View metadata, citation and similar papers at core.ac.uk

brought to you by CORE

provided by Research Papers in Economics

Marine Resource Economics, Volume 17, pp. 35–46
0738-1360/00 $3.00 + .00
Printed in the U.S.A. All rights reserved
Copyright © 2002 Marine Resources Foundation

On the Dynamics of Commercial Fishing and Parameter Identification

AL-AMIN M. USSIF
LEIF K. SANDAL
Norwegian School of Economics and Business Administration

STEIN I. STEINSHAMN
Institute for Research in Economics and Business Administration

Abstract *This paper has two main objectives. The first is to develop a dynamic model of commercial fisheries different from most existing models that assume optimizing behavior. The industry is assumed to have a well-defined index of performance. Based upon this index, the decision to invest or not is made. We do not, however, assume that the industry or firm is efficient or optimal in its operations. The second is that a new approach of fitting model dynamics to time series data is employed to simultaneously estimate the poorly known initial conditions and parameters of nonlinear fisheries dynamics. The approach is a data assimilation technique known as the variational adjoint method. Estimation of the poorly known initial conditions is one of the attractive features of the variational adjoint method.*

Key words Data assimilation, index of performance, nonlinear dynamics, open access, variational adjoint method.

Introduction

The most common approaches of modeling the dynamics of a natural resource system are the routine application of the sophisticated techniques of the calculus of variations or optimal control theory and dynamic programming (Kamien and Schwartz 1984; Clark 1990). The economic theory of an optimally managed fishery has been advanced by several researchers. For example, Clark (1990) discussed various models in some detail. Sandal and Steinshamn (1997a,b,c) made some of the most recent contributions in this area. This framework explicitly assumes that agents

Dr. Al-Amin M. Ussif is research associate at the Norwegian School of Economics and Business Administration, Bergen, Norway, email: aussif@hotmail.com. Dr. Leif K. Sandal is an associate professor at the Norwegian School of Economics and Business Administration, Bergen, Norway, email: leif.sandal@nhh.no. Dr. Stein I. Steinshamn is researcher at the Institute for Research in Economics and Business Administration, Bergen Norway, email: Stein.steinshamn@snf.no.
Financial support from Norwegian School of Economics and Business Administration and Institute for Research in Economics and Business Administration is appreciated.
We want to thank professor J.J. O'Brien of FSU Tallahassee, Florida, for his hospitality. We also wish to thank professor Rognvaldur Hannesson, professor Gerhard Berge, professor Jon Conrad, and Dr. Steve Morey for reading this work, and Drs. Gilbert and Lemarechal of INRIA in France for letting us use their optimization routine. Financial support from the Norwegian School of Economics and Business Administration and Institute for Research in Economics and Business Administration is appreciated. Finally, we thank the two referees whose comments and suggestions helped to improve this manuscript.

are optimal and efficient. However, most real-world fisheries have historically not been optimally managed.

Commercial models of fisheries have previously been discussed by Crutchfield and Zellner (1962) and Smith (1969). The latter provided a model of theoretical nature, which transforms specific patterns of assumptions about cost conditions, demand externalities, and biomass growth technology into a pattern of exploitation of the stock. Smith also discussed the three main features of commercial fishing and mentioned the various types of external effects representing external diseconomies to the industry. In two earlier papers, Gordon (1954) and Scott (1955) noted that all of these externalities arise fundamentally because of the unappropriated "common property" character of most ocean fisheries (Smith 1969).

In this paper, we develop a commercial fishing model that does not necessarily assume optimal behavior of fishers. The goal is to develop a model that is quite general and has potentially broader applications.

This paper also focuses on a very important aspect of fisheries management that has largely been ignored. Deacon *et al.* (1998) noted that much of the information managers need is empirical; *i.e.*, measurements of vital relationships and judgments about various impacts. This area of the economics of fishing has not been adequately explored by economists, probably due to lack of data and computational power in the past. Much of the research effort was used in search of qualitative answers to management problems. This paper employs an efficient method of advanced data assimilation, known as the variational adjoint technique, to analyze real fisheries data (Smedstad and O'Brien 1991). In data assimilation, mathematical or numerical models are merged with observational data in order to improve the model or its predictions. The former application is known as model fitting. Using the variational adjoint technique, appropriate initial conditions and parameters of the nonlinear fisheries dynamics are estimated. Nonlinear fisheries dynamics are highly sensitive to the initial conditions and the parameters, which are often exogenously given inputs to the system. These inputs are very crucial in simulation studies. Researchers who conduct simulation analysis often rely on previous estimates for most of their input parameters. This work will provide some of these estimates for the North East Arctic cod, which may enhance future simulation results.

Data assimilation problems are often ill-posed; *i.e.*, they are characterized by nonuniqueness and instability of the identified parameters (Yeh 1986). Thus, it may be worthwhile to search for best initial and/or boundary conditions when using these models in analysis. The reader may have noticed that this approach has major advantages compared to conventional methods. It allows us to estimate initial conditions of the model dynamics as additional control variables on equal footing as the model parameters. Thus, treating the initial biomass level and the initial harvest rate as uncertain inputs in the system. Most recent models and traditional approaches consider the initial biomass and harvest amounts as known and deterministic. It also provides a more efficient way of calculating the gradients of the cost or loss function with respect to the control variables compared with the finite difference approach. To enhance the results of the data assimilation, we provide bootstrap estimates of the standard errors (Efron and Tibshirani 1993). The bootstrap is a computer-based method for estimating standard errors.

The structure of the remainder of the paper is as follows. The next section is a detailed discussion of the dynamics of the commercial fishing model. It presents a more general model without assuming any optimizing behavior. In the following section, we briefly discuss data assimilation, and some basic concepts of the techniques are defined. All technical details appear in the Appendix. The final section is an application to the North East Arctic cod stock (NEACS). It discusses the results and summarizes the work.

Dynamics of Commercial Fishing

The dynamics of the fishing industry are developed and discussed in detail in this section. The population dynamics are described by:

$$\frac{dx}{dt} = f(x) - y, \tag{1}$$

where x is the biomass and y is the harvest in weight.[1] The growth of the stock is represented by $f(\cdot)$. Several forms of the growth model exist. For some species, the empirical law of growth is asymmetric. In this paper, however, we will use the logistic growth law. The Schaefer logistic function takes the form $f = rx(1 - x/K)$, where r is the intrinsic growth rate, and K is the carrying capacity of the biological species if the population is not exploited. It is symmetric about $K/2$ and has the following properties, $f(0) = f(K) = 0$, $f(k/2) = \max f$ (Schaefer 1967).

To model the fishing industry, we define the following relationship between the rate of increase or decrease of the exploitation of the fish biomass, y, and a function, $\Phi(x,y)$, such that:

$$\frac{dy}{dt} = \gamma \; y\, \Phi(x, y), \tag{2}$$

where γ is an adjustment parameter, and Φ is a certain well-defined value function to be discussed shortly. The constant of proportionality reflects the rate at which capital is being put in or removed from the industry or firm. The adjustment parameter has been discussed in detail in Smith (1969). For instance, if Φ is positive one may expect an increase in capital investment in the fishery and a decrease otherwise. The function defined by Φ can take different parametric forms reflecting our hypotheses about the operation of the industry. It may represent short- or long-run average costs of fishing vessels, the marginal or average net revenues of a firm, *etc.*

Let p be the unit price of fish and c be a proportionality constant (see Clark 1990). Assume that costs of fishing are linear in the harvest and inversely related to the stock. Then, the average net revenue is given by:

$$\Phi(x, y) = p - c/x\,. \tag{3}$$

The average cost of harvesting is assumed to depend explicitly on the size of the stock. This takes into account the stock externalities; *i.e.*, fishing costs decrease as the fish population increases. Note that the total costs are given by the function cy/x, which is nonlinear in stock and linear in harvest and reflects the fact the cod is a dermersal species. Given a set of observations on the costs of harvest, it is possible to estimate c for this specification. An implicit assumption made in this model is that the fishery is one of many sources in the market that characterizes an open-access regime consistent with the period of interest (1946 through 1977). In other words, the market is competitive. Otherwise, price should depend on yield; that is, $p = p(y)$ where $p(\cdot)$ is the inverse demand function.

---

[1] Gordon (1954) assumed that harvest depends on stock and efforts, while Smith (1969) assumed that harvest is a function of the number of identical firms and the stock. Here, no such disaggregation has been made.

From the previous definitions of Φ and the industry model, equation (2), it is obvious that the rate of harvesting from the stock for the industry is perceived to vary in proportion to the net revenue; that is, the difference between total revenues and total costs. Put another way, the output growth rate $\dot{y}/y$ of the industry is proportional to the average or marginal net revenues. Thus, it may be perceived that once the stock is increasing, profit margins will also increase, inducing firms to increase production. Substituting equation (3) in equation (2) and combining with the population dynamics model, equation (3), the industry dynamics model is derived. This system of equations, (1)–(2), constitutes a coupled nonlinear ordinary differential equations (ODEs):

$$\frac{dx}{dt} = f(x, y, K) - y \tag{4}$$

$$\frac{dy}{dt} = \gamma\,(p - c/x)\,y$$

The term $(p - c/x)y$ is the annual total profit (total revenues minus total costs). Owing to the linearity of the net revenue in the harvest, the average net revenue is equal to the marginal net revenue.

Incorporated in the model are the hypotheses about the costs and revenues. If the firms are optimizers, they should at least operate at a level where average or marginal profits are positive. In the construction of such a behavioral model, an implicit assumption about the harvest rate being proportional to the number of firms or fishing vessels is made (Smith 1969).

The system of equations contains these input parameters (r, K, γ, p, c). By redefining the parameters; *i.e.*, $\theta = \gamma p$ and $\tau = \gamma c$, we now have these parameters (r, K, θ, τ) to estimate. Note that a more parsimonious model is achieved, thus reducing the chances of overfitting the model. Also, no data on costs and prices are necessary in order to fit the model.

Notice that Gordon (1954) assumed that harvest depends on stock and efforts, while Smith (1969) assumed that harvest is a function of the number of identical firms and the stock. Here, no such disaggregation has been made. The reason being that catch data are more readily available than that on the number of firms, and it may be more suitable for this analysis. In following section, we shall focus on the main goal of this paper, which is the introduction and application of data assimilation in bioeconomics. The technique will be presented and an application to a real-world fishery will also be demonstrated.

Data Assimilation Methods

Data assimilation methods have been used extensively in meteorology and oceanography to estimate the variables of model dynamics and/or the initial and boundary conditions. These methods include the sequential techniques of Kalman filtering (Kalman 1960) and the variational approach (Bennett 1992). The variational adjoint method has been proposed as a tool for estimation of model parameters. It has since proven to be a powerful tool for fitting dynamic models to data (Smedstad and O'Brien 1991). The methods have recently been used to estimate parameters of the predator-prey equation (Lawson *et al.* 1995) and also some high-dimensional ecosystem models (Spitz *et al.* 1998; Matear 1995). The basic idea is that, given a set of observations, a solution of the model that is as close as possible to the data is sought by adjusting the model parameters, such as the initial conditions. The variational method is an automation of the brute force approach, where the model is manually tuned to fit the data. This becomes tedious very quickly, even with a single parameter.

The variational adjoint method has three parts: the forward model, the backward model derived via the Lagrange multipliers, and an optimization procedure.[2] These components and all of the mathematical derivations are discussed in the Appendix. An outline of the technique is also presented for those who may be interested in learning the advanced and efficient method of data analysis.

An Application to the North East Arctic Cod Stock (NEACS)

The commercial fishing model developed in this paper is applied to the NEACS. The fishery has a long history of supporting a large portion of the Norwegian and Russian coastal populations. Data on catches and estimated stock biomass have been collected since immediately after World War II and are published by the International Council for the Exploration of the Sea (ICES) (Anon. 1998). Different techniques for stock assessments exist in fisheries management. The stock size for the NEACS is measured using the statistical Virtual Population Analysis (VPA) method (Anon. 1998). As a result, the catch data and biomass estimates obtained by the VPA runs may be somewhat correlated. This issue will not be dealt with in this paper. From this point forward, actual or real data will refer to the measurements from the ICES 1998 report. See Anon. (1998) for details of the data. The data used in this work are aggregates of total catch and total estimated biomass.

The history of this fishery is similar to other commercial fisheries around the world. It has been managed based on the common policy of maximum sustainable yield (MSY), which is the most commonly employed method over the majority of the last century. The historical data show a decreasing trend for both the stock biomass and the yield (figure 1). Note that the vertical scales are different in figure 1 for the stock and harvest. It is also observed that the data fluctuate a great deal, which depicts the inherent stochastic feature of a fishery resource. The data available on NEACS date back to 1946 (Anon. 1998). It is, however, intuitive to divide the period into pre-quota (1946–77) and quota (1978–96) periods, which represent different management regimes. The first period may be dubbed the open-access period, and the second the regulated, open-access (total allowable catch [TAC]) period. We will apply our model to analyze the data for the first period. To analyze the second period, additional constraints, such as quota restrictions and minimum safe biomass levels, which reflect the regulations imposed by the management authorities are required (Homans and Wilen 1997). We shall, however, concern ourselves with the first period.

In this study, we combine the nonlinear dynamics model developed in the preceding section and the time series of observations to analyze the NEACS. The technique provides a novel and highly efficient procedure for data analysis. Model initial conditions, as well as parameters of the dynamics, are estimated using the variational adjoint method. First, artificial data generated from the model itself using known initial conditions and known parameters were used to test the performance of the adjoint code.

All the parameters were recovered to within the accuracy of machine precision. Both clean and noisy data were used to first study the model. The results are not shown in this paper.

[2] See Lawson *et al.* (1995) for a straightforward derivation of the adjoint system.

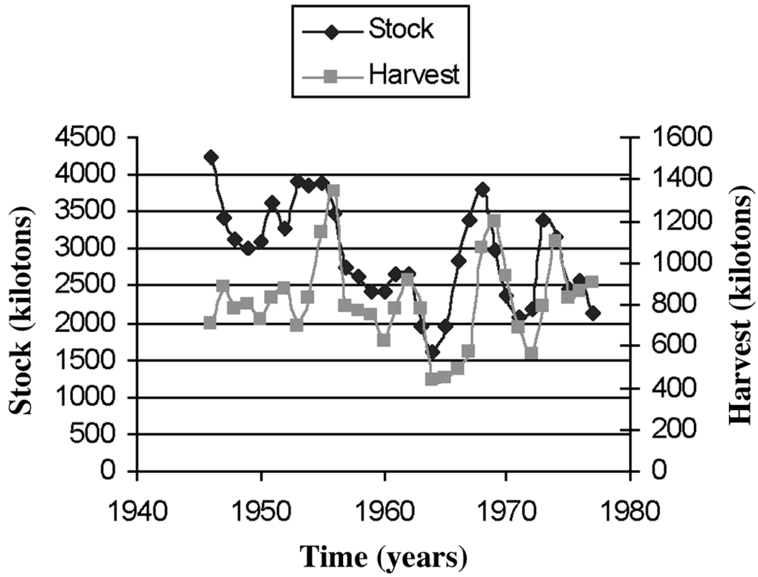

Figure 1. Graph Showing the Time Trends in Estimated Stock and Actual Harvest for the North East Arctic Cod Stock

Source: These data are aggregates of stock and harvest and are taken from International Council for the Exploration of the Sea (Anon. 1998).

To enhance the results of the data assimilation method, we provide standard errors of the estimates using the bootstrap technique. The bootstrap is a computer-based method for estimating standard errors (Efron and Tibshirani 1993). The algorithm is described below.

Let:

$$\hat{x}_t = g(\beta, x_t) + e_t$$

1. Estimate β using the variational method and calculate the residuals $\hat{e}_t$.
2. Draw $N\ e_t^b$'s from $\{e_1, \ldots, e_N\}$ with replacement and compute $\hat{x}_t{}^b = g(\beta; x_t) + e^b{}_t$.
3. Estimate β^b using the bootstrap sample data x_t^b.
4. Repeat steps 2 and 3, B times and calculate the standard errors.

where, x^b is the bootstrap sample, β^b is the bootstrap replication, and B is the number of replications. For this paper, 3,000 replicates (B = 3,000) were estimated.

Next, real data were used to estimate the initial conditions and all the parameters of the model dynamics. Starting from the best guesses of the parameters, the optimization procedure uses the gradient information to find the optimal initial conditions and the model parameters that minimize the penalty function. Note that the observed initial values were taken as the best guesses, and the weights were all set equal to unity.

Table 1
Model Parameter Estimates and the Bootstrap Standard Errors

Parameters	Estimates	(Standard Error)
r	0.4410	(0.2040)
K	5,150.40	(222.43)
θ	0.1863	(0.0984)
τ	738.80	(108.12)
x_0	3,687.11	(573.39)
y_0	902.83	(209.78)

Note: The weights are all set to unity.

Estimates of initial conditions, the parameters, and the bootstrap standard errors are tabulated below (table 1). The units are r per year, x_0, y_0, and K in kilotons.

All the estimated parameters are reasonable and as expected. This is because they all have the right signs; *i.e.*, they are positive. The intrinsic growth rate and the carrying capacity are both within an acceptable range. Note that the carrying capacity is about 20% above the initial stock estimate available to us; *i.e.*, the 1946 estimate (4.2 million tons). Also, statistical tests of significance at the 5% level on all parameters indicate that they are different from zero. These estimates should provide prior guesses for the important parameters (r, K) in the growth model for future research. They may be used as inputs by researchers who develop simulation based models for the NEACS. Since the focus of this paper is the application of the dynamic parameter estimation, we will leave that issue to future research.

Summary and Conclusion

This paper has addressed two major questions in bioeconomic analysis and fisheries management. It developed a dynamic fisheries model in a way that is rare in the literature and employs a powerful approach of efficiently combining the model with available observations collected over a given time domain. The variational adjoint method is used to simultaneously estimate the initial conditions and input parameters of the commercial fishing model.

It is observed that the technique used in this paper has an added virtue compared to the conventional ones used in the literature. Initial conditions of the model dynamics are estimated on equal footing as the model parameters. It is highly versatile in that it enables researchers to include as much information as is available.

The estimates were all reasonable and as expected for the NEACS. Estimates of the intrinsic growth and carrying capacity are both within acceptable ranges and statistically significant. However, caution must be exercised when interpreting the results due to the deterministic nature of the model and the large measurement errors in the data. It has been demonstrated here that dynamic resource models can be combined with real data in order to obtain useful insights about real fisheries. Biological parameters, such as the carrying capacity and economic parameters entering the objective functions of the industry, are identified. These, again, can be used for dynamic optimization in order to improve the economic performance of the fishery. The variational adjoint method has proven to be very promising and deserves further research efforts—not only in resource economics, but economics in general.

References

Anon. 1998. Report of the Arctic Fisheries Working Group, 1999/ACFM:3, ICES, Copenhagen.

Bennett, A.F. 1992. *Inverse Methods in Physical Oceanography*. Cambridge: Cambridge University Press.

Clark, C.W. 1990. *Mathematical Bioeconomics. The Optimal Management of Renewable Resources*. New York: Wiley and Sons.

Crutchfield, J.A., and A. Zellner. 1962. *Economic Aspects of the Pacific Halibut Fishery. Fishery Industrial Research 1(1)*. Washington, DC: US Department of the Interior.

Deacon, R.T., D.S. Brookshire, A.C. Fisher, A.V. Kneese, C.D. Kolstad, D. Scrogin, V.K. Smith, M. Ward, and J. Wilen. 1998. Research Trends and Opportunities in Environmental and Natural Resource Economics. *Journal of Environmental Economics and Management* 11(3-4):383–97.

Efron, B., and R.J. Tibshirani. 1993. *An Introduction to the Bootstrap*. Chapman and Hall.

Gilbert, J.C., and C. Lemarechal. 1991. Some Numerical Experiments with Variable-storage Quasi-newton Algorithms. *Mathematical Programming* 45:405–35.

Gordon, H.S. 1954. The Economic Theory of Common Property Resources. *Journal of Political Economy* LXII (2):124–42.

Harmon, R., and P. Challenor. 1997. Markov Chain Monte Carlo Method for Estimation and Assimilation into Models. *Ecological Modeling* 101:41–59.

Homans, F., and J. Wilen. 1997. A Model of Regulated Open Access Resource Use. *Journal of Environmental Economics and Management* 32(1):1–21.

Kalman, R.E. 1960. A New Approach to Linear Filter and Prediction Problem. *Journal of Basic Engineering* 82:35–45.

Kamien, M.I., and N.L. Schwartz. 1980. *Dynamic Optimization. The Calculus of Variation and Optimal Control in Economics and Management Second Edition*. Amsterdam: North-Holland.

Lawson, L.M., H.Y. Spitz, E.E. Hofmann, and R.B. Long. 1995. A Data Assimilation Technique Applied to Predator-Prey Model. *Bulletin of Mathematical Biology* 57:593–617.

Luenberger, D.C. 1984. *Linear and Nonlinear Programming*. Reading: Addison-Wesley.

Matear, R.J. 1995. Parameter Optimization and Analysis of Ecosystem Models Using Simulated Annealing: A Case Study at Station P. *J. Mar. Res*. 53:571–607.

Sandal, L.K., and S.I. Steinshamn. 1997a. A Stochastic Feedback Model for the Optimal Management of Renewable Resources. *Natural Resource Modeling* 10(1):31–52.

___. 1997b. A Feedback Model for the Optimal Management of Renewable Natural Capital Stocks. *Canadian Journal of Fisheries and Aquatic Sciences* 54:2475–82.

___. 1997c. Optimal Steady States and Effects of Discounting. *Marine Resource Economics* 12(2):95–105.

Schaefer, M.B. 1967. Fisheries Dynamics and the Present Status of the Yellow Fin Tuna Population of the Eastern Pacific Ocean. *Bulletin of the Inter-American Tropical Tuna Commission* 1:25–56.

Scott, A. 1955. The Fishery: The Objectives of Sole Ownership. *Journal of Political Economy*. LXIII(2):116–24.

Smedstad, O.M., and J.J. O’Brien. 1991. Variational Data Assimilation and Parameter Estimation in an Equatorial Pacific Ocean Model. *Progress in Oceanography* 26(10):179–241.

Smith, V.L. 1969. On Models of Commercial Fishing. *Journal of Political Economy* 77:181–98.

Spitz, H.Y., J.R. Moisan, M.R. Abbott, and J.G. Richman. 1998. Data Assimilation and a Pelagic Ecosystem Model: Parameterization Using Time Series Observations. *Journal of Marine Systems* 16:51–68.

Yeh, W.W-G. 1986. Review of Parameter Identification Procedures in Groundwater Hydrology: The Inverse Problem. *Water Resource Research* 22:95–108.

Appendix

Data Assimilation — Background

This section formulates the parameter estimation problem and presents the mathematical aspects of the variational adjoint technique. Numerical issues have also been briefly discussed.

The Model and the Data

For the sake of mathematical convenience, we use compact notation to represent the model dynamics as:

$$\frac{dX}{dt} = F(X, Q) \tag{A1}$$

$$X(0) = X_0 + \hat{X}_0 \tag{A2}$$

$$Q = Q_0 + \hat{Q} \tag{A3}$$

$$\hat{X} = X + \varepsilon, \tag{A4}$$

where $X = (x,y)$ is the state vector, X_0 is the best guess initial condition vector, $\hat{X}$ is the measurement vector, X_0 is the vector of initial misfits, Q is a vector of parameters, and $\hat{Q}$ is the vector of parameter misfits. The dynamics are assumed to exactly satisfy the constraints, while the inputs (the initial conditions and the parameters) are poorly known.

The Loss or Penalty Function

In variational adjoint parameter estimation, a loss function which measures the difference between the data and the model equivalent of the data is minimized by tuning the control variables of the dynamical system. The goal is to find the parameters of the model that lead to model predictions that are as close as possible to the data. A typical penalty function takes the more general form:

$$J[X, Q] = {}^{1}\!/_{2T_f} \int_0^{T_f} (Q - Q_0)^T W_Q (Q - Q_0)\,dt \tag{A5}$$

$$+ \frac{1}{2T_f} \int_0^{T_f} \left[X(0) - X_0\right]^T W \left[X(0) - X_0\right] dt + \frac{1}{2} \int_0^{T_f} (X - \hat{X})^T W (X - \hat{X})\,dt,$$

where the period of assimilation is denoted by T_f and T is the matrix transpose operator. The W's are the weight matrices, which are optimally the inverses of the error covariances of the observations. They are assumed to be positive, definite, and symmetric. The first and second terms in the penalty function represent our prior knowledge of the parameters and the initial conditions, and ensure that the estimated values are not too far away from the first guesses. They may also enhance the curvature of the loss function by contributing positive terms to the Hessian of J (Smedstad and O'Brien 1991). The variational adjoint technique determines an optimal solution by minimizing the loss function J, which measures the discrepancy between the model predictions and the observations. Note that the loss is defined in a more general form. However, for this application, we will assume that all the weighting matrices are identity matrices. This formulation is the ordinary least squares and conforms to the assumption that the errors are white noise. Hence, our estimates are equivalent to the maximum likelihood estimates.

The constrained problem can efficiently be solved by transforming it into an unconstrained optimization problem (Luenberger 1984). Several algorithms for solving the unconstrained nonlinear programming problem are available (Smedstad and O'Brien 1991). Statistical methods, such as the simulated annealing (Matear 1995) and the Markov Chain Monte Carlo (MCMC) (Harmon and Challenor 1997), have recently been proposed as tools for parameter estimation. The most widely used methods are the classical iterative methods, such as the gradient descent and the Newton's methods (Luenberger 1984).

The Variational Adjoint Method

Construction of the adjoint code is identified as the most difficult aspect of the data assimilation technique (Spitz *et al.* 1998). One approach consists of deriving the continuous adjoint equations and then discretizing them (Smedstad and O'Brien 1991). Another approach is to derive the adjoint code directly from the model code (Lawson *et al.* 1995; Spitz *et al.* 1998).

To illustrate the mathematical derivation, we use the first approach. Formulating the Lagrange function, L, by appending the model dynamics as strong constraints, we have:

$$L[X, Q] = J + \tfrac{1}{2} \int_0^{T_f} M \frac{dF}{dX} \, dt, \tag{A6}$$

where M is a vector of Lagrange multipliers, which is computed in determining the best fit. The original, constrained problem is thus reformulated as an unconstrained problem. At the unconstrained minimum, the first order conditions are:

$$\frac{dL}{dX} = 0 \tag{A7}$$

$$\frac{dL}{dM} = 0 \tag{A8}$$

$$\frac{dL}{dQ} = 0. \tag{A9}$$

It is observed that equation (A7) results in the adjoint or backward model, equation (A8) recovers the model equations, while (A9) gives the gradients with respect to the control variables. Using calculus of variations or optimal control theory, the adjoint equation is derived by forming the Lagrange function via the undetermined multipliers $M(t)$. The Lagrange function is:

$$L = J + \int_0^{T_f} M\left[\frac{\partial X}{\partial t} - F(X, Q)\right] dt, \tag{A10}$$

perturbing the function L (A11):

$$L[X + \partial X, Q] = J + \frac{1}{2}\int_0^{T_f} M\left[\frac{\partial (X + \delta X)}{\partial t} - F(X + \delta X, Q)\right] dt, \tag{A11}$$

which implies:

$$L[X + \delta X, Q] = J + \Delta_X J \delta X^T + \int_0^{T_f} M\left[\frac{\partial X}{\partial t} - F(X, Q)\right] dt \tag{A12}$$

$$- 2\int_0^{T_f} M\left(\frac{\partial \delta X}{\partial t} - \frac{\partial F}{\partial X}\delta X^T\right) dt + O(\delta X^2).$$

Taking the difference ($L[X + \delta X, Q] - L[X,Q]$):

$$\partial L = \Delta_X J\, \delta X^T - 2\int_0^{T_f} M\left(\frac{\partial \delta X}{\partial t} - \frac{\partial F}{\partial X}\delta X^T\right) dt + O(\delta X^2). \tag{A13}$$

Requiring that ΔL be of order $O(\delta X^2)$ implies:

$$\Delta_X J \delta X^T - \int_0^{T_f} M\left(\frac{\partial \delta X}{\partial t} - \frac{\partial F}{\partial X}\delta X^T\right) dt = 0. \tag{A14}$$

By integrating the second term of the LHS by parts and rearranging, we have:

$$\frac{\partial M}{\partial t} + \left[\frac{\partial F}{\partial X}\right]^T M = W(X - \hat{X}) \tag{A15}$$

$$M(T_f) = 0.$$

which is the adjoint equation together with the boundary conditions and from equation (A8), the gradient relation is:

$$\Delta_Q J = -\int_0^{T_0} M \frac{dF}{dQ}\, dt + W_Q(Q - Q_0). \tag{A16}$$

The term on the RHS in equation (A15) is the weighted misfit, which acts as a forcing term for the adjoint equation. It is worth noting here that we have implicitly assumed that data is continuously available throughout the integration interval. Equations (A7) and (A8), above, constitute the Euler-Lagrange (E-L) system and form a two-point boundary value problem.

The algorithm is outlined below:

- Choose the first guess for the control parameters.
- Integrate the forward model over the assimilation interval.
- Calculate the misfits and, hence, the loss function.
- Integrate the adjoint equation backward in time forced by the data misfits.
- Calculate the gradient of L with respect to the control variables.
- Use the gradient in a descent algorithm to find an improved estimate of the control parameters, which makes the loss function move towards a minimum.
- Check if the solution is found based on a certain criterion (for example, $J \le \delta$, $\Delta J \le \delta$ may be appropriate convergence criteria).
- If the criterion is not met, repeat the procedure until a satisfactory solution is found.

The optimization step is performed using standard optimization procedures. In this paper, a limited memory quasi-Newton procedure is used (Gilbert and Lemarechal 1991).

The success of the optimization depends crucially on the accuracy of the computed gradients. Any errors introduced while calculating the gradients can be detrimental and the results misleading. To avoid this, it is always advisable to verify the correctness of the gradients (Smedstad and O'Brien 1991; Spitz *et al.* 1998).