DATA ADJUSTING USING VARIATIONAL METHODS

by

Qing Zhang

B.Sc., Shanghai University of Science and Technology, 1982
M.Eng., Shanghai University of Technology, 1987

A Thesis Submitted in Partial Fulfillment of the
Requirements for the Degree of

MASTER OF SCIENCE

in the Department of Computer Science

We accept this thesis as conforming
to the required standard

[redacted]
Dr. M. Danard, Supervisor (Department of Computer Science)

[redacted]
Dr. D. D. Olesky, Departmental Member (Department of Computer Science)

[redacted]
Dr. D. Hewgill, Outside Member (Department of Mathematics)

[redacted]
Dr. R. Illner, External Examiner (Department of Mathematics)

Supervisor: Dr. Maurice Danard

Abstract

The σ-coordinate system in the numerical modeling of the atmosphere has the advantage that the earth surface is always at the coordinate surface $\sigma = 1$. However, when over steeply sloping mountains, the horizontal pressure gradient P in the σ-coordinate system in the equation of motion consists of two terms of almost equal magnitude but of opposite sign. Unless special care is taken, the so called catastrophic cancellation will occur which results in a sudden loss of accuracy.

Two major factors contribute to the poor accuracy of P: the inconsistency of the original data and an inadequate finite difference scheme. Through numerical experiments, we found that the data inconsistency is the main source of error among the two. The following procedures are employed to solve the problem:

- Performing data adjustment to make it consistent with the hydrostatic condition, meanwhile, the adjustment is minimized in the least squares sense using variational methods;

- Using Corby's finite difference scheme to evaluate P, which has the effect of doubling the horizontal resolution, and therefore is helpful in improving the accuracy of P.

Thorough numerical experiments are conducted to investigate the effects of the above two procedures over grids of either mountainous or flat terrains. Results show that more than 20% reduction of errors is achieved on average for both grids, but in terms of absolute

value of error reduction, the procedures are more effective for mountainous terrain than for flat terrain. Significant improvement is also achieved for reducing the maximum value of errors.

Our numerical experiments also show that Corby's finite difference scheme, which is supposed to be superior to other schemes in the evaluation of P in the σ-coordinate system, is only effective when combined with the data adjustment.

Examiners:

[redacted]

Dr. M. Danard, Supervisor (Department of Computer Science)

[redacted]

Dr. D. D. Olesky, Departmental Member (Department of Computer Science)

[redacted]

Dr. D. Hewgill, Outside Member (Department of Mathematics)

[redacted]

Dr. R. Illner, External Examiner (Department of Mathematics)

Contents

Abstract ii

Contents v

List of Figures viii

List of Tables ix

Chapter **Page**

1 Introduction 1

1.1 σ Coordinate System 1

1.2 Problems with the σ-Coordinate System 2

1.3 Improving Accuracy of the Pressure Gradient 3

1.4 Our Work 5

2 Finite Difference Scheme 6

2.1 Hydrostatic Equation 6

2.2 Vertical Discretization 7

2.3 Corby's Finite Difference Scheme 9

2.4 An Explanation of Corby's Finite Difference Scheme 12

2.5 The Application in an N-Level Model 13

3 Variational Method **14**

3.1 The Need for Data Adjusting 14

3.2 Variational Method 16

3.3 Objective Analysis 18

3.4 The Application in Our Case 19

4 Data Adjusting **21**

4.1 The Constraint Conditions 21

4.2 Adjusting Only Temperature 22

4.3 Adjusting Both Temperature and Geopotential 23

4.4 Weights Assignment 25

4.5 Variable Rescale 26

4.6 Computational Results Analysis 28

4.7 Sensitivity Analysis 29

5 Numerical Experiments **31**

5.1 Geostrophic Wind 31

5.2 Evaluation of Geostrophic Wind 33

5.3 Design of the Numerical Experiments 35

5.4 Sample Data 37

5.5 Some Computational Details 38

5.6 Interpolation and Computation 40

6 Error Analysis **42**

6.1 Error Sources Analysis 42

6.2 Error Analysis for Experiment 1 43

6.3 Error Analysis for Experiment 2 44

6.4 Concluding Remarks 48

Bibliography **49**

Appendix

A List of Symbols **51**

B Analysis for Adjustment with only Temperature **53**

C Detailed Experimental Results **56**

List of Figures

Figure **Page**

2.1 The relationship among pressure, temperature and height 7

2.2 Vertical discretization along the σ coordinate 8

3.1 An example to show the problem of data inconsistency 15

5.1 An illustration how geostrophic wind is formed on a horizontal surface . . . 32

5.2 σ surfaces are set to coincide with p surfaces at point i 35

List of Tables

Table		Page
4.1	A typical result for adjusting only T	23
4.2	A typical result for adjusting both T and φ	29
4.3	Weight sensitivity analysis	30
6.1	Experiment 1: comparison of the results over *flat terrain*	45
6.2	Experiment 1: comparison of the results over *mountainous terrain*	45
6.3	Experiment 2: comparison of the results over *flat terrain*	47
6.4	Experiment 2: comparison of the results over *mountainous terrain*	47
C.1	Experiment 1: mean absolute errors over flat terrain	57
C.2	Experiment 1: R.M.S. of errors over flat terrain	58
C.3	Experiment 1: mean absolute errors over mountainous terrain	59
C.4	Experiment 1: R.M.S. of errors over mountainous terrain	60
C.5	Experiment 1: the largest absolute errors at each level	61
C.6	Experiment 2: mean absolute errors over flat terrain	62
C.7	Experiment 2: R.M.S. of errors over flat terrain	63
C.8	Experiment 2: mean absolute errors over mountainous terrain	64
C.9	Experiment 2: R.M.S. of errors over mountainous terrain	65

Chapter 1

Introduction

1.1 σ Coordinate System

Numerical weather prediction is mainly concerned with air movement in the atmosphere. Horizontal air movement, which is much more intense than vertical air movement, has been for a long time a major subject of numerical modeling of the atmosphere. In the study of horizontal air movement, the horizontal pressure gradient force is of great importance; its computational accuracy has a direct influence on the effectiveness of numerical weather prediction.

In the conventional coordinate systems, such as the Z-coordinate system which uses geometric height Z as the vertical coordinate, and the p-coordinate system which uses pressure p as the vertical coordinate, the horizontal pressure gradient P can be written as

$$P = -\frac{1}{\rho}\nabla_z p = -\nabla_p \varphi, \tag{1.1}$$

where ρ is the air density, $\varphi = gZ$ is the geopotential, g is the gravitational acceler-

ation, and ∇ is the horizontal gradient operator, subscripts z and p indicate that the differentiations are taken along constant height and pressure surfaces.

However, if we intend to take consideration of orographic effects in the numerical models, the conventional coordinate systems appear to have difficulties because the earth surface is no longer a coordinate surface. Therefore, complicated lower boundary conditions are often required when doing numerical integration, such as in the case of the p-coordinate system. To solve the problem, a so called σ-coordinate system [12] (also known as the transformed pressure coordinate system) is introduced which uses the independent variable σ as the vertical coordinate,

$$\sigma = \frac{p}{p_s}, \tag{1.2}$$

where p_s is surface pressure . In the σ-coordinate system, the earth surface is always at the coordinate surface $\sigma = 1$, and no special procedure is required to deal with the lower boundary condition. This is the chief advantage of employing the σ-coordinate system.

The horizontal pressure gradient P in the equation of motion can thus be written as

$$P = -\nabla_\sigma \varphi - RT \nabla_\sigma ln p_s, \tag{1.3}$$

where R is the gas constant for air and T is temperature. The subscript σ indicates that the differentiation is taken along a constant σ surface.

1.2 Problems with the σ-Coordinate System

A problem arises when we intend to use (1.3) to evaluate the pressure gradient. The two terms on the right hand side of (1.3) are almost of equal magnitude but opposite sign;

they are usually much larger than their sum. This proved to have significant effect on the accuracy of the result.

Smagorinsky *et al* [19] and Kurihara [10] reported that large computational errors occurred in the evaluation of the pressure gradient over steeply sloping terrain with (1.3). The reason is simple: as one travels up a sloping σ surface, the geopotential increases monotonically with height while the air pressure decreases with height. Therefore, the two terms $\nabla_\sigma \varphi$ and $RT\nabla_\sigma ln\, p_s$ have opposite sign. Furthermore, when over steeply sloping mountains, the values of both terms tend to be much larger than their sum, thus the so called catastrophic cancellation [22] occurs which signals a sudden loss of accuracy.

On the other hand, the evaluation of (1.3) requires data on σ surfaces which is usually obtained by interpolating from data on other coordinate surfaces. This turned out to be also a big problem. Sundqvist [20] pointed out that, in the evaluation of temperature on the transformed coordinate surfaces, if the magnitude of individual terms is ten times larger than a typical value of their sum, 1% (2-3 ^oC) error in temperature T may cause 10% error in the pressure gradient P. This situation is also applicable to geopotential error. Because of the inconsistency of the original data in temperature and geopotential, it is quite possible to produce such errors.

1.3 Improving Accuracy of the Pressure Gradient

One remedy to reduce the computational errors mentioned above is suggested by Smagorinsky *et al* [19] and Kurihara [10], that is, instead of using (1.3), the pressure gradient P is evaluated directly from the original expression (1.1) in the p-coordinate system, with the

pressure surface being interpolated to coincide with the σ surface at the places where the pressure gradient is wanted. However, it is difficult to estimate the magnitude of errors in this procedure because of the vertical interpolation [9, p53]; and extrapolation is often required when dealing with the highest and lowest levels of the model over steep terrain. In any case, abandoning the chosen coordinate system for the evaluation of a crucial term seems to be a defeatist policy [4].

Corby *et al* [4] proposed another finite difference scheme which is relatively simple and effective in reducing the computational errors in the σ-coordinate system. Corby *et al* showed that in a barotropic atmosphere[1] in which temperature varies linearly with the logarithm of pressure and the isobaric surfaces are horizontal , the following finite difference scheme

$$-(\frac{\partial \varphi}{\partial x})_\sigma - R\,T\,\frac{\partial ln\,p_s}{\partial x} \simeq -\delta_x \overline{\varphi}^x - R\overline{\overline{T}^x \delta_x ln\,p_s}^x \tag{1.4}$$

will result in the truncation errors in the two terms on the right hand side cancelling each other. In (1.4), we employ the difference operator

$$\delta_x() = [()_{x+\Delta x/2} - ()_{x-\Delta x/2}]/\Delta x \tag{1.5}$$

and the averaging operator

$$\overline{()}^x = [()_{x+\Delta x/2} + ()_{x-\Delta x/2}]/2 \tag{1.6}$$

[1] In a barotropic atmosphere, each parameter of state is a function of only one parameter of state, for example, $T = T(p)$. Therefore the equi-scalar surfaces $p = constant$ and $T = constant$ are parallel rather than intersecting each other.

where Δx is the grid step in the x direction. Note that

$$\delta_x \overline{()}^x = [()_{x+\Delta x} - ()_{x-\Delta x}]/2\Delta x \tag{1.7}$$

is simply the centered finite difference over two grid steps.

It will be shown later that we can derive the expression of (1.4) even if the isobaric surfaces are not horizontal.

Corby's finite difference scheme suggests that in the real atmosphere, the initial temperature should be assumed to vary linearly with the logarithm of pressure in order to improve the computational accuracy of (1.3).

1.4 Our Work

In this thesis, we attempt to improve the accuracy of the pressure gradient by using Corby's finite difference scheme. And based on the assumption of this scheme, we try to adjust the original data accordingly to make it consistent with the hydrostatic relation. At the same time, the adjustment is minimized in the least squares sense using variational methods.

Detailed mathematical analysis is presented to show why the procedure of data adjustment and Corby's finite difference scheme should be employed. Furthermore, thorough numerical experiments are conducted to convince us from every aspect that, by implementing data adjustment and applying Corby's finite difference scheme, we can significantly reduce the computational errors which otherwise we are subject to if the σ-coordinate system is employed.

Chapter 2

Finite Difference Scheme

2.1 Hydrostatic Equation

In numerical modeling of the atmosphere, people are most interested in the lower part of the atmosphere, usually from 0—20 kilometers above sea level, because more than 90% of the mass of the atmosphere lies within this range.

Among various meteorological parameters, three are of most interest: pressure p, temperature T, and height Z. In the real atmosphere, the relationship among these three parameters is very complicated. However, under the assumption of hydrostatic equilibrium[1], they are connected to each other by the hydrostatic equation

$$\frac{\partial Z}{\partial p} = -\frac{RT}{gp} \tag{2.1}$$

which is one of the most important equations in atmospheric dynamics.

[1] Hydrostatic equilibrium is a balance between pressure gradient and gravitational force along the vertical. In practice, this assumption is sufficiently accurate for many purpose.

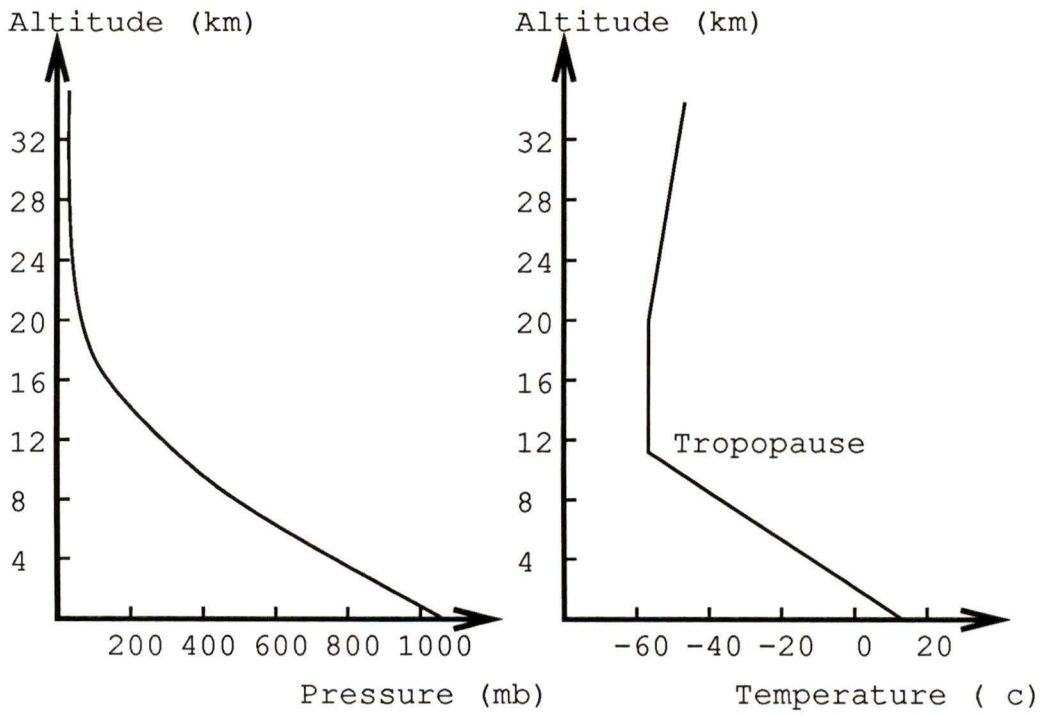


Figure 2.1: A typical relationship among pressure, temperature and height

In Figure 2.1, we can see the relationship between p, T and Z. The model we are going to use covers a range of about 0—16 kilometers above the sea level.

2.2 Vertical Discretization

Since the atmosphere we are dealing with is a three-dimensional space, we have to discretize it in some way horizontally as well as vertically. Usually the horizontal space is represented by two orthogonal curvilinear coordinates x and y; while the vertical height may be represented by Z, p, or some transformed coordinates like σ. Discretization is

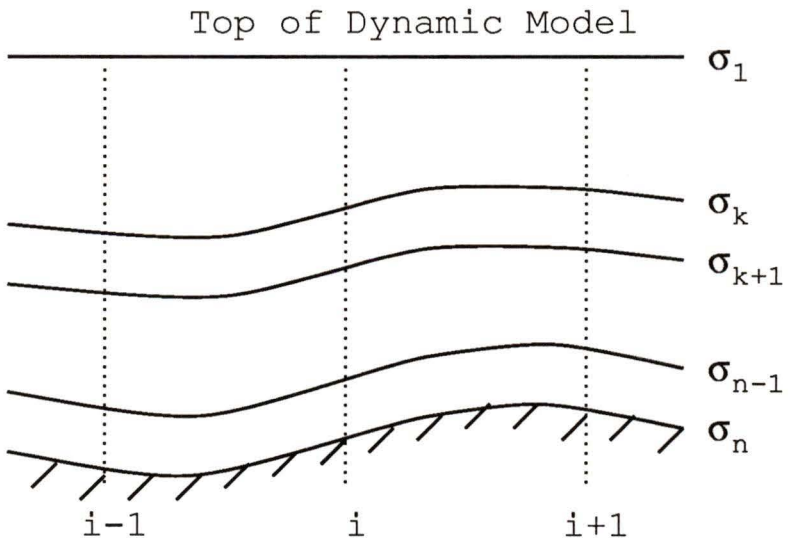


Figure 2.2: Vertical discretization along the σ coordinate

made along all coordinates such that each grid point represents a certain fixed place in the atmosphere.

In the case of the σ-coordinate system, the vertical discretization has a structure like that in Figure 2.2. It can be seen that σ surfaces are influenced heavily by the shape of the earth surface. On the contrary, isobaric surfaces in the p-coordinate system are usually irrelevant to the shape of the earth surface.

The original data is usually obtained from some meteorological organizations, such as the Canadian Meteorological Center (CMC). The CMC distributes meteorological data at a set of regularly spaced grid points with vertical grid points on some isobaric surfaces; the data is computed from a hemisphere spectral model of the atmosphere that is constantly

updated with observations from a variety of sources.

2.3 Corby's Finite Difference Scheme

Corby's finite difference scheme is constructed on the following assumption that the atmosphere is laterally uniform with a vertical temperature given everywhere by

$$T(p) = A\, ln\, p + B \tag{2.2}$$

where A and B are constants. This assumption is acceptable at least in thin layers of the atmosphere.

Since in such an atmosphere structure the isobaric surfaces are horizontal, the pressure gradient, which is the negative of the gradient of φ along isobaric surfaces as in (1.1), is always zero. Then we can try to find approximations to the two components in the right hand side of (1.3) such that they combine to give a zero pressure gradient. This is the initial motivation of Corby's finite difference scheme.

In fact, the condition that the isobaric surfaces are horizontal is not necessary, and we can still get Corby's scheme, as is proved as follows by Danard (CSC 584 notes).

We can rewrite the hydrostatic equation (2.1) in the following way:

$$\frac{\partial \varphi}{\partial ln\, p} = -RT. \tag{2.3}$$

Substituting (2.2) into (2.3), and integrating it from the earth surface φ_s to an isobaric surface $\varphi_0 = \varphi(p_0)$, we get

$$\varphi_0 = \varphi_s + \frac{R\,A}{2}[(ln\, p_s)^2 - (ln\, p_0)^2] + R\,B\,(ln\, p_s - ln\, p_0). \tag{2.4}$$

Then performing the difference operation $\delta_x()$ defined by (1.5) on (2.4), we get

$$\delta_x \varphi_0 = \delta_x \varphi_s + \frac{R A}{2} \delta_x (ln\, p_s)^2 + R B \delta_x ln\, p_s. \tag{2.5}$$

All operations here are at constant $\sigma = 1$ except in the term on the left which is at constant $p = p_0$.

For any functions F and G, the following identities hold:

$$\delta_x(FG) = \overline{F}^x \delta_x G + \overline{G}^x \delta_x F \tag{2.6}$$

and

$$\delta_x F^2 = 2 \overline{F}^x \delta_x F. \tag{2.7}$$

Using (2.7) in (2.5) gives

$$\begin{aligned} \delta_x \varphi_0 &= \delta_x \varphi_s + R A \overline{ln\, p_s}^x \delta_x ln\, p_s + R B \delta_x ln\, p_s \\ &= \delta_x \varphi_s + R (A \overline{ln\, p_s}^x + B) \delta_x ln\, p_s \\ &= \delta_x \varphi_s + R \overline{T_s}^x \delta_x ln\, p_s \end{aligned} \tag{2.8}$$

using (2.2). Now we apply the averaging operation $\overline{()}^x$ defined by (1.6) on (2.8) and obtain

$$\delta_x \overline{\varphi_0}^x = \delta_x \overline{\varphi_s}^x + R \overline{\overline{T_s}^x \delta_x ln\, p_s}^x. \tag{2.9}$$

The left hand side is the negative of the pressure gradient at $p = p_0$. In the case that isobaric surface is horizontal, $\delta_x \overline{\varphi_0}^x = 0$, the right hand side should be zero too. It can be shown that, if φ_s has the same profile as T_s and $ln\, p_s$, the two terms on the right hand side of (2.9) do combine to give a value of zero, that means, the truncation errors of the two terms cancel each other exactly. Therefore, if we intend to approximate $\frac{\partial \varphi_s}{\partial x}$ by $\delta_x \overline{\varphi_s}^x$,

we are obliged to approximate $RT\frac{\partial ln\, p_s}{\partial x}$ by $R\overline{\overline{T_s}^x\,\delta_x ln\, p_s}^x$. In this way, we can achieve Corby's original goal.

Similar to (2.4), we can also define geopotential φ_k on the k^{th} σ-surface

$$\varphi_k = \varphi_s + \frac{RA}{2}[(ln\, p_s)^2 - (ln\, \sigma_k p_s)^2] + RB\,(ln\, p_s - ln\, \sigma_k p_s). \tag{2.10}$$

We apply the difference and averaging operations again on (2.10), but this time on constant σ-surface, and obtain

$$\delta_x\,\overline{\varphi_k}^x = \delta_x\,\overline{\varphi_s}^x - RA\,ln\,\sigma_k\,\delta_x\,\overline{ln\, p_s}^x. \tag{2.11}$$

Combining (2.11) with (2.9) to eliminate $\delta_x\,\overline{\varphi_s}^x$, we obtain

$$\begin{aligned}
\delta_x\,\overline{\varphi_0}^x &= \delta_x\,\overline{\varphi_k}^x + RA\,ln\,\sigma_k\,\delta_x\,\overline{ln\, p_s}^x + R\overline{\overline{(A\,ln\, p_s + B)}^x\,\delta_x\, ln\, p_s}^x \\
&= \delta_x\,\overline{\varphi_k}^x + R\overline{\overline{(A\,ln\,\sigma_k\, p_s + B)}^x\,\delta_x\, ln\, p_s}^x \\
&= \delta_x\,\overline{\varphi_k}^x + R\overline{\overline{T_k}^x\,\delta_x\, ln\, p_s}^x
\end{aligned} \tag{2.12}$$

which is just (1.4). (2.12) gives us a finite difference scheme to compute the pressure gradient in the σ-coordinate system which enables us to avoid large computational error in the results. In the next section, we will give an more direct explanation for Corby's finite difference scheme. In the above deduction, the only assumption we have used is (2.2). That means (2.12) holds even if the isobaric surfaces are not horizontal, as long as the temperature and pressure satisfy (2.2).

In general, the real atmosphere may be extremely complicated. So when applying Corby's finite difference scheme, it is difficult to estimate how large the computational error is. However, it was reported by Corby, *et al* , that the use of (2.12) does avoid the

very large computational error which arises with other finite difference schemes, such as $R\,T\,\delta_x \overline{ln\,p_s}^x$.

2.4 An Explanation of Corby's Finite Difference Scheme

The above mathematical deduction, though quite straight forward, may not be very helpful to understand why Corby's finite difference scheme is effective. Here I would like to explain it in a more explicit way.

Expanding (2.12) by (1.5) and (1.6) at grid point i, we can see that

$$\begin{aligned}
\delta_x \overline{\varphi_{o,i}}^x &= \delta_x \overline{\varphi_{k,i}}^x + R\,\overline{\overline{T_{k,i}}^x\,\delta_x\,ln\,p_{s,i}}^x \\
&= \frac{\varphi_{k,i+1}-\varphi_{k,i-1}}{2\,\Delta x} + \frac{R}{2}[\frac{T_{k,i+1}+T_{k,i}}{2}\,\frac{ln\,p_{s,i+1}-ln\,p_{s,i}}{\Delta x} \\
&\quad + \frac{T_{k,i}+T_{k,i-1}}{2}\,\frac{ln\,p_{s,i}-ln\,p_{s,i-1}}{\Delta x}] \\
&= \frac{1}{2}[(\frac{\varphi_{k,i+1}-\varphi_{k,i}}{\Delta x} + R\,\frac{T_{k,i+1}+T_{k,i}}{2}\,\frac{ln\,p_{s,i+1}-ln\,p_{s,i}}{\Delta x}) \\
&\quad + (\frac{\varphi_{k,i}-\varphi_{k,i-1}}{\Delta x} + R\,\frac{T_{k,i}+T_{k,i-1}}{2}\,\frac{ln\,p_{s,i}-ln\,p_{s,i-1}}{\Delta x})] \\
&= \frac{1}{2}[(\delta_x\varphi_{k,i+1/2} + R\,\overline{T_{k,i+1/2}}^x\,\delta_x\,ln\,p_{s,i+1/2}) \\
&\quad + (\delta_x\varphi_{k,i-1/2} + R\,\overline{T_{k,i-1/2}}^x\,\delta_x\,ln\,p_{s,i-1/2})] \qquad (2.13)
\end{aligned}$$

where k denotes the k^{th} σ-surface.

It can be seen from (2.13) that the two terms within the square brackets are simply the finite difference expressions of the pressure gradients at points $i+1/2$ and $i-1/2$ respectively, with a half length grid step. $\delta_x \overline{\varphi_{o,i}}^x$ is then expressed as the average of the two.

Sundqvist [20] found that the σ-coordinate system requires a higher resolution (smaller

grid step) than the p-coordinate system to achieve the same computational accuracy. Corby's finite difference scheme actually has the effect of doubling the horizontal resolution, and therefore helps to reduce the computational error of (1.3).

2.5 The Application in an N-Level Model

The assumption that temperature varies linearly with the logarithm of pressure is acceptable when applied to the lower part of the troposphere. However, as the height goes above the tropopause, there is a sudden change in the way the temperature varies with height (as well as pressure), as can be seen from Figure 2.1. Therefore, it becomes unreasonable to assume that (2.2) hold throughout the whole range when the height of the range is above the tropopause.

An alternative way is to divide the whole range into some smaller segments, and assume that (2.2) holds within each segment. Of course, we could no longer expect that the truncation errors of the two terms on the right hand side of (2.12) will cancel each other exactly. However, as is supported by our numerical experiments, Corby's finite difference scheme does help to reduce the computational errors which we are subject to with other difference schemes.

In an n-level model of the atmosphere, a segment is naturally defined as the layer between two adjacent levels.

Chapter 3

Variational Method

3.1 The Need for Data Adjusting

The computation of the pressure gradient in the σ-coordinate system requires data on σ-surfaces which is usually interpolated from data on pressure surfaces, as is available from CMC. However, due to various error sources, the original data is usually not consistent with the hydrostatic equation. This inconsistency is also responsible for the large computational error of (1.3).

When analyzing the computational errors in the σ-coordinate system, Sundqvist [20] pointed out that, in the calculation of the pressure gradient forces using (1.3), an error will appear as a horizontal error, while in fact it originates from vertical truncation yielding inconsistencies between T and φ in conjunction with transfer of data from p to σ coordinates.

A simple example may help to understand this problem. Given data on the k^{th} and

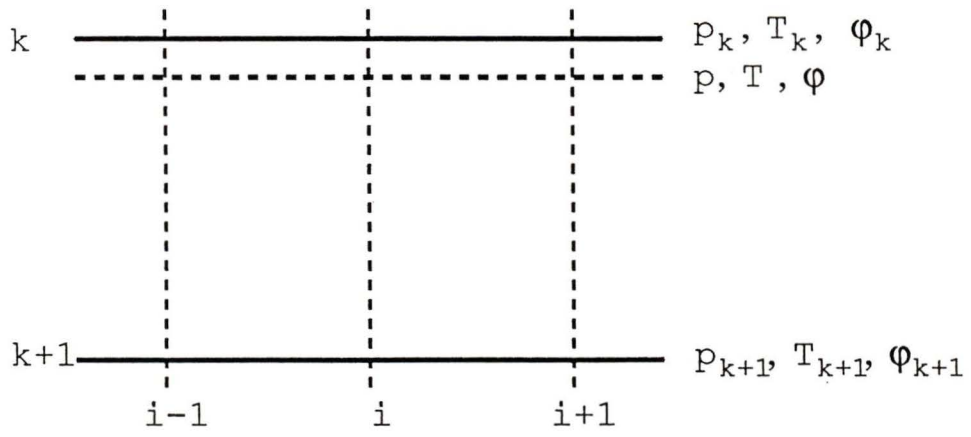


Figure 3.1: An example to show the problem of data inconsistency

$k+1^{st}$ level at the i^{th} grid point, we want to interpolate from it to the pressure surface p to obtain T and φ, as shown in Figure 3.1.

We assume that the data on the k^{th} level and $k+1^{st}$ levels do not satisfy the hydrostatic relation by an amount ϵ,

$$\epsilon = \varphi_k - \varphi_{k+1} - \frac{R(T_k + T_{k+1})}{2} ln(p_{k+1}/p_k). \tag{3.1}$$

We adopt the interpolation method of Kurihara [10] of using the hydrostatic relation to get the geopotential difference $\Delta\varphi$ between two pressure levels p_a and p_b, with the mean temperature being $\overline{T}$,

$$\Delta\varphi = R\overline{T}\, ln(p_a/p_b). \tag{3.2}$$

In our case, we can get φ by

$$\varphi = \varphi_{k+1} + \frac{R(T_k + T_{k+1})}{2} ln(p_{k+1}/p). \tag{3.3}$$

When p is sufficiently close to p_k, we will find that φ differs from φ_k by an amount ϵ. Obviously, this difference comes from the inconsistency of the original data, and is eventually imposed on the interpolated data. Considering the sensitivity of (1.3), this error may be significant.

One way to reduce this error is to adjust the data in advance so as to make it consistent with the hydrostatic condition, with the adjustment being minimized in the least squares sense. Though the adjustment itself brings some errors, it in fact has the effect of distributing the error evenly over the entire range so as to avoid large inconsistency being imposed on data at a single level when interpolating.

3.2 Variational Method

Sasaki [15, 16] has developed an initialization method based on the calculus of variations with which differences between the observed values of the meteorological elements and the corresponding objectively modified values are minimized in the least-squares sense subject to some dynamical constraints.

It is desirable to review some aspects of the calculus of variations first. Suppose $F(x, y, y')$ is a twice differentiable function. We wish to determine a function $y = f(x)$ such that the following integral is a minimum (or maximum):

$$I = \int_a^b F(x, y, y')dx. \tag{3.4}$$

Here the end points a and b are considered fixed with $f(a) = A$ and $f(b) = B$. Now suppose there is a small change from $y = f(x)$ to $y + \delta y = f(x) + \varepsilon g(x)$, where ε is a

parameter and $g(x)$ is arbitrary except that $g(a) = g(b) = 0$.

Then the integral (3.4) will be modified to

$$I + \delta I = \int_a^b (F + \delta F) dx \tag{3.5}$$

where δ is the variational operator, and

$$\delta F = F_y \delta y + F_{y'} \delta y',$$

the subscripts denote partial derivatives.

Thus a necessary condition for a minimum is

$$F_y - \frac{d}{dx} F_{y'} = 0 \tag{3.6}$$

which is referred to as the Euler-Lagrange equation.

Suppose F contains two functions, $y(x)$ and $z(x)$, that is, $F = F(x, y, z, y', z')$, and furthermore, the integral is subject to a constraint, say, $G(x, y, z) = 0$. In this case, a Lagrange multiplier μ may be introduced, and (3.4) is now changed to

$$I = \int_a^b [F(x, y, z, y', z') + \mu\, G(x, y, z)]\, dx. \tag{3.7}$$

The resulting Euler-Lagrange equations are

$$F_y - \frac{d}{dx}(F_{y'}) + \mu G_y = 0 \tag{3.8}$$

$$F_z - \frac{d}{dx}(F_{z'}) + \mu G_z = 0, \tag{3.9}$$

together with the constraint condition

$$G(x, y, z) = 0. \tag{3.10}$$

3.3 Objective Analysis

Sasaki [15] introduced the variational method into numerical weather prediction to determine the initial conditions of the primitive equations; this technique is known as objective analysis. It can be stated as follows.

Let f_{oi} represent observed values of meteorological parameters at a point in space $S(x_1, x_2, x_3, t)$ and f_{ai} the corresponding adjusted values, where the subscript i refers to different meteorological parameters. The principle is expressed by

$$\delta \int_S \sum_i \alpha_i^2 (f_{ai} - f_{oi})^2 dS = 0, \qquad i = 1, 2, \cdots, \tag{3.11}$$

where α_i^2 is a weighting factor reflecting the degree of accuracy of the meteorological parameter f_{oi} (α_i corresponds to Gauss precision modulus).

Since numerical modeling using finite difference or finite element techniques is done at a set of discrete points within S, the equation (3.11) can be written as

$$\delta \sum_S \sum_i \alpha_i^2 (f_{ai} - f_{oi})^2 = 0, \qquad i = 1, 2, \cdots. \tag{3.12}$$

The equation (3.12) can be solved by imposing some constraints

$$G_i(f_{a1}, f_{a2}, \cdots, f_{aj}, \cdots) = 0, \qquad i = 1, 2, \cdots, \tag{3.13}$$

where G_i represents the i^{th} functional relation of f_{aj}'s.

One way of imposing constraints can be written as

$$\delta \sum_S \sum_i [\alpha_i^2 (f_{ai} - f_{oi})^2 + \mu_i G_i] = 0, \qquad i = 1, 2, \cdots, \tag{3.14}$$

which are called "strong constraints" in the sense that the constraint conditions will be satisfied exactly by the solution, as opposed to the "weak constraints" which can be written

as

$$\delta \sum_S \sum_i [\alpha_i^2 (f_{ai} - f_{oi})^2 + \pi_i G_i^2] = 0, \qquad i = 1, 2, \cdots. \tag{3.15}$$

It is called "weak constraints" because the constraint conditions may not be exactly satisfied by the solution.

It should be noted that G is linear in (3.14) and quadratic in (3.15), the coefficient μ_i is the Lagrange multiplier and subject to the variational operator, while π_i is weight and not subject to the variational operator.

Since our purpose is to adjust the data to satisfy the hydrostatic condition exactly, obviously, the "strong constraints" is the suitable one we need.

3.4 The Application in Our Case

The meteorological parameters we are going to deal with are mainly pressure, temperature and geopotential, where pressure parameter represents a set of fixed surfaces on which temperature and geopotential are given. What we want to achieve by our data adjusting is to make the above three parameters satisfy the hydrostatic condition vertically between two adjacent pressure surfaces, so as to avoid large interpolation error arising due to the inconsistency of the observed data.

Since the only constraint imposed on the parameters is the hydrostatic condition along the vertical coordinate, therefore, the adjustment at one grid point is independent of that at other grid points. However, when computing the pressure gradient at one grid point, we need data at other grid points nearby as well. It is interesting to observe how independently

adjusted data affects the accuracy of the pressure gradient evaluated from data at several adjacent grid points.

Chapter 4

Data Adjusting

4.1 The Constraint Conditions

By adopting Corby's suggestion, we assume that temperature varies linearly with the logarithm of pressure within two adjacent pressure surfaces. Then for p_{k-1} and p_k, $k = 2, 3, \cdots, n$, we have

$$T(p) = A_k \, ln \, p + B_k, \qquad p_{k-1} \leq p \leq p_k \tag{4.1}$$

where A_k and B_k are coefficients.

The procedure to derive the constraint conditions is similar to (2.4): we substitute (4.1) into (2.3), integrate it from φ_k to φ_{k-1}, and get

$$\varphi_{k-1} = \varphi_k + \frac{R\,A_k}{2}[(ln\, p_k)^2 - (ln\, p_{k-1})^2] + R\,B_k\,(ln\, p_k - ln\, p_{k-1}). \tag{4.2}$$

Substituting (4.1) back into (4.2), we obtain

$$T_k + T_{k-1} = \frac{2(\varphi_{k-1} - \varphi_k)}{R\, ln\,(p_k/p_{k-1})}. \tag{4.3}$$

Note that equation (4.3) is in fact the finite difference form of (2.3) over two adjacent points along the vertical. If we want (4.3) to hold for any two points within the given range, we are forced to impose (4.1) on temperature and pressure.

4.2 Adjusting Only Temperature

In order that (4.3) is satisfied, there may be several different approaches of data adjusting. We can adjust only temperature, or adjust only geopotential, or we can adjust both of them. However, in case only temperature is adjusted, the result is not satisfactory.

In a 10-level numerical model ($n = 10$), the variational equation for adjusting only temperature is

$$I = \sum_{k=1}^{10} \frac{1}{2}[T_{a,k} - T_{o,k}]^2 + \sum_{k=2}^{10} \mu_k[T_{a,k} + T_{a,k-1} - f_k] \tag{4.4}$$

where $T_{a,k}$ and $T_{o,k}$ are adjusted and observed temperatures respectively, μ_k is the Lagrange multiplier, $f_k = 2(\varphi_{k-1} - \varphi_k)/Rln(p_k/p_{k-1})$, $k = 2, 3, \cdots, 10$. By differentiating I with respect to $T_{a,k}$ and μ_k, and letting the result be zero, we can obtain 19 equations for 19 unknowns, and then solve them for $T_{a,k}, k = 1, 2, \cdots, 10$. Table 4.1 shows the result at a typical grid point.

From Table 4.1 , we can check that the hydrostatic condition is satisfied by the adjusted temperatures. However, from the last column, we can see that there exists an obvious oscillation in the differences of adjusted and observed temperatures, and the absolute value of the difference is very large. It is of course not desirable to introduce a source of large error in the stage of data adjusting. We have shown in *Appendix B* that this error

Table 4.1: A typical result for adjusting only T

k	p_k	φ_k	$T_{o,k}$	$T_{a,k}$	$T_{a,k} - T_{o,k}$
1	100.	154683.2	-55.9	-52.7	3.2,
2	150.	129428.6	-55.5	-59.7	-4.2,
3	200.	111553.4	-57.3	-53.8	3.5,
4	250.	97794.2	-57.4	-63.0	-5.6,
5	300.	86553.6	-53.6	-53.8	-0.2,
6	400.	67757.2	-40.7	-37.3	3.4,
7	500.	52449.6	-29.3	-31.1	-1.8,
8	700.	28047.6	-13.2	-10.0	3.2,
9	850.	13406.4	-7.3	-11.0	-3.7,
10	1000.	715.4	4.5	8.7	4.2,

is the direct result of the structure of the variational equation (4.4), and can't be avoided when only temperature is adjusted. So we have to seek another way for implementing data adjustment.

4.3 Adjusting Both Temperature and Geopotential

When we adjust one parameter and keep the other unchanged, we are implying that the other one is accurate. However, it is not the case for temperature and geopotential. Both temperature and geopotential are subject to a variety of error sources in the process of measurements and data manipulations, so it is more reasonable for us to adjust both of them.

The variational equation for adjusting both temperature and geopotential subject to

the hydrostatic constraint (4.3) for the 10-level numerical model is as follows:

$$\begin{aligned} I &= \sum_{k=1}^{10} \frac{\alpha_k^2}{2}[T_{a,k} - T_{o,k}]^2 + \sum_{k=1}^{10} \frac{\beta_k^2}{2}[\varphi_{a,k} - \varphi_{o,k}]^2 \\ &+ \sum_{k=2}^{10} \mu_k [T_{a,k} + T_{a,k-1} - \frac{2(\varphi_{a,k-1} - \varphi_{a,k})}{R\, ln\,(p_k/p_{k-1})}] \end{aligned} \tag{4.5}$$

where α_k^2 and β_k^2 are weights, and μ_k is the Lagrange multiplier.

Barker [3] obtained a very similar expression to (4.5). However, there are some major differences between the two:

- The initiative of Barker's data adjustment was to reduce the error of temperature when interpolating from pressure surfaces to other coordinate systems. Since merely adjusting temperatures without geopotentials resulted in large errors, similar to what we had shown in the previous section, he had to resort to adjusting both parameters.
- Barker admitted that it is ideal to use the hydrostatic constraint in the form of (4.3), but he deliberately chose to use another much more complicated form of the hydrostatic condition, which makes the problem more difficult to solve. We did not find any side-effects of using (4.3) in our case.
- The weight assignment is different. In Barker's expression, the weights for the data at different levels are the same, and the value of the weights were chosen merely by making experiments on data from an idealized model. However, we assign different weights for data at different levels, and the values of the weights are decided from the standard deviations of the data.
- The experimental methods are different. Barker used the data from an idealized model, and added some disturbance on it. Then he tried to use the data adjustment to recover from the disturbance. On the other hand, we use the data from the real world, and regard it as accurate. Therefore we want to be truthful to the observed data as much as possible, instead of recovering it. We evaluate the effectiveness of the data adjustment by looking at the errors in the pressure gradients computed from the adjusted data and compared with the best results available to us.

In the following part of this chapter, we show how to solve (4.5) to achieve our goal. Differentiating I with respect to $T_{a,k}$, $\varphi_{a,k}$ and μ_k, and letting the result be zero, we get

$$\delta I = \sum_{k=1}^{10} \alpha_k^2 [T_{a,k} - T_{o,k}]\delta T_{a,k} + \sum_{k=1}^{10} \beta_k^2 [\varphi_{a,k} - \varphi_{o,k}]\delta \varphi_{a,k}$$

$$+ \sum_{k=2}^{10} \mu_k [\delta T_{a,k} + \delta T_{a,k-1} - \frac{2(\delta\varphi_{a,k-1} - \delta\varphi_{a,k})}{R\, ln\,(p_k/p_{k-1})}]$$

$$+ \sum_{k=2}^{10} \delta\mu_k [T_{a,k} + T_{a,k-1} - \frac{2(\varphi_{a,k-1} - \varphi_{a,k})}{R\, ln\,(p_k/p_{k-1})}] = 0. \qquad (4.6)$$

From (4.6), we can derive a linear system of equations with 29 unknowns and 29 equations; a unique solution exists.

To solve (4.6), some additional considerations are needed, such as how to choose the weights α_k and β_k, how to avoid large round-off errors to occur, and so on.

4.4 Weights Assignment

It is reasonable to directly relate the weights α_k and β_k with the standard deviations of temperature T_k and geopotential height φ_k such that the larger the error , the smaller the weight.

When computing geopotential φ, the following formula is used:

$$\varphi = \varphi_s + R \int_p^{p_s} T\, d\, ln\, p, \qquad (4.7)$$

where subscript s denotes surface level. Suppose φ_s and T have errors $\varepsilon(\varphi_s)$ and $\varepsilon(T)$ respectively, with the mean square deviations being $\sigma^2_{\varphi_s}$ and σ^2_T, $\varepsilon(\varphi_s)$ is independent of $\varepsilon(T)$, and $\varepsilon(T)$ is independent of p. Then, we have

$$\varepsilon(\varphi) = \varepsilon(\varphi_s) + R \int_p^{p_s} \varepsilon(T)\, d\, ln\, p = \varepsilon(\varphi_s) + \varepsilon(T) R\, ln(p_s/p) \qquad (4.8)$$

and

$$\sigma^2_\varphi = \sigma^2_{\varphi_s} + [R\, ln(p_s/p)]^2 \sigma^2_T \qquad (4.9)$$

We assume that σ_T^2, which mainly comes from the measuring system errors, is the same at different heights. It can be seen from (4.9) that, given $\sigma_{\varphi_s}^2$ and σ_T^2, we can get σ_φ^2 at the place where pressure equals p. Note that σ_φ^2 becomes larger with the increase of the height because the factor $[Rln(p_s/p)]^2$ is growing with height. This implies that geopotentials at higher levels are less accurate than that at lower levels, therefore, smaller weights should be assigned to them. Now we can use σ_T^2 and $\sigma_{\varphi_k}^2$ to define α_k^2 and β_k^2:

$$\alpha^2 = \alpha_k^2 = 1/\sigma_T^2 \tag{4.10}$$

$$\beta_k^2 = 1/\sigma_{\varphi_k}^2 \qquad k = 1, 2, ..., 10. \tag{4.11}$$

In this way, we can determine the weights for temperatures and geopotentials at each level.

4.5 Variable Rescale

In our data adjustment, the data is ranging from p-surface 1000 mb to 100 mb, which is equivalent to a range from somewhere near the earth's surface to a height of about 16 kilometers above the sea level. That means the value of geopotential (the product of the gravitational force and height) will range from the order of 10^2 to 10^5 (m^2/s^2). On the other hand, the weights for geopotentials at top levels become very small. For typical values $\sigma_{\varphi_s} = 10.0$ and $\sigma_T = 0.5$, the value of β_k^2 will vary from $(0.1)^2$ (at $k = 10$) to $(0.00302)^2$ (at $k = 1$). Therefore, the solution of the linear system of equations derived from (4.6) may be subject to severe roundoff errors due to the extremely large and small values of its parameters, which will result in poor computational accuracy.

Fortunately, we can solve the problem by rescaling the variables. From the fact that $\frac{1}{2}\beta_k^2[\varphi_{a,k}-\varphi_{o,k}]^2 = \frac{1}{2}[\beta_k\varphi_{a,k}-\beta_k\varphi_{o,k}]^2$, we define

$$\phi_{o,k} = \beta_k\varphi_{o,k} = \varphi_{o,k}/\sigma_{\varphi_k} \tag{4.12}$$

$$\phi_{a,k} = \beta_k\varphi_{a,k} = \varphi_{a,k}/\sigma_{\varphi_k}. \tag{4.13}$$

It can be shown that now the variable ϕ_k is properly ranged (from the order of 10^1 to 10^2) with the weight being 1.

Thus, by rescaling, (4.5) becomes

$$\begin{aligned} I = & \sum_{k=1}^{10}\frac{\alpha^2}{2}[T_{a,k}-T_{a,k}]^2 + \sum_{k=1}^{10}\frac{1}{2}[\phi_{a,k}-\phi_{o,k}]^2 \\ & + \sum_{k=2}^{10}\mu_k[T_{a,k}+T_{a,k-1} - \frac{2}{Rln\,(p_k/p_{k-1})}(\frac{\phi_{a,k-1}}{\beta_{k-1}} - \frac{\phi_{a,k}}{\beta_k})] \end{aligned} \tag{4.14}$$

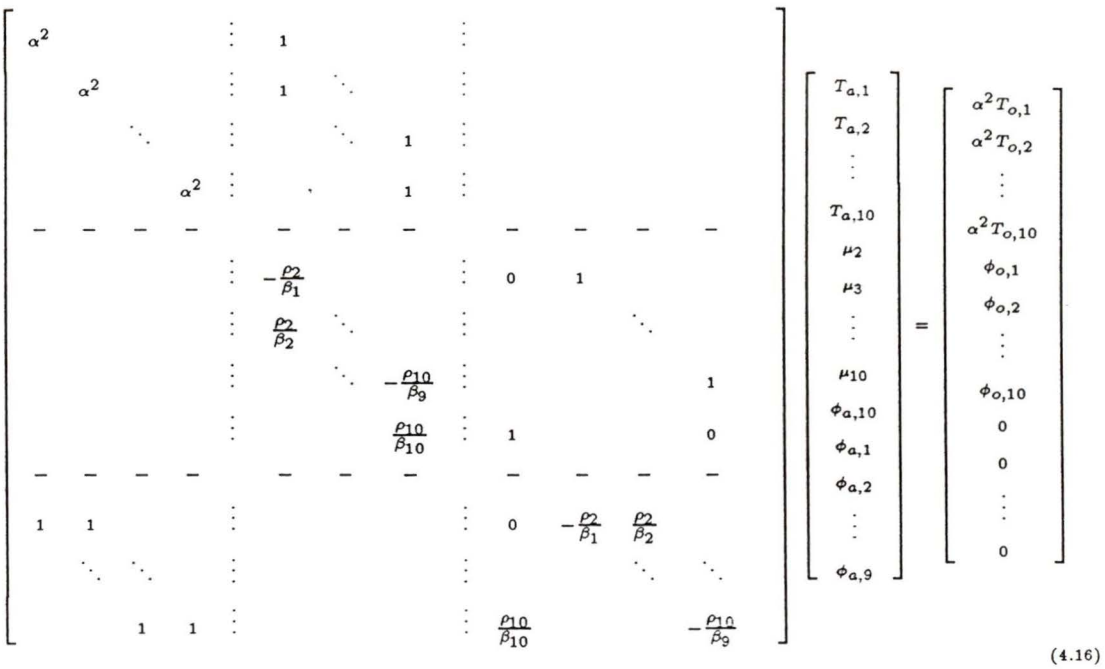
We can get the linear system of equations by differentiating I and letting it be zero:

$$\begin{array}{llll} \alpha^2T_{a,1} & +\mu_2 & & = \alpha^2T_{o,1} \\ \alpha^2T_{a,2} & +\mu_3 & +\mu_2 & = \alpha^2T_{o,2} \\ \vdots & \vdots & \vdots & \vdots \\ \alpha^2T_{a,10} & & +\mu_{10} & = \alpha^2T_{o,10} \\ \phi_{a,1} & -(\rho_2/\beta_1)\mu_2 & & = \phi_{o,1} \\ \phi_{a,2} & -(\rho_3/\beta_2)\mu_3 & +(\rho_2/\beta_2)\mu_2 & = \phi_{o,2} \\ \vdots & \vdots & \vdots & \vdots \\ \phi_{a,10} & & +(\rho_{10}/\beta_{10})\mu_{10} & = \phi_{o,10} \\ T_{a,1}+T_{a,2} & -(\rho_2/\beta_1)\varphi_{a,1} & +(\rho_2/\beta_2)\varphi_{a,2} & = 0 \\ T_{a,2}+T_{a,3} & -(\rho_3/\beta_2)\varphi_{a,2} & +(\rho_3/\beta_3)\varphi_{a,3} & = 0 \\ \vdots & \vdots & \vdots & \vdots \\ T_{a,9}+T_{a,10} & -(\rho_{10}/\beta_9)\varphi_{a,9} & +(\rho_{10}/\beta_{10})\varphi_{a,10} & = 0 \end{array} \tag{4.15}$$

where $\rho_k = 2/[Rln(p_k/p_{k-1})]$.

With proper arrangement, (4.15) can be expressed in the matrix form which could be

easily solved by any standard numerical method, such as LU decomposition.



(4.16)

The adjusted temperatures can be obtained directly from $\{T_{a,k}\}$, and the adjusted geopotentials can be obtained from $\{\varphi_{a,k} = \phi_{a,k}/\beta_k\}$.

4.6 Computational Results Analysis

Table 4.2 shows the result at the same grid point as Table 4.1.

Compared with the results in Table 4.1, we can see that both temperatures and geopotentials have minor adjustments (in percentage), but no oscillation exists any more.

What we intend to achieve at this stage is to adjust the observed data such that the hydrostatic constraints are satisfied, while at the same time, the adjustment is minimized in the least squares sense. However, at this moment, it is still difficult to judge the effectiveness of data adjustment. In later chapters, we will show through numerical experiments

Table 4.2: A typical result for adjusting both T and φ

k	p	T_o	φ_o	T_a	φ_a	$(T_a - T_o)$	$(\varphi_a - \varphi_o)$
1	100.	-55.9	154683.2	-55.9	154792.0	0.0	108.8
2	150.	-55.5	129428.6	-55.5	129480.2	0.0	51.6
3	200.	-57.3	111553.4	-57.3	111579.8	0.0	26.4
4	250.	-57.4	97794.2	-57.4	97755.8	0.0	-38.4
5	300.	-53.6	86553.6	-53.6	86361.9	0.0	-191.7
6	400.	-40.7	67757.2	-40.6	67687.7	0.1	-69.5
7	500.	-29.3	52449.6	-29.1	52419.0	0.2	-30.6
8	700.	-13.2	28047.6	-13.0	28064.8	0.2	17.2
9	850.	-7.3	13406.4	-7.2	13401.3	0.1	-5.1
10	1000.	4.5	715.4	4.6	716.8	0.1	1.4

that this adjustment does improve the accuracy of the horizontal pressure gradient forces computed in the σ-coordinate system.

4.7 Sensitivity Analysis

Obviously, the results in Table 4.2 depend on the values of weighting factors α and $\{\beta_k\}$, which in turn are calculated from σ_T and σ_{φ_s} using (4.10) and (4.11). However, we wish that the results are not sensitive to the change of values in σ_T and σ_{φ_s}. Fortunately, this is indeed the case, as the results in Table 4.3 show.

The reason for obtaining this desirable consequence is that we assign the weights by exploring the intrinsic relationship of the standard deviation of T and φ at different levels, as expressed in (4.9). At upper levels where $\sigma^2_{\varphi_s}$ can be neglected, σ^2_{φ} is almost proportional to σ^2_T, therefore the relative weights for T and φ at upper levels are not affected by the

Table 4.3: Weight sensitivity analysis

	$(T_a - T_o)/(\sigma_a - \sigma_o)$					
σ_T	0.5	0.5	0.1	2.0	0.1	2.0
σ_φ	1.0	30.0	10.0	10.0	30. 0	1.0
1	0.0/108.8	0.0/109.1	0.0/109.6	0.0/108.8	0.0/112.2	0.0/108.8
2	0.0/51.5	0.0/51.9	0.0/52.3	0.0/51.6	0.0/54.5	0.0/51.5
3	0.0/26.3	0.0/26.6	0.0/27.1	0.0/26.4	0.0/28.9	0.0/26.4
4	0.0/-38.4	0.0/-38.2	0.0/-37.7	0.0/-38.4	0.0/-36.2	0.0/-38.4
5	0.0/-191.8	0.0/-191.5	0.0/-190.9	0.0/-191.8	0.0/-188.9	0.0/-191.8
6	0.1/-69.6	0.1/-68.9	0.1/-67.8	0.1/-69.6	0.1/-62.5	0.1/-69.6
7	0.2/-30.7	0.2/-29.4	0.2/-27.4	0.2/-30.7	0.1/-17.0	0.2/-30.7
8	0.2/17.0	0.2/19.3	0.2/23.1	0.2/17.0	0.1/42.9	0.2/17.0
9	0.1/-5.8	0.1/-1.8	0.1/3.0	0.1/-5.7	0.0/26.4	0.1/-5.8
10	0.1/0.0	0.0/6.5	0.0/12.0	0.1/0.1	0.0/36.7	0.1/0.0

choice of σ_T and σ_{φ_s}. On the other hand, at lower levels, the relative weights do change with different σ_T and σ_{φ_s}, however, because of the constraint conditions of (4.3), the values of T_a and φ_a at lower levels are not just determined by the weights alone, they also change the data at upper levels which tends to be not affected by different σ_T and σ_{φ_s}. Therefore, in overall, the results of T_a and φ_a are very insensitive to the change of values in σ_T and σ_{φ_s}.

Chapter 5

Numerical Experiments

5.1 Geostrophic Wind

Our initial purpose of doing the data adjustment is to improve the accuracy of the pressure gradient evaluated in the σ-coordinate system over mountainous terrain. The pressure gradient in turn can be used to compute geostrophic wind. Since geostrophic wind is easier for us to understand (in terms of m/sec) than the pressure gradient (in terms of m/sec^2 per unit mass), therefore, we choose geostrophic wind as the measurement of the accuracy of pressure gradient.

Geostrophic wind is proportional to the pressure gradient. The existence of pressure gradients in the atmosphere causes air to accelerate from regions of high pressure to regions of low pressure. The acceleration tends to continue unless it is balanced by other forces, such as friction force or the Coriolis force. The friction force at relatively higher levels in the atmosphere is usually very small, and therefore can be neglected. The Coriolis

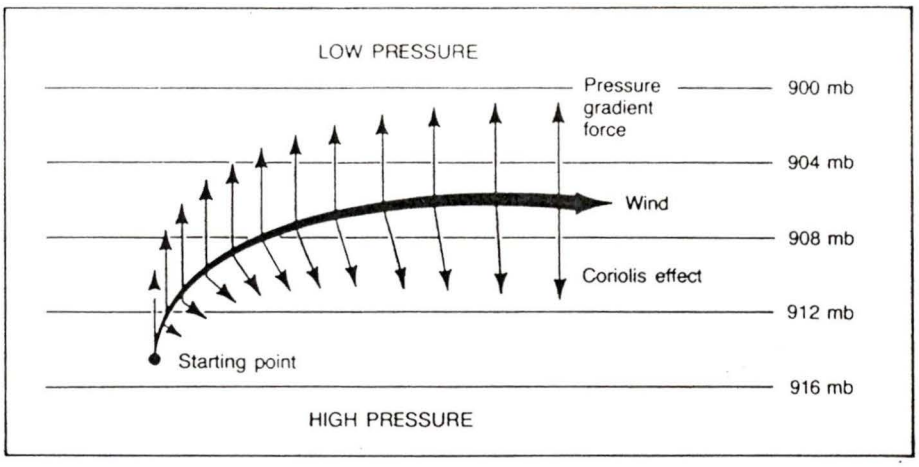


Figure 5.1: An illustration how geostrophic wind is formed on a horizontal surface

force, on the other hand, is proportional to wind speed, and tends to deflect wind to the right in the northern hemisphere. The faster wind speed will produce stronger Coriolis force, therefore result in further deflection until the pressure gradient force is balanced by the Coriolis force. Figure 5.1 illustrates this process. Geostrophic wind is the wind which flows parallel to isobars when the pressure gradient force and Coriolis force reach the state of equilibrium. When studying air flows at heights several kilometers above the earth's surface where friction forces can be neglected, geostrophic wind is a fairly good approximation of actual winds in many situations.

Geostrophic wind can be computed in the following way:

$$\begin{aligned} u_g &= -P_y/f \\ v_g &= P_x/f \end{aligned} \tag{5.1}$$

where u_g and v_g are the x and y components of geostrophic wind, P_x and P_y are pressure gradients in the x and y directions respectively, $f = 2\,\Omega\,sin\,\phi$ is the Coriolis parameter, Ω is the rate of rotation (angular velocity) of the earth, ϕ is latitude.

5.2 Evaluation of Geostrophic Wind

The evaluation of geostrophic wind is quite easy in the p-coordinate system. When the observed data is available on pressure surfaces, geostrophic wind can be computed directly using (1.1) and (5.1). This result can be regarded as accurate in the sense that no additional interpolation is needed, and the truncation error of the centered difference scheme applied is of the order of $O(d^2)$ where d is the grid step. Therefore, all the error analysis made in the following part is based on this result.

Should we not introduce the σ-coordinate system, all the troubles we are going to deal with can be avoided. Unfortunately, it is not the case. For achieving the benefits of the σ-coordinate system we had discussed in *Chapter 1*, we have to sacrifice some accuracy. But we want to make the error as small as possible.

One problem arises when we study the computational accuracy of geostrophic wind evaluated in the σ-coordinate system: how to make the result comparable with the accurate one we have mentioned above which is on pressure surfaces. The surfaces of these two systems usually do not coincide with each other; in mountainous areas, they differ

from each other substantially.

To solve the problem, we devised a computing procedure. Its main idea is to set the σ levels to coincide with the pressure levels on a point-to-point basis. Then the pressure gradients evaluated in the two different coordinate systems at one grid point should be the same if no error is introduced. This computing procedure is applied to all points within the grid we are interested in, and finally, we can get the result which is comparable with the accurate one. The following is a description of our computing procedure (with data being either adjusted or unadjusted):

1. Compute the earth's surface pressure p_s and temperature T_s at each grid point concerned.
2. For a given grid point, set σ levels such that each σ_k coincides with a specific pressure level p_k, that is, $\sigma_k = p_k/p_s,\ k = 1, 2, \cdots, n$.
3. Interpolate temperatures and geopotentials to the specified σ levels at the given grid point as well as the nearby ones. Geostrophic wind can then be computed using (1.4) and (5.1) for that grid point.

Figure 5.2 may help to understand the procedure. At point i, the σ levels are set to coincide with the pressure levels. We interpolate data at points $i-1$ and $i+1$ to the same σ levels as point i, and then the pressure gradient and geostrophic wind can be computed for point i. As can be seen that, at point i, result obtained in the σ-coordinate system is at the same places as that obtained in the p-coordinate system, and therefore, the two are made comparable.

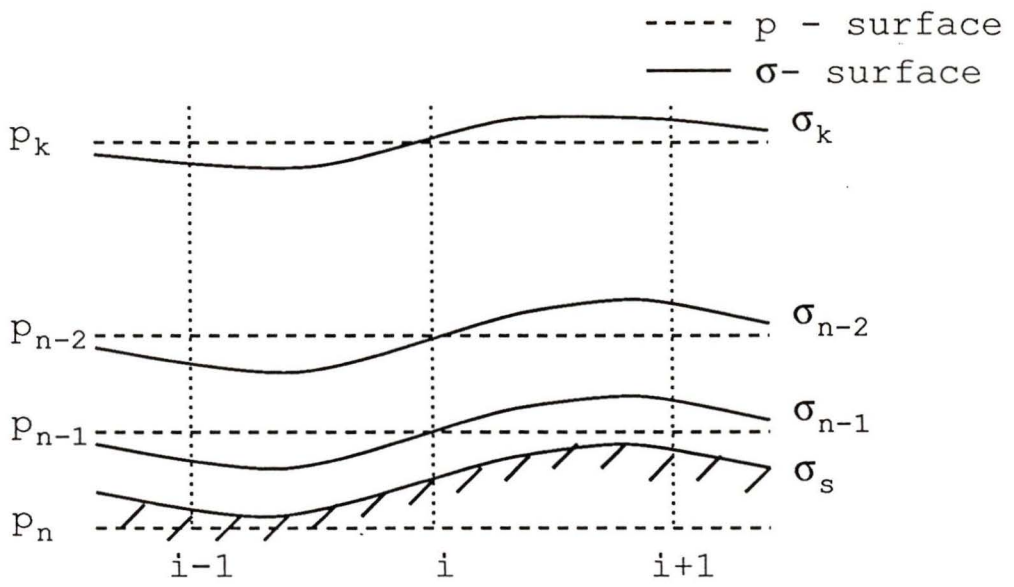


Figure 5.2: σ surfaces are set to coincide with p surfaces at point i

5.3 Design of the Numerical Experiments

In our numerical experiments, we mainly investigate two factors which affect the accuracy of geostrophic wind evaluated in the σ-coordinate system. One is the data inconsistency, the other is the finite difference scheme. To study the individual effects of these two factors, we use the traditional engineering approach of varying one factor and keeping the other fixed in each experiment. We also study the combined effects of these two factors.

Design of Experiment 1

This experiment is to investigate the effects of data adjustment. We designed two approaches to evaluate geostrophic wind in the σ-coordinate system, both using Corby's

finite difference scheme:

- **Direct Approach:** With the observed data on pressure surfaces, compute geostrophic wind directly with our computing procedure described in the previous section.
- **Data Adjusting Approach:** First adjust the observed data. Then with the adjusted data, apply the same procedure to evaluate geostrophic wind.

Since the only difference in these two approaches is that the data used is either adjusted or unadjusted, therefore, it is fair for us to justify the effectiveness of data adjustment by comparing the results of these two approaches. Furthermore, by looking into each step of the second approach, we can also get some information about where the errors mainly come from.

Design of Experiment 2

The second experiment is to investigate the effects of Corby's finite difference scheme, which is supposed to be superior to the centered difference scheme in the evaluation of pressure gradient in the σ-coordinate system. We made the experiments for two different cases:

- **Case 1:** With the data used being adjusted, we evaluate geostrophic wind using our computing procedure, but one employs Corby's finite difference scheme, the other employs the centered difference scheme.
- **Case 2:** With the data used being unadjusted, we repeat the same experiment as case 1.

By comparing two different schemes for each case, we can fairly justify whether Corby's finite different scheme is really superior to the centered difference scheme, how much, and under what conditions. Some interesting results are discovered.

5.4 Sample Data

The data we have chosen to use is given on a regularly spaced grid covering a large part of the North Pacific, a part of the west coast from northern California to Alaska. The grid consists of 28 points along the x (east-west) axis and 24 points along the y (north-south) axis. The points within the grid are equidistant along both axes, with the distance $d = 190.5$ km. This grid is obtained by first making the polar stereographic projection with the cutting plane at 60^o north latitude, and the source of the projection at the source pole, and then interpolating from (14 x 12) points into (28 x 24) points to double the horizontal resolution. The actual distance between grid points on the earth's surface varies with latitude ϕ, and differs from the distance on the map grid by a factor $(1 + sin\,\phi)/(1 + sin\,60^o)$.

At each grid point, we need the following data:

- The terrain elevation Z_s which is used to compute surface pressure p_s and temperature T_s;
- The temperatures $\{T_{o,k}\}$ and heights $\{Z_{o,k}\}$ given at the pressure levels (milibar) $\{p_k\} = \{100, 150, 200, 250, 300, 400, 500, 700, 850, 1000\}$. The height Z_k can be converted into the geopotential φ_k by multiplying the gravitational force g:

$$\varphi_k = g\, Z_k.$$

However, a grid of (28 x 24) seems to be unnecessarily too large for experimental purpose in our case, since the data adjustment is made independently at each point, and the computation of geostrophic wind at one point involves only the data at nearby points. Thus a grid of (5 x 5) will produce results at 9 points (excluding boundary points) in both x and y directions; at each gird point, we compute geostrophic wind at up to 10 levels.

Therefore, at totally up to 180 places the geostrophic wind is computed and compared in each experiment for a (5×5) grid, which is enough to give us convincing statistical results.

It is also of interest and importance to observe how these procedures affects the computational accuracy when applied to different terrains. For this reason, we select two (5×5) sample grids within the (28×24) grid, one is on the Pacific Ocean, and therefore can be regarded as flat; the other is somewhere inside the west coast and can be regarded as mountainous. Each sample grid covers an area of about 760×760 (km^2). For the grid over mountainous terrain, the overall slope from end to end is about 1500 m/760 km on average in north-south direction, and about 300 m/760 km on average in east-west direction. The steepest slope between two grid points is about 900 m/190 km.

5.5 Some Computational Details

Some computational details are explained in this section for computing geostrophic wind in the σ-coordinate system with the procedure and sample data we have described above. The first step of computation is to evaluate the surface pressure p_s, based on p_s obtained, we can set the σ levels for each grid point, and then do the interpolation and computation.

The Computation of the Surface Pressure

For convenience, we consider at this moment the case that surface elevation is below the 850 mb height. As assumed in (4.1) that, in the layer between 850 mb and 1000 mb, temperature T and pressure p should satisfy

$$T(p) = A_{10} \ln p + B_{10}. \tag{5.2}$$

We also assume that (5.2) still holds in the layer below the 1000 mb level if the surface elevation is below the 1000 mb height. Then we can get A_{10} and B_{10} by

$$A_{10} = \frac{T_{1000} - T_{850}}{ln\,(p_{1000}/p_{850})} \tag{5.3}$$

$$B_{10} = T_{1000} - A_{10}\,ln\,p_{1000}. \tag{5.4}$$

Following (4.2), we integrate from φ_{1000} to φ_s,

$$\varphi_s - \varphi_{1000} = -\frac{R\,A_{10}}{2}[(ln\,p_s)^2 - (ln\,p_{1000})^2] - R\,B_{10}\,(ln\,p_s - ln\,p_{1000}). \tag{5.5}$$

Substituting (5.4) into (5.5), and letting $p' = p_{1000}/p_s$, we can obtain the following equation

$$(ln\,p')^2 - (\frac{2\,T_{1000}}{A_{10}})\,ln\,p' + \frac{2}{R\,A_{10}}\,(\varphi_s - \varphi_{1000}) = 0 \tag{5.6}$$

which is a quadratic in $ln\,p'$. By solving (5.6), we get

$$ln\,p' = \frac{2\,T_{1000}/A_{10} \pm \sqrt{(2\,T_{1000}/A_{10})^2 - 8(\varphi_s - \varphi_{1000})/(R\,A_{10})}}{2} \tag{5.7}$$

where $A_{10} \neq 0$. We choose the root with the opposite sign before the radical to the sign of A_{10}, then we can have $p_s = p_{1000}$ when $\varphi_s = \varphi_{1000}$.

If A_{10} is very small, (5.6) and (5.7) cannot be used. In that case, we can get $ln\,p'$ directly from (5.5) and (5.4) by letting $A_{10} = 0$,

$$ln\,p' = \frac{\varphi_s - \varphi_{1000}}{R\,T_{1000}}. \tag{5.8}$$

After we get $ln\,p'$, the surface pressure can be obtained by

$$ln\,p_s = ln\,p_{1000} - ln\,p'. \tag{5.9}$$

The surface temperature can also be obtained by

$$T_s = A_{10} \ln p_s + B_{10}. \tag{5.10}$$

If the surface elevation is in a layer above the 850 mb surface, we can apply the same procedure except that the parameters used is from that layer.

Set the σ levels

The computation of the pressure gradient force at point i using (1.4) requires that the σ-levels are set to coincide with the pressure levels for the purpose of comparability. Also we need to set the σ levels at nearby points $i-1$ and $i+1$ to be the same as point i. This can be done in the following way (given the surface pressures at points i and $i \pm 1$):

- For point i, set the σ levels:

$$\sigma_k = \begin{cases} p_k/p_{s,i} & if \;\; p_k \le p_{s,i} \\ -1 & if \;\; p_k > p_{s,i} \end{cases} \qquad for \;\; k = 1, 2, \cdots, 10,$$

  where $\sigma_k = -1$ is a flag indicating that the k^{th} pressure surface is below the earth surface, and therefore can be skipped.

- For points $i-1$ and $i+1$, we want to locate the position of these σ levels in the pressure coordinate so that we can interpolate data onto them:

$$p_{i\pm1}(\sigma_k) = \sigma_k \, p_{s,i\pm1}, \qquad k = 1, 2, \cdots, 10 \;\; and \;\; \sigma_k \neq -1.$$

5.6 Interpolation and Computation

When the σ levels are set for a given point, we can then interpolate data (T and φ) from the pressure surfaces to the σ surfaces for computing geostrophic wind in the σ-coordinate system.

As assumed that, within a segment between two adjacent pressure levels, temperature varies linearly with the logarithm of pressure. Thus the interpolation of temperature is straight forward as long as we know the location of the σ level in the pressure coordinate. Suppose we know that σ_k is in the segment between p_l and p_{l+1}, that is, $p_l < p(\sigma_k) \leq p_{l+1}$. Then the linear interpolation can be applied:

$$T(\sigma_k) = T_{l+1} + \frac{T_{l+1} - T_l}{ln\,(p_{l+1}/p_l)}\, ln\,[p(\sigma_k)/p_{l+1}] \tag{5.11}$$

where T_l and T_{l+1} are the temperatures at p_l and p_{l+1} respectively.

The geopotential at σ_k is calculated iteratively through the numerical integration along the piecewise linearized curve from the earth's surface upward. Suppose we have got the data for σ_{k+1} level, and want to calculate $\varphi(\sigma_k)$ from $p(\sigma_k)$ and $T(\sigma_k)$. There are basically two cases needed to be considered:

- σ_{k+1} and σ_k are in the same segment between p_{l+1} and p_l, then $\varphi(\sigma_k)$ can be calculated directly:

$$\varphi(\sigma_k) = \varphi(\sigma_{k+1}) + \frac{R}{2}\,[T(\sigma_{k+1}) + T(\sigma_k)]\, ln\,[\frac{p(\sigma_{k+1})}{p(\sigma_k)}]. \tag{5.12}$$

- σ_{k+1} and σ_k are not in the same segment, suppose one is between p_{l+1} and p_l, while the other is between p_l and p_{l-1}, then the integration should be made from one segment to the next because different segment may have different slope. In this case, $\varphi(\sigma_k)$ can be computed in the following way:

$$\begin{aligned} \varphi(\sigma_k) &= \varphi(\sigma_{k+1}) + \frac{R}{2}\,[T(\sigma_{k+1}) + T_l]\, ln\,[\frac{p(\sigma_{k+1})}{p_l}] \\ &+ \frac{R}{2}\,[T_l + T(\sigma_k)]\, ln\,[\frac{p_l}{p(\sigma_k)}]. \end{aligned} \tag{5.13}$$

  If σ_{k+1} and σ_k are separated by more than one segment, the integration can be made in a similar way using iteration.

Once the data is ready on σ surfaces, the computation of geostrophic wind can be done directly using (1.4) and (5.1).

Chapter 6

Error Analysis

6.1 Error Sources Analysis

Errors introduced during the computation of geostrophic wind in our numerical experiments mainly come from two sources: the errors from data inconsistency, and the error from the approximation by finite difference schemes. We have shown in *Chapter 3* that the inconsistency of data will impose errors on the interpolated data. If we adjust data in advance, then the interpolation will be made consistent, but the data adjustment itself will be subject to some errors. However, attention should be paid to the fact that, no matter whatever method is used, these two sources of errors could not be avoided if we intend to employ the σ-coordinate system. We are going to see how much error can be reduced by applying data adjustment and using Corby's finite difference scheme.

All experiments are conducted over two different grids separately, with one over flat terrain and the other over mountainous terrain, for investigating the effects of the earth's

surface on the computational accuracy.

6.2 Error Analysis for Experiment 1

The following notation is used in our analysis for experiment 1 described in *Section 5.3*:

GW: Geostrophic wind;
$GW_{u,p}$: GW computed in the p system with unadjusted data;
$GW_{a,p}$: GW computed in the p system with adjusted data;
$GW_{a,\sigma}$: GW computed in the σ system with adjusted data;
$GW_{u,\sigma}$: GW computed in the σ system with unadjusted data.

From what we had explained in *Section 5.2*, $GW_{u,p}$ can be regarded as the accurate solution based on which comparisons are made.

For the direct approach, $GW_{u,\sigma}$ is computed directly from the observed data, and the errors are calculated by comparing $GW_{u,\sigma}$ and $GW_{u,p}$.

The data adjusting approach actually consists of two steps: data adjustment and computation of GW. It is interesting to investigate in detail the errors that arise in each of the two steps, as well as the actual errors in the final results. The errors in the first step are purely from data adjustment, which are calculated by comparing $GW_{a,p}$ with $GW_{u,p}$, both in the p-coordinate system, but one with the adjusted data. Then with the adjusted data, we calculate the errors in the second step by comparing $GW_{a,\sigma}$ with $GW_{a,p}$, same data but in different coordinate systems. The actual errors of the approach are calculated by comparing $GW_{a,\sigma}$ directly with $GW_{u,p}$, i.e., the accurate solution. This result is of the most importance to us since it is the actual error we are subject to. It is reasonable to expect that the final errors should be larger than that in a single step, but smaller than the sum of errors in the two steps.

By looking at the results from Table 6.1 and Table 6.2, some remarks can be concluded about the two approaches:

1. Compared with the direct approach, the data adjusting approach achieves 25.6% and 21.3% reduction of errors, in terms of mean absolute error, over flat and mountainous terrains respectively; and achieves 25.4% and 22.3% reduction of errors, in terms of root mean squares of error, over flat and mountainous terrains respectively.
2. Surprisingly, the errors in the data adjusting approach mainly come from the first step, which means that the main errors result from the inconsistency of the data. The errors in the second step are much less compared with that in the first step. Over flat terrain, the errors in the second step can almost be neglected.
3. Not surprisingly, the sum of the errors in the first two steps are greater than the actual errors. This indicates that, more or less, there is some cancellation in the errors produced in the first two steps.
4. By comparing the worst cases of the two approaches at each pressure level, we can see 32.3% and 21.3% reduction on average over flat and mountainous terrains respectively; and the maximum errors of 5.397 and 5.049 over the two terrains are reduced to 3.282 and 3.462 respectively (from Table C.5).

The above conclusions obviously reveal the effectiveness of the data adjustment in every aspect: absolute errors, root mean squares of errors and the largest errors. These conclusions are on the average basis. At some individual points or levels, the results of the data adjusting approach may be worse than that of the direct approach, which is understandable.

6.3 Error Analysis for Experiment 2

The following notations are used in our analysis for experiment 2 described in *Section 5.3*:

$\overline{GW}^{Cby}$: GW computed using Corby's finite difference scheme;
$\overline{GW}^{Cnt}$: GW computed using Centered difference scheme.

The subscripts will have the same meaning as that for experiment 1.

Table 6.1: Experiment 1: comparison of the results over *flat terrain*

	Data Adjust Approach			Direct Approach	Reduction of Error
Overall Mean $\|err\|$	A	B	C	D	$\frac{D-C}{D}$
in x-component	0.900	0.010	0.902	1.118	19.3%
in y-component	0.805	0.002	0.804	1.179	31.8%
Overall RMS of err	AA	BB	CC	DD	$\frac{DD-CC}{DD}$
in x-component	1.196	0.013	1.199	1.462	18.0%
in y-component	1.109	0.003	1.112	1.656	32.9%
Avg. Max. $\|err\|$			E	F	$\frac{F-E}{F}$
in x-component			1.647	2.215	25.6%
in y-component			1.698	2.725	37.7%

Table 6.2: Experiment 1: comparison of the results over *mountainous terrain*

	Data Adjust Approach			Direct Approach	Reduction of Error
Overall Mean $\|err\|$	A	B	C	D	$\frac{D-C}{D}$
in x-component	1.174	0.279	1.273	1.567	18.7%
in y-component	0.916	0.112	0.952	1.251	23.9%
Overall RMS of err	AA	BB	CC	DD	$\frac{DD-CC}{DD}$
in x-component	1.488	0.360	1.544	1.953	20.9%
in y-component	1.184	0.150	1.197	1.567	23.6%
Avg. Max. $\|err\|$			E	F	$\frac{F-E}{F}$
in x-component			2.236	2.875	22.2%
in y-component			1.812	2.271	20.2%

Notes: The data in Table 6.1 and Table 6.2 is extracted from Table C.1 to Table C.5 in *Appendix C.*

$A = \frac{1}{n}\sum_1^n |GW_{a,p} - GW_{u,p}|$, $B = \frac{1}{n}\sum_1^n |GW_{a,s} - GW_{a,p}|$,

$C = \frac{1}{n}\sum_1^n |GW_{a,s} - GW_{u,p}|$, $D = \frac{1}{n}\sum_1^n |GW_{u,s} - GW_{u,p}|$,

$AA = \sqrt{\frac{1}{n}\sum_1^n (GW_{a,p} - GW_{u,p})^2}$, $BB = \sqrt{\frac{1}{n}\sum_1^n (GW_{a,s} - GW_{a,p})^2}$,

$CC = \sqrt{\frac{1}{n}\sum_1^n (GW_{a,s} - GW_{u,p})^2}$, $DD = \sqrt{\frac{1}{n}\sum_1^n (GW_{u,s} - GW_{u,p})^2}$,

$E = \max\{|GW_{a,s} - GW_{u,p}|\}$, $F = \max\{|GW_{u,s} - GW_{u,p}|\}$.

The purpose of experiment 2 is to investigate the effectiveness of Corby's finite difference scheme vs. the conventional centered difference scheme in computing GW in the σ-coordinate system. To achieve this purpose, we use the same data and the same procedure for each case, except that the finite difference scheme is different.

For case 1, we use the adjusted data to compute $\overline{GW}_{as}^{Cby}$ and $\overline{GW}_{as}^{Cnt}$, and compare them with GW_{ap}. For case 2, we use the unadjusted data to compute $\overline{GW}_{us}^{Cby}$ and $\overline{GW}_{us}^{Cnt}$, and compare them with GW_{up}. We did this over both flat and mountainous terrain to check the effects of the earth's surface on the accuracy of finite difference scheme.

By looking at the results from Table 6.3 and Table 6.4, some remarks can be concluded about the two finite difference schemes:

1. In both cases, Corby's finite difference scheme achieves better computational accuracy than the centered difference scheme for data over either flat or mountainous terrains.
2. In terms of percentage of error reduction, Corby's finite difference scheme is much more superior to the centered difference scheme in the case that data is adjusted (more than 50% reduction of truncation error over both flat and mountainous terrains). While in the case that data is unadjusted, the improvement is minor.
3. In terms of absolute value of error reduction, Corby's finite difference scheme achieves more improvement for data over mountainous terrain than over flat terrain. That is why it is important to employ Corby's finite difference scheme when the terrain is not flat.
4. For both finite difference schemes, the accuracy is greatly improved when data is adjusted. This implies that the computational error of finite difference scheme is heavily dependent on the consistency of data.

Experiment 2 showed that Corby's finite difference scheme is effective in error reduction, especially over mountainous terrain. It also helps to improve accuracy in flat terrain. However, all these are achieved only after the data is adjusted.

Table 6.3: Experiment 2: comparison of the results over *flat terrain*

	CASE 1 Data Adjusted			CASE 2 Data Unadjusted		
Overall Mean $\|err\|$	Cby	Cnt	$\frac{Cnt-Cby}{Cnt}$	Cby	Cnt	$\frac{Cnt-Cby}{Cnt}$
in x-component	0.010	0.021	52.4%	1.118	1.122	0.3%
in y-component	0.002	0.005	60.0%	1.179	1.179	0.0%
Overall RMS of err	Cby	Cnt	$\frac{Cnt-Cby}{Cnt}$	Cby	Cnt	$\frac{Cnt-Cby}{Cnt}$
in x-component	0.013	0.027	51.9%	1.462	1.465	0.2%
in y-component	0.003	0.006	50.0%	1.656	1.655	0.0%

Table 6.4: Experiment 2: comparison of the results over *mountainous terrain*

	CASE 1 Data Adjusted			CASE 2 Data Unadjusted		
Overall Mean $\|err\|$	Cby	Cnt	$\frac{Cnt-Cby}{Cnt}$	Cby	Cnt	$\frac{Cnt-Cby}{Cnt}$
in X-component	0.279	0.750	62.8%	1.567	1.841	14.9%
in y-component	0.112	0.227	50.7%	1.251	1.289	2.9%
Overall RMS of err	Cby	Cnt	$\frac{Cnt-Cby}{Cnt}$	Cby	Cnt	$\frac{Cnt-Cby}{Cnt}$
in x-component	0.360	0.962	62.6%	1.953	2.165	9.8%
in y-component	0.150	0.300	50.0%	1.567	1.534	-2.2%

Notes: The data in Table 6.3 and Table 6.4 is extracted from Table C.6 to Table C.9 in *Appendix C.*

Cby — Using Corby's finite difference scheme:

$$-(\tfrac{\partial \varphi}{\partial x})_\sigma - RT\,\tfrac{\partial ln\, p_s}{\partial x} \simeq -\delta_x \overline{\varphi}^x - R\overline{\overline{T}^x \delta_x ln\, p_s}^x$$

Cnt — Using the centered difference scheme:

$$-(\tfrac{\partial \varphi}{\partial x})_\sigma - RT\,\tfrac{\partial ln\, p_s}{\partial x} \simeq -\delta_x \overline{\varphi}^x - RT\delta_x \overline{ln\, p_s}^x$$

6.4 Concluding Remarks

The above described numerical experiments strongly suggest that, in order to achieve better computational accuracy, both data adjustment and Corby's finite difference scheme should be used, which will reduce the overall error we are subject to by more than 20%. The experiments also show that these two procedures are also very helpful to improve the accuracy of geostrophic wind in flat terrain, though our initial purpose is to reduce the errors in mountainous terrain.

Theoretically speaking, the procedure of data adjustment and Corby's finite difference scheme do not guarantee the improvement of the worst cases. That means that the maximum possible value of error may be as large as that without the data adjustment and using the centered finite difference scheme. However, we do see obvious reduction of the maximal value of error at each level on average.

The first experiment also reveals that the errors which appear in the σ-coordinate system mainly come from the inconsistency of the observed data, while the errors produced by finite difference scheme are relatively small.

The second experiment reveals that Corby's finite difference scheme, which is superior to the centered difference scheme, is much more effective when applied to the adjusted data and over mountainous terrain.

The numerical experiments also imply that merely seeking improvement of finite difference scheme without solving the problem of data inconsistency will not produce desirable results.

Bibliography

[1] Arakawa, A. and Lamb, V. (1977): "Computational Design of the Basic Dynamical Processes of the UCLA General Circulation Model", *Methods in Computational Physics, Vol.17, pp 174-267.*

[2] Atkinson, B.W. (1981) edited: "Dynamical Meteorology, An Introductory Selection" *Methuen, London and New York.*

[3] Barker, E.H., (1980): "Solving for Temperature Using Unnaturally Latticed Hydrostatic Equation", *Monthly Weather Review, Vol.108, pp 1260-1268.*

[4] Corby,G.A., Gilchrist, A. and Newson, R.L. (1972): "A General Circulation Model of the Atmosphere Suitable for Long Period integrations", *Quarterly Journal of the Royal Meteorological Society, Vol 98, pp 809-832.*

[5] Danard, M., (1989): "On Computing the Surface Horizontal Pressure Gradient over Elevated Terrain", *Monthly Weather Review, Vol.117, pp 1-7.*

[6] Danard, M., and Galbraith, J., (1987): "Meso Gamma Wind Modeling for the 1988 Calgary Winter Olympics", *Technical Report, Atmospheric Dynamics Corp., Victoria, B.C., Canada.*

[7] Gary, J.M., (1973): "Estimate of Truncation Error in Transformed Coordinate, Primitive Equation Atmospheric Models", *Journal of the Atmospheric Sciences, Vol.30, pp 223-233.*

[8] Haltiner, G.J. and Williams, R.T. (1980): "Numerical Prediction and Dynamic Meteorology", *John Wiley and Sons, New York.*

[9] Kasahara, A. (1977): "Computational Aspects of Numerical Models for Weather Prediction and Climate Simulation", *Methods in Computational Physics, Vol.17, pp 2-66.*

[10] Kurihara, Y., (1968): "Note on Finite Difference Expressions for the Hydrostatic Relation and Pressure Gradient force", *Monthly Weather Review, Vol.96, No.9, pp 654-656.*

[11] McEwen, J.N., Danard, M. and Davidson, G.A. (1981): "Variationally Adjusted Surface Winds", *Boundary-Layer Meteorology, Vol.20, pp 473-483.*

[12] Phillips, N.A.,(1957): "A Coordinate system having some special advantages for numerical forecasting". *Journal of Meteorology, Vol.14, pp 184-185.*

[13] Phillips, N.A., (1973): "Principles of Large Scale Numerical Weather Prediction", Dynamic Meteorology, edited by Morel, P., pp 1-96. *D.Reidel Publishing Co., Dordrecht, Holland.*

[14] Sangster, W.E., (1987): "An Improved Technique for Computing the Horizontal Pressure-Gradient Force at the Earth's Surface", *Monthly Weather Review, Vol.115, pp 1358-1369.*

[15] Sasaki, Y., (1958): "An Objective Analysis Based on the Variational Method", *Journal of the Meteorological Society of Japan, Vol.36, No.3, pp 77-88.*

[16] Sasaki, Y., (1960): "An Objective Analysis for Determining Initial Conditions of the Primitive Equations", *Technical Report, The A&M College of Texas, USA.*

[17] Sasaki, Y. (1970): "Some Basic Formalisms in Numerical Variational Analysis", *Monthly Weather Review, Vol.98, pp 875-883.*

[18] Saucier, W.J. (1965): "Principles of Meteorological Analysis", *The University of Chicago Press, Chicago & London.*

[19] Smagorinsky, J., Strickler, R.F., Sangster, W.E., Manabe, S., Holloway, J.L. and Hembree, G.D., (1967): "Prediction Experiments with a General Circulation Model", *Proceedings of the International Symposium on Dynamics of Large Scale Atmospheric Processes, Moscow, pp 70-134.*

[20] Sundqvist, H., (1975): "On Truncation Errors in Sigma-System models", *Atmosphere, Vol.13, No.3, pp 81-95.*

[21] Thompson, P.D. (1961): "Numerical Weather Analysis and Prediction", *Macmillan, New York.*

[22] Watkins, D.S., (manuscript, to be published): "Fundamentals of Matrix Computations", *Washington State University.*

Appendix A

List of Symbols

Symbol	Meaning	Equation number of first appearance
∇	horizontal gradient operator	(1.1)
$\delta\,()$	variational operator	(3.5)
$\delta_x\,()$	difference operator	(1.4)
$\overline{()}^x$	averaging operator	(1.4)
$\Delta\, x$	grid step	(1.5)
$\alpha^2,\ \beta^2$	weighting factor	(4.5)
φ	geopotential height, $\varphi = g\, Z$	(1.1)
φ_s	earth's surface geopotential height	(2.4)
ϕ	rescaled variable of φ	(4.5)
ρ	the air density	(1.1)
ρ_k	coefficient	(4.15)
σ	the vertical coordinate, defined as p/p_s	(1.2)
σ^2	square deviation	(4.9)
μ	Lagrange multiplier	(3.7)
a	subscript standing for adjusted data	(3.11)
f	Coriolis parameter	(5.1)
g	the gravitational acceleration	(1.1)

o	subscript standing for observed data	(3.11)
p	pressure	(1.1)
p_s	earth's surface pressure	(1.2)
P	horizontal pressure gradient	(1.1)
R	the gas constant	(1.3)
T	temperature	(1.3)
T_s	earth's surface temperature	(2.8)
u_g, v_g	the x and y component of geostrophic wind	(5.1)
x	horizontal coordinate	(1.4)
Z	geometrical height	(1.1)

Appendix B

Analysis for Adjustment with only Temperature

Let

$$\epsilon_k = T_{o,k} - T_{a,k} \qquad k = 1, 2, \cdots, n. \tag{B.1}$$

By fitting the observed temperature $\{T_{o,k}\}$ into (4.3), we define

$$\eta_k = T_{o,k-1} + T_{o,k} - f_k \qquad k = 2, 3, \cdots, n \tag{B.2}$$

being the differences by which (4.3) is not satisfied. It can be shown that there is a relationship between $\{\epsilon_k\}$ and $\{\eta_k\}$

$$\eta_k = \epsilon_{k-1} + \epsilon_k, \qquad k = 2, 3, \cdots, n. \tag{B.3}$$

For a simple case where $\eta_j \neq 0$, and $\eta_k = 0$, for all $k \neq j$, (B.3) becomes

$$\epsilon_{k-1} + \epsilon_k = \begin{cases} \eta_j & \text{if } k = j \\ 0 & \text{otherwise} \end{cases} \tag{B.4}$$

In this case, the problem of data adjusting becomes to minimize

$$I = \sum_{k=1}^{n} \frac{1}{2}\epsilon_k^2 \tag{B.5}$$

subject to (B.4).

From (B.4), we can get

$$|\epsilon_1| = |\epsilon_2| = \cdots = |\epsilon_{j-1}| = |\eta_j - \epsilon_j| \tag{B.6}$$

and

$$|\epsilon_{j+1}| = \cdots = |\epsilon_n| = |\epsilon_j| \tag{B.7}$$

Substituting (B.6) and (B.7) into (B.5), differentiating it with respect to ϵ_j, and letting it be zero, we can then solve for ϵ_j

$$\epsilon_j = \frac{j-1}{n}\eta_j.$$

Then the optimal solution to (B.5) subject to (B.4) has the absolute value

$$|\epsilon_k^*| = \begin{cases} (n-j+1)\,|\eta_j|\,/n & k < j \\ (j-1)\,|\eta_j|\,/n & \text{if } k \geq j \end{cases} \tag{B.8}$$

Since from the fact that the equations in (B.4) imply

$$\epsilon_{k-1} = -\epsilon_k \qquad \text{for all } k \neq j,$$

the errors caused by η_j will propagate from j toward both ends of the index, with the same magnitude, but opposite sign for each next one in each direction . This may be responsible for the oscillation occurred in $\{T_{a,k}\}$.

For the more general case where all $\eta_k \neq 0, k = 2,3,\cdots,n$ in (B.3), it can be shown that

$$\epsilon_k = \sum_{j=2}^{n} \epsilon_k^{(j)}, \qquad k = 1,2,\cdots,n, \tag{B.9}$$

where $\epsilon_k^{(j)}$ is the error caused individually by η_j in the same manner as the above simple case. Though there may be some cancellation in $\sum \epsilon_k^{(j)}$, generally speaking, some larger $|\eta_j|$ will dominate the values of $\{\epsilon_k\}$, an oscillation could be expected. The result in Table 1 may suggest that the largest η_j is at $j = 4$.

Appendix C

Detailed Experimental Results

Table C.1: Experiment 1: mean absolute errors over flat terrain
in x-component of geostrophic wind

k	p(mb)	M.Wind	Data Adjust Approach			Direct Approach
			A	B	C	D
1	100.	28.111	0.546	0.002	0.547	0.837
2	150.	32.667	0.619	0.005	0.617	0.976
3	200.	35.331	0.507	0.012	0.498	0.597
4	250.	38.598	1.361	0.017	1.377	1.545
5	300.	39.573	2.268	0.011	2.278	2.328
6	400.	36.764	1.571	0.007	1.565	1.471
7	500.	32.466	0.627	0.012	0.624	1.136
8	700.	25.675	1.021	0.009	1.029	1.524
9	850.	21.692	0.412	0.012	0.408	0.747
10	1000.	18.214	0.071	0.015	0.077	0.014
Overall Mean $\|err\|$			0.900	0.010	0.902	1.118

in y-component of geostrophic wind

k	p(mb)	M.Wind	Data Adjust Approach			Direct Approach
			A	B	C	D
1	100.	19.457	1.551	0.003	1.551	2.249
2	150.	19.113	1.060	0.002	1.058	1.767
3	200.	19.084	0.913	0.002	0.911	1.352
4	250.	18.253	0.608	0.001	0.607	1.306
5	300.	18.139	0.731	0.001	0.730	0.529
6	400.	15.932	0.593	0.003	0.594	0.914
7	500.	14.471	0.802	0.002	0.803	1.303
8	700.	11.577	1.516	0.001	1.516	2.011
9	850.	12.408	0.219	0.004	0.217	0.354
10	1000.	11.476	0.052	0.004	0.055	0.004
Overall Mean $\|err\|$			0.805	0.002	0.804	1.179

Notes: A, B, C and D are the same as in Table 6.1 and Table 6.2.

Table C.2: Experiment 1: R.M.S. of errors over flat terrain

x-component of geostrophic wind

k	p(mb)	M.Wind	Data Adjust Approach			Direct Approach
			AA	BB	CC	DD
1	100.	28.111	0.659	0.003	0.660	0.930
2	150.	32.667	0.713	0.006	0.712	1.181
3	200.	35.331	0.619	0.013	0.610	0.618
4	250.	38.598	1.484	0.020	1.502	1.832
5	300.	39.573	2.345	0.012	2.356	2.452
6	400.	36.764	1.729	0.010	1.721	1.830
7	500.	32.466	0.746	0.015	0.743	1.305
8	700.	25.675	1.145	0.010	1.153	1.795
9	850.	21.692	0.505	0.014	0.496	0.917
10	1000.	18.214	0.089	0.017	0.100	0.016
Overall R.M.S. of *err*			1.196	0.013	1.199	1.462

y-component of geostrophic wind

k	p(mb)	M.Wind	Data Adjust Approach			Direct Approach
			AA	BB	CC	DD
1	100.	19.457	1.829	0.003	1.830	2.723
2	150.	19.113	1.371	0.002	1.369	2.278
3	200.	19.084	1.157	0.002	1.155	1.780
4	250.	18.253	0.674	0.002	0.673	1.561
5	300.	18.139	0.804	0.002	0.803	0.677
6	400.	15.932	0.769	0.003	0.771	1.062
7	500.	14.471	1.095	0.003	1.096	1.470
8	700.	11.577	1.644	0.002	1.643	2.219
9	850.	12.408	0.265	0.005	0.263	0.490
10	1000.	11.476	0.070	0.005	0.074	0.005
Overall R.M.S. of *err*			1.109	0.003	1.112	1.656

Notes: AA, BB, CC and DD are the same as in Table 6.1 and Table 6.2.

Table C.3: Experiment 1: mean absolute errors over mountainous terrain

x-component of geostrophic wind

k	p(mb)	M.Wind	Data Adjust Approach			Direct Approach
			A	B	C	D
1	100.	8.310	1.937	0.198	1.952	2.583
2	150.	5.903	1.451	0.222	1.557	2.252
3	200.	4.986	1.476	0.107	1.429	1.474
4	250.	3.181	1.273	0.234	1.312	1.890
5	300.	0.115	0.695	0.280	0.878	0.914
6	400.	-3.525	1.413	0.089	1.391	1.620
7	500.	-1.891	1.768	0.430	1.813	2.032
8	700.	5.989	0.222	0.495	0.475	0.564
9	850.	9.897	0.327	0.453	0.646	0.778
Overall Mean $\|err\|$			1.174	0.279	1.273	1.567

y-component of geostrophic wind

k	p(mb)	M.Wind	Data Adjust Approach			Direct Approach
			A	B	C	D
1	100.	29.601	1.328	0.110	1.365	1.894
2	150.	24.930	1.334	0.054	1.291	1.466
3	200.	23.984	1.672	0.080	1.705	2.208
4	250.	21.606	0.650	0.066	0.672	0.876
5	300.	23.153	1.020	0.068	0.997	1.315
6	400.	23.354	0.978	0.110	0.974	1.358
7	500.	18.826	0.660	0.185	0.664	0.856
8	700.	10.373	0.489	0.222	0.700	0.915
9	850.	6.899	0.115	0.116	0.200	0.371
Overall Mean $\|err\|$			0.916	0.112	0.952	1.251

Notes: A, B, C and D are the same as in Table 6.1 and Table 6.2.

Table C.4: Experiment 1: R.M.S. of errors over mountainous terrain

x-component of geostrophic wind

k	p(mb)	M.Wind	Data Adjust Approach			Direct Approach
			AA	BB	CC	DD
1	100.	8.310	2.218	0.212	2.252	3.058
2	150.	5.903	1.575	0.245	1.618	2.363
3	200.	4.986	1.619	0.129	1.599	1.837
4	250.	3.181	1.515	0.262	1.602	2.012
5	300.	0.115	0.866	0.317	1.010	1.071
6	400.	-3.525	1.445	0.105	1.425	1.793
7	500.	-1.891	2.086	0.450	2.099	2.445
8	700.	5.989	0.259	0.564	0.566	0.683
9	850.	9.897	0.351	0.611	0.819	0.907
Overall R.M.S. of *err*			1.488	0.360	1.544	1.953

y-component of geostrophic wind

k	p(mb)	M.Wind	Data Adjust Approach			Direct Approach
			AA	BB	CC	DD
1	100.	29.601	1.630	0.120	1.672	2.353
2	150.	24.930	1.461	0.067	1.428	1.660
3	200.	23.984	1.882	0.113	1.880	2.421
4	250.	21.606	0.793	0.081	0.788	0.991
5	300.	23.153	1.223	0.076	1.197	1.479
6	400.	23.354	1.030	0.130	1.024	1.463
7	500.	18.826	0.744	0.202	0.767	1.078
8	700.	10.373	0.606	0.278	0.789	1.021
9	850.	6.899	0.144	0.149	0.236	0.413
Overall R.M.S. of *err*			1.184	0.150	1.197	1.567

Notes: AA, BB, CC and DD are the same as in Table 6.1 and Table 6.2.

Table C.5: Experiment 1: the largest absolute errors at each level

geostrophic wind over flat terrain

k	p(mb)	x-direction		y-direction	
		E	F	E	F
1	100.	1.182	1.662	3.282	5.397
2	150.	1.474	1.849	2.720	4.727
3	200.	1.218	0.866	1.890	3.788
4	250.	2.433	3.787	1.161	2.957
5	300.	3.183	3.998	1.320	1.539
6	400.	2.725	3.252	1.259	1.834
7	500.	1.442	2.107	2.279	2.478
8	700.	1.816	3.021	2.506	3.478
9	850.	0.819	1.576	0.409	1.043
10	1000.	0.177	0.030	0.152	0.009
Avg. Max. $\|err\|$		1.647	2.215	1.698	2.725

geostrophic wind over mountainous terrain

k	p(mb)	x-direction		y-direction	
		E	F	E	F
1	100.	3.462	5.049	3.294	4.697
2	150.	2.533	3.377	2.304	2.305
3	200.	2.665	3.338	2.667	3.006
4	250.	2.385	2.673	1.414	1.655
5	300.	1.924	1.942	2.357	2.068
6	400.	1.789	2.937	1.493	2.476
7	500.	2.809	4.055	1.197	2.030
8	700.	1.040	1.079	1.160	1.538
9	850.	1.521	1.422	0.419	0.661
Avg. Max. $\|err\|$		2.236	2.875	1.812	2.271

Notes: E and F are the same as in Table 6.1 and Table 6.2.

Table C.6: Experiment 2: mean absolute errors over flat terrain

x-component of geostrophic wind

k	p (mb)	M.Wind	CASE 1 Data Adjusted		CASE 2 Data Unadjusted	
			Cby	Cnt	Cby	Cnt
1	100.	28.111	0.002	0.004	0.837	0.839
2	150.	32.667	0.005	0.009	0.976	0.975
3	200.	35.331	0.012	0.025	0.597	0.588
4	250.	38.598	0.017	0.035	1.545	1.562
5	300.	39.573	0.011	0.023	2.328	2.340
6	400.	36.764	0.007	0.014	1.471	1.465
7	500.	32.466	0.012	0.025	1.136	1.141
8	700.	25.675	0.009	0.020	1.524	1.535
9	850.	21.692	0.012	0.025	0.747	0.744
10	1000.	18.214	0.015	0.029	0.014	0.026
Overall Mean $\|err\|$			0.010	0.021	1.118	1.122

y-component of geostrophic wind

k	p (mb)	M.Wind	CASE 1 Data Adjusted		CASE 2 Data Unadjusted	
			Cby	Cnt	Cby	Cnt
1	100.	19.457	0.003	0.005	2.249	2.248
2	150.	19.113	0.002	0.004	1.767	1.766
3	200.	19.084	0.002	0.003	1.352	1.351
4	250.	18.253	0.001	0.003	1.306	1.305
5	300.	18.139	0.001	0.002	0.529	0.529
6	400.	15.932	0.003	0.006	0.914	0.916
7	500.	14.471	0.002	0.004	1.303	1.304
8	700.	11.577	0.001	0.002	2.011	2.011
9	850.	12.408	0.004	0.009	0.354	0.353
10	1000.	11.476	0.004	0.007	0.004	0.007
Overall Mean $\|err\|$			0.002	0.005	1.179	1.179

Notes: Cby and Cnt are the same as in Table 6.3 and Table 6.4.

Table C.7: Experiment 2: R.M.S. of errors over flat terrain

x-component of geostrophic wind

k	p (mb)	M.Wind	CASE 1 Data Adjusted		CASE 2 Data Unadjusted	
			Cby	Cnt	Cby	Cnt
1	100.	28.111	0.003	0.006	0.930	0.931
2	150.	32.667	0.006	0.011	1.181	1.181
3	200.	35.331	0.013	0.027	0.618	0.607
4	250.	38.598	0.020	0.041	1.832	1.848
5	300.	39.573	0.012	0.026	2.452	2.464
6	400.	36.764	0.010	0.020	1.830	1.822
7	500.	32.466	0.015	0.032	1.305	1.307
8	700.	25.675	0.010	0.022	1.795	1.806
9	850.	21.692	0.014	0.031	0.917	0.909
10	1000.	18.214	0.017	0.033	0.016	0.032
Overall R.M.S. of *err*			0.013	0.027	1.462	1.465

y-component of geostrophic wind

k	p (mb)	M.Wind	CASE 1 Data Adjusted		CASE 2 Data Unadjusted	
			Cby	Cnt	Cby	Cnt
1	100.	19.457	0.003	0.006	2.723	2.723
2	150.	19.113	0.002	0.004	2.278	2.276
3	200.	19.084	0.002	0.004	1.780	1.778
4	250.	18.253	0.002	0.004	1.561	1.559
5	300.	18.139	0.002	0.003	0.677	0.676
6	400.	15.932	0.003	0.006	1.062	1.064
7	500.	14.471	0.003	0.005	1.470	1.471
8	700.	11.577	0.002	0.004	2.219	2.218
9	850.	12.408	0.005	0.011	0.490	0.486
10	1000.	11.476	0.005	0.009	0.005	0.008
Overall R.M.S. of *err*			0.003	0.006	1.656	1.655

Notes: Cby and Cnt are the same as in Table 6.3 and Table 6.4.

Table C.8: Experiment 2: mean absolute errors over mountainous terrain

x-component of geostrophic wind

k	p (mb)	M.Wind	CASE 1 Data Adjusted		CASE 2 Data Unadjusted	
			Cby	Cnt	Cby	Cnt
1	100.	8.310	0.198	0.410	2.583	2.622
2	150.	5.903	0.222	0.431	2.252	2.351
3	200.	4.986	0.107	0.228	1.474	1.324
4	250.	3.181	0.234	0.540	1.890	1.950
5	300.	0.115	0.280	1.006	0.914	1.487
6	400.	-3.525	0.089	0.859	1.620	1.670
7	500.	-1.891	0.430	0.779	2.032	2.063
8	700.	5.989	0.495	1.009	0.564	1.407
9	850.	9.897	0.453	1.485	0.778	1.693
Overall Mean $\|err\|$			0.279	0.750	1.567	1.841

y-component of geostrophic wind

k	p (mb)	M.Wind	CASE 1 Data Adjusted		CASE 2 Data Unadjusted	
			Cby	Cnt	Cby	Cnt
1	100.	29.601	0.110	0.215	1.894	1.897
2	150.	24.930	0.054	0.115	1.466	1.422
3	200.	23.984	0.080	0.152	2.208	2.235
4	250.	21.606	0.066	0.150	0.876	0.838
5	300.	23.153	0.068	0.197	1.315	1.272
6	400.	23.354	0.110	0.188	1.358	1.425
7	500.	18.826	0.185	0.270	0.856	0.937
8	700.	10.373	0.222	0.433	0.915	1.058
9	850.	6.899	0.116	0.323	0.371	0.517
Overall Mean $\|err\|$			0.112	0.227	1.251	1.289

Notes: Cby and Cnt are the same as in Table 6.3 and Table 6.4.

Table C.9: Experiment 2: R.M.S. of errors over mountainous terrain

x-component of geostrophic wind

k	p (mb)	M.Wind	CASE 1 Data Adjusted		CASE 2 Data Unadjusted	
			Cby	Cnt	Cby	Cnt
1	100.	8.310	0.212	0.451	3.058	3.129
2	150.	5.903	0.245	0.466	2.363	2.460
3	200.	4.986	0.129	0.277	1.837	1.657
4	250.	3.181	0.262	0.610	2.012	2.173
5	300.	0.115	0.317	1.286	1.071	1.925
6	400.	-3.525	0.105	1.114	1.793	2.108
7	500.	-1.891	0.450	0.934	2.445	2.911
8	700.	5.989	0.564	1.300	0.683	1.685
9	850.	9.897	0.611	1.784	0.907	2.007
Overall R.M.S. of *err*			0.360	0.962	1.953	2.165

y-component of geostrophic wind

k	p (mb)	M.Wind	CASE 1 Data Adjusted		CASE 2 Data Unadjusted	
			Cby	Cnt	Cby	Cnt
1	100.	29.601	0.120	0.238	2.353	2.388
2	150.	24.930	0.067	0.140	1.660	1.640
3	200.	23.984	0.113	0.197	2.421	2.431
4	250.	21.606	0.081	0.168	0.991	0.997
5	300.	23.153	0.076	0.227	1.479	1.490
6	400.	23.354	0.130	0.269	1.463	1.528
7	500.	18.826	0.202	0.392	1.078	1.201
8	700.	10.373	0.278	0.580	1.021	1.273
9	850.	6.899	0.149	0.389	0.413	0.635
Overall R.M.S. of *err*			0.150	0.300	1.567	1.534

Notes: Cby and Cnt are the same as in Table 6.3 and Table 6.4.

VITA

Surname: Zhang | Given Names: Qing

Place of Birth: Shanghai, China | Date of Birth: March 25, 1958

Educational Institutions Attended:

Shanghai University of Science and Technology	1978 to 1982
Shanghai University of Technology	1984 to 1987
University of Victoria	1989 to 1991

Degrees Awarded:

B.Sc.	Shanghai University of Science and Technology	1982
M.Eng.	Shanghai University of Technology	1987

PARTIAL COPYRIGHT LICENSE

I hereby grant the right to lend my M.S. Thesis to users of the University of Victoria Library, and to make single copies only for such users or in response to a request from the Library of any other university, or similar institution, on its behalf or for one of its users. I further agree that permission for extensive copying of this M.S. Thesis for scholarly purposes may be granted by me or a member of the University designated by me. It is understood that copying of this M.S. Thesis for financial gain shall not be allowed without my written permission.

Title of Thesis:

DATA ADJUSTING USING VARIATIONAL METHODS

Author: [redacted]

Qing Zhang

April 24, 1991