THE ADJOINT MODEL TECHNIQUE AND METEOROLOGICAL APPLICATIONS

O. Talagrand
Laboratoire de Météorologie Dynamique
Paris, France

Summary: The theory of adjoint equations is presented, and is shown to provide a powerful tool for studying the sensitivity of a numerical model of the atmospheric flow to initial and boundary conditions, and to physical parameters. Various possible meteorological applications are discussed. It is shown in particular how the adjoint model technique can be used for adjusting a high resolution model simultaneously to observations distributed in time and to large scale fields produced by a lower resolution model.

1. INTRODUCTION

The adjoint equations constitute a very powerful tool which can be applied to many theoretical and numerical problems. Their main interest for numerical applications is that they provide efficient algorithms for explicitly computing the gradient, or partial derivatives, of a "complicated" compound scalar function of a set of arguments. They can in particular be used in problems of optimal control, in which one wants to minimize some scalar function $u \rightarrow \mathcal{J}(u)$ of a set of arguments $u = (u_1, u_2, \ldots, u_n)$. For given u, one integration of appropriate adjoint equations will determine the gradient of $\mathcal{J}$ with respect to u. Successive gradients computed in this way can then be used in an iterative descent process in order to determine the minimizing u. A possible meteorological application of optimal control is assimilation of observations, which can be stated in the following terms. Given a forecasting numerical model, a set of observations and a scalar function $\mathcal{J}$ which, for any model solution, measures the "distance" between that solution of the observations, find the model initial conditions at some time t_o such that the corresponding solution minimizes the function $\mathcal{J}$.

Adjoint equations can also be used in meteorological situations where one is interested in knowing a gradient for itself, independently of any minimization process. A typical example is the following. A general circulation model is used for climatic studies, and one wants to evaluate the sensitivity of some particular climatic indicator $\mathcal{E}$ (e.g. the temperature averaged over some spatial and temporal domain) to the model's physical parameters. One integration of the adjoint equations will suffice in order to determine the gradient of $\mathcal{E}$ with respect to all physical parameters, thus defining an extremely efficient way to evaluate the required sensitivity to the various parameters.

The idea of applying adjoint equations to meteorological problems is by no means new. It has long been advocated by various authors, especially from the Soviet Union (Marchuk , 1974, 1982; Kontarev, 1980). In recent years, various works have been performed in order to study the sensitivity of numerical models with respect to physical parameters (Hall and Cacuci, 1983, 1984). However, probably because of their somewhat sophisticated mathematical technicality, adjoint equations have not so far been much applied to meteorological problems, and a precise assesment of the possibilities they could offer in practical situations has still to be made.

In this article, the general principle of adjoint equations is first described (Section 2) and applied (Section 3) to the problem of data assimilation, considered in the terms stated above. The theory is then extended (Section 4) to the problem of determining both the initial and boundary conditions of a limited area model in the case where one wants to adjust that model, not only to a set of observations, but also to a large scale forecast already produced by another model. Other possible meteorological applications of adjoint equations are briefly described (Section 5) and conclusions are presented (Section 6).

The theory presented in Section 3 is at the basis of the numerical experiments described in the article by Courtier (this volume) which will hereafter referred to as C 85.

2. GENERAL PRINCIPLE OF ADJOINT EQUATIONS

The appropriate mathematical framework in which to use adjoint equations is provided by the theory of Hilbert spaces, the basic elements of which can be found in many text books on linear analysis, for instance in Reed and Simon (1980). A Hilbert space is a linear space in which an inner product has been defined and which has the additional property of being complete with respect to the distance defined by the inner product. Although this last theoretical requirement does have some practical implications, we will only mention that a finite dimensional space on the field of the real numbers, and on which an inner product has been defined, is always complete, and that no difficulty can therefore arise in practical situations as to the existence of adjoint equations.

We will first briefly recall the definitions of two general notions which are basic to what will follow, namely the notions of a gradient and of an adjoint operator.

Let $\mathcal{F}$ be a Hilbert space on which an inner product noted (,) has been defined, and let $v \rightarrow \mathcal{J}(v)$ be a scalar differentiable function defined on $\mathcal{F}$. At any point v in $\mathcal{F}$, there exists a uniquely defined vector $\nabla_v \mathcal{J}$ such that the first order variation $\delta \mathcal{J}$ of $\mathcal{J}$ resulting from a variation δv of v is equal to the scalar product

$$\delta \mathcal{J} = (\nabla_v \mathcal{J} , \delta v) \tag{2.1}$$

$\nabla_v \mathcal{J}$ is the <u>gradient</u> of $\mathcal{J}$. As for ordinary gradients in physical space, $\nabla_v \mathcal{J}$ is directed along the direction of fastest variation of $\mathcal{J}$, and its modulus is equal to the rate of variation of $\mathcal{J}$ per unit length along that direction. In particular, the function $\mathcal{J}$ is stationary at a given point of $\mathcal{F}$ if and only if $\nabla_v \mathcal{J}$ is equal to zero at that point. When $\mathcal{F}$ has finite dimension and is described by orthonormal coordinates x_i, it is well known that the components of $\nabla_v \mathcal{J}$ are the partial derivatives $\frac{\partial \mathcal{J}}{\partial x_i}$. But the concept of a gradient is much more general, and we will use in the sequel the general abstract expression (2.1) rather than explicit coordinates.

Let $\mathcal{E}$ be another Hilbert space, with inner product noted < , >, and $u \rightarrow Lu$ a continuous linear operator of $\mathcal{E}$ into $\mathcal{F}$. There always exists a uniquely defined continuous linear operator L^* of $\mathcal{F}$ into $\mathcal{E}$ such that for any u belonging to $\mathcal{E}$ and any v belonging to $\mathcal{F}$, the following equality between inner products holds

$$(Lu, v) = < u, L^*v > \tag{2.2}$$

L^* is called the <u>adjoint</u> of L. When $\mathcal{F}$ and $\mathcal{E}$ have finite dimensions N and M respectively, and are described by orthonormal coordinates, it is well known that the matrix which represents L^* is the transpose (or transconjugate in case of complex components) of the matrix (l_{ij}) which represents L. Equality (2.2) only corresponds in that case to a change in the order of summation indices

$$\sum_{j=1}^{N} \left[\sum_{i=1}^{M} l_{ji} u_i \right] v_j = \sum_{i=1}^{M} \left[\sum_{j=1}^{N} l_{ji} v_j \right] u_i$$

But again the concept of an adjoint operator is much more general, and we will use in the following the general abstract expression (2.2) rather than explicit indices.

$\mathcal{E}$, $\mathcal{F}$ and $\mathcal{J}$ still having the same significance, we now consider a (normally non-linear) differentiable function $u \rightarrow v = G(u)$ of $\mathcal{E}$ into $\mathcal{F}$. Through G, $\mathcal{J}(v) = \mathcal{J}[G(u)]$ is a compound function of u, and we consider the problem of numerically determining, for given u, the components of the gradient $\nabla_u \mathcal{J}$. In many cases (as in meteorological applications, where the operation $u \rightarrow G(u)$ will denote the temporal integration of a dynamical model of the atmospheric flow), it will be impossible to find a usable analytical expressions for the components of $\nabla_u \mathcal{J}$. One conceivable way for determining $\nabla_u \mathcal{J}$ would be to perturb in turn all components of u and, for each perturbation, to explicitly compute

$v = G(u)$ and the resulting perturbation on $\mathcal{J}$. This would lead to finite difference approximations for the components of $\nabla_u \mathcal{J}$, but would require as many explicit computations of G as there are components in u, an obviously impossible task for systems with large dimensions.

Now, the first-order variation of $\mathcal{J}$ is given by (2.1), while the first order variation δv of v is equal to

$$\delta v = G' \delta u \tag{2.3}$$

where G' is the linear operator obtained by differentiating G with respect to u. Carrying (2.3) into (2.1) and introducing the adjoint G'^* of G' leads to

$$\delta \mathcal{J} = \langle G'^* \nabla_v \mathcal{J} , \delta u \rangle$$

which, by the definition of a gradient, shows that the gradient $\nabla_u \mathcal{J}$ is equal to

$$\nabla_u \mathcal{J} = G'^* \nabla_v \mathcal{J} \tag{2.4}$$

Therefore, if $\nabla_v \mathcal{J}$ can be easily determined (which will be the case whenever $\mathcal{J}$ is an analytically "simple" function of v) and if a program is available which computes $G'^* w$ for given w, (2.4) defines a practical way for computing $\nabla_u \mathcal{J}$.

The computations which, starting from u, lead to G(u) being decomposed as the product of a number of elementary operations

$$G = C_n \ldots . C_2 \; C_1$$

Differentiation leads to

$$G' = C'_n \ldots . C'_2 \; C'_1$$

where, for each i, C'_i is the linear operator obtained by differentiating C_i. It is readily obtained from the adjointness relationship (2.2) that the adjoint of a product of operators is the product of the corresponding adjoints, taken in reverse order. Therefore

$$G'^* = C_1'^* C_2'^* \ldots . C_n'^*$$

which shows that, in the computation of G'^*w , the basic operations which make up G will have to be performed (after they have been linearized and their adjoints have been taken) in reverse order. In particular, if the computations which make up G include a temporal integration of a dynamical model, the corresponding adjoint computations will include some form of reversed time integration.

The computation of $\nabla_u \mathcal{J}$ for a given value of u will successively require the direct computation of v=G(u) , the determination of the corresponding value of $\nabla_v \mathcal{J}$ and the adjoint computation (2.4). In practice, the numerical cost of the adjoint computation will be of the same order of magnitude as the cost of the direct computation. The determination of one gradient $\nabla_u \mathcal{J}$ will therefore require a few times at most the cost of one direct computation of v=G(u). It is seen that considerable gain is achieved over the other way considered above for determining $\nabla_u \mathcal{J}$, namely explicit perturbations of the components of u.

3. APPLICATION TO ASSIMILATION OF METEOROLOGICAL OBSERVATIONS

The general principle described in the previous section will now be applied to the problem of assimilation of meteorological observations. The vector u will denote the set of initial conditions from which a dynamical model of the atmospheric flow is integrated. The model equations are written symbolically as

$$\frac{dx}{dt} = F(x) \tag{3.1}$$

where, for any time t, x(t) is the vector describing the state of the flow, which belongs to a space $\mathcal{E}$, with inner product < , >. Equation (3.1) will therefore be integrated from the initial condition $x(t_o) = u$. The vector v will denote the entire history $x(t)$, $t_o \leqq t \leqq t_1$, produced by the integration of (3.1). As for the scalar function $\mathcal{J}$, it will measure the "distance" between the solution x(t) and a set of available observations. For instance, $\mathcal{J}$ can be a weighted sum of squared difference between the model values and the observed values, which we write symbolically as

$$\mathcal{J} = \sum_i \alpha_i \left[x(\tau_i) - x_{obs}(\tau_i) \right]^2 \tag{3.2}$$

where the τ_i are the times at which observations have been performed. To (3.2) can be added, for example, a term measuring the amount of small scale noise in the solution x(t). The effect of this term will be to reduce the amount of noise in the minimization of $\mathcal{J}$. We will assume for $\mathcal{J}$ the general form

$$\mathcal{J}[x(t)] = \int_{t_o}^{t_1} H[x(t),t]\, dt \qquad (3.3)$$

where for any t, $x \rightarrow H[x,t]$ is a regular scalar function defined on $\mathcal{E}$.

The first-order variation $\delta\mathcal{J}$ resulting from a perturbation δu on the initial condition is equal to

$$\delta\mathcal{J} = \int_{t_o}^{t_1} < \nabla_x H(t),\ \delta x(t) > dt \qquad (3.4)$$

where for any t, the argument t in $\nabla_x H(t)$ means that this gradient is taken at point $[x(t),t]$, and where $\delta x(t)$ is the first-order variation resulting from δu through the integration of (3.1). The variation $\delta x(t)$ is obtained from δu by integration of the tangent linear equation

$$\frac{d\,\delta x}{dt} = F'_x(t)\ \delta x \qquad (3.5)$$

where $F'_x(t)$ is the operator obtained by differentiating F with respect to x, and taken at point x(t). Equation (3.5) being linear, its solution $\delta x(t)$ at time t is obtained from the initial perturbation δu at time t_o through a linear operator, called the resolvent of (3.5) between

t_o and t, which will be noted $R(t,t_o)$

$$\delta x(t) = R(t,t_o)\ \delta u \qquad (3.6)$$

More generally, the resolvent $R(t,t')$ is defined for any two instants t and t' and possesses the following properties

$$R(t,t) = I \qquad \text{for any } t \qquad (3.7\text{ a})$$

where I is the unit operator of $\mathcal{E}$, and

$$\frac{\partial}{\partial t} R(t,t') = F'_x(t)\ R(t,t') \qquad \text{for any } t \text{ and } t' \qquad (3.7\text{ b})$$

Carrying (3.6) into (3.4), introducing the adjoint $R^*(t,t_o)$ of $R(t,t_o)$ and taking the integral into the inner product leads to

$$\delta \mathcal{J} = < \int_{t_o}^{t_1} R^*(t,t_o)\ \nabla_x H(t),\ \delta u >$$

which, by the definition of a gradient, shows that the gradient of with respect to the initial condition u is equal to

$$\nabla_u \mathcal{J} = \int_{t_o}^{t_1} R^*(t,t_o)\ \nabla_x H(t)\ dt \qquad (3.8)$$

We introduce at this point the <u>adjoint tangent equation</u>

$$\frac{d\ \delta^* x}{dt} = -\ F'^*_x(t)\ \delta^* x \qquad (3.9)$$

whose variable $\delta^* x$ also belongs to $\mathcal{E}$, and where $F_x'^*(t)$ is, for any t, the adjoint of $F_x'(t)$. System (3.9) is linear, and we will denote $S(t,t')$ its resolvent between times t' and t. The inner product $< \delta x(t), \delta^* x(t) >$ of any two solutions of (3.5) and (3.9) is constant with time since

$$\frac{d}{dt} < \delta x(t), \delta^* x(t) > = < \frac{d\,\delta x}{dt}(t), \delta^* x(t) > + < \delta x(t), \frac{d\,\delta^* x}{dt}(t) >$$

$$= < F_x'(t)\,\delta x(t), \delta^* x(t) > + < \delta x(t), -F_x'^*(t)\,\delta^* x(t) >$$

$$= 0$$

Denoting by y and z any two elements of $\mathcal{E}$, the solution of the direct equation (3.5) defined by the condition $\delta x(t_o) = y$ takes at time t the value $R(t,t_o)y$, while the solution of the adjoint equation (3.9) defined by the condition $\delta^* x(t) = z$ takes at time t_o the value $S(t_o,t)z$. The corresponding equality between inner products reads

$$< y , S(t_o,t)z > = < R(t,t_o)y , z >$$

which, being valid for any y and z, shows that the adjoint of $R(t,t_o)$ is $S(t_o,t)$. Equality (3.8) accordingly becomes

$$\nabla_u \mathcal{J} = \int_{t_o}^{t_1} S(t_o,t)\,\nabla_x H(t)\,dt \tag{3.10}$$

We now consider the <u>inhomogeneous adjoint equation</u>

$$\frac{d\,\delta^* x}{dt} = -F_x'^*(t)\,\delta^* x - \nabla_x H(t) \tag{3.11}$$

The solution of that equation defined by the condition $\delta^*x(t_1) = 0$ is equal to

$$\delta^*x(t) = \int_t^{t_1} S(t,\tau)\ \nabla_x H(\tau)\ d\tau$$

as can easily be verified from the resolvent properties (3.7), applied to the resolvent S(t,t'). Equation (3.10) now shows that $\nabla_u \mathcal{J}$ is equal to $\delta^*(t_o)$.

In summary, the gradient $\nabla_u \mathcal{J}$ can be obtained, for a given value u of the initial condition, by performing the following operations

i) starting from $x(t_o) = u$, integrate the basic equation (3.1) from t_o to t_1. Store the values thus computed for $x(t)$, $t_o \leqq t \leqq t_1$.

ii) starting from $\delta^*x(t_1) = 0$, integrate the inhomogeneous adjoint equation (3.11) backwards in time from t_1 to t_o, the operator $F'^*_x(t)$ and the "forcing term" $\nabla_x H(t)$ being determined, at each time t, from the values x(t) computed in the direct integration of (3.1). The final value $\delta^*x(t_o)$ is the gradient $\nabla_u \mathcal{J}$.

In any practical situation, where the function F in (3.1) will contain spatial derivatives, equation (3.1) will be discretized with respect to both space and time. If one wants to adjust a solution of the discretized model to the observations, it is of course the adjoint of the discretized model which will have to be used. In particular, the derivation presented above, obtained under the implicit hypothesis of a continuous time coordinate, will have to be performed again for the particular temporal discretization used in the model. Similarly, the adjoint F'^*_x will depend on the particular scheme used for spatial discretization of F.

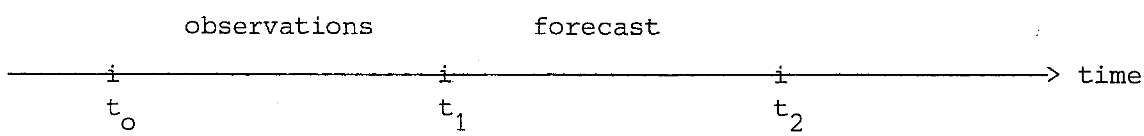
Once the gradient $\nabla_u \mathcal{J}$ is available, it will be used in a descent process in order to determine the minimizing u . A large number of descent algorithms have now been developed, and are available in routine libraries. We will mention the steepest descent algorithm, the conjugate gradient algorithm, and the quasi-Newton, or variable metric, algorithm. These algorithms are described, and their properties discussed, in Gill et al. (1982).

The approach to data assimilation which has just been described seems to have been first suggested by Penenko and Obraztsov (1976). More recently, it has been used by Lewis and Derber (priv. com.), and by Courtier (this volume) who has assimilated radiosonde observations of wind and geopotential with two different models, based respectively on the vorticity equation and the shallow-water equations.

4. APPLICATION TO THE DETERMINATION OF THE INITIAL AND BOUNDARY CONDITIONS OF A LIMITED AREA MODEL

A specific problem encountered when using a limited area model (LAM) for mesoscale numerical weather forecasting is the definition of appropriate lateral boundary conditions in the course of the temporal integration of the model. The procedure most commonly adopted at present is to extract those boundary conditions from a forecast previously produced by a large scale model(LSM). Now, if a large scale forecast is already available, there is no reason to trust it particularly along the boundary of the limited area model. By definition, the large scale forecast must be most reliable in the largest scales of the flow. So, instead of requiring coincidence between the two forecasts along the boundary of the LAM, one may wish to require coincidence, over the entire spatial domain of the LAM, of those spectral scales which are common to both models. This can be achieved, at least in principle, by using the adjoint equations of the LAM, as will be now shown.

We consider the following situation. We want to make a LAM forecast for some time interval (t_1, t_2), using observations available over some time interval (t_o, t_1) anterior to t_1, and a large scale forecast already available for the period (t_1, t_2) (see figure).



For each solution x(t) of the LAM, we define a functional

$$\mathcal{J} = \int_{t_o}^{t_1} P\left[x(t)\right] dt + \int_{t_1}^{t_2} Q\left[x(t)\right] dt$$

where, for t comprised between t_o and t_1, $P\left[x(t)\right]$ is a measure of the "distance" between x(t) and the observations available at that time and, for t comprised between t_1 and t_2, $Q\left[x(t)\right]$ is a measure of the discrepancy between x(t) and the LSM forecast over the largest scales resolved by the LAM. We want to determine the initial and boundary conditions of the LAM such that the corresponding solution minimize the functional $\mathcal{J}$.

We rewrite the function $\mathcal{J}$ as

$$\mathcal{J} = \int_{t_o}^{t_2} H\left[x(t)\right] dt$$

where $H\left[x(t)\right]$ stands for $P\left[x(t)\right]$ or $Q\left[x(t)\right]$ depending on whether $t < t_1$ or $t > t_1$. The evolution equation will be written as

$$\frac{dx}{dt} = F\left[x, c(t)\right]$$

where, for any t, c(t) denotes the boundary condition at that time. A solution x(t) over the time interval (t_o, t_2) is completely defined by the specification of the corresponding initial condition $x(t_o) = u$ and of the boundary condition c(t), for all t, $t_o \leqslant t \leqslant t_2$.

Our purpose is to determine the <u>gradient</u> of $\mathcal{J}$ with respect to the ensemble $(x(t_o), c(t), t_o \leqslant t \leqslant t_c)$. We will proceed along lines very similar to those already followed in the previous section. For a perturbation δu and $\delta c(t)$ $(t_o \leqslant t \leqslant t_2)$ of the initial and boundary conditions respectively, the corresponding first order variation $\delta\mathcal{J}$ of $\mathcal{J}$ is equal to

$$\delta\mathcal{J} = \int_{t_o}^{t_1} < \nabla_x H(t), \delta x(t) > dt \qquad (4.1)$$

where the perturbation $\delta x(t)$ is obtained by integration of the tangent linear equation which now reads

$$\frac{d\delta x}{dt} = F'_x(t)\delta x + F'_c(t)\delta c(t) \qquad (4.2)$$

The solution of (4.2) defined by the initial condition $\delta x(t_o) = \delta u$ is equal to

$$\delta x(t) = R(t,t_o)\ \delta u + \int_{t_o}^{t} R(t,\tau)\ F_c'(\tau)\delta c(\tau)\ d\tau \qquad (4.3)$$

where, as in the previous section, R denotes the resolvent of the tangent equation (3.5). Carrying (4.3) into (4.1), taking adjoints and changing the order of the integrations with respect to t and τ, one obtains

$$\delta \mathcal{J} = \langle \delta^* x(t_o),\ \delta u\rangle + \int_{t_o}^{t_2} \langle F_c'^*(t)\,\delta^* x(t), \delta c(t)\rangle\ dt \qquad (4.4)$$

where for any t, $\delta^* x(t) = \int_{t}^{t_2} R^*(t,\tau)\nabla_x H(\tau)\ d\tau$

is the solution at time t of the inhomogeneous adjoint equation (3.11), integrated from $\delta^* x(t_2)=0$. The first term of the right hand side of (4.4) is the same as in (3.10), and expresses that the gradient of $\mathcal{J}$ with respect to the initial condition u is equal to $\delta^* x(t_o)$. The second term expresses that, for any t, the gradient of $\mathcal{J}$ with respect to c(t) is equal to $F_c'^*(t)\delta^* x(t)$. In order to determine this gradient, we must therefore, in the course of the backward integration of the inhomogeneous adjoint (3.11), apply to the current value $\delta^* x(t)$ the adjoint of the operator $F_c'(t)$ which expresses the dependence of the dynamical evolution with respect to the boundary condition (see 4.2).

5. ADDITIONAL EXAMPLES

Given a scalar function $\mathcal{J}$ whose explicit computation requires the temporal integration of a dynamical model of the atmospheric flow, the determination of the gradient of $\mathcal{J}$ with respect to initial conditions, boundary conditions and/or physical parameters will always require the backward temporal integration of the corresponding adjoint model. The forcing term of the adjoint integration ($\nabla_x H(t)$ in the examples treated above) will depend on the particular function $\mathcal{J}$ which is considered. Similarly the result $\delta^* x(t)$ of the integration of the inhomogeneous adjoint will have to be treated differently according to whether one wants to know the gradient of $\mathcal{J}$ with respect to initial conditions (as in section 3), lateral boundary conditions (as in Section 4), or physical parameters (see e.g. Hall and Cacuci, 1984, for this last case). But the basic adjoint equation (3.9), which makes up the core of the complete adjoint computations leading to $\nabla\mathcal{J}$, will always be the same for a given numerical model.

We will mention a few of the many possible meteorological applications of adjoint methods.

1) Let us assume that a particular numerical forecast has produced some erroneous feature, such as for instance an abnormally low surface pressure in some region, and that one wants to trace back the origin of that error in the forecast initial conditions. Taking as function $\mathcal{J}$ the value of the surface pression at the time and location where it has been abnormally low, one can integrate the adjoint model in order to determine the gradient of $\mathcal{J}$ with respect to initial conditions (and possibly also with respect to other parameters). Although this will not

of course automatically identify the "cause" of the error in the forecast, it will certainly contain useful information as to its origin.

2) The second example is borrowed from Marchuk (1982). Let us suppose that industrial plants have to be located in some area, and that one is concerned with reducing the amount of resulting atmospheric pollution. A mesoscale circulation model is available which, given parameters such as the geographical locations of the plants, the amount of industrial effluents emitted by each plant into the atmosphere, etc, computes some measure $\mathcal{J}$ of the nuisance produced by the resulting pollution. The adjoint model can be used in order to determine the gradient of $\mathcal{J}$ with respect to those parameters. A minimization process can then be used in order to determine the parameter values which minimize $\mathcal{J}$ within some prescribed constraints.

3) Many observing system simulation experiments (OSSE) have been performed in order to assess the impact of modifications of observing systems on the quality of subsequent forecasts. Once an assimilation program and a forecasting model are available, it is possible, at least in principle, to take the adjoint, not only of the model itself, but also of the assimilation program (this is possible whether or not the assimilation program uses the adjoint of the forecasting model). The two adjoints, combined in reverse order, make up together the adjoint of the global assimilation-forecasting procedure. This global adjoint can then be used in order to determine the gradient of some quantity measuring the quality of the forecast with respect to the parameters defining the accuracy, nature and spatio-temporal distribution of the observations.

Other examples of possible applications of adjoint equations are given in Kontarev (1980) and Hall and Cacuci (1983, 1984) and in articles referenced therein.

6. CONCLUDING REMARKS

The various examples discussed in this article show that adjoint techniques can potentially be applied to a large variety of meteorological problems. Some additional remarks are in order.

It has been emphasized in Section 2 that using adjoint equations for computing the gradient of a scalar function achieves considerable numerical gain over direct perturbations of the arguments of that function. More generally, after one integration of the basic model equation (3.1) has been performed, one integration of the adjoint equation will produce the gradient of one output parameter with respect to all input parameters (initial and boundary conditions, physical parameters) of the model. Direct perturbation of one input parameter, on the other hand, will determine the sensitivity of all output parameters with respect to that particular input parameter. Therefore, if one is interested in studying the sensitivity of p output parameters of the model with respect to n input parameters, it will be more economical to use the adjoint equations if $p < n$, while it will be more economical to perform direct perturbations if $p > n$.

It is possible to integrate the adjoint model only if a direct integration of the basic model has been performed previously. This direct integration will produce the components of the terms such as $F'^{*}_{x}(t)$, $\nabla_x H(t)$, which are required by the adjoint integration. The gradient produced by an adjoint integration is therefore local in the space of input parameters of the model. This must be clearly realized since the localness of the gradient, depending on the particular problem at hand, may be an advantage (as in minimization procedures for instance) or on the contrary a disadvantage (as possibly in some climatological studies, for which a gradient varying rapidly in parameter space may not be significant).

Now, if adjoint techniques make it possible to perform computations which it would be absolutely unthinkable to perform otherwise, their computational cost remains high. Not only does the determination of one gradient require one integration of the adjoint model, but that integration itself requires in principle the previous storage of the complete model history produced by the direct integration. At present, adjoint techniques could possibly be used only for applications which are not submitted to strict limitations in computing time. This probably excludes, for the time being, any operational application. But one can be sure that the continuous progress which has been observed in computing power in the last decades will continue in the coming years, and detailed study of adjoint methods, from both a theoretical and a practical point of view, is certainly of great interest.

7. ACKNOWLEDGMENTS

The author thanks Dr. François-Xavier Le Dimet, who first pointed out clearly to him how adjoint equations could be used for data assimilation, and stimulated his further interest in adjoint techniques.

8. REFERENCES

Gill, P.E., W. Murray and M.H. Wright, 1982 : Practical optimization, London, Academic Press, 401 pp.

Hall, C.G. and D.G. Cacuci, 1984 : Systematic analysis of climatic model sensitivity to parameters and processes, in Climatic Processes and Climate Sensitivity, Washington D.C., American Geophysical Union, 171-179.

Hall, C.G. and D.G. Cacuci, 1983 : Physical interpretation of the adjoint functions for sensitivity analysis of atmospheric models, J. Atmos. Sci., 40, 2537-2546.

Kontarev, G., 1980 : The adjoint equation technique applied to meteorological problems. ECMWF Tech. Ref. N°21, 21 pp.

Marchuk, G.I., 1974 : Numerical solution of the problems of the dynamics of the atmosphere and ocean (in Russian) Leningrad, Gidrometeoizdat, 303 pp.

Marchuk, G.I., 1982 : Mathematical issues of industrial effluent optimization. J. Met. Soc. Japan, 60, 481-485.

Penenko, V.V. and Obraztsov, N.N., 1976 : A variational initialization method for the fields of the meteorological elements. Soviet Meteorology and Hydrology, N° 11, 1-11.