AN EMPIRICAL ENSEMBLE MESOSCALE FORECASTING SYSTEM

by

JOSHUA P. HACKER

B.S., The University of California, Los Angeles, 1994

A THESIS SUBMITTED IN PARTIAL FULFILLMENT OF

THE REQUIREMENTS FOR THE DEGREE OF

MASTER OF SCIENCE

in

THE FACULTY OF GRADUATE STUDIES

Atmospheric Sciences Programme

We accept this thesis as conforming
to the required standard

THE UNIVERSITY OF BRITISH COLUMBIA

February 1997

©Joshua P. Hacker, 1997

In presenting this thesis in partial fulfilment of the requirements for an advanced degree at the University of British Columbia, I agree that the Library shall make it freely available for reference and study. I further agree that permission for extensive copying of this thesis for scholarly purposes may be granted by the head of my department or by his or her representatives. It is understood that copying or publication of this thesis for financial gain shall not be allowed without my written permission.

Department of _Geography_

The University of British Columbia
Vancouver, Canada

Date _Feb, 26 1997_

ABSTRACT

Seven short-term numerical weather prediction experiments test the feasibility of an ensemble mesoscale forecasting system that is designed to minimize the effects of analysis errors in the North Pacific. Each experiment consists of a five-member ensemble (4 + control) run once per day for eight days for the period 21-31 January 1990. The Mesoscale Compressible Community (MC2) model, run at $\Delta x = 100$ km over the North Pacific Ocean and western North America, serves as the experimental platform. Empirical perturbations called the Selective Introduction of Hazardous Modes (SIHM) define six of the experiments. The seventh experiment uses perturbations that are bred within the forecast cycle, and serves as a benchmark. Standard root mean square error (RMSE) statistics and surface cross sections are used for verification.

SIHM perturbations are incipient cyclones that are added or subtracted from the initial analyses. Resolvable structures in the flow, such as low-level or stratospheric potential vorticity and jet-stream divergence, determine the locations of the perturbations. Perturbation size is set to match the most energetic wavelength in a particular latitude band, as derived from a spectral analysis. The strength of the incipient cyclones is designed to be within reasonable analysis errors published previously.

One measure of the likely success of ensemble methods is the spread between different forecast members of the ensemble. The system lacks a desirable spread in RMSE development, and the curves converge at the end of many of the forecasts because of dominance of the common imposed boundary conditions. Spreads in sea-level pressure adequate to envelope most observations exist when the model predicts a storm system well. Precipitation consistently shows the most spread along the cross section. When the model completely misses an event, spreads are negligible.

TABLE OF CONTENTS

List of Tables

List of Figures

1. Introduction

a. Background

Errors arise in numerical weather forecasts from approximation and discretization, numerics, and observational limitations. A numerical weather-prediction model integration produces one guess of the evolution of the atmosphere. It represents our incomplete understanding of the atmosphere, and it can only approximate nature.

To make a numerical forecast, a model requires initial conditions describing the state of the atmosphere, and boundary conditions throughout an integration period. Normally, the initial conditions are gridded atmospheric data, interpolated from an analysis and then conditioned. An analysis is a statistical combination of a previous model forecast and actual observations. In a typical analysis cycle, an earlier forecast is the first guess, and observations adjust the analysis toward what might be reality. Even in the most dense observational networks that operate full time, the picture of the state of the atmosphere contains ambiguity on horizontal scales smaller than the observation spacing.

Research and computing advances are combining to minimize error in numerical forecasts. More powerful computers enable higher resolution representation of dynamics, physics, and surface interaction. They also reduce round-off error. Better understanding of unresolvable physics leads to improved parameterizations. New technology allows more complete coverage of observations, given sufficient funding to deploy the instruments. Despite such improvements, uncertainty in initial conditions remains.

In some regions, initial-condition ambiguity is greater because of the lack of observations, resulting in a major source of error. The data-sparse region over the North Pacific Ocean is known as the North Pacific data void. Observations there are much sparser than over continental North America. Merchant ships and commercial aircraft provide some dynamic and thermodynamic data, as do satellite soundings and wind retrievals (Stull 1995, Fig. 12.8). Buoys are widely spaced

1

except near the coast, and rawinsonde observations are virtually nonexistent.

The North Pacific data void causes analysis errors beyond instrument uncertainty, which present difficulties for forecasters. Its effects are most severe for coastal British Columbia (BC). Not only do forecasters have to contend with important local topographic effects, but the lack of upstream observations mean that algorithms that create analyses for numerical weather prediction models must rely primarily on previous model forecasts. To compound the problem, the data void coexists with a region of frequent cyclogenesis, so forecasters often have inadequate warning of potential hazards. Many of the subtleties of developing cyclones are not resolved in an operational analysis, therefore the strength and speed of these systems is largely unknown a reasonable time before striking the coast. Satellite imagery can give evidence of cyclones already established, and provides limited information to positively affect the analysis in these regions.

b. Goals and methods

The goal of this study is to improve the accuracy of numerical weather forecasts, particularly in areas such as the North Pacific data void, which could have a direct impact on forecast skill. Improved forecasts have the potential to affect many aspects of life in BC. Coastal marine weather impacts the natural resource industries of logging and fishing. Rain and snow forecasts are important to highway ministries and hydrologic power industries, and the outdoor lifestyles of the public require accurate mountain and marine forecasts.

One approach for improving forecast accuracy, taken here, is to improve the modeling systems to compensate for the lack of data. Otherwise, the ideal method would be to increase observations within the Pacific data void and generally everywhere in the midlatitudes, which is not likely in the current fiscal environment. Another alternative currently under consideration for research by the United States Weather Research Program (USWRP) is an adaptive observation system that responds to model guidance by increasing observations density using portable platforms (i.e. aircraft) only when and where needed. This is costly and its benefits are still unproven.

2

The most obvious way to improve a modeling system is to improve the model itself, which should improve single forecasts at little or no extra computing cost. Physical parameterizations could be improved to reduce some sources of error. Maintaining state-of-the-art physics increases the accuracy of subgrid scale phenomena itself as well as their effects on the overall forecast. Representation of geophysical forcing could also be improved.

Improved grid resolution will increase the accuracy and understanding of small-scale features, including orographic forcing and mesoscale characteristics of cyclogenesis, but it cannot improve all aspects of the model forecast. As long as the model resolution is higher than the resolution of the observations, uncertainty remains in the initialization data. This uncertainty creates errors in the integrations that likely far surpass any other errors existent in the truncation, approximations to the dynamics, or even the physics.

Another alternative for improving modeling systems is to make multiple numerical forecasts. This system, called an ensemble system, attempts to minimize the problems with inaccurate analyses by beginning several integrations from slightly different initial conditions, thereby accounting for most of the possibly important errors. An important error is one which significantly changes the outcome of the forecast. Applying perturbations to the analysis can accomplish this. The group of forecasts initialized with the perturbed analyses along with the original (control) analysis comprise the ensemble. The average of the individual members is often a better forecast than any single integration (e.g. Toth and Kalnay 1993; Mureau et al. 1993; Houtekamer and Derome 1995). The result has many advantages and can provide useful information beyond a single forecast, including probability and confidence measures.

The promise of an ensemble system is even realizable for researchers using desk-top workstations to perform model integrations. Though it is necessary to make multiple runs during each forecast period, improvement is evident even in ensembles with few members and on lower resolution model grids (Toth and Kalnay 1993; Mureau et al. 1993; Houtekamer and Derome 1995).

The key to implementing a successful ensemble prediction system lies in choosing the appropriate initial perturbation. It should be representative of the errors in the analysis, it should create a spread in the forecast that is adequate to bracket the truth, and it should result in both growing and non-growing features in the flow. Knowledge of the growing features indicates what could happen as a result of analysis errors. This is the basis for the systems implemented at the National Center for Environmental Prediction (NCEP) and the European Centre for Medium Range Weather Forecasting (ECMWF) (Toth and Kalnay 1993, Buizza 1992). Alternately, the knowledge of non-growing features indicates what probably will not happen as a result of realistic analysis errors. Analysis errors are difficult to measure, and understanding of them is limited (e.g. Molteni et al. 1996). Some literature describes attempts to isolate the magnitude of errors, but only experimentation will reveal what size and shape, and in which variables, the most important errors exist.

This project investigates the feasibility of an empirical ensemble method applied to short-range forecasting, the Selective Introduction of Hazardous Modes (SIHM, Stull 1996, personal communication). Because it is empirical, the best way to determine its potential is through experimentation. To explore the potential of the SIHM method, 7 numerical experiments were completed, which are explained below. This is a natural extension to the current state of ensemble forecasting, where few methods have been attempted – and even fewer in the short range – but the possibilities may be vast.

The model platform is the Mesoscale Compressible Community (MC2) model, described below. In this context, it is assumed to be a perfect model, and no information concerning model deficiencies arises from the forecasts.

Section 2 will review the current state of ensemble research. Section 3 describes the experimental platform, which includes a description of the model and the domain. The perturbation technique is explained in section 4. The ensemble system is described in section 5,

4

and verification methods are presented in section 6. Section 7 gives the results and section 8 is a discussion of these results, including analyses of the successes and failures of this experiment. Section 9 presents a few concluding remarks.

2. Ensemble Research Evolution

The work of Ed Lorenz (1965) shows the impossibility of achieving an accurate forecast with inaccurate initial conditions, given the currently-accepted equations of motion. He further explains the unpredictability of the atmospheric system and ties it to the nonlinear processes that are known to exist. Small, unknown features in an analysis can grow throughout a model forecast period to dominate the results, leading to forecast error.

Ensemble forecasting resulted as an offshoot of Lorenz' work. Over the last couple of decades, sporadic research demonstrated the potential of this approach (e.g. Leith 1974, Epstein 1969). It has blossomed at major weather forecasting centers during the last few years, such as NCEP in Washington D.C., ECMWF in Reading, Great Britain, and the Canadian Meteorological Centre (CMC). An ensemble – or group – of forecasts, each begun from slightly different initial conditions surrounding an analysis, can yield a forecast envelope that surrounds the true evolution of the weather. By perturbing the analysis in several appropriate locations, directions, shapes, or magnitudes, the true initial state of the atmosphere may be more accurately represented as some combination of the perturbed conditions.

We distinguish between a single deterministic or categorical forecast, which is the forecast most commonly distributed by operational modeling centers, and a probabilistic forecast, which results from an ensemble system. If the correct range of initial conditions is chosen, the true evolution of the atmosphere should lie within the range of the ensemble forecasts. The spread of the ensemble provides information on the likelihood of specific events. It also produces a measure of confidence in different model forecasts and in the ensemble average.

Several methods for ensemble weather prediction are published. The simplest, called a lagged-

average forecast (LAF), uses successive forecasts in the operational cycle (Hoffman and Kalnay 1983). Forecasts that are valid at the same time, but were initiated at different times, comprise the ensemble. The advantage is that it is cost-free, because sequential forecasts are a normal part of operations. As noted by Toth and Kalnay (1993), a disadvantage of this approach is that the forecast error magnitude is not the same for each ensemble member – a result of the different forecast periods. Some researchers introduced error scaling to battle this, resulting in a scaled LAF (SLAF).

Introducing random perturbations to the analysis leads to another simple ensemble, where a forecast from each differently perturbed analysis is an ensemble member. This Monte Carlo approach does not have a scaling problem as long as the perturbations are of approximately the same magnitude, but it requires many simultaneous forecasts to be effective, which can substantially increase computational expense. Mullen (1994b) derived useful information about the importance of analysis errors by isolating the random component of the errors. His averaged forecasts for ensembles of between 8 and 16 members were often better than the control forecasts.

Houtekamer and Derome (1995) described an observation system simulation experiment (OSSE) where a Monte Carlo approach (OSSE-MC) was taken to simulate the effects of random observational errors in the data assimilation cycle. To arrive at the perturbations, they applied random perturbation to an analysis 4.0 days prior to the forecast period of interest. The perturbation vertical-error correlations mimic the radiosonde network, which was described by a vertical-error covariance matrix for the height and winds (Lönnberg and Hollingsworth 1986). Other constraints ensured a realistic error field. An analysis cycle then assimilated observations, as the integration proceeded, every 0.5 days until the initialization time. The other members of the ensemble resulted from different initial random perturbations. The OSSE also showed promising results, but as the authors indicated later (Houtekamer and Derome 1996), the most valuable information gained from this experiment is the evaluation of different physical parameterizations.

Further improvements may be possible if specific errors in the analysis can be identified before

6

the forecast period, thereby reducing the number of required forecasts. Under this paradigm, the initial conditions of the ensemble members should contain perturbations that grow to produce the largest forecast errors. Teams at ECMWF and NCEP are building upon this premise.

Researchers at ECMWF are pursuing an Optimal Perturbation method (Buizza 1992). To isolate the important analysis errors, they use the forward tangent version and the adjoint tangent version of the Integrated Forecasting System, a medium range forecast model. They run this spectral model at a triangular truncation of T21 with 19 vertical levels, in contrast to the operational run of the ECMWF model at T63. Each model run defines an operator which acts on the state vectors in phase space. The operator then determines a propagator matrix, which describes the evolution of the variables. By calculating the norm of the state vector, it is then an eigenvalue problem, with the largest eigenvalues identifying the fastest growing modes. Each ensemble member contains a perturbation that is one of the 16 fastest growing modes. For more details, refer to Buizza (1992) or Mureau et al. (1993). Their method shows much promise, though it is expensive since additional CPU time must be spent on determining the optimal perturbations.

Scientists at NCEP operating a global spectral model (T62 and 18 vertical levels) developed the Breeding of Growing Modes (BGM) method to isolate the fastest growing modes (Toth and Kalnay, 1993). It also shows much promise. In this study, the BGM provides a benchmark for SIHM because of its simplicity and its documented success. Toth and Kalnay demonstrated improvements over a control forecast using averaged forecasts from as low as 2 ensemble members. Houtekamer and Derome (1995) reported similar success with 8 member ensembles.

The method generates perturbations as follows: a) add a random perturbation to the analysis; b) integrate both the perturbed analysis and the unperturbed (control) analysis forward in time; c) subtract the control forecast from the perturbed forecast at some prescribed time in the period; and d) scale the difference. A new perturbation, equal to the scaled difference, is then added to the new analysis and b) - d) are repeated.

The BGM method is applied slightly differently here. Previous studies maintained two forecasts within each analysis cycle, allowing one from the addition of the perturbation and one from the subtraction. This study uses an initial perturbation derived by subtracting an unperturbed forecast from a verifying analysis. By both adding and subtracting this perturbation to the analysis, a two-member ensemble results. After one repetition of the procedure, again adding and subtracting, four members result. Subsequent repetitions only require addition. Thus a four-member ensemble is maintained, each within its own cycle (fig. 1). All forecast periods for determining the perturbations are 24 hours, as opposed to 6 hours used by other researchers. This is because the forecast cycle used here incorporates new data every 24 hours. Thus, features which exhibit the most error growth in a 24 hour period determine the next perturbations. Nothing suggests that a 24 hour time scale for the growth of errors is inferior to a 6 hour time scale, but errors that initially grow rapidly and then recede by 24 hours would be omitted from the perturbation. The perturbations are scaled to 2.5% of the spatial variance of the control forecast geopotential heights, yielding smaller perturbations when the geopotential height fields are flat, and larger perturbations when the atmosphere contains more vertical variations.

SIHM, as proposed, attempts to identify critical areas in the flow and perturb them regionally. It operates under the premise that to produce a viable ensemble to alleviate a specific problem such as the North Pacific data void, structures that both grow and decay provide useful information. The method requires one mechanism for the identification of the critical areas and one for generating the perturbation. Potential vorticity diagnosis provides insight to flow diagnostics crucial to determining particularly sensitive regions of the flow. Alternately, the location of resolved features such as the jet stream can indicate sensitive regions. Once identified, an incipient cyclone of some prescribed magnitude and horizontal scale is superimposed in multiple locations on the analysis, creating the ensemble. By identifying critical areas and also choosing the appropriate size of the perturbation, SIHM could be an improvement over pure Monte Carlo

methods, and it is not as expensive as the Optimal Perturbation method. More importantly, it specifically addresses the situation over the North Pacific.

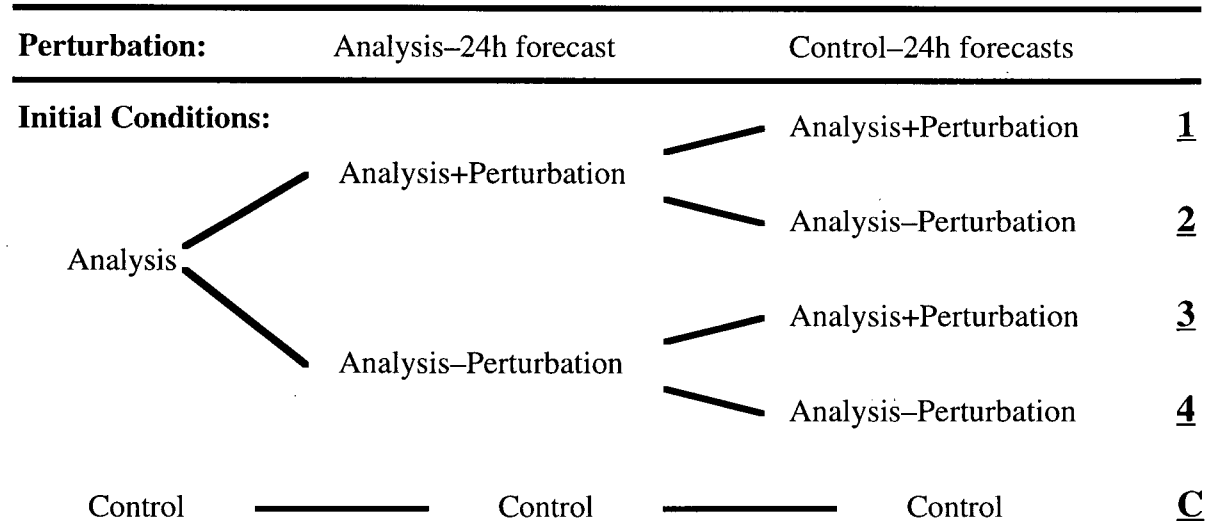| Perturbation:	Analysis–24h forecast	Control–24h forecasts	
Initial Conditions:		Analysis+Perturbation	**1**
	Analysis+Perturbation	Analysis–Perturbation	**2**
Analysis		Analysis+Perturbation	**3**
	Analysis–Perturbation	Analysis–Perturbation	**4**
Control ———— Control	———— Control		**C**

FIG. 1. Development of the four-member BGM ensemble for this project. Subsequent initial conditions result from the addition of the perturbations. All perturbations are scaled to 2.5% of the 24h forecast variance.

This study consists of two ensemble methods that are tested and compared. The BGM method is a benchmark since it is simple and well-established. Its use for short-range forecasting is still being investigated (Hamill and Colucci 1997a, b). Six variations of SIHM complete the experiment.

3. Experiment Description

a. Domain and Initialization Data

The domain is represented on a polar stereographic grid bordered approximately by 90° west longitude at the right edge and 170° west longitude at the top edge (Fig. 2). Two opposite corners are oriented north to south, stretching from near the North Pole to about 10° north latitude to the south. The development of maritime cyclones often occurs within this region.
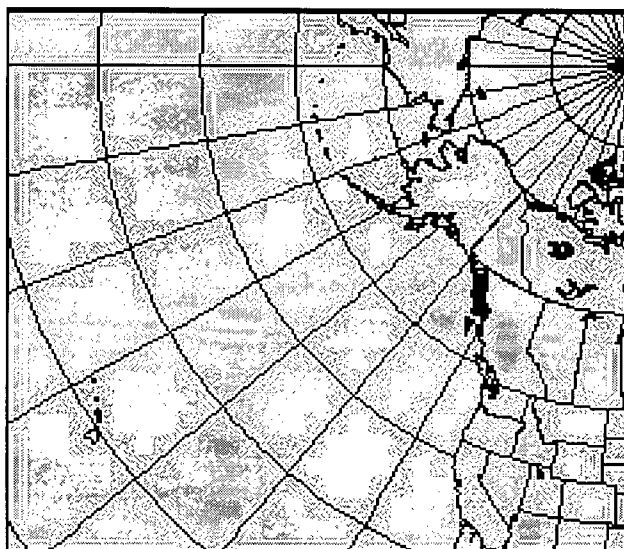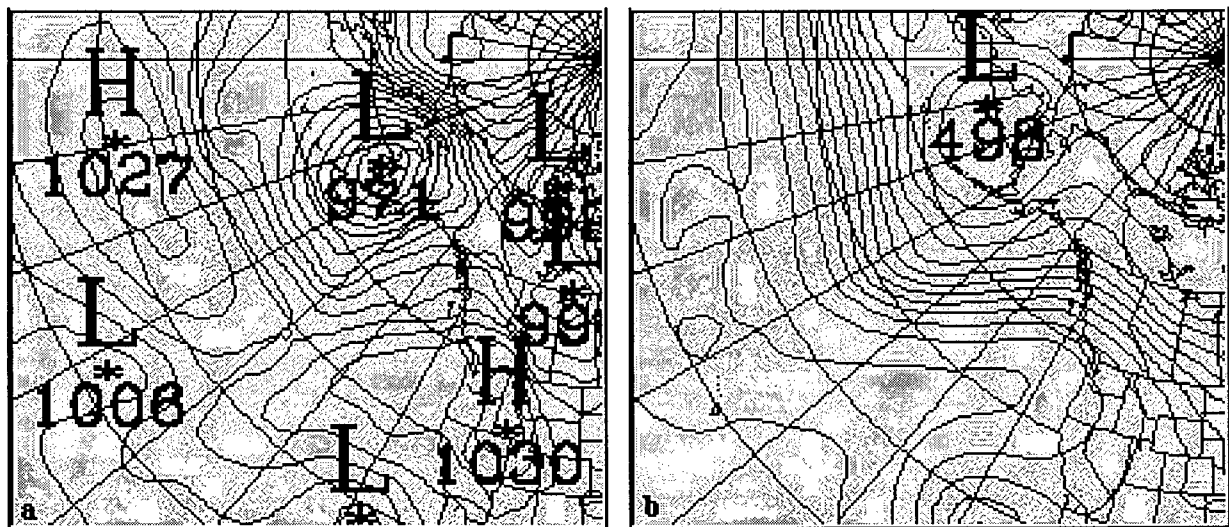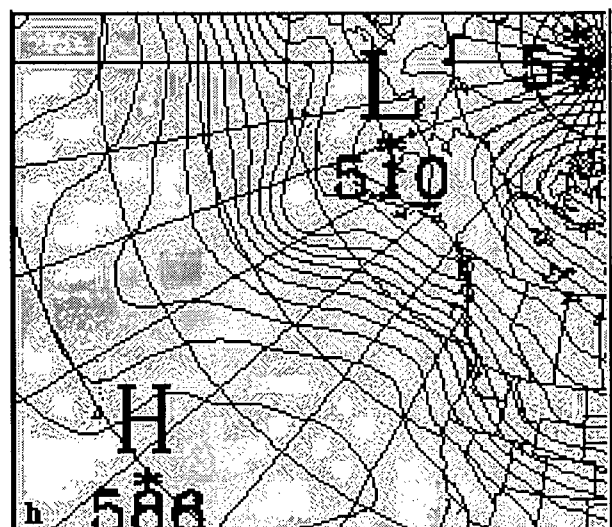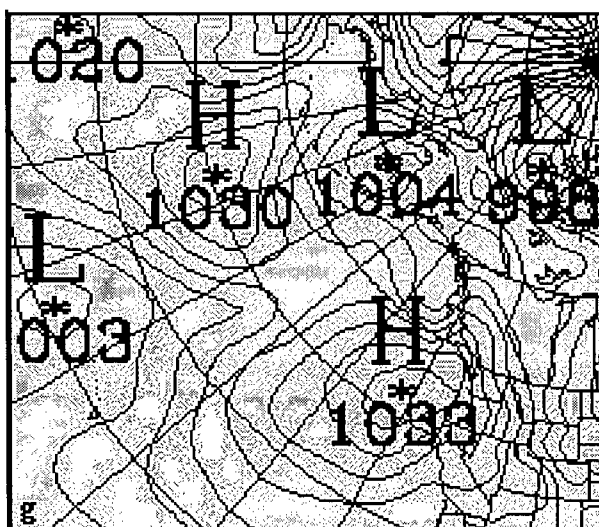
FIG 2. Experimental domain. The grid consists of 81 x 71 horizontal points. Grid points are spaced at 100 km, true at 60° North.

The domain provides adequate coverage beyond the primary area of interest, BC. An 81 x 71 point horizontal grid fills the domain, with 100 km spacing true at 60° North (Fig. 2). The upstream distance to the boundary is great enough that only features evident in the initial conditions can typically affect the Vancouver area within the forecast period of 72 hours. The x-axis of the grid is rotated 10° clockwise with respect to the Greenwich meridian, so that the west coast of North America is approximately aligned with the y-axis. Consequently the major mountain ranges, such as the Cascades and the Rockies, lie well within the domain and their effects do not interact with the edges. This is particularly important because of the steep topographic gradient from the Pacific to the Coastal Range and Cascades. Barnes et al. (1996) noted the development of gravity waves in Eta model forecasts that extended the length of North America. They apparently originated from westerly flow aloft of the Cascades. MC2 also experienced instabilities with heating and vertical velocity in the BC Coastal Mountain region when a preliminary domain was located with mountains too close to one edge.

The test period is from January 21-31 1990. It was chosen because it exhibits characteristics typical of the winter season in the North Pacific, and at least three substantial winter storms developed during the period. The 50.0 kPa height field was dominated by weak longwaves within strong zonal flow, and punctuated by several shortwaves of varying strength which affected the west coast of North America. Many of the shortwaves propagated rapidly and had significant surface low-pressure centers associated with them. Figure 3 shows the synoptic situation for 00 UTC 21-31 January 1990. The left column depicts sea -level pressure at 0.4 kPa (4 mb) intervals, and the right 50 kPa geopotential height at 60 m (6 dam) intervals. Each row is the analysis for a different day, chronologically throughout the period. The deepest surface low within the domain was analyzed at 97.1 kPa at 00 UTC 21 January (Fig. 3a), and the deepest low to strike the coast of BC was 97.8 kPa, 30 January (Fig. 3s). The discussion in section 8 focuses on the period 23-28 January because it offers a good contrast in performance of the ensemble system.

o 1020 989 1041 L 987 L H 1001 1034

p L 504 541 L 75 H 88

q L L 1041 995 978 H 1024 H 1030 H 992 H 1030 1000

r 500 533 H 109 589

s 1027 L 1048 989 H 1030 L 972

t H 535 H H 88 588

14
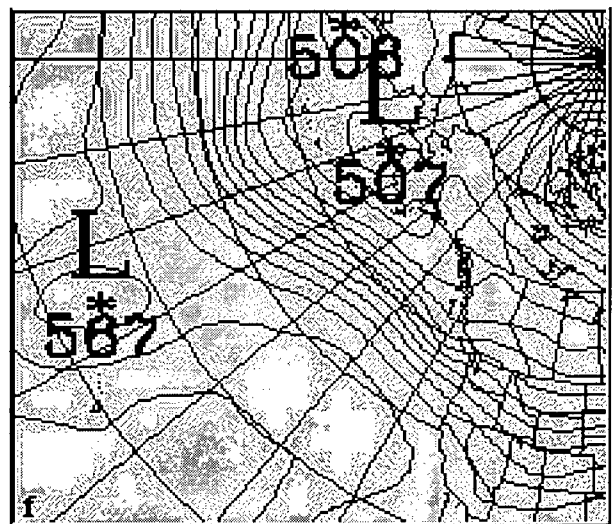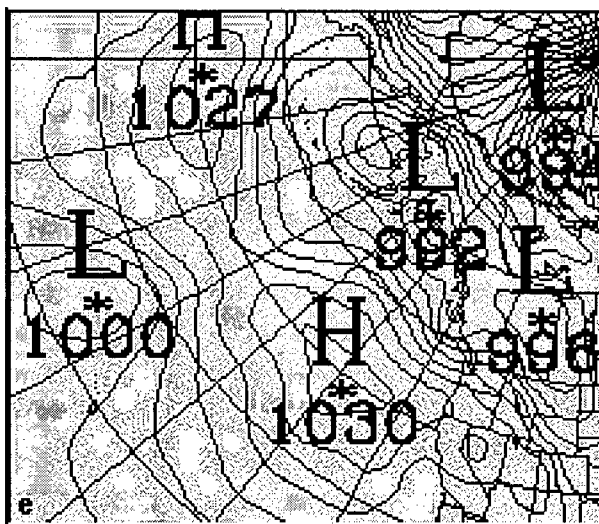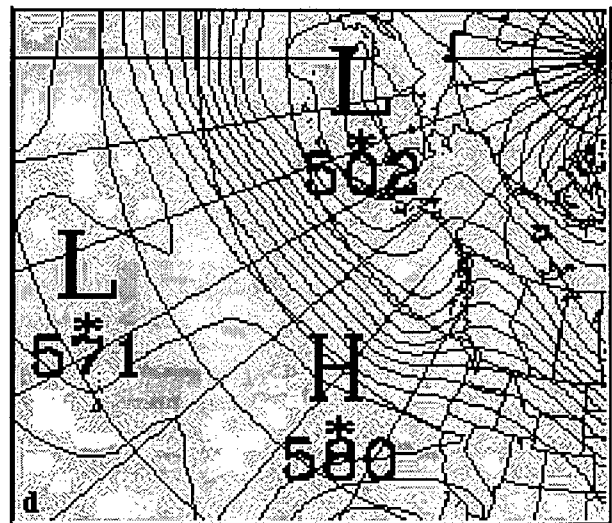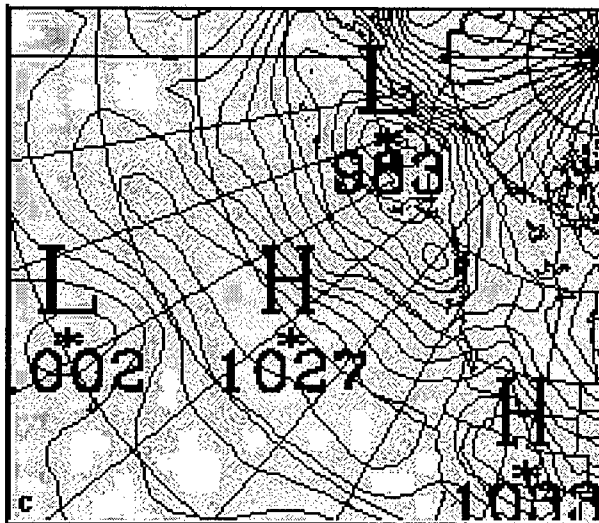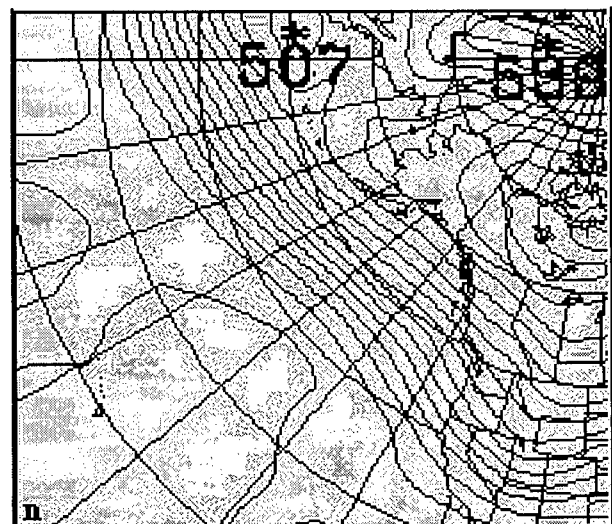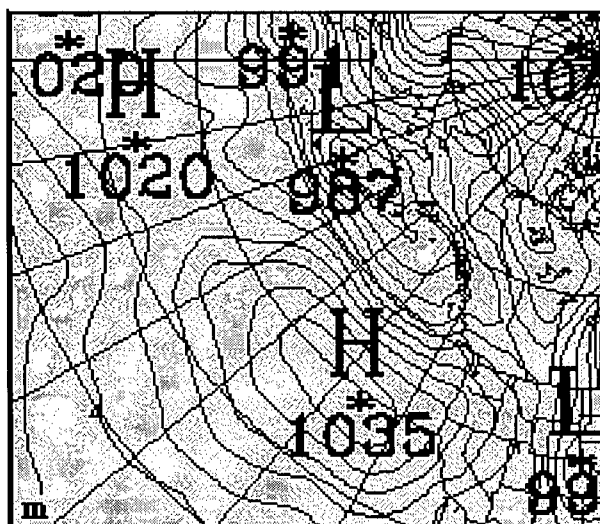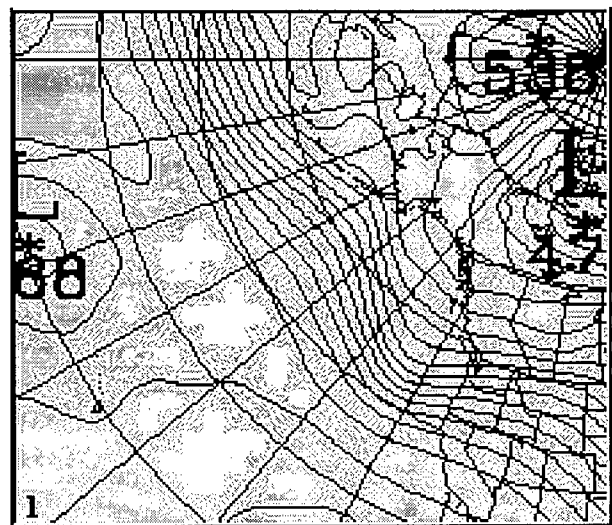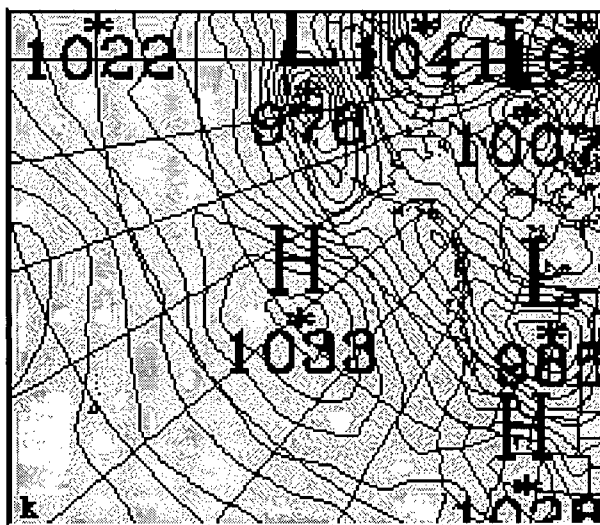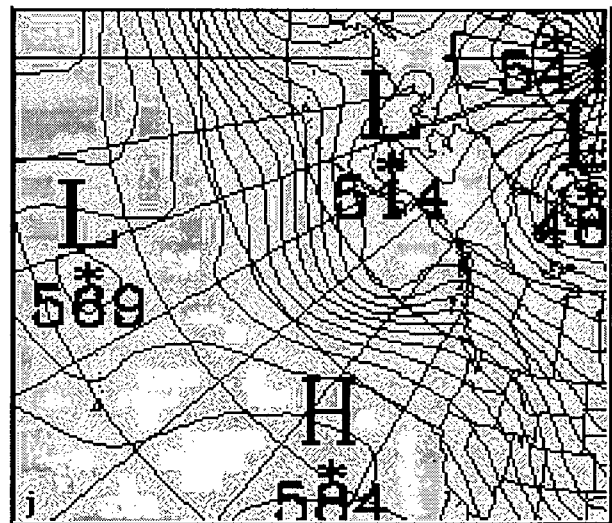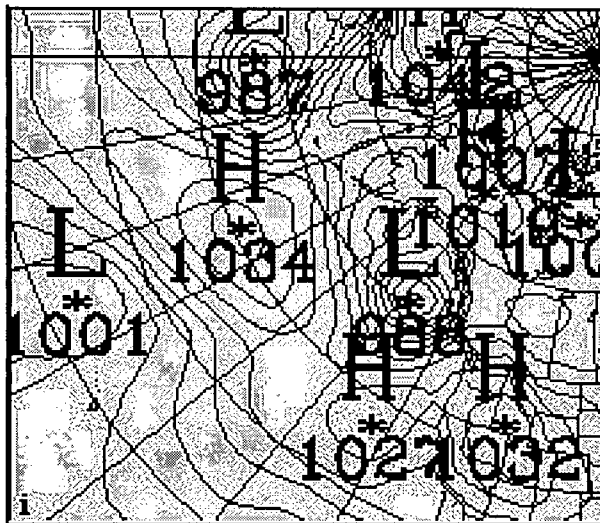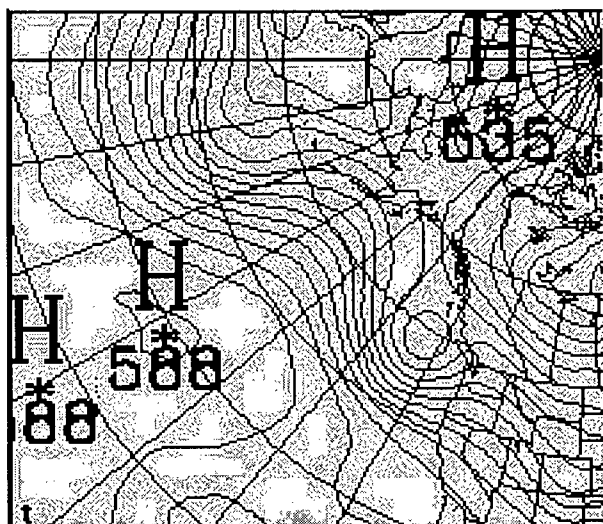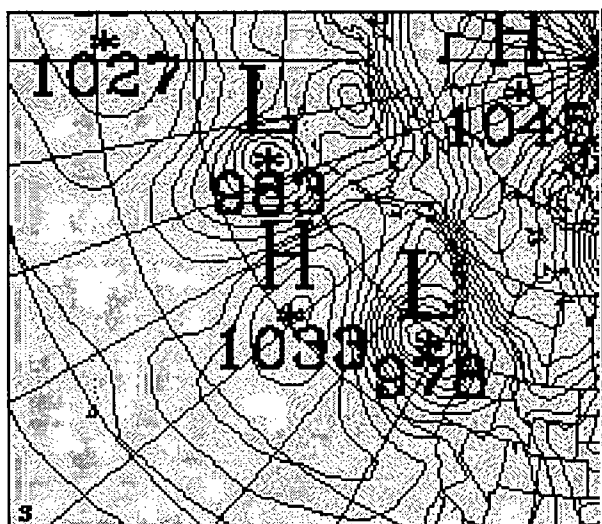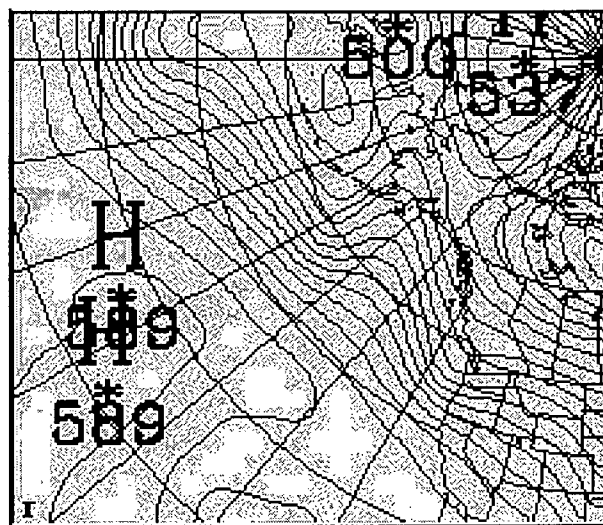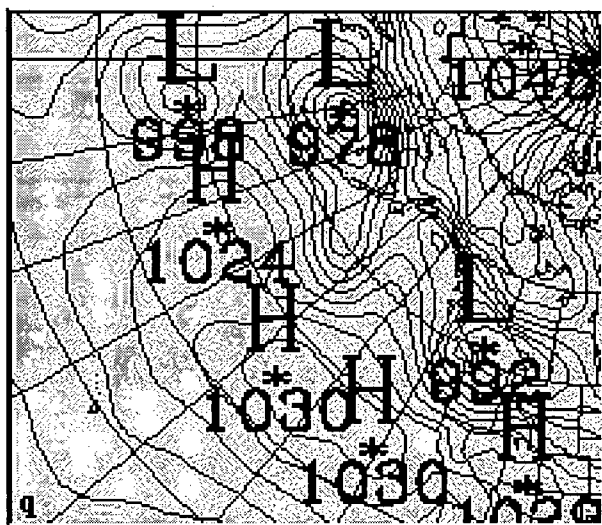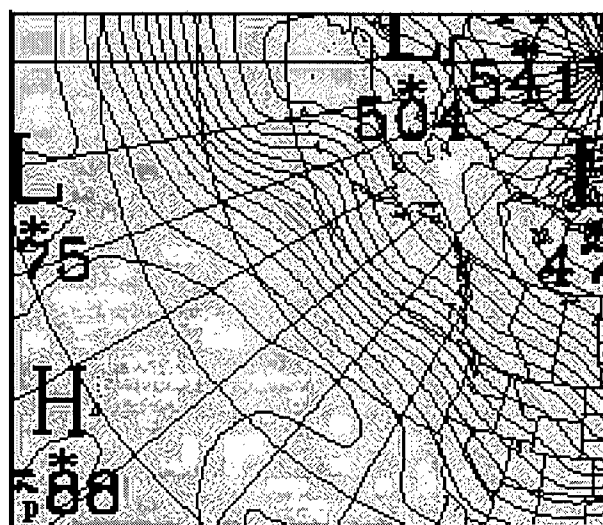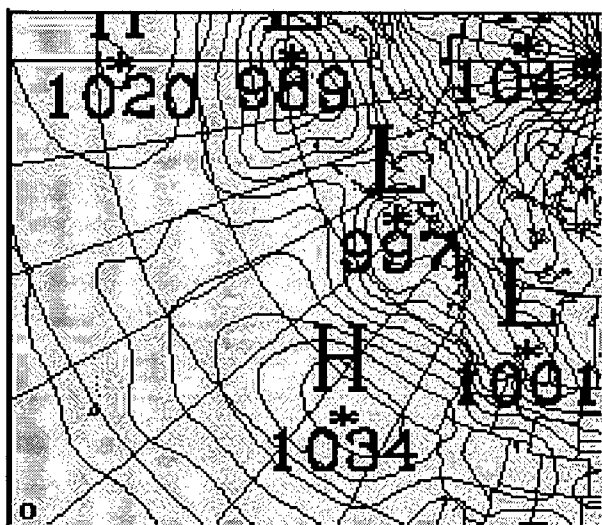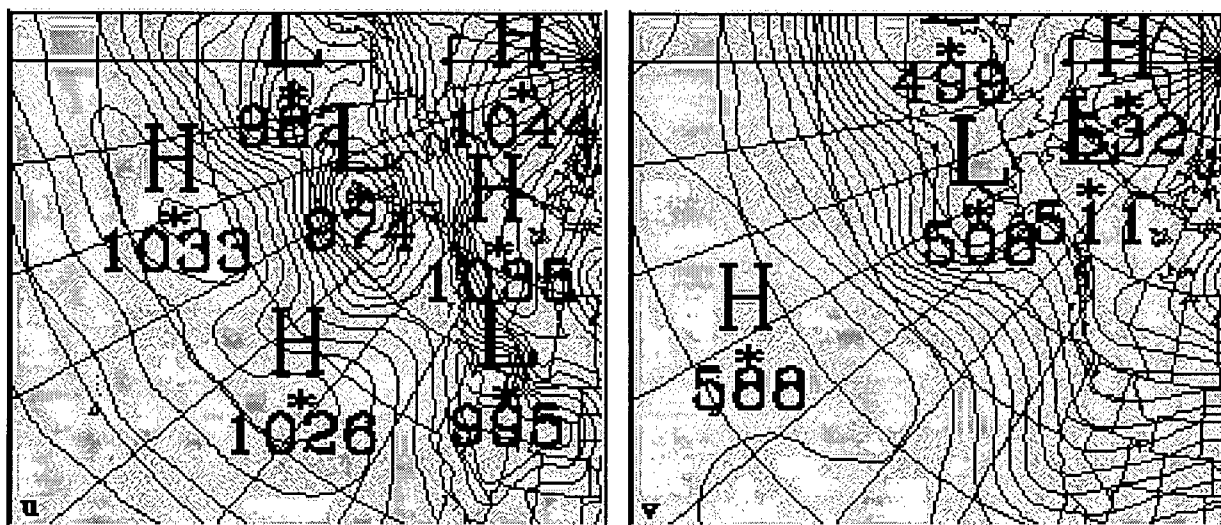
FIG. 3. Sea-level pressure (left column) and 50 kPa geopotential height (right column) analyses valid 00 UTC January 21-31 1990. Contour intervals are 0.4 kPa (4 mb) and 60 m (6 dam).

b. Mesoscale Compressible Community Model

1) MC2

The MC2 model, developed by Recherche en Prévision Numérique (RPN) and the Cooperative Centre for Research in Mesometeorology (CCRM), provides the experimental platform. It grew from André Robert's (1985) efforts to implement the semi-Lagrangian finite differencing approach. Fully Canadian, it is being used successfully by researchers worldwide. Table 1 is a summary of the configuration used for this study.

2) MODEL EQUATIONS

The dynamical system is fully compressible and nonhydrostatic. The model solves the equations on a polar stereographic grid and modified Gal-Chen vertical coordinates (σ-type) using successive corrections. It solves for a particular variable while considering only the dynamic terms of each prognostic equation, including the linear part of the gravity and elastic waves. It then corrects the prognosis by successively including the vertical physics, horizontal diffusion, nesting,

a time filter, and other numerical treatments.

3) FINITE DIFFERENCING

MC2 integrates the Euler equations with a semi-implicit, semi-Lagrangian finite differencing scheme. A second-order centered time discretization along the Lagrangian trajectory, off-centered time averages using an iterative solution of the Lagrangian displacements, and time averages of the nonlinear terms yield a system of equations for the prognostic state at time $t+\Delta t$. When decoupled, this system gives a Helmholtz equation that is solvable. To complete the time step, MC2 updates the variables with a back-substitution.

Manipulation of the Eulerian equations to arrive at a Helmholtz problem also achieves vertical and horizontal discretization. The vertical derivative is a second-order centered difference. Thermodynamic levels alternate with momentum levels throughout the domain. Any interpolations use vertical averages. Horizontal derivatives are again centered second order differenced on an Arakawa 'C' grid, accounting for the magnification and rotation of the polar stereographic projection. A decoupling procedure yields another Helmoltz equation. For more details and specific boundary considerations, see Bergeron, et al. (1994).

4) TREATMENT OF INITIAL CONDITIONS

Canadian Meteorological Center (CMC) analysis data from the study period provided initial and boundary conditions for the model runs. It came on a 1.0° cylindrical equidistant grid, 11 vertical levels from 100 to 50 kPa, at 6 hour intervals. Geophysical fields such as albedo, roughness, and soil moisture availability were from CMC monthly climatology.

5) TREATMENT OF BOUNDARY CONDITIONS

Because MC2 is a limited-area model, boundary conditions must be supplied throughout the

integration period. MC2 uses its one-way nesting algorithms to incorporate them into the model domain. CMC analyses valid every 6 hours satisfy the boundary requirements. The model linearly interpolates in time for boundary conditions between the intervals. Among the user-level configuration switches are several that control the lateral "sponge" used by the boundary and nesting algorithms (Robert and Yakimiw, 1986). To keep the boundary values from nudging the interior of this domain toward the analysis, thereby allowing perturbations to develop undisturbed within the domain, the sponge zone is set to include no points in the interior. Thus the boundary conditions directly affect only the boundaries themselves. While nesting a model domain creates discontinuity problems at the grid boundaries, Robert and Yakimiw (1986) reported that the sponge zone technique they developed minimized nesting errors by reducing or eliminating them. Thus, eliminating the sponge zone could result in spurious modes. For this project it is assumed that the amplitude of these modes is smaller than the amplitude of the waves resolved by the grid. The time filter in the model and also the horizontal grid staggering and unstaggering contribute to controlling any computational modes. Every forecast is subject to the same boundary conditions, which eliminates these errors as a degree of freedom within the experiment.

The vertical domain extends 20.0 km. A 4 km sponge at the top – corresponding to more than two grid points – absorbs vertically propagating waves to inhibit reflection at the lid.

6) PHYSICAL PARAMETERIZATIONS

For a complete description of the physics used with the MC2, consult Mailhot (1994). This study uses the model in its fully compressible mode, and includes the full RPN physics package. The Kuo Cumulus parameterization satisfies the requirements for convection. The radiation scheme executes every 1.5 hours during the forecast, which is adequate to resolve the diurnal cycle while remaining computationally inexpensive.

TABLE 1. MC2 characteristics for this project. There are choices for some of the physical parameterizations and the nesting treatment.

Feature		Model
Formulation		Non-hydrostatic
		Fully-compressible
		Non-Boussinesq
Coordinates	• *Horizontal*	Polar stereographic
	• *Vertical*	Modified Gal-Chen
Closure		1.5 Order with TKE prediction
Nesting		One way
Finite Differencing	• *Dynamics*	Semi-Lagrangian, semi-implicit time
	• *Thermodynamics*	Same as dynamics, separate diabatic processes
	• *Grid*	Arakawa 'C', vertically modified
Initial Conditions		Prescribed analysis or earlier forecast
Boundary Conditions	• *Top*	Absorbing sponge layer
	• *Lateral*	Nested with lateral sponge zone (Robért and Yakimiw 1986); set to zero interior points here
	• *Surface*	Momentum, heat, and moisture fluxes, parameterized physics
Physical Parameterizations	• *Surface layer*	Bulk transfer and PBL diffusion
	• *Soil*	Force-restore (Deardorff 1978)
	• *Surface energy*	Force-restore
	• *Convection*	Kuo (1974)
	• *Phase microphysics*	Sudqvist (1989)
	• *Radiation*	Fouquart-Bonnel (1980) - solar Garand (1983) - IR

7) TIME STEP AND COURANT NUMBER

A time step length of 15 minutes ensures both stability and accuracy. The Courant number varies between 0.4 and 0.6. The semi-Lagrangian semi-implicit scheme of MC2 remains

18

unconditionally stable for Courant numbers that even exceed 1.0, but accuracy can suffer since the scheme may improperly handle forced stationary gravity waves (Héreil and Laprise, 1996).

4. The Generation of Perturbations

a. Locating the perturbations

Locating perturbations in the appropriate places within the flow is one essential element of a successful ensemble system. The working premise of ensemble forecasting is that the analysis provides inadequate information about flow structures at scales large enough to introduce significant error into a model integration. This is precisely the reason that many ensemble forecast centers (i.e. ECMWF and NCEP) are using versions of the models themselves to determine the critical areas of the flow. They use these tools in an attempt to isolate reasonable analysis errors that will yield the largest errors in the forecast.

This method instead uses the flow of the day to isolate local regions that are sensitive to cyclone development, and introduces coherent structures that might be absent from the analysis. This may increase or decrease the error in the integration, rather than define a maximum error, but the ensemble has the potential to eliminate surprise events by bracketing most of the observations during a forecast period.

To do this under the assumption of insufficient analysis information, two approaches are taken. The first utilizes resolved structures to make a first guess of the sensitive areas of the flow. In particular, jet stream location and strength is likely to be well represented, even over the North Pacific, because of its horizontal extent and its relatively slow evolution. Jet divergence indicates locations where the flow may be sensitive to changes in the initial conditions that could spawn cyclone development.

Second, a derived quantity such as potential vorticity can also reveal structure in the atmosphere that is not apparent from the examination of any first-order variables (Hoskins et al.

1985). An isobaric potential vorticity diagnosis provides the required information, using the following equation:

$$PV = -g(f\mathbf{k} + \nabla_p \times \mathbf{v}) \cdot \nabla_p \theta$$

PV is the potential vorticity, the subscript p indicates the ∇ operator is applied isobarically, g is gravitational acceleration, f is the Coriolis parameter, and θ is potential temperature. The vectors $\mathbf{k}$ and $\mathbf{v}$ are the vertical unit vector and velocity vector, respectively. This is not the best or most useful estimate of PV (compared to isentropic PV diagnosis), but it gives the desired results.

Low-level PV can occur where a cyclone is beginning to develop, or it may exist immediately downstream from maturing cyclones. An algorithm searches the 85 kPa pressure level to determine the highest values of PV and locates perturbations there.

When PV values are represented in potential vorticity units (1 PVU=10^{-6} m^2 s^{-1} K kg^{-1}), 1.5 PVU can represent the level of the tropopause. A PV diagnosis shows locations of those stratospheric PV intrusions into the troposphere that are resolved. Knowledge of the coarse PV structure of the atmosphere cannot provide exact locations of unresolvable structures that may contribute significantly to model error, because it is derived from first-order variables that are not resolved accurately. Rather, larger areas of baroclinity can be revealed. Given that analyses everywhere and particularly in the North Pacific are inaccurate, general knowledge of the most sensitive and baroclinically unstable regions of the flow is a useful starting point for determining where to insert perturbation structures.

The PV diagnosis is on a hemispherical Mercator grid of one degree increments, for calculation simplicity. To locate the perturbation centers, an algorithm searches a specified domain above the altitude of 70 kPa for any values of potential vorticity greater than 1.5 PVU, indicating the intrusion of stratospheric PV. When it finds the first PV intrusion, it searches the same pressure

20

level again, and records the locations of as many of the highest values of PV as desired. By searching only the level where the first PV intrusion is detected, the algorithm minimizes the effects of the tropopause slope toward the northern latitudes.

The preferred domain to search is only slightly larger than the model domain, and includes areas of the North Pacific critical to cyclogenesis. For each of the locating methods, the perturbation centers must be at least 20 degrees (latitude or longitude) away from any other perturbation centers, thereby eliminating the possibility of locating more than one structure at the same PV intrusion, low-level PV region, or region of jet divergence. This also prevents overlapping of the structures to produce an unrealistically strong perturbation.

b. Perturbation Size

Perturbation size and shape are critical to its growth and role in determining the overall flow. Precisely the problem with Monte Carlo methods, many random perturbations will not grow and may be filtered or propagate out of the model domain. An examination of the actual flow in the region yields useful perturbation size information. A spectral analysis indicates the wavelengths that contain the most energy. Empirical knowledge of the most energetic wavelengths can lead to a perturbation that projects part of its energy onto the most-sensitive modes and part of its energy on less-sensitive modes.

In order to apply this algorithm universally, a spectral analysis for several different flow regimes is required. In an operational mode, a climatology could be easily built to accomplish this. Here, an average of the daily flow over the time period of interest will suffice. A Fast Fourier Transform (Press, 1986) gives information for calculating the spectral energy density of the flow along latitudes from $36°$ N to $66°$ N, every $2°$ at levels of 85 kPa and 50 kPa, which is then averaged over the ten-day period of interest (Fig. 4).

The second most energetic wavelength are used to prescribe the diameter of the cyclones that comprise the perturbations. The third was calculated in the event that the second does not affect the

flow enough. It is shown here for completeness. Results shown here are from the 50 kPa analysis, but the 85 kPa shows similar characteristics. Within this spectrum, local maxima at each latitude are evident, as the same wave number changes wavelength with latitude. The first maximum varies linearly with latitude throughout the domain, with one exception. These wavelengths, from 11 515 km at 36°N to 6235 km at 64° N, are too long to accommodate the scales of this study, so they were not included. The exception is at 66°N, where the first peak corresponds to a wavelength of 11 570 km. As the wavelengths decrease, the second and third spectral energy density maxima also display a linear relationship to latitude. Here one line can not describe the entire domain, but the characteristics of different latitude bands can be described by a few slightly different lines. This suggests different wave numbers are important within different latitude bands, and also that the peak energy may fall between discrete modes. At specific latitudes, a peak does not exist that falls on one of the lines. To fill in these gaps, wavelengths were selected that displayed more spectral energy, which meant in most cases the longer wavelength. The exception is in the far south on the second peak plot, where for simplicity the line extending from 42°N to 52°N was extrapolated.

See table 2 for a summary of the lines that describe the peak wavelengths. The intercept refers to the wavelength resulting from extrapolating the line to the equator. Figure 5 shows plots of the second and third peaks. Beyond these, noise in the spectrum makes it difficult to derive any relationship between the peaks and the wavelengths. The magnitude of the spectral density of the peaks with smaller wavelengths indicate that these are not as important, in agreement with the overall spectral energy decrease with wavelength.
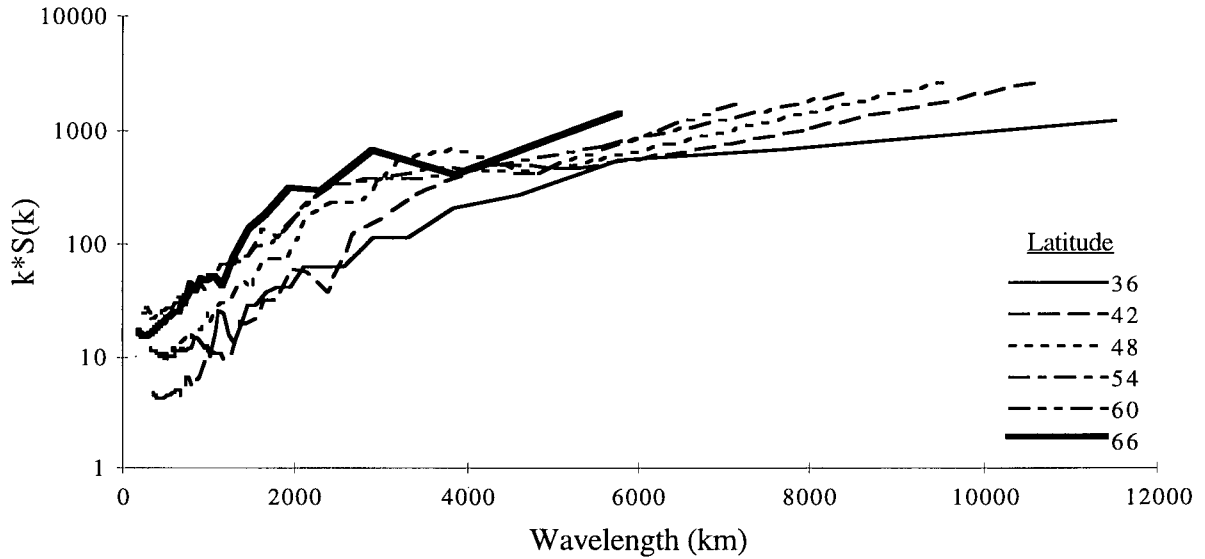
FIG. 4. Spectra analyzed at 50 kPa along the given latitudes, averaged over 21-31 January 1990. S(k) is the spectral energy density, and k is the wavenumber.

TABLE 2. Summary of the lines that describe the peak wavelengths, as shown in figure 5. The intercept is the wavelength that would lie on these lines at the equator.

Second Peak

Latitude (°)	Slope (km/°)	Intercept (km)
36-54	-72.9	7290.7
56-58	-69.5	6544.4
60-66	-110.7	10201.6

Third Peak

Latitude (°)	Slope (km/°)	Intercept (km)
36-44	-30.6	3405.6
48-48	-45.4	4562.7
50-52	-35.1	3419.7
54-60	-27.8	2616.5
62-66	-40.6	3733.0

Second Peak Wavelength Full Period Averages
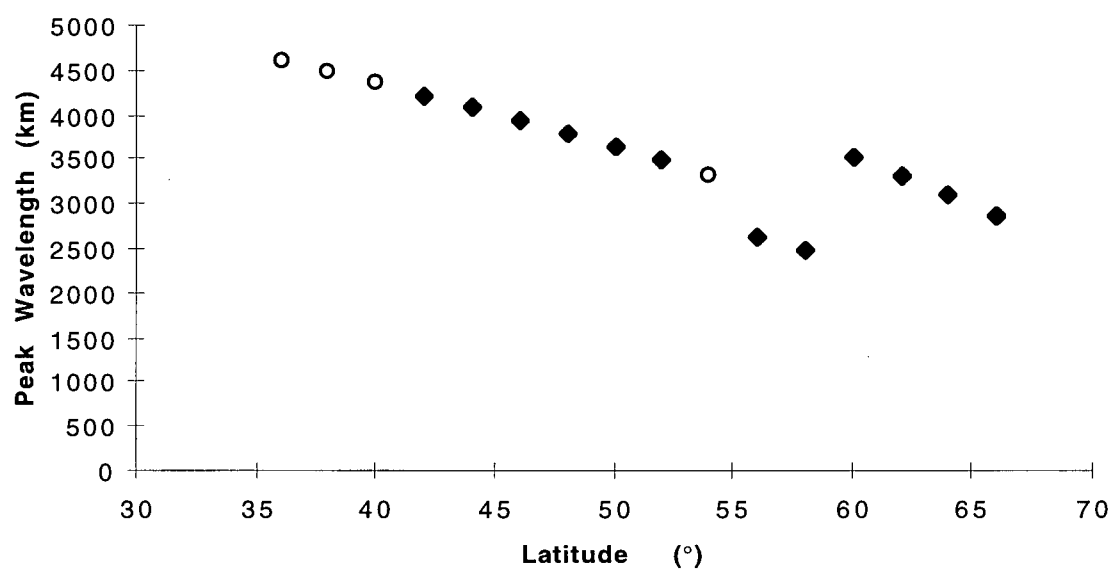
Third Peak Wavelength Full Period Averages
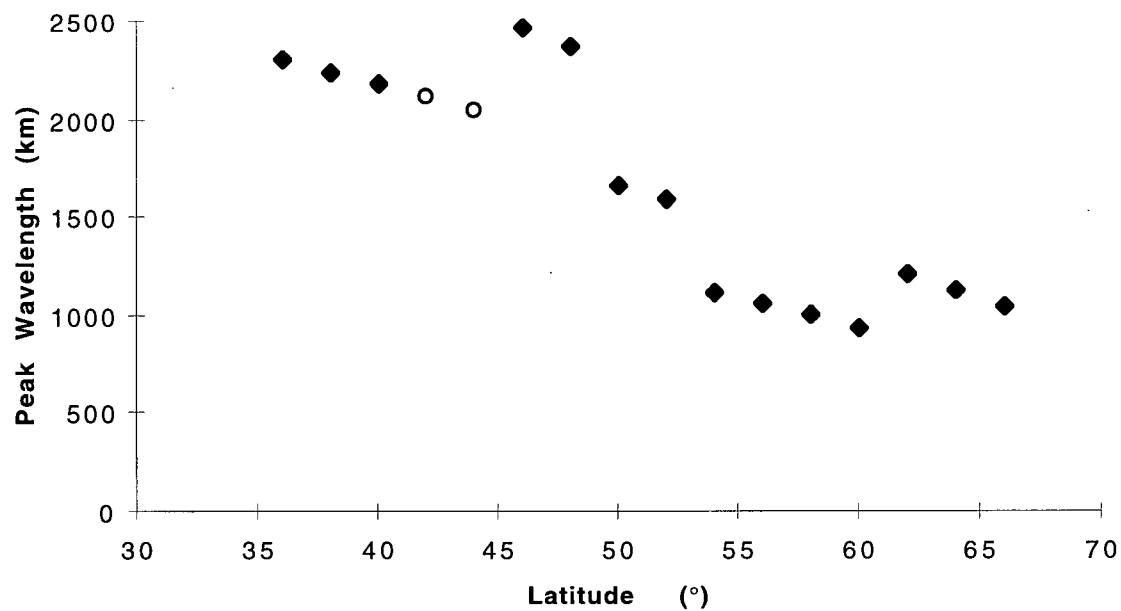
FIG. 5. Relationship between latitude and peak energy wavelengths for 50 kPa level, averaged over the 10 day study period. The solid diamonds are from real

24

data and the open circles are chosen to fill the gaps (see text for details).

c. Perturbation Shape

The shape of the perturbation is equally as important as the size. Atmospheric dynamics constrain the shape, but also provide a picture of what a realistic shape is. Geostrophic theory restricts anticyclones to weak gradients, but cyclones can exhibit stronger gradients, more isobar curvature, and a cusp at the center (Stull, 1995). Certain functions present themselves as physically reasonable descriptions of geopotential gradients that exist in a cyclone or anticyclone. For example, a high pressure region might exhibit an idealized horizontal structure like:

$$\Phi_h = \Phi_0 \left[\cos\left(\frac{\pi r}{r_0}\right) + 1 \right],$$

and a low pressure region might follow:

$$\Phi_h = \Phi_0 \left(\frac{r - r_0}{r_0}\right)^2.$$

The subscript h indicates that the equation describes the horizontal structure, Φ_0 is a prescribed central geopotential deviation, and r_0 is a prescribed perturbation radius determined empirically from spectral analysis, as explained above. Figure 6 shows plots of these functions. A cosine shape for an anticyclone was chosen because of its smooth transition to zero. The quadratic structure predicted by geostrophic theory would create a kink in the pressure pattern at the perturbation extremes.

A vertical structure further describes the perturbations within a horizontally homogeneous domain. It determines the perturbation temperature structure. If a layer is thicker than its surroundings, it is warmer. The opposite is true if a layer is thinner. Thus, to cool the entire

atmospheric column in the location of a cyclone, the perturbation magnitude must increase with height throughout the domain. An example of a function that accomplishes this is

$$\Phi_v = \Phi_h \left[1 + \sin\left(\frac{p_0 - p}{A} \right) \right].$$

The subscript v indicates vertical structure, p_0 locates the perturbation vertically, and A is the prescribed vertical extent of the perturbation. If p_0 is high enough in the domain and A is large enough, the entire column is cooler than its surroundings. Alternately, this function allows for low-level cool air below upper–level warm air if p_0 is close enough to the surface and A is set appropriately. Note that this vertical structure has the property that the maximum perturbation is not located at p_0 because it continues to increase with height throughout the depth of the perturbation.

Cyclone

$$\Phi_h = \Phi_0 \left(\frac{r - r_0}{r_0} \right)^2$$

Anticyclone

$$\Phi_h = \Phi_0 \cos\left(\frac{\pi r}{r_0} + 1 \right)$$
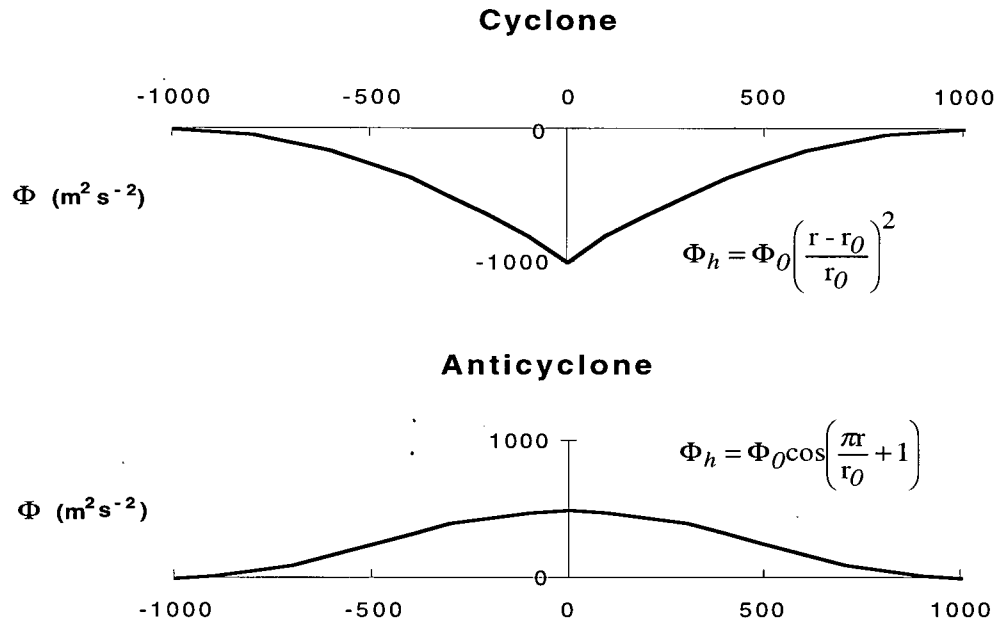
FIG. 6. Horizontal cross-sections of geopotential perturbation shape on one pressure level. The subscript h indicates that the equation describes the horizontal structure, Φ_0 is a prescribed maximum central geopotential deviation, and r_0 is a prescribed perturbation radius determined empirically.

The perturbation to geopotential height, temperature, and the stream function that results from the solution to Charney's equation (see Appendix A) is shown in figure 7. It is easy to relate these structures to the characteristics present in a well-developed extratropical cyclone (e.g. Hoskins et al. 1985). The gradient of the stream function gives the wind speed.

27
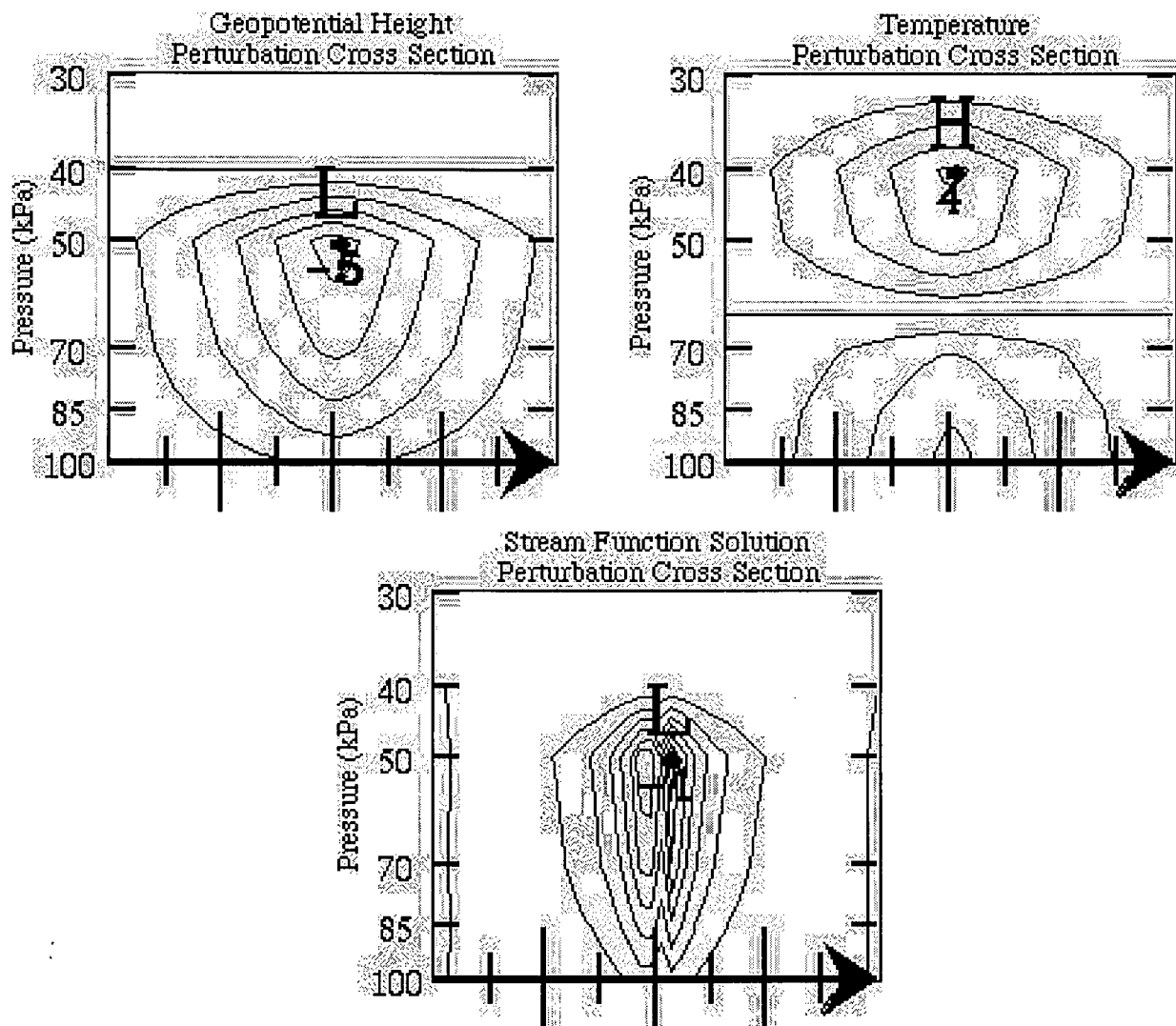
FIG. 7. Cross-sections of a perturbation placed according to p_0=85 kPa, Φ_0=30 m, and a perturbation depth of 80 kPa. The vertical structure is such that the largest magnitude of the perturbation occurs at 50 kPa.

d. Perturbation Strength

Perturbations were chosen to create a substantial spread in the ensemble members while remaining within the constraints of realistic analysis error. Reasonable analysis errors are difficult to determine because observations do not exist for comparison, and literature on the subject is sparse. Estimates have been derived artificially, or involve many assumptions (Daley and Mayer 1986).

Daley and Mayer (1986) used an OSSE technique to get an idea of analysis errors. They used a long model integration, assuming climatic stability, to create a "truth" reference state. From this run, they extracted artificial observations approximating real observations, such as soundings and satellite radiances. They also accounted for error from instruments and from not resolving the entire spectrum. By subjecting the raw input data to optimal interpolation and data assimilation schemes, they produced an artificial analysis set. Subtraction of the analysis from the "truth" gave an estimate of the analysis error.

Their results, when averaged zonally and over several forecast periods, indicate the order of magnitude of the error. Between 30° and 60° North latitude, magnitudes of temperature error were approximately 0 - 2 K; magnitudes of geopotential height error were approximately 10 m, and zonal and meridional wind components showed error of about 0.5 m/s (Daley and Mayer 1986, Fig. 3). These values are evident throughout the troposphere. The noteworthy exception is a distinct maximum in the geopotential height errors between 50 and 30 kPa of approximately 30 meters. One region of the North Pacific also had errors greater than 30 meters in the 50 kPa geopotential height field.

Studies such as Daley and Mayers' cannot provide precise error values because of the assumptions involved. Rather, they are valuable for estimating the order of the errors. Because of the uncertainty in analysis errors, a degree of freedom remains when choosing the magnitude of errors to apply in the SIHM method, while still remaining within a reasonable range. The Fourier analysis allows the potential for some projection onto growing modes, and the order of the errors ensures realism.

More empirical approaches yield realistic analysis errors for specific features. Mullen (1994) gave an account of the systematic error in Aviation Model (AVN) Spectral Statistical Interpolation (SSI), which serves as initial conditions. He reported that the SSI analysis overestimates cyclone central pressures by 0.39 kPa for surface lows below 98 kPa. It misses the location of the cyclone

centers by about 140 km for the same cyclones, and by 298 km for less intense cyclones - those with central pressures above 100 kPa.

Table 3 shows the perturbation magnitude for this study when the perturbation is introduced to the geopotential height fields. Values of the other variables result from the iterative solution to Charney's quasi-geostrophic approximation (see Appendix A). Shown are the maxima, which are above p_0 and depend on the vertical extent of the perturbation. Thus the low-level and upper-level perturbations have maximum geopotential deflections at 50 kPa and the top of the domain, respectively, resulting in significant mid-tropospheric perturbations in both cases. Z' is the maximum geopotential height deflection, which is a function of Φ_0 and p_0. T' and M' are maximum perturbation strengths of temperature and wind speed. The low perturbation wind speed is a weakness of this method and is addressed in section 8.

TABLE 3. Parameters of perturbations. Shown are maxima, based on the vertical location parameter p_0. Z' is the maximum geopotential height perturbation, and T' and M' are derived temperature and wind speed perturbations.

Experiment	p_0 (kPa)	Z' (m)	T' (K)	M' (ms^{-1})
SPVH	50	60	-2	<1
SPVL	85	50	-4	<1
LPVH	50	60	-2	<1
LPVL	85	50	-4	<1
JDH	50	60	-2	<1
JDL	85	50	-4	<1

5. The Ensemble System

The criteria for placing the perturbations and the vertical location of the perturbation characterize each of the six experiments performed. Experiments SPVH and SPVL place cyclonic perturbations according to the intrusions of stratospheric PV. LPVH and LPVL finds low-level

(85 kPa) PV anomalies, while jet stream divergence at 30 kPa dictates the placement for JDH and JDL. The parameter p_0 is set to 50 kPa for the first of each pair of experiments, while p_0 is in the lower troposphere at 85 kPa for the second. Table 4 summarizes the experiments.

TABLE 4. Experimental configuration. Information includes the criteria for locating the perturbations and the vertical location of the center of the perturbations.

Experiment	Name	Location Criteria	p_0
1	SPVH	Stratospheric PV	50 kPa
2	SPVL	Stratospheric PV	85 kPa
3	LPVH	Low-level PV anomaly	50 kPa
4	LPVL	Low-level PV anomaly	85 kPa
5	JDH	Jet divergence	50 kPa
6	JDL	Jet divergence	85 kPa

For each experiment, four members comprise the ensembles, which are initialized at every 24 hours throughout the study period (00 UTC). Members one and two result from the addition and subtraction of an incipient cyclone at only the most sensitive location, as determined by the locating criteria. Members three and four locate cyclones at the two most sensitive regions, which are separated by at least 20 degrees latitude or longitude.

6. Verification Methods

Verification of the ensemble forecasts include statistical verification and comparison with selected surface observations. Root mean square error (RMSE) provides a measure of forecast skill over the forecast period. While the sample is too small to have statistical significance, it should give a preliminary indication of the viability of the system. Geopotential height and temperature RMSE values were calculated at six-hour intervals over the critical levels of 85, 70, 50, and 25 kPa. An error spread that is too narrow means it is more likely that all of the possible weather scenarios will not lie within the ensemble bounds, while a large spread is more likely to capture all of the possible weather events while increasing the risk of a false alarm.

In the short-term forecasting mode, surface variables are the most important to verify since events such as flooding or snowfall directly affect the most people. Verification against raobs would be useful to aviation, but soundings are even more widely spread than surface observations, limiting its value. Thus, surface observations along the west coast of Canada and the U.S. provide the second means of verification. This study looks at sea-level pressure, surface temperature, and 12 hour accumulated precipitation for verification. Observations on the coast are not tainted by local effects, such as topography and roughness changes, as much as those further inland. This agrees with the 100 km grid resolution, which does not resolve such local effects well. In addition, the coastal stations form a rough cross-section that is approximately 1400 km in length, across which storm tracks often contain a dominant normal component. This allows insight to the evolution of the parameters, along this cross-section, throughout the forecast period. See table 5 for the location of these observations.

TABLE 5. Location of the weather observation stations used for forecast verification, listed from north to south.

Station	Location	West Latitude	North Longitude	Elevation (m)
WFV	Cape St. James, BC	51.93	131.02	91.16
WFG	Sartine Is., BC	50.82	128.90	100.91
WRU	Solander Is., BC	50.12	127.93	88.11
WEB	Estevan Pt., BC	49.38	126.53	7.01
UIL	Quillayute, WA	47.95	124.55	62.50
AST	Astoria, OR	46.15	123.88	6.71
OTH	North Bend, OR	43.42	124.25	5.18
ACV	Arcata, CA	40.98	124.10	68.60

The only modifications to the observed data is a conversion to similar units, and the only modification to model forecast data is an altitude-based temperature correction. The topographic elevation of the model is subtracted from the listed elevation of each weather station. The difference is multiplied by the dry adiabatic lapse rate if the forecast relative humidity is less than 95%, and by an approximate moist adiabatic lapse rate (5.8 °C/km) if the forecast relative humidity is greater than 95%.

Once corrected for elevation, a weighted average of grid points within 115 km of the station gives a value for each parameter at the location and elevation of the station. The Cressman method (Haltiner and Williams 1980) determines the weights applied to each gridpoint.

Additionally, model biases for this group of weather observations were calculated. The biases are averaged over the control (unperturbed) forecasts during the 10-day period and over the stations chosen for verification. This results in a bias value for each forecast length from zero to 72 hours at six hour intervals. These numbers are valid only for this particular time period and averaged over these particular observations, and can in no way be generalized to the model in all scenarios. The model biases appear in table 6.

TABLE 6. Model biases calculated for the verification period and the eight observation stations chosen for verification. Precipitation is accumulated over the 12 hours prior to measurement. ** indicates values that could not be determined because of insufficient integration time.

Forecast Hour	Precipitation (m)	Temperature (°C)	Sea-Level Pressure (kPa)
0	**	0.0953	-0.04308
6	**	0.7824	-0.05630
12	-0.0019	1.2026	-0.04497
18	-0.0016	0.6095	-0.06337
24	-0.0017	0.6663	-0.00501
30	-0.0012	1.4912	-0.06660
36	-0.0020	1.7291	0.01034
42	-0.0020	1.1104	0.05268
48	-0.0020	1.2028	1.04772
54	-0.0008	2.0246	0.00220
60	-0.0012	2.4731	0.00952
66	-0.0014	1.5903	-0.04211
72	-0.0005	1.6296	-0.06179

Obvious characteristics of the biases are the consistent under-prediction of accumulated precipitation and overprediction of temperature. At this grid scale, precipitation and surface temperature are determined primarily by the physics. It is not clear, however, whether the majority of the bias is from the physics or from the dynamics. Further study and bias calculations over an extended period may reveal the largest source of bias.

It was decided that a correction for the biases should not be applied here for two reasons. First, since the bias is calculated only for this period, it may improve the results unfairly. Second, any improvement would be small and perhaps negligible, given the small values.

7. Results

When MC2 utilized the dynamic initialization option, RMSE error increased compared to the analyses throughout the forecast period. To provide a more difficult standard, all integrations were performed without it. This reduces the value of the first few hours of the forecast, but that is unimportant here.

a. Bred Modes

As explained, the BGM method serves as the benchmark. While this method demonstrates positive results when applied to global, medium-range forecasting (3-10 days), it has not been as successful in limited area, short-range (0-48 hours) applications. (Hamill and Colucci 1997a, b). The breeding method here also demonstrates limited success. Two forecast days are required to establish the four-member ensemble, so results are available for the six days of 23-28 January.

The first thing to notice is that on several days, the RMSE evolves as two pairs of the ensemble members (Fig. 8a, b). The couples change for each forecast, and sometimes couples switch later in the forecast, showing that similar initial error characteristics do not necessarily lead to similar forecast errors. The pairs often develop along the high and low edges of the error spectrum, with the control and the average error curves lying near the middle. This result is similar to that of Hamill and Colucci (1997a), where forecasts at the extreme edges of the distribution are much more heavily populated than the rest. The spread between the BGM extremes is generally larger than for any of the other experiments – a consequence if its nonlocal nature.
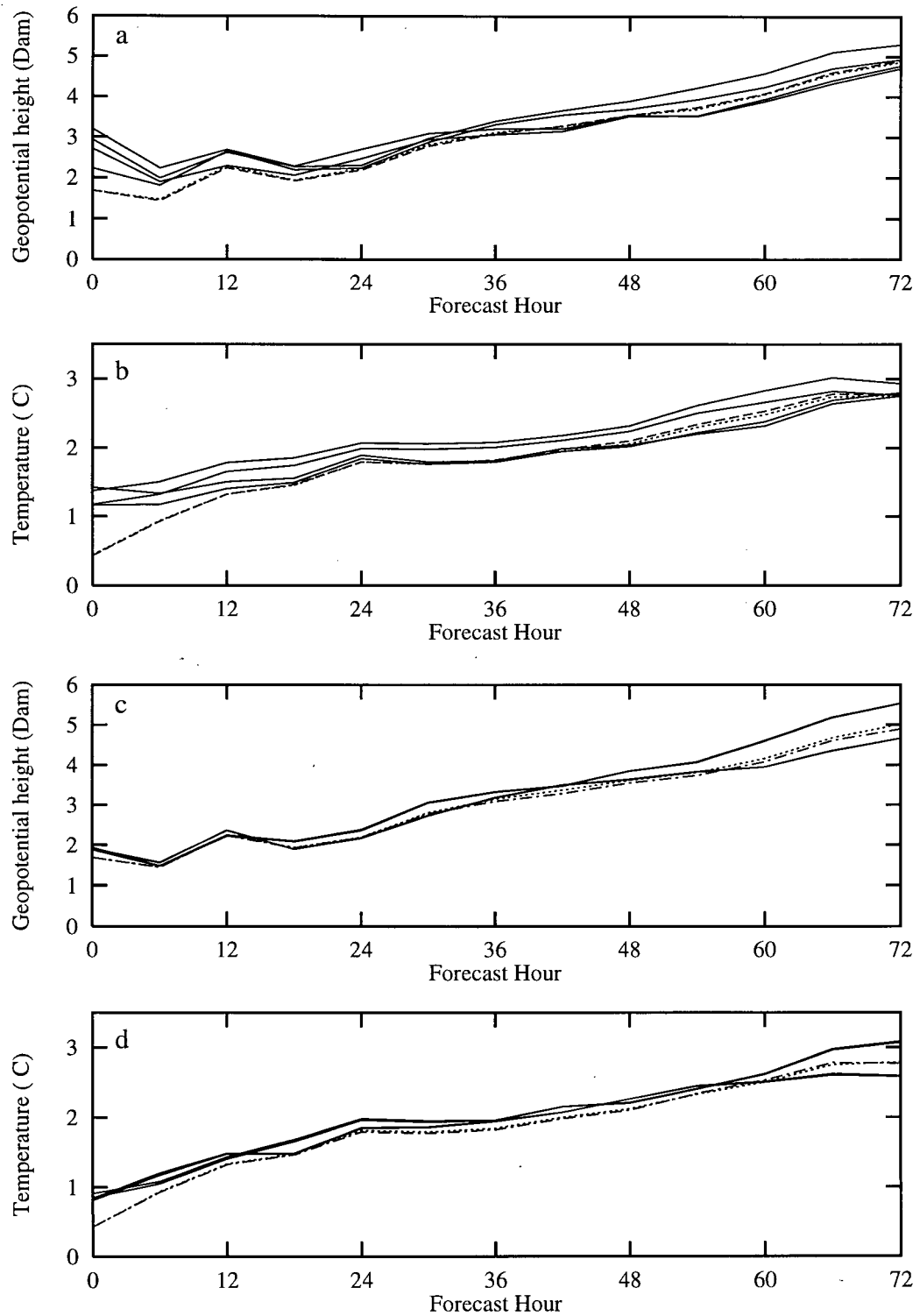
FIG. 8. RMSE development for forecasts initialized 00 UTC 23 January 1990.
Dotted line is the control forecast, dashed line is the ensemble-averaged forecast,
and solid lines represent the four perturbed ensemble members. (a) BGM and (c)
JDL geopotential height; (b) BGM and (d) JDL temperature.

36

At the beginning of each forecast, the error of the ensemble average and the error of the control forecast are nearly identical. They then evolve along similar paths near the middle of the ensemble. On 27 and 28 January, however, the ensemble-averaged forecast is slightly better than all of the other forecasts, including the control. Toth and Kalnay (1993) reported that four forecast cycles were necessary before the ensemble began to outperform the control forecast, which is consistent with these results. More experimentation could determine whether this is a consequence of the weather on these two days or a characteristic of the bred modes.

When comparing the forecast data to a cross-section of observations, we look for answers to two simple questions. Is the spread in the forecasts such that all of the observations lie within its envelope? Also, is the average of the perturbed forecasts and the control forecast closer to the observations than the control forecast alone?

To address the first, the spread in forecasts evident in this application of bred modes does not surround all of the observations. Though the RMSE spread is larger than the other experiments, the spread of surface verification variables along the cross-section is similar. The spread generally extends further, however, given the nonlocal nature of these perturbations. The largest spread in the 12 hour precipitation occurs for the 60 hour forecast initialized 23 January, exceeding 1 cm. The largest sea-level pressure spread, approximately 1.2 kPa, coincides with this event, occurring at 54 hours into the forecast period. The temperature spread from the coincident 60 hour forecast is about 2.5 °C. Higher temperature spreads occur elsewhere, but only at the initialization time, and the perturbation itself is directly responsible for this. Later in the forecasts, the temperature cross-sections are characterized by little spread, not exceeding 2 °C.

A deep low-pressure center passes through the cross-section very near the time of maximum spread in precipitation and sea-level pressure. This provides some indication that even in the first complete breeding cycle, the perturbation targeted a sensitive area. The large spread in the precipitation shows the sensitivity of the microphysics and the convective parameterization, given

37

their nonlinear nature. Also, the precipitation occasionally peaks in different locations along the cross-section.

Conversely, another low-pressure center that tracks through the cross-section with a central low of approximately 98.6 kPa, at 18 UTC 28 January, is barely evident at the 42 hour forecast from 27 January or the 66 hour forecast from 26 January. Moreover, the spread of the forecasts in the vicinity of the low does not exceed 0.3 kPa. The 18 hour forecast from 28 January, which corresponds to this event, captures the low-pressure center better, but the spread in the forecasts over the low is still less than 0.4 kPa. Rather, the spread is greater in the high pressure areas of the cross-section, nearing 1.0 kPa.

If the breeding method were entirely successful here, we would expect the ensemble-averaged forecast to be closer to the observations later in the forecast, but the two do not deviate significantly from one another for the sea-level pressure and temperature. As expected, the ensemble-averaged forecast and the control forecast are essentially identical at zero hours. The precipitation forecasts do show more deviation, given the larger spread and the different locations of the peaks, but no clear evidence suggests that the average is better than the control forecast alone.

b. Experiment SPVH

One or two weak incipient cyclones placed according to the largest stratospheric PV intrusions define the perturbation in experiment SPVH. The parameter p_0 is 50 kPa and therefore defines a cold column. The spread in the RMSE is the smallest of any experiment. Ensemble-averaged RMSE curves demonstrate negligible deviation from the control forecast RMSE curve. The perturbations placed at these locations are doing little to affect the overall flow.

Cross-sectional verification also shows the lack of spread between the different ensemble members. The maximum spread in 12 hour precipitation accumulation does not exceed 4 mm. Similarly, sea-level pressure and temperature spreads are always less than 0.3 kPa and 3 °C,

respectively. They are more often lower than that. The perturbations are not sustained within the flow, and they do not induce enough of an increase in the instability of the flow to spawn differences elsewhere.

This perturbation, placed where the tropopause is already intruding, further deflects the tropopause. This may result in a cyclone structure that is unsustainable by the surrounding atmosphere, and begins to deteriorate immediately.

c. Experiment SPVL

The RMSE evolution of these runs, perturbed with $p_0 = 85$ kPa according to intrusions of stratospheric PV, also lacks a desirable spread. For each forecast period, the behavior of the average of the ensemble is nearly identical to that of the control forecast. As for the other experiments, the errors of the individual ensemble members usually form pairs, with one pair at each extreme. The error development suggests that the perturbed forecasts are not developing their own features that deviate substantially from the control forecast, but rather that the perturbations are maintaining their relative magnitudes.

Cross-sectional verification reveals that the individual members of the ensemble never spread satisfactorily. Even spreads in the precipitation, which is the most sensitive given its nonlinear nature, do not exceed a few millimeters for the 12 hour period. Furthermore, the sea-level pressure ensemble does not spread near the low pressure centers, indicating that the resolved tropopause folds at this scale are not collocated with the low-pressure centers, which intuition tells us are important areas to forecast.

d. Experiment LPVH

The spread in RMSE for this experiment, which locates upper-level perturbations where high PV exists in the lower troposphere, is generally greater than in experiment SPVH, but it is still undesirably narrow. Spreads for 26 January are negligible. The characteristic evolution in pairs is

still evident. Also, the ensemble-averaged forecast evolves nearly identically to the control forecast.

Like the other experiments, the largest spread in the surface verification variables occurs in the latter stages of the forecast initialized on 23 January. Sea-level pressure spreads to a maximum of near 0.8 kPa for both the 54 hour and the 60 hour forecast, and decreases to about 0.6 kPa by hour 66. Twelve hour precipitation accumulation varies the most at the 72 hour forecast, extending from 24 to 29 mm – a smaller spread than what is forecast in other experiments. The location of the peaks do not vary as much as some other experiments. All of the forecasts peak near station UIL, but the observed peak is closer to station AST. Thus, the spread in the ensemble could not accommodate the missing information in the analysis.

Relatively little spread is evident elsewhere in these results, suggesting that the cold column perturbation does not interact favorably with the low-level PV maxima.

e. Experiment LPVL

RMSE for the ensembles created by placing low-level perturbations according to low-level PV anomalies again exhibit a lack of spread throughout the forecast period. Also, while these statistics do not show an increase in error for the ensemble-averaged forecast, neither do they show a decrease in error. For each forecast, the ensemble-average forecast and the control forecast display nearly identical error values and trends.

The largest spread in geopotential height and temperature errors occurred on the first forecast day, 21 January. The spread between members of the RMSE of geopotential height and temperature reached 6.0 m and 0.31 °C, respectively, at 48 hours.
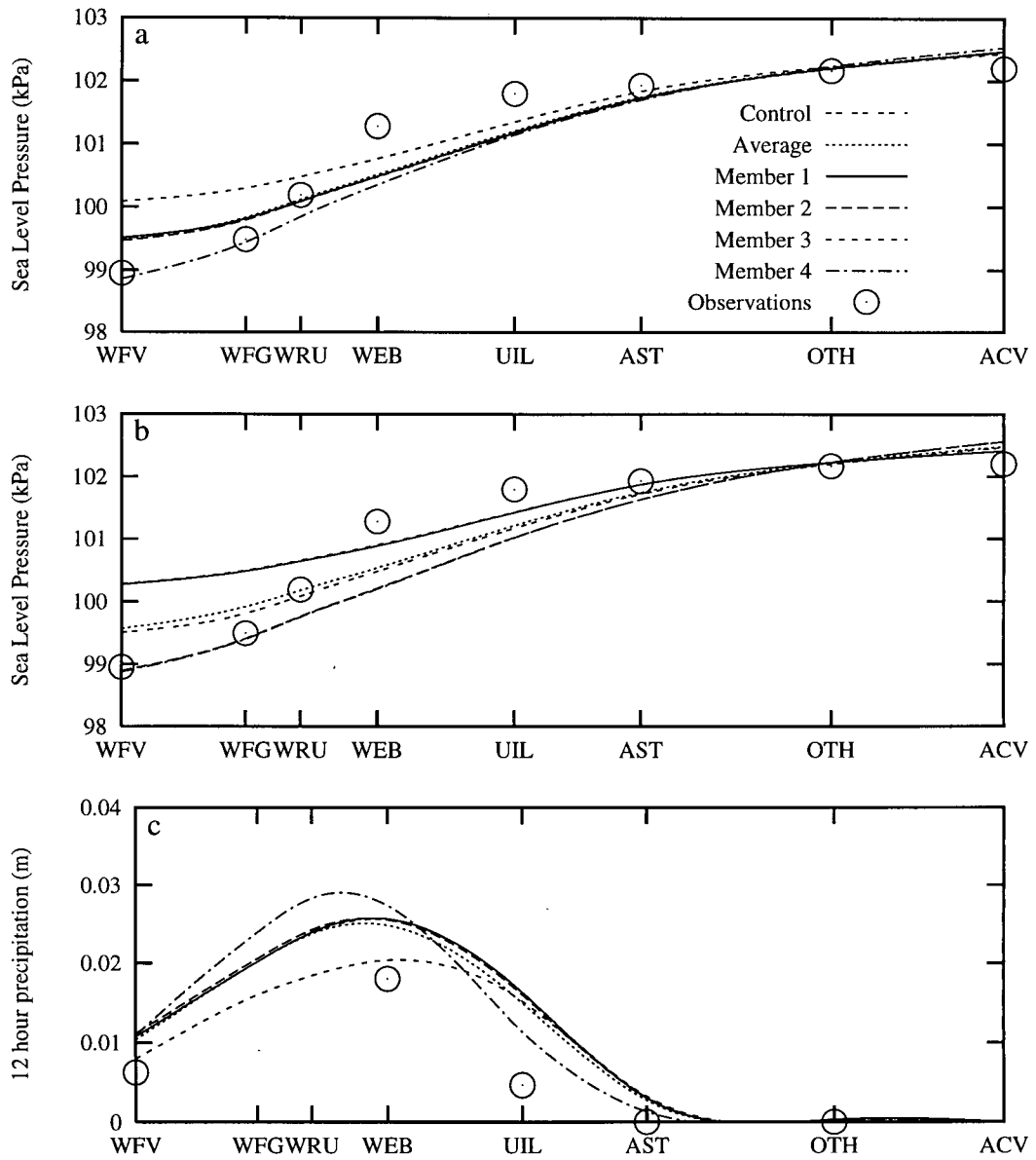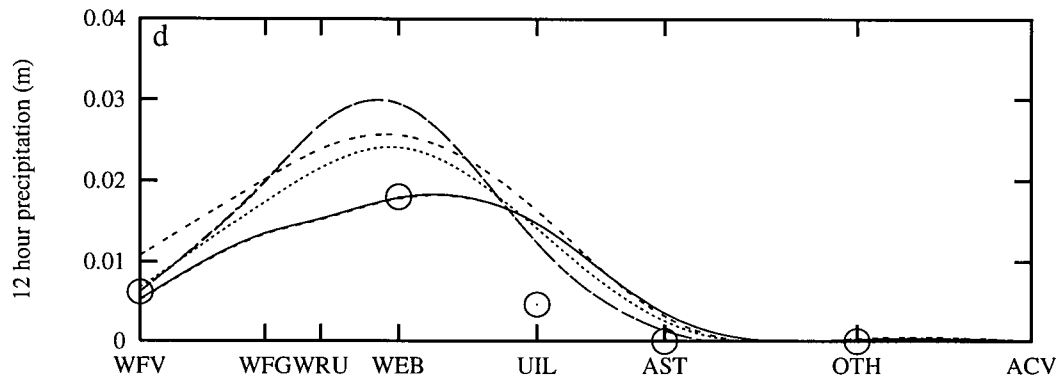
FIG. 9. Forecasts initialized 00 UTC 23 January 1990. (a) LPVL and (b) JDL sea-level pressure valid at 54 hours; (c) LPVL and (d) JDL 12 hour accumulated precipitation valid at 60 hours. Station identifiers are represented along the x-axis. The legend shown in (a) is valid for the entire figure.

Examining the verification with real observations, it is clear that the largest spread in the ensemble occurs in the regions of lowest pressure. This implies first that the location of the perturbations are in the regions of developing or developed low-pressure centers, and second that the perturbations do not die during the forecast period.

The first consequence is no surprise since the regions in the analysis with the lowest pressure near the surface – and thereby the thinnest regions – will also perhaps have the most relative vorticity in the wind field. The PV diagnosis will take both of these factors into account and show a region of high PV at low levels in the atmosphere.

For example, the observations valid 06 UTC on 25 January indicate a relatively deep low pressure region centered near the north end of the observation cross-section. The minimum observed low was approximately 98.9 kPa at station WFV. The ensemble of the forecasts initialized on 23 January led to a 54 hour forecast with a spread between ensemble members reaching 1.2 kPa (Fig. 9a), also at the northern end of the observational cross-section. As a result, the spread is enough to cover the observations at the northernmost three stations. Six hours later, at 12 UTC, the observation at WFV is missing, but the steeper gradient between WFG and WRU

42

suggests that the central pressure may have been even lower. Indeed, the forecast central pressure lowers, but the spread of the ensemble members decreases to below 1.0 kPa as the perturbed regions advect eastward beyond the cross-section. The spread of the ensemble members for the 30 hour forecast initialized a day later is only about 0.4 kPa, which is not enough to surround the observations. In both cases, the spread diminishes southward toward a region of high pressure.

The later forecast of this event appears to do a better job of matching the observations throughout the cross-sections, with the ensemble-average fitting the magnitude and gradient of the observed pressure more accurately, but the spread is smaller. By 36 hours, when the lowest pressure is suggested in the observations, the spread reaches about 0.5 kPa. Twelve hours later, the spread has not changed significantly, but all of the ensemble members retain a central pressure that is far below the observations, which indicate that this particular low pressure system moved on-shore quicker than forecasted by the model. The forecast pressures are significantly lower than the observed pressures until near the end of the forecast period. Notably the longer forecast from the previous day raises pressures more rapidly to a closer match with the observations, but it is still too low.

The precipitation forecasts in this experiment do display significant spread for some days. For example, the spread of the forecasts initialized on 23 January reaches about 1 cm in 12 hour accumulated precipitation at 60 hours (Fig. 9c). Additionally, the latitude of the peaks deviate from each other, with two members more closely matching the observations and two of the members deviating further. The average appears slightly closer to the observations than the control.

f. Experiment JDH

This experiment consisted of upper-level perturbations placed according to areas of maximum jet-stream divergence. Never does the average forecast deviate noticeably from the unperturbed forecast, and the narrow spreads are similar to the other experiments.

When examining sea-level pressure, this set of ensembles has a slightly smaller spread in

forecasts than in experiment LPVL. Forecasts from 23 January show the most spread with the passage of a deep low-pressure center across the line of observations. The lowest observed pressure along the cross-section is approximately 98.9 kPa, at station WFV, concurrent with the 54 hour forecast, but a data point is missing where a lower pressure may have been observed concurrent with the 60 hour forecast. By the 60 hour forecast, the spread has narrowed to less than 1.0 kPa, indicating that the perturbation has moved on shore. By the 66 hour forecast, the spread has narrowed further to less than 0.5 kPa, and the corresponding observed pressures have risen to a central low of around 99.9 kPa.

Not surprisingly given the 100 km resolution of the model grid, the north-south pressure gradient of this vigorous low pressure center is not accurately captured by the forecasts. But the three northernmost stations reported the three lowest pressures, which all fell within the 1.0 kPa spread of the forecasts.

The forecasts for the deepest low-pressure centers also produced the most precipitation and the warmest temperatures to the south. The precipitation forecasts exhibit a large spread nearing 1 cm for each 12 hour accumulation measured at 60, 66, and 72 hours. This is the largest spread of experiment JDH. Similarly, the temperature cross-section shows a relatively large spread of 2 °C near the southern end.

g. *Experiment JDL*

This experiment was made up of low-level perturbations placed according to high values of divergence aloft. The RMSE of these forecasts exhibited significantly larger spread than those of experiment JDH. Specifically, spreads in the temperature RMSE are often near 0.5 °C, and geopotential-height spreads ranged from 5 to 10 m (Fig. 8c, d). The evolution in pairs described above is observable, as is some convergence near the end of some of the forecast periods. No substantial evidence exists in these data suggesting that the average of the ensemble members

perform better or worse than the control run.

Similar to experiment JDH, the precipitation, sea-level pressure, and temperature demonstrate the most spread of all the forecasts during the last few verification times of the forecast initialized 23 January, 1990. Sea-level pressure spreads reach 1.2 kPa, precipitation spreads exceed 1.5 cm (Fig. 9b, d), and temperature spreads are near 3 °C.

The sea-level pressure derived from the ensemble-averaged forecast is further away from the observations than the control forecast or negligible differences are present until 66 hours into the forecast. The temperature spread is largest toward the south as in experiment JDH, but the spread is also larger throughout the cross-section. This is expected because the initial temperature perturbation at low levels is much larger here than for experiment JDH.

The precipitation also shows more variability in the peak location along the cross-section. Notably, the 12 hour accumulated precipitation ensemble-averaged forecasts at 60, 66, and 72 hours more closely match the observations than the control forecast. The average and the control at station WFV at 60 hours differs by 4 mm, with the average almost exactly duplicating the observation. This is the largest deviation of the average and the control of any of the experiments.

8. Discussion

a. General performance

The ensemble members necessarily have a greater RMSE than the control for the zero hour forecast, where the analysis is the basis for comparison. MC2 performs horizontal and vertical grid staggering for the integration, and then the post-processing routines performs unstaggering. The interpolation required to perform the staggering produces the error. Ideally, the RMSE of the control runs would exceed the RMSE of the ensemble average at some later time in the forecast.

Throughout one forecast period, individual ensemble members and the control forecast often exhibit similar trends in the error growth. Furthermore, they tend to evolve in pairs (Fig. 8). This

can be explained by the fact that two of the four members of the ensemble differ by only one small perturbation in the form of a weak incipient cyclone. The largest differences between members is between the addition and subtraction of the perturbations. In this case, since these opposing perturbations are centered around the analysis, the error is the same. Thus at initialization time, the two members that contain two local perturbations have similar error, and the two members that contain one local perturbation have similar error. During the forecast, the pairs switch so that the two members with the positive perturbations develop similarly as do the two with negative perturbations.

For this domain, convergence of solutions is apparent near the end of the forecast period in nearly all of the forecasts. This may also be understood as the seeded structures that compose the perturbations, or the bred modes, advecting out of the domain. When the perturbations are placed close to the upstream boundary, some divergence is evident at the end of the forecast period.

An optimal result would exhibit a good spread of ensemble members between the zero hour forecast and the time of solution convergence. This requires modes that grow or die quickly and differ between ensemble members. Initializing ensembles that differ more (i.e. using more than one location algorithm) would vary the spread, but it would not increase the range of extremes to envelope more observations.

The BGM forecasts produce the largest spread in RMSE curves (Fig. 8). Recall the perturbation is applied over the entire domain, while the other six experiments only apply local perturbations. Using the six-hour forecast errors (e.g. Toth and Kalnay 1993), rather than the 24-hour errors, will create a larger perturbation near the upstream boundaries, which may add to the spread in RMSE between the ensemble members. The boundary conditions are part of the "truth" that the BGM method uses to determine perturbations, so common values near the upstream boundaries lead to small perturbations. Using 24-hour differences to generate perturbations allows more common values. Subsequent forecasts then display convergence earlier in the forecast period, a characteristic that is evident in forecasts from 27 and 28 January. Further experiments

can assess the value in using six-hour bred modes.

The different structure of the two different perturbations become more obvious when verifying against observations. Though both induce significant mid-tropospheric disturbances, the low-level perturbation ($p_0 = 85$ kPa) produces a much larger initial surface perturbation than the upper-level perturbation ($p_0 = 50$ kPa). Thus, to produce a large spread in the surface variables of sea-level pressure and temperature, the upper-level perturbation must develop low-level structures during the integration, which occurs in some cases. The upper-level perturbations develop a spread that is consistently less than the spread induced by the lower-level perturbations throughout the 72 hour forecast period. Precipitation, on the other hand, can deviate much sooner since the mid-tropospheric vertical velocities, present in both perturbations, more often control this parameter.

Both types of perturbations are vertically stacked, but as they evolve with the model integration, the baroclinity in the initial conditions induces tilt. The initial surface disturbance present in the low-level perturbations is advected ahead of the upper-level disturbance, with a larger surface magnitude than what is visible for the upper-level perturbation. Both types also develop a significant surface disturbance to the upstream side of the initial perturbation, which is sometimes larger than what appears downstream. Thus, addition of a mid-tropospheric cyclonic circulation causes development of a surface low downstream and a surface high upstream, and subtraction causes development of a surface high downstream and a surface low upstream. Dynamics tells us this will help maintain the strength of the perturbation.

b. Analysis of specific events

1) 25-28 JANUARY, 1990

By examining one particular event during the study period, we may gain more physical insight to the performance of each of the perturbations. As a first example, consider a low-pressure system that tracked across the model domain 26-28 January. The analysis at 00 UTC 26 January

indicates a surface low pressure of approximately 97.5 kPa located near the tip of the Aleutian Islands (Fig. 3k). It was the later stages of a vigorous system that formed over the Kuroshio current in the western Pacific. As it slowly moved along the Aleutians, it weakened and filled so that by 00 UTC 27 January, the central pressure was about 98.7 kPa, and it was combining with a weak, persistent low over the Gulf of Alaska to form an elongated region of low pressure aligned with the southern coast of Alaska (Fig. 3m). By 00 UTC 28 January, the low pressure was digging into a weak low-level ridge with an axis cutting north to south through the Gulf of Alaska (Fig. 3o). Over the following day, the low strengthened and tracked rapidly southeastward and intersected the BC coast just south of the Queen Charlotte Islands (Fig. 3q). The observations along the coastal cross-section suggest a minimum central low along the coast of about 98.6 kPa at 18 UTC 28 January. The system then continued across southern BC.

In general, the model does a poor job of predicting the evolution of this system. Virtually no evidence of it exists in the sea-level pressure for the control forecasts initialized 26 and 27 January, when the lowest forecasted pressure along the cross-section is near 100.8 kPa. The control forecast initialized 28 January did a better job, giving a minimum pressure of around 99.8 kPa at 12 UTC, but the model moved the low onshore too quickly and it still did not deepen it enough. The analysis apparently did not contain the crucial ingredients of this developing cyclone.

The spread in the ensemble members for this system is negligible for all seven of the experiments, suggesting that the developing cyclone was not sufficiently perturbed. When the analyses are examined closely, it is clear that local maxima of tropopause deformation, low-level PV, and divergence aloft do exist associated with this system. The constraint that prevents perturbation overlap sometimes prevented their placement in important areas. On 26 January, only the search for low-level PV maxima resulted in a perturbation to the area because this was *the maximum* within the domain. Higher values of tropopause deformation and divergence aloft within 20 degrees of this disturbance disallowed the placement of perturbations there for the other

ensembles, even though the values at this location were the second highest in the search domain.

On 27 January, experiments JDH and JDL did not contain perturbations to this system for the same reason. Finally, on 28 January, none of the perturbations were connected with this low because of a stronger low-pressure center spinning up due south of the Aleutians, which was too close to the center of this low. This problem may be corrected by relaxing the spacing constraint to allow some overlap, and reducing the magnitude of the perturbations so that overlapping structures do not combine to become unrealistic.

When located correctly, these experiments are unsuccessful in perturbing the correct mode. As mentioned, experiments LPVH and LPVL placed perturbations near this system for the forecast initialized 26 January. The perturbation survives the integration, but it does not develop the rapid movement displayed in the observations and analyses. It moves slowly along the Aleutians just as in the control forecast. The BGM forecasts were equally as unsuccessful.

Of the forecasts initialized 27 January, LPVH, LPVL, SPVH, and SPVL all located perturbations in the vicinity of this system, but none of the members accurately predicted it. SPVH and SPVL held the low center too far north, combining with the weaker, stagnant low in the Gulf of Alaska. In these cases, a second low shown in the observations for 00 UTC, 30 January falls to around 98.1 kPa. All of these experiments show a larger spread associated with this low than with the first one, with LPVL being the largest at 0.7 kPa. Note that SPVH and SPVL did not locate second perturbations where this second low was developing. JDH and JDL did locate perturbations there, and JDL displays a spread of around 0.8 kPa at the 72 hour forecast. The small difference between JDL, which altered the second low-pressure center directly, versus LPVL, which did not, points to the ability of these local perturbations to affect a more extensive area of the flow. The surface high created upstream of the initial perturbation contributes to the deepening of a surface low further upstream.

2) 23-25 JANUARY, 1990

A different low-pressure center crossed the coast about 12 UTC 25 January, just north of the Queen Charlotte Islands (Fig. 10f, labelled C). The system began as a shortwave propagating along the southern edge of a more persistent low north of the Aleutian Islands (Fig. 10a, labelled A). It deepened as it interacted with an inverted shortwave trough associated with a low-pressure center northwest of Hawaii (Fig. 10a, labelled B). By 06 UTC 24 January, the sea-level pressure analysis indicates a cut-off structure with a central low still above 100.7 kPa. Its cut-off structure was distinct by 18 UTC as it tracked east-northeast through the North Pacific. The analysis valid at 06 UTC on 25 January contains the lowest central pressure of the system at 97.7 kPa. The system then filled as it crossed the shoreline and rapidly tracked to the Alberta-Saskatchewan border by 00 UTC, 26 January (Fig. 3k). By this time little evidence of the system existed along the BC coast as a shortwave ridge built into the region. Figure 10 shows the evolution of this system, using analyses valid every 12 hours.
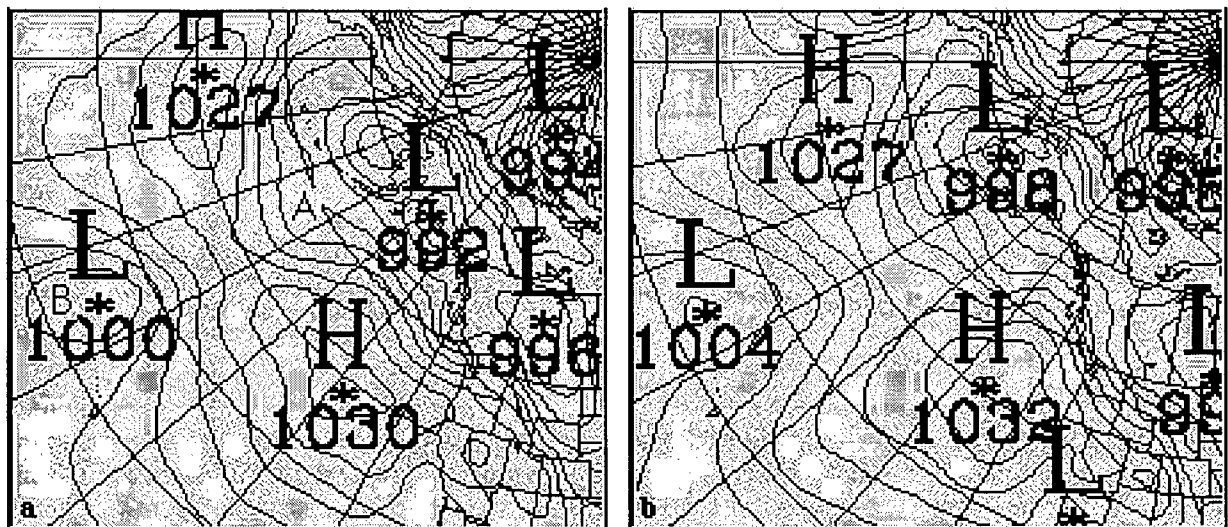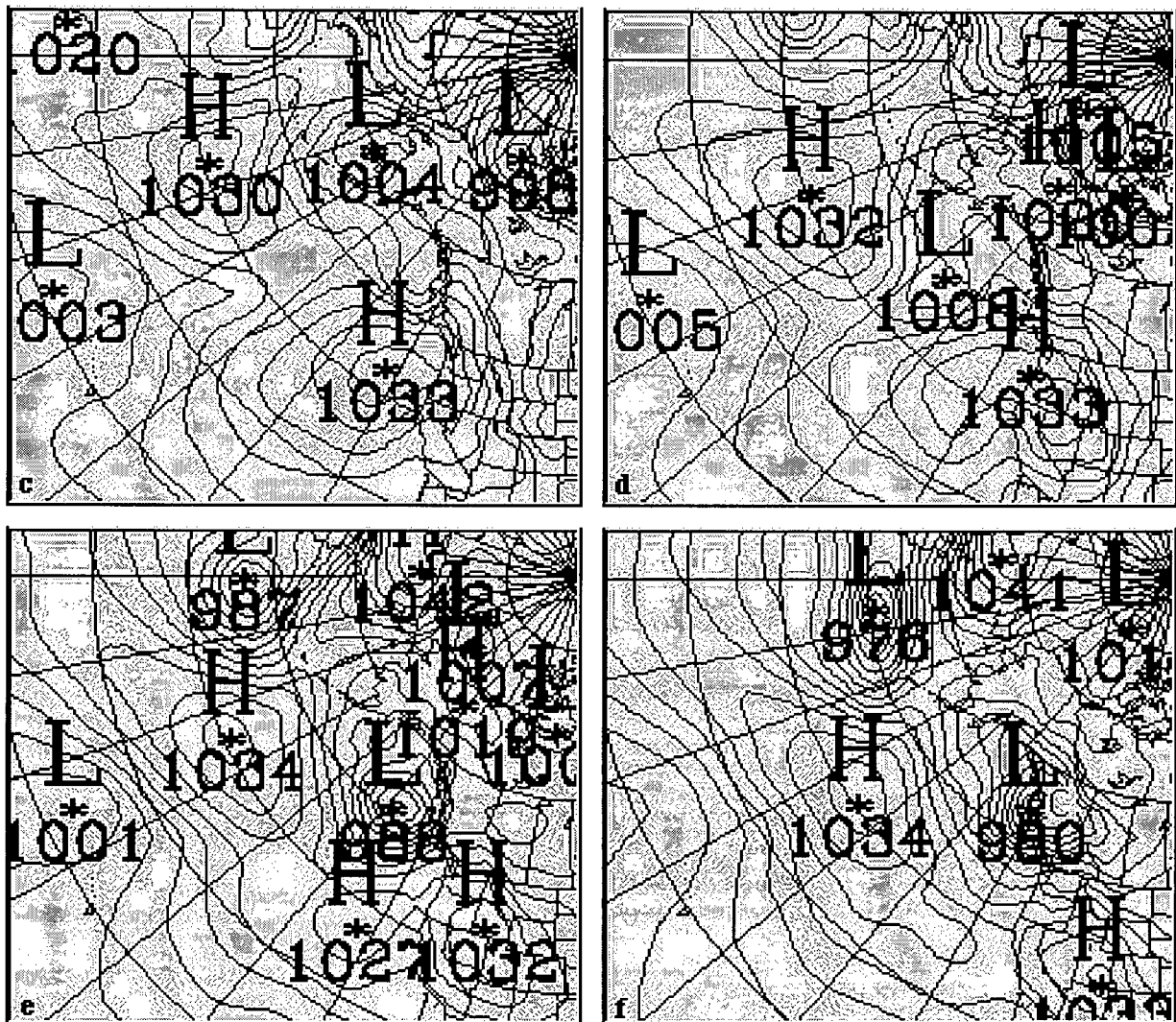
FIG 10. Sea-level pressure analyses valid every 12 hours from 00 UTC 23 January 1990 through 12 UTC 25 January. Pressures are in mb (1 mb=0.1 kPa), and contour interval is 4 mb.

MC2 predicts this event better, though not perfectly. The control forecast initialized 23 January deepens the sea-level pressure of this system to only 98.9 kPa before it intersects the coast, deepening it further to 98.4 kPa by the time it reaches the BC-Alberta border. The forecasted low-pressure center is more zonally elongated than the analysis indicates. Once it reaches the coast, it does not continue to move over land rapidly enough, and it instead stalls near the BC-Alberta border. Low pressure lingers over the Pacific and as far south as Vancouver.

51

The control forecast initialized 24 January deepens to 98.2 kPa just as the low-pressure center moves onshore adjacent to the Queen Charlotte Islands. The location of the low before it makes landfall is closer to the analysis for this forecast than for the previous one, which is expected given the shorter forecast time. But this forecast maintains a trough axis extending to the southwest of the low center, and the corresponding analysis is more circular. As the forecasted low moves onshore, this trough lags behind and keeps low pressures over the coast. The model moves the low center inland too slowly again, and cross-sectional verification pressures rise much faster than those forecasted.

Even the forecast initialized 25 January, which deepens the low to 974 as it crosses the Queen Charlotte Islands, does not move it inland fast enough. It is not clear whether this is a model deficiency (i.e. topography interaction) or the result of some characteristic that is repeatedly absent from the analysis. Further experimentation could yield an answer.

Each of the experiments successfully places an incipient cyclone very near this disturbance, resulting in encouraging spreads in some of the forecasts. Along the verification cross-section, the ensembles initialized 23 January display a larger spread in sea-level pressure than the shorter ensemble forecasts initialized 24 and 25 January. This occurs despite the control forecasting a central low pressure of only 98.9 kPa as the low crosses the shoreline. In each experiment, the lowest pressure extreme of the ensemble is lower than the two observations closest to the low-pressure center.

The low-level perturbations normally induce a larger spread in the surface variables. Also, precipitation spreads are consistently the largest and show the most variation in peak locations. But forecasts of the cyclone affecting BC on 25 January, initialized with upper-level perturbations, all display larger spreads in sea-level pressure and precipitation in the very short range (0-12 hours) than the low-level perturbations in the same places. The upper-level perturbations decrease the static stability, allowing the upper-level perturbations to develop an associated low-level

structure rapidly. Below 70 kPa, the perturbation cools with height, but with a lapse rate less than moist adiabatic. The low-level perturbation temperature increases with height to near the tropopause.

Refer to figures 11-14, which depict the difference between the third and fourth members of experiments LPVL and JDL. Recall the third member is the addition of incipient cyclones, and the fourth is the subtraction, causing a negative difference in each of the initial sea-level pressure and 50 kPa geopotential-height perturbations. One row of any of the figures is from a forecast initialized on either 23, 24, or 25 January. The left column shows the perturbations at initialization time on the respective days, and the right column is the appropriate forecast hour verifying at 06 UTC, 25 January. In this manner, we can see how the perturbations handle the storm system described above.
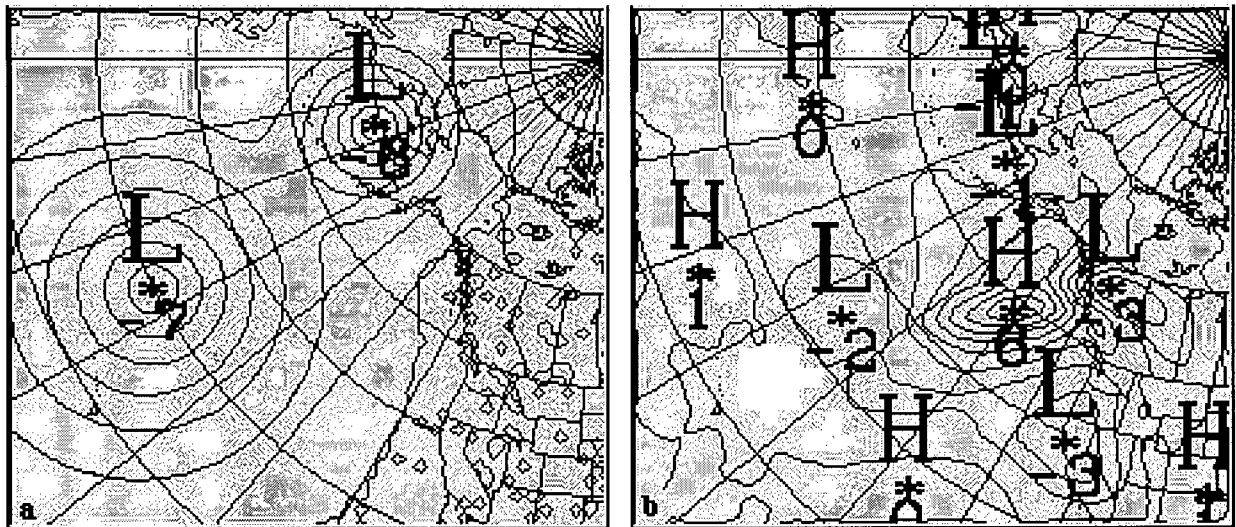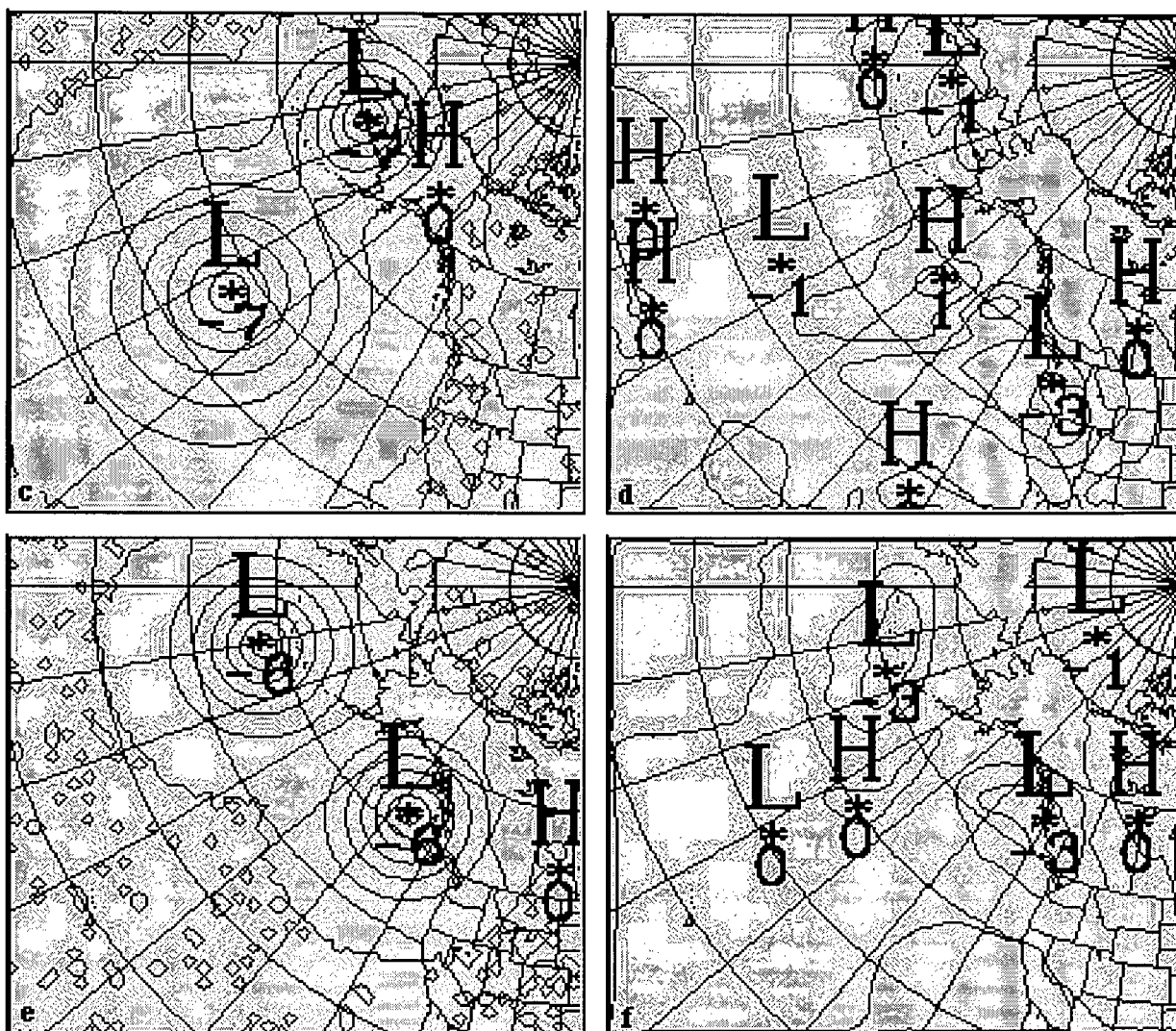
Figure 11. Calculated differences between members three and four of experiment LPVL, 50 kPa geopotential heights. Contour interval is 0.01 km (1 dam), and central values are given in decameters. The left column is at initialization time, 00 UTC on (a) 23, (c) 24, and (e) and 25 January. The noise away from the perturbation is caused by interpolation to the grid. The right column is valid at 06 UTC, 25 January, represented by the (b) 54, (d) 30, and (f) 6 hour forecasts from the corresponding initialization.
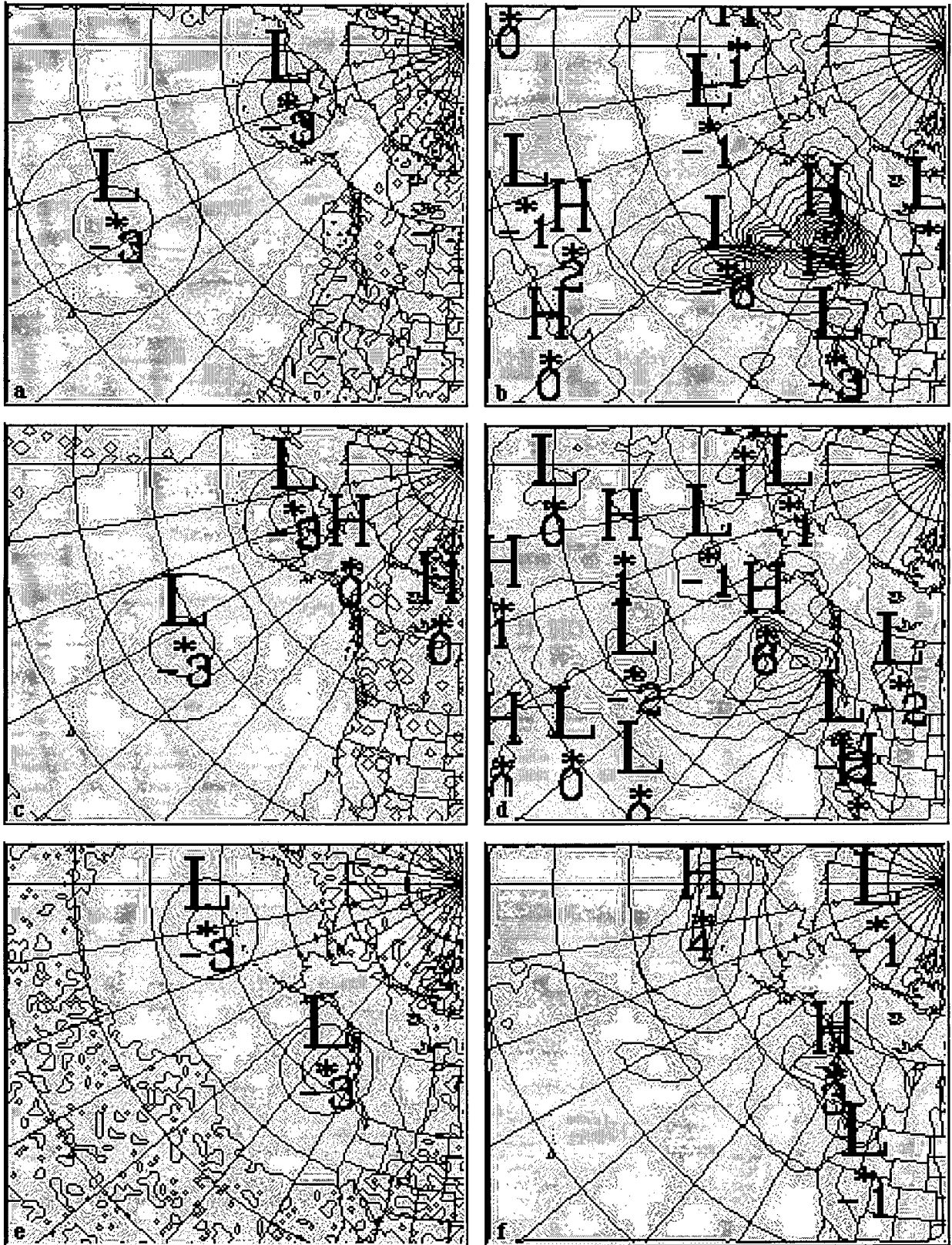
FIG. 12. Same as figure 11 for sea-level pressure. Interval is 1 mb (0.1 kPa)
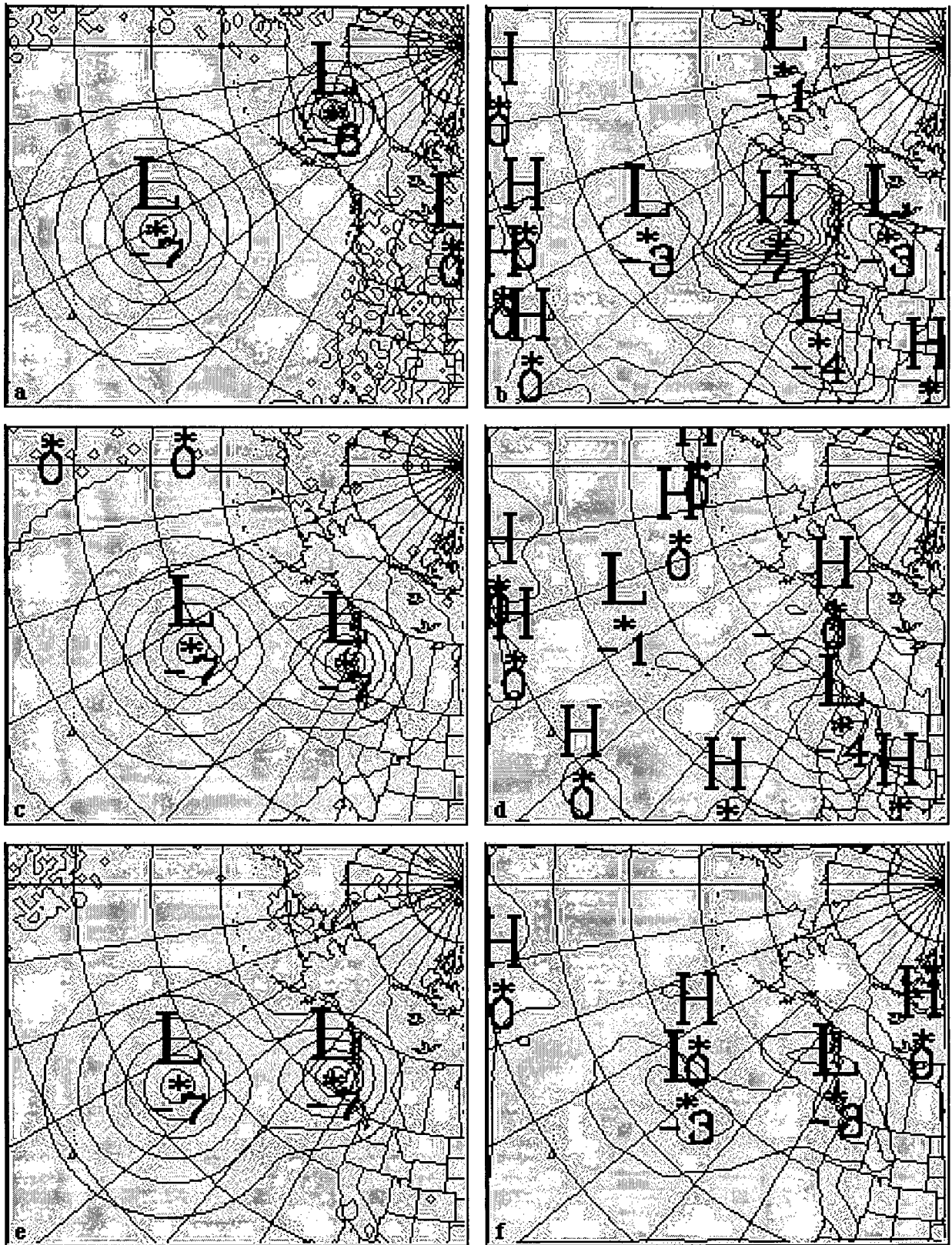
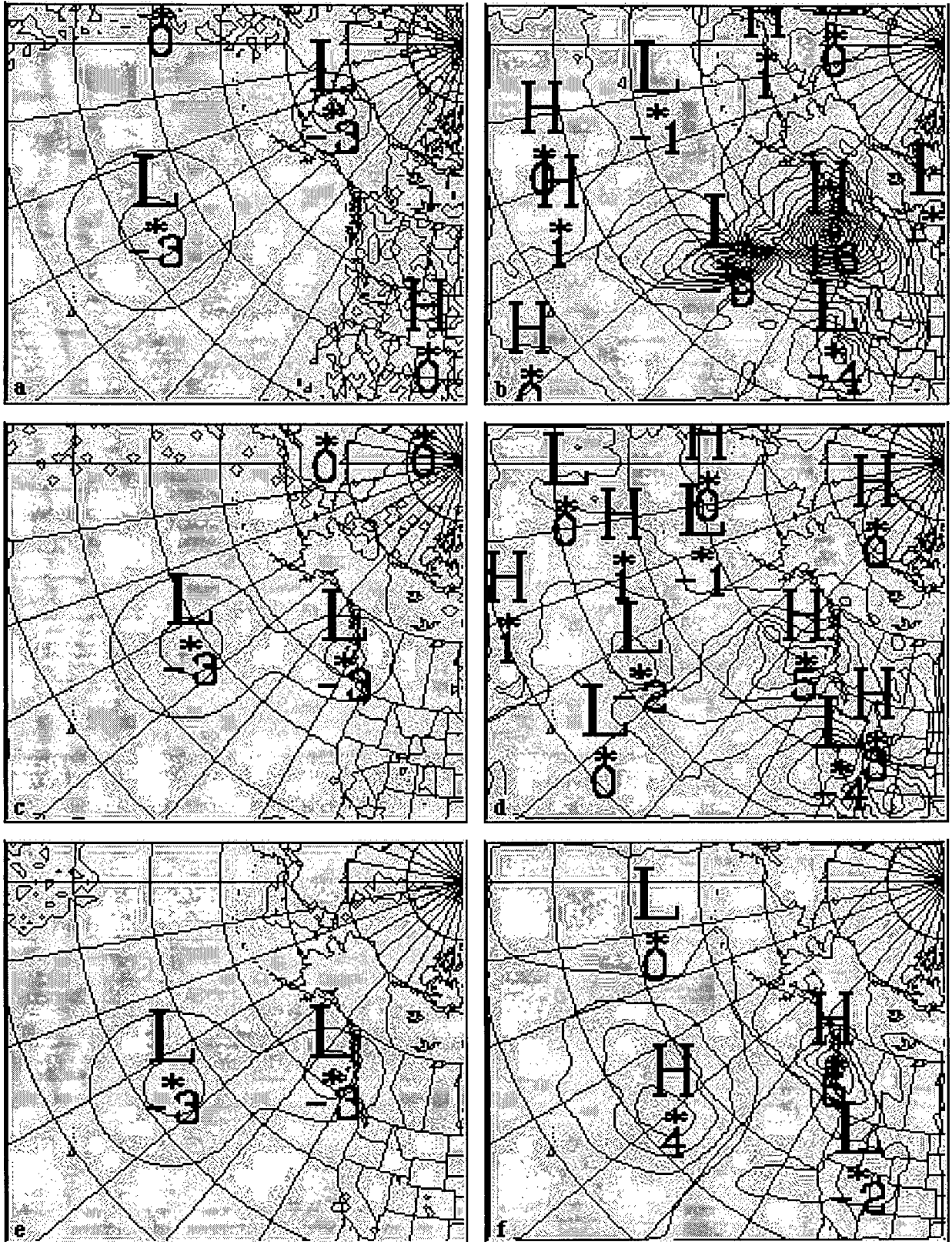Figure 13. Same as figure 11, but for experiment JDL.

Figure 14. Same as figure 12, but for experiment JDL.

Comparison with figure 10 reveals that LPVL and JDL (LPVH and JDH) located perturbations over a low-level inverted trough axis associated with the low-pressure center northwest of Hawaii. As this trough begins to combine with the shortwave to the north, the perturbation tracks to the northeast where the cut-off cyclone forms. The perturbation is strongest in the mid-troposphere, and the convergence associated with adding the incipient cyclone here causes subsidence near the surface, which gives higher sea-level pressures. Conversely, the subtraction of the cyclone causes causes upward vertical velocity near the surface and lower sea-level pressures. Thus, the difference between the sea-level pressures becomes positive by 6 hours into the forecast (Fig. 12f and 14f). As the forecasts progress, this difference normally increases, becomes the strongest overall feature in the difference fields, and lags slightly upstream from the upper-level perturbation. The maximum positive differences are in the region just behind the surface trough axis throughout the forecast period. Downstream from this feature, a weaker region of negative differences is often present as the mid-tropospheric cyclone develops a surface cyclone ahead of it. Similarly, the mid-tropospheric anticyclone, created by subtracting the cyclone, develops a surface anticyclone ahead of it. In both LPVL and JDL, the longer forecast time of 54 hours provides for more development of the perturbation and thus more forecast spread (Fig. 12b and 14b), which we saw in the cross-sectional verification.

In many of the forecasts, regions of alternating high and low differences develop both upstream and downstream from the original perturbations. These are evident in both the sea-level pressure fields and the upper-level geopotential height fields. Since a sensitive wavelength has been excited, in some cases a larger spread in forecasts develops away from the original perturbation, associated with some other developing system which was not perturbed.

Experiments SPVL and SPVH collocated perturbations with the northern shortwave component of this system. The difference fields for this experiment show good spread as well, but further north. This explains the relative lack of spread in the verification cross-sections. The

58

perturbation initialized 23 January strengthened at 50 kPa throughout the forecast period as its center remained just downstream of the trough axis. The perturbations initialized 24 and 25 January do not induce sufficient spread throughout the forecast period.

Comparison of the two low-pressure systems discussed above suggest that the spread between the forecasts increases when the control forecast better predicts an event. At times, this results from the perturbation location algorithms failing to place a perturbation near a developing disturbance. For other experiments, the ensemble members simply do not diverge rapidly enough to accommodate all possible scenarios. The analysis is lacking some element in the flow that is a major contributor to the cyclone development, and within 72 hours the forecasts may not reach a stage for which nonlinear processes dominate the evolution. Boundary-condition contamination limits the utility of an experiment with a longer integration to determine if the solutions would continue to diverge, as we would expect.

c. Suggestions for improvement

Before we can try to use the ensemble-averaged forecast in place of a single deterministic forecast, we must be able to envelope all of the observations in the forecast period. This is clearly not the case in these experiments. Potential improvements can be made by increasing the winds associated with the perturbation, changing the vertical structure of the perturbations, introducing tilt in the perturbation, and providing perturbed boundary conditions during the integrations.

The weakness of the wind circulation associated with the perturbation was noted in section 4. Gyakum et al. (1992) discuss the importance of low-level vorticity in the development of explosive cyclones. Increasing the wind circulation associated with the perturbations here may contribute to a greater spread between forecasts in the short term. To check the potential, an over-relaxation factor of 1.9 was used to iterate to a tolerance of 50% of the initial change, instead of the 90% used for these experiments. Appendix B shows these results in more detail. The computing time on an

R5000 processor increased sixfold from about 30 minutes. The associated winds increased to around 1.4 ms^{-1}, and a greater increase is possible. To keep the computing time short enough to maintain a viable ensemble system, the iterations would have to begin with an initial guess much closer to the final solution. This may be accomplished by making a first guess of an analytic solution incorporating the linear terms, and then iterating to include the nonlinear effects.

The vertical structure can be altered to introduce more static instability throughout the column by adjusting p_0 and the vertical extent of the perturbation. The factor 1 that is added to the vertical sine structure can also be changed to ensure mass conservation in the column.

To increase the initial baroclinity associated with the perturbations, a westward tilt with height can be included to account for baroclinic theory. This may have small effects, however, given the tilting evident in the model forecasts. Recall this occurs when the vertically stacked perturbation is modified by the baroclinity already present in the analysis.

The boundary conditions could be perturbed throughout the integration period by perturbing a larger domain, to avoid common boundary condition contamination. Perturbing boundary conditions for the BGM method would require an ensemble on a larger domain, within which this domain could be nested. Obviously, this results in a much larger computing investment, and the effects on short-range forecasts over a domain of this size may be inconsequential to western Canada.

As mentioned, the spacing restriction which prevents overlap of individual incipient cyclones should be relaxed to allow more sensitive regions to be perturbed. This requires a reduced perturbation magnitude so that the combined perturbation is not unrealistically strong.

All of these possible improvements must be subject to further experimentation. The large spreads in precipitation forecasts indicate that these methods have the potential to yield useful information for BC weather forecasters, largely because precipitation is such an important parameter. More spread is desirable for other variables such as sea-level pressure and temperature, which help determine the speed and strength of cyclones and also the difference between liquid and solid precipitation.

9. Conclusion

Six weather prediction experiments tested the feasibility of an empirically-based ensemble forecasting system designed to improve numerical guidance to forecasters on the west coast of Canada. A seventh employed the BGM method for use as a benchmark. Verification included standard statistical comparison with analyses and a cross-section of surface observations along the coast. Forecast spreads in RMSE were never significant within the 72 hour forecast period. The ensemble-averaged forecast and the control forecast rarely deviated from each other. Cross-section analysis showed the most spread in precipitation forecasts, with some changes in the location of precipitation peaks, a characteristic likely attributable to the nonlinear nature of the microphase physics and convective parameterization in MC2. Sea-level pressure at times spread adequately, but never enough to accommodate all of the observations. Temperature forecasts did not spread adequately.

The perturbations placed at lower levels ($p_0 = 85$ kPa) had stronger initial surface characteristics, but the upper-level perturbations ($p_0 = 50$ kPa) demonstrated the ability to develop low-level structures. Overall, experiments LPVL (low-level perturbations placed according to low-level PV values) and JDL (low-level perturbations placed according to jet divergence) yielded the largest forecast spreads in the best locations. The low-level perturbations were stronger in the mid-troposphere than at the surface, causing either convergence or divergence which immediately

61

changed the nature of the surface structure to conform to baroclinic theory. The perturbations placed according to tropopause deformation developed the least, and resulted in the least spread between members.

The use of common boundary conditions for each of the ensemble members was a limiting factor in this project. Configuring the domain with a large distance upstream from western Canada minimized these effects. Convergence in the RMSE curves occurred near the end of the forecast period in many of the ensembles, but cross-sectional verification showed no decrease in spread near the end of the forecast period if a perturbed region was crossing the coastline at that time. Longer integrations would require perturbed boundary conditions, so their effect was not tested here.

In its current state, the results of this study are not useful in an operational mode. The forecast spread must envelope the observations more frequently before the value of the probabilistic information can be assessed. It is not clear whether the dominant factor is a limitation of ensemble forecasting techniques applied to short-range forecasting, or simply flaws in these systems.

Several possible improvements were suggested, aimed at increasing the spread of the ensemble in the short range. These included increasing the initial vorticity of the perturbations, changing the perturbation structure, perturbing the boundary conditions, and changing the spacing of the incipient cyclones. Experimentation will continue to evaluate the performance of these modifications.

Additionally, future research will investigate other perturbation methods. One potential approach is to generally sharpen parameter gradients throughout the domain in an effort to improve the baroclinic preconditioning. This is conducive to perturbing the boundary conditions. Details would focus on determining appropriate magnitudes and variables to perturb.

REFERENCES

Bergeron, Guy et al., February 1994: *Formulation of the Mesoscale Compressible Community (MC2) Model.* Cooperative Centre for Research in Mesometeorology.

Barnes, Stanley L., Fernando Caracena, and Adrian Marroquin., 1996: Extracting synoptic scale diagnostic information from mesoscale models: the Eta model, gravity waves, and quasigeostrophic diagnostics. *Bulletin of the American Meteorological Society,* **77** (3), 519-528.

Buizza, Roberto, 1992: *Unstable perturbations computed using the adjoint technique.* Research Department Technical Memorandum No. 189, European Center for Medium-Range Weather Forecasting.

Charney, J., 1955: The use of the primitive equations of motion in numerical prediction. *Tellus,* **7**, 22-26.

Davis, Christopher A. and Kerry A. Emanuel, 1991: Potential vorticity diagnostics of cyclogenesis. *Monthly Weather Review,* **119**, 1929-1953.

Daley, Roger and Thomas Mayer, 1986: Estimates of global analysis error from the Global Weather Experiment observational network. *Monthly Weather Review,* **114**, 1642-1653.

Deardorff, J.W., 1978: Efficient prediction of ground surface temperature and moisture with inclusion of a layer of vegetation. *Journal of Geophysical Research,* **83**, 1889-1903.

Epstein, Edward S., 1969: Stochastic dynamic prediction. *Tellus,* **21**, 739-759.

Fouquart, Y. and B. Bonnel, 1980: Computations of solar heating of the earth's atmosphere: A new parameterization. *Contributions to Atmospheric Physics,* **53**, 35-62.

Garand, L., 1983: Some improvements and complements to the infrared emmisivity algorithm including a parameterization of the absorption in the continuum region. *Journal of the Atmospheric Sciences,* **40**, 230-244.

Gyakum, John A., Paul J. Roebber, and Timothy A. Bullock, 1992: The role of antecedent surface vorticity development as a conditioning process in explosive cyclone intensification. *Monthly Weather Review,* **120**, 1465-1488.

Hamill, Thomas J. and Stephen J. Colucci, 1997a: Verification of Eta/RSM short-range ensemble forecasts. Accepted by *Monthly Weather Review.*

_____, 1997b: Evaluation of Eta/RSM ensemble probabilistic forecasts. Submittted for publication 7 Oct., 1996.

Héreil, Philippe and René Laprise, 1996: Sensitivity of internal gravity waves solutions to the time step of a semi-implicit semi-Lagrangian nonhydrostatic model. *Monthly Weather Review,* **124**, 972-999.

Hoffman, Ross N. and Eugenia Kalnay, 1983: Lagged average forecasting, an alternative to Monte Carlo forecasting. *Tellus,* **35A**, 100-118.

Hoskins, B. J., M. E. McIntyre and A. W. Robertson, 1985: On the use and significance of isentropic potential vorticity maps. *Quarterly Journal of the Royal Meteorological Society,* **111**, 877-946.

Houtekamer, P.L. and Jacques Derome, 1995: Methods for ensemble prediction. *Monthly Weather Review,* **123**, 2181-2196.

Kuo, H.L., 1974: Further studies on the parameterization of the influence of cumulus convection on large-scale flow. *Journal of the Atmospheric Sciences,* **31**, 1232-1240.

Leith, C. E., 1974: Theoretical skill of Monte Carlo forecasts. *Monthly Weather Review,* **102**, 409-418.

Lönnberg, P., and A. Hollingsworth, 1986: The statistical structure of short-range forecast errors as determined from radiosonde data. Part II: The covariance of height and wind errors. *Tellus,* **38A**, 137-161.

Lorenz, E. N., 1965: The predictibility of flow which possesses many scales of motion. *Tellus,* **17**, 289-307.

Mailhot, Jocelyn, March 1994: *The Regional Finite-Element (RFE) Model Scientific Description - Part 2: Physics.* RPN, Atmospheric Environment Service.

Mullen, Steven L., 1994: An estimate of systematic error and uncertainty in surface cyclone analysis over the North Pacific Ocean: some forecasting implications. *Weather and Forecasting,* **9**, 221-227.

_____ and David P. Baumhefner, 1994: Monte Carlo simulations of explosive cyclogenesis. *Monthly Weather Review,* **122**, 1548-1567.

Press, William H., *Numerical Recipes: The Art of Scientific Computing.* Cambridge University Press, 1986, 390-395.

Robert, André and Evhen Yakimiw, 1986: Identification of an inflow boundary computational solution in limited area model integrations. *Atmosphere-Ocean,* **24**(4), 369-385.

Robert, André, Tai Loy Yee and Harold Ritchie, 1985: A semi-Lagrangian and semi-implicit numerical integration scheme for multilevel atmospheric models. *Monthly Weather Review,* 388-394.

Stull, Roland B., *Meteorology Today for Scientists and Engineers.* West Publishing Company, 1995, 150-151.

Sundqvist, H., E. Berge and J.E. Kristjansson, 1989: Condensation and cloud parameterization studies with a mesoscale numerical weather prediction model.. *Monthly Weather Review,* **117**, 1641-1657.

Thorpe, A. J., 1985: Diagnosis of balanced vortex structure using potential vorticity. *Journal of the Atmospheric Sciences,* **42**, 397-406.

Toth, Zoltan and Eugenia Kalnay, 1993: Ensemble forecasting at NMC: The generation of perturbations. *Bulletin of the American Meteorological Society,* **74**, 2317-2330.

Appendix A

The perturbations originate in the geopotential height field. The domain on which the perturbations are specified is larger than the domain of the model integrations, covering most of the hemisphere, but they can be located only in the North Pacific Region. Once specified, a balance condition is imposed to yield the other dynamic and thermodynamic fields. Charney (1955) derived the condition used here:

$$\nabla^2 \Phi = \nabla \cdot (f \nabla \Psi) + 2 \left[\frac{\partial^2 \Psi}{\partial x^2} \frac{\partial^2 \Psi}{\partial y^2} - \left(\frac{\partial^2 \Psi}{\partial x \partial y} \right)^2 \right]$$

where Φ is the geopotential and Ψ is the stream function, defined in the usual way. If the geopotential field and boundary values of the stream function are specified, a well-posed Poisson problem results. A standard successive over-relaxation (SOR) procedure, as described in Haltiner and Williams (1980), iterates toward the solution to the stream function within the domain.

The background geopotential is logarithmic vertically, with no horizontal gradients. The stream function is set to zero at the boundaries. Because the perturbation is regional and far away from the edges of the iteration domain, the boundaries do not interfere with the solution. Also, the associated temperature perturbation only depends on the geopotential departure from its surroundings, so the precise background geopotential is inconsequential. The only requirement is that it increases with height to remain physically reasonable.

Second-order, centered finite differences provided discretization for all of the interior points when iterating. The same served to get zonal and meridional components of the wind from the stream-function solution. The hypsometric relation determined the temperature structure from the geopotential height field. Once a solution is reached, the geopotential, temperature, and winds can be added or subtracted to the analysis.

CPU time on all calculations except the iteration procedure is inconsequential. The iteration time depends on the number of cyclones and anticyclones that comprise the perturbation, and it is approximately 30 minutes for 900 iterations on an R5000 processor. In all cases, an over-relaxation factor of 1.3 was used, and it was assumed a satisfactory solution was reached when the maximum change in the geopotential for each iteration was 90% of the change for the first iteration. This was usually on the order of 10-100 $m^2 s^{-2}$, which translates to centimeters of geopotential height. The number of iterations could be reduced dramatically since the solution converges to the same order on the first iteration. The 90% criteria was arbitrary and many more iterations would be necessary to reach an accuracy of millimeters.

Figure B.1 is a plot of the maximum change in the value of the stream function over the

perturbed domain versus the number of iterations. It can be seen that the curve is becoming quite

flat at when iterations approach 4000, where the stream function solution reaches half of its initial

maximum change.

It was found that iterating to 50% of the initial change produces higher perturbation wind

speeds, but negligibly affects the forecasts. The left column in Figure B.2 shows the calculated

differences between the third and fourth members of experiment LPVL (low-level perturbations

placed according to low-level PV values). The right shows the same perturbation iterated longer to

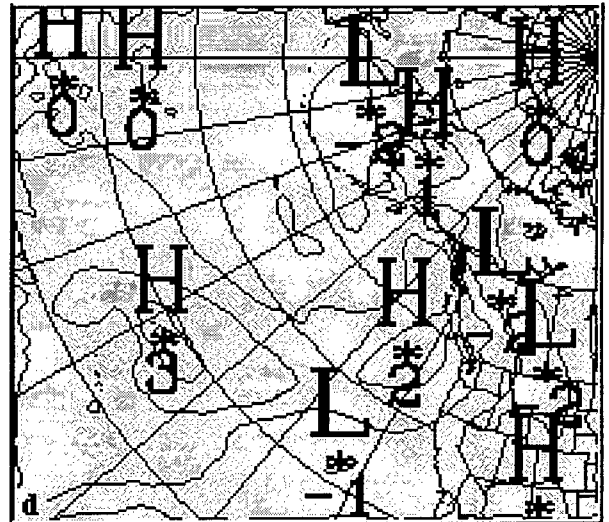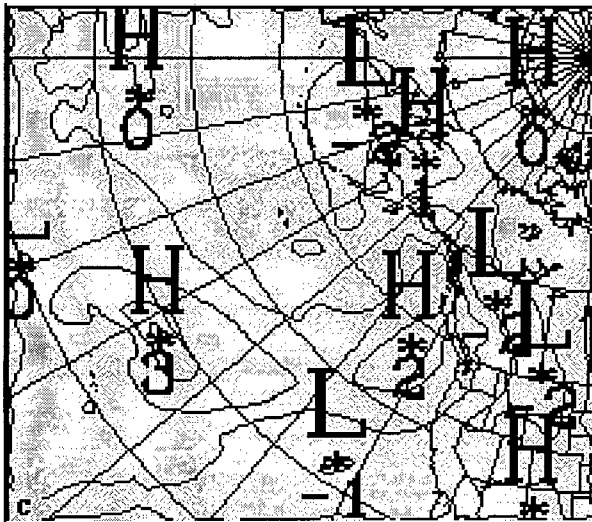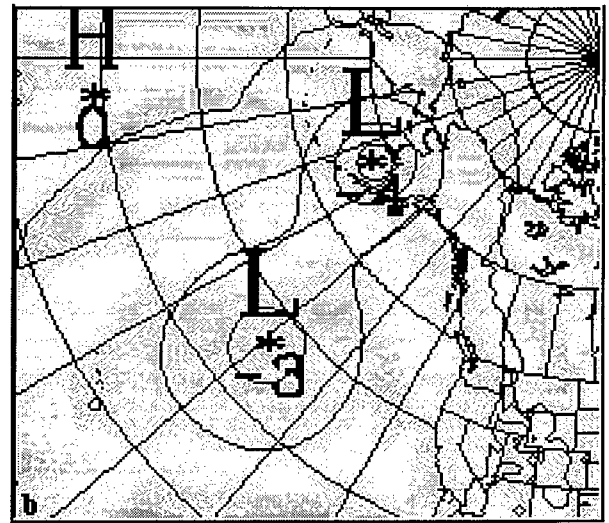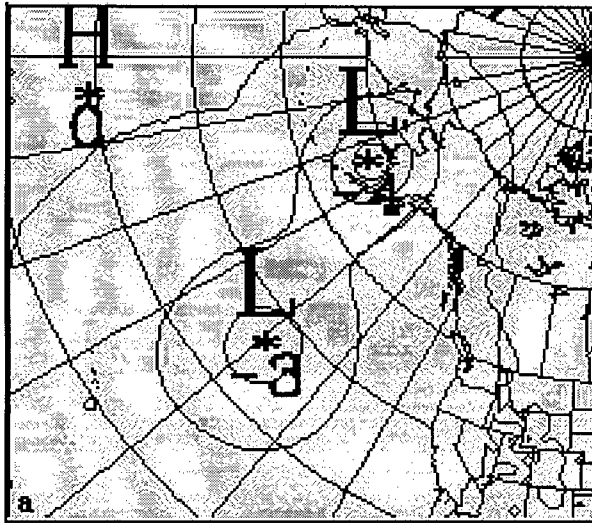increase the winds. These forecasts are initialized 21 January 1990.
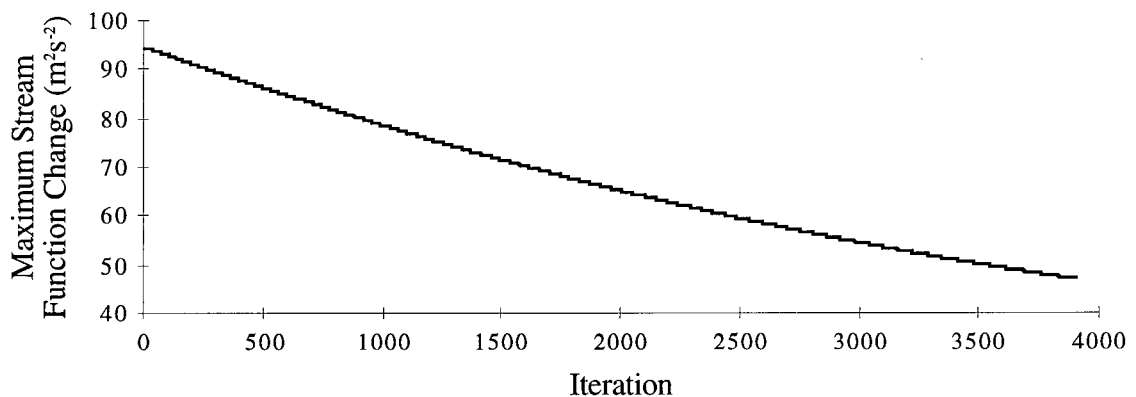
FIG. B.1. Maximum change in the stream function over the perturbed domain
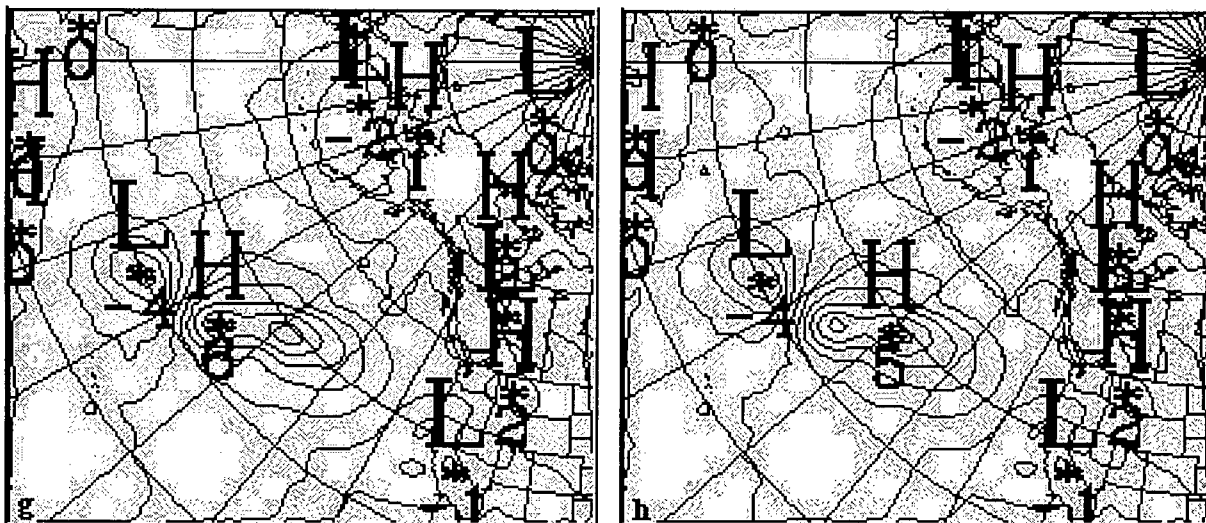versus the iteration number. For this curve, the over-relaxation factor was 1.9.
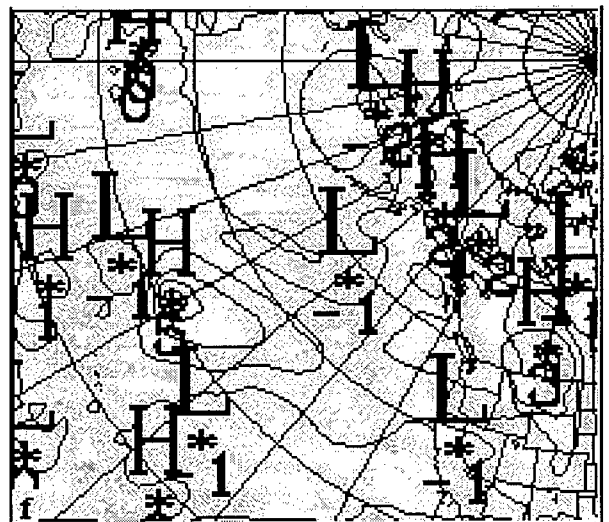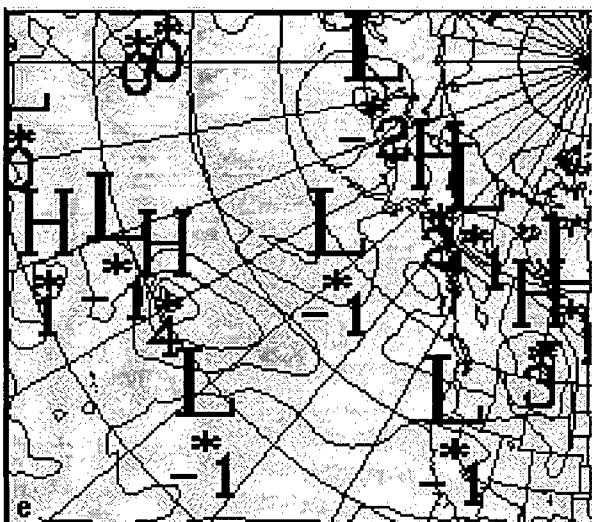
FIG. B.2. Sea level pressure differences between members three and four of experiment LPVL a perturbation iterated to a lower tolerance. (a) 00, (c) 24, (e) 48, and (g) 72 hour forecasts for experiment LPVL. (b) 00, (d) 24, (f) 48, and (f) 72 hour forecasts with higher perturbation wind speeds. The shape and locations of the perturbations are otherwise identical. Contour interval is 1 mb (0.1 kPa).