



Parameter Identification in Tidal Models with Uncertain Boundaries*

ARUNABHA BAGCHI† and PAUL TEN BRUMMELHUIS‡

A simultaneous estimation of states and unknown parameters by the method of maximum likelihood is used for calibrating large-scale tidal models with uncertain boundary conditions.

Key Words—Maximum likelihood; modeling uncertain boundaries; tidal models.

Abstract—In this paper we consider a simultaneous state and parameter estimation procedure for tidal models with random inputs, which is formulated as a minimization problem. It is assumed that some model parameters are unknown and that the random noise inputs only act upon the open boundaries. The hyperbolic nature of the governing dynamical equations is exploited in order to determine the smoothed states efficiently. This enables us to also apply the procedure to nonlinear tidal models without an excessive computational load. The main aspects of this paper are that the method of Chavent (*Identification and System Parameter Estimation. Proc. 5th IFAC Symp.* Pergamon, Oxford, pp 85–97, 1979), used to calculate the gradient of a criterion that is to be minimized, is now embedded in a stochastic environment and that the estimation method can also be applied to practical, large-scale problems.

INTRODUCTION

TIDAL MODELS are used to describe the water movement in a restricted area, enforced by a tidal wave that enters this area through one of its open boundaries. The propagation of the wave is governed by the dynamical equations, whereas the driving tidal force is represented by the open boundary condition(s). Examples of such models are a model that describes the flow in a tidal estuary in the southwestern part of the Netherlands (Fig. 1) and CSM-16, a model of the continental shelf of the northwestern part of Europe (Fig. 2). Although a lot of effort has

been taken to increase the accuracy of these deterministic models, there still remain uncertainties in this representation. If these uncertainties are described by means of stochastic noise processes, data assimilation techniques can be applied in order to find solutions of the stochastic model that meet certain optimality demands. The technique used if the underlying deterministic model is one-dimensional is the Kalman filter. In this case there is no need to impose linearity restrictions on the dynamics, and the favorable properties of the filter can be fully exploited to:

- estimate simultaneously uncertain parameters and states (both on-line and off-line),
- process the data up to the actual time, to determine an optimal initial condition that is used for a prediction model (this application has a very great practical relevance),
- optimize measurement networks by analyzing the amount of information that is associated with a certain measurement point.

The usefulness of the Kalman filter is illustrated by ten Brummelhuis *et al.* (1988) where it is applied to an operational model of the tidal estuary (Fig. 1). However, if the model is two-dimensional, computational problems arise and a direct application of Kalman filtering techniques is out of the question, unless one is satisfied with the steady state approach to process the data. Therefore, if we focus our attention on these stochastic, two-dimensional nonlinear tidal models, alternative algorithms have to be developed to accomplish the goals mentioned above.

In this paper we describe a method that is suited for the simultaneous estimation of uncertain parameters and states in large-scale models in the presence of uncertainties in the open boundaries. This estimation problem is

* Received 16 July 1991; revised 5 August 1992; revised 15 May 1993; received in final form 3 September 1993. The original version of this paper was presented at the 9th IFAC/IFORS Symposium on Identification and System Parameter Estimation which was held in Budapest, Hungary during 8–12 July 1991. The Published Proceedings of this IFAC Meeting may be ordered from Elsevier Science Limited, The Boulevard, Langford Lane, Kidlington, Oxford OX5 1GB, U.K. This paper was recommended for publication in revised form by Associate Editor Bo Wahlberg under the direction of Editor P. C. Parks. Corresponding author Professor A. Bagchi. Tel. +31 53 893452; Fax +31 53 340733; E-mail a.bagchi@math.utwente.nl.

† University of Twente, Department of Applied Mathematics, P.O. Box 217, 7500 AE Enschede, The Netherlands.

‡ Present address: Delft Hydraulics, P.O. Box 152, 8300 AD Emmeloord, The Netherlands.

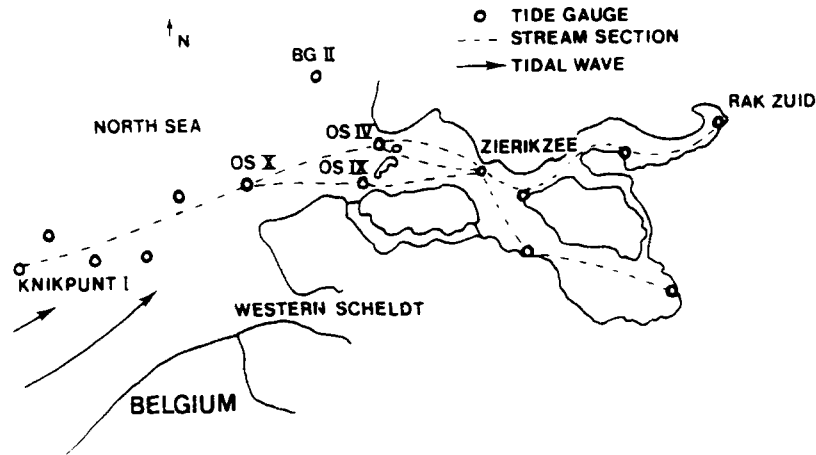


FIG. 1. The model of the tidal estuary in The Netherlands.

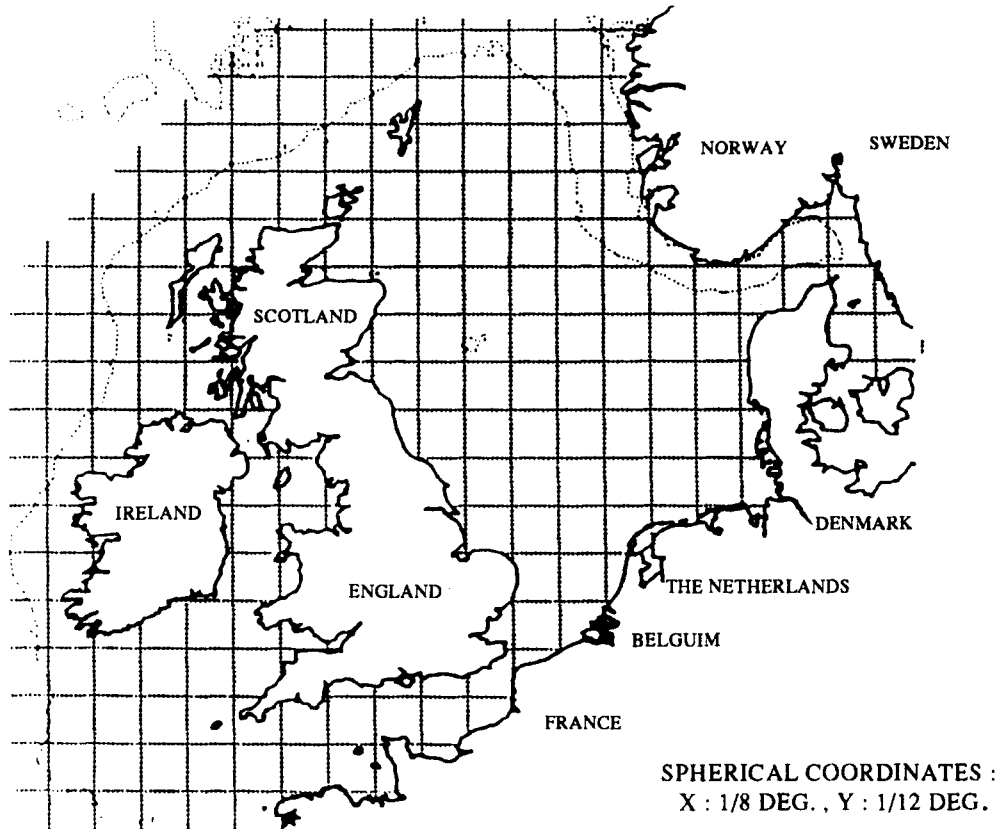


FIG. 2. The Continental Shelf Model.

treated as a constrained minimization problem, based on the Maximum Likelihood (ML) method. The original idea of using the ML method for the estimation of the state of a lumped dynamical system goes back to the 1960s (Cox, 1964; Detchmendy and Sridhar, 1965; Mortensen; 1968). A big disadvantage is that it leads to a two-point boundary value problem (TPBVP) that can usually only be solved at relatively high (computational) cost if large-scale systems are involved. Therefore we make use of two important facts that enable us to find the solution of the TPBVP (which is the smoothed

state) in a direct way:

- Due to the hyperbolic nature of the model equations, the solution is (after some time) completely determined by the boundary conditions.
- In case the noise sequences are assumed to be Gaussian and the model is linear, the solution of the ML equations can also be derived using a Kalman filter. This is based on the fact that, for linear systems, the mean and the mode of a probability density function coincide.

The parameter estimation is performed with a

gradient-based search algorithm. It will be shown in one of the next sections that such an algorithm is well suited to use in combination with the ML method to estimate the states, because the gradient of the criterion is completely determined by the solution of the TPBVP and the original model simulation: a result of the application of the method of Chavent (1979). This method is well known in the field of deterministic control theory to derive the exact gradient of some constrained cost function. Here, the method can be used in a stochastic environment if the original stochastic model is transformed first into a 'deterministic' one, where the new 'deterministic' variables are the estimates of the states in the stochastic model.

We start in the next section with a brief description of numerical tidal models in general, and the specific stochastic aspects that are considered here. In the sections thereafter the state and parameter estimation procedures are presented in a formal way. Subsequently, the solution method of the TPBVP and the parameter estimation problem are treated. The paper ends with a large-scale application and some conclusions.

NUMERICAL TIDAL MODELS

The dynamical equations of tidal models are the so-called shallow water equations, stating the conservation of mass and momentum. In differential form these equations are:

$$\frac{\partial h}{\partial t} + u \frac{\partial [D+h]u}{\partial x} + v \frac{\partial [D+h]v}{\partial y} = 0 \quad (1)$$

$$\frac{\partial u}{\partial t} + u \frac{\partial u}{\partial x} + v \frac{\partial u}{\partial y} + g \frac{\partial h}{\partial x} + fv + gS_f \cdot n_x - C_d \frac{\rho_a V^2 \cos \phi}{\rho_w D+h} - \frac{1}{\rho_w} \frac{\partial p_a}{\partial x} = 0 \quad (2)$$

$$\frac{\partial v}{\partial t} + u \frac{\partial v}{\partial x} + v \frac{\partial v}{\partial y} + g \frac{\partial h}{\partial y} - fu + gS_f \cdot n_y - C_d \frac{\rho_a V^2 \sin \phi}{\rho_w D+h} - \frac{1}{\rho_w} \frac{\partial p_a}{\partial y} = 0, \quad (3)$$

with:

- x, y, t the space and time coordinates,
- h the water elevation above some reference level,
- u, v the vertically averaged velocity components in the x, y directions,
- C_d the wind stress coefficient,
- $D(x, y)$ the water depth below the reference level,
- g the gravitational acceleration,

- f the Coriolis coefficient,
- n_x, n_y unit vectors in the x and y directions,
- S_f the resistance slope,
- ϕ the wind angle,
- V the wind velocity,
- p_a the atmospheric pressure,
- ρ_a the density of air,
- ρ_w the density of water.

Boundary conditions must be imposed to obtain a unique solution of this tidal model. At closed boundaries the velocity perpendicular to the coast is set equal to zero and at open boundaries the water level elevation is given as an explicit function of time. Because of the nonlinearity of the dynamics, an additional boundary condition has to be imposed in case of inflow to ensure the well-posedness of the problem. This condition states that the velocity parallel to the boundary vanishes. The interested reader is referred to Stelling (1983) for a detailed treatment of the boundaries.

With respect to the discretization of the dynamical model just described, two aspects should be emphasized. Because we are dealing with operational tidal models, the dimension of the state is up to 20,000. Therefore, a lot of effort has already been taken to optimize the numerical codes for these models. One of the ways to reduce the dimension of the state without affecting the accuracy of the computation is to use a staggered grid where the location of the computed water level and the two velocity components do not coincide (Fig. 3; Hansen, 1956; Patankar and Spalding, 1970).

The discretization scheme that is used to represent equations (1)–(3) in discrete time is known as an ADI-type scheme (Alternate Direction Implicit; Mitchell and Griffiths, 1980).

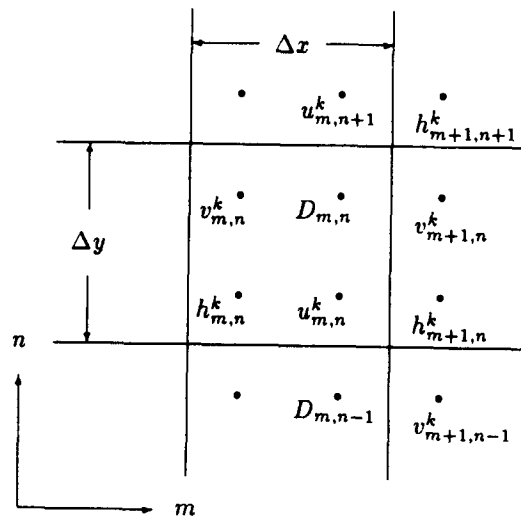


FIG. 3. The computational grid.

Characteristic of this representation is the splitting of the time step from k to $k + 1$ in two half time steps. In the first part the velocity in the y -direction $v_{m,n}^{k+1/2}$ is determined explicitly, based on $u_{m,n}^k, v_{m,n}^k$ and $h_{m,n}^k$, while the remaining variables, the velocity in the x -direction $u_{m,n}^{k+1/2}$ and the water level elevation $h_{m,n}^{k+1/2}$ are found as the solution of implicit equations. In the second part the roles of m and n are changed: the velocity component $u_{m,n}^{k+1}$ is based on $u_{m,n}^{k+1/2}, v_{m,n}^{k+1/2}$ and $h_{m,n}^{k+1/2}$ and $v_{m,n}^{k+1}$ and $h_{m,n}^{k+1}$ are again the solution of implicit equations. It is not our purpose to go into detail about the characteristics of this discretization method; we will only remark that this method has proven its usefulness in the area of numerical tidal modeling (Stelling, 1983). To simplify the notation, the time splitting will be ignored and the numerical model is formally written as:

$$\mathbf{X}_{k+1} = \mathbf{f}_1(\mathbf{p}, \mathbf{X}_k) + B_1 \mathbf{u}_{k+1} \quad (4)$$

with \mathbf{u}_k a deterministic input representing the boundary conditions and \mathbf{p} the vector of unknown parameters.

The next step is to introduce the uncertainty into the model to make it stochastic, which becomes imperative when the open boundary condition is described. In dealing with tidal motion problems, the water level and the velocity components can be expressed in terms of a number of harmonic constituents with known frequencies (Fig. 4).

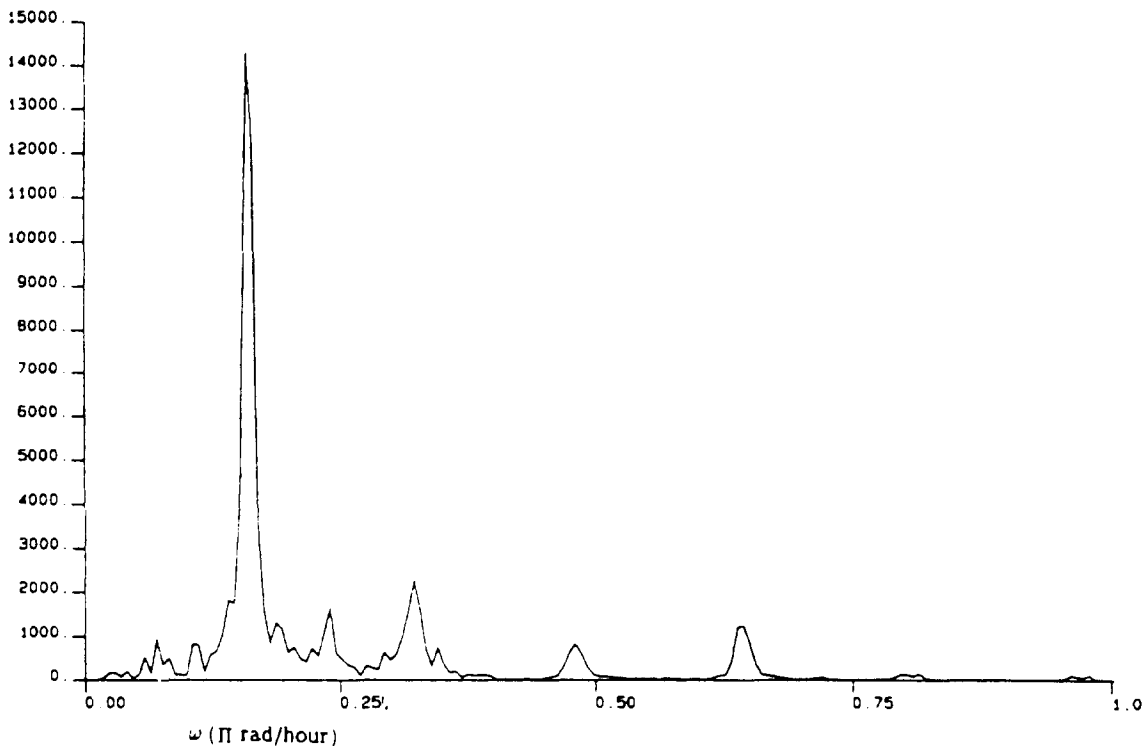


FIG. 4. The harmonic analysis of the tide.

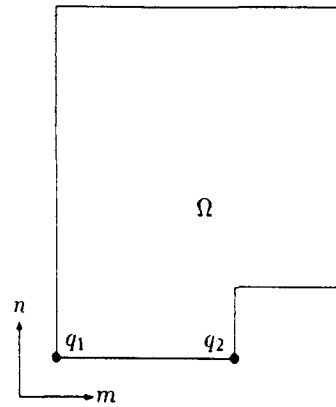


FIG. 5. A simple geometry Ω with an open boundary at $n = 0$. The other boundaries are closed.

This means that the uncertainty in the open boundary condition can also be interpreted in terms of harmonic components. For convenience, consider the open boundary of a simple basin (Fig. 5). If we represent the open boundary by equation (5)

$$h_{m,0}^k = u_k + \text{'unknown part'} \quad (5)$$

we may model this 'unknown part' in its simplest form, as an AR(1)-model [see equations (7)-(8)]:

$$h_{m,0}^k = u_k + \mu_m q_1^k + (1 - \mu_m) q_2^k, \quad \text{with } 0 \leq \mu_m \leq 1 \quad (6)$$

$$q_1^k = \rho q_1^{k-1} + \eta_1^{k-1}, \quad \eta_1^k = N(0, Q_1) \quad (7)$$

$$q_2^k = \rho q_2^{k-1} + \eta_2^{k-1}, \quad \eta_2^k = N(0, Q_2), \quad (8)$$

with ρ a correlation coefficient $0 \leq \rho \leq 1$. From equation (6) one can see that the processes q_1 and q_2 are linearly interpolated along the boundary where $n=0$ by taking $\mu_m = 1 - m/m_{\max}$, for $m = 0, 1, \dots, m_{\max}$. The idea behind this description is that the 'unknown part' of the open boundary condition consists of harmonic constituents, similar to u_k . By choosing ρ close to 1, the low frequency waves are taken into account. On the other hand, the high frequency waves are represented by the noise sequences η_1 and η_2 . The introduction of the processes q_1 and q_2 at the edges of the open boundary allows us to account for (linear) variations of both the amplitude and the phase of the 'unknown part' in the open boundary condition. This is of course a very simple approach. However, the extension to higher order AR-models (and the introduction of additional processes q) is straightforward. In matrix notation this stochastic model now reads:

$$\begin{bmatrix} \mathbf{X}_{k+1} \\ q_1^{k+1} \\ q_2^{k+1} \end{bmatrix} = \begin{bmatrix} \mathbf{f}_1(\mathbf{p}, \mathbf{X}_k) + \rho G_1 q_1^k + \rho G_2 q_2^k \\ \rho q_1^k \\ \rho q_2^k \end{bmatrix} + \begin{bmatrix} B_1 \\ 0 \\ 0 \end{bmatrix} \mathbf{u}_{k+1} + \begin{bmatrix} G_1 & G_2 \\ 1 & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} \eta_1^k \\ \eta_2^k \end{bmatrix}. \quad (9)$$

Now, if the noise sequences η_1 and η_2 are denoted by the system noise \mathbf{w}_k and a new state vector \mathbf{Y}_k is defined as $\mathbf{Y}_k = [\mathbf{X}_k, q_1^k, q_2^k]^T$ the stochastic model becomes:

$$\mathbf{Y}_{k+1} = \mathbf{f}(\mathbf{p}, \mathbf{Y}_k) + B\mathbf{u}_{k+1} + G\mathbf{w}_{k+1} \quad (10)$$

$$\mathbf{z}_k = H\mathbf{Y}_k + \mathbf{v}_k \quad (11)$$

with:

- \mathbf{u}_k a deterministic input,
- \mathbf{z}_k the measurements,
- H the measurement matrix,
- $\mathbf{w}_k = N(0, Q)$ the system noise,
- $\mathbf{v}_k = N(0, R)$ the measurement noise.

This prescription will be the basic model formulation, used throughout this paper.

The method that will be introduced in the next sections is a tool in calibrating a simulation model in the sense that it tries to find a trajectory, being the solution of equations (10) and (11) that fits best to the observed data series. Here, the difference between the model outcome and the data is measured, taking into account the covariances of the measurement and system noise processes. This trajectory can be influenced by adaptations of the unknown parameters and the open boundary condition. In our model we assume that the values of the

parameters are constant in time, whereas the errors in the open boundary prescription that affect the trajectory have a time-dependent nature: the 'unknown part' of this boundary is a function of time. Despite the different character of these two error sources, the estimation of the parameters and the estimation of the open boundary will be performed simultaneously, by a method that is based on the maximum likelihood approach.

THE MAXIMUM LIKELIHOOD APPROACH

In the sequel a method will be discussed to estimate the parameters and the states simultaneously by maximizing a likelihood function $L(\mathbf{p}, \mathbf{Y}_k)$. Usually, this is treated as the minimization of $J(\mathbf{p}, \mathbf{Y}_k)$ with $J(\mathbf{p}, \mathbf{Y}_k) = -\log L(\mathbf{p}, \mathbf{Y}_k)$ and $J(\mathbf{p}, \mathbf{Y}_k)$ is referred to as the error criterion. The criterion that is considered here consists of two terms, one of them representing the measurement error, the other the system error. So, the minimization of J with respect to states and parameters can also be interpreted as finding the smallest realization of the covariance of the measurement and system noise sequences. For J we take:

$$J(\mathbf{p}, \mathbf{Y}_k) = \frac{1}{2} \sum_{k=N_0}^N \|\mathbf{z}_k - H\mathbf{Y}_k\|_{R^{-1}}^2 + \frac{1}{2} \sum_{k=0}^{N-1} \|\mathbf{Y}_{k+1} - \mathbf{f}(\mathbf{p}, \mathbf{Y}_k) - B\mathbf{u}_{k+1}\|_{(GQG^T)^{-1}}^2. \quad (12)$$

As already mentioned, we will derive \mathbf{p}^* and \mathbf{Y}_k^* , which minimize $J(\mathbf{p}, \mathbf{Y}_k)$, in two steps:

- (1) First to find the minimizing trajectory $\{\mathbf{Y}_k^*(\mathbf{p}), k = 0, 1, \dots, N\}$ for a certain value of \mathbf{p} .
- (2) The second step selects from this class of minimizing trajectories the trajectory, parametrized by \mathbf{p}^* , which minimizes J globally. In this sense, the class of minimizing trajectories acts as a constraint for the minimization of J w.r.t. \mathbf{p} .

State estimation

In order to find the minimizing trajectory w.r.t. the state, equation (12) is rewritten as:

$$J(\mathbf{p}, \mathbf{Y}_k) = \frac{1}{2} \sum_{k=N_0}^N \|\mathbf{z}_k - H\mathbf{Y}_k\|_{R^{-1}}^2 + \frac{1}{2} \sum_{k=0}^{N-1} \|G\mathbf{w}_k\|_{(GQG^T)^{-1}}^2 + \sum_{k=0}^{N-1} \mathbf{v}_{k+1}^T [\mathbf{Y}_{k+1} - \mathbf{f}(\mathbf{p}, \mathbf{Y}_k) - B\mathbf{u}_{k+1} - G\mathbf{w}_k], \quad (13)$$

where the adjoint states $\mathbf{v}_k, k = 1, 2, \dots, N$ are

introduced to incorporate the constraint in the minimization. At this point the discrete Pontryagin minimum principle (Boltjanski, 1976), may be applied to find a realization of the system noise sequence \mathbf{w}_k . This leads to the requirement that $G\mathbf{w}_k = GQG^T\mathbf{v}_{k+1}$. For the minimization of J , $\partial J/\partial \mathbf{Y}_k$ and $\partial J/\partial \mathbf{v}_k$ are set equal to zero, so the smoothed state \mathbf{Y}_k^* and the adjoint state \mathbf{v}_k satisfy:

$$\mathbf{Y}_{k+1}^* = \mathbf{f}(\mathbf{p}, \mathbf{Y}_k^*) + B\mathbf{u}_{k+1} + GQG^T\mathbf{v}_{k+1},$$

$$k = 0, \dots, N-1 \quad (14)$$

$$\mathbf{v}_k = F(\mathbf{p}, \mathbf{X}_k)^T\mathbf{v}_{k+1} + H^TR^{-1}$$

$$\times [\mathbf{z}_k - H\mathbf{Y}_k^*], \quad k = N_0, \dots, N \quad (15)$$

$$\mathbf{v}_k = F(\mathbf{p}, \mathbf{X}_k)^T\mathbf{v}_{k+1}, \quad k = 1, 2, \dots, N_0 - 1 \quad (16)$$

$$\mathbf{v}_{N+1} = \mathbf{0} \quad (17)$$

$$\mathbf{Y}_0^* = \mathbf{0}, \quad (18)$$

where F is the Jacobian $\partial \mathbf{f}/\partial \mathbf{X}_k$. These equations must be satisfied by the states \mathbf{Y}_k^* in order to minimize the criterion J for a certain value of the parameter \mathbf{p} . The second part of the estimation procedure now consists of finding the minimum of J w.r.t. \mathbf{p} .

Parameter estimation

As already mentioned in the Introduction, the parameter estimation is performed by a gradient-based algorithm, where the gradient of $J(\mathbf{p}, \mathbf{Y}_k)$ w.r.t. \mathbf{p} is determined by the method of Chavent (1979). In trying to find \mathbf{p}^* we are minimizing $J(\mathbf{p}, \mathbf{Y}_k)$ under the restriction that the equations (14)–(16) are satisfied, including the conditions (17) and (18). Then, \mathbf{p}^* can be estimated iteratively, using a gradient-based algorithm. The way to derive ∇J is (ten Brummelhuis and Heemink, 1990; Chavent; 1979):

(1) Introduce Lagrange multipliers ξ_k and ζ_k to incorporate the TPBVP as the new constraints:

$$J(\mathbf{p}, \mathbf{Y}_k) = \frac{1}{2} \sum_{k=N_0}^N \|\mathbf{z}_k - H\mathbf{Y}_k\|_{R^{-1}}^2$$

$$+ \frac{1}{2} \sum_{k=0}^{N-1} \|GQG^T\mathbf{v}_{k+1}\|_{GQG^T}^2$$

$$+ \sum_{k=0}^{N-1} \xi_{k+1}^T [\mathbf{Y}_{k+1} - \mathbf{f}(\mathbf{p}, \mathbf{Y}_k)$$

$$- B\mathbf{u}_{k+1} - GQG^T\mathbf{v}_{k+1}]$$

$$+ \sum_{k=0}^{N_0-1} \zeta_k^T [\mathbf{v}_k - F(\mathbf{p}, \mathbf{X}_k)\mathbf{v}_{k+1}]$$

$$+ \sum_{k=N_0}^N \zeta_k^T [\mathbf{v}_k - F(\mathbf{p}, \mathbf{X}_k)\mathbf{v}_{k+1}$$

$$- H^TR^{-1}[\mathbf{z}_k - H\mathbf{Y}_k]]. \quad (19)$$

(2) The method of Chavent (1979) considers the effect of infinitesimal changes Δp in one of the components of \mathbf{p} on J , \mathbf{Y}_k , \mathbf{v}_k , ξ_k and ζ_k . This is done by linearizing the expression given above:

$$\Delta J = \frac{\partial J}{\partial p} \Delta p = \sum_{k=N_0}^N -H^TR^{-1} \Delta \mathbf{Y}_k^T$$

$$\times [\mathbf{z}_k - H\mathbf{Y}_k]$$

$$+ \sum_{k=0}^{N-1} \Delta \mathbf{v}_{k+1}^T GQG^T \mathbf{v}_{k+1}$$

$$+ \sum_{k=0}^{N-1} \xi_{k+1}^T [\Delta \mathbf{Y}_{k+1} - F(\mathbf{p}, \mathbf{Y}_k) \Delta \mathbf{Y}_k$$

$$- GQG^T \Delta \mathbf{v}_{k+1}]$$

$$+ \sum_{k=0}^{N-1} \Delta \xi_{k+1}^T [\mathbf{Y}_{k+1} - \mathbf{f}(\mathbf{p}, \mathbf{Y}_k)$$

$$- (GQG^T)\mathbf{v}_k]$$

$$+ \sum_{k=0}^{N_0-1} \zeta_k^T [\Delta \mathbf{v}_k - F(\mathbf{p}, \mathbf{Y}_k)^T \Delta \mathbf{v}_{k+1}]$$

$$+ \sum_{k=0}^{N_0-1} \Delta \zeta_k^T [\mathbf{v}_k - F(\mathbf{p}, \mathbf{Y}_k)^T \mathbf{v}_{k+1}]$$

$$+ \sum_{k=N_0}^N \zeta_k^T [\Delta \mathbf{v}_k - F(\mathbf{p}, \mathbf{Y}_k)^T \Delta \mathbf{v}_{k+1}$$

$$+ H^TR^{-1}H \Delta \mathbf{Y}_k]$$

$$+ \sum_{k=N_0}^N \Delta \zeta_k^T [\mathbf{v}_k - F(\mathbf{p}, \mathbf{Y}_k)^T \mathbf{v}_{k+1}$$

$$+ 2H^TR^{-1}[\mathbf{z}_k - H\mathbf{Y}_k]]$$

$$+ \sum_{k=0}^{N-1} \xi_{k+1}^T \frac{\partial \mathbf{f}}{\partial p} \Delta p - \sum_{k=0}^N \zeta_k^T \frac{\partial F^T}{\partial p} \mathbf{v}_{k+1} \Delta p. \quad (20)$$

(3) In equation (20) the terms with $\Delta \xi_k$ and $\Delta \zeta_k$ can be neglected because the factors between brackets are the left-hand sides of equations (14)–(16) which are equal to zero. Then, by rearranging the remaining terms, one finds:

$$\Delta J = \frac{\partial J}{\partial p} \Delta p$$

$$= \sum_{k=1}^{N_0-1} \Delta \mathbf{Y}_k^T [\xi_k - F(\mathbf{p}, \mathbf{Y}_k)^T \xi_{k+1}]$$

$$+ \sum_{k=N_0}^N \Delta \mathbf{Y}_k^T [\xi_k - F(\mathbf{p}, \mathbf{Y}_k)^T \xi_{k+1}$$

$$- H^TR^{-1}[\mathbf{z}_k - H\mathbf{Y}_k - H\zeta_k]]$$

$$+ \sum_{k=1}^N \Delta \mathbf{v}_k^T [\zeta_k - F(\mathbf{p}, \mathbf{Y}_k)\zeta_{k-1}$$

$$- GQG^T[\xi_k - \mathbf{v}_k]]$$

$$\begin{aligned}
 & + \sum_{k=0}^{N-1} \xi_{k+1}^T \frac{\partial f}{\partial p} \Delta p \\
 & + \sum_{k=0}^N \zeta_k^T \frac{\partial F^T}{\partial p} \mathbf{v}_{k+1} \Delta p.
 \end{aligned} \tag{21}$$

(4) The expressions under the summation signs are quite similar to equations (14)–(16), including the boundary conditions, except that ξ_k now plays the role of \mathbf{v}_k whereas ζ_k can be associated with \mathbf{Y}_k . In looking for a simple expression for ∇J we can exploit this similarity: if $\zeta_k = \mathbf{0}$ and $\xi_k = \mathbf{v}_k$ for all k , then

$$\frac{\partial J}{\partial p} = - \sum_{k=0}^N \mathbf{v}_{k+1}^T \frac{\partial f(\mathbf{p}, \mathbf{Y}_k)}{\partial p}. \tag{22}$$

This implies that ∇J is simply derived from the solution of the TPBVP; or with a slightly different interpretation: the only problem for the simultaneous estimation of states and parameters is to find the solution of the TPBVP.

The fact that the condition for \mathbf{Y}^* is given for the initial time and the condition for \mathbf{v} at the final time, eliminates the possibility to find a solution of equations (14)–(16) in a direct way. Although a number of methods are available in the literature to solve two-point boundary value problems, such as iterative shooting methods, none of them seems very appropriate here, because they require an excessive computational load. Therefore, it is necessary to develop an alternative strategy that makes use of the special characteristics of the problem in order to make procedure applicable also to large-scale problems.

THE ESTIMATION PROCEDURE

In the previous sections, all the ingredients of the complete estimation procedure are described, and they are now put together. It was shown there that the ML estimate of the state and ∇J are both determined by the solution of the adjoint model. This is the main reason to choose a gradient-based algorithm for estimating the parameters. In our approach, an iterative quasi-Newton method is used to find the value \mathbf{p}_0^* that minimizes J . Therefore we need, besides $\nabla J(\mathbf{p})$, information about the Hessian $H(\mathbf{p})$ of $J(\cdot)$. By approximating $J(\cdot)$ locally by a quadratic function $J^q(\Delta \mathbf{p})$:

$$J^q(\Delta \mathbf{p}) = J(\mathbf{p}) + \Delta \mathbf{p}^T \nabla J(\mathbf{p}) + \frac{1}{2} \Delta \mathbf{p}^T H(\mathbf{p}) \Delta \mathbf{p}. \tag{23}$$

The value for $\Delta \mathbf{p}$ that minimizes J^q is found from $\nabla J^q = \mathbf{0}$, or, $\Delta \mathbf{p} = -H^{-1}(\mathbf{p}) \nabla J(\mathbf{p})$. Generally, the criterion J is not locally quadratic w.r.t. \mathbf{p} . The vector $\mathbf{s}(\mathbf{p}) = -H^{-1}(\mathbf{p}) \nabla J(\mathbf{p})$ is therefore

interpreted as a search direction in the parameter space. The new parameter estimate is now found as the line minimum in the direction of $\mathbf{s}(\mathbf{p})$.

The exact calculation of $H^{-1}(\mathbf{p})$ requires a large amount of computation time. To avoid these types of problems, the calculation of $H^{-1}(\mathbf{p})$ is replaced by a recursive update of a positive definite symmetric matrix for which we used the BFGS method (Fletcher, 1980). It makes the convergence of the iteration nearly second order and it can be proved that in case the function J is purely quadratic w.r.t. the parameter \mathbf{p} , the number of iterations equals the dimension of \mathbf{p} plus one. The whole procedure is schematically shown in Table 1.

SOLUTION OF THE TPBVP FOR LINEAR MODELS

In this section, the strategy to determine the solution of the TPBVP is described for linear tidal models where the random input for the open boundaries is modelled by an AR(1) model. Fundamental to this alternative strategy for solving the TPBVP is the fact that the components of $GQG^T \mathbf{v}_{k+1}$ in equation (14) only contribute to \mathbf{Y}_{k+1} through the components q_1^{k+1} and q_2^{k+1} as is the case in the example with the simple geometry (Fig. 3). This means that the smoothed states directly follow from equation (14) once the smoothed boundary is known. The second crucial point is the linearity of the model. This property implies the equality of the mean, which can be found by the Kalman filter, and the mode, the maximum of the likelihood function (or the minimum of the criterion J).

Proposition 1. The smoothed states $\{\mathbf{Y}_k^* | k = 0, 1, \dots, N\}$ can be calculated directly, without using equations (14)–(16).

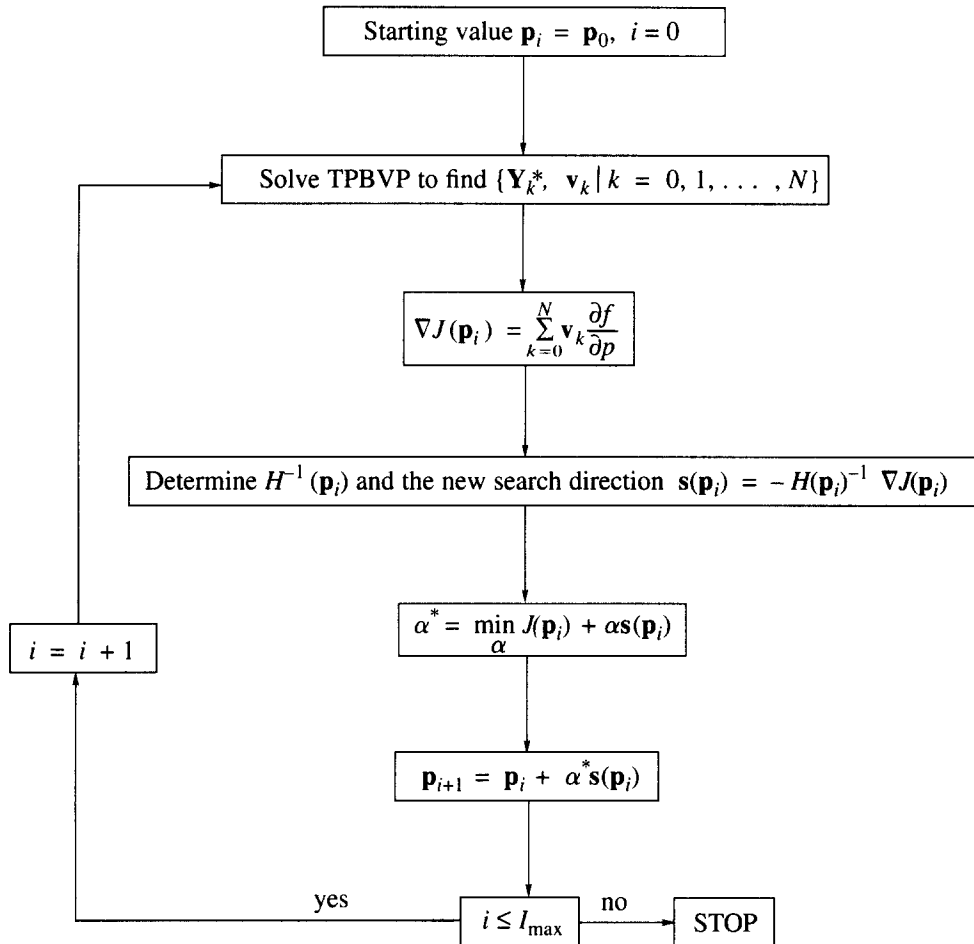
The proof of this proposition is given in the Appendix. With this result, the solution of the TPBVP is found after the following steps:

- (1) Introduce a new state $\mathbf{Y}_k = [\mathbf{Y}_k, q_1^{k-1}, q_2^{k-1}, q_1^{k-2}, q_2^{k-2}, \dots, q_1^0, q_2^0]^T$, which contains the history of the stochastic processes q_1 and q_2 from the starting time until the present time k .
- (2) If a Kalman filter is used to process the data $\{z_l, l = 1, 2, \dots, N-1, N\}$ one finds at the final time, $k = N$, the filtered estimate of the state

$$\mathbf{Y}'(N | N) = [\mathbf{Y}(N | N), q_1(N-1 | N), q_2(N-1 | N), \dots, q_1(0 | N), q_2(0 | N)]^T.$$

- (3) Based on the proposition, equation (9) may

TABLE 1. THE ESTIMATION PROCEDURE



be rewritten as

$$\mathbf{X}_{k+1}^* = \mathbf{f}_1(\mathbf{p}, \mathbf{X}_k^*) + B_1 \mathbf{u}_{k+1} + G_1 [q_1(k+1|N), q_2(k+1|N)]^T.$$

(4) Finally, equations (15)–(16) are solved, substituting \mathbf{X}_k^* , q_1^{*k} and q_2^{*k} .

The essence of this proposition is that the backwards integration of equations (15) and (16) is avoided while the smoothed states can still be determined. In case we are only interested in the estimation of unknown parameters (and therefore not in the smoothed states), the backward integration can also be avoided by making use of the observation in Söderström (1989) that for linear systems the likelihood function is a known function of the one-step ahead prediction error and its covariance. This implies that the smoothed states do not have to be computed explicitly. In this paper, we in fact do compute the smoothed states explicitly because only in this way can the uncertainties of the open boundary prescription be analyzed properly.

In practice, the whole history of the stochastic processes $\{q_{1,2}^l \mid l = 0, 1, \dots, k-1, k\}$ is not included in \mathbf{Y}_{k+1}^* but only the limited history

$\{q_{1,2}^l \mid l = k - k_0, k - k_0 + 1, \dots, k - 1, k\}$. The length of this time lag k_0 is chosen so large that $q_{1,2}^{k-k_0}$ do not influence the state \mathbf{Y}_k anymore. This is a direct consequence of the fact that the dynamical equations are hyperbolic, implying that every signal entering the model has only a limited residential time.

NONLINEAR MODELS

If the parameter estimation is considered for a nonlinear model with uncertain boundary inputs, the TPBVP is similar to the one in the linear case. The only difference is the occurrence of a nonlinear dynamics term $\mathbf{f}(\mathbf{p}, \mathbf{Y}_k)$ instead of $F(\mathbf{p})\mathbf{Y}_k$. The solution strategy for nonlinear models is again based on the fact that for hyperbolic systems, the smoothed states can be easily computed once the smoothed boundaries are known. For linear models this could be achieved by first using the Kalman filter to determine the smoothed boundaries by augmenting the state vector with the history of the boundary variables q_k .

In the case the model is nonlinear, the natural extension would be to replace the steady state

filter by an extended filter. This straightforward adaptation, however, requires too much computational power. So, besides the argument that an extended filter assumes a (local) linearization of the dynamics and can therefore only generate an approximated smoothed boundary, it cannot be used for large-scale models. For solving the TPBVP fast, we have to include some approximations. For example:

- linearize the dynamics globally and apply a linear filter to determine a smoothed boundary,
- derive a smoothed boundary by recursively minimizing a criterion $J(\mathbf{p}, \mathbf{Y}'_k)$: $(\mathbf{Y}'_k)^*$ is the minimizing solution of

$$\text{minimize } J_k(\mathbf{p}, \mathbf{Y}'_k) = \|\mathbf{z}_k - H\mathbf{Y}'_k\|_{R^{-1}}^2 + \|\mathbf{Y}'_k - \mathbf{f}(\mathbf{p}, \mathbf{Y}'_{k-1}) - B\mathbf{u}_k\|_{P^{-1}}^2 \quad (24)$$

w.r.t. \mathbf{Y}'_k if $(\mathbf{Y}'_{k-1})^*$ is given. In equation (23) the matrix \mathbf{p} can be interpreted as the covariance of \mathbf{Y}'_k . This gives the same solution as the linear filter if the prediction step is based on the nonlinear dynamics:

$$(\mathbf{Y}'_{k+1})^* = \mathbf{f}(\mathbf{p}, (\mathbf{Y}'_k)^*) + K\{\mathbf{z}_{k+1} - H\mathbf{f}(\mathbf{p}, \mathbf{Y}'_k)^*\}. \quad (25)$$

The justification of an approximation can be done by a close physical inspection of the dynamics, as is done in the application that is examined in the next section.

APPLICATION TO THE CONTINENTAL SHELF MODEL

The Continental Shelf Model (CSM) is one of the most important models for the simulation and prediction of the tidal flow in the North Sea. It covers an area of 2.310^6 km^2 [25 (western) by 14 (northern) degrees in latitude]. Because of the large size of the domain, the dynamical equations have to be considered within a spherical coordinate system (ten Brummelhuis, 1992). The state vector has a dimension of approximately 15,000.

Due to its importance, a lot of effort has already been invested to calibrate this model. This is done manually. In the sequel, this manually calibrated model is referred to as the operational model. In order to test our method to simultaneously estimate states and parameters, it is applied to CSM. Therefore we consider the estimation of some parameters in this tidal model where the open boundary is assumed to be uncertain.

In the operational model, $C_{m,n}$ and $D_{m,n}$ are the bottom friction coefficient and the depth in a gridpoint with coordinates (m, n) . Obviously,

not every $C_{m,n}$ and $D_{m,n}$ can be conceived as unknown parameters of the simulation model. Hence, we assume that the numerical domain Ω is divided into subdomains Ω_i , $i = 1, \dots, I$. For each of these subdomains a correction coefficient c_i is defined that is related to $C_{m,n}$ by

$$C_{m,n} = C_{m,n}^{op} + c_i \quad \text{for } (m, n) \in \Omega_i, \quad (26)$$

with $C_{m,n}^{op}$ the value of the coefficient in the operational model. The new coefficients c_i are treated as the unknown parameters that are to be estimated: they act as a correction for the mean level of the $C_{m,n}$ in a subdomain Ω_i and leave the spatial dependence inside Ω_i unaltered.

The same can be done with respect to the depth. Again we introduce correction coefficients d_j , $j = 1, \dots, J$ that are related to the depth by

$$D_{m,n} = D_{m,n}^{op} + d_j \quad \text{for } (m, n) \in \Omega_j. \quad (27)$$

Obviously, the division $\Omega = \bigcup \Omega_i$ (correction coefficients for the bottom friction) may differ from $\Omega = \bigcup \Omega_j$ (correction for the depth).

The open boundary condition of CSM is a prescribed water level elevation (see the section on Numerical Tidal Models), given in the form of a series of 10 harmonic constituents

$$h_{m_b, n_b}^k = \sum_{l=1}^{10} A_{m_b, n_b}^l \cos(\omega_l k \Delta t + \phi_{m_b, n_b}^l). \quad (28)$$

The uncertainty in the open boundaries [denoted by coordinates (m_b, n_b)] is represented by a number (=5) of stochastic processes q_i satisfying AR(1) models

$$q_i^{k+1} = \rho q_i^k + \eta_i^{k+1} \quad (29)$$

that are located at the edges of the open boundary (Fig. 6). In between, the correction of the boundary condition is found by linear interpolation. This is a straightforward extension of the boundary treatment in the example with the simple geometry (Fig. 3). The increased number of stochastic processes q_i^k is due to the more complex form of the open boundary.

The data are processed by means of a steady state Kalman filter. Although the dynamical equations of CSM are nonlinear this approximation is legitimate; first because the nonlinearities are weak, and second because the open boundaries are located in deep water regions just outside the continental shelf. This implies that the propagation of the tidal waves in the area along the open boundaries is a rather linear process and admits the steady state approximation for data processing in order to determine the smoothed boundaries. In other parts of the geometry, the smoothed states are determined from the original dynamical model with the

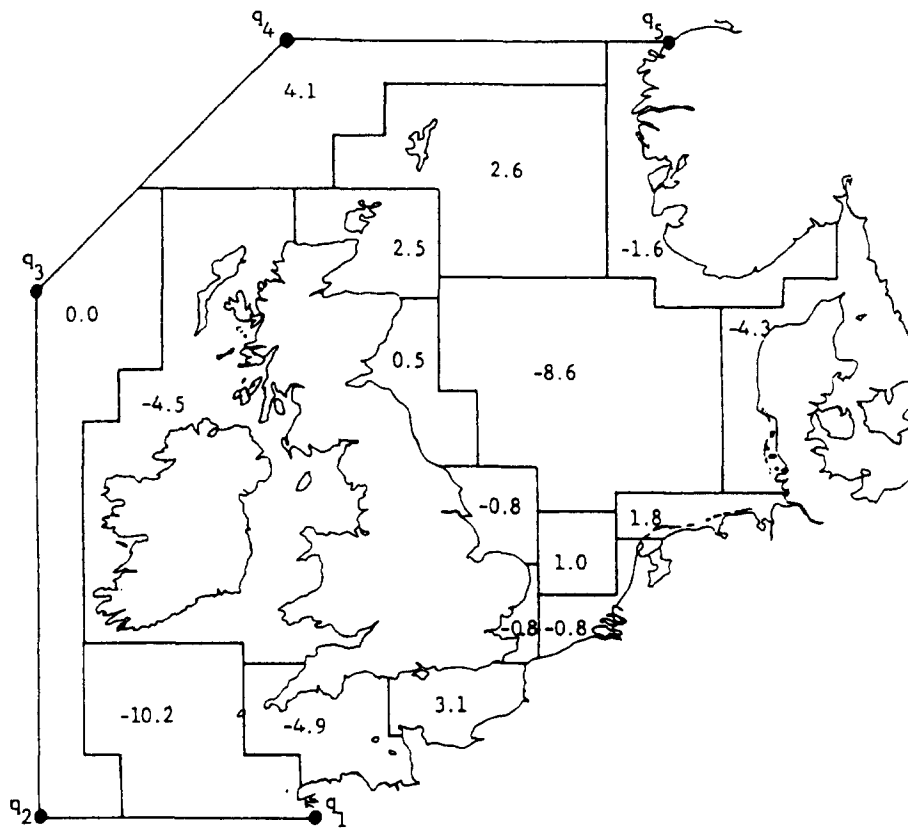


FIG. 6. The adjustment of the depth.

smoothed open boundaries as input. The nonlinearity of the dynamical model is then fully taken into account. However, a prerequisite for this approach is the availability of data observation points in the deep water areas along the open boundary.

The effectiveness of our simultaneous state and parameter estimation procedure will now be illustrated by comparing it to the performance of

TABLE 2. THE RMS ERRORS [M] IN SOME DATA OBSERVATION POINTS

Station	Operational model	After calibration
Wick	0.157	0.089
North Shields	0.152	0.115
Lowestoft	0.104	0.076
Dover	0.171	0.144
Vlissingen	0.261	0.200
H.v.Holland	0.124	0.117
Ijmuiden	0.157	0.131
Den Helder	0.121	0.101

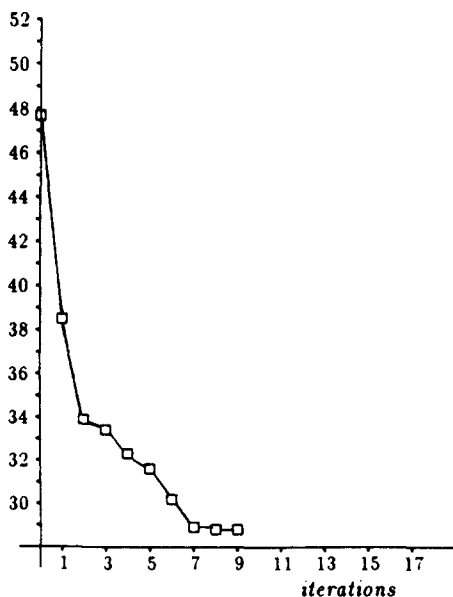


FIG. 7. The reduction of the output least squares function.

CSM (before using the operational model). One of the first quantities to consider is the reduction of the output least squares function (Fig. 7)

$$\sum_{k=N_1}^N \|\mathbf{z}_k - H\mathbf{Y}_k^*\|^2. \quad (30)$$

For the operational model we found a value of 45.7 where we used data of a one-day period. We see from Table 2 that the RMS error is substantially reduced, the average reduction is 22% compared to the RMS errors in the operational model.

The smoothed boundary (Fig. 8), indicates that only a small additional forcing is needed. The realization of the processes q_2 and q_5 are very minor; in absolute sense they do not exceed 0.01 [m]. The realizations of the remaining

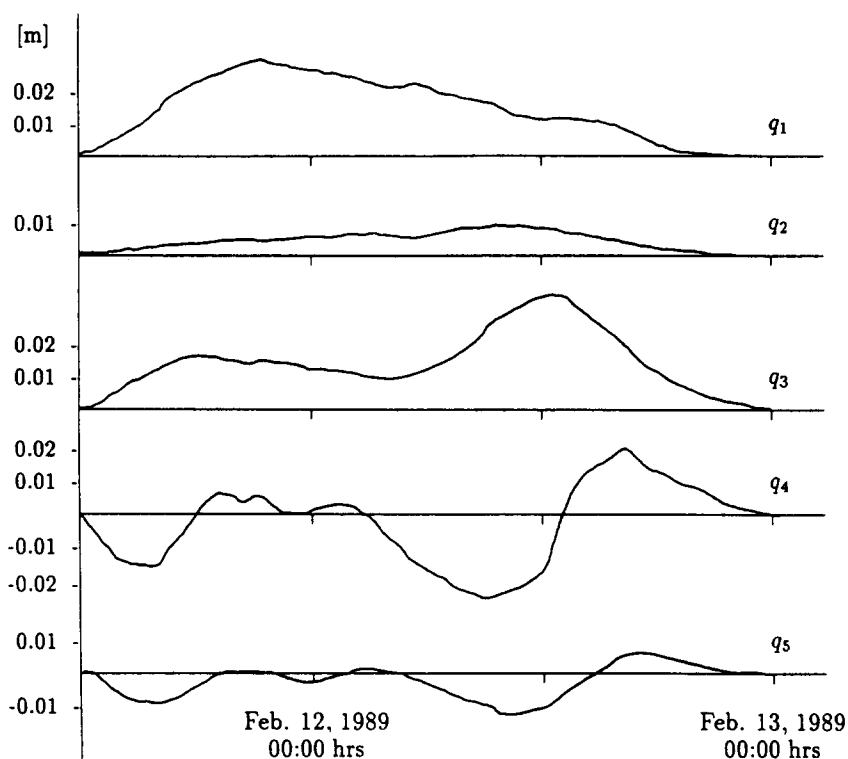


FIG. 8. The smoothed correction at the edges of the open boundaries.

processes are somewhat larger, but still they do not exceed a 0.035 [m] bound. It is remarkable that q_1, q_2 and q_3 are strictly positive. We may conclude that the mean level of the prescribed open boundary must be increased. Looking at

the realizations of the processes $q_i, i = 1, \dots, 5$, we notice a time dependence that can be interpreted in terms of tidal constituents. Especially from the realizations of q_3, q_4 and q_5 , an oscillation with a period of approximately

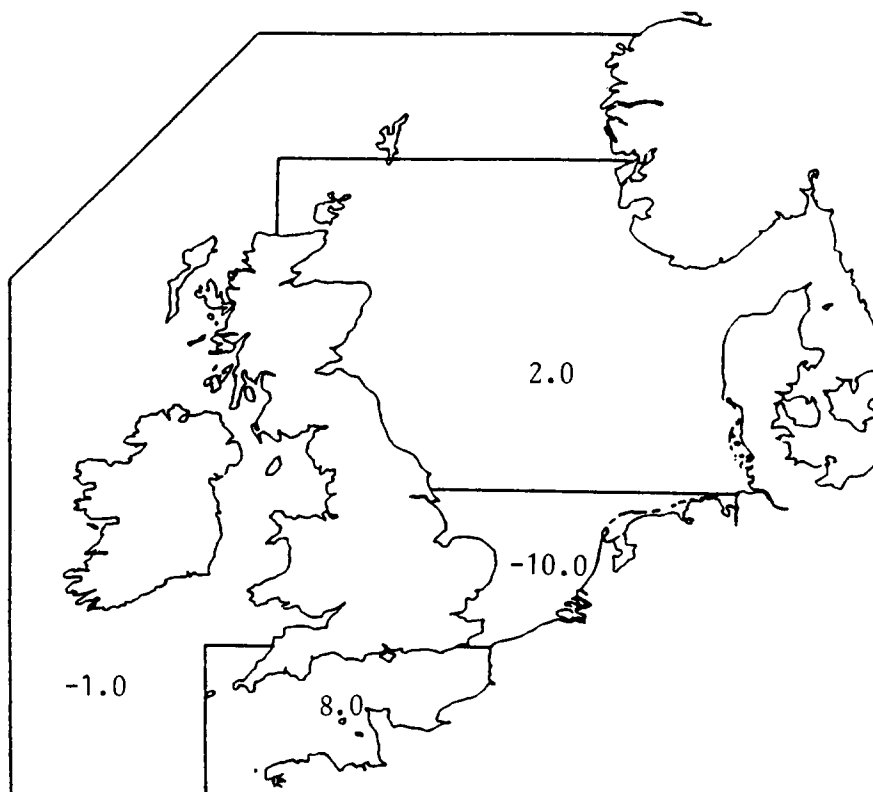


FIG. 9. The adjustment of the bottom friction.

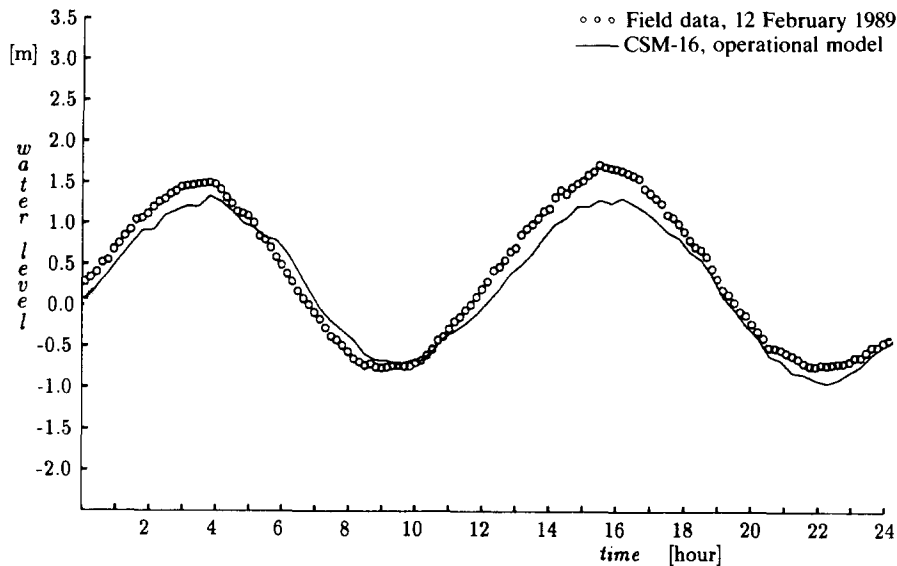


FIG. 10. Water level elevation according to the manually calibrated operational version of CSM in station Wick.

18 h can be recognized. This can be interpreted as an interaction between the M_2 and O_1 tidal constituents. In general, a harmonic analysis could be performed to derive q_i adjustments for the amplitude and phases of the various components. They, in their turn, could be used to adjust the open boundary forcing for the operational model. The adjustment of the bottom friction is shown in Fig. 9.

Without the estimation of the open boundary, the depth was partly corrected for the neglected open boundary errors. This implies that the depth is adjusted for very specific open boundary errors and it makes these adjustments only useful for that particular situation in order to reduce the RMS error. If, for instance, the adjusted depth was used in a simulation model to represent the tidal motion in a different

period (which means with different open boundary errors), the simulation results could be worsened.

The effect of the calibration on the simulation of the water level elevation is illustrated in Figs 10–13. Figures 10 and 12 represent the water level elevation according to the manually calibrated operational version of CSM. By estimating some parameters and a correction on the open boundary condition, the performance of CSM is improved substantially. The fact that a lot of effort was already invested to calibrate the operational version of CSM emphasizes the strength of our approach.

CONCLUDING REMARKS

In this paper we discussed a method to estimate simultaneously uncertain parameters

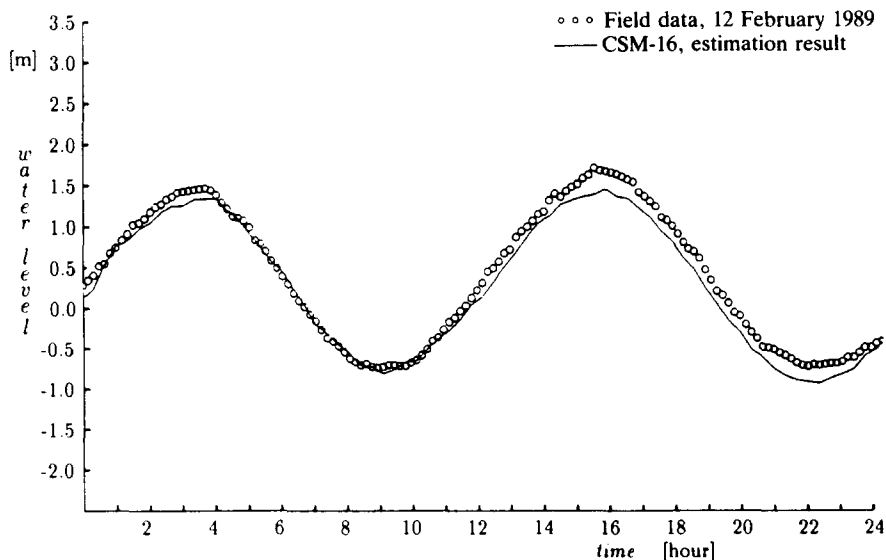


FIG. 11. Water level elevation based on the simultaneous estimation of state and parameters in station Wick.

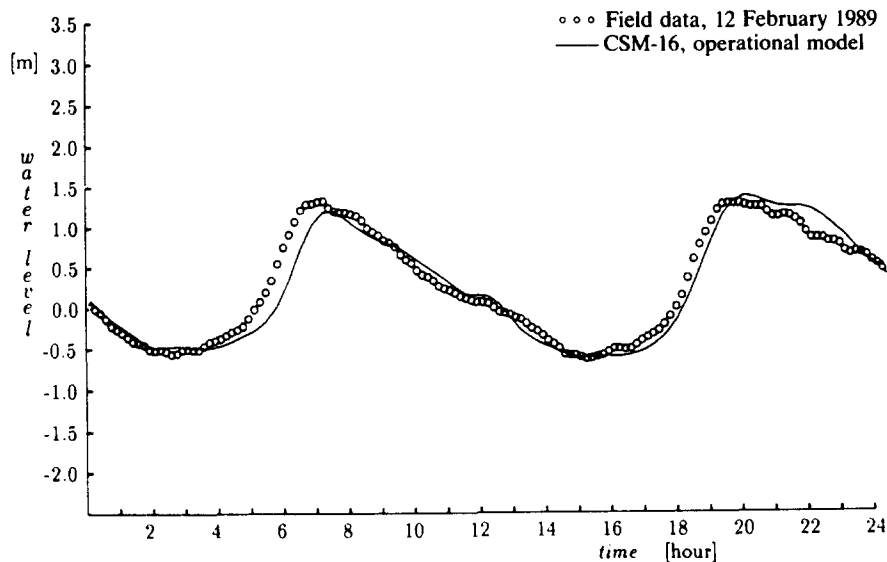


FIG. 12. Water level elevation according to the manually calibrated version of CSM in station Ijmuiden.

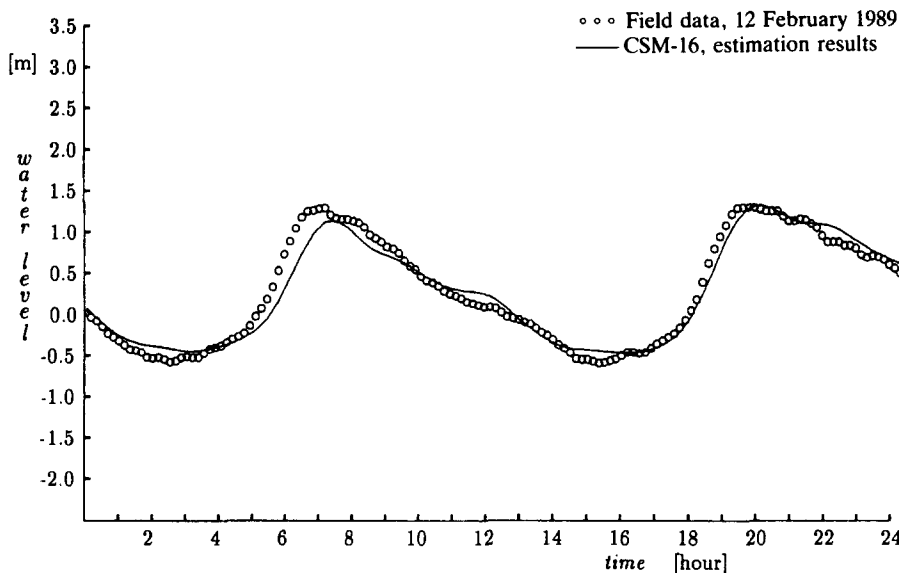


FIG. 13. Water elevation based on the simultaneous estimation of state and parameters in station Ijmuiden.

and states in tidal models with uncertainties in the prescription of the open boundary conditions. By exploiting

- the power of the method of Chavent (1979) to determine gradients of cost functions,
- the equivalence of the Kalman filter and ML solutions in the case where the model is linear and disturbed by Gaussian white noise, and
- the special characteristics of the dynamics.

A method could be developed that is applicable for large-scale models, without an excessive computational load. However, if the dynamics contain nonlinearities, approximations have to be made. The best way to choose the approximations is a subject for further research.

REFERENCES

- Boltjanski, V. (1976). *Optimale Steuerung Diskreter Systeme*. Akademische Verlagsgesellschaft Geest und Portig K.-G., Leipzig.
- ten Brummelhuis, P. G. J. (1992). Parameter estimation in tidal models with uncertain boundary conditions. Ph.D. Thesis, University of Twente, Enschede.
- ten Brummelhuis, P. G. J. and A. W. Heemink (1990). Parameter identification in tidal models with uncertain boundary conditions. *Stochastic Hydrol. Hydraulics*, **4**, 193–208.
- ten Brummelhuis, P. G. J., A. W. Heemink and B. de Jong (1988). Stochastic dynamic approach to predict water levels in estuaries. *J. Hydraulic Engng*, **114**, 1339–1358.
- Chavent, G. (1979). Identification of distributed parameter systems: about the least square method, its implementation and identifiability. *Identification and Syst. Parameter Estimation, Proc. 5th IFAC Symposium*, Darmstadt. Pergamon, Oxford, pp. 85–97.
- Cox, H. (1964). On the estimation of state variables for noisy dynamical systems. *IEEE Trans. Aut. Control*, **AC-9**, 5–12.
- Detchmندی, D. M. and R. Sridhar (1965). Sequential estimation of states and parameters in noisy non linear systems. *Proc. Joint Automatic Control Conf.* Troy, New York, pp. 56–63.
- Fletcher, R. (1980). *Practical Methods in Optimization*, 1 and 2. Wiley, New York.
- Hansen, W. (1956). Theorie zur Errechnung des Wasserstandes und der Strömungen in Randmeeren nebst Anwendungen. *Tellus*, **8**.

Mitchell, A. R. and D. F. Griffiths (1980). *The Finite Difference Method in Partial Differential Equations*. Wiley, Chichester.
 Mortensen, R. E. (1968). Maximum likelihood recursive nonlinear filtering. *J. Optimiz. Theory Applic.* **2**, 386-394.
 Patankar, S. V. and D. B. Spalding (1970). *Heat and Mass Transfer in Boundary Layers: A General Calculation Procedure*. Intertext, London.
 Söderström, T. (1989) *System Identification*. Prentice-Hall, New York.
 Stelling, G. S. (1983). *On the Construction of Computational Methods for Shallow Water Flow Problems*. Rijkswaterstaat Communications, The Hague.

APPENDIX: THE PROOF OF PROPOSITION 1

To prove the proposition we consider the related system where only one stochastic process q^k is introduced to account for uncertainties in the open boundary prescription, which is done to simplify the notation. The state of this model is

$$Y'_k = [X_k, q^k, q^{k-1}, \dots, q^0]^T, \quad (A.1)$$

so the entire history, up to the actual time index k , is included in the state. It evolves in time according to equation (A.2)

$$\begin{bmatrix} \cdot \\ \cdot \\ X_{k+1} \\ q^{k+1} \\ q^k \\ q^{k-1} \\ \cdot \\ \cdot \\ \cdot \\ \cdot \\ q^0 \end{bmatrix} = \begin{bmatrix} \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & F(\mathbf{p}) & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & 1 & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & 1 & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & 1 & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot & 1 & \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & 1 & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & 1 & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & 1 & \cdot & \cdot \end{bmatrix} \begin{bmatrix} \cdot \\ \cdot \\ X_k \\ q_k \\ q_{k-1} \\ \cdot \\ \cdot \\ \cdot \\ \cdot \\ q^0 \end{bmatrix} + B \begin{bmatrix} \cdot \\ u_{k+1} \\ \cdot \\ 0 \\ 0 \\ \cdot \\ \cdot \\ \cdot \\ \cdot \\ 0 \end{bmatrix} + G \begin{bmatrix} \cdot \\ 0 \\ w_{k+1} \\ 0 \\ \cdot \\ \cdot \\ \cdot \\ \cdot \\ \cdot \\ 0 \end{bmatrix} \quad (A.2)$$

To abbreviate the notation, $\psi(k, \mathbf{p})$ is introduced as the new state transition matrix to describe the stochastic model by equations (A.3) and (A.4):

$$Y'_{k+1} = \psi(k, \mathbf{p})Y'_k + B(k+1)u_{k+1} + G(k+1)w_{k+1}, \quad (A.3)$$

$$z_k = H(k)Y'_k + v_k. \quad (A.4)$$

Hence, if the smoothed states $\{Y'_k^* | k=0, 1, \dots, N\}$

minimize the analogon of the criterion, defined by equation (12), then they satisfy the TPBVP, equations (A.5)-(A.9):

$$Y'_{k+1}^* = \psi(k, \mathbf{p})Y'_k{}^* + B(k+1)u_{k+1} + \frac{1}{2}G(k+1)QG(k+1)^T v_{k+1}, \quad k=0, 1, \dots, N-1 \quad (A.5)$$

$$v_k = \psi(k, \mathbf{p})^T v_{k+1} + 2H(k)^T R^{-1} [z_k - H(k)Y'_k{}^*], \quad k=N_0, N_0+1, \dots, N \quad (A.6)$$

$$v_k = \psi(k, \mathbf{p})v_{k+1}, \quad k=0, 1, \dots, N_0-1 \quad (A.7)$$

and the conditions

$$Y'_0{}^* = 0, \quad (A.8)$$

$$v_{N+1} = 0. \quad (A.9)$$

Proposition 1 states that the smoothed states can be determined using Kalman filter solutions. In the following Lemma, the smoothed states Y^* of the TPBVP are first related to the Kalman filter solutions.

Lemma. If $Y'(k | k)$ is defined by

$$Y'(k | k) = Y'_k{}^* - \frac{1}{2}C(k)\psi(k, \mathbf{p})^T v_{k+1}, \quad k=0, 1, \dots, N, \quad (A.10)$$

with

$$C(k) = [I + P(k)H(k)^T R^{-1} H(k)]^{-1} P(k) \quad (A.11)$$

$$P(k) = \psi(k, \mathbf{p})C(k)\psi(k, \mathbf{p})^T + G(k)QG(k)^T \quad (A.12)$$

then

$$\begin{cases} Y'(k+1 | k+1) = \psi(k, \mathbf{p})Y'(k | k) + C(k+1)H(k+1)^T R^{-1} \times [z_{k+1} - H(k+1)\psi(k, \mathbf{p})Y'(k | k)], \\ k = N_0-1, N_0, \dots, N-1 \\ Y'(k+1 | k+1) = \psi(k, \mathbf{p})Y'(k | k), \quad k=0, 1, \dots, N_0-2 \\ Y'(0 | 0) = 0. \end{cases} \quad (A.13)$$

Proof.

$$Y'_{k+1}{}^* - \psi(k, \mathbf{p})Y'_k{}^* - \frac{1}{2}G(k)QG(k)v_{k+1} = 0 \quad (A.14)$$

\Rightarrow (A.10)-(A.11)

$$Y'(k+1 | k+1) + \frac{1}{2}C(k+1)\psi(k+1, \mathbf{p})^T v_{k+2} - \psi(k, \mathbf{p})Y'(k | k) - \frac{1}{2}P(k+1)v_{k+1} = 0 \quad (A.15)$$

\Rightarrow (A.5)-(A.6)

$$\begin{cases} Y'(k+1 | k+1) - \psi(k, \mathbf{p})Y'(k | k) + \frac{1}{2}[C(k+1) - P(k+1)]v_{k+1} - C(k+1) \times H(k+1)^T R^{-1} [z_{k+1} - H(k+1)\psi(k, \mathbf{p})Y'_k{}^*] - \frac{1}{2}H(k+1)G(k+1)QG(k+1)^T v_{k+1} = 0 \\ Y'(k+1 | k+1) - \psi(k, \mathbf{p})Y'(k | k) + \frac{1}{2}[C(k+1) - P(k+1)]v_{k+1} = 0, \end{cases} \quad (A.16)$$

where the equation above is valid for $k=N_0-1, N_0, \dots, N-1$ and the equation below for $k=0, 1, \dots, N_0-2$.

\Rightarrow (A.5)-(A.6)

$$\begin{cases} Y'(k+1 | k+1) - \psi(k, \mathbf{p})Y'(k | k) - C(k+1) \times H(k+1)^T R^{-1} [z_{k+1} - H(k+1)\psi(k, \mathbf{p})Y'(k | k)] + \frac{1}{2}[C(k+1) - P(k+1) + C(k+1) - H(k+1)^T R^{-1} H(k+1)P(k+1)]v_{k+1} = 0 \\ Y'(k+1 | k+1) - \psi(k, \mathbf{p})Y'(k | k) + \frac{1}{2}[C(k+1) - P(k+1)]v_{k+1} = 0. \end{cases} \quad (A.17)$$

The fact that $P(k+1)$ and $C(k+1)$ are the covariances of $\psi(k, \mathbf{p})\mathbf{Y}_k + G\mathbf{w}_k$ and \mathbf{Y}_{k+1} respectively makes $P(k+1)$ and $C(k+1)$ symmetric. Moreover, in the absence of data for $k < N_0$, $P(k) = C(k)$. If we apply this to the last equations, we can complete the proof:

$$\begin{cases} \mathbf{Y}'(k+1 | k+1) - \psi(k, \mathbf{p})\mathbf{Y}'(k | k) \\ - C(k+1)H(k+1)^T R^{-1} \\ \times [\mathbf{z}_{k+1} - H(k+1)\psi(k, \mathbf{p})\mathbf{Y}'(k | k)] = 0 \\ \mathbf{Y}'(k+1 | k+1) - \psi(k, \mathbf{p})\mathbf{Y}'(k | k) = 0. \end{cases} \quad (\text{A.18})$$

Now, suppose that the data $\{\mathbf{z}_k | k = N_0, N_0 + 1, \dots, N\}$ are recursively processed by equation (43), we finally establish the equivalence of $\mathbf{Y}'(N | N)$ and \mathbf{Y}'_N^* because $\mathbf{v}_{N+1} = \mathbf{0}$. This, on its turn, immediately implies that $q(k | N) = q^{*k}$ for all k .

The analogon of equation (14), in case only one stochastic process is introduced, is

$$\begin{bmatrix} X_{k+1}^* \\ q_{k+1}^* \end{bmatrix} = \begin{bmatrix} F_1(\mathbf{p}) & \rho G_1 \\ 0 & \rho \end{bmatrix} \begin{bmatrix} X_k^* \\ q_k^* \end{bmatrix} + \begin{bmatrix} B_1 \\ 0 \end{bmatrix} \mathbf{u}_{k+1} + \frac{1}{2} \begin{bmatrix} G_1 \\ 1 \end{bmatrix} Q [G_1^T \ 1] \begin{bmatrix} (\mathbf{v}_x)_{k+1} \\ (\mathbf{v}_q)_{k+1} \end{bmatrix} \quad (\text{A.19})$$

from which it follows that

$$q(k+1 | N) = q^{*k+1} = \rho q^{*k} + \frac{1}{2} Q G_1^T (\mathbf{v}_x)_{k+1} + \frac{1}{2} Q (\mathbf{v}_q)_{k+1} \quad (\text{A.20})$$

and

$$\mathbf{X}^{*k+1} = F_1(\mathbf{p})\mathbf{X}^{*k} + B_1 \mathbf{u}_{k+1} + G_1 q(k+1 | N), \quad (\text{A.21})$$

which completes the proof of the proposition. \square