Chapter- 1

Introduction

Data adaptation techniques are used to exploit information contained in observational data, previous forecasts and atmospheric dynamics for the purpose of weather forecasting. By statistically weighting this contributing information, data adaptation produces the best estimate of the current state of the atmosphere; this is used as the initial conditions for a model forecast. The weighted significance of each aspect in the adaptation is determined by its associated error. The confused nature of the atmosphere requires that the initial conditions can be exactly specified to take no notice of quick error growth and thus the correct specification of the weighting errors is fundamentally important.

ECMWF is a world leader in data adaptation research and development. The quality of our forecasts depends on how well we use information received in real-time from the global observing system, which consists of numerous satellite instruments, weather stations, ships, buoys, and other components.

The purpose of data adaptation is to determine a best possible atmospheric state using observations and short range forecasts. Data adaptation is typically a sequential time-stepping procedure, in which a previous model forecast is compared with newly received observations, the model state is then updated to reflect the observations, a new forecast is initiated, and so on. The update step in this process is usually referred to as the analysis; the short model forecast used to produce the analysis is called the background.

We also use data adaptation to monitor climate change based on past observations – this is called reanalysis.

The atmosphere is chaotic, meaning that even small differences in its state can lead to very different weather patterns occurring several days later – this is sometimes referred to as the butterfly effect. To account for the chaotic nature of the atmosphere and the
associated uncertainty in prediction, I run an ensemble of 51 forecasts simultaneously; the forecast using the best possible initial state plus 50 other forecasts with slight variations to the initial state. Our ensembles provide a probabilistic forecast which is an estimate of how predictable a particular weather situation is.

The atmosphere is a fluid. The idea of numerical weather prediction is to sample the state of the fluid at a given time and use the equations to estimate the state of the fluid at some time in the future. The process of entering observation data into the model to generate initial conditions called initialisation.

1.1 Inspiration for research

Numerical weather prediction uses mathematical models of the atmosphere and oceans to forecast the weather, based on present weather condition. Mathematical models based on the same physical principles can be used to generate either short-term weather forecasts or longer-term climate predictions; latter on it can be used for widely applied for understanding and projecting climate change. The improvements made to local models have allowed for significant improvements in tropical cyclone track and air quality forecasts; however, atmospheric models do weakly at handling processes that occur in a relatively rigid area, such as wildfires.

In Numerical weather prediction (NWP) is the governing equations used to describe the behaviour of the atmosphere contain approximately $10^7$ variables, and are sampled by order $10^6$ observations in a 6 hour synoptic period (Period of time during which the essential characteristics of a particular synoptic situation persist over a large area of the globe). The observations are provided by the Global Observing System (GOS)[1] and remotely sensed measurements, all with an associated error structure. We treat observation errors as independent with their type, i.e, radar data errors are independent of aeroplane data errors, but dependency is regularly exists between observations measured by the equal device. Satellite observations typically have horizontally and vertically
correlated errors. Origins of these errors include observation spatial proximity, dissimilar model and data resolutions, and data pre-processing. Surface based data are also affected by correlated errors but their usually lower density means the special effects of the correlated error are less significant. The size of the difficulty in NWP restricts the storage of the extra information provided by the error correlations. In operational weather prediction (OWP) centres around the world, the data adaptation is most frequently performed under the assumption of uncorrelated satellite observation errors.

The assumption of zero correlations is frequently used in combination with data thinning weather prediction centres around the world[2], the data adaptation is most frequently used methods. This one reduces the density of data by getting an averaging the properties of observations in a area, and assigning this mean as a single observation value. Under such assumptions, increasing the observation density beyond some entrance value has been shown to yield little or there is no improvement in analysis accuracy[3],[4]. Although abandonment accessible information may be appropriate when the spatial, resolution of the observations is denser than the model grid, recent technological advances have challenged the practicality of such methods. The Unified Model at the Met Office is run operationally at a 4 km horizontal resolution, but the increasing demand for now casting and convective scale modeling has encouraged the move towards a UK area model with resolution of order 1.5km. [5] Under such conditions there is a requirement to store in all the available data to provide details on the appropriate scales, and thus an alternative approach to dealing with observation error correlations is needed.

Approximating observation error correlation is a comparatively new path of research but progress has been made in future. In circulate matrices were used to approximate a Toeplitz inspection error covariance matrix. Results showed that incorrectly assuming uncorrelated observation errors gave baffling estimates of information content. In Fisher proposed assigning a block-diagonal structure to the observation error covariance matrix, with (uncorrelated) blocks corresponding to different devices or channels. Using this technique, individual block matrices were approximated by a truncated deigned composition. On a simple domain, there is no long-range correlations were pragmatic.
One of the greatest challenges in meteorology is quantifying and communicating the uncertainty in weather forecasts, especially for those forecasts based on numerical weather prediction (NWP) information. Weather decision aids, widely used in military operations, are often exploited to "translate" the NWP information into actionable weather intelligence. However, unrepresentative or erroneous environmental observations, errors in the data adaptation process, and inaccurate or imprecise microphysics or numerical analysis techniques can lead to uncertainty in the NWP output data. Extensive, on-going research on each segment of the NWP process is leading to the development of better analytical techniques, but the issue of quantifying the uncertainty remains a challenge. Ensemble techniques applied to NWP have proven to be valuable forecast tools but interpreting the output is often difficult, and they have demonstrated sensitivities to initial conditions. Bayes' Theorem and Bayesian Hierarchical Model (BHM) techniques have recently been applied to the NWP process with documented success at quantifying weather forecast uncertainty.

The first successful numerical prediction was performed using the ENIAC digital computer in 1950 by a team composed of American meteorologists Jule Charney Philip Thompson, Larry Gates, and Norwegian meteorologist Ragnar and applied mathematician John von newmann They used a simplified form of atmospheric dynamics based on solving the barotropic vorticity equation over a single layer of the atmosphere, by computing the geopotential height of the atmosphere's 500 millibars (15 inHg) pressure surface[5]. This simplification greatly reduced demands on computer time and memory, so the computations could be performed on the relatively primitive computers of the day[6]. When news of the first weather forecast by ENIAC was received by Richardson in 1950, he remarked that the results were an "enormous scientific advance[2]. The first calculations for a 24-hour forecast took ENIAC nearly 24 hours to produce[2], but Charney's group noted that most of that time was spent in "manual operations", and expressed hope that forecasts of the weather before it occurs would soon be realized
Forecast models based upon the equations for atmospheric dynamics do not perfectly determine weather conditions near the ground, statistical corrections were developed to attempt to resolve this problem. Statistical models were created based upon the three-dimensional fields produced by numerical weather models, surface observations, and the climatologically conditions for specific locations. These statistical models are collectively referred to as model output statistics (MOS), and were developed by the National Weather Service (NWS) for their suite of weather forecasting models by 1976. The United States Air Force developed its own set of MOS based upon their dynamical weather model by 1983.

The form of model data we are most familiar with on a day-to-day basis comes from Numerical Weather Prediction (NWP). NWP is focused on taking current observations of weather and giving out these data with computer models to forecast the future state of weather. Meaningful the current state of the weather is just as important as the numerical computer models processing the data. Current climate observations serve as input to the numerical computer models through a process known as data adaptation to produce outputs of temperature, precipitation, and hundreds of other meteorological elements from the oceans to the top of the atmosphere.

1.2 Research Objectives

In this Research I expand on the existing body of work on modeling observation correlation structure. I first compute observation error correlation structure for an operational satellite device. The statistical results I get or I obtain are new and helpful to the need to include satellite observation error correlations in data adaptation algorithms. By performing variational data adaptation experiments in a 3-D (dimensional) and 4-D framework, I then analyse the impact of new and existing approximations to error correlation construction. For this activity I wish to address the following questions:
• What is the real format of the observation error correlations?

In order to generate a good approximation, I must first have a precise estimate of the true error correlation structure. Satellite data are typical; and its carry correlated observation errors; the size and structure of the error covariances can be difficult to compute.

• What Type of approximations are available to model error correlation structure? What is their impact on data adaptation nature?

Data adaptation provides techniques for combining observations and prior Observation error correlation matrix structures impact on the accuracy of the observed field itself time storms, there is a requirement to preserve all the available data to provide detail linear approximation to the forward model. For an approximate error correlation structure to be implemented operationally it is compulsory computationally viable. I therefore express present approximations which will not be overly demanding on computational time and storage. These approximations are then ranked on their impact on different adaptation retrieval measures.

1.3 Thesis Summary

The research structured is as follows. In a Chapter 2 I try introduces the concepts of data adaptation and remote sensing. Although data adaptation is a relatively new science, fast development has been made. I focus on the two methods used in the thesis: three-dimensional variational data adaptation (3D-Var) and four-dimensional variational data adaptation (4D-Var). Observing System Experiments (OSEs) at the European Centre for Medium Range Weather Forecasting (ECMWF) and elsewhere have shown that the inclusion of satellite data in a 4D-Var algorithm results in the greatest positive forecast impact over all observation types. Here I assess the physics and operational behavior of satellite data, and accentuate its importance in current NWP. Details on the nature and origin of observation error covariances are then given. The chapter is concluded with a description of the techniques used to quantify these error
In Chapter 3 I address the second question posed in and present possible matrix representations of the observation error covariance structure. Three various approximating structures are described one by one: diagonal, circulant, and eigen composition approximations. I carefully pay attention to the feasibility of including these matrices operationally. The impact of each approximation can be evaluated using several parameters. In the second part of the chapter describe the following retrieval diagnostics: analysis error covariance matrix, information content, and matrix and vector norms. These properties can be used to decide how well each approximation performs in a data adaptation system.

Chapter 4 addresses the first question posed in Section 1.2 and contains new results on quantifying error correlation structure. We first set up the Infrared Atmospheric Sounding Interferometer (IASI) satellite instrument, and explain how its measurements are processed in the Met Office incremental 4D-Var adaptation scheme. Using a post-analysis diagnostic based on variational data adaptation theory and statistics from the Met Office system, we successfully calculate the cross-channel error correlations between IASI measurements. Diagnosed error covariances are given for the pre-processing 1D-Var adaptation and the core 4D-Var adaptation. Comparisons can be made with the current operational error variances.

More clean outcomes are presented in Chapter 5, where I will consider modeling correlation structure in a 3D-Var framework. Being a simple system than the four-D-Var framework, the results can be analyses more easily. Using information content measures, I quantify the accomplishment of each matrix approximation described (Circulant approximation, Toiplitz circulate approximation and eigen composition approximation) in Chapter 3 in modeling an empirically resulting on observation error correlation structure. The collision of each approximation can then be evaluated relative to the fact. Conclusions based on numerical evidence are drawn for different background error structures and constructions of the analysis error covariance matrix. The real results in this chapter address the second thesis question discuss in Section 1.2.

Inspired by the results in Chapter 5, I express the mathematical framework needed to make
bigger this investigation to a 4D-Var setting. I commence a set of one-dimensional shallow water equations (SWEs), used to represent simplified atmospheric dynamics, and describe the continuous analytical and discretised numerical models. I then develop a new incremental 4D-Var data adaptation system for the 1D SWEs which models observation error correlation structure using diagonal, Markov and Eigen composition matrix approximations. Lastly I have explained the coding tests used to test the validity of the model assumptions.

In last Chapter 6 I put the work done in entire research and draw conclusion from these experimental results concerning the effectiveness of modelling observation error correlations in data adaptation algorithms. I also make some suggestions for possible further work in this research.