged over both time and space. This method was used by the ECMWF (European Centre for Medium-range Weather forecasts) to evaluate $\mathbf{B}$ for their OI system (Hollingsworth and Lönnberg (1986), Lönnberg and Hollingsworth (1986)) and also to

diagnose information about their B-matrix in Var (Järvinen (2001)). This method is obviously limited by the number and quality of independent and unbiased in situ observations available. In Chapter 1 it was seen that the radiosonde network is too sparse to allow for the success of this method. In principle satellite observations could be used but it is more complicated in this case to separate out the errors from the observations as the assumption of uncorrelated errors is usually not valid.

A second popular method is known as the *NMC* method whose name is taken from the National Meteorological Centre now called NCEP (National Centre for Environmental Prediction). This method quantifies the background error by comparing forecasts of varying lengths valid at the same time (Parrish and Derber (1992)). For example writing 48hr and 24hr forecasts in terms of their errors

$$\begin{aligned} \mathbf{x}^{48} &= \mathbf{x}^{\mathrm{t}} + \boldsymbol{\eta}^{48} + \mathbf{b}^{48} \\ \mathbf{x}^{24} &= \mathbf{x}^{\mathrm{t}} + \boldsymbol{\eta}^{24} + \mathbf{b}^{24} \end{aligned} \tag{2.30}$$

where $\mathbf{x}^{48/24}$ represents a 48hr/24hr forecast, $\boldsymbol{\eta}$ are their errors and $\mathbf{b}$ their biases. Assuming the biases are the same, partly based on the lack of diurnal interference using forecasts separated by 24 hours, allows us to remove $\mathbf{x}^{\mathrm{t}}$ from the expression for the background errors

$$\delta\mathbf{x} = \mathbf{x}^{48} - \mathbf{x}^{24} = \boldsymbol{\eta}^{48} - \boldsymbol{\eta}^{24} \tag{2.31}$$

and so if the errors of the different length forecasts are uncorrelated we can derive an expression for $\mathbf{B}$:

$$\begin{aligned} \left\langle \delta\mathbf{x}\delta\mathbf{x}^{\mathrm{T}} \right\rangle &= \left\langle (\boldsymbol{\eta}^{48} - \boldsymbol{\eta}^{24})(\boldsymbol{\eta}^{48} - \boldsymbol{\eta}^{24})^{\mathrm{T}} \right\rangle \\ &= \left\langle \boldsymbol{\eta}^{48}(\boldsymbol{\eta}^{48})^{\mathrm{T}} \right\rangle + \left\langle \boldsymbol{\eta}^{24}(\boldsymbol{\eta}^{24})^{\mathrm{T}} \right\rangle \approx 2\mathbf{B}. \end{aligned} \tag{2.32}$$

Assuming 24 hour forecast errors have the same statistics as 48 hour forecasts. Again the angled brackets represent averaging over space and time.
A large number of forecasts are necessary, to give an adequate sample to average over, for this equality to approximately hold. These forecasts are taken over a long period of a few months and so can only give an estimate of climatological covariances. The NMC method has also been shown to incur problems in poorly observed regions where the forecasts vary very little, and at large scales (Berre, 2000; Berre et al., 2006) when the assumption that the forecasts are uncorrelated breaks down. To compensate for the known problem of underestimating the background error variance in regions of low observation density and overestimation of correlation length scales, adjustments are necessary (Derber and Bouttier, 1999; Ingleby, 2001).

The third main method for estimation of the background error covariance matrix discussed here makes use of an *ensemble* of forecasts which have been generated to have the same spread as the true forecast error PDF (e.g. Houtekamer et al. 1996; Anderson 2001). $\mathbf{B}$ can be constructed as

$$\mathbf{B} = \left\langle \left(\mathbf{x}^{\mathrm{b}} - \left\langle \mathbf{x}^{\mathrm{b}} \right\rangle\right) \left(\mathbf{x}^{\mathrm{b}} - \left\langle \mathbf{x}^{\mathrm{b}} \right\rangle\right)^{\mathrm{T}} \right\rangle, \tag{2.33}$$

where this time we are averaging over ensemble members. This allows us to get a proper assessment of what the background should be for different states of the atmosphere, assuming that the ensemble is able to fully represent different sources of error within the forecast. This method is used implicitly by the ensemble Kalman filter (EnKF) (e.g. Evensen, 1994; Hamil and Snyder, 2000).

In equation (2.33) it is assumed that the mean of the ensemble, $\langle \mathbf{x}^{\mathrm{b}} \rangle$, is an estimate of the true state of the atmosphere. An alternative estimate of the 'truth' could be the control member, as this is the ensemble member which has been run with our best available estimates of the initial conditions but this may no longer be the the central value of the background. The necessity to estimate the truth can also be avoided by instead looking at the difference between pairs of ensemble members.

In order to give a full rank estimate of $\mathbf{B}$ it is necessary that the number of ensemble members, n, is at least the length of the background vector, N. When $n << N$ false long-range correlations may appear (which need to be filtered in some way) and variances can be underestimated. The effective number of ensemble members can be increased by introducing either spatial or temporal averaging over neighbouring points although these will of course lose some information about the flow dependency of the errors that are being measured.

The advantages of flow dependence of the ensemble method are hindered by the expense of calculating the errors in this way and may need to be restricted to regions in which the flow dependency of the background errors is known to have the biggest impact, for example in quickly evolving storms.

2.2.2 Operational B-matrices

At the Met Office the NMC method (see Section 2.2.1) is currently used to diagnose the B-matrix used in their operational 4DVar system. As mentioned this is a climatological estimate of the background errors, and so it can only capture large scale and slowly forming background error structures. The correlation structure of the B-matrix used by the Met Office's 1DVar system (available as a deliverable of the NWPSAF (Numerical Weather Prediction Satellite Application Facility) EUMETSAT, 2007) for temperature and humidity is shown in Figure 2.4. It is difficult to quantify meaningful covariances between temperature and humidity when the error statistics are averaged over such a long time period and so to avoid including erroneous information in the matrix it is often assumed, as has been done in Figure 2.4, that temperature and humidity are uncorrelated.

Most of the discussion in this Chapter has been concerned with vertical aspects of covariances but a B-matrix used in 3D and 4DVar has horizontal aspects too. Basic methods of representing the B-matrix make assumptions of homogeneity (3D correlation functions are independent of horizontal position) and isotropy (3D correlation functions are independent of the orientation between two points in the same horizontal plane). There have been attempts to introduce inhomogeneous and non-isotropic statistics into the assimilation. Desroziers (1997) tried coordinate transforms, Riishøjgaard (1998) took into account gradient information, Fisher (2003) and Deckmyn and Berre (2005) have both looked at the use of wavelets, and recursive filters were studied by Purser and Wu (2003).

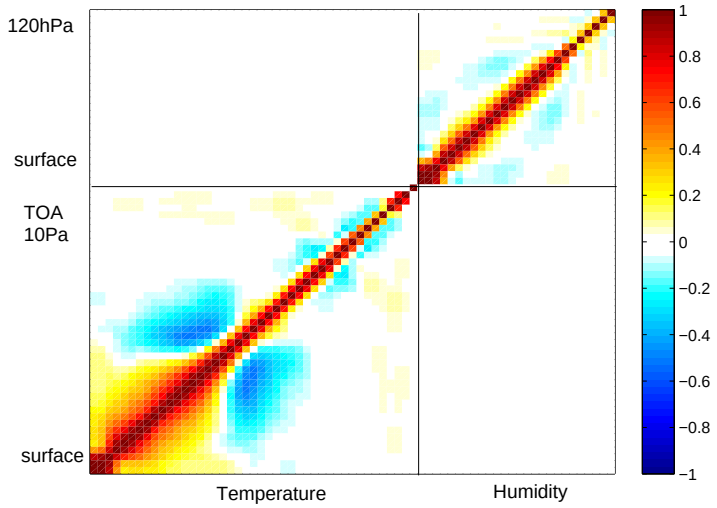


Figure 2.4: Background error correlation matrix as used by the Met Office's 1DVar system: the first 43 elements represent temperature at the 43 levels within the model, the next 26 variables are humidity for the lowest 26 levels of the model.

The non-linear nature of background errors also needs to be represented. This has been achieved by non-linear change of variables (e.g. Dee and da Silva, 2003) and introducing non-linear balance relationships (Fisher 2003). Recently the idea of combining static statistics with Extended KF (EKF) approaches (Hamill and Snyder, 2002; Etherton and Bishop, 2004; Buehner, 2005) has also been developed.
In the Met Office's variational system, J is minimised in terms of variables which are assumed to be uncorrelated (Lorenc et al., 2000; Met Office, 1995; Ingleby, 2001; Rawlins et al. 2007). These variables are not necessarily model variables, but alternative control variables, namely stream function, velocity potential, unbalanced pressure and relative humidity. The B-matrix needed can be specified in terms of the control variables, this is simplified due to their assumed uncorrelated property.

2.3 Ensemble derived B-matrix for the boundary layer

Accurate estimates of the background error statistics for the boundary layer are particularly important when attempting to analyse the boundary layer capping inversion. Due to the variability of the structure of the temperature inversion (see Chapter 1) the background error statistics for this region are highly flow dependent. A climatological estimate of the errors is therefore unsatisfactory and will lead to an erroneous spread in information from the observations and a sub-optimal analysis.

Since 2005 the UK Met Office (MetO) have been running an ensemble of 24 forecasts over the NAE domain (Bowler et al., 2008) at 24km horizontal resolution. If the ensembles are generated so that they have the same spread as the true forecast error PDF then looking at the spread in the ensembles (as described in Section 2.2.1) for particular times and

locations can give us valuable information about the flow dependence of the background errors.

The variability in the height and structure of the BL capping inversion leads to some very different correlation structures demonstrated in Figure 2.5 for temperature. These have been calculated using (2.33) with 6hr forecasts. In Figure 2.5, background error correlation matrices are shown for three different locations all valid at the same time. Figure 2.5a (CASE A) shows a background error correlation matrix for the case when a strong inversion is present at approximately 1000m, Figure 2.5b (CASE B) similarly shows the error correlation structure when there is a weaker inversion at approximately 1300m and Figure 2.5c (CASE C) shows the correlation structure when no inversion is resolved. In Figures 2.5a and 2.5b the decoupling between the errors in the boundary layer and the air above is evident, and stronger when the inversion is stronger. This is due to the inversion preventing the two masses of air from mixing, see Chapter 1. The accurate representation of this correlation structure is very important for the assimilation of the inversion in order to provide a correct account of how information from the observations should be spread. If in CASE A the B-matrix for CASE B were used, the observations made within the BL would be allowed to strongly affect the analysis up to 300m above the BL, effectively losing information about the BL. A worse scenario would be if the B-matrix for CASE C were used. Here the two air masses are able to mix and so there is no decoupling between the errors. Therefore if this B-matrix were used for CASE A then information from observations made in the BL would be allowed to spread throughout the depth of the troposphere, up to approximately 5km, and an even more significant amount of information about the BL would be lost. This is equivalent to what would happen if a climatological B-matrix were used which has averaged out the variable structure of the capping inversion.

Variability in the correlation structure above the BL between the three cases given in Figure 2.5 is also noticeable. In cases B and C regions of strong correlation are seen at the top of the troposphere which are decoupled with the air below. This structure is not present in case A. The differences in the correlation structures away from the BL may be due to turbulence caused by wind shear and the presence of fronts separating air masses. Again an averaged B-matrix, such as that shown in Figure 2.4, would not be capable of representing this variability in the correlation structure. These features above the BL will not be studied further within this Thesis.

2.3.1 Models of the B-matrix for the boundary layer

Alternative representations of the B-matrix, that represent the structure of the BL, may be considered by making a sensible change of control variable. This has been successfully tried for a similar problem of assimilating frontal information (Desroziers and Lafore (1993) and Desroziers (1997)). A front separates two different masses of air and is characterised by large gradients in wind and temperature in the horizontal. This is similar to the inversion separating the BL air from the free tropospheric air. Forecast errors are often assumed isotropic in the horizontal due to practical difficulties in implementing them otherwise. However since a front divides two different air masses, information from observations should not be spread across the front but instead should be spread along the front, and

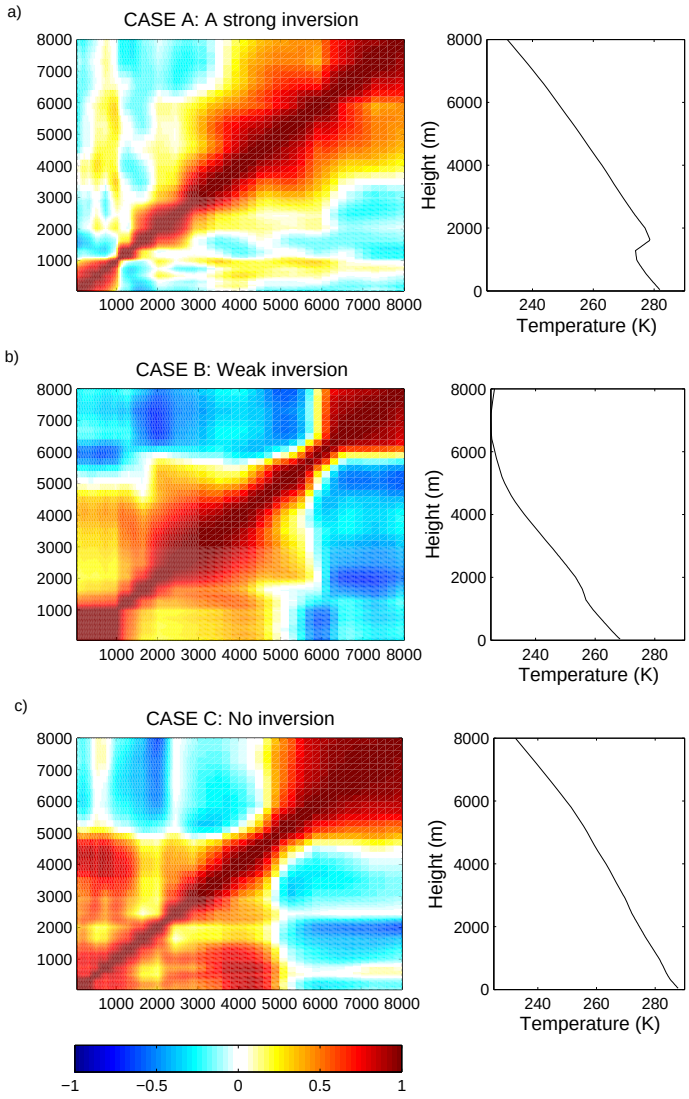


Figure 2.5: Examples of three different background error correlation matrices and profiles for temperature in the BL. CASE A: A strong low inversion. CASE B: A high inversion. CASE C: A weak inversion. The x and y coordinates of the matrices are altitude (m).

similarly information should not be spread across an inversion. The use of a change of coordinates via a control variable transform was shown to introduce the necessary flow dependence into the forecast errors to solve this frontal problem. Using semigeostrophic theory (Hoskins, 1975) they changed to geostrophic coordinates for the minimisation of the cost function. In these coordinates the assumption of isotropic forecast errors is more valid and introduces more realistic anisotropic error correlations in physical space.
Introducing a change of variables to describe the BL into the data assimilation scheme may allow us to consequently model the flow-dependent background error structure for the BL temperature profile as illustrated in Figure 2.5. A model of the BL temperature can be implemented using a change of control variable:

$$\mathbf{x} = \mathbf{U}\boldsymbol{\chi} \tag{2.34}$$

where $\mathbf{U}$ is the control variable transform and $\boldsymbol{\chi}$ is the new control variable. If $\mathbf{U}$ is dependent upon the atmospheric state then the implied background errors in model space, $\mathbf{B}_{\mathrm{x}}$, will also depend on the atmospheric state and in this way become flow dependent:

$$\mathbf{B}_{\mathrm{x}} = \mathbf{U}\mathbf{B}_{\chi}\mathbf{U}^{\mathrm{T}}. \tag{2.35}$$

As seen from Figure 2.5 it is important that the background errors are dependent upon the BL structure so that observed information is spread in the correct way, and if ensembles are not available it is useful to be able to describe the errors in terms of variables which are not dependent upon the BL structure. Here we illustrate this idea for temperature using a very simplified model of the BL temperature based on the bulk model as described by Stull (1988). It is assumed the mixed layer has a constant potential temperature, $\bar{\theta}_{\mathrm{ML}}$, the inversion can be described using a single value for its gradient, $\frac{\mathrm{d}\theta}{\mathrm{d}z}|_{\mathrm{inv}}$, and the lapse rate above the inversion is constant up to 4km, $\frac{\mathrm{d}\theta}{\mathrm{d}z}|_{\mathrm{FA}}$. There is also a separate value for the potential temperature at the surface, θ_{s}, and values of potential temperature are calculated for each model level, j, above 4km, $\boldsymbol{\theta}_{j\uparrow}$, i.e.

$$\boldsymbol{\chi} = \begin{pmatrix} \theta_{\mathrm{s}} \\ \bar{\theta}_{\mathrm{ML}} \\ \frac{\mathrm{d}\theta}{\mathrm{d}z}|_{\mathrm{inv}} \\ \frac{\mathrm{d}\theta}{\mathrm{d}z}|_{\mathrm{FA}} \\ \boldsymbol{\theta}_{j\uparrow} \end{pmatrix}. \tag{2.36}$$

In this way the model separates the BL air from the FA air. If $\mathbf{x}$ is now just a vector of temperature on 20 levels describing the troposphere we want to find a $\mathbf{U}$ which satisfies

$$\mathbf{x} = \mathbf{U} \begin{pmatrix} \theta_{\mathrm{s}} \\ \bar{\theta}_{\mathrm{ML}} \\ \frac{\mathrm{d}\theta}{\mathrm{d}z}|_{\mathrm{inv}} \\ \frac{\mathrm{d}\theta}{\mathrm{d}z}|_{\mathrm{FA}} \\ \boldsymbol{\theta}_{j\uparrow} \end{pmatrix}. \tag{2.37}$$

$\mathbf{U}$ is a linear transformation and can be shown to be a matrix of partial derivatives of $\mathbf{x}$ with respect to $\boldsymbol{\chi}$. In practice it is necessary to define the top and bottom of the inversion in order to be able to calculate the mean potential temperature within the mixed layer, $\bar{\theta}_{\mathrm{ML}}$, the gradient of potential temperature with respect to height within the inversion, $\frac{\mathrm{d}\theta}{\mathrm{d}z}|_{\mathrm{inv}}$, and the gradient of potential temperature with respect to height above the inversion up to 4km, $\frac{\mathrm{d}\theta}{\mathrm{d}z}|_{\mathrm{FA}}$.

For the control temperature profile given by Figure 2.6a in red, $\mathbf{x}$ has 20 elements and $\boldsymbol{\chi}$ has 10 elements (one each for θ_{s}, $\bar{\theta}_{\mathrm{ML}}$, $\frac{\mathrm{d}\theta}{\mathrm{d}z}|_{\mathrm{inv}}$ and $\frac{\mathrm{d}\theta}{\mathrm{d}z}|_{\mathrm{FA}}$ and six for $\boldsymbol{\theta}_{j\uparrow}$). In Figure 2.6b the potential temperature profile, θ, (blue) calculated from the control temperature profile, T, using $\theta = T\left(\frac{p_o}{p}\right)^{R/c_p}$ (where p_o is the reference pressure, R is the gas constant of air and c_p is the specific heat capacity at a constant pressure) is plotted in red. Also plotted is the potential temperature profile reconstructed from $\boldsymbol{\chi}$ (blue). It can be seen that in this case the assumption of a constant potential temperature within the BL is fairly poor and so at the top of the ML the potential temperature is too low which then causes the profile up to 4km to be too cold compared to the actual potential temperature profile. The $\mathbf{U}$ which satisfies (2.37), in this case a 20×10 matrix, can be calculated as illustrated in Figure 2.6c. The fit of $\mathbf{U}\boldsymbol{\chi}$ to $\mathbf{x}$ is seen in Figure 2.6d which has captured the height and structure of the inversion, but is too cool between 1km and 4km due to the assumption of a constant potential temperature in the BL.

We can now assess the success of this model by comparing the correlation part of $\mathbf{U}\mathbf{B}_{\chi}\mathbf{U}^{\mathrm{T}}$ to the background error correlation structure calculated via the ensemble method. The ensemble members for this example are plotted in Figure 2.6a. Figure 2.7a shows the background error correlation matrix for $\boldsymbol{\chi}$, $\mathbf{B}_{\chi}$, which is assumed to be static as it is not sensitive to the inversion height. This has been calculated by transforming each temperature profile of the ensemble, shown in Figure 2.6a, to $\boldsymbol{\chi}$ and then calculating $\mathbf{B}_{\chi} = \left\langle (\boldsymbol{\chi} - \langle\boldsymbol{\chi}\rangle)(\boldsymbol{\chi} - \langle\boldsymbol{\chi}\rangle)^{\mathrm{T}} \right\rangle$. Figure 2.7b shows $\mathbf{U}\mathbf{B}_{\chi}\mathbf{U}^{\mathrm{T}}$, the implied background error correlation matrix for $\mathbf{x}$, which can be used to approximate Figure 2.7c the actual background error correlation matrix for $\mathbf{x}$. We can see that the change of variables has been successful in recreating the region of strong positive correlation seen within the BL and the decoupling of errors with the air above. However the implied B-matrix lacks the finer structure seen in the actual B-matrix due to the simplicity of this model and the strength of the correlations has been greatly increased. We also see in Figure 2.8 that the error variances have been greatly increased above the inversion height up to 4km where the lapse rate was assumed to be constant. This is also seen as a region of strong correlation between levels 9 and 15 in Figure 2.7b, which is not very accurate as this region is not well mixed.

If no inversion is present it is not possible to calculate the top of the BL in this way and a climatological B-matrix may be used. Modelling humidity for the BL is not as simple and it is not so obvious how the cross correlations between temperature and humidity could be modelled. As computational expense is not an issue for these 1D experiments, in this thesis the ensemble-derived B-matrices will be used in the theoretical studies. Later when real, past data are used and ensembles are not available, the climatological B-matrix derived from the NMC method will be used.

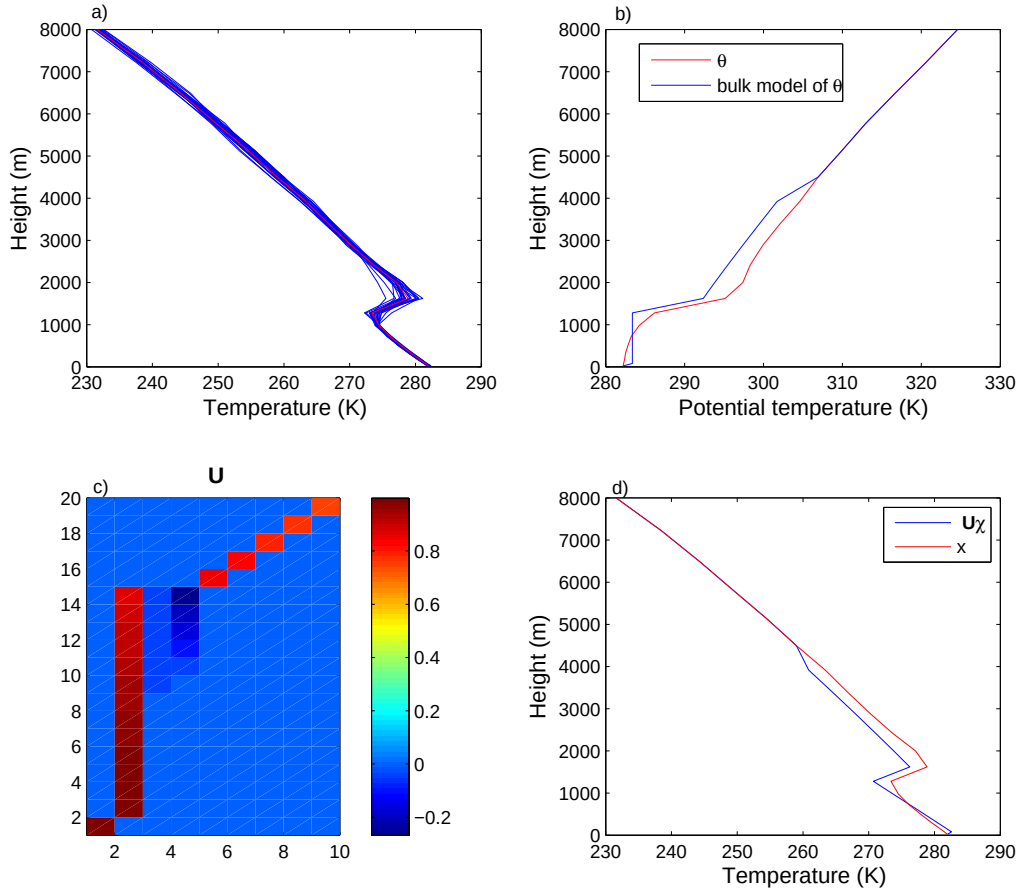


Figure 2.6: a) An ensemble of 24 temperature profiles given by MOGREPS. The red profile is the control ($\mathbf{x}$). b) The potential temperature calculated from $\mathbf{x}$ given in a) (red), and the bulk model of the potential temperature based on this, $\boldsymbol{\chi}$ (blue). c) $\mathbf{U}$ calculated for the control profile of temperature. d) The original temperature profile, $\mathbf{x}$, (red) and reconstructed from $\mathbf{U}\boldsymbol{\chi}$ (blue).

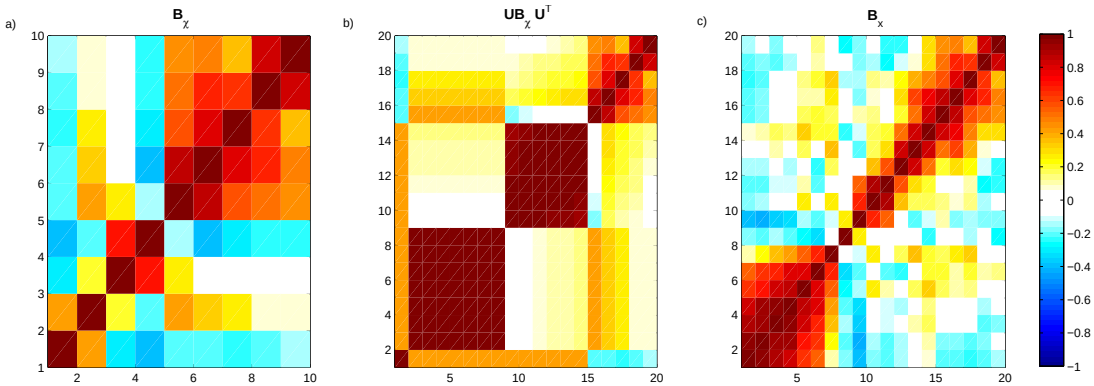


Figure 2.7: a) background error correlation matrix for $\boldsymbol{\chi}$. b) Implied background error correlation matrix for $\mathbf{x}$. c) Actual background error correlation matrix for $\mathbf{x}$ calculated from the spread in ensembles plotted in Figure 2.6a).

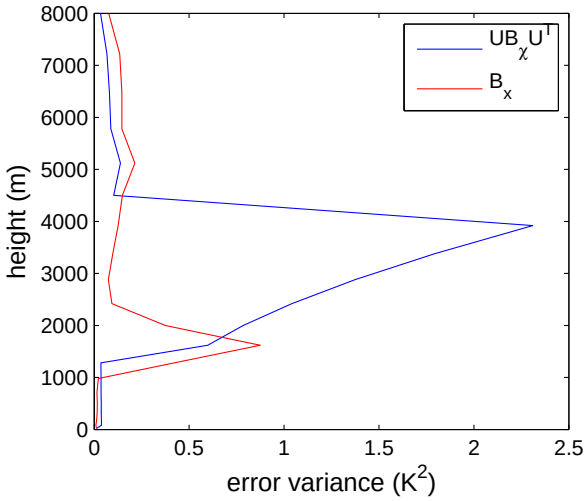


Figure 2.8: The ensemble derived (red) and implied (blue) background error variances for temperature.

In order to have a full rank estimate of $\mathbf{B}$ the number of ensembles, n, needs to at least equal the length of the state vector, N. In the examples shown in Figure 2.5, $n = 24$ and $N = 20$ and so we have confidence that the correlation structure calculated represents the true error correlations. However, in later chapters of this thesis, N is artificially increased to 100 using interpolation, in order to separate out the problems associated with resolution when studying the temperature inversion. In order to look at both temperature and humidity, N will later be increased to 200. To avoid spurious correlations, the estimate of $\mathbf{B}$ from the ensembles can be averaged spatially and temporally as well as over the ensemble members. The more $\mathbf{B}$ is averaged, the more it will begin to lose its flow-dependence and the sharp structure of the BL capping inversion will begin to be smoothed out. Hence spatial and temporal averaging will not be used in this Thesis as the error correlations caused by the BL capping inversion is the feature we are most interested in. It can be shown that this feature is less sensitive to sampling noise than smaller scale correlation features by comparing estimates of $\mathbf{B}$ for different sample sizes.

2.4 The problem in Var with assimilating the boundary layer capping inversion

Improvements in NWP forecasts have been accredited to the introduction of 3D and 4DVar at various numerical weather centres (e.g. Simmons and Hollingsworth, 2002). Var ensures that the analysis field is smooth and combines observations and the background at each model grid point by taking their weighted averages, as seen in equation (2.11). Unfortunately this method is not ideal in all situations. Problems can arise in particular when there are disagreements in the position of sharp features between the observations and background which are not described by $\mathbf{B}$ or $\mathbf{R}$. The result can be an analysis which loses the structure of the feature present, as it attempts to agree with both the observations and background.

One such feature is the BL capping inversion which is critical for its impact on the for-

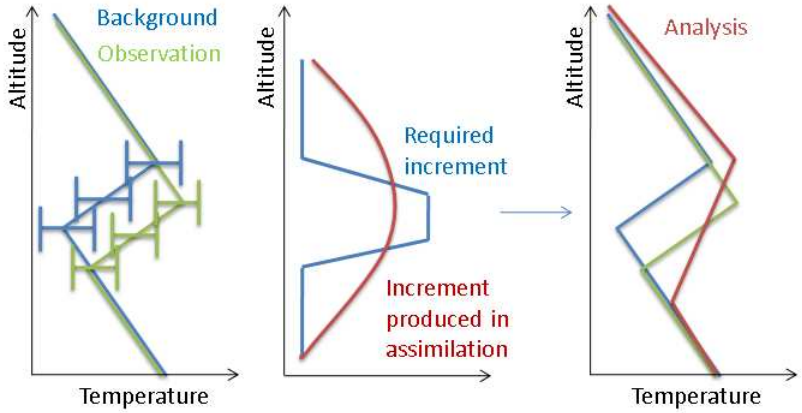


Figure 2.9: A Schematic of the the difficulty in assimilating the BL capping inversion. (i) Often there may be a disagreement in the height of the inversion in the background and the observation but otherwise the structures are similar. Due to the large gradients characterising the inversion the temperature values of the background and observation may have a large disagreement and be outside the errors diagnosed (one standard deviation of the errors is represented by bars). (ii) In standard 1DVar a discrete increment is necessary to give an analysis which agrees with the observation. This is inhibited by large correlation length scales in the B-matrix. (iii) Therefore 1DVar increases the temperature in too broad a region in an attempt to agree with the observations and smears out the structure of the inversion.

mation of Sc (stratocumulus) in the BL. The BL capping inversion is characterised by a sharp gradient in the vertical temperature profile (refer to Chapter 1). If the height of the inversion is different within the observations and background but otherwise similar in structure then there can be a large difference in the observed and background temperature in the region between the two estimates of the inversion height. It is possible that the difference will be larger than the errors given by $\mathbf{B}$ or $\mathbf{R}$ and so it is difficult to find an estimate of the true temperature profile which is consistent with both the background and observation error statistics. This is illustrated in the simple schematic of Figure 2.9(i), where the background and observations have the same inversion structure but the background places the inversion higher than the observations. This is seen as a systematic error in the background temperature values in the region between the background estimate of the inversion height and the observed inversion height.

If it is assumed that the observed inversion height is correct and there is a good agreement in the structure of the inversion between the background and observations it is desirable to end up with an analysis in close agreement to the observed profile. The analysis increment, $\mathbf{x}^{\mathrm{a}} - \mathbf{x}^{\mathrm{b}}$ (see equation (2.13)) then needs to be both large and sharp, as represented by the blue profile in Figure 2.9(ii). This increment is inhibited in Var not only by the smaller than required error variances in $\mathbf{B}$ but also by the covariances which spread information from an observation ensuring that the profile is smooth and so the resulting increment is more similar to that of the red profile in Figure 2.9(ii). Adding this to the background produces an analysis which has smeared out the inversion structure relative to both the observations and background, and in the example of Figure 2.9(iii), where the background

appears to be systematically too cool in this region, the assimilation will also spread the warming of the background profile to a region outside that which was necessary via the background error covariances. This has a big impact on the ability to diagnose realistic cloud and the poor temperature profile will then be propagated forward with the model resulting in a poor forecast. This was highlighted as one of the most important problems in NWP at the MetO at the end of 2006 when fog lead to large disruptions at Heathrow during the Christmas period (Lorenc, 2007b). In this instance radiosonde observations were not able to improve the analysis of the inversion and so the fog was not accurately forecast. This has the potential to have a large detrimental impact on the utility of the high resolution models currently being developed at the MetO.

This problem is very sensitive to the way information is spread from the observations due to the large differences between the background and observation temperatures. The flow-dependent B-matrices, as calculated in Section 2.3, decorrelate errors above and below a sharp inversion. In the schematic given in Figure 2.9 the correlation length scales were very large and so the inversion structure was smeared out, which is similar to what would be expected if a static B-matrix were used. However, if a flow-dependent B-matrix were used, the information from the observations would not spread across the inversion in the background (if the ensemble that produced the covariance matrix contained members with a strong and consistent inversion structure, like in Figure 2.5a).

One consequence of flow-dependent background errors is that the problem is sensitive to whether the background inversion is too high or too low. In Figure 2.10 a background and observation profile of temperature have been assimilated using 1DVar. The background is the control temperature profile as in CASE A of Figure 2.5 and the background error covariances are those calculated from the ensemble in CASE A. In the left hand panel the observations are synthetically produced from the control with the inversion shifted down and warmed to allow for the effects of adiabatic warming. In the right hand panel the observations are produced from the control with the inversion shifted up and cooled to allow for the effects of adiabatic cooling. In the left hand panel an analysis increment is needed to warm the background between 900m and 1400m. However the background inversion begins at 1100m and below this point the background errors are very small and so it is only above this level the background can be significantly warmed and because of the decorrelation between the background errors above and below 1100m the warming that does occur can only be spread upwards. The resulting analysis has failed to shift the inversion towards the observed height and has sharpened the inversion so that it is no longer consistent with the inversion structure seen in the observations or background. In the right hand panel the background needs to be cooled between 1100m and 1600m. This region is above the background BL and so the error bars are large enough to allow a significant amount of cooling and are free to spread this cooling vertically in both directions. The assimilation has been able to effectively shift the inversion up in the analysis. However the cooling above and below the inversion is not required and so again the structure of the inversion has been altered. The decoupling in the errors above and below the inversion in the background is not consistent with a positional error in the background inversion.

The value of the cost function at the analysis differs quite significantly between the two examples. When the background inversion is too high the value is 78.1 and when the back-

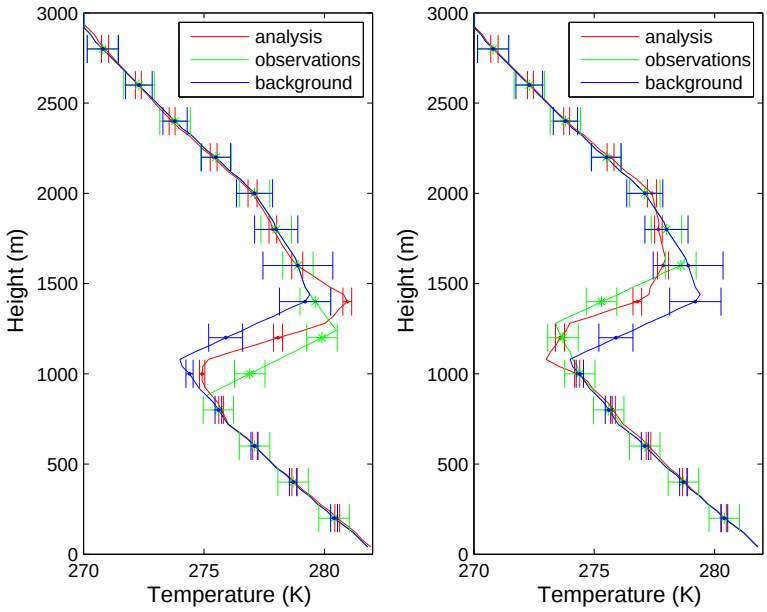


Figure 2.10: 1D variational assimilation of background temperature (blue) and observed temperature (green). A flow dependent B-matrix is used in these examples which decorrelates errors above and below the background inversion. In the left hand panel the background has the inversion 200m higher than the observations, the resulting analysis is given in red. The value of J at the minimum, in this case is 78.1 with contributions from $J^o = 115.6$ and $J^b = 40.5$. In the right hand panel the background has the inversion 200m lower than the observations, the resulting analysis is given in red. The value of J at the minimum, in this case is 27.7 with contributions from $J^o = 34.3$ and $J^b = 21.0$.

ground is too low J is 27.7. This suggests that in the second case the analysis gives a much better fit to the data given with respect to their errors. The cost function may be split into background and observation terms, $J = \frac{1}{2}(J^b + J^o)$, where $J^b = (\mathbf{x} - \mathbf{x}^{\mathrm{b}})^{\mathrm{T}}\mathbf{B}^{-1}(\mathbf{x} - \mathbf{x}^{\mathrm{b}})$ and $J^o = (\mathbf{y} - \mathcal{H}(\mathbf{x}))^{\mathrm{T}}\mathbf{R}^{-1}(\mathbf{y} - \mathcal{H}(\mathbf{x}))$ (see (2.11)). Looking at the individual components of J, J^o and J^b, we see that both values are smaller for the case when the background inversion is too low but the analysis gives a better fit to the observations in particular. However, according to the Bennett-Talagrand theorem we should expect the value of J at the analysis to be half the number of observations (see appendix A). In the examples given here P=100, and so we would expect J to equal 50 at the analysis if the assumptions made by Var were true. Therefore in both cases it can be assumed that the background and observed profiles are not consistent with the errors given and the assumption that the errors are non-biased.

The most obvious solution to this problem (and easiest to implement for this idealised case) would be to inflate the background error variances around the inversion. This has the effect of giving more weight to the observations and so the background inversion is effectively ignored. This is illustrated by Figure 2.11: the left hand panel shows the analysis which results from assimilation with an averaged B-matrix where the error variances are much smaller than the difference between the background and the observations. As expected the inversion in the analysis is weakened and inconsistent with the observation and

background errors. The right hand panel illustrates the effect of inflating the background error variances so that information about the inversion from the background has little impact on the final analysis. The analysis inversion now has the structure and position which is in good agreement with that observed. However valid information about the structure of the background inversion is also lost, and the amplitude of the observed temperatures on the model levels have no other information to correct them, and if the observations do not resolve the inversion structure then again the inversion would be weakened in the analysis.
The effect of displacement on only temperature has been discussed so far. Although the univariate problem is an important first step for study, multivariate aspects are also important. For instance it is also important that the humidity profile is altered so that cloud within the BL is consistently shifted with the change of inversion height. This problem will be dealt with in Chapter 5.

For this problem it is useful to split the background errors into *amplitude errors* and *positional errors*. In this case, the amplitude error is the small scale error associated with a model variable (e.g. temperature) on the model levels and the positional error is the large scale error in the positioning of a coherent feature (i.e the BL capping inversion). It is difficult to solve the 'displacement error' problem with just amplitude errors because position errors introduce bias due to the background mean being displaced from the truth (Ravela et al., 2007). Variational methods are generally derived from the assumption that forecast errors are strictly random and non biased. If a bias is present in the background then this will lead to a bias in the analysis (Dee and da Silva, 1998).

2.4.1 Other work that considers position error

Positional errors are a common problem for features other than the BL capping inversion such as thunderstorms, squall lines, hurricanes, precipitation, and fronts (e.g. Alexander et al., 1998; Thiebaux et al., 1990; Jones and MacPherson, 1997). Hoffman et al. (1995) demonstrated that splitting forecast errors into displacement and amplitude errors for certain features can give valuable information. This method has been used extensively for the validation of forecasts (e.g. Seo et al., 2009).
The cause of positional errors is difficult to deduce. They can arise from errors in the 'background flow, model parameters, inadequate model resolution, existence of multi-scale interactions, and approximations of governing equations and parameterisations' (Ravela et al., 2007). For errors in the forecast inversion height, inadequate model resolution is certainly a contributor but an increase in resolution would not solve this problem due to the inversion's complex relationship with turbulence in the BL, front location and subsidence.
A few attempts to include the positional error in the assimilation have been made. One such proposed technique developed by Alexander et al. (1998), for the improvement of forecasts of features associated with marine cyclones, is based on manually identifying corresponding features. A warping technique was then used to warp the entire field to match the observations. Mariano (1990) used contour analysis to calculate the average position of contours which was shown to avoid the weakening and smearing of geophys-

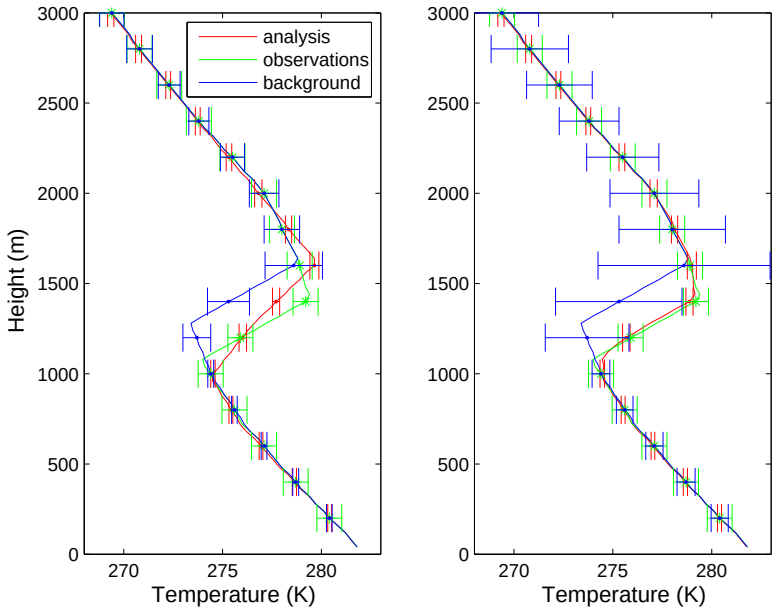


Figure 2.11: 1D variational assimilation of background temperature (blue) and observed temperature (green), where the inversion in the observed profile is 200m lower than that in the background. In the left hand panel a static B-matrix is used to calculate the analysis (red). In the right hand panel the background error variances are inflated to give the analysis shown in red. When the background errors are too small on the left hand side $J = 27.6$ at the minimum with contributions from $J^o = 34.3$ and $J^b = 20.9$. When the background errors are increased on the right hand side $J = 7.3$ at the minimum with contributions from $J^o = 3.6$ and $J^b = 11.0$.

ical features. Inclusion of position errors has also been used by Brewster (2003) for the assimilation of thunderstorms. Brewster (2003) defined a field of translation vectors in order to shift and distort the forecast field to match the observations, the aim being to minimise a displacement-based cost function and a penalty term based on the inverse of a second order autoregressive term to prevent displacing too far from the original position. Ravela et al., (2007) also developed a similar method making use of displacement vectors but without the autoregressive constraint. Because this method involves defining a field of displacement vectors for the entire field (the statistical error covariances for which are not known) it is very expensive and so the positional and amplitude adjustments were done separately. However this will not find the optimal 'balance' between the amplitude and positional error adjustment. A much less expensive one dimensional technique for dealing with positional errors associated specifically with the inversion will be developed and tested in what follows. This technique will be able to simultaneously adjust for position and amplitude errors.

2.5 Concluding remarks

The formulation of background errors is a fundamental part of Var. They should be representative of the error in the forecast in order to produce an optimal analysis.
The BL is a very dynamic region of the atmosphere and so the associated background errors are highly dependent upon the structure of the BL. The use of an ensemble of forecasts allows for a measurement of the background errors for particular atmospheric conditions. Section 2.3 showed that when a strong inversion is present the background errors are uncorrelated between the BL and FA. Within the BL where the air is well mixed there is a region of strong correlation. When no inversion is present, this feature in the B-matrix is no longer evident. It is important that these correlations are represented within the assimilation to ensure the correct spreading of information from the observations.
Variational assimilation, as it stands, does not deal well with disagreements in the position of sharp features between the background and observations. It cannot recognise features present in the background and observations which may be so simple to spot by eye. Instead, the assimilation 'sees' large amplitude errors in this region and so the assimilation of the observations can destroy the feature instead of correcting it. For the BL capping inversion this results in a weaker, deeper inversion which can have a significant (erroneous) impact on the BL cloud within a forecast. The exploration of this problem forms the majority of this thesis. The intention is to explicitly include the positional error associated with the background inversion into the assimilation scheme. Including a positional error as well as the amplitude error on each of the model levels gives a more complete description of the background errors.
In the next chapter a new scheme for including positional error in the background inversion height is formulated. In Chapter 4 this new scheme is tested for idealised temperature profiles and then a wider range of real life cases are look at using radiosonde observed temperature profiles. With the use of radiosonde data the characteristics of the positional error are also studied. This scheme will then be extended to include the assimilation of

moisture observations in Chapter 5 and the scheme's effect on the diagnosis of cloud will be looked at.

Chapter 3

The floating boundary layer scheme I: Formulation of the univariate problem

In the previous chapter we saw the importance of the background errors in the formulation of variational data assimilation. It is usual to describe these errors as amplitude errors alone, however when assimilating coherent structures such as the BL capping inversion it is advantageous to describe the positional error too. While the amplitude errors allow for the variation of the magnitude of a field on the model levels, a positional error should allow for the variation of the position of a structure within that field. The aim, still, to minimise a cost function given the background and observations.

Separating the background inversion error into a positional and amplitude error allows for the inversion structure in the background to contribute to the analysis of the inversion but also means it is possible to represent the uncertainty we have in the background inversion's height. Using this idea we aim, in this chapter, to formulate a new one dimensional variational scheme to give an improved analysis of the BL capping inversion. In this approach it is important that the error correlations associated with the inversion (as seen in 2.3) are maintained. This may be achieved in the assimilation by shifting the vertical positions of the underlying model levels themselves.

Perhaps the most general way of allowing the levels to shift in height is to assign one extra control variable per level, each describing how each level height is perturbed by the assimilation. This is similar to the approach taken by Hoffman et al. (1995) and Ravela et al. (2007). Besides dramatically increasing the cost of the assimilation, this may lead to problems such as the potential for the assimilation to allow levels to interchange or to move below the ground. In this work only one new control variable is introduced, denoted a (see Section 3.1). This single variable controls the coherent vertical shift of a range of levels in the region of the inversion in such a way that the allowed shift approaches zero away from the inversion, ensuring levels near the ground remain fixed. Moving the levels in this way helps to prevent the levels from switching places as all levels can move only in one direction. The new levels, which are able to shift in height in the minimisation of the cost function by varying the extra variable a, are referred to as *floating levels*.

This scheme will from now on be referred to as the *floating BL scheme*. The scheme that

calculates the analysis in the fixed model space with only a description of the amplitude errors, as given in Chapter 2, will be referred to as the *standard scheme*.

3.1 Description

A fundamental part of the floating boundary layer scheme is the definition of the floating levels given by their heights. The heights of the floating levels are represented by the vector $\tilde{\mathbf{z}} = (\tilde{z}_1, ..., \tilde{z}_i, ..., \tilde{z}_N)$ and are defined as

$$\tilde{z}_i = z_i + aD(z_i), \tag{3.1}$$

where z_i are the original fixed model level heights. a is varied within the assimilation in order to minimise a cost function and has units of metres. $D(z_i)$ is a prescribed function that describes which levels are allowed to move, and by relatively how much. It is called the displacement function and in this work it is chosen such that it is unity within the centre of the background inversion and tends to zero symmetrically away from the background inversion. In this way the model levels in the centre of the inversion, where $D(z) = 1$, are shifted a meters and the levels far from the inversion, where $D(z) = 0$, are not shifted at all. An illustration of the effect of a on calculating the floating model levels is given in Figure 3.1. In this illustration the fixed model levels are distributed evenly (separated by a distance d) for simplicity. The displacement function is triangular and only non-zero for the 2nd and 3rd levels (located at $z = 2d$ and $3d$)[1]. Examples of the shift in the levels are given for $a = d$ and $a = -d$. When $a = d$ levels where $D(z)$ is non-zero shift up proportionally to the value of D at the fixed model level height. When $a = -d$ the levels shift in a similar way but downwards.

The choice of σ_a^2 and D together translate into a vertical error in the position of the background levels, given by $\sigma_a D(z)$. In Figure 3.2 the displacement function is chosen to be triangular and centred on 1400m and $\sigma_a = 200$m. Therefore there is maximum positional error (represented by the vertical error bars) at 1400m, where the vertical positional error standard deviation equals σ_a (=200m). This error linearly decreases to zero, 600m either side of the centre, so that levels above 2000m and below 800m have no positional error. The choice of the function which should be used in practice to describe $D(z)$ is discussed later in Section 3.1.3.

In order to utilise a, a new cost function is written in terms of a new augmented control vector, $\tilde{\mathbf{v}}$. This cost function has the same form as the standard cost function for $\mathbf{x}$ (2.11)

$$J(\tilde{\mathbf{v}}) = \frac{1}{2}(\tilde{\mathbf{v}} - \tilde{\mathbf{v}}^{\mathrm{b}})^{\mathrm{T}}\tilde{\mathbf{B}}_{\mathrm{v}}^{-1}(\tilde{\mathbf{v}} - \tilde{\mathbf{v}}^{\mathrm{b}}) + \frac{1}{2}(\mathbf{y} - \mathcal{H}(\tilde{\mathbf{v}}))^{\mathrm{T}}\mathbf{R}^{-1}(\mathbf{y} - \mathcal{H}(\tilde{\mathbf{v}})), \tag{3.2}$$

[1]In the application of this method the number of levels associated with non-zero D is higher than that shown in the schematic 3.1.

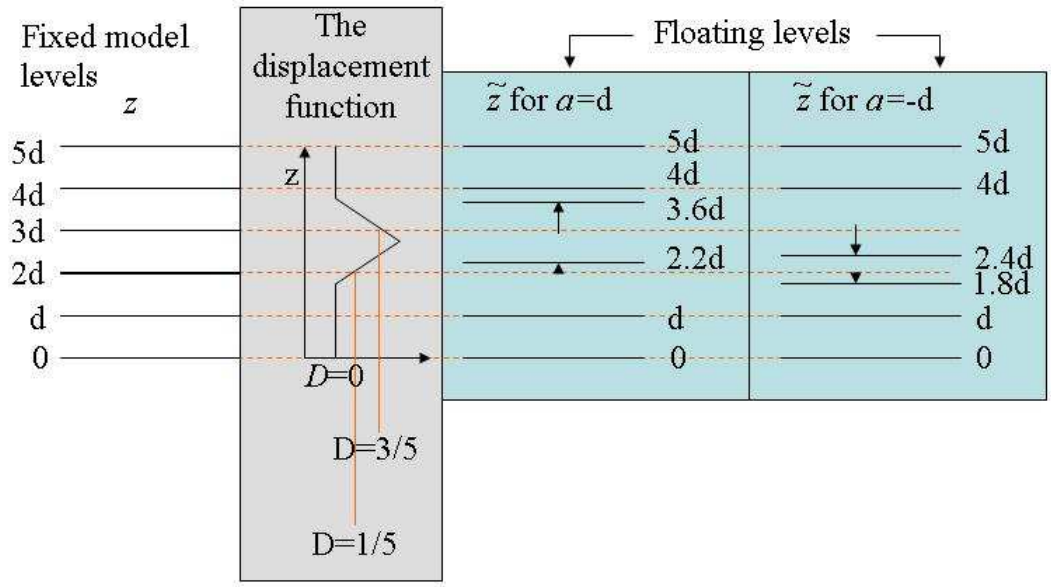


Figure 3.1: Example of how the new levels would be calculated for a triangular displacement function, when a is positive and when a is negative.

where $\tilde{\mathbf{v}}$ is the new control vector comprising the level-by-level values of e.g. temperature, $\tilde{\mathbf{x}}$ now represented on floating levels, and a as follows

$$\tilde{\mathbf{v}} = \begin{pmatrix} \tilde{\mathbf{x}} \\ a \end{pmatrix}. \tag{3.3}$$

A tilde over-bar indicates a vector that is represented on the floating model levels rather than on fixed model levels. The level indices (1,2,3,...,N) of the floating levels are the same as those of the fixed model levels, but the heights of the floating levels are allowed to vary as discussed above.

The a-priori value of $\tilde{\mathbf{v}}$ is $\tilde{\mathbf{v}}^{\mathrm{b}} = \begin{pmatrix} \tilde{\mathbf{x}}^{\mathrm{b}} \\ a = 0 \end{pmatrix}$, where $\tilde{\mathbf{x}}^{\mathrm{b}}$ is the background state, $\mathbf{x}^{\mathrm{b}}$, on floating model levels. $\tilde{\mathbf{B}}_{\mathrm{v}}$ is the background error covariance matrix for $\tilde{\mathbf{v}}^{\mathrm{b}}$ and so is also described on the floating model levels.

$$\tilde{\mathbf{B}}_{\mathrm{v}} = \begin{pmatrix} \tilde{\mathbf{B}} & 0 \\ 0 & \sigma_a^2 \end{pmatrix}, \tag{3.4}$$

where $\tilde{\mathbf{B}}$ is numerically the same B-matrix as in the standard Var scheme as described in Chapter 2, but applied to floating rather than fixed levels, and σ_a^2 is the variance of the extra variable. Note that when $a = 0$, $\tilde{\mathbf{x}}$ represents the same state as $\mathbf{x}$ and $\tilde{\mathbf{B}}$ represents the same error covariances as $\mathbf{B}$.

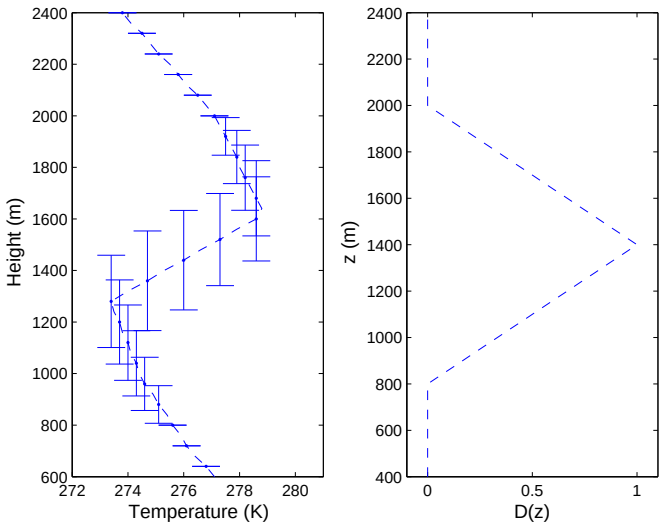


Figure 3.2: Left hand panel: Vertical error bars are given by $\sigma_a D(z)$. The choice of $D(z)$ is given in the right hand panel.

3.1.1 The observation operator

The observation operator computes the control variable's version of the observations, $\mathbf{y}^{\text{model}}$:

$$\mathbf{y}^{\text{model}} = \mathcal{H}(\tilde{\mathbf{v}}). \tag{3.5}$$

Radiosonde observations of temperature (at fixed heights) are considered here and so $\mathcal{H}$ consists of interpolation from floating model levels to the fixed observation levels. The floating model level heights are dependent upon the value of a, as shown in (3.1) and so the observation operator for the floating BL scheme is now non-linear (see below). This means that the cost function is not quadratic and so its minimum must be found iteratively. Under each iteration the observation operator is expanded as a Taylor series to linear order $\mathcal{H}(\tilde{\mathbf{v}}_n) \approx \mathcal{H}(\tilde{\mathbf{v}}_{n-1}) + \mathbf{H}_n \delta\tilde{\mathbf{v}}_n$, where subscript n indicates the iteration index.

The value of the j^{th} element of $\mathcal{H}$ acting on $\tilde{\mathbf{v}}_n$ can be expressed via linear interpolation as a weighted combination of the values of $\tilde{\mathbf{x}}_n$ either side of the j^{th} observation level.

$$\mathcal{H}(\tilde{\mathbf{v}}_n)^j = \left(\frac{\tilde{z}_n^{\text{II}} - z_o^j}{\tilde{z}_n^{\text{II}} - \tilde{z}_n^{\text{I}}}\right) \tilde{x}_n^{\text{I}} + \left(\frac{z_o^j - \tilde{z}_n^{\text{I}}}{\tilde{z}_n^{\text{II}} - \tilde{z}_n^{\text{I}}}\right) \tilde{x}_n^{\text{II}}, \tag{3.6}$$

where z_o^j is the height of the j^{th} observation, and $\tilde{x}_n^{\text{I}}$ $\tilde{x}_n^{\text{II}}$ are the values of $\tilde{\mathbf{x}}_n$ either side of the j^{th} observation level and $\tilde{z}_n^{\text{I}}$ and $\tilde{z}_n^{\text{II}}$ are their level heights respectively. As a is varied the floating heights $\tilde{z}_n^{\text{I}}$ and $\tilde{z}_n^{\text{II}}$ used to calculate $\mathcal{H}(\tilde{\mathbf{v}}_n)^j$ will change and so the levels needed may be different. This is illustrated in Figure 3.3 for the case when the assimilation has decided to move a floating level across an observation location between two iterations. In the $n+1^{th}$ iteration the level used to define $\tilde{x}_{n+1}^{\text{II}}$ was previously used to define $\tilde{x}_n^{\text{I}}$.

Even though linear interpolation is used in (3.6), $\mathcal{H}(\tilde{\mathbf{v}})$ is non-linear in components of $\tilde{\mathbf{v}}$ due to the presence of a. The linearised observation operator, $\mathbf{H}_n$, can be calculated for

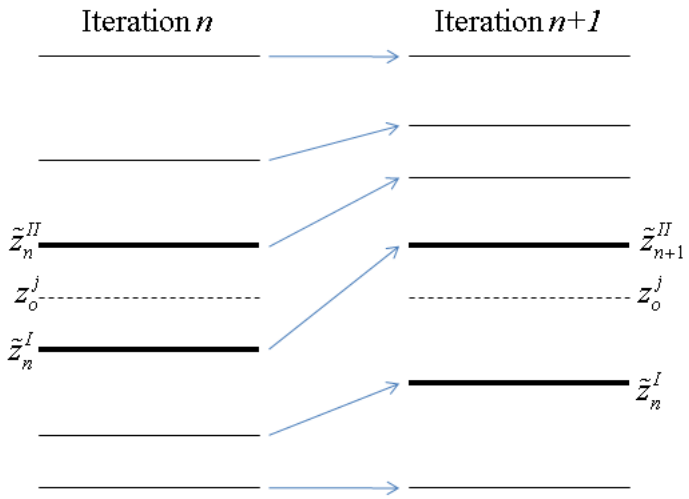


Figure 3.3: Example of how the floating model levels can change between iterations respective to the observation level heights. The two floating levels that are used in the interpolation for the location of one observation (dashed) are shown in bold.

each iteration as,

$$\mathbf{H}_n = \frac{\partial \mathcal{H}(\tilde{\mathbf{v}})}{\partial \tilde{\mathbf{v}}}\Big|_{\tilde{\mathbf{v}}=\tilde{\mathbf{v}}_n}. \tag{3.7}$$

In the example given in Figure 3.3, there is one observation and so $\mathbf{y}$ and $\mathbf{z}_o$ are scalars. $\tilde{\mathbf{x}}$ is described on 6 levels and so the length of $\tilde{\mathbf{v}}$ is 7. This means that $\mathbf{H}$ is a 1×7 matrix. For this example $\mathbf{H}_n$ and $\mathbf{H}_{n+1}$ can be written as:

$$\mathbf{H}_n = \begin{pmatrix} 0 & 0 & 1/2 & 1/2 & 0 & 0 & \frac{\partial \mathcal{H}(\tilde{\mathbf{v}})}{\partial a}\Big|_{\tilde{\mathbf{v}}=\tilde{\mathbf{v}}_n} \end{pmatrix}. \tag{3.8}$$

$$\mathbf{H}_{n+1} = \begin{pmatrix} 0 & 1/3 & 2/3 & 0 & 0 & 0 & \frac{\partial \mathcal{H}(\tilde{\mathbf{v}})}{\partial a}\Big|_{\tilde{\mathbf{v}}=\tilde{\mathbf{v}}_{n+1}} \end{pmatrix}. \tag{3.9}$$

The evaluation of $\frac{\partial \mathcal{H}(\tilde{\mathbf{v}})}{\partial a}\Big|_{\tilde{\mathbf{v}}=\tilde{\mathbf{v}}_n}$ is dependent upon $\tilde{\mathbf{v}}_n$ and the choice of displacement function (see Appendix C). In general P observations will be present and there will be N floating levels. The $P \times (N+1)$ $\mathbf{H}_n$ matrix will have rows of a similar structure to (3.8) and (3.9). The non-linearity of the observation operator is discussed further in Section 4.2.

3.1.2 Calculating the analysis with the floating BL scheme

It has already been noted that the floating BL scheme must be solved iteratively as in (2.18). For the floating BL scheme this is equivalent to calculating

$$\delta\tilde{\mathbf{v}}_n = (\tilde{\mathbf{B}}_{\mathrm{v}}^{-1} + \mathbf{H}_n^{\mathrm{T}}\mathbf{R}^{-1}\mathbf{H}_n)^{-1}\left(\tilde{\mathbf{B}}_{\mathrm{v}}^{-1}(\delta\tilde{\mathbf{v}}_n^{\mathrm{b}}) + \mathbf{H}_n^{\mathrm{T}}\mathbf{R}^{-1}(\delta\mathbf{y}_n)\right), \tag{3.10}$$

where $\delta\tilde{\mathbf{v}}_n = \tilde{\mathbf{v}}_{n+1} - \tilde{\mathbf{v}}_n$, $\delta\tilde{\mathbf{v}}_n^{\mathrm{b}} = \tilde{\mathbf{v}}_n - \tilde{\mathbf{v}}^{\mathrm{b}}$ and $\delta\mathbf{y}_n = \mathbf{y} - \mathcal{H}(\tilde{\mathbf{v}}_n)$. At each iteration, n, the updated estimate of $\tilde{\mathbf{v}}$ is $\tilde{\mathbf{v}}_{n+1} = \tilde{\mathbf{v}}_n + \delta\tilde{\mathbf{v}}_n$. The iterations stop when the relative change in the norm of the control vector from one iteration to the next is appropriately small. Convergence is assumed when $c = \left|\frac{1}{N+1}\sum_{i=1}^{N+1}\frac{\tilde{v}_n^i - \tilde{v}_{n-1}^i}{\tilde{v}_n^i}\right|$ is less than 0.01, however in practice c is often less than 10^{-6} when only temperature is assimilated and less than 10^{-5} when temperature and humidity are assimilated in Chapter 5. Once the iterations

have stopped the analysis is given by $\tilde{\mathbf{v}}_{n+1} = \begin{pmatrix} \tilde{\mathbf{x}}_{n+1} \\ a_{n+1} \end{pmatrix}$, $\tilde{\mathbf{x}}_{n+1}$ may then be interpolated back to the original model levels if necessary, however this will result in an interpolation error. An algorithm for the floating BL scheme can be found in Appendix B along with the Fortran code for the main program[2].

3.1.3 Choosing the displacement function

The choice of $D(z)$ is important if the scheme is to successfully maintain the background inversion structure. Before we examine the workings of the floating BL scheme in detail, we consider the effect of the specification of the displacement function on some trial data. A more detailed study of the mechanics of the floating BL scheme and how it compares to the standard scheme is left until Section 3.2 and Chapter 4. For now we assume that we have a working floating BL scheme, a given background and set of observations and appropriately specified error covariance matrices $\tilde{\mathbf{B}}_{\mathrm{v}}$ and $\mathbf{R}$.

As already stated it is intended that $D(\mathbf{z})$ be centred on the inversion and symmetrical about this centre. The choice of how wide the function is dictates how many levels should be allowed to move. It is only the inversion that we wish to shift, however, in order to ensure that this doesn't cause any discontinuities in our analysis temperature profile, the model levels describing the region between the inversion in the background and the inversion in the observations must also be allowed to move so that we do not end up with levels switching their order. In other words the width of $D(\mathbf{z})$ must be at least twice the final shift which is equal to the final value of a. It is unknown prior to the assimilation what this value will be but we do know σ_a.

In Figures 3.4 to 3.7 the effect of the choice of the displacement function on the final analysis is illustrated. In these examples the background (blue) has the inversion 200m too high compared to the observations (green) as seen in the left hand panels. In the right hand panels, the error in the model level heights, given by $\sigma_a D(\mathbf{z})$, is shown for the background (blue dashed) and the analysis (red). The blue dashed line also allows us to see the shape of $D(\mathbf{z})$ as σ_a is fixed.

In Figure 3.4 $D(z)$ has the value of 1 throughout the inversion so that its structure is maintained during the shift and then $D(z)$ slowly tends to zero away from the inversion allowing for a smooth analysis (red). In this example the displacement function is given by

$$D(z) = \begin{cases} \frac{z-950}{250} & \text{for } 950 < z \leq 1200 \\ 1 & \text{for } 1200 < z < 1600 \\ 1 - \frac{z-1600}{250} & \text{for } 1600 \leq z < 1850 \\ 0 & \text{elsewhere} \end{cases} \tag{3.11}$$

It is seen that the analysis has an accurate and realistic temperature profile with the inversion structure retained and its height in good agreement with that observed ($a = -187$m). This appears to be close to the optimum displacement function.

The floating scheme may cause unrealistic features in the temperature profile if the displacement function is not uniform across the inversion, such as the triangular function in

[2]In Appendix **B**, **U** is a transform (which will be used later in Chapter 5). For now **U** is the identity and $\tilde{\mathbf{v}} = \tilde{\mathbf{w}}$.

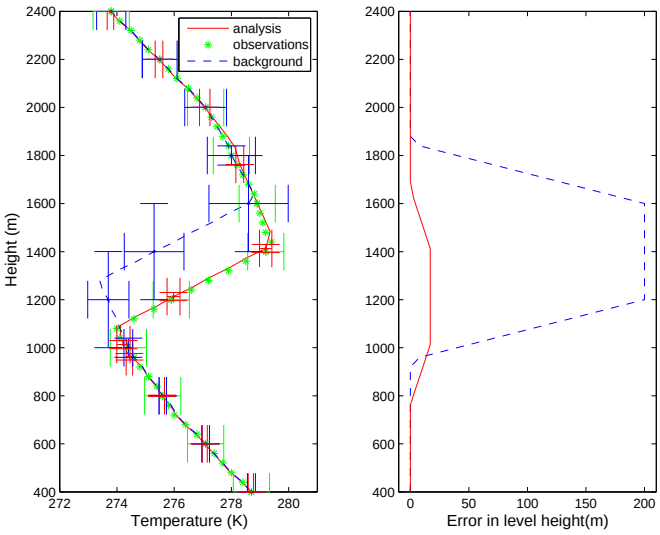


Figure 3.4: The analysis (red profile in left hand panel) given by the floating BL scheme when the displacement function is unity across the inversion depth and tends to zero away from this. a=-187m, J_{min}=11.2. Error bars are given for every 5th level. The right hand panel gives $\sigma_a D(\mathbf{z})$, which represents the error in the model levels, for the analysis (red) and background (blue). It is assumed there is no error in the height of the observation levels.

Figure 3.5. In this example the displacement function is given by

$$D(z) = \begin{cases} \frac{z-950}{450} & \text{for } 950 < z \leq 1400 \\ 1 - \frac{z-1400}{450} & \text{for } 1400 \leq z < 1850 \\ 0 & \text{elsewhere} \end{cases} \tag{3.12}$$

In this case the levels interchange position below the maximum of D.
If the displacement function is too narrow such as in Figure 3.6, the analysis is smoother but the floating BL scheme is less effective as a is only -74m, less than half the displacement needed, and the structure of the background inversion cannot be kept. In this example the displacement function is given by

$$D(z) = \begin{cases} \frac{z-1350}{25} & \text{for } 1350 < z \leq 1375 \\ 1 & \text{for } 1375 < z < 1450 \\ 1 - \frac{z-1425}{25} & \text{for } 1425 \leq z < 1450 \\ 0 & \text{elsewhere} \end{cases} \tag{3.13}$$

If the displacement function does not tend gradually to zero away from the inversion, as in Figure 3.7 where a uniform function is used, discontinuities are caused between the background and observed inversion as it is only the levels describing the background inversion that are moved and so again we get an interchanging of levels. In this example

the displacement function is given by

$$D(z) = \begin{cases} 1 & \text{for } 1200 < z < 1600 \\ 0 & \text{elsewhere} \end{cases} \tag{3.14}$$

The effectiveness of the floating BL scheme can also be jeopardised if the displacement function is not centred at the correct height (Figure 3.8). In this example the displacement function is the same as the optimal one given in Figure 3.4 but is now centred at the top of the inversion instead of at its centre. The assimilation now reverts back to the standard scheme (a=0) as a shift of these levels where $D(z)$ is non zero does not minimise J.
In summary, in order for the floating BL scheme to have the greatest positive effect, the displacement function must:

- Be centred on the background inversion, so that the correct levels are shifted and the structure of the inversion is kept.
- Gradually tend to zero away from the background inversion to reduce the chance of levels interchanging position.
- Be broad enough, again to ensure that levels do not interchange position. This would occur if levels are shifted into a region where there is no movement in the model levels (i.e. where $D(z)$ is zero). If $D(z)$ is too broad, though, then too large a region would be shifted and so it would not only be the inversion structure that would be shifted when attempting to minimise J, and so the scheme would not be as effective. There would also be an increased danger that levels could be shifted below the surface.

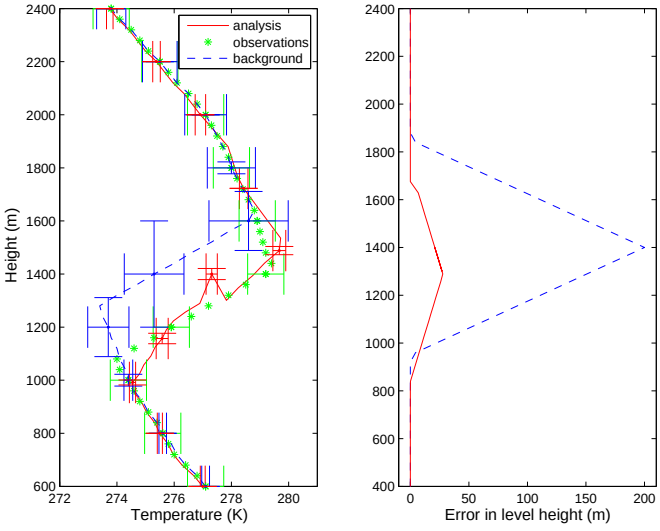


Figure 3.5: As in Figure 3.4 but with triangular displacement function; a=-112m, J_{min}=36.

If there is no inversion in the background, $D(z)$ should equal zero for all heights. This means that the position error is then zero and so the scheme can revert back to the standard scheme where only amplitude errors are taken into account.

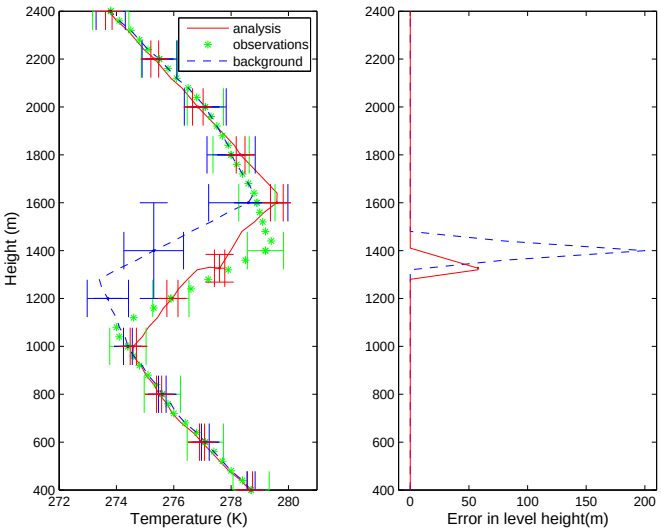


Figure 3.6: As in Figure 3.4 but now the displacement function is too narrow; a=-74m, J_{min}=25.

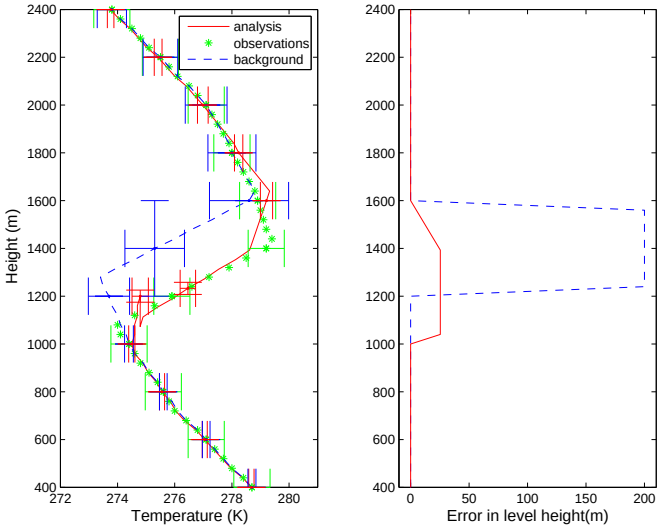


Figure 3.7: As in Figure 3.4 but with uniform displacement function; a=-167m, J_{min}=22.

3.1.4 Choice of variable for displacement

The choice of variable for which the inversion is shifted may also be important for the utility of this proposed scheme. It may be important that this scheme does not alter the stability of the BL, and so any variable minimised on the floating levels should be conserved when shifted vertically. In particular it is reasonable to consider that the adiabatic effects of varying ambient pressure on each level needs to be taken into account as the layers change height (and hence change pressure). In Section 4.3, when real data is studied, it is seen that performing the assimilation in temperature means that as the inversion is shifted up/down it also needs to be cooled/warmed. Potential temperature, θ, is therefore

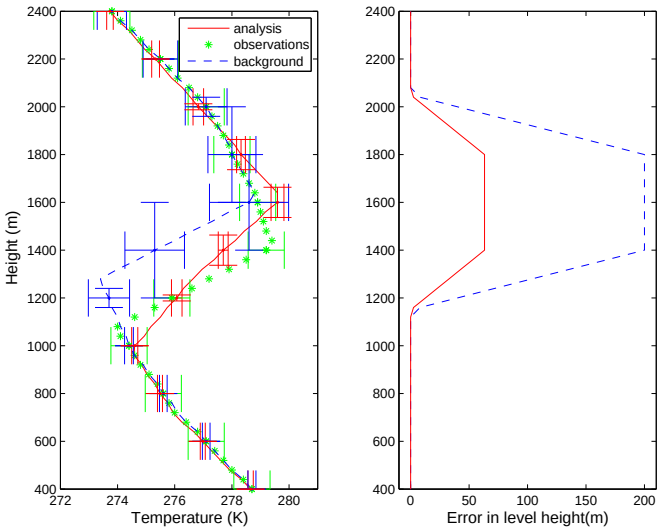


Figure 3.8: As in Figure 3.4 but the displacement function is now centred on the wrong location; a=0m, J_{min}=33.

a much better choice of control variable. Potential temperature is insensitive to changes in pressure, p, and so is conserved adiabatically as the model levels in the vicinity of the inversion are shifted vertically.

$$\theta = T\left(\frac{p_o}{p}\right)^{R/c_p}, \tag{3.15}$$

where T is temperature, p_o is the reference pressure, p is pressure, R is the specific gas constant and c_p is the isobaric specific heat capacity.

In Figure 3.9 the assimilation is repeated for Figure 3.4, where the displacement function was near optimal, however this time potential temperature is used to minimise the cost function (3.2) instead of temperature. In order to achieve this the background, observations, and their errors have been converted to potential temperature. The value of a at the end of the assimilation is -172m compared to -187m when the assimilation is performed for temperature. The error in the analysis (calculated as $\mathbf{P}^{\mathrm{a}} = (\mathbf{I} - \mathbf{KH})\mathbf{B}$) of a is reduced marginally for potential temperature than for temperature; for potential temperature the standard deviation of a is 14.0m compared 17.0m for temperature. The amplitude error for potential temperature is greater in the region of the inversion as seen in Figure 3.10 due to the level by level values of potential temperature being varied more in order to fit to the observations. Generally there is very little difference between the two analyses.

3.2 Illustration

We can demonstrate how the floating BL scheme would work for an idealised, statistically consistent example. Starting with a profile taken to be the truth, observation and background profiles can be simulated which are consistent with their specified errors. For the background this includes a vertical shift corresponding to the error for a. These profiles can then be used to test the floating BL scheme. In this example both the observation and background measure potential temperature at level spacing of 40m within the lowest

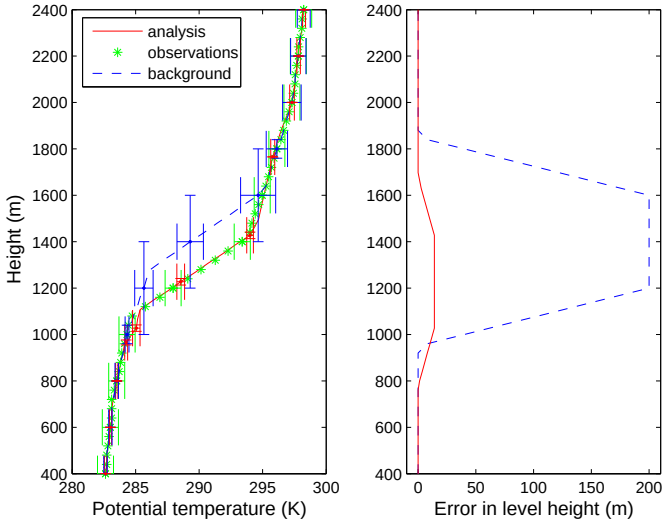


Figure 3.9: As in Figure 3.4 but assimilation performed in terms of potential temperature instead of temperature; a=-172m, J_{min}=18.

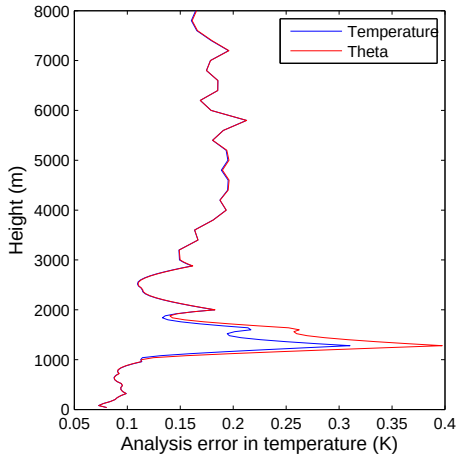


Figure 3.10: The standard deviation of the analysis error for temperature (blue) and potential temperature (red) versus height.

3km of the atmosphere, and 200m between 3km and 8km altitude, giving a total of 100 levels.

In this work it is useful to have an estimate of the flow dependent B-matrix. $\tilde{\mathbf{B}}$ can be calculated from the spread in an ensemble of forecasts (see Chapter 2). The control member of the ensemble or the mean of the ensemble represents the state for which $\tilde{\mathbf{B}}$ is valid, depending on which has been used to estimate the truth. In this illustration the ensemble is the same as that used in CASE A of Figure 2.5. From the control member, profiles representing the truth, observations and background can be simulated in the following way:

- **Truth:** The truth, $\mathbf{x}^{\mathrm{t}}$, is constructed by taking the control member plus a shift which is consistent with the standard deviation of a and the displacement function (see below). The control member is presented in terms of temperature and is

firstly converted from temperature to potential temperature to allow for the effects of adiabatic cooling. A new set of heights are assigned to each level,

$$\tilde{\mathbf{z}} = \mathbf{z} + \sigma_a e_z D(\mathbf{z}), \tag{3.16}$$

where e_z is a random number taken from a normal distribution with mean 0 and standard deviation 1, $N(0,1)$. σ_a at this point has been chosen as 200m (the error statistics of a based on real data will be discussed in Chapter 4). The displacement function is chosen to be uniform across the depth of the inversion in the control member, and to linearly tend to zero 600m either side of the inversion. The shifted profile, given in potential temperature, described on these levels is then interpolated back to the original fixed levels to give $\mathbf{x}^t$.

- **Observations:** The imposed observation errors are not correlated in the vertical and so can be described solely by their variances (assuming their errors are Gaussian and unbiased). Observations can then be simulated from the truth as follows

$$\mathbf{y} = \mathcal{H}(\mathbf{x}^t) + \mathbf{R}^{1/2}\mathbf{e}_y, \tag{3.17}$$

where $\mathbf{e}_y$ is an observation space vector of random numbers, each generated from the Gaussian distribution with mean zero and unit variance, $N(0,1)$. In this case $\mathcal{H}(\mathbf{x}) = \mathbf{x}$ i.e. observations of potential temperature are given for each fixed level.

- **Background:** The background error covariances, $\tilde{\mathbf{B}}_v$, include covariances between potential temperature variables at different levels but not between the errors for a and potential temperature, as seen in (3.4). The background state is found from the truth plus an amplitude error, plus a vertical shift. Since the control member and truth already differ by a shift, $\mathbf{x}^b$ can be found directly from the control member as follows

$$\mathbf{x}^b = \mathbf{x}^{\text{control}} + \mathbf{L}\boldsymbol{\Lambda}^{\frac{1}{2}}\mathbf{e}_x, \tag{3.18}$$

where $\mathbf{L}$ is a matrix of eigenvectors describing $\tilde{\mathbf{B}}$, $\boldsymbol{\Lambda}$ is a diagonal matrix of their corresponding eigenvalues and $\mathbf{e}_x$ is a model space vector of random numbers, each taken from $N(0,1)$. In this way, $\mathbf{B} = \mathbf{L}\boldsymbol{\Lambda}^{\frac{1}{2}}\boldsymbol{\Lambda}^{\frac{1}{2}}\mathbf{L}^T$ represents a decomposition of $\mathbf{B}$ and it can be shown that $\mathbf{L}\boldsymbol{\Lambda}^{\frac{1}{2}}\mathbf{e}_x$ is normally distributed with covariance given by $\mathbf{B}$. There are infinitely many ways a positive definite matrix may be decomposed (see Lewis et al. (2006) for more examples), this method has been chosen for convenience.

An example of a background and an observation profile generated in this way, converted back to temperature, can be seen in Figure 3.11. The standard deviation of the error in the observations has been taken to be the level-by level average of the background standard deviations, and so does not vary with height (one error standard deviation is represented by the horizontal bars). The truth is shown in black. The background (blue) has the inversion height 197m too high. The observations are given by the green stars. Due to the displacement of the background inversion, the background values appear systematically too cool in this region.

The red profiles in Figure 3.11 show the analysis (also converted from potential temper-

ature to temperature) calculated when only amplitude errors are taken into account (the standard scheme) for the case of a static B-matrix, where the errors have been averaged temporally and spatially (left hand panel) and a flow-dependent B-matrix, where the errors are dependent on the background state (right hand panel). The B-matrix used affects the way information from an observation is spread. The static B-matrix is found to have large correlation length scales and so the analysis calculated is too warm above and below the inversion. The flow-dependent B-matrix has shorter length-scales. When the flow dependent B-matrix is used, the information from observations do not spread across the inversion in the background and so the analysis is too warm above the inversion but not below. In both cases the analysis has lost some of the sharp and coherent structure of the inversion which is present in both the background and the observations. These analyses could alter the BL dynamics in a subsequent forecast initialised from either of these analyses, and could have an erroneous effect on the diagnosis of cloud, for example.
The background errors are not consistent for this example because the positional error is not included. If the positional error associated with the inversion were present within the ensemble of forecasts by the range of possible shifts sampled by the ensemble then the amplitude error would be effectively increased in the region, and the amplitude errors would be more consistent (Fowler et al., 2010). However this would be interpreted in the assimilation as having little knowledge about the inversion from the forecast (seen in Section 2.4, Figure 2.11). In conjunction with this the height of the decoupling between errors in the BL and above would be smeared as there would be no consensus from the ensemble as to which height the inversion should be. We would also have to rely on the ensemble being able to capture the uncertainty in the inversion which due to under sampling may not possible. These points will be illustrated later in Section 4.1 when the floating BL scheme is compared to the standard assimilation scheme using more consistent background amplitude error statistics when a positional error is present.

We can now recalculate the analysis by introducing the positional error in the inversion via the floating BL scheme. The results are shown in Figures 3.12 and 3.13. When a flow-dependent B-matrix is used in Figure 3.12, a has been minimised to give a final value of -193m causing the background inversion structure to move down in order to agree with the observed structure. Similarly, when a static B-matrix is used in Figure 3.13, a has been minimised to give a final value of -185m. In this case there is very little difference between using a static and flow dependent B-matrix as it is not necessary to produce large amplitude increments which are then spread vertically.
The displacement function is the same as that used to generate the truth. The movement of the levels, and the positional errors associated with the background and analysis (calculated from the inverse of the Hessian) levels are shown in right hand panels of Figures 3.12 and 3.13. We see that we now have much greater trust in the position of the inversion in the analysis than in the background, due to a smaller analysis error. The presence of the vertical error bars in this scheme, which are associated with the background state, allows the background and observation profiles to be consistent, despite the height shift. This is a novel aspect of this work. This consistency is reflected in the value of J at the analysis, which is 48.3 when the floating BL scheme is used with a flow-dependent B-matrix and

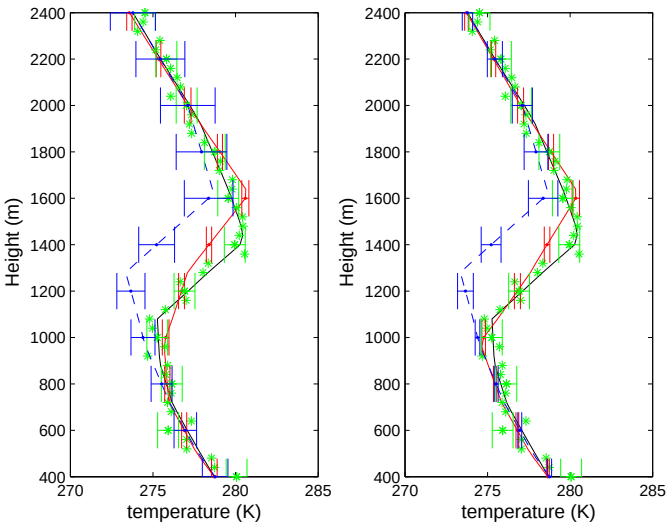


Figure 3.11: Assimilation results using the standard scheme when there is a disagreement between the background (blue) and the observed (green) inversion height. The true temperature profile is shown in black. The resulting analyses are shown in red. In the left hand panel a static B-matrix has been used $J_{min} = 74.4$, $J_o = 113.6$, $J_b = 35.1$. In the right hand panel a flow-dependent B-matrix has been used $J_{min} = 73.1$, $J_o = 101.1$, $J_b = 45.2$. Error bars are shown for every 5th level.

73.1 for the standard scheme. As we have 100 observations, 48.3 is more consistent with that expected from the Bennett-Talagrand theorem (see appendix A).

In this chapter we have devised a new scheme to work within the variational framework, which explicitly accounts for positional error. This new scheme has introduced an extra variable, a, into the control vector to represent the uncertainty we have in the inversion height within the background. As a is varied, in the minimisation of the cost function, it distorts the background space so that the background inversion is shifted towards the observed inversion. This distortion is modulated by a displacement vector which is chosen subjectively. It has been shown that the effectiveness of the floating BL scheme is sensitive to the choice of this displacement function.

In the next chapter we shall look deeper at how the floating BL scheme alters the assimilation. We shall look at the effect of the non-linear observation operator, the shape of the cost function, and the effect the scheme has on the floating background error covariance structure projected back on to the fixed model levels. Finally we shall look at some real radiosonde and model data to investigate the characteristics of these positional errors associated with the inversion.

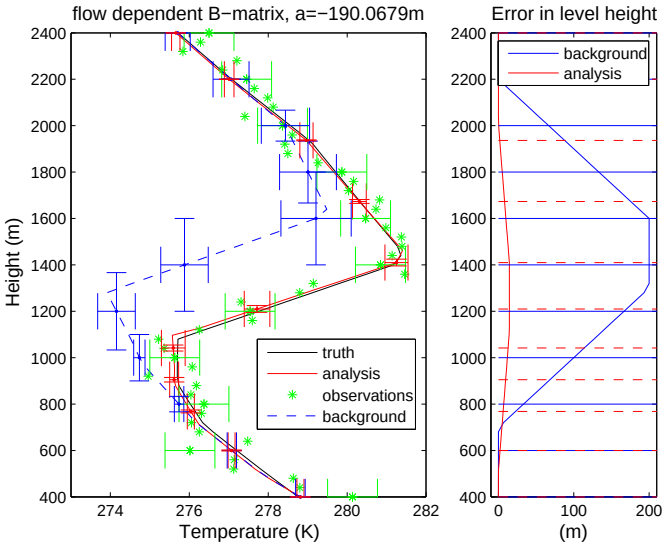


Figure 3.12: Left hand panel: Analysis (red) recalculated for the floating BL scheme using a flow dependent B-matrix. The truth (black), observations (green) and background (blue) are the same as in Figure 3.11. This scheme has minimised the cost function by moving the background inversion down 193m as well as varying the values of temperature on each model level. J_{min}=48.3, $J_o = 75.2$, $J_b = 21.8$. The movement of the levels is seen in the right hand panel. The fixed model levels are represented by the blue horizontal lines and the final floating positions as given by a in the analysis are represented by the red dashed horizontal lines. Also plotted in the right hand panel are the positional errors as calculated by $a\sigma_a$ for the background (blue) and analysis (red).

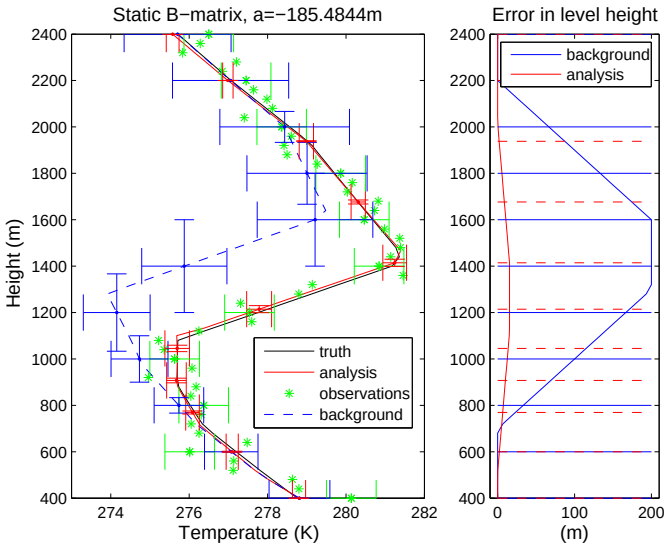


Figure 3.13: As in Figure 3.12 but with a static B-matrix.a=-185m, J_{min}=87.9, $J_o = 81.9$, $J_b = 93.8$.

Chapter 4

The floating boundary layer scheme II: Results for the univariate problem

In the previous chapter a one dimensional variational scheme was formulated capable of explicitly representing the positional error associated with the BL capping inversion in the background. We shall now study the impact that the inclusion of the positional error has on the assimilation. In the first part of this chapter comparisons will be made between the floating BL scheme -which has correct amplitude and positional background error statistics- and the standard scheme -which has been arranged to have correct amplitude error statistics consistent with a positional error. This will highlight the importance of the explicit inclusion of the positional error. In the second part of this chapter we shall look at the effect of the non-linear observation operator. A non-linear observation operator can have a big impact on the necessary observation error, the accuracy of the estimate of the analysis error covariances, given by (2.22), and the shape of the cost function. Finally we shall see what this positional error looks like in reality using a large sample of radiosonde profiles to estimate the statistics of a.

4.1 Comparisons to the standard scheme

By introducing a into the control vector (see (3.3)), the floating BL scheme allows an extra degree of freedom into the assimilation. In order to give a fair comparison with the standard scheme we need to be able to allow for the existence of positional errors within the standard assimilation. This has been done in two ways; both involving an alteration of the B-matrix description of the amplitude errors alone. In the first method we artificially create an ensemble of forecasts to have a spread in the inversion height consistent with that given by the error variance for a. This ensemble is then used to create a new B-matrix using (2.33). In the second method we use the B-matrix implied from the floating BL scheme using the final value of a for a given assimilation. This allows us to look at the equivalent amplitude errors generated by the floating BL scheme for a specific misplacement of the background inversion. Each method is explained in more detail below.

4.1.1 Calculating an amplitude B-matrix for the standard scheme using an ensemble with positional error

An ensemble of forecasts that represents accurately the variability due to both the amplitude and positional error may not necessarily be seen in practise. This may be due to under sampling or a positional bias in each forecast which may be the same for all members. It is therefore expected that the variability of the inversion height within an ensemble would not necessarily accurately represent the positional error.
However it is possible to generate an ensemble of forecasts to have a positional error (and amplitude error) consistent with that allowed for in the floating BL scheme. From this ensemble we can then calculate a new B-matrix which has error statistics consistent with the background state used in Chapter 3 (whose BL height is known to be in error).
In the left hand panel of Figure 4.1 the same ensemble is used as CASE A in Section 2.3, but modified to have a spread of Gaussian positional errors with a standard deviation of 200m (this is the standard deviation used for a in Chapter 3). The shift of the inversion in each ensemble member has been performed in the same way as that achieved by the floating BL scheme; using a displacement function which has a value of 1 in the region of the inversion to shift the model levels and their potential temperature values, interpolating back to the original grid to give the final profile. The correlation matrix calculated from this new ensemble is plotted in the middle panel of Figure 4.1. By using the B-matrix found from this procedure, a positional error in the background inversion will be interpreted in a similar way to a weak inversion when studying the amplitude error in this way, and we see that the decoupling of errors above and below the inversion is weakened due to the uncertainty as to which level the decoupling should happen. Figure 4.1 may be compared to the correlation matrices in Figure 2.5 for CASE A, without the positional modification and for CASE B, a weak inversion. We also see an increase in correlation length scale in the region of inversions between 900m and 2200m. Plotted in the right hand panel of Figure 4.1 is the level-by-level standard deviations of T, with (blue) and without (red) the height modification. We can see that the positional error results in an inflation of the errors in the region of the background inversion.

Using this B-matrix to assimilate the background and observation profiles in Figure 3.12 (Section 3.2) using the standard scheme gives the analyses seen in Figure 4.2. The inflation of the background errors means that little trust is given to the background in the region of the background inversion. This means that the analysis now gives more weight to the observations than was seen previously when the ensemble did not represent the positional error (right hand panel of Figure 3.11, Section 3.2). The analysis has been calculated for observations made at every model level (left), every 5th model level (middle) and every 10th model level (right). In each case the analysis now has near complete agreement with the observations in this region due to large error associated with the background. However, if the observations are unable to completely resolve the true inversion then the inversion in the analysis is again smoothed, this can be seen in particular in the right hand case. So even though the analysis is now in good agreement with the observations it cannot necessarily represent the true temperature profile. The smoothness of the profile is due to the error correlations in $\mathbf{B}$. If were it not for these correlations the analysis would appear far less realistic.

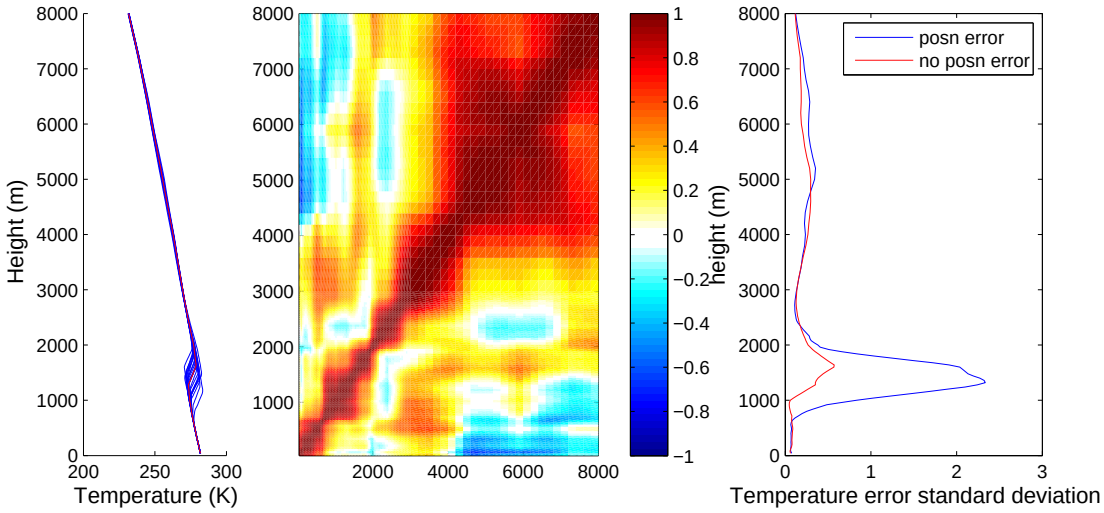


Figure 4.1: Left: An ensemble of forecasts have been generated with magnitude error given by the B-matrix in CASE A (Section 2.3) and a positional error assumed to be Gaussian with a standard deviation equal to 200m, consistent with that calculated in Section 4.3. Middle: Background error correlation matrix calculated from the ensemble generated in the left hand panel. The axes give height (m). Right: comparison of error standard deviation when a positional error is (red) and is not (blue) present in the ensemble.

Although this method certainly gives an improved analysis it does not solve the problem. In practise it may be difficult to obtain an ensemble which describes the positional error accurately. Furthermore it is dependent on dense observations giving an accurate description of both the height and structure of the inversion. As discussed in Section 2.4, inflating the background errors gives too little weight to the background in this region, which loses valuable information about the inversion structure which actually may be subject to a comparatively small amplitude error.

Also plotted in dashed magenta in Figure 4.2 is the analysis that would be given from the floating BL scheme for temperature when only these observations are available. As expected, the analysis is able to keep the inversion structure and shift its height in accordance with the observations given, even if the observations do not completely resolve the true inversion.

The floating BL scheme with its extended control vector and non-linear observation operator is still therefore a much more sophisticated solution, the effect of which the ensemble cannot replicate. We could, instead however, use the B-matrix implied from the floating BL scheme to see what kind of description of the background amplitude error covariances may be used with the standard scheme in order to try to replicate the floating BL scheme. This is investigated in the next subsection.

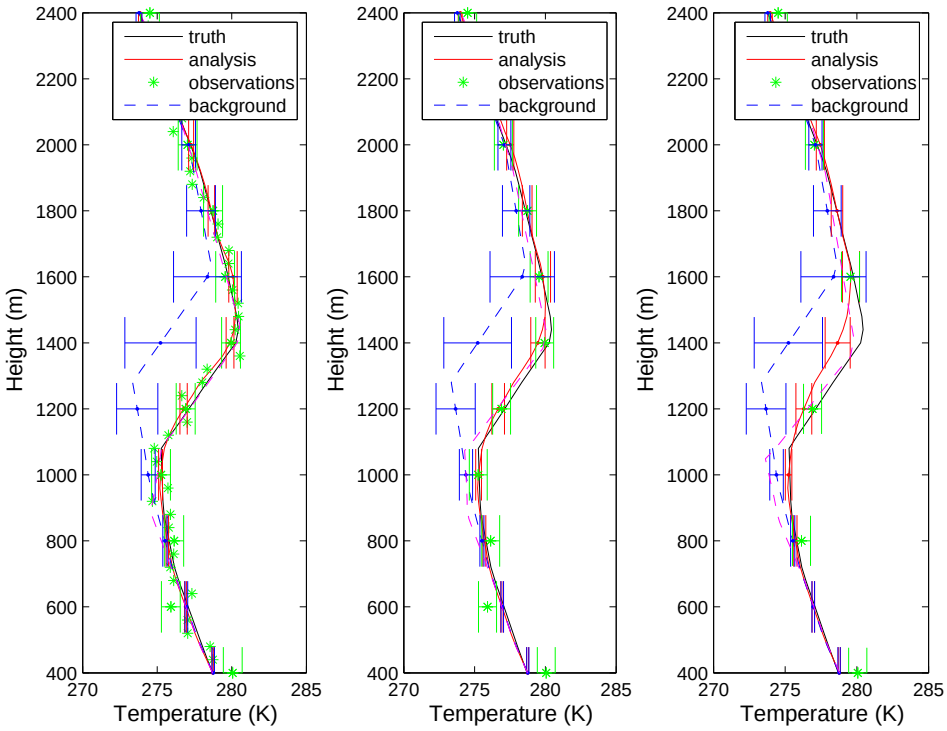


Figure 4.2: Assimilation using the standard scheme with the B-matrix produced in Figure 4.1 when the truth is observed at every level (left), every 5th model level (middle) and every 10th model level (right). In each case the analysis calculated using the floating BL scheme is plotted in dashed magenta.

4.1.2 Calculating an amplitude B-matrix for the standard scheme using the floating BL scheme

The introduction of the extra control variable, a, effectively permits a vertical position error in $\mathbf{x}^{\mathrm{b}}$ as seen in the previous chapter. In fixed model space this has the effect of inflating the implied background amplitude errors in the region of the inversion in a similar way to that seen when an ensemble of forecasts represents the positional error in the inversion (see Subsection 4.1.1). It is possible to study these implied errors through the transformation of the background error covariance matrix $\tilde{\mathbf{B}}_{\mathbf{v}}$, used in the floating BL scheme. The transformation matrix, $\mathbf{Z}$, which maps from floating to fixed space is made up of the partial derivatives of $\mathbf{x}$ (temperature on fixed model levels) with respect to $\tilde{\mathbf{v}}$ (temperature on floating model levels and a) and so the matrix is a linear interpolation from the floating levels to the fixed levels which has a contribution from a (see Figure 4.3), i.e. $\delta\mathbf{x} = \mathbf{Z}\delta\tilde{\mathbf{v}}$. $\mathbf{Z}$ is a diagnostic tool and has not been used in the assimilation. $\mathbf{Z}$ can be compared to the linearised observation operator, $\mathbf{H}$, which is used within the assimilation. $\mathbf{H}$ represents the linear interpolation of values from the floating model levels to fixed observation levels (see appendix C). $\mathbf{Z}$ is a $N \times (N+1)$ matrix where as $\mathbf{H}$ is a $P \times (N+1)$ matrix.

$$\left(\begin{array}{c|c} \dfrac{\partial \mathbf{x}}{\partial \tilde{\mathbf{x}}} & \dfrac{\partial \mathbf{x}}{\partial a} \end{array} \right)$$

Figure 4.3: The transform from floating to fixed model levels, $\mathbf{Z}$. $\frac{\delta \mathbf{x}}{\delta a}$ is non-zero when $D(z)$ is non-zero.

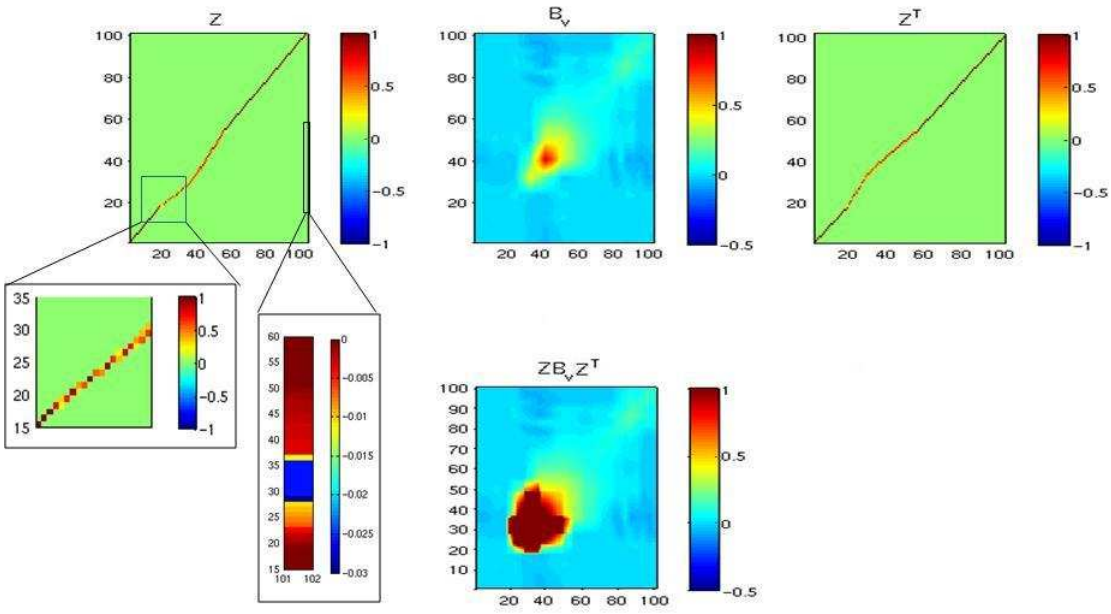


Figure 4.4: A schematic of how $\mathbf{B}_{\text{impl}}$ has been calculated for the analysis shown in Figure 3.12. $\mathbf{Z}$ is a $N \times (N+1)$ matrix, $\tilde{\mathbf{B}}_{\mathbf{v}}$ is a $(N+1) \times (N+1)$ matrix and therefore $\mathbf{B}_{\text{impl}}$ is a $N \times N$ matrix. $\mathbf{Z} = \partial \mathbf{x} / \partial \tilde{\mathbf{v}}$ and so is predominately an interpolation matrix with a small contribution from $\partial \mathbf{x} / \partial a$ in the region where the displacement function is non zero.

Operating on $\tilde{\mathbf{B}}_{\mathbf{v}}$ with $\mathbf{Z}$ gives the implied background errors in fixed model space,

$$\mathbf{B}_{\text{impl}} = \mathbf{Z} \tilde{\mathbf{B}}_{\mathbf{v}} \mathbf{Z}^{\mathrm{T}}. \tag{4.1}$$

$\mathbf{B}_{\text{impl}}$ has been calculated using the value of a in the analysis of Figure 3.12 (the illustration in Section 3.2 where the floating BL scheme was used with a flow-dependent B-matrix). A diagram of the calculation of $\mathbf{B}_{\text{impl}}$ is shown in Figure 4.4, with the specific $\mathbf{Z}$ and $\mathbf{B}_{\text{impl}}$ for the case of Figure 3.12. We see that $\mathbf{Z}$ resembles the identity where $D(z)$=0 (between levels 1 and 18, and levels 55 to 100), and represents linear interpolation where $D(z) \neq 0$. It is difficult in this Figure to see the contribution from $\frac{\delta \mathbf{x}}{\delta a}$ as the values are much less than one but not insignificant, and so have been blown up and plotted on a separate scale.

The implied background error standard deviations (square-roots of diagonal elements $\mathbf{B}_{\text{impl}}$) are plotted in Figure 4.5. Due to the structure of the displacement function giving

rise to non-zero values for $\partial \mathbf{x}/\partial a$ in the region where the displacement function is non-zero only, the standard deviations for $\mathbf{B}_{\text{impl}}$ are inflated only in that region. We see that the greatest increase in error is between 1km and 1.4km which is the main region where the background and observed inversions disagree. This gives the necessary amplitude of background errors in order for the standard scheme to mirror the floating BL scheme's movement of the background inversion. We note that this inflation is more than double that seen in the previous subsection (Figure 4.1). This effective error increase is a function of a and so is dependent not only on the height of the inversion in the background (as was the case in the previous subsection) but also on the true height of the inversion as observed.

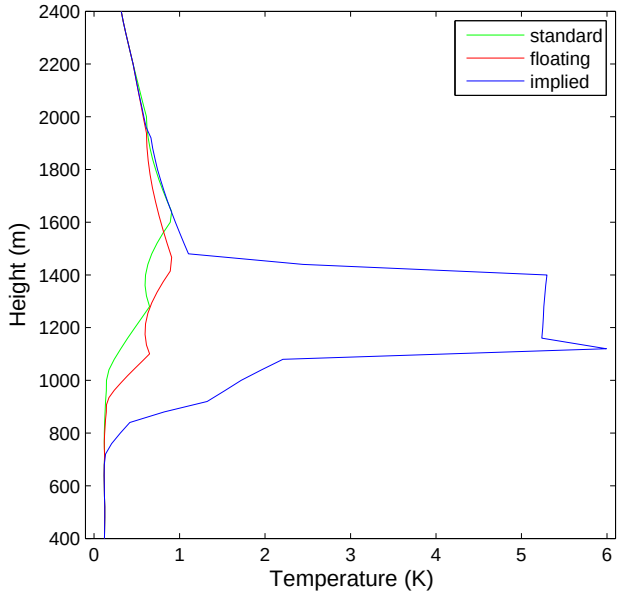


Figure 4.5: Background error standard deviations on fixed levels as would be used in standard 1DVar (green) and on floating levels (red) as used by the floating BL scheme. The implied standard deviations derived from (4.1) are given in blue. The implied errors have been calculated using the value of a given by the analysis in Figure 3.12.

In Figure 4.6 the temperature correlation structures for $\tilde{\mathbf{B}}_{\mathbf{v}}$ (left hand Figure) and $\mathbf{B}_{\text{impl}}$ for the analysis in Figure 3.12 (right hand Figure) are shown (labelled $\tilde{\mathbf{C}}_{\mathbf{v}}$ and $\mathbf{C}_{\text{impl}}$ respectively). In the left hand Figure it is possible to identify the BL as the region of strong positive correlation spanning the lowest 25 or so levels. In the analysis of Figure 3.12 the assimilation has moved the inversion down by the equivalent of approximately 5 model level spacings and so the BL is shallower and the region of strong positive correlation seen in $\tilde{\mathbf{C}}_{\mathbf{v}}$ has shrunk slightly in $\mathbf{C}_{\text{impl}}$.

In Figure 4.7, $\tilde{\mathbf{C}}_{\mathbf{v}}$ (left hand Figure) and $\mathbf{C}_{\text{impl}}$ (right hand Figure) have been plotted against the model level heights so that it is possible to see that the region of strong positive correlation marking the BL are similar in size.

In addition to the resizing of the correlations associated with the BL a change in the correlation structure can also be seen in $\mathbf{B}_{\text{impl}}$ where $D(z)$ is non-zero (700m$< z <$2200m). The correlation length scales have increased so that the levels just above the inversion are

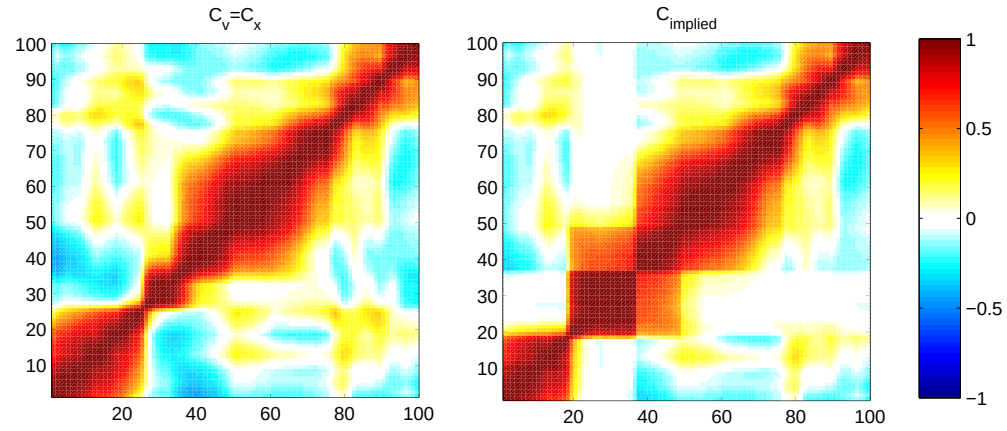


Figure 4.6: The background error correlation matrix used in the floating BL scheme on floating model levels (left) and the correlation matrix implied on the original fixed model levels (right). Axes are given in terms of model levels (on the left these are the floating model levels and on the right these are the fixed model levels).

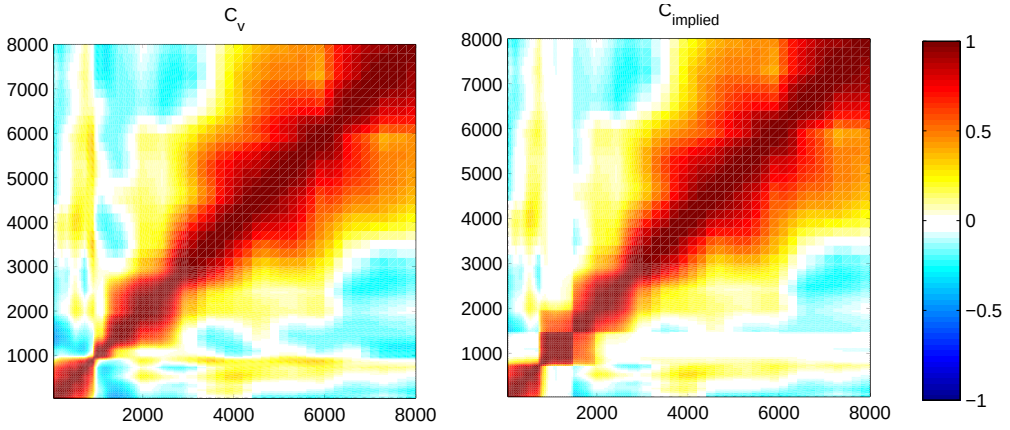


Figure 4.7: Same as in Figure 4.6 but now plotted against the model level heights.

correlated up to 2km although the decoupling of errors above and below the inversion has remained. Along with inflating the background error variances this is the correlation structure necessary for the 1DVar scheme on fixed levels to give an analysis close to that seen in Figure 3.12.

Comparing the floating BL scheme with the standard 1DVar using $\mathbf{B}_{\text{impl}}$ shows the effect of the non-linear observation operator compared to the linear observation operator with an optimally specified B-matrix. The result from this can be seen in Figure 4.8. The solid red line shows the standard 1DVar analysis using $\mathbf{B}_{\text{impl}}$ and the magenta dashed line shows the analysis given by the floating BL scheme. The background error bars are now much larger all the way down to the position of the true inversion, allowing consistency between the background and the observations, but remain small above the background inversion. Important issues to note are:

- The amplitude errors have been inflated much more than is necessary and so do not make the most of the background information available

Observation density	floating BL scheme	standard scheme	
		B from § 4.1.1	**B** from § 4.1.2
every model level (100 obs)	48.3	40.9	51.0
every 5th model level (20 obs)	9.08	10.8	8.33
every 10th model level (10 obs)	6.59	6.06	4.09

Table 4.1: The value of J at the analysis.

- The standard 1DVar scheme with $\mathbf{B}_{\text{impl}}$ is less smooth than the floating BL scheme, due to the high relative weight given to the observations in the standard scheme and the sharp correlation structure produced in $\mathbf{B}_{\text{impl}}$ (Figure 4.6 right hand panel).
- The floating BL scheme has kept the structure of the background, but the standard scheme cannot.
- The value of J at the analysis is comparable and consistent with the Bennett-Talagrand result, $J(x^{\text{a}}) \approx$ number of observations/2 (see appendix A), see Table 4.1, suggesting that in each case the background and observed profiles given are now consistent with the error statistics assumed.

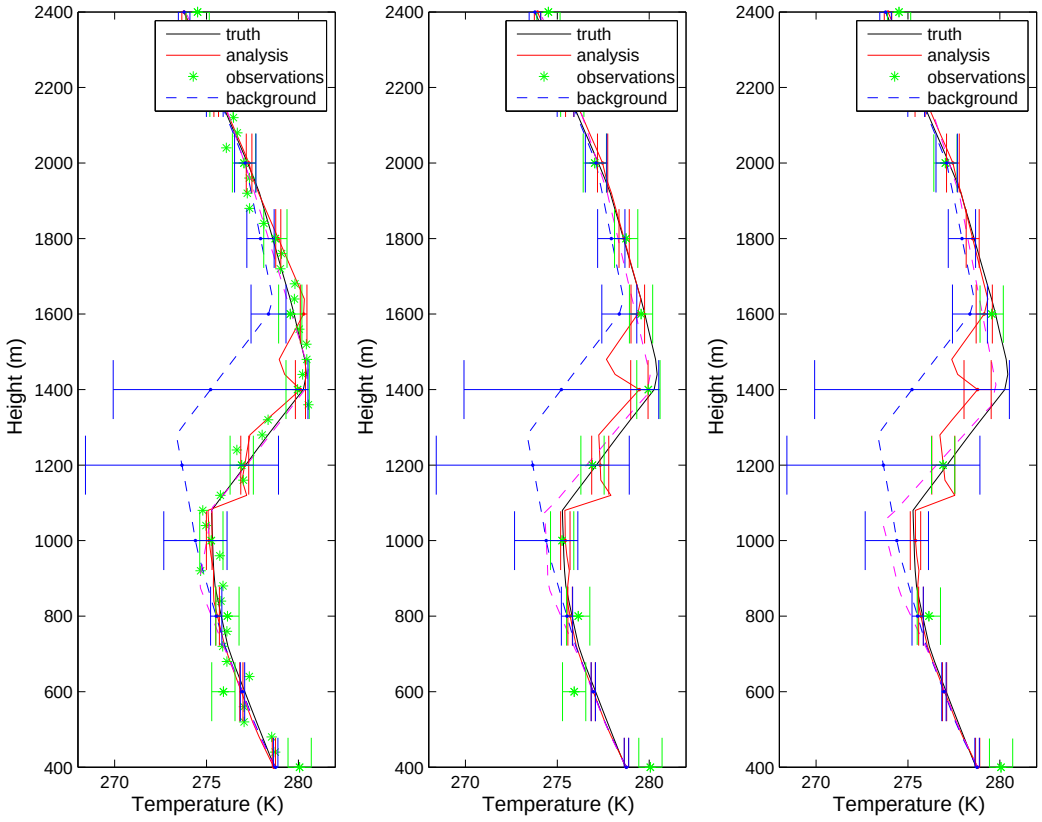


Figure 4.8: Assimilation using the standard scheme with the implied B-matrix given by (4.1) when the truth is observed at every model level (left), every 5th model level (middle) and every 10th model level (right). In each case the analysis calculated using the floating BL scheme is plotted in dashed magenta.

With this improved description of background errors the standard 1DVar scheme is able to give an analysis more consistent with observations. The inflation of the background errors in the inversion means that the background inversion structure is assumed to be wrong, unlike the floating BL scheme where it is the height and/or structure that may be assumed wrong. In addition to this the analysis, as in Subsection 4.1.1, is sensitive to the observation density. We can therefore conclude from these results seen that the floating BL scheme cannot be satisfactorily reproduced by the standard scheme, even with a different description of background amplitude errors.

We shall now look at the observation operator for the floating BL scheme which is a unique aspect of this scheme over the standard scheme.

4.2 Non-linearity of the observation operator

One way the floating BL scheme differs from the standard scheme is that the observation operator, $\mathcal{H}$, is no longer linear even with linear interpolation ($\mathcal{H}$ for the floating BL scheme has been defined in (3.6)). A linearisation of $\mathcal{H}$ about a state is necessary for the minimisation of the cost function as seen in Chapter 2 (2.18). Defining the tangent linear operator, $\mathbf{H}_n$, as a matrix of partial derivatives of $\mathcal{H}(\tilde{\mathbf{v}})$ with respect to $\tilde{\mathbf{v}}$ evaluated at $\tilde{\mathbf{v}}_n$, the analysis can be found iteratively by solving

$$\delta\tilde{\mathbf{v}}_{n+1} = (\tilde{\mathbf{B}}_{\mathrm{v}}^{-1} + \mathbf{H}_n^{\mathrm{T}}\mathbf{R}^{-1}\mathbf{H}_n)^{-1}\left(\tilde{\mathbf{B}}_{\mathrm{v}}^{-1}\delta\tilde{\mathbf{v}}_n^{\mathrm{b}} + \mathbf{H}_n^{\mathrm{T}}\mathbf{R}^{-1}\delta\mathbf{y}_n\right), \tag{4.2}$$

where $\delta\tilde{\mathbf{v}}_{n+1} = \tilde{\mathbf{v}}_{n+1} - \tilde{\mathbf{v}}_n$, $\delta\tilde{\mathbf{v}}_n^{\mathrm{b}} = \tilde{\mathbf{v}}^{\mathrm{b}} - \tilde{\mathbf{v}}_n$ and $\delta\mathbf{y}_n = \mathbf{y} - \mathcal{H}(\tilde{\mathbf{v}}_n)$. The evaluation of $\mathbf{H}$, for the floating scheme, is given in appendix C [1] .

4.2.1 Linearisation error

Estimating $\mathcal{H}(\tilde{\mathbf{v}}_{n+1}) - \mathcal{H}(\tilde{\mathbf{v}}_n)$ by $\mathbf{H}_n\delta\tilde{\mathbf{v}}_n$ and ignoring higher order terms introduces a linearisation error. The linearisation error can be quantified as

$$\mathcal{H}(\tilde{\mathbf{v}}_{n+1}) - (\mathcal{H}(\tilde{\mathbf{v}}_n) + \mathbf{H}_n\delta\tilde{\mathbf{v}}_n) \tag{4.3}$$

If there is no linearisation error then (4.3) will equal zero, as is the case here when D is equal to zero. It is expected that the linearisation error should increase with an increasing $|\delta\tilde{\mathbf{v}}|$. This is illustrated in Figure 4.9a where the linearisation error has been calculated for $\delta\tilde{\mathbf{v}}$ (illustrated in Figure 4.9b) equal to the standard deviation of the background errors (blue solid). a_n already has a value of 200m in this example. The linearisation error has then been calculated for half this increment (red solid) and a fifth (green solid). It is seen that the linearisation error decreases proportionally to the size of the increment.
There are peaks in the linearisation error close to the the top and bottom of the inversion in $\tilde{\mathbf{v}}_n$ (Figure 4.9d, black line) due to the interpolation of this sharp feature from floating levels which have a varying resolution, to observed levels which have a resolution of 40m.

[1] These iterations are analogous to the 'outer loop' iterations in 3D and 4DVar (see e.g. Rabier et al., 2000) (inner loop iterations are not required here as the approximate Hessian $\tilde{\mathbf{B}}_{\mathrm{v}}^{-1} + \mathbf{H}_n^{\mathrm{T}}\mathbf{R}^{-1}\mathbf{H}_n$ is inverted directly.)

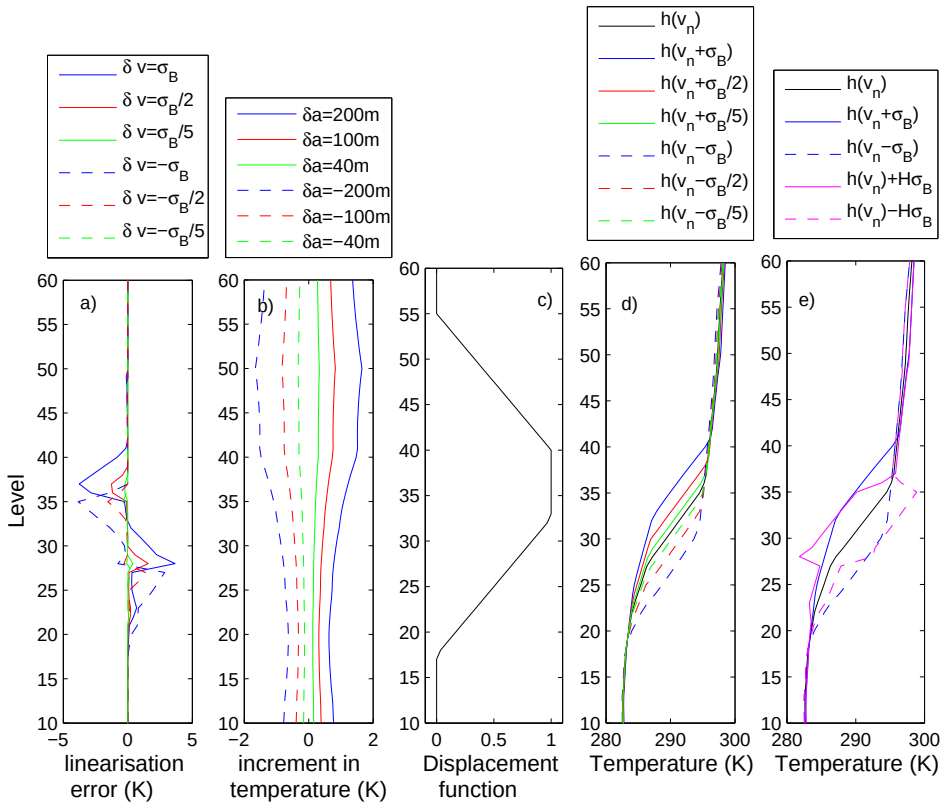


Figure 4.9: a) The linearisation error calculated from (4.3) for varying increment size. b) the size of the amplitude increment. c) The displacement function used. d) Potential temperature profile for n^{th} iteration in observation space (black) and for the $n+1^{th}$ iteration depending on the increment. e) Comparison of $\mathcal{H}(\tilde{\mathbf{v}}_n + \delta\tilde{\mathbf{v}}_{n+1})$ (blue) and the linear estimate $\mathcal{H}(\tilde{\mathbf{v}}_n) + \mathbf{H}\delta\tilde{\mathbf{v}}_n$ (magenta).

For negative increments (dashed lines) the error structure is very similar but shifted down. In Figure 4.9e the linearised estimate of $\mathcal{H}(\tilde{\mathbf{x}}_{n+1})$ when $\delta\tilde{\mathbf{v}}_n$ equals +/- the standard deviation of the background errors (magenta) is given for comparison with the non-linear $\mathcal{H}(\tilde{\mathbf{x}}_{n+1})$ (blue). The linearised estimate is clearly poor and the need to solve the assimilation iteratively (where $\delta\tilde{\mathbf{v}}_n$ becomes smaller with progressive iterations) is evident.

The linearisation error may be thought of as a contribution to the representativeness error (see Lorenc, 1986 and Subsection 1.2.3), as it produces an error in the estimate of the state when transformed to observation space for comparison with the observation. The linearisation error should be included in $\mathbf{R}$, which could significantly increase the observation error where the displacement function is non-zero. The value of σ_o^2 used for all observation levels in this experiment is 0.4K^2.

The numerical nature of the linearisation error may be studied by looking at the Taylor expansion of $\mathcal{H}(\tilde{\mathbf{v}})$ about $\tilde{\mathbf{v}}^n$ for higher order terms:

$$\mathcal{H}(\tilde{\mathbf{v}}_{n+1}) = \mathcal{H}(\tilde{\mathbf{v}}_n) + \mathbf{H}\delta\tilde{\mathbf{v}} + \delta\tilde{\mathbf{v}}^{\mathrm{T}}\nabla^2\mathcal{H}\delta\tilde{\mathbf{v}} + ... \tag{4.4}$$

where $\delta\tilde{\mathbf{v}} = \begin{pmatrix} \delta\tilde{\mathbf{x}} \\ \delta a \end{pmatrix}$ and $\delta\tilde{\mathbf{v}}^{\mathrm{T}}\nabla^2\mathcal{H}\delta\tilde{\mathbf{v}} = \begin{pmatrix} \delta\tilde{\mathbf{v}}^{\mathrm{T}}\nabla^2\mathcal{H}^1\delta\tilde{\mathbf{v}} & ... & \delta\tilde{\mathbf{v}}^{\mathrm{T}}\nabla^2\mathcal{H}^j\delta\tilde{\mathbf{v}} & ... & \delta\tilde{\mathbf{v}}^{\mathrm{T}}\nabla^2\mathcal{H}^P\delta\tilde{\mathbf{v}} \end{pmatrix}^{\mathrm{T}}$ is a vector of length P. The second derivative of each component of $\mathcal{H}$ with respect to $\tilde{\mathbf{v}}$

can be expressed as:

$$\nabla^2\mathcal{H}^j = \begin{pmatrix} \frac{\partial^2\mathcal{H}^j}{\partial\tilde{\mathbf{x}}^2} & \frac{\partial^2\mathcal{H}^j}{\partial\tilde{\mathbf{x}}\partial a} \\ \frac{\partial^2\mathcal{H}^j}{\partial a\partial\tilde{\mathbf{x}}} & \frac{\partial^2\mathcal{H}^j}{\partial a^2} \end{pmatrix} \quad \text{for } j = 1, ..., P. \tag{4.5}$$

This is a $(N+1)\times(N+1)$ matrix. It is clear from (3.6) that $\frac{\partial^2\mathcal{H}^j}{\partial\tilde{\mathbf{x}}^2}$, a $N\times N$ matrix, is zero for each element and $\left(\frac{\partial^2\mathcal{H}^j}{\partial\tilde{\mathbf{x}}\partial a}\right)^{\mathrm{T}} = \frac{\partial^2\mathcal{H}^j}{\partial a\partial\tilde{\mathbf{x}}} = (0, ..., 0, \frac{\partial^2\mathcal{H}^j}{\partial\tilde{x}^{\mathrm{I}}\partial a}, \frac{\partial^2\mathcal{H}^j}{\partial\tilde{x}^{\mathrm{II}}\partial a}, 0, ..., 0)$, where $\tilde{x}^{\mathrm{I}}$ and $\tilde{x}^{\mathrm{II}}$ are the values of $\mathbf{x}$ in floating level space associated with the levels $\tilde{z}^{\mathrm{I}}$ and $\tilde{z}^{\mathrm{II}}$. $\tilde{z}^{\mathrm{I}}$ and $\tilde{z}^{\mathrm{II}}$ are the floating model levels either side of the location of the j$^{\mathrm{th}}$ observation. We can now express $\delta\tilde{\mathbf{v}}^{\mathrm{T}}\nabla^2\mathcal{H}^j\delta\tilde{\mathbf{v}}$ as

$$\delta\tilde{\mathbf{v}}^{\mathrm{T}}\nabla^2\mathcal{H}^j\delta\tilde{\mathbf{v}} = 2\delta a\left(\frac{\partial^2\mathcal{H}^j}{\partial\tilde{x}^{\mathrm{I}}\partial a}\delta\tilde{x}^{\mathrm{I}} + \frac{\partial^2\mathcal{H}^j}{\partial\tilde{x}^{\mathrm{II}}\partial a}\delta\tilde{x}^{\mathrm{II}}\right) + \delta a^2\frac{\partial^2\mathcal{H}^j}{\partial a^2} \quad \text{for } j = 1, ..., P. \tag{4.6}$$

where $\frac{\partial^2\mathcal{H}^j}{\partial\tilde{x}^{\mathrm{I}}\partial a}$, $\frac{\partial^2\mathcal{H}^j}{\partial\tilde{x}^{\mathrm{II}}\partial a}$ and $\frac{\partial^2\mathcal{H}^j}{\partial a^2}$ are functions of a and the choice of $D(z)$ (see appendix C). Adding this error to $\mathbf{R}$ would mean $\mathbf{R}$ should become dependent upon the presence, location and misplacement (given by a) of the background inversion, as well as the size of $\delta\tilde{\mathbf{v}}$.

If the linearisation error is not included in $\mathbf{R}$ then the analysis error will be subject to additional uncertainty; this can be evaluated as (Rodgers, 2000)

$$\mathbf{K}\{\mathcal{H}(\tilde{\mathbf{v}}^{\mathrm{t}}) - (\mathcal{H}(\tilde{\mathbf{v}}^{\mathrm{a}}) + \mathbf{H}_{\mathbf{n}}(\tilde{\mathbf{v}}^{\mathrm{t}} - \tilde{\mathbf{v}}^{\mathrm{a}}))\}, \tag{4.7}$$

where $\mathbf{K}$ is the Kalman gain matrix given in Chapter 2, $\tilde{\mathbf{v}}^{\mathrm{t}}$ is the truth and $\tilde{\mathbf{v}}^{\mathrm{a}}$ is the analysed value. Assuming the solution $\tilde{\mathbf{v}}^{\mathrm{a}}$ has converged to the truth this value will be small and so our estimate of the analysis error may not need to be altered.

4.2.2 Analysis error

For a linear observation operator and unbiased Gaussian error statistics the analysis error covariance matrix, $\mathbf{P}^{\mathrm{a}}$ can be shown to be equal to the inverse of the Hessian of the cost function, $\mathbf{A}$,

$$\mathbf{P}^{\mathrm{a}} = \mathbf{A}^{-1} = (\mathbf{B}^{-1} + \mathbf{H}^{\mathrm{T}}\mathbf{R}^{-1}\mathbf{H})^{-1}. \tag{4.8}$$

The Hessian of the cost function is a positive definite matrix and measures the convexity of the cost function. Therefore the convexity of the cost function gives us insight into the accuracy of the solution. However, when the observation operator is non-linear (4.8) is only an approximation to $\mathbf{P}^{\mathrm{a}}$ and the analysis errors can no longer be assumed to be Gaussian.

PROOF: $\mathbf{P}^{\mathrm{a}} = \mathbf{A}^{-1}$ when $\mathcal{H}$ is linear and there is no bias

---

$\mathbf{A} = \nabla^2 J = \mathbf{B}^{-1} + \mathbf{H}^{\mathrm{T}}\mathbf{R}^{-1}\mathbf{H}$ (the second derivative of the cost function (2.11))

At the analysis $\nabla J = 0$

implying that $\mathbf{B}^{-1}(\mathbf{x}^{\mathrm{a}} - \mathbf{x}^{\mathrm{b}}) - \mathbf{H}^{\mathrm{T}}\mathbf{R}^{-1}(\mathbf{y} - \mathcal{H}(\mathbf{x}^{\mathrm{a}})) = 0$ (see (2.12))

Using the second order Taylor expansion of $\mathcal{H}(\mathbf{x}^{\mathrm{a}})$ about $\mathbf{x}^{\mathrm{t}}$, this is equivalent to $\mathbf{B}^{-1}(\mathbf{x}^{\mathrm{a}} - \mathbf{x}^{\mathrm{t}} - \mathbf{x}^{\mathrm{b}} + \mathbf{x}^{\mathrm{t}}) - \mathbf{H}^{\mathrm{T}}\mathbf{R}^{-1}(\mathbf{y} - \mathcal{H}(\mathbf{x}^{\mathrm{t}}) - \mathbf{H}(\mathbf{x}^{\mathrm{a}} - \mathbf{x}^{\mathrm{t}}))$ where $\mathbf{x}^{\mathrm{t}}$ is the truth.

Let $\boldsymbol{\eta}^{\mathrm{a}} = \mathbf{x}^{\mathrm{a}} - \mathbf{x}^{\mathrm{t}}$, $\boldsymbol{\eta}^{\mathrm{b}} = \mathbf{x}^{\mathrm{b}} - \mathbf{x}^{\mathrm{t}}$ and $\boldsymbol{\eta}^{\mathrm{o}} = \mathbf{y} - \mathcal{H}(\mathbf{x}^{\mathrm{t}})$

Then: $(\mathbf{B}^{-1} + \mathbf{H}^{\mathrm{T}}\mathbf{R}^{-1}\mathbf{H})\boldsymbol{\eta}^{\mathrm{a}} = \mathbf{B}^{-1}\boldsymbol{\eta}^{\mathrm{b}} + \mathbf{H}^{\mathrm{T}}\mathbf{R}^{-1}\boldsymbol{\eta}^{\mathrm{o}}$

Multiplying each side by its transpose and taking the expectant with the assumption that $\left\langle \boldsymbol{\eta}^{\mathrm{b}}\boldsymbol{\eta}^{\mathrm{oT}} \right\rangle = \left\langle \boldsymbol{\eta}^{\mathrm{o}}\boldsymbol{\eta}^{\mathrm{bT}} \right\rangle = 0$, i.e. the background and observation errors are uncorrelated, gives:

$$(\mathbf{B}^{-1} + \mathbf{H}^{\mathrm{T}}\mathbf{R}^{-1}\mathbf{H}) \left\langle \boldsymbol{\eta}^{\mathrm{a}}\boldsymbol{\eta}^{\mathrm{aT}} \right\rangle (\mathbf{B}^{-1} + \mathbf{H}^{\mathrm{T}}\mathbf{R}^{-1}\mathbf{H})^{\mathrm{T}} =$$
$$\mathbf{B}^{-1} \left\langle \boldsymbol{\eta}^{\mathrm{b}}\boldsymbol{\eta}^{\mathrm{bT}} \right\rangle \mathbf{B}^{-1} + \mathbf{H}^{\mathrm{T}}\mathbf{R}^{-1} \left\langle \boldsymbol{\eta}^{\mathrm{o}}\boldsymbol{\eta}^{\mathrm{oT}} \right\rangle \mathbf{R}^{-1}\mathbf{H}$$

Substituting in the definitions of $\mathbf{P}^{\mathrm{a}} = \left\langle \boldsymbol{\eta}^{\mathrm{a}}\boldsymbol{\eta}^{\mathrm{aT}} \right\rangle$, $\mathbf{B} = \left\langle \boldsymbol{\eta}^{\mathrm{b}}\boldsymbol{\eta}^{\mathrm{bT}} \right\rangle$ and $\mathbf{R} = \left\langle \boldsymbol{\eta}^{\mathrm{o}}\boldsymbol{\eta}^{\mathrm{oT}} \right\rangle$ gives:

$$(\mathbf{B}^{-1} + \mathbf{H}^{\mathrm{T}}\mathbf{R}^{-1}\mathbf{H})\mathbf{P}^{\mathrm{a}}(\mathbf{B}^{-1} + \mathbf{H}^{\mathrm{T}}\mathbf{R}^{-1}\mathbf{H}) = \mathbf{B}^{-1} + \mathbf{H}^{\mathrm{T}}\mathbf{R}^{-1}\mathbf{H}.$$

Therefore: $\mathbf{P}^{\mathrm{a}} = (\mathbf{B}^{-1} + \mathbf{H}^{\mathrm{T}}\mathbf{R}^{-1}\mathbf{H})^{-1} = \mathbf{A}^{-1}$ Q.E.D.

An estimate of $\mathbf{P}^{\mathrm{a}}$ which does not assume that the observation operator is near linear can be obtained in a similar way to our flow-dependent estimates of $\mathbf{B}$ in Section 2.3. An ensemble of analyses can be obtained by running the floating BL scheme multiple times with observations measuring the truth plus a random noise and backgrounds taken from an ensemble of forecasts, which each have the inversion height shifted from the truth by a random amount with an error standard deviation of 200m. $\mathbf{P}^{\mathrm{a}}$ can then be calculated as

$$\mathbf{P}^{\mathrm{a}} = \left\langle \left(\mathbf{x}^{\mathrm{a}} - \mathbf{x}^{\mathrm{t}}\right) \left(\mathbf{x}^{\mathrm{a}} - \mathbf{x}^{\mathrm{t}}\right)^{\mathrm{T}} \right\rangle \tag{4.9}$$

where $\mathbf{x}^{\mathrm{t}}$ is the truth and the expectation is taken over ensemble members.

In Figure 4.10 a 100-member ensemble of observations (green stars) and backgrounds (blue profiles) have been created in this way from the truth (black). For each background and 'observed' temperature profile pair, an analysis has also been calculated using the floating BL scheme (red). The spread in the analyses is much smaller than the spread in the background and observations as expected, and the analyses give a good representation of the true inversion.

The ensemble of analyses allows us to look at the distribution of each variable and assess the validity of the Gaussian error assumption. In Figure 4.11 histograms are given for the random shift used to create the ensemble of backgrounds from the truth (left), modelled as $\sim N(0, 200\mathrm{m})$, the value of a given by the analysis (middle) and the shift plus a (right). If this scheme only sought to minimise the positional error then we would expect the value of the background inversion was shifted plus a to equal zero. The difference between the shift in the background inversion and a may then be thought of as an estimate of the positional error in the analysis inversion height. The standard deviation of this error is 71.4m (in Figure 4.10 the positional error in the analysis appears smaller due to the additional amplitude error allowing for some effective shift in the inversion height). We see that the distribution of analysed a is biased and the mean is not equal to the mode. Even though the background errors of a were initially modelled as Gaussian, the non-linear observation operator has meant that they have not remained Gaussian in the analysis.

From the ensemble of analyses we can also see how the statistics of the analysis error have been altered by the non-linear observation operator. The analysis error has been

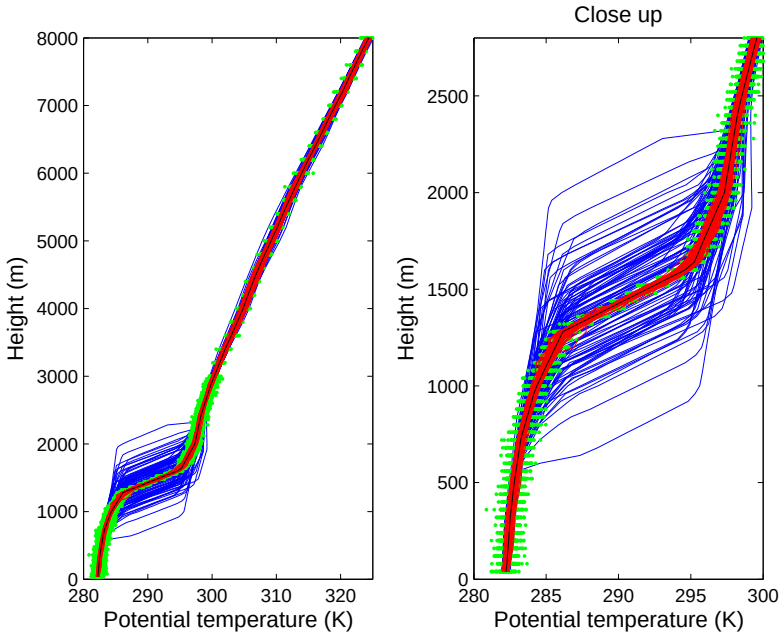


Figure 4.10: An ensemble of analyses (red) calculated for background (blue) calculated from the truth (black) plus amplitude and positional error and observations (green) measuring the truth plus noise. The true temperature profile is the same as that used to illustrate the floating BL scheme in Chapter 3. The right hand panel is a close-up of the inversion region in the left hand panel.

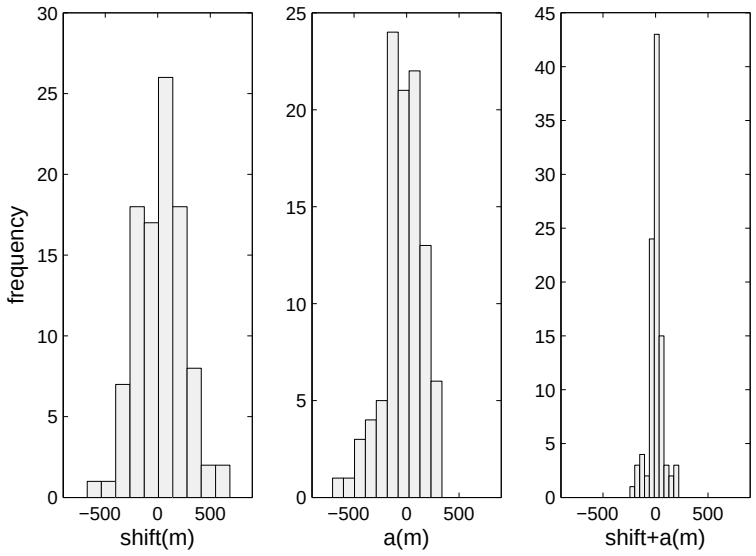


Figure 4.11: Histograms of the amount the 100 backgrounds were shifted (left), the final value of a in the 100 analyses (middle) and the shift in the background inversion height plus a (right).

approximated using (4.9) and compared to a linear estimate using (4.8) averaged over each ensemble member. The analysis error standard deviation is plotted in Figure 4.12 and the analysis error correlation matrix ($\mathbf{C} = \Sigma^{-1}\mathbf{P}^{\mathrm{a}}\Sigma^{-1}$, where Σ is a diagonal matrix containing the standard deviations of $\mathbf{P}^{\mathrm{a}}$) is plotted in Figure 4.13.

In Figure 4.12 we can see that the analysis error standard deviations are comparable however the error is slightly larger calculated using the non-linear estimate, (4.9), than using the linear estimate, (4.8), above and below the inversion (<1km and >2km). The standard deviation of the error in a is also slightly larger using the inverse Hessian method, 81.9m compared to 71.4m for the ensemble method.

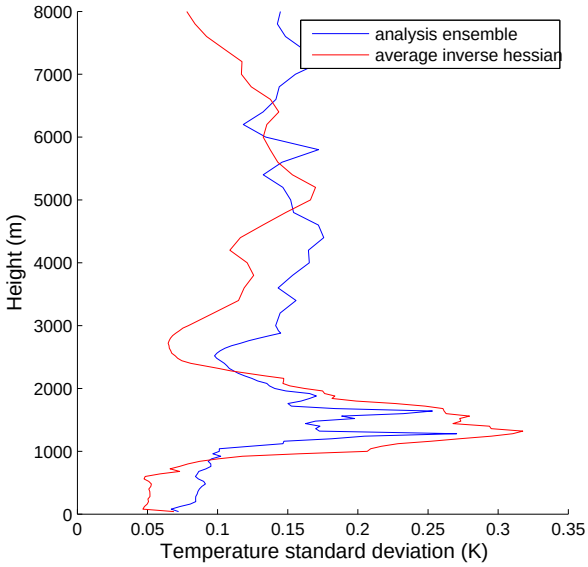


Figure 4.12: The analysis error standard deviation calculated using the analysis ensemble method in (4.9) (blue) and the inverse Hessian method (4.8) (red).

In Figure 4.13 we also see that the correlation structure of $\mathbf{P}^{\mathrm{a}}$ is similar for the two estimates, however again there are small differences. In the left hand panel the correlation structure is shown for the analysis ensemble method, in the right hand panel it is shown for the inverse Hessian method. The regions of strong positive correlation (shown in dark red) are very similar for the two estimates. Both have a strong region of positive correlation within the BL which is consistent with having well mixed air capped by an inversion. Above approximately 3km the inverse Hessian method has slightly longer correlation length scales seen as the broader diagonal band of oranges and reds.

Figures 4.12 and 4.13 appear very similar for the two methods, verifying the validity of the linear theory. However, from Figure 4.11 we see that assumption of Gaussian statistics made by the inverse Hessian method is no longer valid and a measure of the mean and variance is not enough to describe the true PDF of the analysis error statistics. The analysis given by Var represents the mode of $P(\mathbf{x}|\mathbf{y})$ (see Section 2.1), assumed to be Gaussian. The mode may no longer be the mean if the errors are not truly Gaussian and so we may no longer be finding the state of the atmosphere with the smallest error standard deviation as was assumed. Investigating this issue further is beyond the scope of this work but would be a very important future issue to consider, especially as many DA methods involve non-linear aspects.

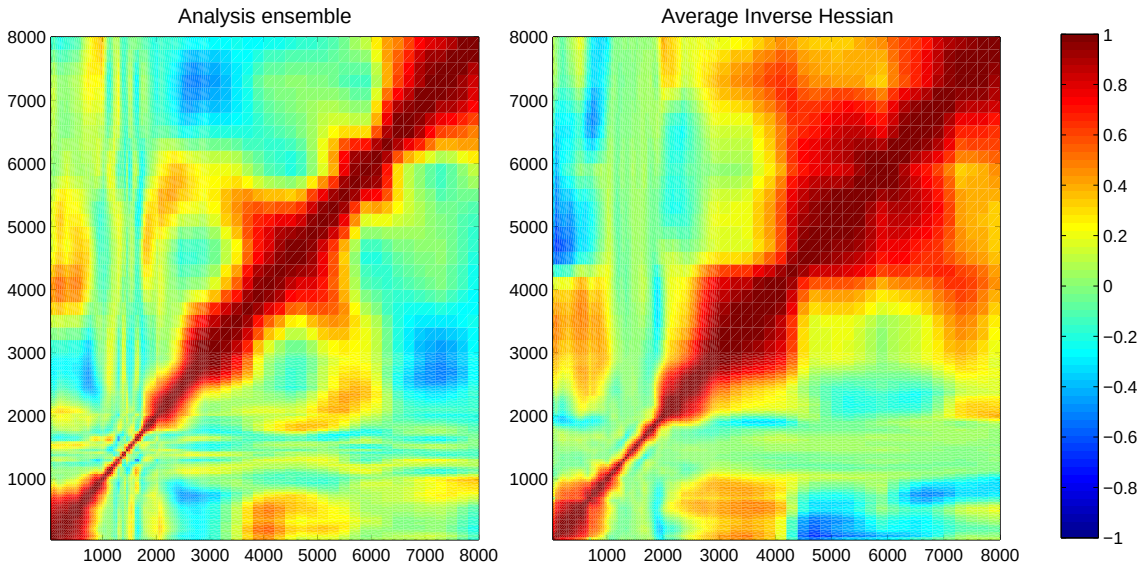


Figure 4.13: The analysis error correlation matrix calculated using the analysis ensemble method in (4.9) (left) and the inverse Hessian method (4.8) (right). The variance of a calculated using the analysis ensemble method is 5010m^2 and using the inverse Hessian method is 6710m^2. The axis for each plot give height in meters.

4.2.3 Shape of the cost function

The effect that the non-linear observation operator has on the statistics of $P(\mathbf{x}|\mathbf{y})$ is reflected in the shape of the cost function and can make it more difficult to find its minimum. When $\mathcal{H}$ is linear and the cost function is quadratic, the convexity of the cost function, $\nabla^2 J$, is $\mathbf{A}$ and is related to the analysis error as shown by (4.8). Therefore the less convex J is, the more unreliable the estimate of $\mathbf{x}^{\mathrm{a}}$, which is intuitive as a stronger convergence criterion would be necessary to pin point the same value at the minimum of a function which is more convex.

In Figure 4.14 the cost function has been calculated for the setup in Figure 3.12 (the illustration of the floating BL scheme with a flow-dependent B-matrix in Chapter 3) as a function of a with the rest of the elements in $\tilde{\mathbf{v}}$ equal to their background value. J^{b} (the background part of the cost function) and J^{o} (the observation part of the cost function) are also plotted separately.

It is clear to see that the quadratic nature of J^{b} is unaffected by the introduction of a and remains small as a is varied. The shape of J is dominated by the observation part of the cost function and its minimum is shifted towards a =-200m. This is the distance that the background inversion is displaced in the set up of Figure 3.12. Within this region of a-space there are no other maxima or minima confirming that the floating BL scheme is able to calculate the optimum displacement in this case.

Single Observations

We are now going to show it is possible for double minima to appear with the introduction of the floating BL scheme, especially if the resolution of the observations is not adequate to

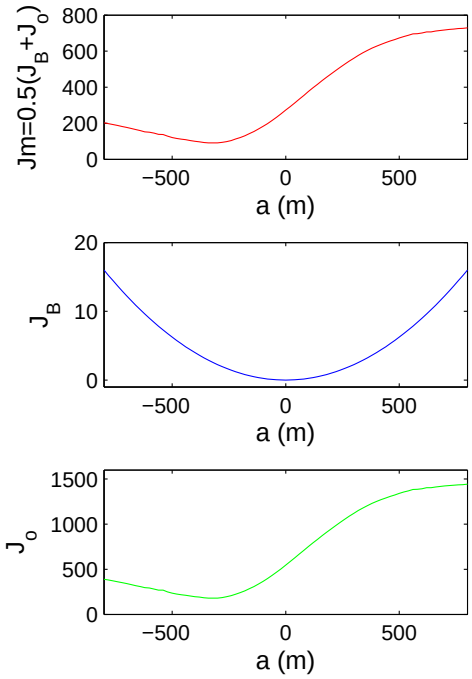


Figure 4.14: The cost function plotted versus a for a fully observed inversion when the background inversion height is displaced from the true inversion height by 200m. The top panel shows the full cost function and the middle and bottom panels show the background and observation cost function terms respectively.

represent the true structure and position of the inversion. In this case it is possible that the background inversion could be moved to different positions in order to give an improved agreement with the observations and can make it questionable whether the particular analysis calculated is truly the most probable state of the atmosphere. An example of this is shown in Figure 4.15. The background inversion is misplaced from the truth by 200m and only one observation of the centre of the true inversion is available at 1400m. The observation no longer resolves the inversion and the floating scheme has a choice to move the background inversion either up or down to give a better fit to this observation. The cost function as a function of a has also been plotted in Figure 4.15. The ambiguity in which way the background should be shifted introduces a second local minimum at approximately 300m in addition to the global minimum at -200m. This could result in the scheme moving the levels in the wrong direction and giving a worse analysis than the standard 1DVar scheme. In this example the maximum value of J separating the two minima is at $a = 0$ which is the background value of a and hence the starting point of the minimisation. This means it is difficult to predict which direction the minimisation of J will choose to take a, using standard Gauss-Newton minimisation. The red profile in Figure 4.15 shows the analysis when a is minimised (incorrectly) to 300m and the green profile shows the analysis when a has been minimised (correctly) to -200m. The only difference between these two assimilations is the B-matrix that has been used. The displacement function in this instance was triangular with a width of 1600m. We see that

in the second analysis (green) the inversion height is in very good agreement with the true inversion and the analysis would have been near-perfect if it took into account adiabatic warming as the structure was shifted down (i.e. if the cost function was minimised in terms of potential temperature instead of temperature). In the first analysis (red), amplitude errors have prevented the inversion from shifting too far up and the positional error has insured that the inversion structure is kept.

The location of the maximum in the cost function can be seen to shift along the a-axis as

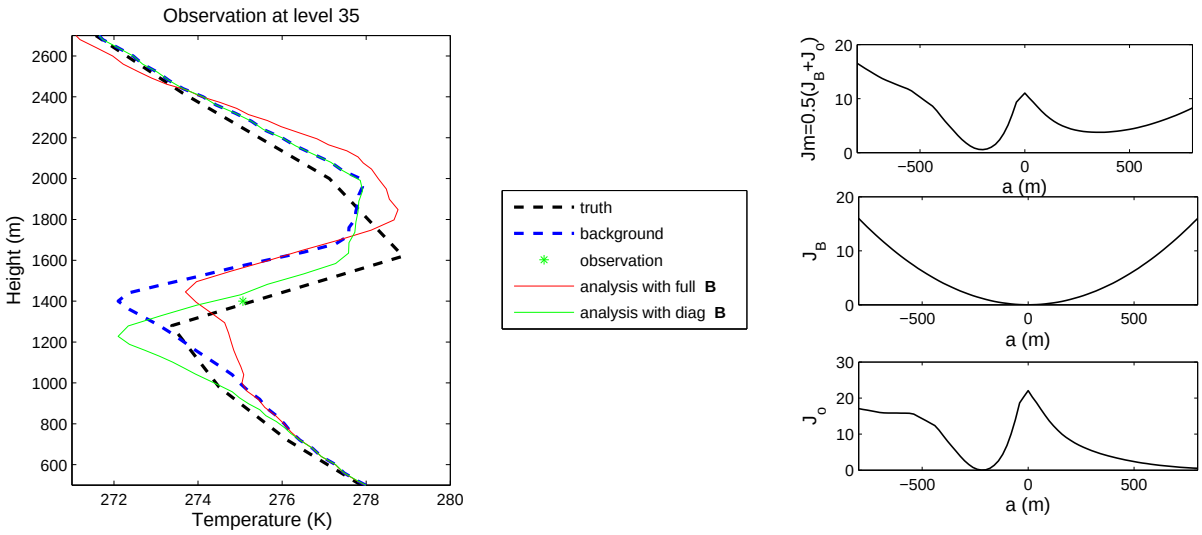


Figure 4.15: Left hand panel: Assimilation of a single observation at 1400m (green star) measuring the centre of a true inversion which is 200m lower than the inversion in the background (blue dashed). The red and green profiles in the left hand panel represent two possible analyses for these background and observation profiles. The red profile evaluates a to be 300m and the green profile evaluates a to be -200m. Right hand panel: The cost function versus a.

the position of the observation measuring the true inversion is shifted. For example if we observe the inversion slightly higher than 1400m it is less likely that the floating scheme will try to shift the profile in the region of the background inversion up in order to meet the observation as it will have further to go. This means that the starting point of the minimisation, $a = 0$, is within the well of the global minimum and hence the assimilation will converge towards the correct value of a. However if we have an observation slightly lower than 1400m we instead move the maximum in the other direction so that $a = 0$ is within the well of the second local minima and the temperature profile would instead erroneously be shifted up. This is shown in Figure 4.16 for an observation of the true inversion at 1360m in the top panel and for the case when only one observation is available at 1440m in the bottom panel. It is also clear that as well as a shift in the location of the maximum, the value of J in the second local minima is also altered, so that for the case of an observation at 1360m (when $a = 0$ is within the well of the second (local) minimum) the value of J at the second minimum is less than in the case of an observation at 1440m. This suggests that when an observation is at 1360m, shifting the background profile up gives a better fit to the observation than shifting the background profile up when the observation is at 1440m.

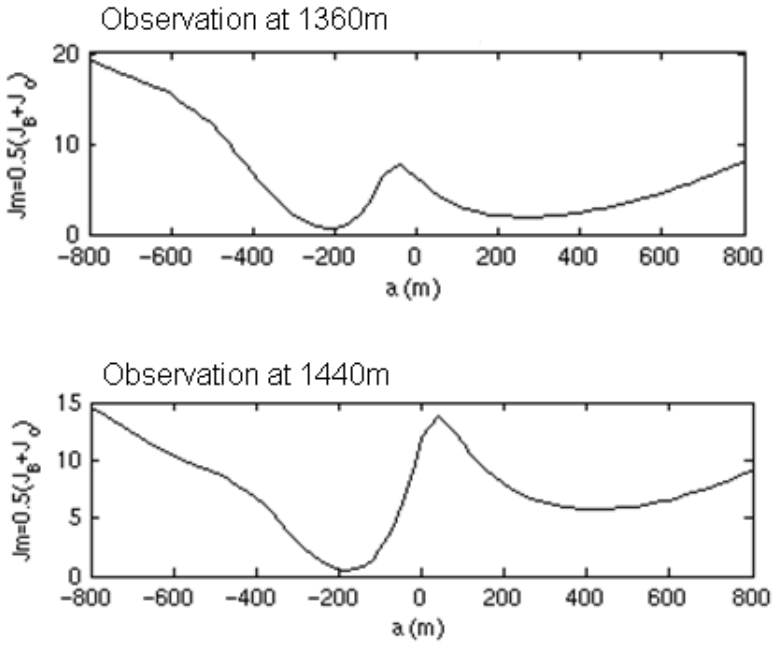


Figure 4.16: Top: Cost function versus a for the assimilation of a single observation at 1360m for the same background and true profiles given in Figure 4.15. Bottom: Cost function versus a for the assimilation of a single observation at 1440m.

Although Figures 4.15 and 4.16 are contrived examples to demonstrate the possible problems with the scheme, it does highlight the importance of an adequate observation density to remove ambiguities. However, the analysis given by the floating BL scheme is far less sensitive to the observation density than the standard scheme (see Figures 4.2 and 4.8), and irrespective of which minima the floating BL scheme finds for the value of a, the background inversion structure always remains, if the observations do not contradict this (see Figure 4.15). It is also important to note that for the two analyses given in Figure 4.15, red and green profiles, could give very different TOA radiances; and so if satellite observations were assimilated as well as this point observation we would be more likely to minimise J correctly, finding the global minimum and not the second local minimum.

4.3 Statistics of the positional error for the background boundary layer capping inversion

In Chapter 2, different methods of estimating the background amplitude error statistics were discussed and we found a flow-dependent estimate could be given by studying the spread in an ensembles of forecasts. To use the floating BL scheme with real data we also need to understand the true error statistics of a.

The standard deviation of a, σ_a, describes the likely range of distances that the inversion found in the background may be allowed to move vertically in order to give a fit to the observed inversion. σ_a may then be thought of as equivalent to the standard deviation of the error in the height of the background inversion. This error can be estimated by comparing the heights of the inversion in a population of collocated actual forecasts

and radiosonde observations similar to a simplification of the analysis of innovations (see Section 2.2). We assume that the observed inversion height gives the true inversion. In practice the observed inversion height could be subject to a representativeness error if the vertical resolution is not high enough to capture the change from a negative to a positive gradient in temperature, but we neglect this effect here.

The height of the inversion in the background, $z^{\mathrm{b}}_{\mathrm{inv}}$, may be described in terms of the true height of the inversion, $z^{\mathrm{t}}_{\mathrm{inv}}$:

$$z^{\mathrm{b}}_{\mathrm{inv}} = z^{\mathrm{t}}_{\mathrm{inv}} - \eta_{\mathrm{inv}}, \tag{4.10}$$

where η_{inv} is the error in the background height of the inversion. Assuming the observed inversion height, $z^{\mathrm{y}}_{\mathrm{inv}}$, gives the true inversion height, we can substitute $z^{\mathrm{y}}_{\mathrm{inv}}$ for $z^{\mathrm{t}}_{\mathrm{inv}}$ in equation (4.10), giving an expression for η_{inv} in terms of calculable and observed quantities

$$\eta_{\mathrm{inv}} = z^{\mathrm{y}}_{\mathrm{inv}} - z^{\mathrm{b}}_{\mathrm{inv}}. \tag{4.11}$$

These data allow, e.g., the variance of the error of the height of the background inversion to be calculated

$$\sigma_a^2 = \left\langle \eta^2_{\mathrm{inv}} \right\rangle \tag{4.12}$$

The triangular brackets represent an average over a suitable population when an inversion is present in both the background and the observations. Ideally the error distribution of any control variable should be Gaussian and non-biased so that the assumptions that the variational assimilation theory is based on are true.

Figure 4.17 shows the distribution of η_{inv} when the height of the inversion is measured in terms of ln(pressure) and so η_{inv} is dimensionless ($\eta_{\mathrm{inv}} = \ln\left(\frac{\text{observed inversion pressure}}{\text{background inversion pressure}}\right)$). Note that pressure is related to height using the equation for hydrostatic balance: $\ln(p/p_o) = -gz/RT$ where p_o is a reference pressure, g is acceleration due to gravity, z is height, R is the specific gas constant and T is temperature. Therefore $\Delta \ln p = -\frac{g}{RT}\Delta z$.

The background and collocated observations have been taken from a population comprising 50 days of data covering Europe over land within the period of 24/12/2006 to 13/10/2008 all valid at midday. During these 50 days an elevated inversion was present in the background approximately 40% of the time, giving a sample of about 1000. Both the radiosonde and background profiles were interpolated vertically to the same high resolution grid with levels separated by 40m within the lowest 3km of the atmosphere.

The error in ln(pressure) ranges from -0.06 to 0.05 (∼-400m to 500m), a positive number (negative distance) indicates the background inversion is higher than that observed. The standard deviation of the error is 0.0162 (∼130m) and the mean of the distribution is -0.008 (∼65m); so the error is biased towards the inversion being too low.

An example of an extreme case where the background temperature inversions height differs from that observed by 400m ($\eta_{\mathrm{inv}} \sim -0.05$) is given in Figure 4.18. This example comes from Christmas 2006 near Belfast when there was widespread fog which was not correctly forecast. This was largely due to the assimilation not being able to utilise the radiosonde observations to correct the inversion height in the forecast. We see in this example that the background and observed inversion structures are very similar and that

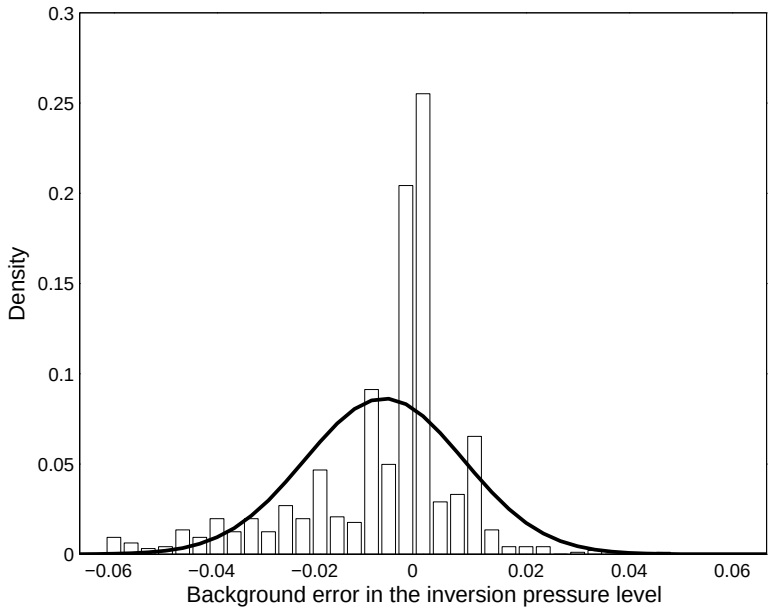


Figure 4.17: A histogram of observation minus background log inversion pressure. The black line gives a Gaussian approximation to this distribution $\sim N(-0.008, 0.0162)$.

a large proportion of the error in the background can be attributed to a positional error.

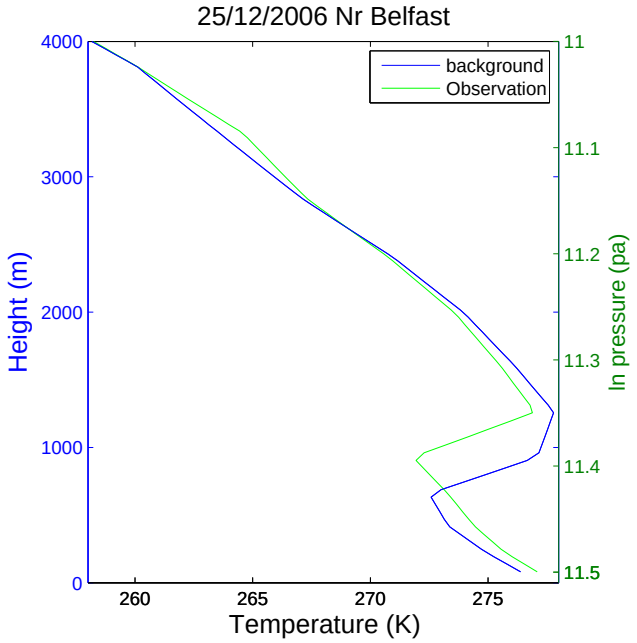


Figure 4.18: Example of radiosonde profile (green) and the equivalent 1D background (blue) profile from Christmas 2006, near Belfast.

The distribution of the errors of a is clearly non-Gaussian and asymmetric. Modelling the distribution of a as Gaussian (black line) underestimates the probability that the difference between the observation and background inversion heights is small in the

assimilation. This would mean that the floating BL scheme would move the levels of the background more often than is necessary. However in the cases when there is a large difference between the observation and background inversion height, the floating BL scheme would be able to move the heights adequately and give an improved analysis. This effect is seen when the background and observed profiles used to produce the statistics in Figure 4.17 are assimilated using both the standard and floating BL scheme. A climatological estimate for the background amplitude errors has been used for simplicity and $\sigma_a = 0.016$. A histogram of the resulting error in analysis inversion pressure level (measured against the observed inversion pressure levels) is shown in the left hand panel in Figure 4.19. As expected there is a higher probability that the standard scheme (blue stems) gives an analysis inversion in very good agreement with that observed (seen as the peak in the central bin). However on average the floating BL scheme (red bars) gives a better inversion height in the analysis as the standard deviation of the errors is lower (0.007 ($\sim$ 55m) compared to 0.009 ($\sim$ 75m)). This is emphasised in the right hand panel of Figure 4.19 where the cumulative density of the absolute error in the analysis pressure level is plotted. In this Figure we see that the cumulative density reaches 1 at a smaller error for the floating BL scheme ($\sim$ 0.02 or 160m) than for the standard scheme ($\sim$ 0.04 or 320m).

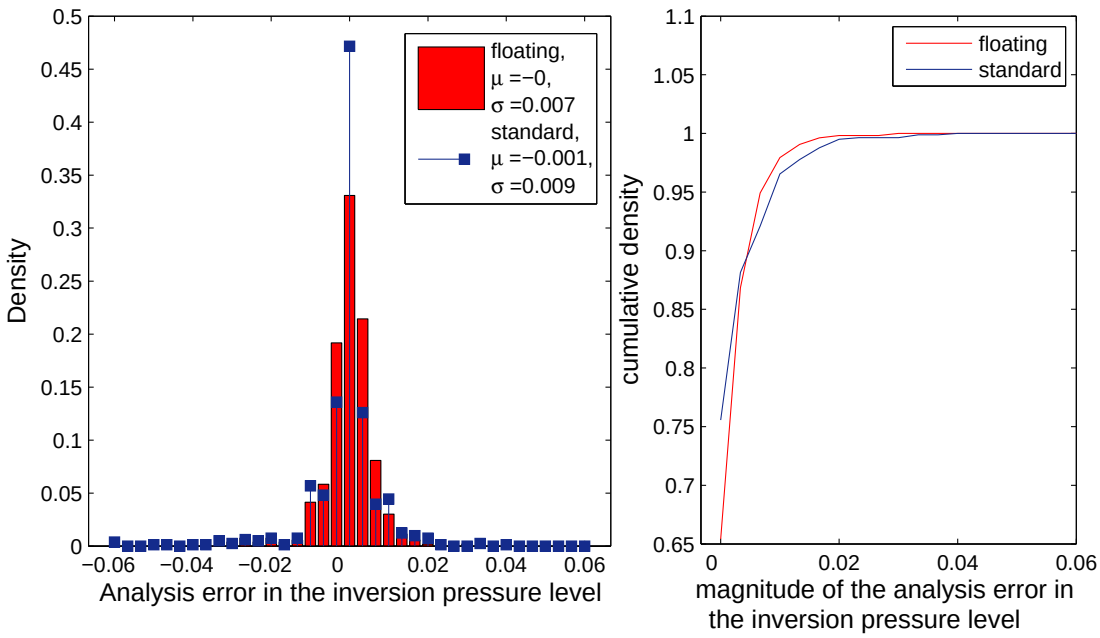


Figure 4.19: Left: A histogram of the error in the analysis inversion pressure level when the floating BL scheme is used (red bars) and when the standard scheme is used (blue stems). Right: Cumulative density as a function of the absolute error in the analysis pressure level when the floating BL scheme is used (red) and when the standard scheme is used (blue).

4.3.1 Allowing for non-Gaussian positional error statistics

A better description of the prior error distribution of a could improve the floating BL scheme. There are two obvious routes which could be taken. The first is the transformation from a to a new control variable to describe the movement of the background levels. This new control variable could be chosen such that its errors are better described by a Gaussian distribution. Work done by Hólm (2002) has shown that any variable, g, can be transformed to a variable, γ, with a Gaussian error distribution by finding a transform, f, such that the cumulative distribution of the probability of g less than a fixed value $\delta\phi$, $P(g \leq \delta\phi)$, equals the cumulative distribution of the probability of γ less than the transformed fixed value $f(\delta\phi)$, $P(\gamma \leq \mathrm{f}(\delta\phi))$.
The second route which could be investigated is to reformulate the cost function for non-Gaussian errors and to simply choose a distribution which more accurately describes the errors seen.
The main down-fall of the current Gaussian distribution for a is that it does not give high enough probability to the event of the background inversion height being in close agreement to that observed. Another incredibly important feature of the distribution are the long tails representing the events when the background height is in a large disagreement with the observation. It is necessary to accurately represent these two characteristics of the error distribution in order for our floating BL scheme to accept these events and to be able to move the background accordingly. These criteria are satisfied by a sum of two Gaussian curves. One Gaussian distribution with a large standard deviation in order to capture the long tails which represent the cases when the background has a large error in the height of inversion. The second Gaussian with a small standard distribution to accurately represent the probability that the background inversion height is good; this will reduce the chance of the floating BL scheme moving the background inversion more often than is necessary. This is similar to ideas given by Andersson and Järvinen (1999) for the description of observation errors used for quality control purposes. The distribution of $\eta_{\mathrm{inv}} = a$ can be given by P_a,

$$P_a = (1 - \mathrm{A})P_1 + \mathrm{A}P_2 \tag{4.13}$$

where P_i is a Gaussian distribution ($N(0, \sigma_i)$) and A is a constant between 0 and 1. An example of this is shown in Figure 4.20, where A=0.5, $P_1 \sim N(0, 0.003)$ and $P_2 \sim N(0, 0.035)$. Note that in using (4.13), we account for the sharp peak and long tails of the known error distribution but not its asymmetry.

The sum of two Gaussians is no longer Gaussian (as seen in Figure 4.20) and so using this distribution to represent the error distribution of a means the cost function is no longer quadratic (even if H were now linear). The background part of the cost function can now be written as the sum of two terms (see appendix D), as shown below:

$$J_B = \frac{1}{2}(\tilde{\mathbf{x}} - \tilde{\mathbf{x}}^{\mathrm{b}})^{\mathrm{T}}\tilde{\mathbf{B}}_{\mathrm{x}}^{-1}(\tilde{\mathbf{x}} - \tilde{\mathbf{x}}^{\mathrm{b}}) - \ln\left\{\frac{1-\mathrm{A}}{\sigma_1}\exp\left\{-\frac{a^2}{2\sigma_1^2}\right\} + \frac{\mathrm{A}}{\sigma_2}\exp\left\{-\frac{a^2}{2\sigma_2^2}\right\}\right\} \tag{4.14}$$

The first term represents an amplitude description of the background profile (as before) and the second term represents the (now non-Gaussian) positional description of the background inversion. This separation is possible due to the assumption that the amplitude

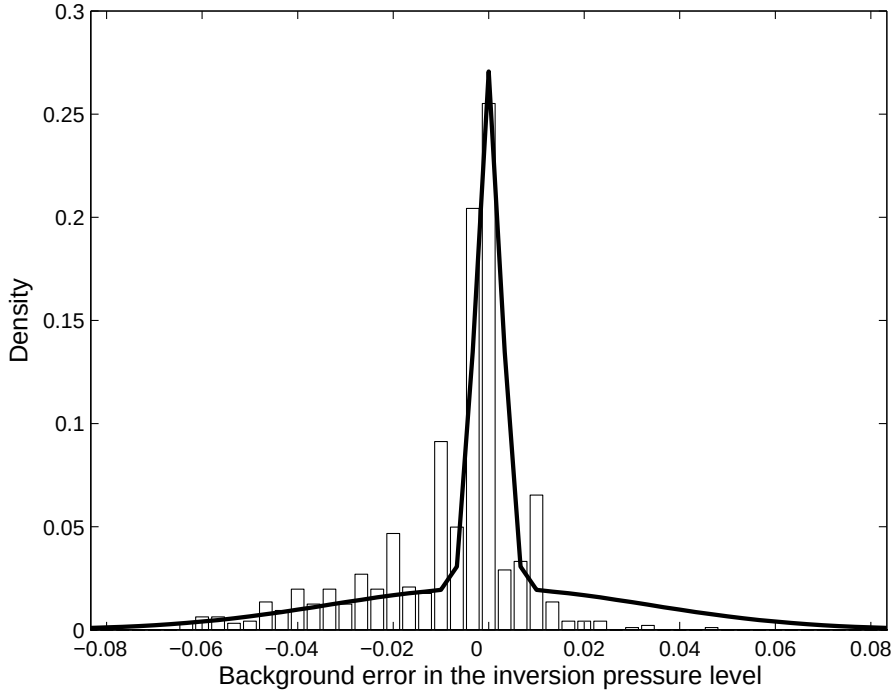


Figure 4.20: The error in the background inversion height approximated as the sum of two Gaussian distributions.

and positional errors are uncorrelated i.e $\tilde{\mathbf{B}}_{\mathrm{v}} = \begin{pmatrix} \tilde{\mathbf{B}}_{\mathrm{x}} & 0 \\ 0 & \sigma_a \end{pmatrix}$.
The minimisation of this new cost function gives significant results, seen when assimilating the same population of backgrounds and observations as used in Figures 4.17 and 4.19. The results are shown in Figure 4.21. In the top left panel, the distribution of the error in the analysis inversion height from the non-Gaussian floating BL scheme (grey bars) is compared to the standard scheme (blue stems). Unlike in Figure 4.19 the two distributions now have a similar probability of the analysis having a very close agreement to the observed inversion height but the floating BL scheme still has a smaller standard deviation in the error (0.006 ($\sim$ 50m) compared to 0.009 ($\sim$ 75m)). In the bottom left panel the non-Gaussian floating BL scheme (grey bars) is compared to the Gaussian floating BL scheme (red stems). The standard deviation of the two distributions are very close (0.006 ($\sim$ 50m) compared to 0.007 ($\sim$ 55m)) but their shapes are very different due to the assumed distributions of the inversion height error in the background. The improvement in the floating BL scheme using these non-Gaussian statistics is also seen in the cumulative distributions (right hand panels). We can still see that the cumulative density reaches 1 for a lower error in the analysis position of the inversion compared to the standard scheme (top, blue). Also the floating BL scheme with non-Gaussian statistics (black) reaches unity at a marginally smaller error than that for the floating BL scheme with the Gaussian statistics (bottom, red).

4.4 Discussion

In this chapter, I have shown some of the aspects which make the floating BL scheme unique. In Subsection 4.1.2 we saw that the floating BL scheme has the effect of increasing the amplitude errors in fixed model space in the region of the inversion in order to make the assimilation more consistent when there is a large positional displacement. It also increases the correlation length scales above the true inversion up to the region where $D(z)$ is non-zero. Using this effective amplitude correlation with the standard 1DVar scheme does not produce an analysis as close to the the truth or as realistic as the floating BL scheme which uses the non-linear observation operator. Therefore the analysis given by the floating BL scheme is advantageous and so a description of the positional error is essential for solving this problem.

The main way this scheme differs from the standard scheme is through the observation operator which has to take into account the floating model levels when calculating $\mathbf{y}^{\text{model}}$ (Section 4.2). The non-linearity of the observation operator can impact on the assimilation in a variety of ways. One way is through the error it introduces when approximating $\mathcal{H}$ through a truncated Taylor expansion. The size of this error can be significant (up to 5K for early iterations) and is dependent upon the presence, location and misplacement (given by a) of the background inversion and, of course, the size of the increment. It is difficult to account for this in $\mathbf{R}$ and so instead it should be noted that the analysis error will be subject to additional uncertainty. We also showed that the analysis error is no longer Gaussian due to the non-linear observation operator and that the mode of the distribution can no longer be assumed to be the mean. The non-linear observation operator has also been shown to alter the shape of J as a function of a. This could potentially lead to difficulty in finding the minimum of J if the observations are ambiguous in the location of the inversion. However, generally the minimum of J is well defined and the floating BL scheme can reproduce the true temperature profile with relatively few observations.

In Section 4.3 the characteristics of the errors of a were calculated by looking at a large population of radiosonde profiles, when an inversion was present, and comparing to the equivalent background profile. This was shown to give a distribution which was non-Gaussian and asymmetric. Assuming that the error was Gaussian meant that the floating BL scheme shifted the background inversion more often than necessary and so even though it did a better job when there was a gross error in the background inversion height, the standard scheme gave a better analysis inversion height when the error was very small. This was improved by describing the error distribution of a as a sum of two Gaussians. This showed that for a significant number of cases the floating BL scheme can give an improved analysis inversion height when there is both a small and large error in the background inversion height.

In the next chapter we shall look at how this scheme can be extended to the multivariate case of temperature and humidity. It is hoped that this scheme will allow for a more consistent analysis of temperature and humidity when there is a positional error in the background inversion, which in turn will give an improved diagnosis of cloud.

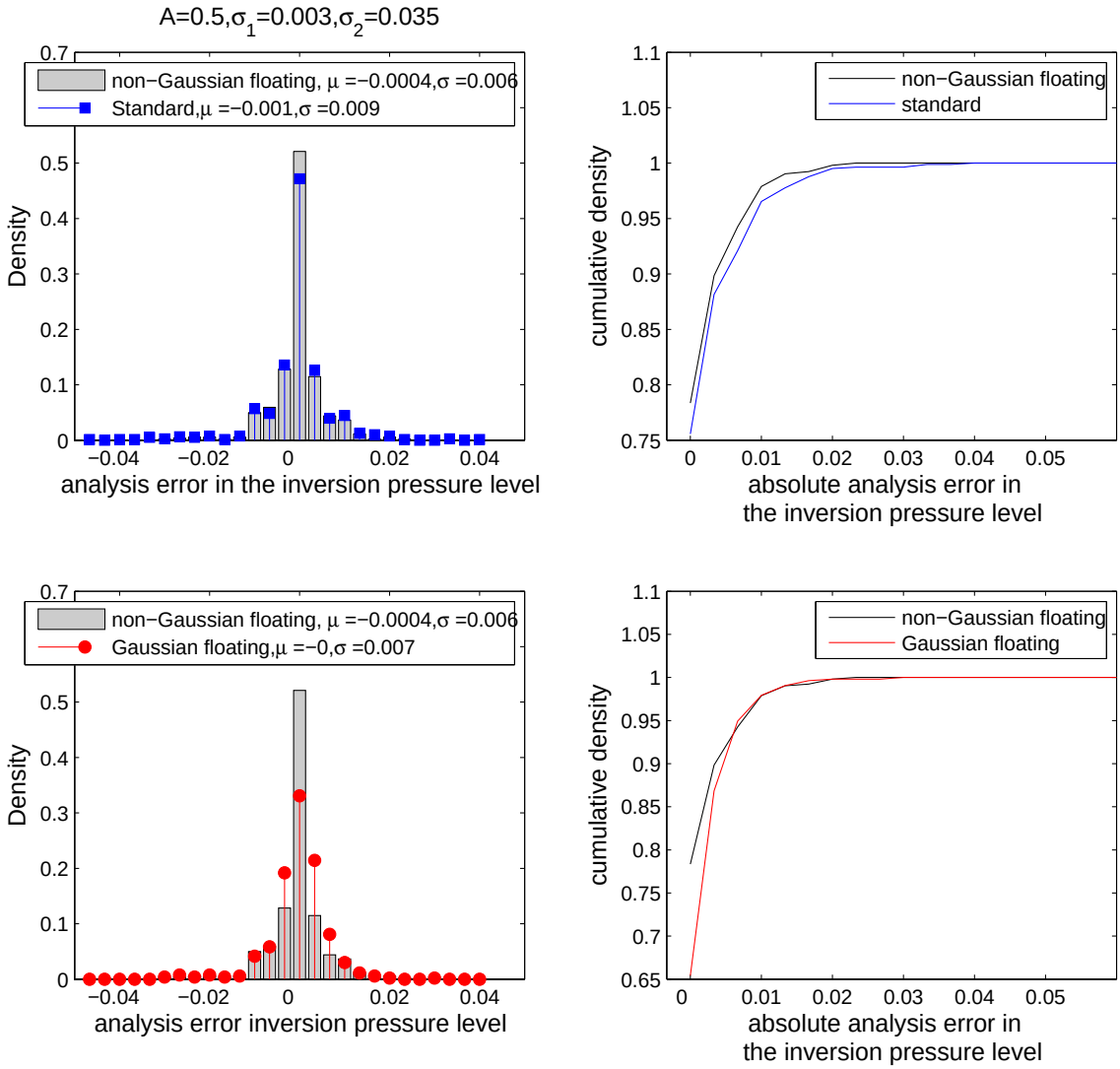


Figure 4.21: Top left: Distribution of the error in the analysis inversion height when the floating BL scheme is used and the error in the background inversion height is approximated as the sum of two Gaussian distributions (grey bars) and the error in the analysis inversion height when the standard scheme is used (blue stems). Top right: Cumulative density as a function of the absolute error in the analysis pressure level when the non-Gaussian floating BL scheme is used (black) and when the standard scheme is used (blue). Bottom left: Grey bars same as top left, red stems are when the floating BL scheme is used and the error in the background inversion height is approximated as Gaussian. Bottom right: Black line same as top right, red line is when the Gaussian floating BL scheme is used.

Chapter 5

Diagnosing cloud

So far, the success of the floating BL scheme has been demonstrated in giving an improved analysis of the BL capping inversion. This improved representation of the temperature profile in the analysis should have a positive impact on the forecast of features directly affected by the the height and strength of the inversion, e.g. the formation of BL clouds and the dispersion of pollutants (see Chapter 1). BL clouds are of particular interest due to their large impact on the propagation of radiation which in turn affects the BL dynamics. However, the diagnosis of clouds is dependent upon humidity as well as temperature. It is therefore intended in this chapter to extend the floating BL scheme to assimilate humidity as well as temperature and to assess the impact the floating BL scheme may have on the diagnosis of clouds.

Firstly we shall determine the requirements for a general assimilation scheme to give an accurate analysis of both temperature and humidity. In the second section we shall look at the behaviour of the floating BL scheme when assimilating both temperature and humidity. In the third section we shall introduce a cloud scheme which may be used to diagnose cloud from temperature and humidity. In the fourth section we shall see how this cloud scheme could be used within the assimilation and then finally in the fifth section we shall compare the diagnosis of cloud given by the floating BL scheme to that given by the standard scheme when the BL capping inversion is misplaced in the background.

5.1 Requirements (multivariate error correlations)

For the diagnosis of cloud it is important we have an accurate analysis of both temperature and humidity. Firstly we shall consider the multivariate assimilation using standard Var as derived in Chapter 2, which accounts for only the background amplitude errors.
In order to assimilate both temperature and humidity, the control vector, $\mathbf{x}$, used within Var to minimise the cost function should be extended to include humidity, so that

$$\mathbf{x} = \begin{pmatrix} \mathbf{T} \\ \mathbf{q} \end{pmatrix}, \tag{5.1}$$

where $\mathbf{T}$ is a vector of temperature on vertical model levels and $\mathbf{q}$ is a vector of humidity (here it is taken to be specific humidity) on the same vertical model levels, so that the

length of the state vector, N, is now twice the number of model levels.
When observations of temperature and/or humidity are assimilated it is important that the analysis increments are consistent in both the temperature and humidity fields. As discussed in Chapter 2, one way of enforcing this in the assimilation, is through multivariate correlations in the B-matrix. These multivariate correlations should be related to known physical relationships between temperature and humidity.

Temperature and the saturation level of humidity are related through the Clausius-Clapyeron relation (e.g. Rogers, 1979):

$$\frac{\mathrm{d}e_s}{\mathrm{d}T} = \frac{L_v e_s}{R_v T^2}, \tag{5.2}$$

where e_s is the saturation water vapour pressure, T is temperature, L_v is the latent heat of evaporation ($2.5 \times 10^6 Jkg^{-1}$) and R_v is the specific gas constant of water vapour ($461.51 Jkg^{-1}K^{-1}$). If we approximate latent heat to be constant, (5.2) can be integrated to give:

$$\ln\frac{e_s}{e_{s0}} = \frac{L_v}{R_v}\left(\frac{1}{T_0} - \frac{1}{T}\right), \tag{5.3}$$

where e_{s0} is the value of saturation vapour pressure at a reference temperature, T_0. This shows that the saturation vapour pressure increases exponentially as temperature increases, shown in Figure 5.1 (e_{s0} = 611Pa and T_0 = 273K), i.e. as the air is warmed it can hold more water vapour.

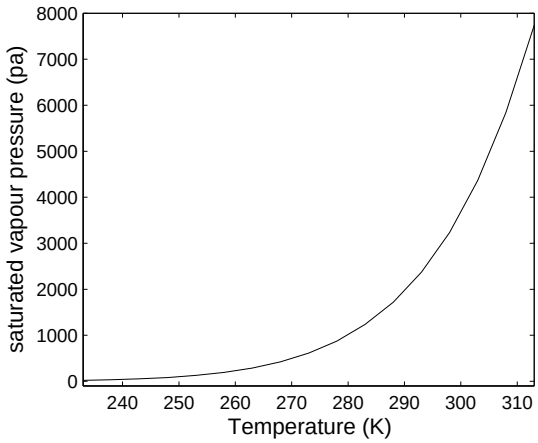


Figure 5.1: Saturated vapour pressure versus temperature.

Relative humidity is the percentage ratio of vapour pressure, e, to the saturated vapour pressure:

$$RH = 100\frac{e}{e_s} \qquad [\%]. \tag{5.4}$$

Therefore as temperature increases and e_s increases (but e remains fixed), relative humidity decreases. Therefore for a fixed vapour pressure, temperature and relative humidity are negatively correlated.
Specific humidity is defined as the mass of water vapour, m_v, per unit mass of moist air,

m.

$$q = \frac{m_v}{m} = \frac{\frac{R_d}{R_v} e}{p - e\left(1 - \frac{R_d}{R_v}\right)} \qquad \text{[kg/kg]}, \tag{5.5}$$

where V is the volume of moist air, T is the temperature, R_d is the specific gas constant of dry air ($287.05 \text{Jkg}^{-1}\text{K}^{-1}$) and p is the pressure of moist air Specific humidity is unaffected by a change in temperature unless the air is saturated. Then a decrease in temperature causes the water vapour to condense out, and m_v to decrease. At this point cloud will form and q will decrease. According to this process specific humidity and temperature are positively correlated in regions of cloud.

The correlation between temperature and specific humidity may also be affected by local sources and sinks of water vapour (Berre, 2000). For example in regions where saturated air is advected in from another region. In this case, vapour within this advected air may condense (decreasing q) and release latent heat leading to an increase in T locally. Therefore for this particular example T and q are negatively correlated. Another case that may lead to a correlation between temperature and specific humidity errors is in a region of subsiding air. In this case T and q may be negatively correlated as dry descending air warms adiabatically.

Multivariate error correlations for the background between temperature and specific humidity are highly flow dependent. The atmospheric state, including the synoptic situation and particular sources and sinks of water vapour, all affect the way errors in temperature and specific humidity are correlated.

We can study the nature of these multivariate correlations by looking at the spread in an ensemble of forecasts as described in Section 2.3.

$$\mathbf{B} = \left\langle \left(\mathbf{T} - \langle \mathbf{T} \rangle\right) \left(\mathbf{q} - \langle \mathbf{q} \rangle\right)^{\mathrm{T}} \right\rangle \tag{5.6}$$

This has been done for two examples at midday when the BL is capped by a temperature inversion (see Figures 5.2b and 5.3b). In each example, the BL is capped by a region of high relative humidity (see Figure 5.2d and Figure 5.3d) indicating the presence of cloud. The local correlation between temperature and specific humidity for a given height is shown in Figure 5.2a and in 5.3a.

In both examples, we see that at the top of the BL, temperature and specific humidity are positively correlated. This is consistent with water vapour condensing at high values of RH as the temperature cools leading to a decrease in q. In Figure 5.2a there is also a significant amount of positive correlation throughout the whole depth of the BL. This example is located off the east coast of Ireland and so the positive correlation may be explained by the surface acting as a source of water vapour and a source of heating. The second example, Figure 5.3, comes from over land in South Wales and we see that at the surface there is a small negative correlation between temperature and specific humidity.

At the inversion, in both cases, the correlation between temperature and specific humidity switches sign to become strongly negatively correlated with a value of -1. This is due to the simultaneous warming and drying of the air as we move into a different air mass; the amount of warming that occurs at this level defines the strength of the inversion which in turn controls the depth of the transition between moist to dry air. Therefore when an

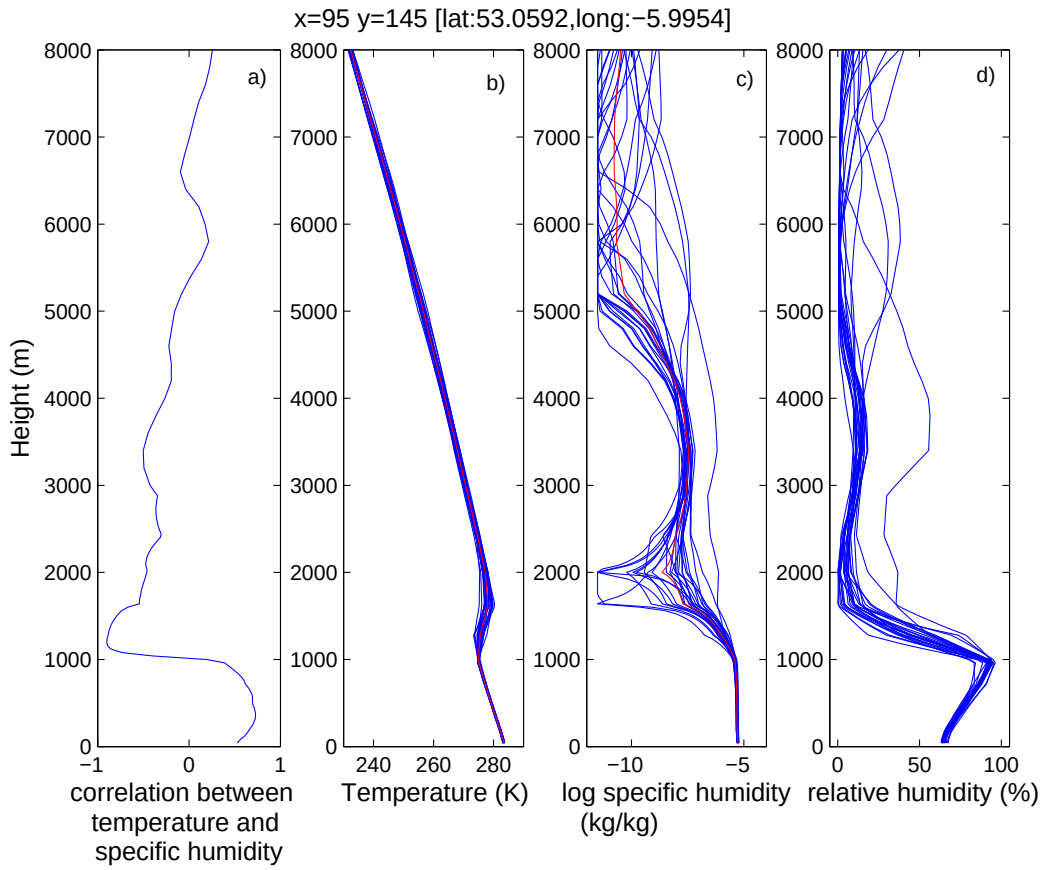


Figure 5.2: Irish sea: a) Forecast error correlation between temperature and specific humidity at the same height. b) Ensemble of temperature profiles. c) Ensemble of log specific humidity profiles. d) Ensemble of relative humidity profiles derived from the temperature and humidity profiles.

ensemble member is warmer than average at this level it will also be drier than average. Well above the inversion there is much less water vapour and the correlations are much weaker, fluctuating around zero. This is similar to the result given by Berre (2000) for the mean vertical cross-covariance between specific humidity and temperature estimated using a spectral approach and from a set of forecast differences.

In Figures 5.4 and 5.5 the background error correlation matrices are given between temperature and specific humidity at different heights. The profile given in Figure 5.2a is the diagonal of the matrix given in 5.4 and the profile given in Figure 5.3a is the diagonal of the matrix given in 5.5. The multivariate error correlation structures given by these two matrices are very different emphasising the flow-dependence of the background error correlations between temperature and specific humidity. In Figure 5.4 the correlations between temperature and specific humidity are larger than in Figure 5.5, especially within the BL. This difference is most likely related to an increase in the moisture flux over the sea.

Above the BL there are also major differences between Figures 5.4 and 5.5. This emphasises the sensitivity of the correlation between errors in temperature and specific humidity

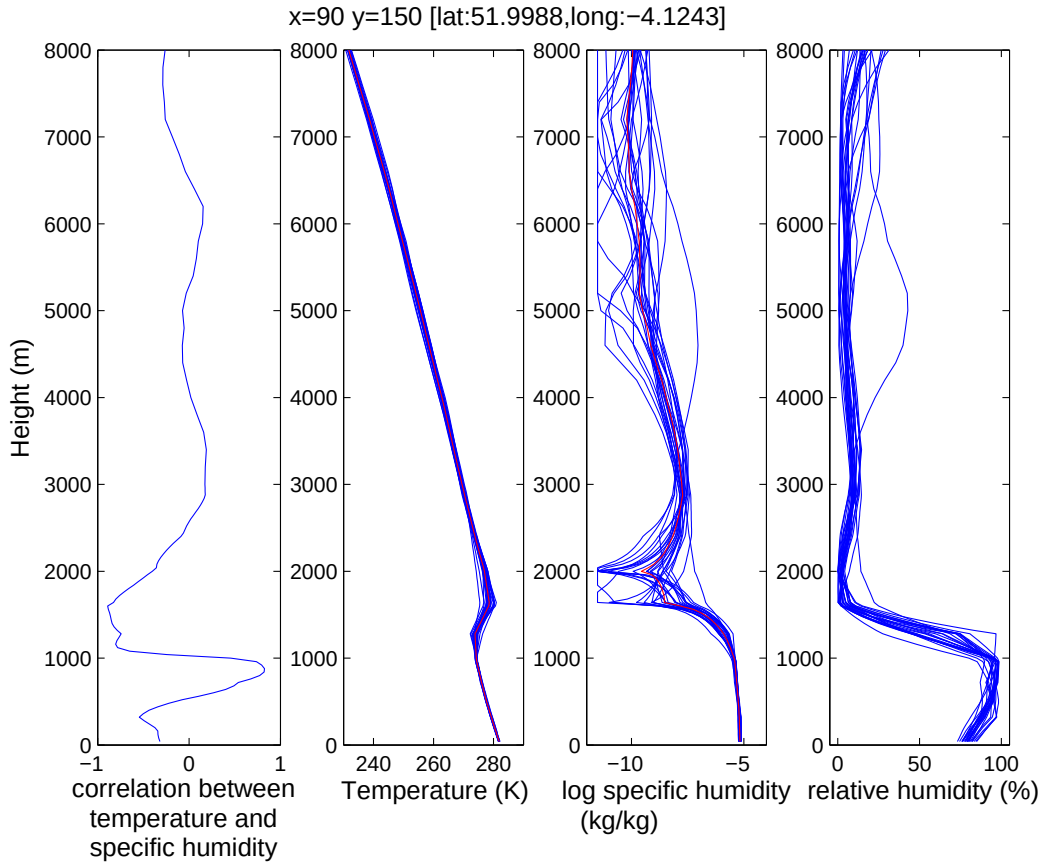


Figure 5.3: South Wales: a) Forecast error correlation between temperature and specific humidity at the same height. b) Ensemble of temperature profiles. c) Ensemble of log specific humidity profiles. d) Ensemble of relative humidity profiles derived from the temperature and humidity profiles.

to the meteorology as previously discussed. It is difficult to identify reasons for these differences without knowing more about the synoptic situation.

If there is a large positional error in the height of the BL capping inversion in the background, then this too would be a source of correlation between temperature and humidity. We can see from the profiles in Figure 5.2 and Figure 5.3 that a sharp positive gradient in temperature at the top of the BL is mirrored by a sharp negative gradient in specific humidity. Therefore a positional error in the background capping inversion is also likely to be accompanied by a positional error in the humidity feature in the same region. For example if the temperature inversion is too high then the region between it and the true inversion will be too cool and too moist; and so a positional error should cause a negative correlation in this region. Similarly if the background inversion is too low, then the region between the background and the true inversion will be too warm and too dry so again there will be a negative correlation. This is what it seen at the top of the BL in the left hand panels of Figures 5.4 and 5.5, however a greater positional error (more variability in the height of the BL top between ensemble members) would lead to a larger region of negative correlation.

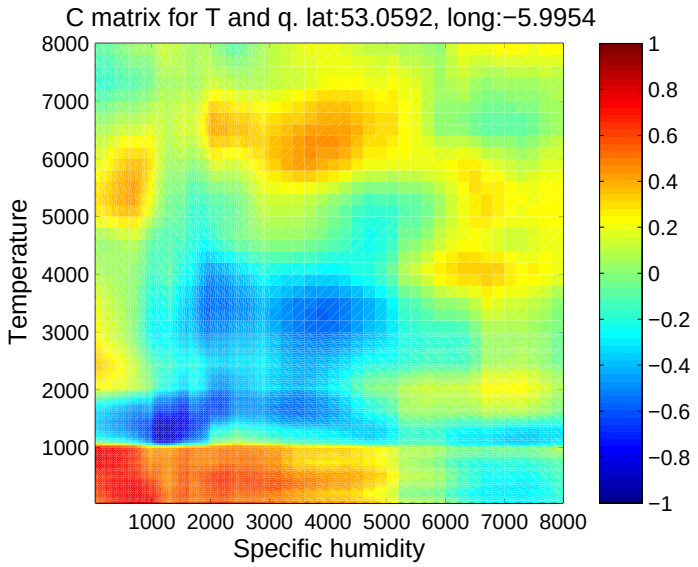


Figure 5.4: Irish Sea: Matrix of multivariate correlations between temperature and specific humidity. Axes are given as altitude (m).

By building a flow-dependent model of the multivariate covariances, we wish to recreate some of these more coherent multivariate correlation structures between temperature and humidity to give a consistent analysis between the two variables. Firstly, in Subsection 5.2, we will look at the effect the floating BL scheme has on the multivariate correlations when information is available about both temperature and humidity. We shall then introduce a cloud scheme into the assimilation to build in more structure to the multivariate covariances.

5.2 Multivariate application of the floating boundary layer scheme

Extending the floating BL scheme to incorporate variables other than temperature is straightforward. All variables can be assimilated on the same floating model levels so that the variable a is still the only extra variable needed to describe the positional error. The control vector for the floating BL scheme may now be given as:

$$\tilde{\mathbf{v}} = \begin{pmatrix} \tilde{\mathbf{x}} \\ a \end{pmatrix} = \begin{pmatrix} \tilde{\mathbf{T}} \\ \ln \tilde{\mathbf{q}} \\ a \end{pmatrix}. \tag{5.7}$$

The natural logarithm of specific humidity is used instead of straightforward specific humidity as it is easier to model its errors as Gaussian (q must be positive, but $\ln q$ is not bounded, it also allows us to model the errors of q when they vary by orders of magnitude, as is seen in the vertical).

Figure 5.6 shows results from an idealised experiment where synthetic observations of T and $\ln q$ are taken from a known true profile. The true profile (black) is given for temperature (left) and log specific humidity (right). These true profiles have been generated from the control profile given in Figure 5.2 (Irish sea) with an additional shift of the inversion

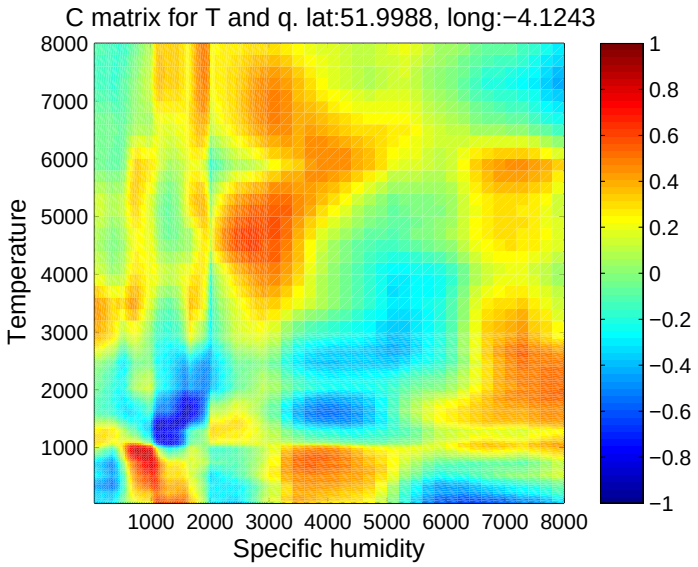


Figure 5.5: South Wales: Matrix of multivariate correlations between temperature and specific humidity. Axes are given as altitude (m).

by 266m. From this truth, background and observation profiles of temperature and specific humidity have been simulated in a similar manner to that described in Section 3.2, taking into account their respective errors. In this case the amplitude part of $\tilde{\mathbf{B}}_{\mathrm{v}}$ (see (3.4) for a reminder of this part of the B-matrix) has been calculated from the spread in the ensemble given in Figure 5.2, however, the multivariate correlations between temperature and log humidity have been taken to be zero so that $\tilde{\mathbf{B}}_{\mathrm{v}}$ is block diagonal. An analysis has then been calculated given these background and observations as input using the standard scheme (magenta dashed) and the floating BL scheme (red).
Using the floating BL scheme (red profile) with these inputs a is minimised to -270m and so gives a very good fit to the true temperature and specific humidity profiles. This can be compared to the standard scheme (magenta dashed). As already seen (in Chapters 2 and 4) the standard scheme does not reproduce the true temperature profile well when there is a positional error in the background. We also see that the standard scheme struggles to correctly shift the feature in the humidity profile associated with the inversion.

When there are no multivariate correlations, as is the case here, the standard scheme treats the assimilation of temperature and humidity completely separately. However, the floating BL scheme performs the assimilation of temperature and humidity both on the same distorted grid, controlled by the variable a. Therefore the floating BL scheme implicitly introduces correlation between temperature and humidity on a fixed grid where $D(z)$ is non-zero. We can appreciate this by performing the assimilation using the floating BL scheme when observations are only available of temperature or humidity.

Idealised univariate example with only T or q observations

In Figure 5.7, the assimilation has been performed again but with only observations of temperature. The analysis (red) gives a good fit to the true profiles of both temperature and humidity despite only temperature being observed. This is because a large proportion

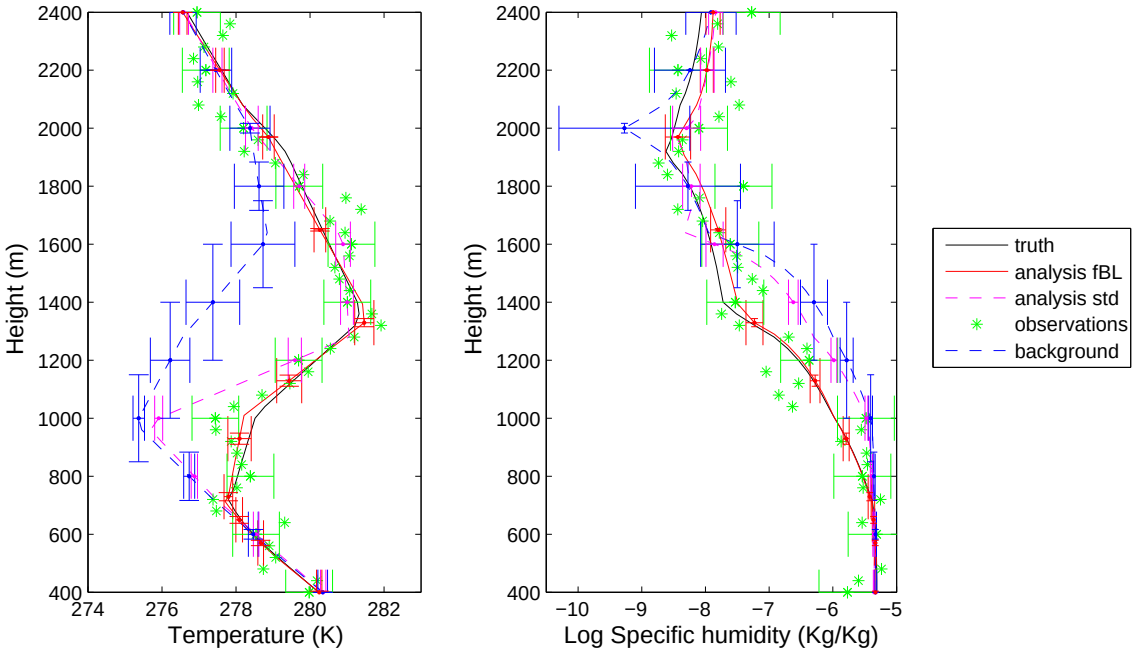


Figure 5.6: Profiles of temperature (left) and log specific humidity (right) are given for the truth (black), background (blue), observations (green). With these background and observations as input, analyses have been calculated using the floating BL scheme (red) and the standard scheme (dashed magenta). B-matrix is block diagonal (no multivariate correlations). Using the floating BL scheme a has been minimised to -270m.

of the background error comes from an error in the vertical positioning of the temperature inversion which causes an error in the humidity profile. A similar picture is given in Figure 5.8 when only observations are available of humidity. Because humidity observations are associated with a larger amplitude error, the feature in humidity is not as well defined and so it is more difficult to get an accurate value of a: in this case a is slightly too large ($a = -290$m compared to the actual shift of -266m).

The correlation introduced between temperature and humidity using the floating BL scheme is dependent upon the choice of D and the final value of a. The implied cross-correlation in fixed model space can be calculated in a similar way to that described in Subsection 4.1.2, by specifying a transform matrix which maps $\tilde{\mathbf{B}}_{\mathrm{v}}$ to $\mathbf{B}_{\mathrm{x}}$. In Figure 5.9 the implied cross-correlations have been plotted for the two examples given in Figures 5.7 and 5.8. For these examples D is positive between 550m and 2050m. In both cases the cross-correlation is strong and negative in this region as expected.

It should be noted that in order to minimise the cost function in the floating BL scheme the variation of a now affects the movement of features in all variables, not just temperature. A typical humidity profile associated with a temperature inversion has a sharp drop in humidity just above the inversion (see Figures 5.2 and 5.3). It is therefore important that if the temperature inversion is shifted this feature in the humidity profile should also be shifted. This may mean that the floating BL scheme may not perform as well as when it

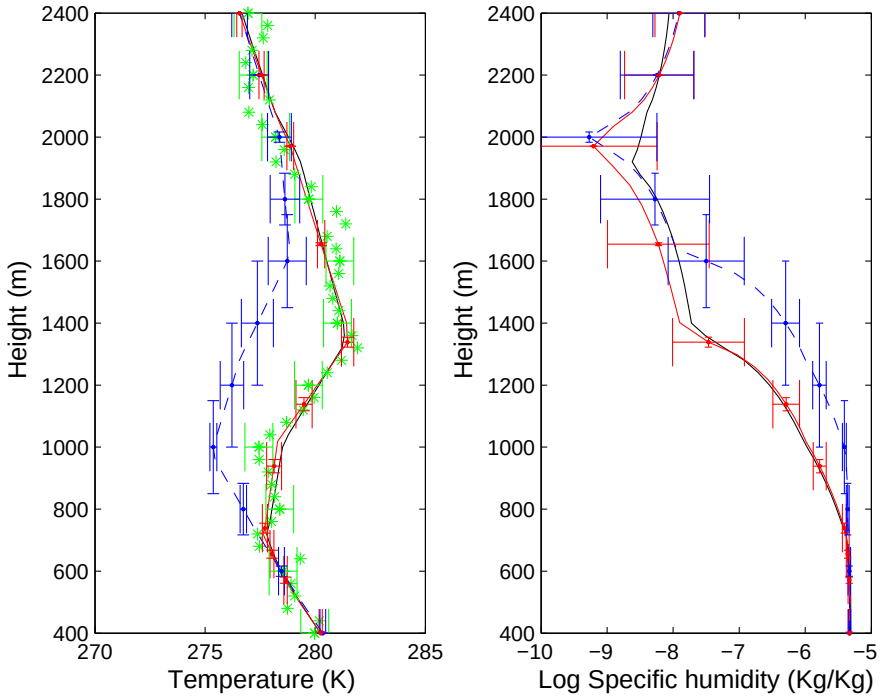


Figure 5.7: As in Figure 5.6 but with no observations of humidity. a has been minimised to -261m.

is only assimilating temperature or the extra variable may help to constrain the problem. In the example given in Figure 5.6 (T and q observations), a was minimised to -270m which is almost identical to the actual displacement of -266m. In Figure 5.7 when observations of temperature alone were assimilated a was minimised to -261m, suggesting that (at least in this idealised example) the additional information given by the observations of specific humidity has not affected greatly the success of the floating BL scheme.

In this section we have seen that the floating BL scheme introduces implicit cross-correlations between temperature and humidity when there is a displacement in the inversion height in the background. These implied multivariate correlations are important for keeping temperature and humidity consistent in the analysis when there is a positional error in the background inversion height. When cloud is forming within the BL we would also expect temperature and specific humidity to be positively correlated. This correlation which is associated with the formation of cloud cannot be described by the floating BL scheme on its own. To describe this correlation a simple model of cloud formation is necessary. In this project we have decided to use the cloud scheme developed by Smith (1990).

5.3 Introduction to the Smith cloud scheme

Currently liquid cloud is diagnosed within the Met Office's Unified Model (UM) using the layer cloud scheme based on that of Smith (1990), referred to from now on as the 'Smith

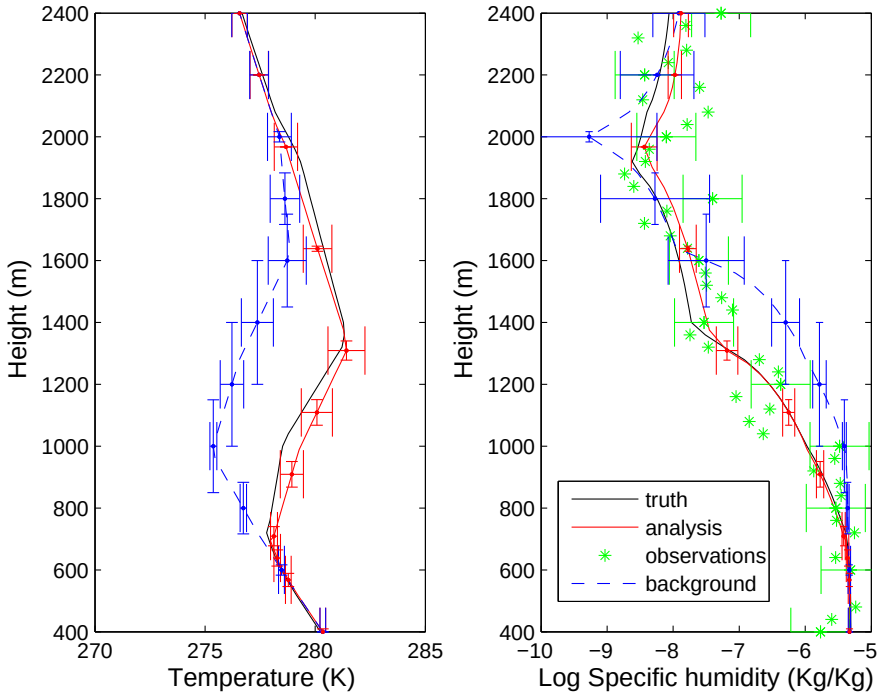


Figure 5.8: As in Figure 5.6 but with no observations of temperature. a has been minimised to -290m.

cloud scheme'[1]. This scheme generates a cloud amount and water content based on a statistical distribution of thermodynamic and water content variables about their grid box mean. The statistical distribution is meant to represent sub-grid scale variability, allowing for parts of an individual grid box to be saturated and other parts not. It reformulates the model's BL mixing in terms of 'cloud conserved variables', variables which are conserved during changes of state of water. These variables are liquid-frozen temperature, T_L, and total water content, q_t, given by:

$$T_L = T - (L_c/c_p)q_L - \{(L_c + L_F)/c_p\}q_F, \tag{5.8}$$

where T is temperature, L_c is latent heat of condensation, L_F is latent heat of fusion, c_p is the specific heat capacity, q_L is specific liquid water content, q_F specific ice content, and

$$q_t = q + q_L + q_F. \tag{5.9}$$

We can simplify (5.8) to take into account only liquid clouds:

$$T_L = T - (L_c/c_p)q_L = T - (L_c/c_p)q_c. \tag{5.10}$$

where q_c is layer cloud water content. T_L can therefore be interpreted as the temperature minus the effect of the heat produced as water vapour condenses. Similarly q_t can be

[1]The key equations of the scheme are given here without derivation. The reader is referred to the original paper for derivation of details

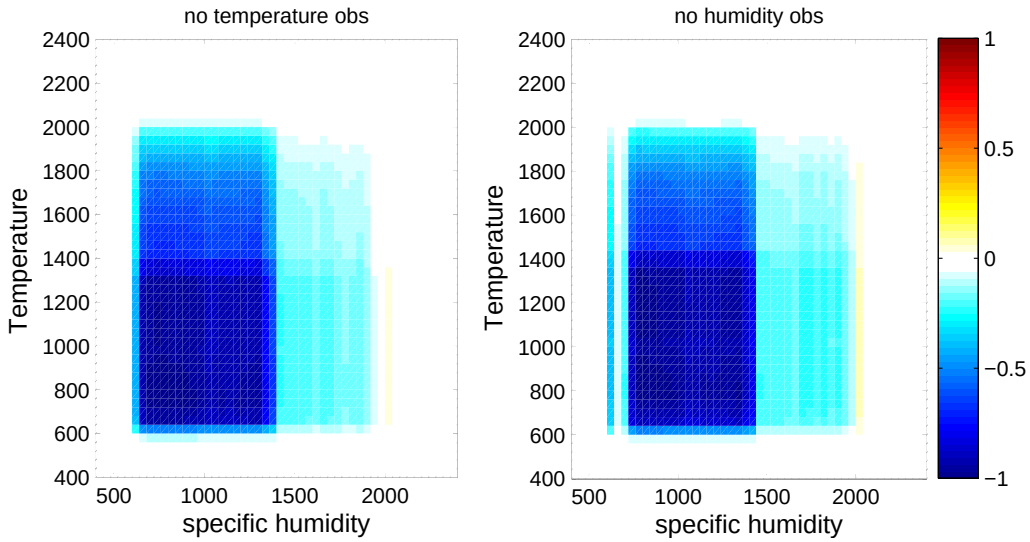


Figure 5.9: Implied correlation between temperature and humidity given by the floating BL scheme when only humidity observations are available (left) and when only temperature observations are available (right). Axes are given as altitude (m).

defined as

$$q_t = q + q_c. \tag{5.11}$$

The grid box mean cloud fraction, C, is given by the fraction of the grid box with cloud water content greater than zero. In the Smith scheme it is calculated as:

$$C = \int_{-Q_c}^{\infty} G(s)ds, \tag{5.12}$$

where Q_c is the grid box mean difference between specific total water content and saturation specific humidity ($q_t - q_{sat}$) with s the local deviation from the mean, Q_c. The grid box mean cloud water content, q_c, is given by

$$q_c = \int_{-Q_c}^{\infty} (Q_c + s)G(s)ds. \tag{5.13}$$

G is a distribution function which was chosen by Smith (1990) to be a symmetric triangular function. The assumption that cloud amount and water content can be assumed to be distributed in a certain manner within a grid box was devised by Sommeria and Deardorff (1977) and Mellor (1977). These subgrid scale fluctuations which are described by G are due not only to small scale turbulence but also to mesoscale effects. Modifications have been made to this scheme for the assimilation of cloud variables in the MetO system by Sharpe (2007).

Sharpe (2007) replaces the triangular distribution for G by a sech^2 distribution

$$G(s) = \frac{\text{sech}^2\left(\frac{s}{2\delta}\right)}{4\delta} \tag{5.14}$$

$$\delta = \frac{\sigma_s \sqrt{6}}{4}, \tag{5.15}$$

where σ_s is the standard deviation of s, which is a function of the critical relative humidity, RH_C, (the minimum value of RH at which cloud may occur, this should be dependent upon the resolution of the model- in this case RH_C had been set to 80%[2]), the saturated specific humidity, q_{sat}, and $\alpha_L = \frac{\partial q_{sat}}{\partial T}$ (see appendix E). $G(s)$ is now a smooth and analytically integrable function and so the scheme is differentiable, although the infinite domain of the sech^2 distribution is not strictly physical. Equations (5.14) and (5.15) now allow equations (5.12) and (5.13) to be expressed analytically

$$C = 1 - e^{\frac{-q_c}{\delta}} \tag{5.16}$$

$$q_c = Q_c + \delta \ln\left(e^{\frac{-Q_C}{\delta}} + 1\right) \tag{5.17}$$

$$Q_c = a_L(q_t - q_{sat}(T_L)), \tag{5.18}$$

where $a_L = \left(1 + \frac{L\alpha_L}{c_p}\right)^{-1}$. It is now possible using (5.10), (5.11), (5.16), (5.17) and (5.18) to calculate q_t and T_L from T and q iteratively and hence to diagnose cloud fraction and cloud water content (the full equations are given in appendix E). Examples of this are given in Figures 5.10 and 5.11.

We see in each case, in the right hand panel, that cloud is diagnosed below the inversion (between 500m and 1000m for Figure 5.10 and between the surface and 1000m for Figure 5.11) and is bounded by the height of the inversion. The values for T and T_L are very similar throughout the profile but T is slightly larger than T_L in the region of the cloud due to $(L_c/c_p)q_c$ being positive as latent heat is released. Likewise, the values for q and q_t are very similar throughout the profile but q_t is slightly larger than q in the region of the cloud due to the addition of water condensate. The difference between T and T_L is proportional to the amount of cloud water present and so in Figure 5.11 when a larger fraction of cloud is diagnosed compared to Figure 5.10, we see a bigger difference between the profiles for the model variables, T and q, and the cloud conserved variables, T_L and q_t.

As well as allowing us to diagnose cloud we will later see that changing to cloud conserved variables may be useful to the floating BL scheme as θ_L and q_t are conserved under adiabatic vertical motion when precipitation is not present (Lock et al., 2000). The conservation of variables during the changes in pressure caused by the vertical shift is important for ensuring that the floating BL scheme does not have an adverse effect on the analysis as discussed in Subsection 3.1.4.

5.4 Including cloud conserved variables (θ_L and $\ln q_t$) in Var

We wish to perform the assimilation in terms of cloud conserved variables which also need to be adiabatically conserved for the floating BL scheme. For these reasons we choose

[2]In the U.M. the value of RH_C varies with height due to the varying resolution. This would be important if the floating BL scheme were made operational as RH_C should be a function of a, the variable which controls the shifting levels and hence the vertical grid size around the inversion height.

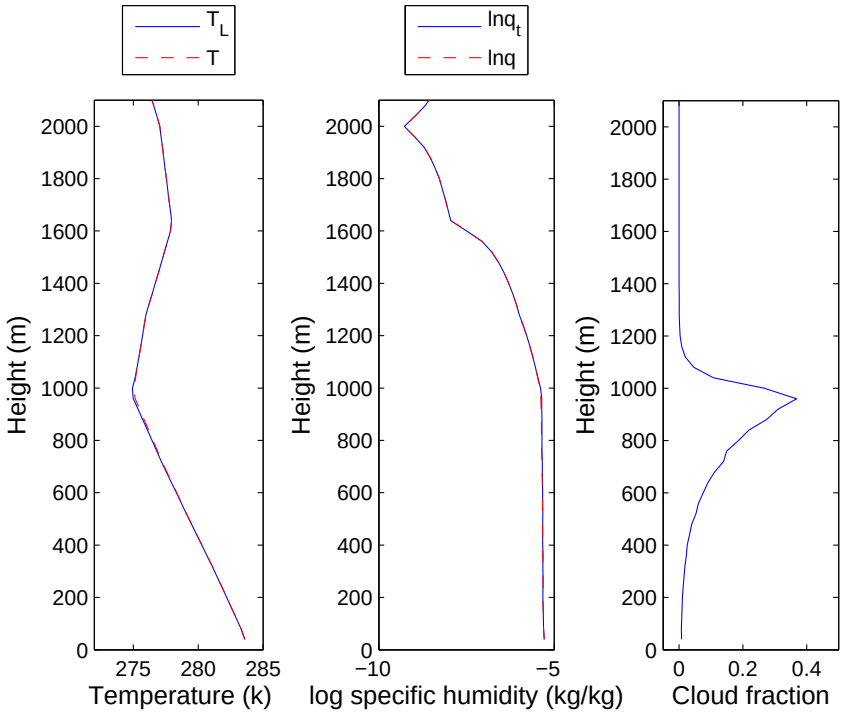


Figure 5.10: Irish sea: Left: temperature (red) and frozen liquid temperature (blue). middle: specific humidity (red dashed) and total water (blue). Right: mean grid box cloud fraction calculated using the smith cloud scheme.

to minimise the cost function in terms of the new control vector $\mathbf{w} = \begin{pmatrix} \boldsymbol{\theta}_L \\ \ln \mathbf{q}_\mathrm{t} \end{pmatrix}$, for the standard scheme and $\tilde{\mathbf{w}} = \begin{pmatrix} \tilde{\boldsymbol{\theta}}_L \\ \ln \tilde{\mathbf{q}}_\mathrm{t} \\ a \end{pmatrix}$, for the floating BL scheme (as in Chapter 3, the tilde overbar indicates that the equations are stored on floating levels whose heights are controlled by a). A change of variables can be implemented by introducing the Smith cloud scheme as a transform, $\mathcal{U}$, between the cloud conserved variables and our model variables, $\mathbf{x} = \begin{pmatrix} \mathbf{T} \\ \ln \mathbf{q} \end{pmatrix}$ for the standard scheme or $\tilde{\mathbf{v}} = \begin{pmatrix} \tilde{\mathbf{T}} \\ \ln \tilde{\mathbf{q}} \\ a \end{pmatrix}$ for the floating BL scheme.

5.4.1 The U-transform

Defining the tangent linear estimate of the Smith cloud scheme at the n^{th} iteration as $\mathbf{U}_n$, an increment in the new control variables, $\delta\mathbf{w}_n$, can now be calculated for the standard scheme as (see (2.18))

$$\delta\mathbf{w}_n = (\mathbf{B}_\mathrm{w}^{-1} + \mathbf{U}_n^\mathrm{T}\mathbf{H}_n^\mathrm{T}\mathbf{R}^{-1}\mathbf{H}_n\mathbf{U}_n)^{-1}(\mathbf{B}_\mathrm{w}^{-1}\delta\mathbf{w}_n^\mathrm{b} + \mathbf{U}_n^\mathrm{T}\mathbf{H}_n^\mathrm{T}\mathbf{R}^{-1}\delta\mathbf{y}_n), \tag{5.19}$$

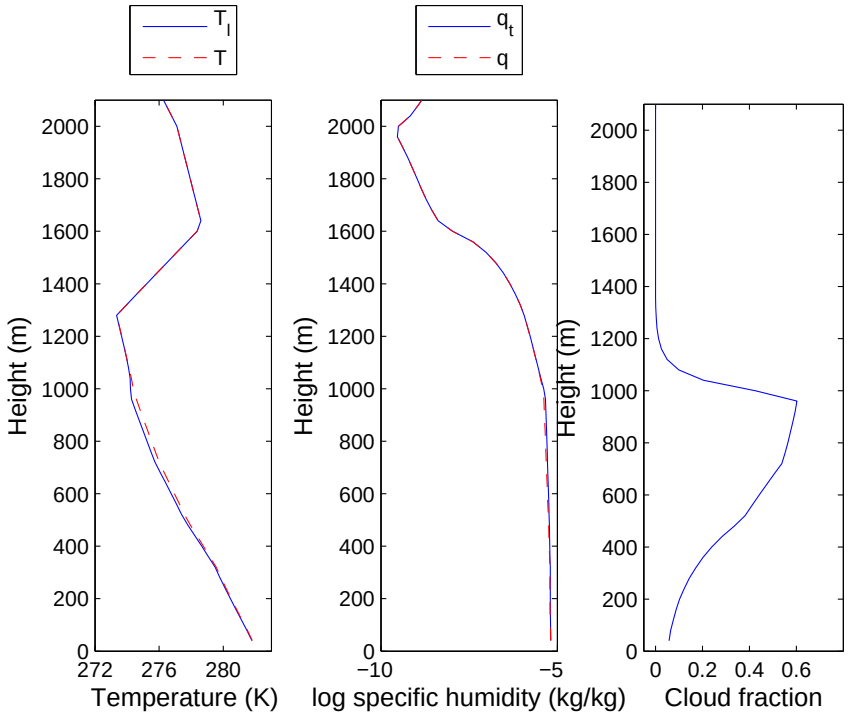


Figure 5.11: South Wales: Left: temperature (red) and frozen liquid temperature (blue). middle: specific humidity (red dashed) and total water (blue). Right: mean grid box cloud fraction calculated using the smith cloud scheme.

and for the floating BL scheme as

$$\delta\tilde{\mathbf{w}}_n = (\tilde{\mathbf{B}}_{\mathrm{w}}^{-1} + \mathbf{U}_n^{\mathrm{T}}\mathbf{H}_n^{\mathrm{T}}\mathbf{R}^{-1}\mathbf{H}_n\mathbf{U}_n)^{-1}(\tilde{\mathbf{B}}_{\mathrm{w}}^{-1}\delta\tilde{\mathbf{w}}_n^{\mathrm{b}} + \mathbf{U}_n^{\mathrm{T}}\mathbf{H}_n^{\mathrm{T}}\mathbf{R}^{-1}\delta\mathbf{y}_n), \tag{5.20}$$

where $\delta\mathbf{w}_n$ is the difference between $\mathbf{w}$ at iteration n and the previous iteration. $\mathbf{B}_{\mathrm{w}}$ is the background error covariance matrix for the cloud conserved variables. $\delta\mathbf{w}_n^{\mathrm{b}}$ is the difference between the current state guess of the cloud conserved variables, $\mathbf{w}_{n-1}$, and the background in control vector space[3]. Other symbols are defined in Chapter 2.

The transform $\mathcal{U}$ is non-linear and so as with the linearised observation operator $\mathbf{H}$, $\mathbf{U}$ is only valid when applied to small increments, for the standard assimilation this can be expressed as

$$\delta\boldsymbol{\theta}_L, \delta\ln\mathbf{q}_t \quad \overset{\mathbf{U}}{\rightarrow} \quad \delta\mathbf{T}, \delta\ln\mathbf{q} \ . \tag{5.21}$$

The structure of $\mathbf{U}$ for the standard scheme is given in Figure 5.12. Each block is a diagonal matrix ($\frac{N}{2} \times \frac{N}{2}$). This can be extended to the floating BL scheme by allowing for the extra element a as shown in Figure 5.13. The evaluation of each of these partial derivatives is given analytically in appendix E.

An example of the diagonal elements of each of the four main blocks of $\mathbf{U}$, shown in Figure 5.12, are shown in Figure 5.14. These have been calculated for the temperature and humidity profiles given in Figure 5.10. In the top left hand panel we see that $\frac{\delta T}{\delta\theta_L}$

[3] $\delta\mathbf{w}_n^{\mathrm{b}} = \delta\mathbf{w}_{n-1}^{\mathrm{b}} - \delta\mathbf{w}_{n-1}$

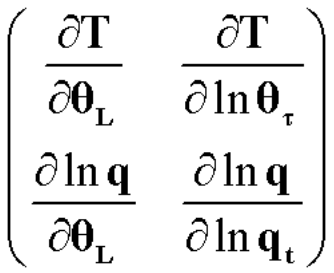


Figure 5.12: The U-transform for the standard scheme.

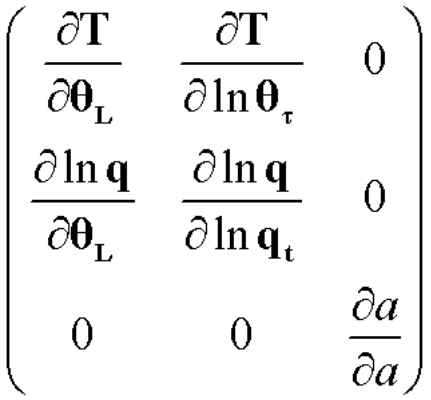


Figure 5.13: The U-transform for the floating BL scheme.

is positive throughout this region, i.e. a change in T is proportional to a change in θ_L. Outside the region of cloud the gradient of $\frac{\delta T}{\delta \theta_L}$ is negative due to the effects of adiabatic cooling on T but not on θ_L. Outside cloud $\frac{\delta T}{\delta \theta_L}$ is equal to $\left(\frac{p}{p_o}\right)^{R/c_p}$, where p_o is a reference pressure. At the level of the cloud we see the effect of latent heating on T which does not affect θ_L, and so $\frac{\delta T}{\delta \theta_L}$ drops but remains positive[4].
In the bottom right hand panel we see the effect that cloud has on $\frac{\delta \ln q}{\delta \ln q_t}$. Outside the region of cloud the values of $\ln q$ and $\ln q_t$ are identical and so the value of $\frac{\delta \ln q}{\delta \ln q_t}$ is 1. At the level of the cloud we see a drop in this value as $\ln q_t$ is conserved but $\ln q$ decreases[5].
In the other two panels $\frac{\delta T}{\delta \ln q_t}$ (top right) and $\frac{\delta \ln q}{\delta \theta_L}$ (bottom left) are given. We see that in regions outside the cloud $\ln q$ is unaffected by changes in θ_L and T is unaffected by changes in $\ln q_t$. However, in the region of cloud a decrease in θ_L will lead to a decrease in $\ln q$ due to a decrease in θ_L occurring when cloud is formed which also means a decrease in the specific humidity as some of the water vapour is converted to condensate. Similarly, an increase in $\ln q_t$ will lead to an increase in T due to the increase in total water in a cloud primarily being due to an increase in cloud condensate which releases latent heat.

In (5.19) we have had to approximate $\mathcal{U}(\mathbf{w}_{n+1}) - \mathcal{U}(\mathbf{w}_n)$ by $\mathbf{U}\delta\mathbf{w}_n$. In a similar way to the error caused by linearisation of our observation operator in Subsection 4.2.1, the linearisation of the Smith cloud scheme, $\mathcal{U}$, may also be a source of error. We can evaluate the magnitude of this error for an increment that would be necessary to shift an inversion in the background down by 270m. The results are shown in Figure 5.15. In the left hand

[4] $\frac{\delta T}{\delta \theta_L} = \frac{\delta T}{\delta T_L}\frac{\delta T_L}{\delta \theta_L} = \left(\frac{p}{p_o}\right)^{R/c_p}\left[1 + \frac{L}{c_p}\frac{\delta q_c}{\delta T_L}\right]$, outside cloud $\frac{\delta q_c}{\delta T_L}$=0, inside cloud $\frac{\delta q_c}{\delta T_L} < 0$.
[5] $\frac{\delta \ln q}{\delta \ln q_t} = \frac{\delta \ln q}{\delta q}\frac{\delta q}{\delta q_t}\frac{\delta q_t}{\delta \ln q_t} = \frac{1}{q}\left[1 - \frac{\delta q_c}{\delta q_t}\right] q_t$, outside cloud $\frac{\delta q_c}{\delta q_t}$=0, inside cloud $\frac{\delta q_c}{\delta q_t} > 0$.

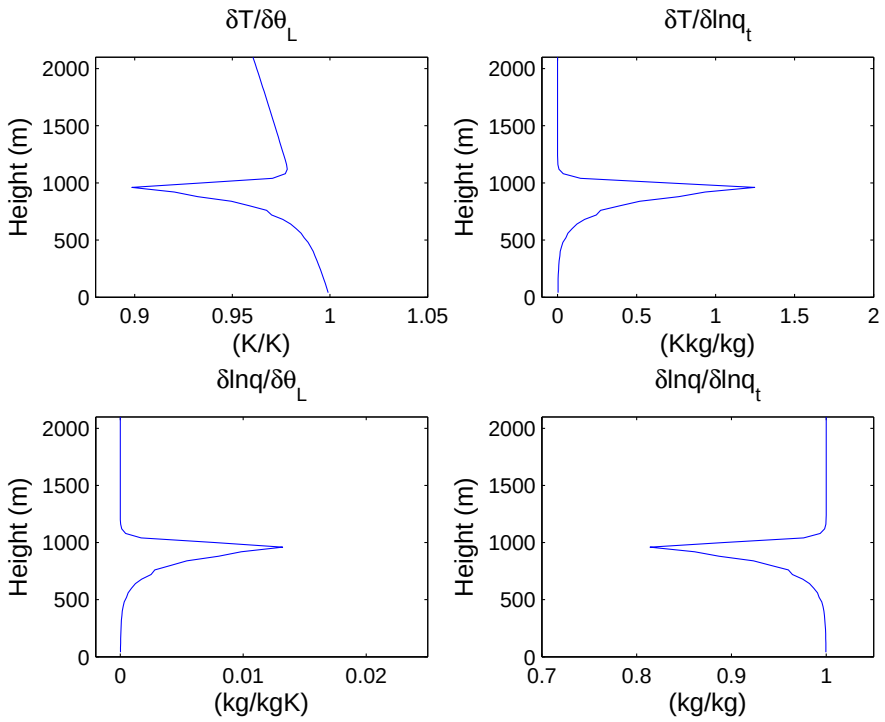


Figure 5.14: U-transform for the example shown in Figure 5.10.

panels of Figure 5.15 profiles of temperature (top) and humidity (bottom) are given for $\mathcal{U}(\mathbf{w})$ at iteration n (blue) and $n+1$ (black). These provide test profiles to examine the linearisation. Also plotted is the linearised estimate $\mathcal{U}(\mathbf{w}_n) + \mathbf{U}\delta\mathbf{w}_n$ (black dashed). The discrepancy between $\mathcal{U}(\mathbf{w}_{n+1})$ and $\mathcal{U}(\mathbf{w}_n) + \mathbf{U}\delta\mathbf{w}_n$ (error due to linearisation) is plotted in the middle panels. The peak in the error comes at the height of the cloud level at iteration n and is more significant for temperature than humidity relative to the typical variation of each quantity.

In the bottom right hand panel the cloud fraction derived from each of these temperature and humidity fields is given. There appears to be very little difference between the cloud calculated from $\mathcal{U}(\mathbf{w}_{n+1})$ (black) and $\mathcal{U}(\mathbf{w}_n) + \mathbf{U}\delta\mathbf{w}_n$ (black dashed) due to the error occurring above the cloud level at iteration $n+1$. Assuming that $\mathbf{x}$ converges to the truth and $\delta\mathbf{w}_n$ is small this linearisation error can be assumed to have little impact on the accuracy of the analysis and its error statistics.

5.4.2 Creating a B-matrix in terms of cloud conserved variables

A B-matrix for the cloud conserved variables, $\mathbf{B}_{\mathrm{w}}$, necessary to solve (5.19), can be approximated in a similar way to those created in Section 2.3 using an ensemble of forecasts. Each of the ensemble members given in Figure 5.2 (Irish sea case) has been converted to the cloud conserved variables from the model variables using the Smith cloud Scheme. This then gives an ensemble of θ_L and $\ln q_t$ forecasts. The spread in this ensemble can then be equated to the background error covariances given by $\mathbf{B}_{\mathrm{w}}$.

$$\mathbf{B}_{\mathrm{w}} = \left\langle (\mathbf{w} - \langle \mathbf{w} \rangle)(\mathbf{w} - \langle \mathbf{w} \rangle)^{\mathrm{T}} \right\rangle. \tag{5.22}$$

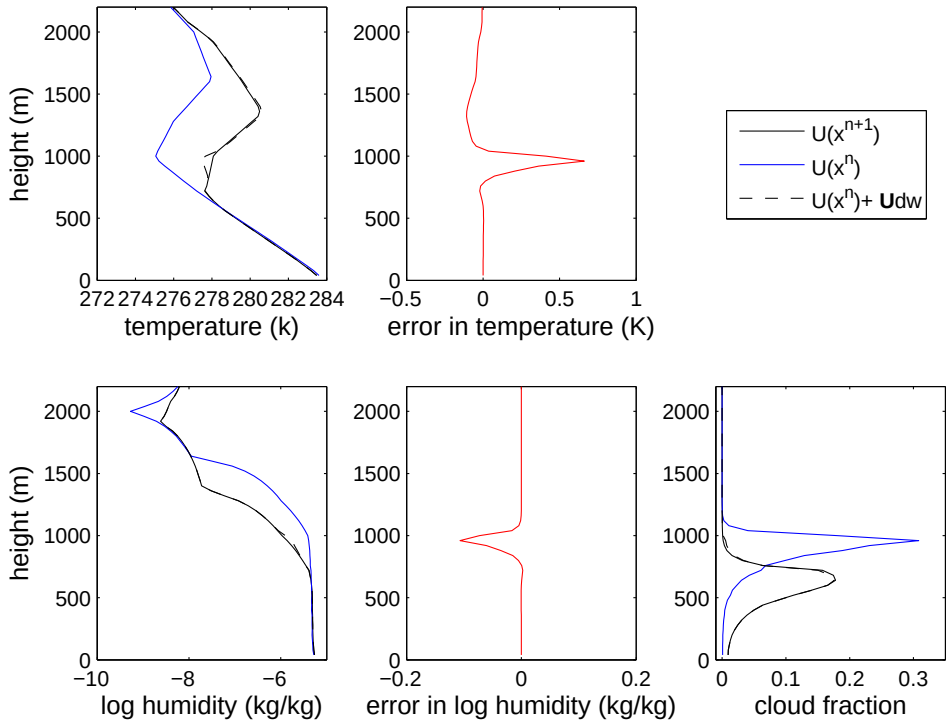


Figure 5.15: Profiles of temperature (top left) and log specific humidity (bottom left) given by $\mathcal{U}(\mathbf{w}_n)$ (blue) for iteration n and by $\mathcal{U}(\mathbf{w}_n + \delta\mathbf{w})$ (black) for iteration $n+1$. T and $\ln q$ are also approximated for iteration $n+1$ by $\mathcal{U}(\mathbf{w}_n) + \mathbf{U}\delta\mathbf{w}$ (black dashed). The difference between $\mathcal{U}(\mathbf{w}_n) + \mathbf{U}\delta\mathbf{w}$ and $\mathcal{U}(\mathbf{w}_n + \delta\mathbf{w})$ is plotted in the middle panels for temperature (top) and log specific humidity (bottom). In the bottom right hand panel cloud has been diagnosed for each of the 3 cases.

Figure 5.16 illustrates the background error for the cloud conserved variables. Both the size of the covariances (left hand Figure) and the structure of the correlations (right hand Figure) are similar to those for the model variables, T and $\ln q$. The region of strong positive correlation for temperature within the BL is still a strong feature and so there is still an appreciable amount of flow dependence and a need for accurately derived flow-dependent error covariances.

5.4.3 The benefits of using cloud conserved variables

A change of variables via a control variable transform may be done for a variety of reasons. In Chapter 2 we attempted to model the covariance structure given when a strong inversion is present through the use of a simple BL model. In this case it was hoped that the errors in the new control variables would be less flow dependent as they would be insensitive to the presence of the inversion. From Figure 5.16 we can see that the errors for the cloud conserved variables are still dependent upon the presence of the inversion and so are just as flow dependent as T and q.

A second reason a control variable transform may be used is to model some aspects of error

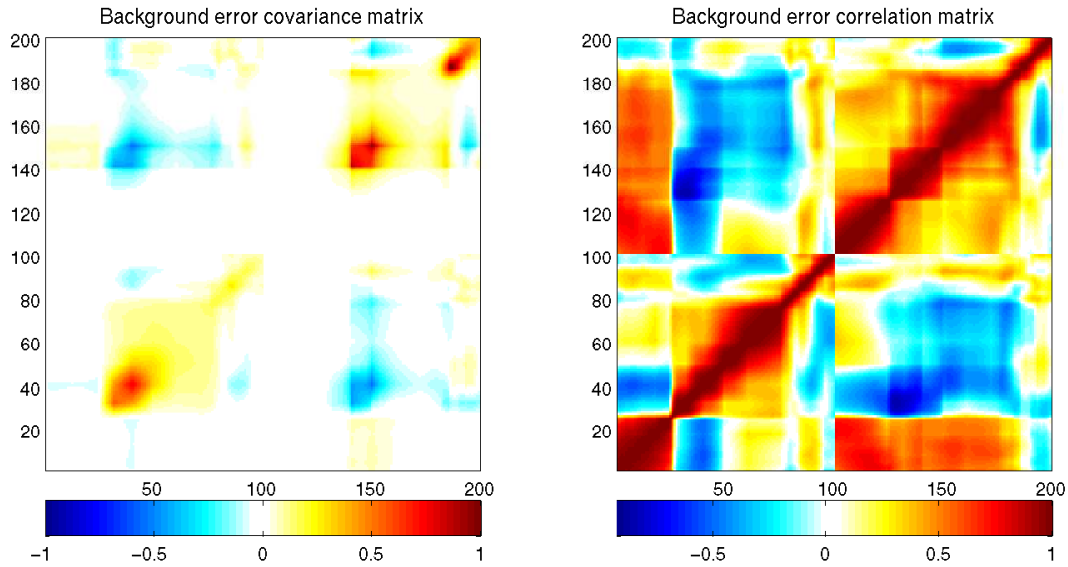


Figure 5.16: Left: The full background error covariance matrix for θ_L (elements 1:100) and $\ln q_t$ (elements 101:200). Right: The full background error correlation matrix for θ_L and $\ln q_t$. Axes are given by index of $\mathbf{x}$.

covariances but not necessarily all of the flow-dependent aspects. In this case we would expect cloud conserved variables to be less sensitive to the presence of cloud allowing us to model part of the correlation between temperature and humidity in the region of cloud. An example of the error correlation matrix for the cloud conserved variables minus that of the model variables for the Irish sea case is seen in Figure 5.17. Plotted separately are the univariate differences for the temperature variables (top left) and humidity variables (top right) and the multivariate differences between the temperature and humidity variables (bottom left). Generally at the cloud layer the magnitude of the correlations are greater for the model variables than for the cloud conserved variable (blue). This implies that the correlations, both univariate and multivariate, are less dependent on the presence of cloud for cloud conserved variables than for the model variables.

This lack of dependency on the presence of cloud means that the U-transform can be used to imply the necessary correlations between T and $\ln q$ for a consistent assimilation between temperature and humidity in the presence of cloud. This can be demonstrated by assuming the B-matrix for the cloud conserved variables is block diagonal, i.e. θ_L and $\ln q_t$ are uncorrelated. This has been done for two examples; Irish sea, left hand panel of Figure 5.18 and South Wales, left hand panel of Figure 5.19. We can then transform this matrix to model space to look at the implied covariances, so that implied $\mathbf{B}_\mathrm{x} = \mathbf{U}\mathbf{B}_\mathrm{w}\mathbf{U}^\mathrm{T}$. This has been performed for each case and plotted in the middle panels of Figure 5.18 and Figure 5.19.

In both cases we see that the implied error correlations between temperature and humidity are non-zero, despite being zero in $\mathbf{B}_\mathrm{w}$, due to the U-transform. These non-zero values occur at the level of the cloud and are positive as expected. In the second example, South Wales, the implied correlations appear more significant due to a larger region of

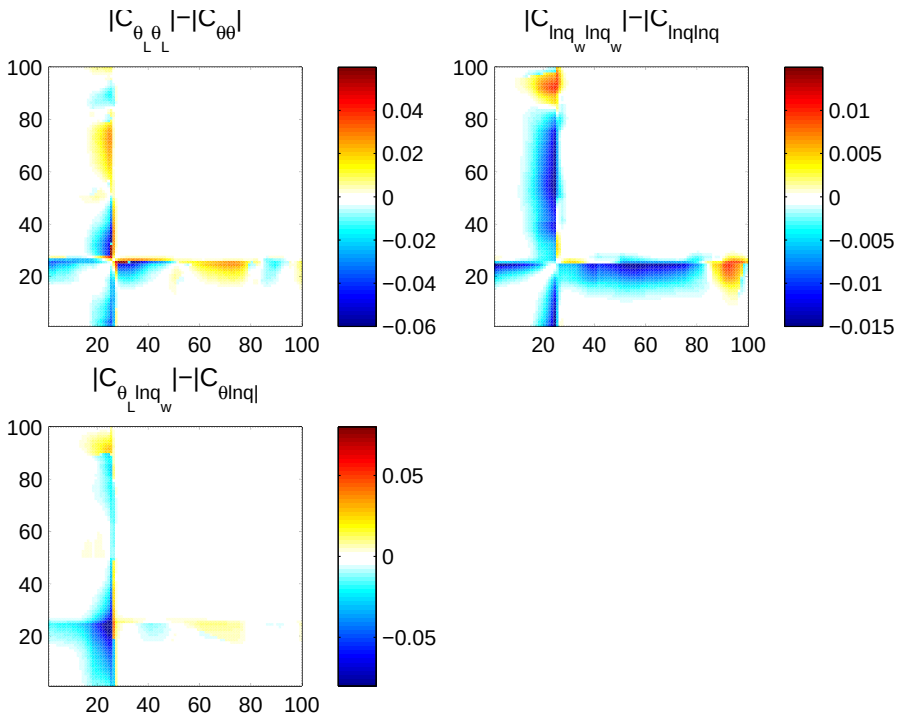


Figure 5.17: The difference in the magnitude of correlations between errors for cloud conserved variables and the model variables. Axes are model levels.

cloud and a greater cloud fraction compared to the Irish sea case, as seen in Figure 5.3. These implied multivariate correlations can be compared to those calculated using the ensemble method in the right hand panels of Figure 5.18 and Figure 5.19. There is obviously a lot of structure within the correlations of T and $\ln q$ that the introduction of the cloud model has been unable to capture, especially for the Irish sea case. However for the diagnosis of cloud, it is the strong signal produced at the cloud level that we are most interested in.

It is now possible to investigate the effect that the cloud scheme has on the assimilation when there is a disagreement in the height of the inversion between the truth (measured by the observations) and background. In this instance we would expect the cloud layer associated with the inversion to also be misplaced in the background.
In the next section we shall look at this problem with the standard scheme with the cloud scheme and then we shall compare this to the floating BL scheme with the cloud scheme. We will then look at real data from a case study and assess if the floating BL scheme has the potential to assimilate real data in the presence of cloud.

5.5 Comparisons between standard 1DVar and the floating BL scheme for the diagnosis of cloud

We can study the effect of the Smith cloud scheme on the standard assimilation scheme, for the background (blue) and observations (green) given in Figure 5.6, simulated from

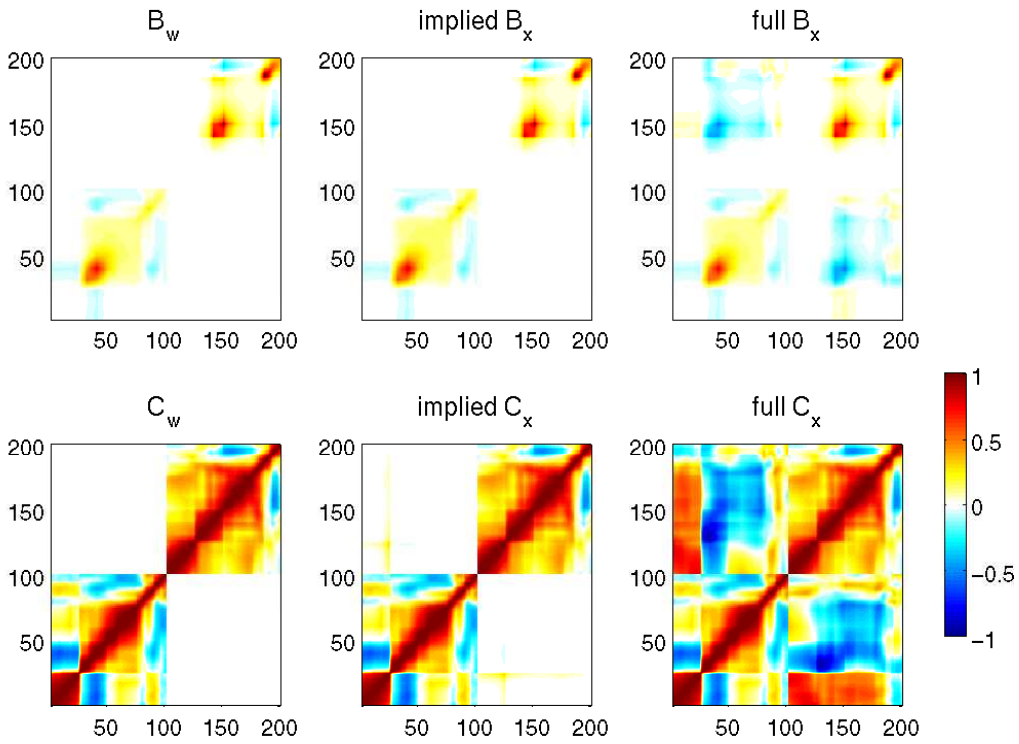


Figure 5.18: The effect of $\mathbf{U}$ on the implied background errors for the model variables, Irish sea example. Left: Background error covariances (top) and correlations (bottom) for the cloud conserved variables, assumed to be block diagonal. Middle: Background error covariances (top) and correlations (bottom) for the model variables implied by $\mathbf{U}\mathbf{B}_{\mathrm{w}}\mathbf{U}^{\mathrm{T}}$. Right: Full background error covariances (top) and correlations (bottom) for the model variables. Axes are the index of the control vector.

the Irish sea case. In Figure 5.20 the analysis given by the standard assimilation scheme without the cloud scheme (magenta dashed) is given for temperature (left) and log specific humidity (middle). This can be compared to the analysis given by the standard scheme when the cloud scheme is included via the U-transform described in the last section (red). For temperature, the profiles are very similar, but the humidity is more moist when the cloud scheme is used. A possible explanation is the large positive T increment and the positive background error correlation between T and $\ln q$, in the region of cloud, so that the multivariate B-matrix would act to prevent a simultaneous large warming and drying of the background.

In the right hand panel of Figure 5.20 cloud has been diagnosed using the Smith cloud scheme for each profile. The true cloud profile is given in black and the background cloud profile is given in blue. We see that the effective lowering of the inversion removes cloud at higher levels. In both cases, with and without the cloud scheme, the analysis has decreased the cloud fraction but has not shifted its height.

This can be compared to the effect the cloud scheme has on the floating BL scheme given in Figure 5.21. This time the effect of the cloud scheme on both the temperature (left) and humidity (middle) is marginal. The level of implied positive correlation between T

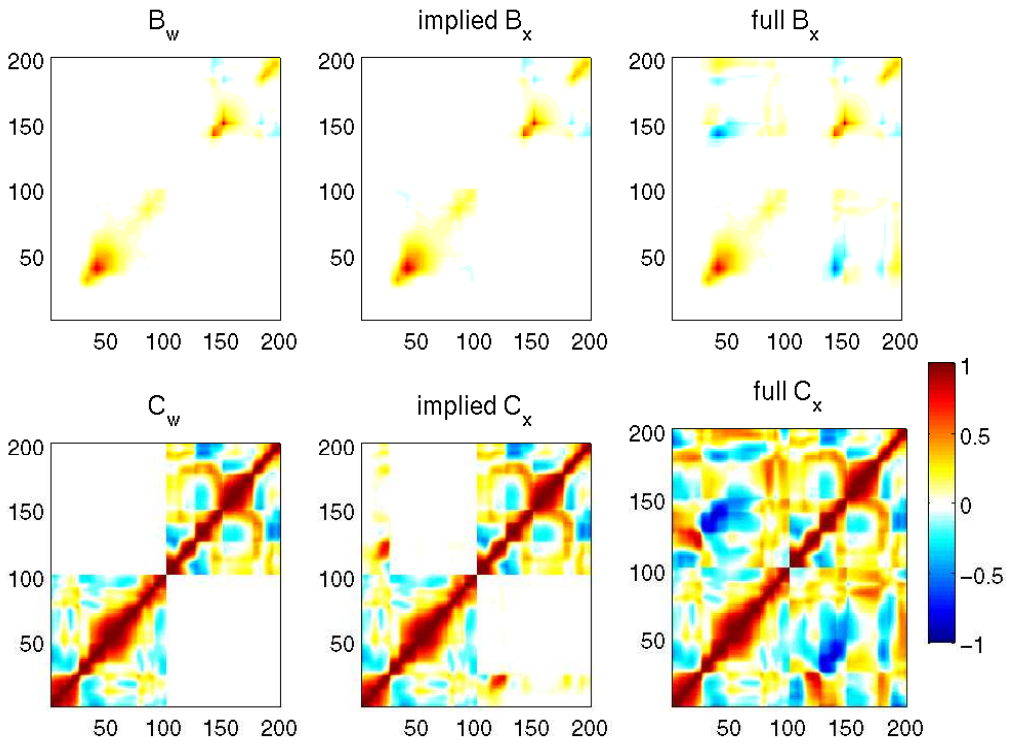


Figure 5.19: As in Figure 5.18 except for the South Wales example.

and $\ln q$ will have shifted, in this case, down with the inversion in the background. At the shifted level of the inversion there is little difference in amplitude between the background and observation values so that the implied multivariate correlations have less of an impact. In the right hand panel of Figure 5.21 the cloud fraction diagnosed from these profiles is given. Again there is little difference between the cases where the cloud scheme is (red) and is not (magenta) included, however when the cloud scheme is included it gives a slightly better fit to the true cloud profile. In both analyses the cloud has successfully been shifted down to remain within the BL.

5.5.1 Case study

We finish this chapter by looking at a real case study, using radiosonde measurements of temperature and relative humidity and the equivalent background profile of temperature and relative humidity. A radiosonde profile from 25/12/2006 in the Czech Republic has been chosen. The analysis chart of constant surface pressure contours for this date is given in Figure 5.22. The location of this radiosonde profile (marked by the blue dot) is close to the centre of a high pressure system.

In Figure 5.23 profiles of temperature (left) and log specific humidity (middle) are plotted for the radiosonde observations (green) and the background (blue). We see that the temperature inversion in the background is about 200m lower than that observed and that the humidity profile in the background is also too dry, compared to the observations, at the level of the background inversion. In the right hand panel of Figure 5.23 cloud

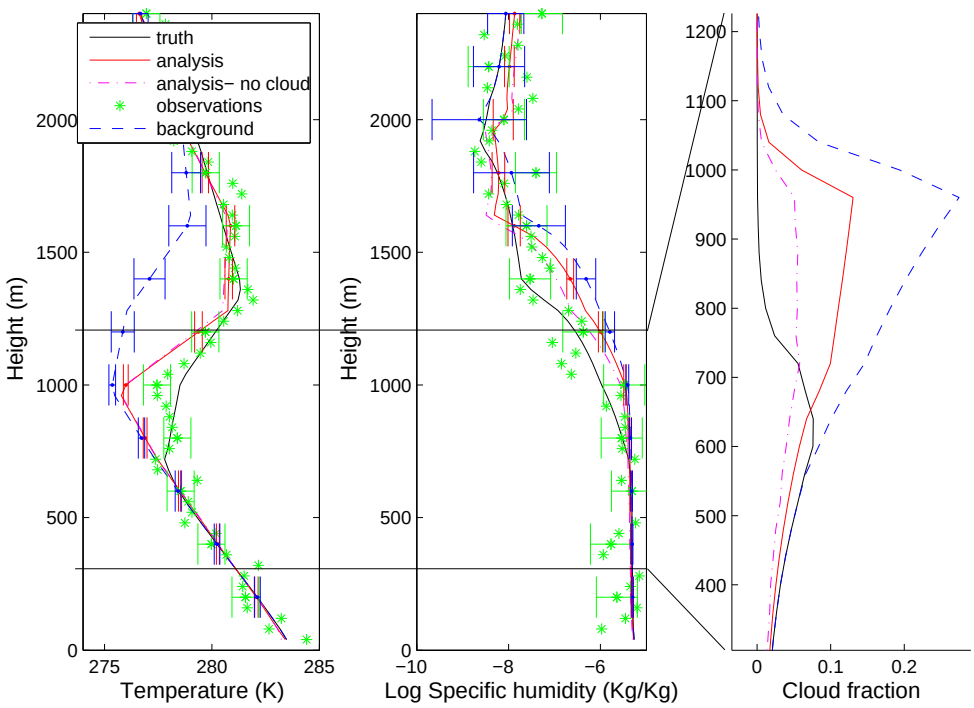


Figure 5.20: Profiles of temperature (left) and log specific humidity (middle) are given for the truth (black), background (blue), observations (green). With these background and observations as input analyses have been calculated using the standard scheme with (red) and without (dashed magenta) the Smith cloud scheme. In the right hand panel the cloud fraction is diagnosed.

diagnosed, using the Smith cloud scheme, from the observation[6] (green) and background (blue) temperature and humidity profile is given. From these profiles of cloud fraction we can note that the cloud in the background is capped by the height of the inversion given by the background and likewise the observed cloud is capped by the height of the inversion observed. This results in the cloud layer being too low in the background.

Using the standard scheme to assimilate these profiles gives the analyses; red - with the Smith cloud scheme and, red dashed dot- without the Smith cloud scheme. An averaged estimate of the background errors for the cloud conserved variables has been used, averaging over 900 ensembles. As seen before, the standard assimilation is unable to give an analysis capable of fully shifting the height of the inversion in the temperature profile and has not altered the humidity profile consistently with the positional error that is present. In this case there is very little difference between the assimilation when the cost function is minimised in terms of cloud conserved variables (red) or the model variables (red dashed). When diagnosing cloud from these analyses we see that the poor temperature and humidity profiles result in a cloud profile which has not adequately shifted the cloud top height to match the cloud fraction consistent with the observations. It has also increased the

[6]It is important to note that the assumptions made in the Smith Cloud scheme, that cloud is distributed within a grid box, is not valid for a point observation such as one from a radiosonde profile, so this profile of observed cloud fraction is only for illustration.

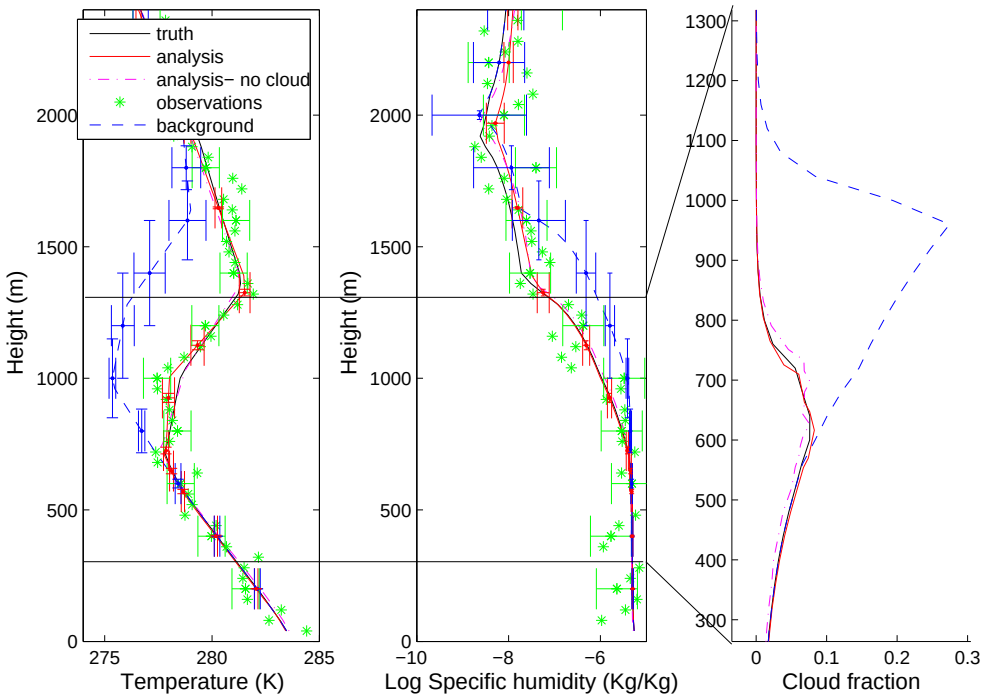


Figure 5.21: As in Figure 5.20 except the analyses have been calculated using the floating BL scheme with (red) and without (dashed magenta) the Smith cloud scheme. With the cloud scheme a has been minimised to -275m.

amount of cloud significantly in both cases.

We can now compare this to the performance of the floating BL scheme given in Figure 5.24. Again the red profile gives the resulting analysis when the cost function has been minimised in terms of cloud conserved variables via the Smith cloud scheme and the red dashed dot profiles give the analysis when the Smith cloud scheme has not been included. The displacement function, $D(z)$ has also been plotted (see (3.1)). In this case a has been minimised to 169m with the Smith cloud scheme and 156m without the Smith cloud scheme. Despite this difference in the final analysed value of a, there is again very little notable difference between the two profiles. In both cases the temperature inversion has successfully been shifted up to agree with the observed inversion whilst maintaining the structure of the inversion. In shifting the inversion in the temperature profile the humidity values in this region have also been shifted up so that in the analysis the profile is no longer too dry in this region.

Whilst the displacement function gives a good description of the positional error in the temperature inversion (=1 where $\frac{dT}{dz} > 0$), it is not a perfect description of the positional error in the transition in the humidity profile from the moist BL to the dry FA. This is primarily due, in this case, to the background being too moist throughout the entire FA and so the shape of the observed humidity profile is not simply a linear shift of the background humidity structure analogous to the shift in the background temperature structure. This accounts for the difference between the analysed humidity values and observed humidity values between approximately 1500m and 2000m.

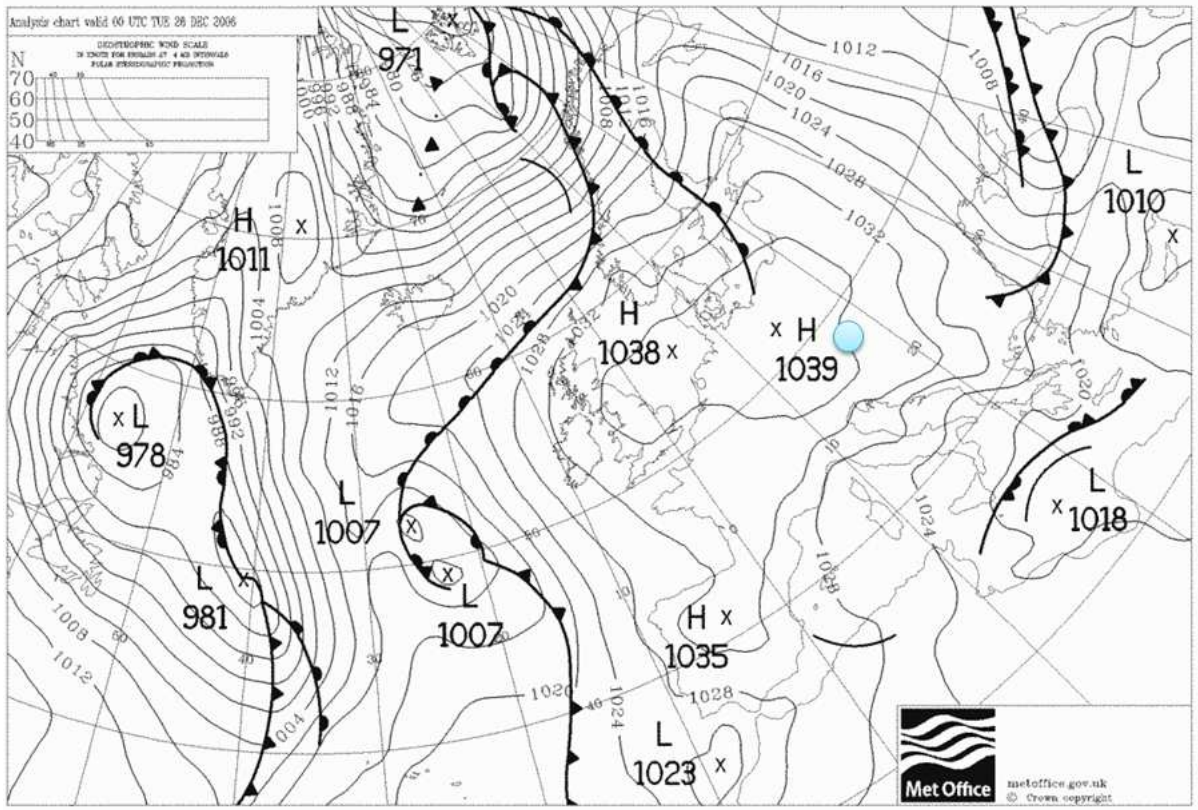


Figure 5.22: Analysis chart from 25th December 2006 of constant surface pressure contours. The blue dot marks the location of the radiosonde data given in Figures 5.23 and 5.24. Courtesy Met Office (Crown Copyright).

In the far right panel of the Figure 5.24 cloud fraction has been diagnosed from the analysed temperature and humidity profiles. We see that the cloud top height has been shifted to give a better agreement with the cloud top height observed and much of the structure of the cloud has been preserved.

5.6 Conclusions

In assimilating both temperature and humidity information it is important that the significant multivariate correlations are accounted for. These depend on the large scale synoptic conditions and local sources and sinks and so can be difficult to model. Above the inversion, the humidity profile rapidly becomes drier, so when there is a misplacement of the inversion in the background we expect the humidity profile to also dry at the wrong level. This gives rise to an effective negative correlation between temperature and humidity errors. This negative correlation can be accounted for automatically with the use of the floating BL scheme which was seen to allow for an accurate analysis of temperature when a large positional error is present in the inversion when only humidity observations are available and vice versa. We can also recreate a positive correlation expected between

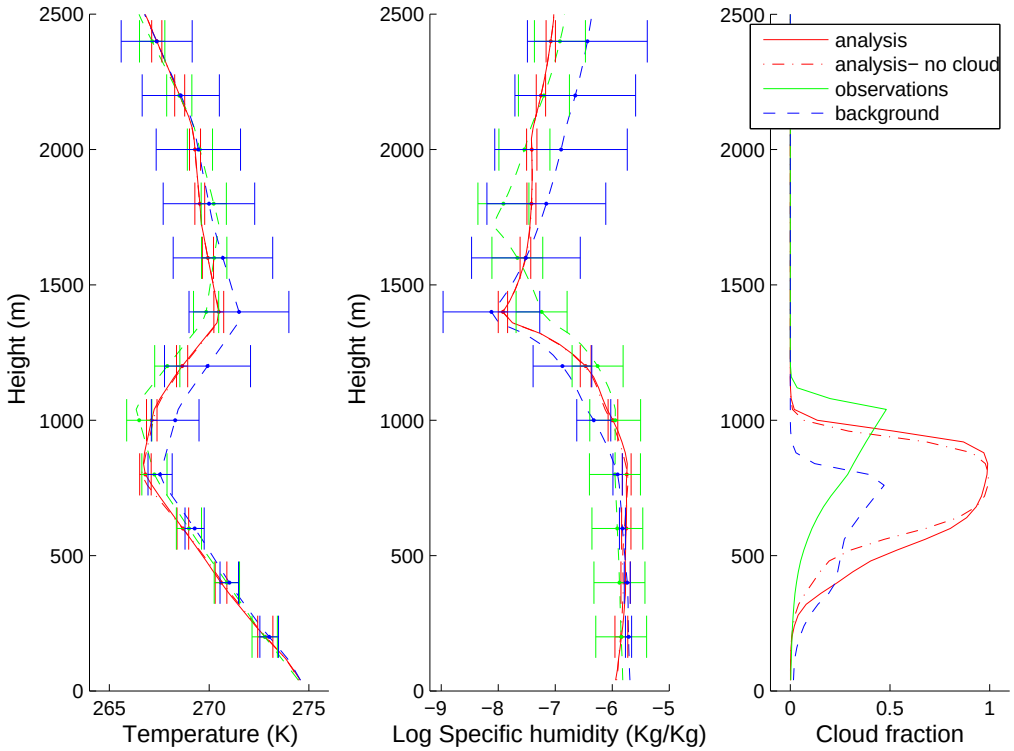


Figure 5.23: Profiles of temperature (left) and log specific humidity (middle) are given for the background (blue) and observations (green). These measurement coincide with the radiosonde released on 25th December 2006 from Czech Republic (marked on Figure 5.22). With these background and observations as input analyses have been calculated using the standard scheme with (red) and without (dashed dot red) the Smith cloud scheme.

temperature and specific humidity in a region of cloud formation through the use of the Smith cloud scheme, however this correlation is small compared to the correlation caused by a large positional error and had only a small effect on the analysis.

In order to implement the Smith cloud scheme we performed the minimisation of the cost function in terms of cloud conserved variables, θ_L and $\ln q_t$. These variables are also conserved under adiabatic motion when precipitation is not present which is beneficial to the floating BL scheme. A tangent linear of the Smith cloud scheme was formulated analytically and was seen to imply a positive correlation between T and q in the region of cloud as required.

A comparison of the diagnosis of cloud from analyses found using the standard and floating BL scheme when the inversion is misplaced in the background was given in Section 5.5. As expected, through giving an analysis of temperature and humidity closer to the truth the floating BL scheme was also able to give an improved representation of cloud; shifting the cloud layer so that it is consistently capped by the inversion and also altering the cloud amount.

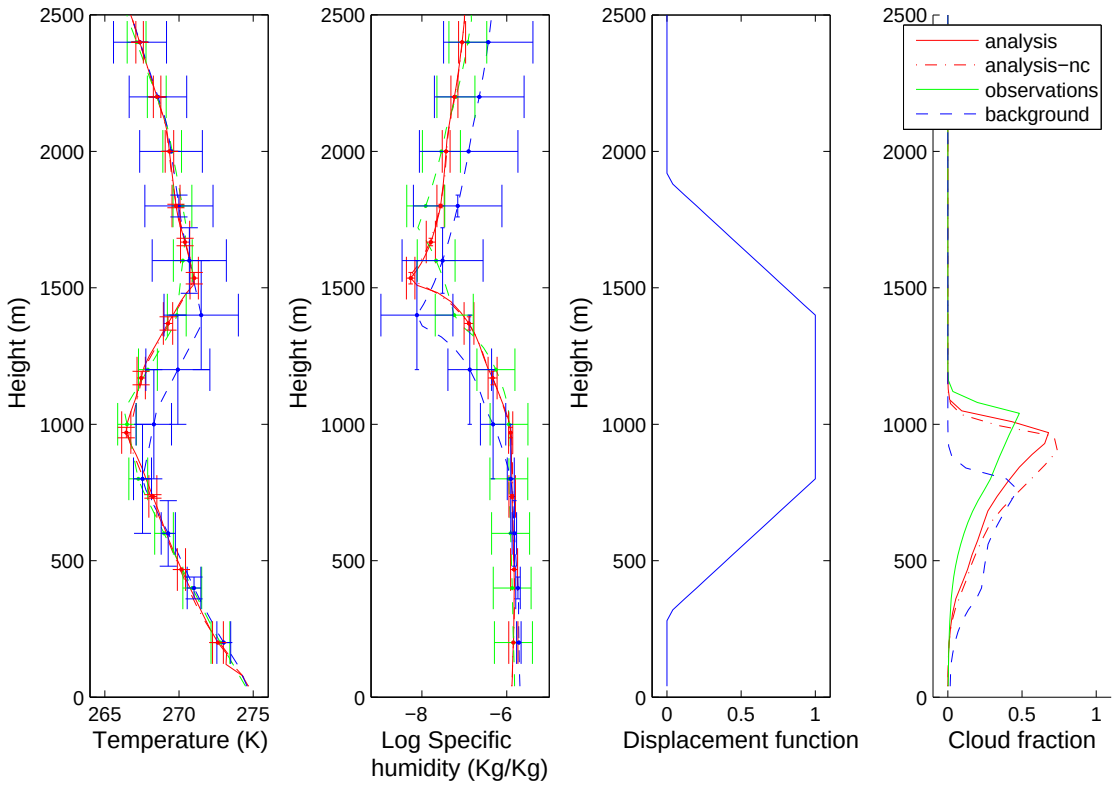


Figure 5.24: As in Figure 5.23 except the analyses have been calculated using the floating BL scheme with (red) and without (dashed dot red) the Smith cloud scheme. With the cloud scheme a has been minimised to 169m.

Chapter 6

Discussion and conclusions

In anti-cyclonic conditions an inversion in the temperature profile often forms at the top of the boundary layer (BL) separating the moist and turbulent BL air from free atmospheric (FA) air above. It has been shown that the BL capping inversion may be subject to a large vertical positional error within the forecast. Assimilating such a forecast as the background state with observed temperature information, using a traditional Variational (Var) technique (designed to minimise amplitude error), can lead to a poor analysis in which the inversion structure is lost.
An accurate and complete description of the background error statistics is important in giving an optimal analysis in which the background and observed information are weighted accordingly (see Chapter 2). The aim of this thesis was to give an improved one dimensional analysis of the BL thermal and moisture vertical structure. This was achieved through the modification of the standard Variational data assimilation scheme to explicitly include a description of the vertical positional error associated with the BL capping inversion in the background.

6.1 Summary

A summary of the key findings in this thesis follows:

- The spread in an ensemble of forecasts for particular times and locations has been used to give a measure of the state-dependent error covariances for the background temperature profile (Section 2.3). It was seen that the error statistics are highly flow-dependent due to the variability in the presence, structure and height of the BL capping inversion, which acts to separate the BL air from the air above. When a strong inversion was present, it was seen that background errors within the BL were strongly correlated, whilst the errors in the BL and the free atmosphere are uncorrelated. A state-dependent estimate of the background error covariances is therefore necessary to ensure that information introduced by the observations is not erroneously spread across these strong inversions.

- Radiosonde observation profiles of temperature have been compared to collocated background temperature profiles. It was concluded that a significant vertical po-

sitional error (of the order of 100m) is often present in the background inversion height (Section 4.3). This error may not necessarily be represented by the ensemble due to under sampling and a possible bias in the forecast (i.e a systematic positional error common to all members). This positional error should be treated separately to the standard amplitude error (which is represented by the ensemble) so that the background inversion structure is able to contribute to the final analysis even if its height is wrong.

- The 'floating BL scheme' was formulated which allows for the inclusion of a positional error in the assimilation (see Chapter 3). This was done by extending the control vector with one extra variable, a, with variance σ_a and the use of a displacement function which is adaptable to the background state. Within the new floating BL scheme the cost function is minimised on model levels which are now allowed to shift according to the value of a in the region of the background inversion, as defined by the displacement function. This enables the background inversion structure to move in accordance with σ_a. The background error correlations between levels are preserved which maintains the decoupling in error correlations above and below the inversion.

- It has been shown that the analysis given by the floating BL scheme could not be emulated by the traditional Var scheme by simply altering the background amplitude error statistics (Section 4.1). This shows that the floating BL scheme adds a fundamentally new perspective to the data assimilation.

- For the floating BL scheme the observation operator changes from one iteration to the next (due to the level heights changing) and so is non-linear (see Subsection 3.1.1). This is the main way in which the floating BL scheme differs from the standard scheme. It was shown that providing the inversion is adequately observed (Section 4.2.3) the shape of the cost function as a function of a remains convex within the region of $-4\sigma_a < a < 4\sigma_a$ and so the non-linear observation operator does not affect the minimisation of the cost function significantly (Subsection 4.2.3). In most cases it is possible to converge close enough towards the global minimum of the cost function within 2 or 3 iterations.

- The performance of the floating BL scheme has been tested for a sample of radiosonde profiles (Section 4.3). In assimilation tests it was seen that overall the floating BL scheme gave an improved analysis of the inversion, including BL height.

- An exception to the previous point was found when there was a small positional error present in the background, in such cases the traditional variational scheme gave an analysis with an inversion height closer to that observed than the analysis produced by the floating BL scheme. This was a consequence of using Gaussian statistics to

describe the errors of a. This was improved on by giving a better description of the error statistics of a; in this case a sum of two Gaussians was used (Section 4.3).

- The floating BL scheme has been extended to include the assimilation of moisture (Chapter 5). When a sharp temperature inversion is present at the top of the BL there is also a large drop in specific humidity, indicating a transition to a different air mass. If the temperature inversion is at the wrong height then it was concluded that this negative gradient in the humidity profile would also be at the wrong height and would therefore need to be shifted vertically as with temperature. The floating BL scheme assimilates both temperature and humidity on the same floating levels and so shifts the temperature inversion and the humidity simultaneously. This ensures that the temperature and humidity profiles remain consistent when a large positional error is present and also allows observations of just temperature or humidity to give a good analysis for both variables. This was then seen to give an improved diagnosis of cloud, shifting the cloud top height so that it remains bounded by the inversion height.

- In assimilating both temperature and humidity it is theoretically important that in regions of cloud the necessary multivariate correlations are accounted for so that the two variables remain consistent and enable realistic cloud amounts to be diagnosed. This was achieved in Chapter 5 by implementing a control variable transform based on the Smith cloud scheme (Smith, 1990). This transform allows the cost function to be minimised in terms of cloud conserved variables. The inclusion of the Smith cloud scheme was seen to have little impact on the analysis in practice when a large positional error was present.

6.1.1 Caveats

The floating BL scheme was shown to be reliable when assimilating a sample of approximately 1000 radiosonde profiles. However, it was shown that the success of the floating BL scheme is sensitive to the choice of displacement function (Subsection 3.1.3). The displacement function must be adaptable to the background temperature profile so that the correct levels are shifted. The choice of displacement function should also ensure that the surface level does not move and levels do not shift below the surface. This criterion is more difficult to guarantee as the value of a is unknown before the observations are assimilated, however, for the results given in this thesis this was not found to be a problem.
It was also shown that care is needed when the inversion is not well observed and the observations are unable to resolve the inversion structure (Subsection 4.2.3). In this case a double minima may form in the cost function where it is possible that the solution will converge to the wrong minimum and the background inversion may be shifted in the wrong direction, although the structure will be maintained. This is not necessarily an issue for radiosonde observations but could cause problems if the floating BL scheme is used to assimilate observations with a lower vertical resolution such as satellite measurements.
This scheme also needs a high vertical resolution in the model which has been assured

artificially within this Thesis. It should then also be noted that the analyses have been given on the final floating levels and not interpolated back to the fixed levels.

6.2 Further work

The results from the floating BL scheme have been limited to the 1D assimilation of radiosonde profiles. In the section below it is speculated how this scheme may be extended to 3D and 4DVar and and some of the problems that may be faced with an operational scheme. The assimilation of different observations and the extension to other variables is also discussed.

6.2.1 Expanding to 3D and 4D Var

Extending the floating BL scheme to work with 3D and 4D Var raises some important research issues. The issues that are anticipated are discussed here.

3DVar

In order to extend the floating BL scheme to 3D there are two elements that need to be considered:
(i) Horizontal correlations of a.
(ii) Horizontal dependence of the displacement function.
If the positional error is correlated horizontally with a large decorrelation length, the observational information about the BL top at one location would imply information about the BL top for the surrounding region, at distances roughly up to the decorrelation length. It would be particularly useful if the error could be assumed isotropic and independent of state as is often done in conventional 3DVar.
It is possible to get an idea of the horizontal decorrelation length scales of a by plotting the ensemble derived correlation matrix for each horizontal grid point. This will allow us to evaluate the validity of an isotropic assumption. This has been done in Figure 6.1 for the North Atlantic and Europe (NAE) domain using MOGREPS (the Met Office's Global and Regional Ensemble Prediction System) data from 25th January 2008 for a point in the centre of the NAE domain (marked by the blue cross). The absence of colour corresponds to grid points where some (or all) members of the ensemble did not have an inversion present. As discussed previously in Subsection 4.1.1 the ensemble may not necessarily be able to capture the full magnitude of the error in the inversion height and the estimate of $\mathbf{B}$ will be rank deficient (and hence noisy) for such a large state (see Section 2.3), but should be able to give us some idea of the horizontal correlation structure.

Figure 6.1 may be compared to Figure 1.4 in which the height of the inversion at this time was plotted for the control forecast member. The variability in the height of the inversion, as shown in Figure 1.4, dictates the variability of the displacement function needed at each horizontal location (the displacement function must relate to the particular background temperature profile). Figure 1.4 shows that the inversion height can be very variable in some regions and so in these regions the displacement function should be calculated for each grid point. This means that even if a itself is not varied by very much over neighbouring grid points, the group of levels that shift (dictated by the local

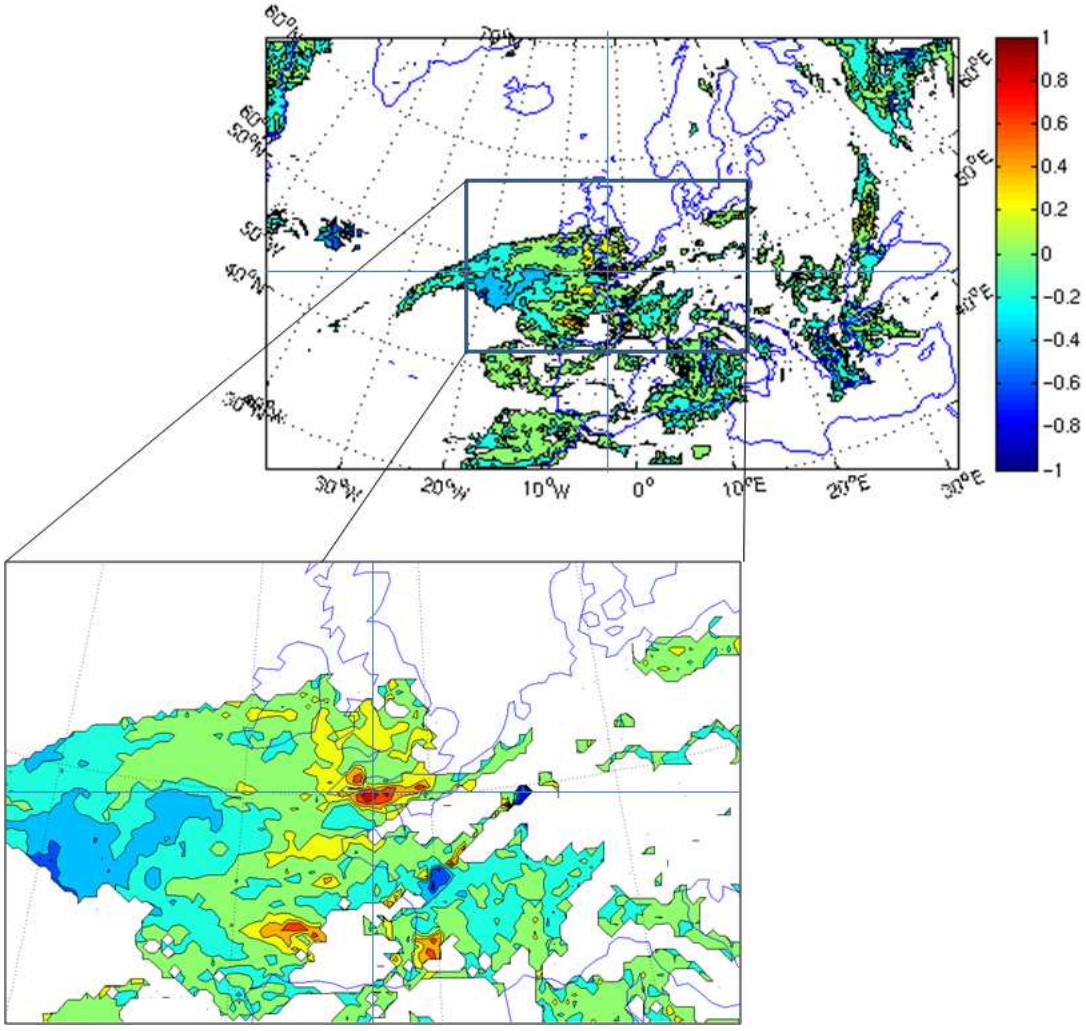


Figure 6.1: Horizontal map of correlations of vertical positional error in the BL capping inversion height, a, at the centre of the NAE domain (marked by the blue cross) as calculated from an ensemble of forecasts for the NAE domain for 25th January 2008, 12Z.

displacement function) may not be identical from one horizontal grid point to the next. If the errors of a are positively correlated, however, the levels will be shifted in the same direction.

When no inversion is present in the background the displacement function may be set to zero at all levels and so varying a would not affect the analysis, therefore the variability in the presence (or not) of the inversion may not necessarily hinder the assumption of an isotropic positional error. For the example shown in Figure 6.1, the positional error does not appear to be isotropic and the decorrelation length scale is relatively small however the pattern should be compared to the synoptic situation and other cases should also be looked at.

In 3D, positional errors in the height of the inversion may be related to positional errors in the horizontal location of fronts and centres of high pressure. This could com-

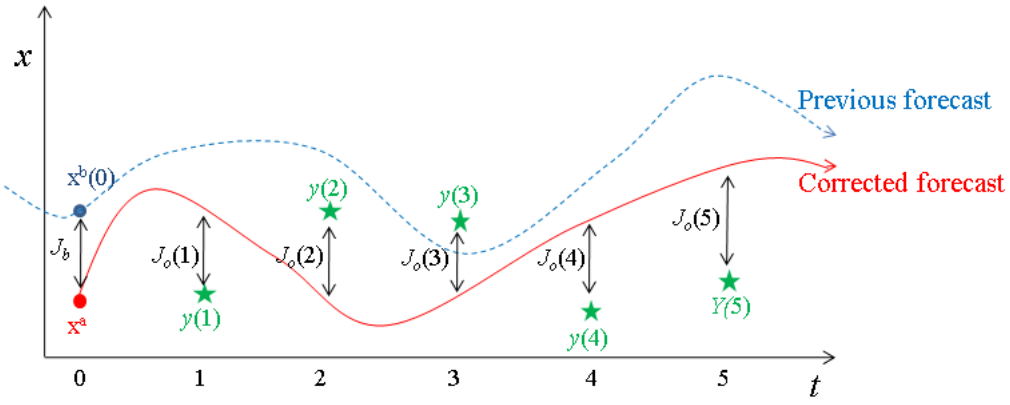


Figure 6.2: Schematic of 4Dvar where an analysis, x^{a}, is calculated at time $t = 0$ from a given background value, x^{b}, also valid at time $t = 0$ and observations, $\mathbf{y}$, valid at times $t = 1, 2, 3, 4, 5$.

plicate matters if there was also a proposal to simultaneously account for these horizontal positional errors, such as in the work of Hoffman et al. (1995) and Ravela et al. (2007).

4DVar

3DVar assimilates observations made only at (or assumed to be at) the validity time of the background. 4DVar allows observations made at different times to be assimilated by including a time dependent model with the observation operator. This is illustrated in Figure 6.2 (the bracketed numbers give the time index). In this illustration we have a background value of x at time $t = 0$, $x^{\mathrm{b}}(0)$, and observations, $y(1, 5)$, made at times $t = 1, 5$. A forecast based on the initial background value of x at time $t = 0$ is run forward in time (blue dashed trajectory) and compared to the observations at the time each observation is made. The initial state $x(0)$ is then perturbed to try to improve the fit to the later observations by minimising a 4DVar cost function (see below) until the solution converges, leading to the analysis x^{a} at time $t = 0$. The forecast run from x^{a} (red trajectory) now gives an improved fit to the observations. The cost function can be defined in a similar way to 1D (see (2.11)) and 3DVar, although the observation part of the cost function is now summed over all observation times, $t = 1, T$:

$$J = J_{\mathrm{b}} + \sum_{t=1}^{T} J_{\mathrm{o}}(t). \tag{6.1}$$

J_{b} is the same kind of background penalty term as shown before in Chapter 2 (2.11) evaluated at time $t = 0$, but includes a 3D state vector $\mathbf{x}(0)$ and 3D covariances in $\mathbf{B}$.

$$J_{\mathrm{b}} = \frac{1}{2}(\mathbf{x}(0) - \mathbf{x}^{\mathrm{b}}(0))^{\mathrm{T}}\mathbf{B}^{-1}(\mathbf{x}(0) - \mathbf{x}^{\mathrm{b}}(0)). \tag{6.2}$$

$\sum_{t=1}^{T} J_{\mathrm{o}}(t)$ is the observation penalty term which, unlike in (2.11), now accounts for the time that the observation is made,

$$J_{\mathrm{o}}(t) = \frac{1}{2}(\mathcal{H}_t(\mathbf{x}(t)) - \mathbf{y}(t))^{\mathrm{T}}\mathbf{R}^{-1}(\mathcal{H}_t(\mathbf{x}(t)) - \mathbf{y}(t)). \tag{6.3}$$

It is this time dimension that distinguishes 4DVar from 3DVar (setting T=1 in (6.1) leads to 3DVar cost function). In operational systems T typically represents a time window of 6 or 12 hours (e.g. Rawlins et al., 2007). $\mathbf{x}(t)$ can be evaluated using the forecast model, $\mathcal{M}_{t\leftarrow 0}$,

$$\mathbf{x}(t) = \mathcal{M}_{t\leftarrow 0}\mathbf{x}(0). \tag{6.4}$$

The forecast model, $\mathcal{M}$, allows state information at time $t = 0$ to propagate forward to the times that the observations are made. In order for the floating BL scheme to be extended to four dimensions it should be understood how a should change with time. As discussed in Chapter 1 the BL has a strong diurnal cycle over land, and so the displacement function should, in principle, be allowed to be dependent upon time. However it may be possible to assume that a is approximately constant. This time dependence of the positional error requires further investigation.

Operational Var

At the MetO an incremental version of 4DVar is operational. Increments of the model variables are calculated as:

$$\begin{aligned} \delta\mathbf{x}_n &= \mathbf{x}_n - \mathbf{x}_{n-1} \\ \delta\mathbf{x}^{\mathrm{b}}_n &= \mathbf{x}^{\mathrm{b}} - \mathbf{x}_{n-1}, \end{aligned} \tag{6.5}$$

where subscript n is the iteration number. This allows (6.1) to be rewritten as:

$$\begin{aligned} J = \tfrac{1}{2}(\delta\mathbf{x}_n(0) - \delta\mathbf{x}^{\mathrm{b}}_n(0))^{\mathrm{T}}\mathbf{B}^{-1}(\delta\mathbf{x}_n(0) - \delta\mathbf{x}^{\mathrm{b}}_n(0)) \\ +\tfrac{1}{2}\textstyle\sum_{t=1}^{T}(\mathcal{H}_t\mathcal{M}_{t\leftarrow 0}\mathbf{x}_n(0) - \mathbf{y}(t))^{\mathrm{T}}\mathbf{R}^{-1}(\mathcal{H}_t\mathcal{M}_{t\leftarrow 0}\mathbf{x}_n(0) - \mathbf{y}(t)) \end{aligned} \tag{6.6}$$

If we also let

$$\delta\mathbf{y}_n(t) = \mathbf{y}(t) - \mathcal{H}(\mathcal{M}_{t\leftarrow 0}(\mathbf{x}_{n-1}(0))), \tag{6.7}$$

we can approximate (6.6), using the Taylor expansion of $\mathcal{H}$ and $\mathcal{M}_{t\leftarrow 0}$ about $\mathbf{x}_n$, as

$$\begin{aligned} J(\delta\mathbf{x}_n) = \tfrac{1}{2}(\delta\mathbf{x}_n(0) - \delta\mathbf{x}^{\mathrm{b}}_n(0))^{\mathrm{T}}\mathbf{B}^{-1}(\delta\mathbf{x}_n(0) - \delta\mathbf{x}^{b}_n(0)) \\ +\tfrac{1}{2}\textstyle\sum_t \left[(\mathbf{H}\mathbf{M}_{t\leftarrow 0}\delta\mathbf{x}_n(0) - \delta\mathbf{y}_n(t)\right]^{\mathrm{T}}\mathbf{R}^{-1}\left[(\mathbf{H}\mathbf{M}_{t\leftarrow 0}\delta\mathbf{x}_n(0) - \delta\mathbf{y}_n(t)\right]. \end{aligned} \tag{6.8}$$

$\mathbf{M}_{t\leftarrow 0}$, the first order Taylor expansion of $\mathcal{M}_{t\leftarrow 0}$ about $\mathbf{x}_n$, is referred to as the *perturbation forecast* (PF) model and can be more expensive to run than $\mathcal{M}$ is.
Each outer loop calculates the incremental value, $\delta\mathbf{x}_n$. The increment $\delta\mathbf{x}_n$ is calculated iteratively using an inner loop via a descent algorithm which requires evaluation of the gradient vector $\partial J/\partial\delta\mathbf{x}_n$ via the adjoint method. The new reference state for outer loop $n+1$ is $\mathbf{x}_{n+1}(0) = \mathbf{x}_n + \delta\mathbf{x}_n$ for which $\mathcal{H}_t$ and $\mathcal{M}_{t\leftarrow 0}$ will need to be re-run. (Note that this procedure is different for 1DVar performed in previous chapters where the inner loop is replaced by an explicit expression for $\delta\mathbf{x}_n$.)

For an operational forecast the length of the model state, $\mathbf{x}$, is approximately of the order 10^7. Which implies the number of elements in $\mathbf{B}$ is of the order 10^{14}. Such a large matrix is difficult to store, let alone operate with. As already seen in Chapters 2 and 5, it is possible to transform to new control variables, χ, via a control variable transform $\mathbf{U}$,

to perform the assimilation scheme. If we let

$$\mathbf{U}\delta\boldsymbol{\chi} = \delta\mathbf{x}, \tag{6.9}$$

then

$$\begin{aligned} J(\delta\boldsymbol{\chi}(0)) = \tfrac{1}{2}(\delta\boldsymbol{\chi}(0) - \delta\boldsymbol{\chi}^{\mathrm{b}}(0))^{\mathrm{T}}(\mathbf{U}^{\mathrm{T}}\mathbf{B}\mathbf{U})^{-1}(\delta\boldsymbol{\chi}(0) - \delta\boldsymbol{\chi}^{\mathrm{b}}(0)) \\ +\tfrac{1}{2}\textstyle\sum_t \left(\mathbf{H}_t\mathbf{M}_{t\leftarrow 0}\mathbf{U}\delta\boldsymbol{\chi}(0) - \delta\mathbf{y}(t)\right)^{\mathrm{T}}\mathbf{R}^{-1}\left(\mathbf{H}_t\mathbf{M}_{t\leftarrow 0}\mathbf{U}\delta\boldsymbol{\chi}(0) - \delta\mathbf{y}(t)\right). \end{aligned} \tag{6.10}$$

The introduction of $\mathbf{U}$ means it is no longer necessary to store $\mathbf{B}$ or calculate its inverse. By carefully choosing $\mathbf{U}$, $\mathbf{B}_\chi = \mathbf{U}^{\mathrm{T}}\mathbf{B}\mathbf{U}$ may become the identity which avoids the need to store 10^{14} elements. At the MetO this is done by splitting the U-transform into three stages (Lorenc et al., 2000; Met Office, 1995; Ingleby 2001; Rawlins et al., 2007).

$$\mathbf{U} = \mathbf{U}_{\mathrm{p}}\mathbf{U}_{\mathrm{v}}\mathbf{U}_{\mathrm{h}} \tag{6.11}$$

$\mathbf{U}_{\mathrm{p}}$ transforms from parameter space to model space. The parameters used are assumed to be uncorrelated with one another, whilst the model variables are strongly correlated[1]. This only leaves spatial correlation amongst the parameters. $\mathbf{U}_{\mathrm{v}}$ and $\mathbf{U}_{\mathrm{h}}$ are the vertical and horizontal transforms respectively. The inverses of $\mathbf{U}_{\mathrm{v}}$ and $\mathbf{U}_{\mathrm{h}}$ remove spatially univariate correlations by use of a spectral transform (Lorenc et al., 2000).

The incremental nature of operational 4DVar and the balances enforced by the control variable transform could lead to problems when implementing the floating BL scheme into this framework. Some suggestions for how to attempt this are given next.

Ideas for including the floating boundary layer scheme into operational var at MetO

Suggestion I: incremental 4DVar where the model runs on floating levels

4DVar with the floating BL scheme may be implemented by minimising the cost function with respect to $\delta\tilde{\boldsymbol{\chi}}(0)$, the control variables on floating model levels.

$$\tilde{\mathbf{U}}\delta\tilde{\boldsymbol{\chi}} = \delta\tilde{\mathbf{v}} \tag{6.12}$$

where $\tilde{\mathbf{U}}$ is defined by (6.11) but is now associated with the floating model levels.

$$\begin{aligned} J(\delta\tilde{\boldsymbol{\chi}}(0)) = \tfrac{1}{2}(\delta\tilde{\boldsymbol{\chi}}(0) - \delta\tilde{\boldsymbol{\chi}}^{\mathrm{b}}(0))^{\mathrm{T}}\tilde{\mathbf{B}}_\chi^{-1}(\delta\tilde{\boldsymbol{\chi}}(0) - \delta\tilde{\chi}^{\mathrm{b}}(0)) \\ +\tfrac{1}{2}\textstyle\sum_t \left(\tilde{\mathcal{H}}_t\tilde{\mathcal{M}}_{t\leftarrow 0}\tilde{\mathbf{U}}\delta\tilde{\boldsymbol{\chi}}(0)) - \mathbf{y}(t)\right)^{\mathrm{T}}\mathbf{R}^{-1}\left(\tilde{\mathcal{H}}_t\tilde{\mathcal{M}}_{t\leftarrow 0}\tilde{\mathbf{U}}\delta\tilde{\boldsymbol{\chi}}(0)) - \mathbf{y}(t)\right), \end{aligned} \tag{6.13}$$

where a tilde overbar indicates a vector or matrix associated with the floating levels. The shifting of levels is included in the observation operator, $\tilde{\mathcal{H}}_t$, as has been done so far within this thesis, in order to calculate the model version of the observations. On floating levels the floating BL scheme may be treated as an incremental scheme.

[1]In formulating $\mathbf{U}_{\mathrm{p}}$ certain balance relations may be assumed, e.g. geostrophic balance, and so when the final solution is transformed back to model variables these balances are implied.

This method could lead to potential imbalances which were previously enforced by $\mathbf{U}$. It would also make $\tilde{\mathcal{M}}$ and $\tilde{\mathbf{M}}$ (the linearisation of $\tilde{\mathcal{M}}$ on floating levels) more complicated due to a varying with position.

Suggestion II: incremental 4DVar where the model runs on fixed levels

The introduction of the time dimension in the 4DVar problem makes the observation term in the cost function a lot more complicated and so it may be simpler to implement the floating scheme as part of the U-transform instead. Consider

$$\mathbf{U}_{\mathrm{FL}}\delta\tilde{\boldsymbol{\chi}} = \delta\mathbf{x}, \tag{6.14}$$

where

$$\mathbf{U}_{\mathrm{FL}} = \mathbf{U}_{\mathrm{d}}\tilde{\mathbf{U}}_{\mathrm{p}}\tilde{\mathbf{U}}_{\mathrm{v}}\tilde{\mathbf{U}}_{\mathrm{h}}. \tag{6.15}$$

$\mathbf{U}_{\mathrm{d}}$ is fundamentally an interpolation of model increments on floating levels to fixed model levels (see below for details). We would now wish to minimise the following cost function:

$$\begin{aligned} J(\delta\tilde{\boldsymbol{\chi}}(0)) = \tfrac{1}{2}(\delta\tilde{\boldsymbol{\chi}}(0) - \delta\tilde{\boldsymbol{\chi}}^{\mathrm{b}}(0))^{\mathrm{T}}(\mathbf{U}_{\mathrm{FL}}^{\mathrm{T}}\mathbf{B}\mathbf{U}_{\mathrm{FL}})^{-1}(\delta\tilde{\boldsymbol{\chi}}(0) - \delta\tilde{\boldsymbol{\chi}}^{\mathrm{b}}(0)) \\ +\tfrac{1}{2}\sum_t \left(\mathbf{H}_t\mathbf{M}_{t\leftarrow 0}\mathbf{U}_{\mathrm{FL}}\delta\tilde{\boldsymbol{\chi}}(0) - \delta\mathbf{y}(t)\right)^{\mathrm{T}}\mathbf{R}^{-1}\left(\mathbf{H}_t\mathbf{M}_{t\leftarrow 0}\mathbf{U}_{\mathrm{FL}}\delta\tilde{\boldsymbol{\chi}}(0) - \delta\mathbf{y}(t)\right), \end{aligned} \tag{6.16}$$

which bears more resemblance to the traditional 4DVar cost function (6.8), since the model and its linearisation act on fixed model levels. In order to calculate an effective increment the transform, $\mathbf{U}_{\mathrm{d}}$, needs to be split into a few steps as it is the full temperature profile, describing the temperature inversion, that needs to be shifted not just the temperature increments:

1. Calculate full model field on floating levels $\tilde{\mathbf{x}}_{n+1} = \delta\tilde{\mathbf{x}}_n + \tilde{\mathbf{x}}_n$
2. Interpolate to fixed levels giving $\mathbf{x}_{n+1}$
3. Remove $\mathbf{x}_n$ to give $\delta\mathbf{x}_n$

Unlike with the first idea the observation operator no longer needs to interpolate from the floating model levels to the fixed observation levels: instead $\mathbf{U}_{\mathrm{d}}$ interpolates from floating to fixed model levels so there is now an interpolation error associated with $\mathbf{U}_{\mathrm{FL}}$, which could be large.

Due to the sharp feature of the inversion a small increment in $\tilde{\mathbf{x}}$ may lead to a large increment in $\mathbf{x}$. This may again cause problems when the perturbation forecast model is calculated. Also $\mathbf{U}_{\mathrm{d}}$ is the last operator to act when calculating $\delta\mathbf{x}$ which means that the balances introduced when operating with $\mathbf{U}_{\mathrm{p}}$, such as geostrophic balance and hydrostatic balance may be destroyed.

6.2.2 Use of observations other than radiosondes

The assimilation of radiosondes is limited due to their poor temporal resolution and spatial coverage (see Subsection 1.2.1 and Figure 1.5), although their vertical resolution is very good. Satellite data such as measured infra-red radiances can give frequent sampling over a large area of the globe. However as discussed in Subsection 1.2.1 it is not easy to infer

from the observed top of the atmosphere (TOA) radiances the height-resolved, thermal and moisture structure of the BL. Despite this, the radiances are sensitive to the height of the BL capping inversion when capped by cloud, especially at high spectral resolution, and so may be used to constrain the value of a in the floating BL scheme. More work is needed to determine how best to assimilate this information for use with the floating BL scheme, e.g. on-line radiance assimilation versus cloud-top height as a derived product.
Measurements of the BL top are also available from active remote sensors such as lidars (see Chapter 1, Figure 1.3). Cloud top temperature measured from a satellite may also be useful if it is known that the cloud is within the BL. These kinds of observations would have to be treated differently to the full information provided by radiosondes due to difficulties with defining the observation operator, see Figure 6.3. Some suggestions for using BL top height observations follow:

1. Assimilate the raw measurement. For lidar this would be a vertical profile of backscattered laser power (sensitive to clouds, precipitation and aerosols). This would require careful forward modelling and an assessment of errors in the model, including linearisation error.

2. Assimilate a temperature profile derived off line from the lidar measurement and the background.

3. Allow for observations of a, by comparing the measured value of the BL top height, or equivalent, to that derived from the background temperature profile. The difference may then be used as the observed value of a.

These ideas may help to define the observation operator needed for the assimilation of BL top height observations (as illustrated in Figure 6.3) and should be investigated further.

6.2.3 The inclusion of other variables

Within this thesis we have explored the use of the floating BL scheme when assimilating thermal and moisture information. It was concluded that the variables used as the control variables should be conserved with adiabatic motion. For this reason potential temperature was used for univariate assimilation of temperature and liquid water potential temperature and total water were used for the multivariate assimilation of temperature and moisture.
As humidity is sensitive to the height and structure of the inversion the assimilation of humidity benefited from the floating BL scheme. In the same way the direct assimilation of other variables which are sensitive to the inversion could also benefit from the floating BL scheme. These variables could include wind speed and aerosol/chemical concentration.

In summary; a new assimilation scheme, the 'floating BL scheme', has been developed within the 1DVar framework. This scheme allows for the explicit inclusion of a positional error associated with the BL capping inversion. It has been shown for the assimilation of radiosonde profiles that this new scheme gives an improved temperature and humidity profile compared to the standard scheme when the inversion is misplaced in the background. Ideas have been presented for how this scheme may be extended for use with

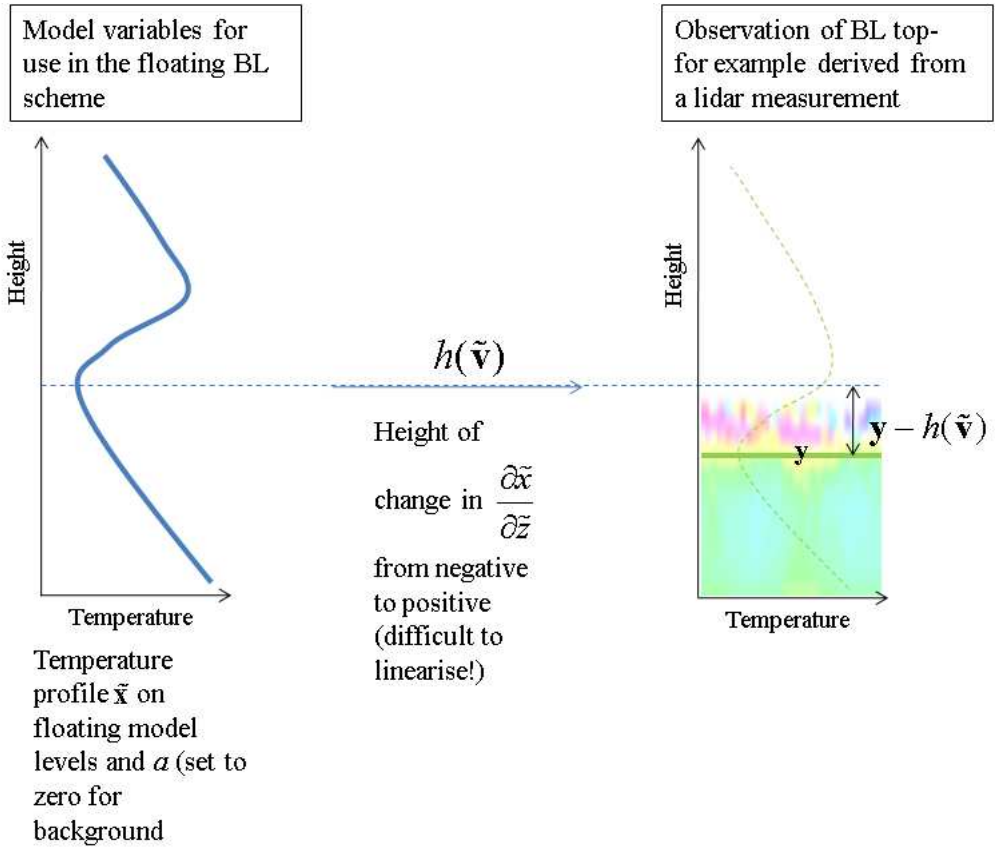


Figure 6.3: Schematic of the difficulty in defining an observation operator needed for use with observations of BL top. $\tilde{\mathbf{v}}$ and $\tilde{\mathbf{z}}$ are defined in (3.3) and (3.1).

3D and 4DVar, and other observations. It is also speculated that the analysis of other variables may also benefit from the use of the floating BL scheme.

Appendix A

Proof of the Bennett-Talagrand Theorem

In statistical theory, the cost function may be thought equivalent to $\frac{1}{2}\chi^2$ (Bennett, 1992; Bennett, 2002), which should equal approximately half the number of degrees of freedom of the system (Barlow, 1989). Provided the assumed error statistics are correct, the number of degrees of freedom should equal the number of data (Bennett, 1992; Bennett, 2002), and so the minimum of an accurately derived cost function equals half the number of observations. A thorough proof of this result is given below.

Assuming Gaussian and unbiased error statistics, the analysis $\mathbf{x}^{\mathrm{a}}$ is found by minimising the cost function, J,

$$J(\mathbf{x}) = \frac{1}{2}(\mathbf{x} - \mathbf{x}^{\mathrm{b}})^{\mathrm{T}}\mathbf{B}^{-1}(\mathbf{x} - \mathbf{x}^{\mathrm{b}}) + \frac{1}{2}(\mathbf{y} - \mathcal{H}(\mathbf{x}))^{\mathrm{T}}\mathbf{R}^{-1}(\mathbf{y} - \mathcal{H}(\mathbf{x})). \tag{A.1}$$

(see Chapter 2). If we assume that the observation operator $\mathcal{H}$ is linear and the background and observation errors are uncorrelated we can show that the Bennett-Talagrand Theorem holds that:

$$\langle J(\mathbf{x}^{\mathrm{a}})\rangle = P/2 \tag{A.2}$$

where P is the length of of the observation vector, i.e. $\mathbf{y}(\epsilon\mathbb{R}^{P\times 1})$.

PROOF: In Chapter 2 we showed that if $\mathcal{H}$ is linear and the background and observation errors are Gaussian the analysis is given by

$$\mathbf{x}^{\mathrm{a}} = \mathbf{x}^{\mathrm{b}} + \mathbf{K}(\mathbf{y} - \mathcal{H}(\mathbf{x}^{\mathrm{b}})). \tag{A.3}$$

where

$$\mathbf{K} = \mathbf{B}\mathbf{H}^{\mathrm{T}}(\mathbf{R} + \mathbf{H}\mathbf{B}\mathbf{H}^{\mathrm{T}})^{-1} \tag{A.4}$$

We also showed in Chapter 2 that for these assumptions the analysis error covariance matrix can be shown to be:

$$\mathbf{P}^{\mathrm{a}} = (\mathbf{I} - \mathbf{K}\mathbf{H})\mathbf{B}. \tag{A.5}$$

Using the identity: $\mathbf{B}\mathbf{H}^{\mathrm{T}}(\mathbf{R} + \mathbf{H}\mathbf{B}\mathbf{H}^{\mathrm{T}})^{-1}\mathbf{R} = (\mathbf{I} - \mathbf{B}\mathbf{H}^{\mathrm{T}}(\mathbf{R} + \mathbf{H}\mathbf{B}\mathbf{H}^{\mathrm{T}})^{-1}\mathbf{H})\mathbf{B}\mathbf{H}^{\mathrm{T}}$ and substituting (A.4) and (A.5) into this gives:

$$\mathbf{K} = \mathbf{P}^{\mathrm{a}}\mathbf{H}^{\mathrm{T}}\mathbf{R}^{-1} \tag{A.6}$$

The error in the analysis, $\boldsymbol{\eta}^{\mathrm{a}}$, can be shown to be

$$\begin{gathered}
\boldsymbol{\eta}^{\mathrm{a}} = \mathbf{x}^{\mathrm{a}} - \mathbf{x}^{\mathrm{t}} \\
= \mathbf{x}^{\mathrm{a}} - \mathbf{x}^{\mathrm{b}} + \mathbf{x}^{\mathrm{b}} - \mathbf{x}^{\mathrm{t}} \\
= \mathbf{K}(\mathbf{y} - \mathcal{H}(\mathbf{x}^{\mathrm{b}})) + \mathbf{x}^{\mathrm{b}} - \mathbf{x}^{\mathrm{t}} \text{ using (A.3)} \\
= \mathbf{K}(\boldsymbol{\eta}^{\mathrm{o}} - \mathbf{H}(\boldsymbol{\eta}^{\mathrm{b}})) + \boldsymbol{\eta}^{\mathrm{b}} \text{ using the tangent linear hypothesis and} \\
\text{definitions of } \boldsymbol{\eta}^{\mathrm{b}} \text{ and } \boldsymbol{\eta}^{\mathrm{o}} \text{ given by (2.6) and (2.7) respectively, assuming there are no biases.} \\
\text{Therefore } \boldsymbol{\eta}^{\mathrm{a}} = (\mathbf{I} - \mathbf{KH})\boldsymbol{\eta}^{\mathrm{b}} + \mathbf{K}\boldsymbol{\eta}^{\mathrm{o}}.
\end{gathered} \tag{A.7}$$

MATRIX IDENTITIES (Lewis et al., 2006)

$$trace(\mathbf{MN}) = trace(\mathbf{NM}) \tag{A.8}$$

$$\mathbf{u}^{\mathrm{T}}\mathbf{Mv} = trace(\mathbf{Mvu}^{\mathrm{T}}) \tag{A.9}$$

From the background part of (A.1) we can use the above identities to show:

$$\begin{gathered}
2\left\langle J^{\mathrm{b}}(\mathbf{x}^{\mathrm{a}})\right\rangle \\
= \left\langle (\mathbf{x}^{\mathrm{a}} - \mathbf{x}^{\mathrm{b}})^{\mathrm{T}}\mathbf{B}^{-1}(\mathbf{x}^{\mathrm{a}} - \mathbf{x}^{\mathrm{b}})\right\rangle \\
= trace(\mathbf{B}^{-1}\left\langle (\mathbf{x}^{\mathrm{a}} - \mathbf{x}^{\mathrm{b}})(\mathbf{x}^{\mathrm{a}} - \mathbf{x}^{\mathrm{b}})^{\mathrm{T}}\right\rangle \text{ using identity (A.9)} \\
= trace(\mathbf{B}^{-1}\left\langle (\boldsymbol{\eta}^{\mathrm{a}} - \boldsymbol{\eta}^{\mathrm{b}})(\boldsymbol{\eta}^{\mathrm{a}} - \boldsymbol{\eta}^{\mathrm{b}})^{\mathrm{T}}\right\rangle) \text{ using definitions of } \boldsymbol{\eta}^{\mathrm{a}} \text{ and } \boldsymbol{\eta}^{\mathrm{b}}
\end{gathered} \tag{A.10}$$

Expanding $\left\langle (\boldsymbol{\eta}^{\mathrm{a}} - \boldsymbol{\eta}^{\mathrm{b}})(\boldsymbol{\eta}^{\mathrm{a}} - \boldsymbol{\eta}^{\mathrm{b}})^{\mathrm{T}}\right\rangle$ gives

$$\begin{gathered}
\left\langle (\boldsymbol{\eta}^{\mathrm{a}} - \boldsymbol{\eta}^{\mathrm{b}})(\boldsymbol{\eta}^{\mathrm{a}} - \boldsymbol{\eta}^{\mathrm{b}})^{\mathrm{T}}\right\rangle = \mathbf{P}^{\mathrm{a}} - 2\left\langle \boldsymbol{\eta}^{\mathrm{a}}\boldsymbol{\eta}^{\mathrm{bT}}\right\rangle + \mathbf{B} \\
= \mathbf{P}^{\mathrm{a}} - 2\left\langle ((\mathbf{I} - \mathbf{KH})\boldsymbol{\eta}^{\mathrm{b}} + \mathbf{K}\boldsymbol{\eta}^{\mathrm{o}})\boldsymbol{\eta}^{\mathrm{bT}}\right\rangle + \mathbf{B} \text{ using (A.7)} \\
= \mathbf{P}^{\mathrm{a}} - 2(\mathbf{I} - \mathbf{KH})\mathbf{B} + \mathbf{B} \\
\text{assuming that background and observation errors are uncorrelated} \\
= (\mathbf{I} - \mathbf{KH})\mathbf{B} - 2(\mathbf{I} - \mathbf{KH})\mathbf{B} + \mathbf{B} \text{ using (A.5)} \\
= \mathbf{B} - \mathbf{P}^{\mathrm{a}} \text{ simplifying and using (A.5)}
\end{gathered} \tag{A.11}$$

Substituting (A.11) into (A.10) gives

$$\begin{gathered}
2\left\langle J^{\mathrm{b}}(\mathbf{x}^{\mathrm{a}})\right\rangle \\
= trace(\mathbf{I} - \mathbf{B}^{-1}\mathbf{P}^{\mathrm{a}}) \\
= trace(\mathbf{KH}) \text{ using (A.5) and identity (A.8)}
\end{gathered} \tag{A.12}$$

Similarly, we can now evaluate the observation part of (A.1)

$$\begin{aligned} & 2\left\langle J^{\mathrm{o}}(\mathbf{x}^{\mathrm{a}})\right\rangle \\ & = \left\langle (\mathbf{y}-\mathcal{H}(\mathbf{x}^{\mathrm{a}}))^{\mathrm{T}}\mathbf{R}^{-1}(\mathbf{y}-\mathcal{H}(\mathbf{x}^{\mathrm{a}}))\right\rangle \\ & = trace(\mathbf{R}^{-1}\left\langle (\mathbf{y}-\mathcal{H}(\mathbf{x}^{\mathrm{a}}))(\mathbf{y}-\mathcal{H}(\mathbf{x}^{\mathrm{a}}))^{\mathrm{T}}\right\rangle) \text{ using identity (A.9)} \\ & = trace(\mathbf{R}^{-1}\left\langle (\boldsymbol{\eta}^{\mathrm{o}}-\mathbf{H}\boldsymbol{\eta}^{\mathrm{a}})(\boldsymbol{\eta}^{\mathrm{o}}-\mathbf{H}\boldsymbol{\eta}^{\mathrm{a}})^{\mathrm{T}}\right\rangle) \text{ using definitions of } \boldsymbol{\eta}^{\mathrm{a}} \text{ and } \boldsymbol{\eta}^{\mathrm{b}} \end{aligned} \tag{A.13}$$

$$\begin{aligned} & \left\langle (\boldsymbol{\eta}^{\mathrm{o}}-\mathbf{H}\boldsymbol{\eta}^{\mathrm{a}})(\boldsymbol{\eta}^{\mathrm{o}}-\mathbf{H}\boldsymbol{\eta}^{\mathrm{a}})^{\mathrm{T}}\right\rangle \\ & = \mathbf{R}+\mathbf{H}\mathbf{P}^{\mathrm{a}}\mathbf{H}^{\mathrm{T}}-\mathbf{H}\left\langle \boldsymbol{\eta}^{\mathrm{a}}\boldsymbol{\eta}^{\mathrm{o}\mathrm{T}}\right\rangle -\left\langle \boldsymbol{\eta}^{\mathrm{o}}\boldsymbol{\eta}^{\mathrm{a}\mathrm{T}}\right\rangle \mathbf{H}^{\mathrm{T}} \end{aligned} \tag{A.14}$$

Using (A.7) and assuming background and observation errors are uncorrelated gives

$$\begin{aligned} \left\langle (\boldsymbol{\eta}^{\mathrm{o}}-\mathbf{H}\boldsymbol{\eta}^{\mathrm{a}})(\boldsymbol{\eta}^{\mathrm{o}}-\mathbf{H}\boldsymbol{\eta}^{\mathrm{a}})^{\mathrm{T}}\right\rangle & = \mathbf{R}+\mathbf{H}\mathbf{P}^{\mathrm{a}}\mathbf{H}^{\mathrm{T}}-\mathbf{H}\mathbf{K}\mathbf{R}-\mathbf{R}\mathbf{K}^{\mathrm{T}}\mathbf{H}^{\mathrm{T}} \\ & = \mathbf{R}-\mathbf{H}\mathbf{K}\mathbf{R} \text{ using symmetry of } \mathbf{H}\mathbf{K}\mathbf{R} \text{ and (A.6)} \end{aligned} \tag{A.15}$$

And so

$$\begin{aligned} & 2\left\langle J^{\mathrm{o}}(\mathbf{x}^{\mathrm{a}})\right\rangle \\ & = trace(\mathbf{I}-\mathbf{R}^{-1}\mathbf{H}\mathbf{K}\mathbf{R}) \\ & = P-trace(\mathbf{K}\mathbf{H}) \text{ using identity (A.8)} \end{aligned} \tag{A.16}$$

Now putting A.12 and A.16 together we get:

$$2\left\langle J(\mathbf{x}^{\mathrm{a}})\right\rangle = 2\left\langle J^{\mathrm{b}}(\mathbf{x}^{\mathrm{a}})\right\rangle + 2\left\langle J^{\mathrm{o}}(\mathbf{x}^{\mathrm{a}})\right\rangle = P \qquad \text{Q.E.D} \tag{A.17}$$

Appendix B

Algorithm for the floating boundary layer scheme

This flow chart shows the algorithm for the floating BL scheme.

- $n = 1$
- $\tilde{\mathbf{v}}_1 = \begin{pmatrix} \mathbf{x}^{\mathrm{b}} \\ a = 0 \end{pmatrix}$
- Set $\delta\tilde{\mathbf{w}}_0$ and $\delta\tilde{\mathbf{w}}_0^b$ to zero

1. Calculate $\mathbf{H}_n = \frac{\partial\mathcal{H}(\tilde{\mathbf{v}})}{\partial\tilde{\mathbf{v}}}|_{\tilde{\mathbf{v}}=\tilde{\mathbf{v}}_n}$ (see Appendix C)
2. Calculate $\delta\mathbf{y} = \mathbf{y} - \mathcal{H}(\tilde{\mathbf{v}}_n)$
3. Calculate $\mathbf{U}_n$ (see Appendix E for the transform to cloud conserved variables)
4. Calculate $\delta\tilde{\mathbf{w}}_n^{\mathrm{b}} = \delta\tilde{\mathbf{w}}_{n-1}^{\mathrm{b}} - \delta\tilde{\mathbf{w}}_{n-1}$
5. Calculate $\delta\tilde{\mathbf{w}}_n = (\tilde{\mathbf{B}}_{\mathrm{w}}^{-1} + \mathbf{U}_n^{\mathrm{T}}\mathbf{H}_n^{\mathrm{T}}\mathbf{R}^{-1}\mathbf{H}_n\mathbf{U}_n)^{-1}(\tilde{\mathbf{B}}_{\mathrm{w}}^{-1}\delta\tilde{\mathbf{w}}_n^{\mathrm{b}} + \mathbf{U}_n^{\mathrm{T}}\mathbf{H}_n^{\mathrm{T}}\mathbf{R}^{-1}\delta\mathbf{y}_n)$
6. Update $\tilde{\mathbf{v}}_{n+1} = \tilde{\mathbf{v}}_n + \mathbf{U}\delta\tilde{\mathbf{w}}_n$
7. n=n+1

- Repeat step 1 to 7 until $\tilde{\mathbf{v}}$ converges. The criteria for convergence is achieved when $\left|\frac{\tilde{\mathbf{v}}_n - \tilde{\mathbf{v}}_{n-1}}{\tilde{\mathbf{v}}_{n-1}}\right| < 0.01$

The Fortran code used to implement this scheme is given in the following pages. Approximately 20 subroutines are used and so only the main program is given here along with the subroutine 'calc_H', which calculates the linearised observation operator, the transform from the floating model levels to the fixed observation space. Also given is 'calc_delta_w', which calculates the increment of the state variables in floating level space between two iterations.

A brief description of the other key subroutines used follows:

calc_back	Reads in background data, 'xB', and model level heights, 'mod_ht'.
calc_bltop	Calculates the middle of the inversion, 'bl_top', and its depth 'inv_depth'.
calc_y	Reads in observation data, 'y', and observation level heights, 'ob_ht'.
read_covariances	Reads in the background error covariance matrix, 'Bw', and the observation error covariance matrix, 'R'.
calc_delta_y	Calculates 'delta_y', the difference between the observation and the current estimate of 'v' in observation space.
calc_Jmin	Calculates the value of the cost function at the analysis.
calc_anal_error	Calculates the analysis error covariance matrix, 'A_matrix' (see (2.22)).

Appendix C

The linearised observation operator for the floating BL scheme

The observation operator, $\mathcal{H}$, transforms the state vector from model space to observation space. For the floating BL scheme the state vector is

$$\tilde{\mathbf{v}} = \begin{pmatrix} \tilde{\mathbf{x}} \\ a \end{pmatrix}. \tag{C.1}$$

Model space is described on floating levels given by

$$\tilde{z} = z + aD(z), \tag{C.2}$$

where z is the original fixed model level height and D is the displacement function (see Chapter 3). Observation space is fixed, and so the observation operator is non-linear. Assuming that the model and observations measure the same variables, for example temperature, the observation operator consists of linear interpolation. The value of the j^{th} element of $\mathcal{H}$ acting on $\tilde{\mathbf{v}}$ gives the model version of the j^{th} observation and is given by:

$$\mathcal{H}(\tilde{\mathbf{v}})^j = \left(\frac{\tilde{z}^{\mathrm{II}} - z_o^j}{\tilde{z}^{\mathrm{II}} - \tilde{z}^{\mathrm{I}}}\right)\tilde{x}^{\mathrm{I}} + \left(\frac{z_o^j - \tilde{z}^{\mathrm{I}}}{\tilde{z}^{\mathrm{II}} - \tilde{z}^{\mathrm{I}}}\right)\tilde{x}^{\mathrm{II}}, \tag{C.3}$$

where z_o^j is the height of the j^{th} observation, and $\tilde{x}^{\mathrm{I}}$ $\tilde{x}^{\mathrm{II}}$ are the values of $\tilde{\mathbf{x}}$ either side of the j^{th} observation level and $\tilde{z}^{\mathrm{I}}$ and $\tilde{z}^{\mathrm{II}}$ are their heights respectively.
For the minimisation of the cost function it is necessary to linearise $\mathcal{H}$. Let $\mathbf{H}$ be the Jacobian of $\mathcal{H}$, a $P \times N$ matrix.

$$\mathbf{H} = \frac{\partial \mathcal{H}(\tilde{\mathbf{v}})}{\partial \tilde{\mathbf{v}}} = \begin{pmatrix} \frac{\partial \mathcal{H}(\tilde{\mathbf{v}})}{\partial \tilde{\mathbf{x}}} & \frac{\partial \mathcal{H}(\tilde{\mathbf{v}})}{\partial a} \end{pmatrix}. \tag{C.4}$$

If $\mathcal{H}$ were linear then $\mathcal{H}(\mathbf{x}) = \mathbf{H}\mathbf{x}$.
An example of a j^{th} row of this matrix where $1 < j < P$ is

$$\begin{pmatrix} 0 & 0 & \ldots & 0 & \frac{\partial \mathcal{H}(\tilde{\mathbf{v}})^j}{\partial \tilde{\mathbf{x}}^I} & \frac{\partial \mathcal{H}(\tilde{\mathbf{v}})^j}{\partial \tilde{\mathbf{x}}^{II}} & 0 & \ldots & 0 & \frac{\partial \mathcal{H}(\tilde{\mathbf{v}})^j}{\partial a} \end{pmatrix}. \tag{C.5}$$

The sum of $\frac{\partial \mathcal{H}(\tilde{\mathbf{v}})^j}{\partial \tilde{\mathbf{x}}^I}$ and $\frac{\partial \mathcal{H}(\tilde{\mathbf{v}})^j}{\partial \tilde{\mathbf{x}}^{II}}$ should equal 1, as they represent interpolation.

$$\frac{\partial \mathcal{H}(\tilde{\mathbf{v}})^j}{\partial \tilde{\mathbf{x}}^I} = \frac{\tilde{z}^{\mathrm{II}} - z_o^j}{\tilde{z}^{\mathrm{II}} - \tilde{z}^{\mathrm{I}}}. \tag{C.6}$$

$$\frac{\partial \mathcal{H}(\tilde{\mathbf{v}})^j}{\partial \tilde{\mathbf{x}}^{II}} = \frac{z_o^j - \tilde{z}^{\mathrm{I}}}{\tilde{z}^{\mathrm{II}} - \tilde{z}^{\mathrm{I}}}. \tag{C.7}$$

(C.6) and (C.7) depend on a through the definitions for $\tilde{z}^{\mathrm{I}}$ and $\tilde{z}^{\mathrm{II}}$. To evaluate $\frac{\partial \mathcal{H}(\tilde{\mathbf{v}})}{\partial a}$ let

$$\mathcal{H}(\tilde{\mathbf{v}})^j = \left(\frac{f}{h}\right)\tilde{x}^{\mathrm{I}} + \left(\frac{g}{h}\right)\tilde{x}^{\mathrm{II}}, \tag{C.8}$$

where $f = \tilde{z}^{\mathrm{II}} - z_o^j$, $g = z_o^j - \tilde{z}^{\mathrm{I}}$ and $h = \tilde{z}^{\mathrm{II}} - \tilde{z}^{\mathrm{I}}$. Then using the chain rule

$$\frac{\partial \mathcal{H}(\tilde{\mathbf{v}})}{\partial a} = \left(\frac{\frac{\partial f}{\partial a}h - \frac{\partial h}{\partial a}f}{h^2}\right)\tilde{\mathbf{x}}^I + \left(\frac{\frac{\partial g}{\partial a}h - \frac{\partial h}{\partial a}g}{h^2}\right)\tilde{\mathbf{x}}^{II} \tag{C.9}$$

where $\frac{\partial f}{\partial a} = D(z^{II})$, $\frac{\partial g}{\partial a} = D(z^I)$ and $\frac{\partial h}{\partial a} = D(z^{II}) - D(z^I)$. If $D(z^I)$ and $D(z^{II}) = 0$ then $\frac{\partial \mathcal{H}(\tilde{\mathbf{v}})}{\partial a} = 0$, i.e. if the observation is outside the inversion region defined by $D(z)$.

C.1 Higher order terms

If $\mathcal{H}$ is non-linear then in the Taylor expansion of $\mathcal{H}$ about $\tilde{\mathbf{v}}$ there exist higher order terms.

$$\mathcal{H}(\tilde{\mathbf{v}}^{n+1}) = \mathcal{H}(\tilde{\mathbf{v}}^n) + \mathbf{H}\delta\tilde{\mathbf{v}} + \delta\tilde{\mathbf{v}}^{\mathrm{T}}\nabla^2\mathcal{H}\delta\tilde{\mathbf{v}} \tag{C.10}$$

where

$$\delta\tilde{\mathbf{v}}^{\mathrm{T}}\nabla^2\mathcal{H}\delta\tilde{\mathbf{v}} = \left(\begin{array}{ccccccc} \delta\tilde{\mathbf{v}}^{\mathrm{T}}\nabla^2\mathcal{H}^1\delta\tilde{\mathbf{v}} & ... & \delta\tilde{\mathbf{v}}^{\mathrm{T}}\nabla^2\mathcal{H}^j\delta\tilde{\mathbf{v}} & ... & \delta\tilde{\mathbf{v}}^{\mathrm{T}}\nabla^2\mathcal{H}^P\delta\tilde{\mathbf{v}} \end{array}\right)^{\mathrm{T}} \tag{C.11}$$

$\nabla^2\mathcal{H}^j$ is a $(N+1)\times(N+1)$ matrix:

$$\nabla^2\mathcal{H}^j = \left(\begin{array}{cc} \frac{\partial^2\mathcal{H}^j}{\partial\tilde{\mathbf{x}}^2} & \frac{\partial^2\mathcal{H}^j}{\partial\tilde{\mathbf{x}}\partial a} \\ \frac{\partial^2\mathcal{H}^j}{\partial a\partial\tilde{\mathbf{x}}} & \frac{\partial^2\mathcal{H}^j}{\partial a^2} \end{array}\right) \tag{C.12}$$

where $\frac{\partial^2\mathcal{H}^j}{\partial\tilde{\mathbf{x}}^2}$, a $N\times N$ matrix, is zero everywhere as seen from the definition of $\mathcal{H}$ in (C.3) and

$$\frac{\partial^2\mathcal{H}^j}{\partial\tilde{\mathbf{x}}\partial a} = \left(\frac{\partial^2\mathcal{H}^j}{\partial a\partial\tilde{\mathbf{x}}}\right)^{\mathrm{T}} = (0,...,0,\frac{\partial^2\mathcal{H}^j}{\partial\tilde{x}^{\mathrm{I}}\partial a},\frac{\partial^2\mathcal{H}^j}{\partial\tilde{x}^{\mathrm{II}}\partial a},0,...,0), \tag{C.13}$$

where $\tilde{x}^{\mathrm{I}}$ and $\tilde{x}^{\mathrm{II}}$ are the values of $\mathbf{x}$ in floating level space associated with levels $\tilde{z}^{\mathrm{I}}$ and $\tilde{z}^{\mathrm{II}}$ which are the levels either side of the location of the j^{th} observation. Therefore

$$\delta\tilde{\mathbf{v}}^{\mathrm{T}}\nabla^2\mathcal{H}^j\delta\tilde{\mathbf{v}} = 2\delta a\left(\frac{\partial^2\mathcal{H}^j}{\partial\tilde{x}^{\mathrm{I}}\partial a}\delta\tilde{x}^{\mathrm{I}} + \frac{\partial^2\mathcal{H}^j}{\partial\tilde{x}^{\mathrm{II}}\partial a}\delta\tilde{x}^{\mathrm{II}}\right) + \delta a^2\frac{\partial^2\mathcal{H}^j}{\partial a^2} \tag{C.14}$$

where

$$\frac{\partial^2\mathcal{H}^j}{\partial\tilde{x}^{\mathrm{I}}\partial a} = \frac{\frac{\partial f}{\partial a}h - \frac{\partial h}{\partial a}f}{h^2} \tag{C.15}$$

$$\frac{\partial^2 \mathcal{H}^j}{\partial \tilde{x}^{\mathrm{II}} \partial a} = \frac{\frac{\partial g}{\partial a} h - \frac{\partial h}{\partial a} g}{h^2} \tag{C.16}$$

$$\frac{\partial^2 \mathcal{H}^j}{\partial a^2} = \frac{2h\frac{\partial h}{\partial a}}{h^4} \left[(\frac{\partial h}{\partial a} f - \frac{\partial f}{\partial a} h) \tilde{x}^I + (\frac{\partial h}{\partial a} g - \frac{\partial g}{\partial a} h) \tilde{x}^{II} \right] \tag{C.17}$$

Appendix D

The non-Gaussian cost function for the floating BL scheme

A Gaussian distribution describes well the PDF of many variables. It is however difficult to model the distribution of the errors of the variable a (which controls the degree of the vertical shifting of the levels in the floating BL scheme) as Gaussian. The variable a has a PDF which simultaneously has long tails and a sharp peak which cannot be modelled as Gaussian (see Figure D.1). The consequences of assimilating the (non-Gaussian) a with a Gaussian PDF are discussed in Section 4.3.

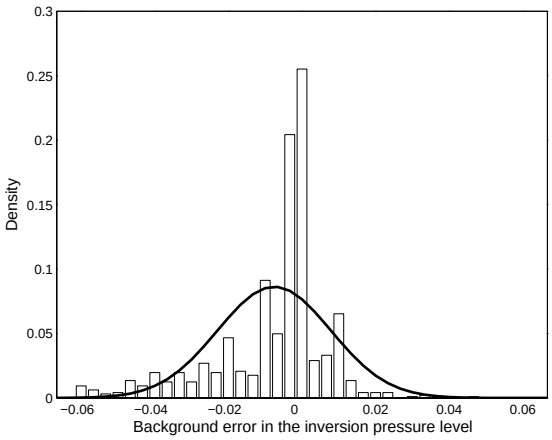


Figure D.1: Histogram of the error in the background inversion height and an approximation to this with a Gaussian distribution $\sim N(-0.008, 0.0162)$ (black line).

A more accurate description of the error distribution of a is the sum of two Gaussian curves. One Gaussian distribution has a large standard deviation (σ_{a1}) to capture the long tails, which represent the cases when the background has a large error in the height of inversion. These values are important as they can lead to a poor forecast of such things as stratocumulus (see Chapter 5). The second Gaussian can be given a small standard distribution (σ_{a2}) to accurately represent the probability that the background inversion height is good; this will reduce the chance of the floating BL scheme moving the background inversion more often than is necessary. An example of the sum is shown in Figure D.2.

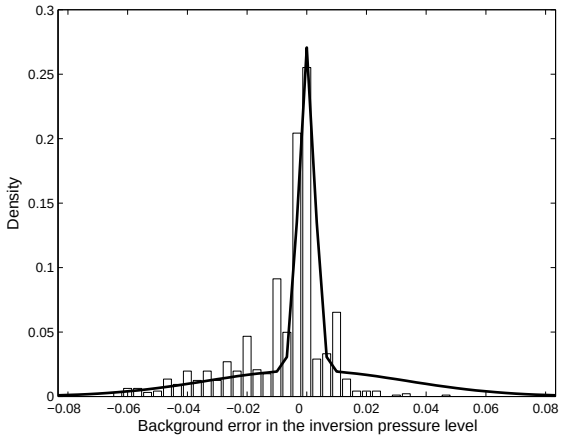


Figure D.2: The error in the background inversion height approximated as the sum of two Gaussian distributions.

The sum of two Gaussians is no longer Gaussian and so using this approximation to the errors of a means the cost function is no longer quadratic.

D.1 The non-Gaussian background cost term

The probability of the background, $\tilde{\mathbf{v}}^{\mathrm{b}} = \begin{pmatrix} \tilde{\mathbf{x}}^{\mathrm{b}} \\ a = 0 \end{pmatrix}$, can now be described as

$$P_{\mathrm{B}} = (1 - \mathrm{A})P_1 + \mathrm{A}P_2, \tag{D.1}$$

where A is a constant: $0 \leq \mathrm{A} \leq 1$ and

$$P_i = \frac{1}{(2\pi)^{\frac{N+1}{2}}|\tilde{\mathbf{B}}_{\mathrm{v}i}|^{1/2}} \exp\left\{-\frac{1}{2}(\tilde{\mathbf{v}} - \tilde{\mathbf{v}}^{\mathrm{b}})^{\mathrm{T}}\tilde{\mathbf{B}}_{\mathrm{v}i}^{-1}(\tilde{\mathbf{v}} - \tilde{\mathbf{v}}^{\mathrm{b}})\right\}. \tag{D.2}$$

$i = 1, 2$ and N is the length of the background vector, $\mathbf{x}^{\mathrm{b}}$. $\tilde{\mathbf{B}}_{\mathrm{v}i}$ is the background error covariance matrix for the background $\tilde{\mathbf{v}}^{\mathrm{b}}$ (see Section 3.1).

$$\tilde{\mathbf{B}}_{\mathrm{v}i} = \begin{pmatrix} \tilde{\mathbf{B}}_{\mathrm{x}} & 0 \\ 0 & \sigma_{ai}^2 \end{pmatrix} \tag{D.3}$$

Since errors in a are uncorrelated with errors in $\tilde{\mathbf{x}}^{\mathrm{b}}$, the part of the PDF concerned with a may be written as a separate part of the PDF as follows (the PDF for errors in $\tilde{\mathbf{x}}^{\mathrm{b}}$ remains Gaussian, while only the PDF for a becomes non-Gaussian):

$$P_i = \frac{1}{(2\pi)^{\frac{N+1}{2}}|\tilde{\mathbf{B}}_{\mathrm{x}}|^{1/2}} \exp\left\{-\frac{1}{2}(\tilde{\mathbf{x}} - \tilde{\mathbf{x}}^{\mathrm{b}})^{\mathrm{T}}\tilde{\mathbf{B}}_{\mathrm{x}}^{-1}(\tilde{\mathbf{x}} - \tilde{\mathbf{x}}^{\mathrm{b}})\right\} \frac{1}{\sigma_{ai}} \exp\left\{-\frac{a^2}{2\sigma_{ai}^2}\right\}. \tag{D.4}$$

Substituting (D.4) into (D.1) gives

$$\begin{aligned} P_{\mathrm{B}} = \frac{1}{(2\pi)^{\frac{N+1}{2}}|\tilde{\mathbf{B}}_{\mathrm{x}}|^{1/2}} \exp\left\{-\tfrac{1}{2}(\tilde{\mathbf{x}} - \tilde{\mathbf{x}}^{\mathrm{b}})^{\mathrm{T}}\tilde{\mathbf{B}}_{\mathrm{x}}^{-1}(\tilde{\mathbf{x}} - \tilde{\mathbf{x}}^{\mathrm{b}})\right\} \times \\ \left\{\tfrac{1-\mathrm{A}}{\sigma_{a1}} \exp\left\{-\tfrac{a^2}{2\sigma_{a1}^2}\right\} + \tfrac{\mathrm{A}}{\sigma_{a2}} \exp\left\{-\tfrac{a^2}{2\sigma_{a2}^2}\right\}\right\}. \end{aligned} \tag{D.5}$$

The background cost function is commonly defined as

$$J_{\mathrm{B}} = -\ln P_{\mathrm{B}} + \mathrm{c}, \tag{D.6}$$

where c is a constant which has no affect on the minimisation of the cost function and so can be added without altering the analysis. Substituting equation (D.5) into (D.6) gives

$$J_B = \frac{1}{2}(\tilde{\mathbf{x}} - \tilde{\mathbf{x}}^{\mathrm{b}})^{\mathrm{T}}\tilde{\mathbf{B}}_{\mathrm{x}}^{-1}(\tilde{\mathbf{x}} - \tilde{\mathbf{x}}^{\mathrm{b}}) - \ln\left\{\frac{1-\mathrm{A}}{\sigma_{\mathrm{a1}}}\exp\left\{-\frac{\mathrm{a}^2}{2\sigma_{\mathrm{a1}}^2}\right\} + \frac{\mathrm{A}}{\sigma_{\mathrm{a2}}}\exp\left\{-\frac{\mathrm{a}^2}{2\sigma_{\mathrm{a2}}^2}\right\}\right\}, \tag{D.7}$$

where the constant c has been added:

$$\mathrm{c} = \ln\left\{\frac{1}{(2\pi)^{\frac{\mathrm{N}+1}{2}}|\tilde{\mathbf{B}}_{\mathrm{x}}|^{1/2}}\right\}. \tag{D.8}$$

The choice of A dictates the shape of the function given in Figure D.2. As A tends to 0, P_{B} tends to a Gaussian with $\sigma_a = \sigma_{a1}$. As A tends to 1, P_{B} tends to a Gaussian with $\sigma_a = \sigma_{a2}$.

D.2 The gradient of the background cost function

The gradient of the background cost function can be split with respect to $\tilde{\mathbf{x}}$ and a:

$$\nabla_{\tilde{\mathbf{v}}} J_{\mathrm{B}} = \begin{pmatrix} \nabla_{\tilde{\mathbf{x}}} J_{\mathrm{B}} \\ \nabla_a J_{\mathrm{B}} \end{pmatrix}. \tag{D.9}$$

Firstly differentiate J_B with respect to $\tilde{\mathbf{x}}$

$$\nabla_{\tilde{\mathbf{x}}} J_{\mathrm{B}} = \tilde{\mathbf{B}}_{\mathrm{x}}^{-1}(\tilde{\mathbf{x}} - \tilde{\mathbf{x}}^{\mathrm{b}}). \tag{D.10}$$

Next differentiate with respect to a

$$\nabla_a J_{\mathrm{B}} = \frac{\frac{(1-\mathrm{A})a}{\sigma_{a1}^3}\exp\left\{-\frac{a^2}{2\sigma_{a1}^2}\right\} + \frac{\mathrm{A}a}{\sigma_{a2}^3}\exp\left\{-\frac{a^2}{2\sigma_{a2}^2}\right\}}{\frac{(1-\mathrm{A})}{\sigma_{a1}}\exp\left\{-\frac{a^2}{2\sigma_{a1}^2}\right\} + \frac{\mathrm{A}}{\sigma_{a2}}\exp\left\{-\frac{a^2}{2\sigma_{a2}^2}\right\}}. \tag{D.11}$$

D.3 Gradient of full cost function

We can now calculate the gradient of the full cost function, bringing in the observation term which has remained quadratic.

$$\nabla_{\tilde{\mathbf{v}}} J = \nabla_{\tilde{\mathbf{v}}} J_{\mathrm{B}} + \nabla_{\tilde{\mathbf{v}}} J_{\mathrm{O}}, \tag{D.12}$$

where (see Section 2.1 and (2.12))

$$\nabla_{\tilde{\mathbf{v}}} J_{\mathrm{O}} = -\mathbf{H}^{\mathrm{T}}\mathbf{R}^{-1}(\mathbf{y} - \mathbf{H}(\tilde{\mathbf{v}} - \tilde{\mathbf{v}}^{\mathrm{b}}) - \mathcal{H}(\tilde{\mathbf{v}}^{\mathrm{b}})). \tag{D.13}$$

This leads to

$$\nabla_{\tilde{\mathbf{v}}} J = \begin{pmatrix} \tilde{\mathbf{B}}_{\mathrm{x}}^{-1}(\tilde{\mathbf{x}} - \tilde{\mathbf{x}}^{\mathrm{b}}) \\ \frac{\frac{(1-A)a}{\sigma_1^3}\exp\left\{-\frac{\mathrm{a}^2}{2\sigma_1^2}\right\} + \frac{\mathrm{Aa}}{\sigma_2^3}\exp\left\{-\frac{\mathrm{a}^2}{2\sigma_2^2}\right\}}{\frac{(1-A)}{\sigma_1}\exp\left\{-\frac{\mathrm{a}^2}{2\sigma_1^2}\right\} + \frac{\mathrm{A}}{\sigma_2}\exp\left\{-\frac{\mathrm{a}^2}{2\sigma_2^2}\right\}} \end{pmatrix} - \mathbf{H}^{\mathrm{T}}\mathbf{R}^{-1}(\mathbf{y} - \mathbf{H}(\tilde{\mathbf{v}} - \tilde{\mathbf{v}}^{\mathrm{b}}) - \mathcal{H}(\tilde{\mathbf{v}}^{\mathrm{b}})). \quad \text{(D.14)}$$

In order to find the approximate value of $\tilde{\mathbf{v}}$ when $\nabla_{\tilde{\mathbf{v}}} J = 0$ we assume

$$\tilde{\mathbf{v}}_{n+1} \approx \tilde{\mathbf{v}}_n - \mathbf{A}^{-1}\nabla_{\tilde{\mathbf{v}}} J, \quad \text{(D.15)}$$

where n is the iteration number, and $\mathbf{A}$ is the Hessian of the cost function.

D.4 The Hessian

To evaluate the Hessian of the cost function to evaluate (D.15), we may split $\mathbf{A}$ into a background and an observation term.

$$\mathbf{A} = \nabla^2_{\tilde{\mathbf{v}}} J = \begin{pmatrix} \nabla_{\tilde{\mathbf{x}}}\nabla_{\tilde{\mathbf{x}}} J_{\mathrm{B}} & 0 \\ 0 & \nabla_a \nabla_a J_{\mathrm{B}} \end{pmatrix} + \nabla^2_{\tilde{\mathbf{v}}} J_{\mathrm{o}}. \quad \text{(D.16)}$$

Again the amplitude error term and positional error term for the background separate out as the errors are uncorrelated. $\nabla_{\tilde{\mathbf{x}}}\nabla_{\tilde{\mathbf{x}}} J_{\mathrm{B}}$ and $\nabla^2_{\tilde{\mathbf{v}}} J_{\mathrm{o}}$ remain the same as in previous Gaussian theory,

$$\nabla_{\tilde{\mathbf{x}}}\nabla_{\tilde{\mathbf{x}}} J_{\mathrm{B}} = \tilde{\mathbf{B}}_{\mathrm{x}}^{-1}, \quad \text{(D.17)}$$

$$\nabla^2_{\tilde{\mathbf{v}}} J_{\mathrm{o}} = \mathbf{H}^{\mathrm{T}}\mathbf{R}^{-1}\mathbf{H}. \quad \text{(D.18)}$$

The non-Gaussian term, $\nabla_a \nabla_a J_{\mathrm{B}}$, is slightly more complicated. Let

$$\nabla_a J_{\mathrm{B}} = \frac{f(a)}{g(a)}, \quad \text{(D.19)}$$

where $f(a)$ and $g(a)$ are the numerator and denominator of (D.11) respectively, then

$$\nabla_a \nabla_a J_{\mathrm{B}} = \frac{f'(a)g(a) - f(a)g'(a)}{g(a)^2}, \quad \text{(D.20)}$$

where $'$ denotes the derivative of the function with respect to a. These derivatives can be evaluated as

$$f'(a) = \frac{1-\mathrm{A}}{\sigma_{a1}^3}\left(1 - \frac{a^2}{\sigma_{a1}^2}\right)\exp\left\{-\frac{a^2}{2\sigma_{a1}^2}\right\} + \frac{\mathrm{A}}{\sigma_{a2}^3}\left(1 - \frac{a^2}{\sigma_{a2}^2}\right)\exp\left\{-\frac{a^2}{2\sigma_{a2}^2}\right\} \quad \text{(D.21)}$$

and

$$g'(a) = \frac{-(1-\mathrm{A})a}{\sigma_{a1}^3}\exp\left\{-\frac{a^2}{2\sigma_{a1}^2}\right\} + \frac{-\mathrm{A}a}{\sigma_{a2}^3}\exp\left\{-\frac{a^2}{2\sigma_{a2}^2}\right\}. \quad \text{(D.22)}$$

Appendix E

The Smith scheme and its linearised transform

The Smith cloud scheme is introduced in Section 5.3. This scheme is based on that of Smith (1990) and gives a diagnosis of cloud amount and water vapour based on a statistical distribution of thermodynamic and water content variables about their grid box mean. In Section E.1 the full equations are given in order to express the cloud conserved variables θ_L and $\ln q_t$ in terms of model variables T and $\ln q$. In Section E.2 it is shown how to linearise the Smith scheme analytically for use in Var as a control variable transform.

E.1 Smith scheme

The basic equations in Smith (1990) describe cloud conserved variables; liquid water temperature, T_L, and total water content, q_t.

$$T_L = T - \frac{Lq_c}{c_p}, \tag{E.1}$$

where q_c is specific cloud water content, L is latent heat of condensation, c_p is isobaric specific heat capacity (1.005×10^3 J kg^{-1} K^{-1})

$$q_t = q + q_c \tag{E.2}$$

where q is specific humidity. (E.1) and (E.2) are given by equations (2.4) and (2.5) in Smith (1990) respectively. The second term in (E.1) represents the temperature change due to water condensing to form cloud, which the value of T_L is unaffected by.

For the assimilation of cloud conserved variables in the floating BL scheme the following variables are used; liquid water potential temperature, θ_L, and natural logarithm of total water content, $\ln q_t$. These variables benefit from being conserved with changes in pressure which is essential for the floating BL scheme (Section 3.1.4) and have error statistics which are unbounded and so may be modelled as Gaussian.
θ_L and $\ln q_t$ may be expressed in terms of the model variables T and $\ln q$ in the following

way:

$$T = \theta_L \left(\frac{p}{p_o} \right)^{R/c_p} + \frac{Lq_c}{c_p} \tag{E.3}$$

where p is the pressure of the level, p_o is the reference pressure (generally the surface pressure), and R is specific gas constant ($287.05 JKg^{-1}K^{-1}$).

$$\ln q = \ln \left(\exp(\ln q_t) - q_c. \right) \tag{E.4}$$

To evaluate the specific cloud water content, q_c, the Smith cloud scheme uses the following equation (equivalent to equation 2.24 in Smith (1990) for the assumed sech^2 distribution of the subgrid scale fluctuations of total water content given by Sharpe (2007), described in Section 5.3):

$$q_c = Q_c + \delta \ln(\exp(-\frac{Q_c}{\delta}) + 1) \tag{E.5}$$

where

$$Q_c = a_L(q_t - q_{sat}(T_L)) \tag{E.6}$$

(the grid box mean difference between specific total water content and the saturation specific humidity, q_{sat}, equivalent to equation 2.19 in Smith (1990))

$$\delta = \frac{\sqrt{6}}{4}\sigma_s \tag{E.7}$$

$$\sigma_s = \frac{1 - RH_c}{\sqrt{6}} q_{sat}(T_L) a_L \tag{E.8}$$

(the standard deviation of the subgrid scale fluctuations of total water content, s, where RH_c is the critical relative humidity value at which water vapour will condense.)

$$a_L = \left(1 + \frac{L\alpha_L}{c_p} \right)^{-1} \tag{E.9}$$

(equation 2.15 in Smith (1990))

$$\alpha_L = \frac{\partial q_{sat}(T)}{\partial T} \tag{E.10}$$

(equation 2.14 in Smith (1990))

$$q_{sat}(T) = \frac{0.62198 e_s}{p - e_s}, \tag{E.11}$$

using the Goff and Gratch (1946) formula for saturated vapour pressure e_s.

E.2 The linearised Smith scheme

The first order tangent linearisation of the Smith scheme may be expressed as a matrix $\mathbf{U}$ of dimension $N \times N$ where N is the length of the state vector, $\mathbf{x}$. $\mathbf{U}$ is made up of the

following block diagonal matrices:

$$\mathbf{U} = \frac{\partial \mathbf{x}}{\partial \mathbf{w}} = \begin{pmatrix} \frac{\partial \mathbf{T}}{\partial \theta_{\mathbf{L}}} & \frac{\partial \mathbf{T}}{\partial \ln \mathbf{q_t}} \\ \frac{\partial \ln \mathbf{q}}{\partial \theta_{\mathbf{L}}} & \frac{\partial \ln \mathbf{q}}{\partial \ln \mathbf{q_t}} \end{pmatrix} \tag{E.12}$$

E.2.1 Partial derivatives

U may be expressed analytically by calculating the partial derivatives given in (E.12).

Main partial derivatives

$$\frac{\partial T}{\partial \theta_L} = \frac{\partial T}{\partial T_L}\frac{\partial T_L}{\partial \theta_L} = \left(\frac{p}{p_o}\right)^{R/c_p}\left[\frac{\partial T_L}{\partial T_L} + \frac{L}{c_p}\frac{\partial q_c}{\partial T_L}\right] \tag{E.13}$$

$$\frac{\partial \ln q}{\partial \theta_L} = \frac{\partial \ln q}{\partial q}\frac{\partial q}{\partial T_L}\frac{\partial T_L}{\partial \theta_L} = -\frac{1}{q}\left(\frac{p}{p_o}\right)^{R/c_p}\left[\frac{\partial q_c}{\partial T_L}\right] \tag{E.14}$$

assuming q_t is independent of T_L.

$$\frac{\partial T}{\partial \ln q_t} = \frac{\partial T}{\partial q_t}\frac{\partial q_t}{\partial \ln q_t} = q_t\left(\frac{L}{c_p}\frac{\partial q_c}{\partial q_t}\right) \tag{E.15}$$

$$\frac{\partial \ln q}{\partial \ln q_t} = \frac{\partial \ln q}{\partial q}\frac{\partial q}{\partial q_t}\frac{\partial q_t}{\partial \ln q_t} = \frac{q_t}{q}\left(\frac{\partial q_t}{\partial q_t} - \frac{\partial q_c}{\partial q_t}\right) \tag{E.16}$$

E.2.2 Secondary partial derivatives

$$\frac{\partial q_c}{\partial T_L} = \frac{\partial q_c}{\partial Q_c}\frac{\partial Q_c}{\partial T_L} + \frac{\partial q_c}{\partial \delta}\frac{\partial \delta}{\partial T_L} \tag{E.17}$$

$$\frac{\partial q_c}{\partial Q_c} = 1 - (e^{\{-\frac{Q_c}{\delta}\}})(e^{\{-\frac{Q_c}{\delta}\}} + 1)^{-1} \tag{E.18}$$

$$\frac{\partial q_c}{\partial \delta} = \ln\left(e^{\{-\frac{\mathbf{Q_c}}{\delta}\}} + 1\right) + \left(Q_c e^{\{-\frac{Q_c}{\delta}\}}\right)\left(\delta\left(e^{\{-\frac{Q_c}{\delta}\}} + 1\right)\right)^{-1} \tag{E.19}$$

$$\frac{\partial Q_c}{\partial T_L} = \frac{\partial Q_c}{\partial a_L}\frac{\partial a_L}{\partial T_L} + \frac{\partial Q_c}{\partial q_{sat}(T_L)}\frac{\partial q_{sat}(T_L)}{\partial T_L} \tag{E.20}$$

$$\frac{\partial Q_c}{\partial a_L} = q_t - q_{sat}(T_L) \tag{E.21}$$

$$\frac{\partial Q_c}{\partial q_{sat}(T_L)} = -a_L \tag{E.22}$$

$$\frac{\partial a_L}{\partial T_L} = -\left(1 + \frac{L\alpha_L}{c_p}\right)^{-2}\frac{L}{c_p}\frac{\partial \alpha_L}{\partial T_L} \tag{E.23}$$

$$\frac{\partial \delta}{\partial T_L} = \frac{\sqrt{6}}{4}\frac{\partial \sigma_s}{\partial T_L} \tag{E.24}$$

$$\frac{\partial \sigma_s}{\partial T_L} = \frac{\partial \sigma_s}{\partial a_L}\frac{\partial a_L}{\partial T_L} + \frac{\partial \sigma_s}{\partial q_{sat}(T_L)}\frac{\partial q_{sat}(T_L)}{\partial T_L} \tag{E.25}$$

$$\frac{\partial \sigma_s}{\partial a_L} = \frac{1 - RH_c}{\sqrt{6}} q_{sat}(T_L) \tag{E.26}$$

$$\frac{\partial \sigma_s}{\partial q_{sat}(T_L)} = \frac{1 - RH_c}{\sqrt{6}} a_L \tag{E.27}$$

$$\frac{\partial q_{sat}(T_L)}{\partial T_L} = \left(\frac{0.612198p}{(p - e_s)^2} \right) \frac{\partial e_s}{\partial T_L} \tag{E.28}$$

$$\frac{\partial q_c}{\partial q_t} = \frac{\partial q_c}{\partial Q_c} \frac{\partial Q_c}{\partial q_t} \tag{E.29}$$

$$\frac{\partial Q_c}{\partial q_t} = a_L \tag{E.30}$$

Appendix F

Notation

symbol	Meaning
$\mathbf{A}$	Hessian of cost function
a	Positional error control variable
$\mathbf{B}$	Background error covariance matrix
B_λ	Energy flux emitted by a black body
$\tilde{\mathbf{B}}_{\mathrm{v}}$	Background error covariance matrix for floating BL control variables
C	Cloud fraction
c_p	Specific heat capacity
$D(z)$	Displacement function
e_s	Saturated water vapour pressure
$\mathcal{H}$	Observation operator
$\mathbf{H}$	Linearised observation operator
J	Cost function
J_b	Background part of cost function
J_o	Observation part of cost function
$\mathbf{K}$	Kalman gain matrix
L_c	Latent heat of condensation
L_c	Latent heat of fusion
L_v	Latent heat of evaporation ($2.5 \times 10^6 Jkg^{-1}$)
L_λ	Radiance
$\mathbf{M}$	Tangent linearisation of the forecast model
$\mathcal{M}$	Forecast model
N	Length of state vector
n	Iteration index (subscript)
P	Length of observation vector
$\mathbf{P}^{\mathrm{a}}$	Analysis error covariance matrix
$\mathbf{P}^{\mathrm{f}}$	Forecast error covariance matrix
p	pressure
p_o	reference pressure
Q	length of control vector
$\mathbf{Q}$	Model error covariance matrix

symbol	Meaning
q	Specific humidity
q_c	Specific cloud water content
q_F	Specific ice content
q_L	Specific liquid water content
q_t	Total water content
$\mathbf{R}$	Observation covariance matrix
R	Specific gas constant ($287.05 Jkg^{-1}K^{-1}$)
R_v	Specific gas constant of water vapour ($461.51 Jkg^{-1}K^{-1}$)
T	Temperature
T_L	Liquid-frozen temperature
$\mathbf{U}$	Control Variable transform
$\tilde{\mathbf{v}}$	Floating variable vector $\begin{pmatrix} \tilde{\mathbf{x}} \\ a \end{pmatrix}$
W_λ	Weighting function
$\mathbf{x}$	State vector
$\tilde{\mathbf{x}}$	State vector on floating model levels
$\mathbf{x}^{\mathrm{a}}$	Analysis
$\mathbf{x}^{\mathrm{b}}$	Background vector
$\mathbf{x}^{\mathrm{t}}$	True state
$\mathbf{y}$	Observation vector
$\mathbf{Z}$	Interpolation matrix used in §4.1.2.
z	model levels
$\tilde{z}$	fixed model levels
η^{a}	Analysis error
η^{b}	Background error
η^{o}	Observation error
λ	wavelength
θ	potential temperature
τ_λ	transmittance

Appendix G

Acronyms

Term	Meaning
1DVar	One dimensional variational data assimilation
3DVar	Three dimensional variational data assimilation
4DVar	Four dimensional variational data assimilation
ATOVS	Advanced TIROS Operational Vertical Sounder
BL	Boundary Layer
CALIPSO	Cloud-Aerosol Lidar and Infrared Pathfinder Satellite
CTH	Cloud Top Height
DA	Data assimilation
ECMWF	European Centre for Medium-range Weather Forecasting
EKF	Extended Kalman Filter
EnKF	Ensemble Kalman Filter
FA	Free atmosphere
IASI	Infrared Atmospheric Sounding Interferometer
KF	Kalman filter
lidar	LIght Detection And Ranging
LCL	Lifting Condensation Level
LW	Long Wave
ML	Mixed Layer
MetO	UK Met Office
NAE	North Atlantic and Europe
NCEP	National Centre of Environmental Prediction
NMC	National Meteorological Centre
NWP	Numerical weather prediction
NWPSAF	Numerical weather prediction satellite application facility
OI	Optimal interpolation
PDF	Probability Distribution Function
radar	RAdio Detection And Ranging
RL	Residual Layer
RTTOV	Radiative Transfer for TOVs

Term	Meaning
SBL	Stable Boundary Layer
Sc	Stratocumulus
SCM	Successive correction method
SL	Stable Layer
SST	Sea Surface Temperature
SW	Short Wave
TIROS	Television and Infrared Observational Satellite
TOA	Top Of Atmosphere
UK	United Kingdom
UM	Unified Model
Var	Variational data assimilation

References

Alexander, G. D., J. A. Weinman, J. L. Schols, 1998: The use of digital warping of microwave integrated water vapour imagery to improve forecasts of marine extratropical cyclones. *Mon. Wea. Rev.*, **126**, 1469-1495.

Anderson, J. L., 2001: An ensemble adjustment Kalman filter for data assimilation. *Mon. Wea. Rev.*, **129**, 2884-2903.

Andersson, E., H. Järvinen, 1999: Variational quality control. *Q. J. R. Meteorol. Soc.*, **125**, 697-722.

Bannister, R. N. 2008a: A review of forecast error covariance statistics in atmospheric variational data assimilation. I: Characteristics and measurements of forecast error covariances. *Q. J. R. Meteorol. Soc.*, **134**, 1951-1970.

Bannister, R. N. 2008b: A review of forecast error covariance statistics in atmospheric variational data assimilation. II: Modelling the forecast error covariances statistics. *Q. J. R. Meteorol. Soc.*, **134**, 1971-1996.

Barlow, R. J., 1989: A guide to the use of statistical methods in physical sciences. John Wiley & Sons Ltd, Chichester.

Barry, R. G., R. J. Chorley, 1968: Atmosphere, weather and climate. Routledge, New York.

Bennett, A. F., 1992: Inverse modelling in ocean and atmosphere. Cambridge Uni. Press.

Bennett, A. F., 2002: Inverse methods in physical oceanography. Cambridge Uni. Press, Cambridge.

Bergthorsson, P., and B. Döös, 1955: Numerical weather map analysis, *Tellus*, **7**, 329-340.

Berre, L., 2000: Estimation of synoptic and mesoscale forecast error covariances in a limited area model. *Mon. Wea. Rev.*, **128**, 644-667.

Berre, L., S. E. Ştefănescu, M. Belo, 2006: The representation of the analysis effect in three error simulation techniques. *Tellus*, **58A**, 196-209.

Bowler , N. E., A. Arribas, K. R. Mylne, K. B. Robertson, S. E. Beare 2008: The MOGREPS short-range ensemble prediction system. *Q. J. R. Meteorol. Soc.*, **134** 703-722

Brewster, K. A., 2003: Phase-correcting data assimilation and application to storm-scale numerical weather prediction. Part I: Method, description and simulation testing. *Mon. Wea. Rev.* **131**, 480-492.

Buehner, M., 2005: Ensemble-derived stationary and flow-dependent background error covariances: Evaluation in quasi-operational NWP setting. *Q. J. R. Meteorol. Soc.*, **131**, 1013-1044.

Cressman, G. P., 1959: An operational objective analysis system. *Mon. Wea. Rev.*, **87**, 367-374.

Cullen, M. J. P., 1993: The unified forecast/climate model. *Meteorol. Mag.*, **122**, 81-94.

Dance, S. L., 2004: Issues in high resolution limited area data assimilation for quantitative precipitation forecasting. *Physica D*, **196**, 1-27.

Daley, R., 1997: Atmospheric data assimilation. *J. Meteorol. Soc. Japan*, **75**, 319-329.

Davies T., M. J. P. Cullen, A. J. Malcolm, M. H. Mawson, A. Staniforth, A. A. White, N. Wood, 2005 A new dynamical core for the Met Office's global and regional modelling of the atmosphere. *Q. J. R. Meteorol. Soc.*, **131**, 1759-1782.

Deckmyn, A., L. Berre, 2005: A wavelet approach to representing background error covariances in a limited-area model. *Mon. Wea. Rev.*, **133**, 1279 -1294.

Dee, D. P., and A. M. da Silva, 1998: Data assimilation in the presence of forecast bias. *Q. J. R. Meteorol. Soc.*, **124**, 269-295.

Dee, D. P., and A. M. da Silva, 2003: The choice of variable for atmospheric moisture analysis. *Mon. Wea. Rev.*, **131**, 155 - 171.

Derber, J., F. Bouttier, 1999: A reformulation of the background error covariance in the ECMWF global data assimilation system. *Tellus*, **51A**, 195-221.

Desroziers, G., J.-P. Lafore 1993: A coordinate transformation for objective frontal analysis. *Mon. Wea. Rev.*, **121**, 1531 - 1533.

Desroziers, G. 1997: A coordinate change for data assimilation of frontal structures. *Mon. Wea. Rev.*, **125**, 3030 - 3038.

Errico, R., 1997: What is an adjoint model? *Bull. Amer. Meteorol. Soc.*, **11**, 2577-2591.

Errico, R., P. Bauer, J. F. Mahfouf, 2007: Issues regarding the assimilation of cloud and precipitation data. *J. Atmos. Sci.*, **64**, 3785-3798.

Etherton, B. J., C. H. Bishop, 2004: Resilience of hybrid ensemble/ 3DVar analysis schemes to model error and ensemble covariance error. *Mon. Wea. Rev.*, **132**, 1065 - 1080.

Evensen, G., 1994: Sequential data assimilation with a nonlinear quasigeostrophic model using Monte Carlo methods to forecast error statistics. *J. Geophys. Res*, **94** (C5), 10143-10162.

Eyre, J. R., 1991: Inversion methods for satellite sounding data. Data assimilation Training Course. ECMWF, Reading, UK.

European Organisation for the Exploitation of Meteorological Satellites (EUMETSAT), 2007: NWPSAF Met Office 1D-Var, Version 3.2.

Fisher, M., M. Leutbecher, G. A. Kelly, 2005: On the equivalence between Kalman smoothing and weak-constraint four-dimensional variational data assimilation. *Q. J. R. Meteorol. Soc.*, **131**, 3235-3246.

Fisher, M., 2003: Background error covariance modelling. pp. 1-28 in Proceedings of seminar on recent developments in data assimilation for atmosphere and ocean, 8-12 September 2003, ECMWF, Reading, UK.

Fowler, A. M., R. N. Bannister, J. Eyre, 2010: Characterising the background errors for the boundary-layer capping inversion. *Austral. Meteorol. Ocean. J.*, **59**, 17-24.

Gandin, L. S., 1963: Objective analysis of meteorological fields, *Gidrometeorologicheskoe Izdatelstvo*, Lenningrad. English translation by Israeli Program for Scientific Translations, Jerusalem, 1965.

Gaspari, G., S. E. Cohn, 1999: Construction of correlation function in two and three dimensions. *Q. J. R. Meteorol. Soc.*, **125**, 723-757.

Ghil, M., S. Cohn, J. Tavantzis, K. Bube, E. Isaacson, 1981: Applications of estimation theory to numerical weather prediction. In *Dynamic Meteorology: Data Assimilation Methods*, L. Bengtsson, M. Ghil and E. Kallen, editors. Springer-Verlag, New York.

Goff, J. A., S. Gratch, 1946: Low-pressure properties of water from -160 to 212 ∘F, in Transactions of the American Society of Heating and Ventilating Engineers, pp 95-122, presented at the 52nd annual meeting of the American Society of Heating and Ventilating Engineers, New York, 1946.

Gyr, A., F.-S. Rys, 1995: Diffusion and transport of pollutants in atmospheric mesoscale flow fields (ECOFTAC Series). Springer.

Hamill, T. M., C. Snyder, 2000: A hybrid ensemble Kalman filter-3D variational analysis scheme. *Mon. Wea. Rev.*, **128**, 2905 - 2919.

Hamill, T. M., C. Snyder, 2002: Using improved background-error covariances from an ensemble Kalman filter for adaptive observations. *Mon. Wea. Rev.*, **130**, 1552 - 1572.

Hoffman, R. N., Z. Liu,J.-F. Louis, C. Grassoti 1995: Distortion representation of forecast errors. *Mon. Wea. Rev.*, **123**, 2758 - 2770.

Hoke, J., R. Anthes, 1976: The initialization of numerical models by a dynamic relaxation technique. *Mon. Wea. Rev.*,**104**, 1551 - 1556.

Hollingsworth, A., P. Lönnberg, 1986: The statistical structure of short-range forecast errors as determined from radiosonde data. Part I: The wind field. *Tellus*, **38A**, 111-136.

Hólm, E., 2002: Revision of the ECMWF humidity analysis: Construction of a Gaussian control variable. Proceedings of the ECMWF/GEWEX Workshop on Humidity Analysis, pp 1- 6.

Hoskins, B. J., 1975: The geostrophic momentum approximation and the semi-geostrophic equations. *J. Atmos, Sci.*, **32**, 233-242.

Houtekamer, P. L., L. Lefaivre, J. Derome, H. Ritchie, H. L. Mitchell, 1996: A system simulation approach to ensemble prediction. *Mon. Wea. Rev.*, **124**, 1225 - 1242.

Ide, K., P. Courtier, M. Ghil, A. C. Lorenc 1997: Unified notation for data assimilation: operational, sequential and variational. *J. Meteorol. Soc. Japan*, **75**, 181-189.

Ingleby, N. B., 2001: The statistical structure of forecast errors and its representation in the Met Office global 3-dimensional variational data assimilation system. *Q. J. R. Meteorol. Soc.*, **127**, 209-231.

Järvinen, H. 2001: Temporal evolution of innovation and residual statistics in the ECMWF variational data assimilation systems. *Tellus*, **53A**, 333-347.

Jones, C. D., B. MacPherson, 1997: A latent heat nudging scheme for assimilation of precipitation data into an operation mesoscale model. *Meteorol. Appl.*, **4**, 269-277.

Kalnay, E. 2003: Atmospheric modelling, data assimilation and predictability. Cambridge University Press, Cambridge, UK.

Leith, C. E., 1978: Predictability of climate. *Nature*, **276**, 352-355.

Leith, C. E., 2006: Objective methods for weather prediction. *Ann. Rev. Fluid Mech.*, **10**, 107-128.

Lewis, J. M., S. Lakshimivarahan, S. K. Dhall, 2006: Dynamic data assimilation: a least squares approach. Cambridge University Press, Cambridge, UK.

Liu, Z.-Q., F. Rabier, 2002: The interaction between model resolution and observation resolution and density in data assimilation: A one-dimensional study. *Q. J. R. Meteorol. Soc.*, **128**, 1367-1386.

Liu, Z.-Q., F. Rabier, 2003: The potential of high density observations for numerical weather prediction. *Q. J. R. Meteorol. Soc.*, **129**, 3013-3035.

Lock, A. P. 1998: The parametrisation of entrainment in cloudy boundary layers, *Q. J. R. Meteorol. Soc.*, **124**, 2729-2753.

Lock, A. P., A. R. Brown, M. R. Bush, G. M. Martin, R. N. B. Smith, 2000: A new boundary layer mixing scheme. Part I: Scheme description and single-column model tests. *Mon. Wea. Rev.*,**128**, 3187 - 3199.

Lönnberg, P., A. Hollingsworth, 1986: The statistical structure of short-range forecast errors as determined from radiosonde data. Part II: The covariance of height and wind errors. *Tellus*, **38A**, 137-161.

Lorenc, A., 1986: Analysis methods for numerical weather prediction. *Q. J. R. Meteorol. Soc.*, **112**, 1177-1194.

Lorenc, A., S. P. Ballard, R. S. Bell, N. B. Ingleby, P. L. F. Andrews, D. M. Barker, J. R. Bray, A. M. Clayton, T. Dalby, D. Li, T. J. Payne, F. W. Saunders, 2000: The Met Office global three-dimensional variational data assimilation scheme. *Q. J. R. Meteorol. Soc.*, **126**, 2991-3012.

Lorenc, A. C., 2007a: Ideas for adding flow-dependence to the Met Office VAR system. ECMWF Workshop on flow-dependent aspects of data assimilation, 11-13 June 2007. http://www.ecmwf.int/ publications/library/do/references/list/14092007

Lorenc, A. C., 2007b: A study of o-b monitoring statistics from radiosondes, composited for low-level cloud layers. Met Office NWP Forecasting Research Technical Report No.504.

Lorenz, E. N., 1969: The predictability of a flow which possesses many scales of motion. *Tellus*, **21**, 289-307.

Mariano, A. J., 1990: Contour analysis: A new approach for melding geophysical fluids. *J. Ocean. Atmos. Technol.*, **7**, 285-295.

Mass, C. F., D. Ovens, K. Westrick and B. A. Colle, 2002: Does increasing horizontal resolution produce more skillful forecasts? *Bull. Am. Meteorol. Soc.*, **83**, 407-430.

Matricardi, M., R. Saunders, 1999 Fast radiative transfer model for simulation of infrared atmospheric sounding interferometer radiances. *Appl. Optics*, **38**, 5679-5691.

McNally, A. P., 2002: A note on the occurrence of cloud in meteorologically sensitive areas and the implications for advanced infrared sounders. *Q. J. R. Meteorol. Soc.*, **128**, 2551-2556.

Mellor, G. L. 1997: The Gaussian cloud model relations. *J. Atmos. Sci.*, **34**, 356-358.

Met Office, 1995: Control variable transforms - parameters. Var Scientific Documentation Paper, **11** (available from Met Office, Fitzroy Road, Exeter, Devon, EX8 3PB, UK).

Parrish, D., F., J. C. Derber, 1992: The National Meteorological Center's spectral statistical interpolation analysis system. *Mon. Wea. Rev.*, **120**, 1747 - 1763.

Pavelin, E.G., S. J. English, J. R. Eyre, 2008: The assimilation of cloud-affected infrared satellite radiances for numerical weather prediction. *Q. J. R. Meteorol. Soc.*, **134**, 737-749.

Purser, R. J., W.-S., Wu, 2003: Numerical aspects of Recursive filters to variational statistical analysis. Part II: Spatially inhomogeneous and anisotropic general covariances. *Mon. Wea. Rev.*, **131**, 1536 - 1548.

Rabier, F. 2005: Overview of global data assimilation developments in numerical weather prediction centers. *Q. J. R. Meteorol. Soc.*, **131**, 3215-3233.

Rabier, F. H. Jarvinen, E. Klinker, J.-F. Mahfouf and A. Simmons 2000: The ECMWF operational implementation of four-dimensional variational assimilation. I: Experimental results with simplified physics. *Q. J. R. Meteorol. Soc.*, **126**, 1143-1170.

Ravela, S., K. Emanuel, D. McLaughlin, 2007: Data assimilation by field alignment. *Physica D*, **230**, 127-145.

Rawlins, F., S. P. Ballard, K. J. Bovis, A. M. Clayton, D. Li, G. W. Inverarity, A. C. Lorenc, T. J. Payne, 2007: The Met Office global four-dimensional variational data assimilation scheme. *Q. J. R. Meteorol. Soc.*, **133**, 347-362.

Riishøjgaard, L. P., 1998: A direct way of specifying flow-dependent background error correlations for meteorological analysis systems. *Tellus*, **50A**, 42-57.

Rodgers, C. D.. 2000: Inverse methods for atmospheric sounding. World Scientific Publishing.

Rogers, R. R. 1979: A short course in cloud physics. Pergamon Press Ltd.

Rutherford, I. D., 1972: Data assimilation by statistical interpolation of forecast error fields. *J. Atmos. Sci.*, **29**, 809-815.

Salby, M. L., 1996: Fundamentals of atmospheric physics. Academic Press.

Seo, K., W. Wang , J. Gottschalck, Q. Zhang, J. Schemm, W. R. Higgins, A. Kumar, 2009: Evaluation of MJO Forecast Skill from Several Statistical and Dynamical Forecast Models. *J. Clim.* **22**, 2372-2388.

Sharpe, M. C., 2007: Incrementing liquid & frozen cloud water and grid-box cloud fractions for VAR observation operators and UM initialisation. Var scientific document paper, **61** (available from Met Office, Fitzroy Road, Exeter, Devon, EX8 3PB, UK).

Simmons, A. J., A. Hollingsworth, 2002: Some aspects of the improvement in skill of numerical weather prediction. *Q. J. R. Meteorol. Soc.*, **128**, 647-677.

Smith, R. N. B. 1990: A scheme for predicting layer clouds and their water content in a general circulation model. *Q. J. R. Meteorol. Soc.*, **116**, 435-460.

Sommeria G., and J.W. Deardorff, 1977: Subgrid-scale condensation in models of non precipitating clouds. *J. Atmos. Sci.*, **34**, 344-355.

Stephens, G. L., D. G. Vane, R. J. Boain, G. G. Mace, K Sassen, Z Wang, A. J. Illingworth, E. J. O'Connor, W. B. Rossow, S. L. Durden, S. D. Miller, R. T. Austin, A. Benedetti, C. Mitrescu, The CloudSat science team, 2002: The CloudSat mission and the A-Train: A new dimension of space-based observations of clouds and precipitation. *Bull. A.M.S.*, **83**, 1771-1790.

Stensrud, D. J., 2007: Parameterization schemes: keys to understanding numerical weather prediction. Cambridge University Press.

Stull, R. B. 1988: An introduction to boundary layer meteorology. Kluwer Academic Publishers.

Thiebaux, H. J., P. R. Julian, G. J. DiMego, 1990: Areal versus collocation data quality control. Intl. Symp. on Assimilation of Observations in Meteorology and Oceanography, Clermont-Ferrand, France, WMO, 1990, pp. 255-260.

Tiedtke M., 1993: Representation of clouds in large-scale models. *Mon. Wea. Rev.*, **121**, 3040-3061.

Wallace, J. M., P. V. Hobbs, 1977: Atmospheric Science. Academic Press.

Warren, S. G., C. J. Hahn, J. London, R. M. Chervine, and R. L. Jenne, 1986: Global distribution of total cloud cover and cloud type amounts over ocean. NCAR/TN-317+STR, NCAR Tech. Note., New York.

Winker, D. M., M. A. Vaughan, A. Omar, Y. Hu, K. A. Powell, Z. Liu, W. H. Hunt, S. A. Young, 2009: Overview of the CALIPSO Mission and CALIOP Data Processing Algorithms. *J. Atmos. Ocean. Technol.*, **26**, 2310-2323.

Zupanski, D., 1997: A general weak constraint applicable to operational 4DVAR data assimilation systems. *Mon. Wea. Rev.*, **125**, 2274-2292.