

Natural Resource Modeling

The Ensemble Kalman Filter for Multidimensional Bioeconomic Models

Journal:	<i>Natural Resource Modeling</i>
Manuscript ID:	NRM2014-12-088.R2
Manuscript Type:	Original Article
Date Submitted by the Author:	n/a
Complete List of Authors:	Kvamsdal, Sturla; SNF-Centre for Applied Research at NHH, Sandal, Leif Kristoffer; NHH Norwegian School of Economics, Department of Business and Management Science
Keywords:	Ensemble Kalman filter, State Space Model, Ecosystem-Based Management, Multidimensional Models, Bioeconomics

1
2
3
4
5
6
7
8
9
10
11 The Ensemble Kalman Filter for
12
13
14 Multidimensional Bioeconomic Models
15
16
17
18
19

20 Sturla Furunes Kvamsdal (corresponding author)

21 SNF-Centre for Applied Research at NHH

22 Helleveien 30, N-5045 Bergen, Norway

23
24
25
26 sturla.kvamsdal@nhh.no

27
28
29 Tel.: 004755959979

30
31
32 Leif Kristoffer Sandal

33
34 NHH Norwegian School of Economics

35
36 Helleveien 30, N-5045 Bergen, Norway
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

July 16, 2015

Abstract

To integrate economic considerations into management decisions in ecosystem frameworks, we need to build models that capture observed system dynamics and incorporate existing knowledge of ecosystems, while at the same time accommodating economic analysis. The main constraint for models to serve in economic analysis is dimensionality. In addition, to apply in long-term management analysis, models should be stable in terms of adjustments to new observations. We use the ensemble Kalman filter to fit relatively simple models to ecosystem or foodweb data and estimate parameters that are stable over the observed variability in the data. The filter also provides a lower bound on the noise terms that a stochastic analysis requires. In the present article, we apply the filter to model the main interactions in the Barents Sea ecosystem. In a comparison, our method outperforms a regression-based approach.

Keywords: Barents Sea, Bioeconomics, Ecosystem-Based Management, Ensemble Kalman Filter, Multidimensional Models, State Space Model

1 Introduction

The ecosystem approach to fisheries management has been broadly embraced at both the political and scientific levels (UN 2002, Olsen *et al.* 2007, Holland *et al.* 2010, Link 2010), but adoption in real world situations is limited (Vert-pre *et al.* 2013, Skern-Mauritzen *et al.* 2015). While both empirical and theoretical evidence demonstrate the insufficiency of management regimes based on single species models (see Link 2010 and Skern-Mauritzen *et al.* 2015 and references therein), single species models remain at the center of attention for a number of reasons. Two central reasons are (i) the limited (in most cases) knowledge of the extent and importance of ecological and economic interactions and (ii) the complexity that arises in models with such interactions. Notwithstanding, the future lies in co-management and multidimensional analysis, and we need methods and techniques to harness multispecies models for use in fisheries management (Link 2010, Peck *et al.* 2014).

Multispecies models provide insights into how a group of populations responds to a process; such models are critical for better management and evaluation of multispecies tradeoffs and tradeoffs between different user sectors (Link 2010, p. 100). Multispecies models can also address dynamics that are hard to deal with consistently in single species models. Examples are interaction-induced critical depensation, effective carrying capacity (which depends on interactions and on the state of the system), dynamic interactions, and consistent system responses. Sandal and Steinshamn (2010) is a case in point, where highly nonlinear harvest profiles are derived for a multispecies model. The harvest profiles have features hardly imaginable in a single species model, such as a declining harvest rate with increasing stock level in parts of the state space.

To construct tractable multispecies fisheries models, we need estimation techniques that deal with parameter and model uncertainty. In our view, relevant models should submit to dynamic decision analysis, which implies a limitation in the number of dynamic variables

1
2
3 (that is, dimensionality). At the same time, we want models to capture as much as possible of
4
5 the system structure and dynamics, including nonlinear and chaotic behavior. We suggest
6
7 applying the ensemble Kalman filter (Burgers *et al.* 1998, Evensen 2003), a data assimilation
8
9 method that has seen wide application in meteorology, oceanography, and other fields that are
10
11 concerned with chaotic, nonlinear, and stochastic dynamic systems (Evensen 2009).
12

13
14 We use the ensemble Kalman filter to fit a marine ecosystem model to data.
15
16 Evensen (2003, 2009) reviews both theoretical developments and applications of the
17
18 ensemble Kalman filter and related methods. This literature extends the Kalman filter to a
19
20 large class of nonlinear models. The fundamental idea is to use a Markov Chain Monte
21
22 Carlo approach to solve the Fokker-Planck equation that governs the time evolution of the
23
24 model. The model is formulated as a stochastic differential equation, and both the model
25
26 and observations are assumed to have error. The method facilitates simultaneous model
27
28 fitting and parameter estimation. Simultaneity is consistent with viewing the problem
29
30 as a combined state and parameter estimation problem, which differs from traditional
31
32 approaches that essentially ignore model error (typically, parameters are first estimated
33
34 and then model predictions are calculated in a deterministic fashion; for further details and
35
36 discussion, see Evensen 2009, pp. 95–ff.). With the ensemble Kalman filter, relatively simple
37
38 models can capture much of the complexity observed in marine ecosystems. We briefly
39
40 describe the ensemble Kalman filter and apply it to a three-species model of the Barents
41
42 Sea ecosystem. We compare our results with those of a regression-based approach.
43
44
45
46

47
48 Several different data assimilation methods, usually variational adjoint methods,
49
50 have been suggested to fit aggregated biomass models to data (see Ussif *et al.* 2003,
51
52 and references therein). Grønnevik and Evensen (2001) applied different ensemble-based
53
54 data assimilation techniques to age-structured fish stock assessment models, the
55
56 ensemble Kalman filter among them. An advantage of the ensemble Kalman filter when
57
58
59
60

1
2
3 compared to variational adjoint methods is that it does not rely on direct optimization, and
4
5 all observations are not processed simultaneously, which reduces computational demand.
6
7 Instead, variable and parameter estimates are updated sequentially according to the
8
9 filtering procedure. The ensemble Kalman filter also facilitates flow-dependent noise
10
11 attribution; flow-dependent (or rather, state-dependent) noise processes, it turns out,
12
13 are fundamental in capturing the dynamics of marine ecosystems.
14
15

16
17 If, as in Ussif *et al.* (2003), there is a known or easily identified functional
18
19 relationship between biological variables and the exploitation strategy, the filter can also
20
21 estimate economic parameters (for example, the exploitation rate). Another feature of the
22
23 ensemble Kalman filter is that it readily applies to both continuous and discrete time
24
25 formulations. This is a useful feature, because, while much bioeconomic modeling uses
26
27 discrete time models, a substantial share of related work in optimization relies on
28
29 continuous time. The difference in models of time may seem innocuous, but the
30
31 corresponding discrete time model of a nonlinear model in continuous time is rather
32
33 complex, while the corresponding continuous time model of a nonlinear model in discrete
34
35 time cannot be uniquely determined because of missing information about the dynamic
36
37 behavior between discrete observation times.
38
39

40
41 The ensemble Kalman filter fits, in an efficient manner, nonlinear aggregated
42
43 biomass ecosystem models to data. It also estimates the model error, which can be
44
45 translated into uncertainty in model predictions. Combined with developments in high-
46
47 dimensional, stochastic optimization, the filter can make bioeconomic analysis relevant
48
49 for real-world fisheries management decisions. Perhaps the main critique of bioeconomics,
50
51 the over-simplification of biological models, loses much of its force when the explanatory
52
53 power of the fitted biomass models matches, and even competes with, that of age-
54
55 structured and synthetic models. The potential of the ensemble Kalman filter reaches
56
57
58
59
60

1
2
3 further. It has the ability to process large amounts of data in high-dimensional systems
4
5 with large numbers of poorly known parameters (see Evensen 2003 and references
6
7 therein) and it should be of interest to researchers working with large and volatile
8
9 systems.
10

14 2 The Ensemble Kalman Filter

16 Our theoretical presentation of the ensemble Kalman filter is based upon Evensen (2003,
17
18 2009). Our point of departure is the continuous time state space model:
19

$$20 \quad dx = f(x) dt + \sigma(x) dB \quad (1)$$

$$21 \quad y = M(x) + v \quad (2)$$

22
23
24
25 An incremental change dx in the state variable (or n -vector) x is the sum of the drift
26
27 term $f(x) dt$ and the stochastic diffusion term $\sigma(x) dB$. The diffusion term represents
28
29 model error, which is composed of inadequacy in $f(x)$, potential parameter uncertainty,
30
31 and stochastic drivers. When x is an aggregated biomass vector, $f(x)$ is the multi-
32
33 dimensional growth function ($f: \mathbb{R}^n \rightarrow \mathbb{R}^n$). $\sigma(x)$ is generally a matrix ($\mathbb{R}^r \rightarrow \mathbb{R}^n$;
34
35 elements may depend on x) and the r -dimensional stochastic, Brownian increments in dB
36
37 are independent, identical, and normally distributed with mean zero and variance dt . That
38
39 is, stochastic drivers in equation (1) are Gaussian white noise. The measurement functional
40
41 $M(x)$ relates the state vector to the observations y . When the state vector is directly
42
43 observed, the measurement functional is the identity operator. v is a normally
44
45 distributed error term with mean zero and covariance R . Equation (1) is called the state
46
47 equation; equation (2) is called the measurement or observation equation.
48
49
50
51

52
53 The ensemble Kalman filter is a sequential method and works as follows. The
54
55 model is integrated forward in time until measurements become available.

56
57 Measurements are used to update the model. The updated model is then further
58
59
60

integrated until the next measurement time. In the theoretical literature, the updating step is called the analysis, thus the notation x^a for the updated state vector. The forward integrated model (the forecast) is denoted x^f . C^f is the covariance of the model forecast; C^a is the covariance of the model analysis.

The ensemble Kalman filter uses, as the name suggests, an ensemble of model states: a cloud of points in state space, to represent the probability density function at any given time. Each ensemble member represents a realization of the state vector from the underlying distribution, and, with a large enough ensemble, it conveys distributional properties. With a Markov Chain Monte Carlo method (meaning that the model can be formulated as a Markov Chain and that a large number of simulated solutions are considered; see Evensen 2009), each ensemble member is integrated forward in time according to (1). Errors are simulated. The integrated ensemble represents a forecast of the probability density and the only approximation is the limited number of ensemble members (Evensen 2009, p. 47). The Markov Chain Monte Carlo method is the backbone of the ensemble Kalman filter and is equivalent to solving the Fokker-Planck equation for the time evolution of the probability density; see Evensen (2003, p. 348) for further details.

When measurements are available, each ensemble member is updated as a linear weighting between the forecast and the measurements:

$$x^a = x^f + K(y - M x^f) \quad (3)$$

The weight K is called the Kalman gain. Assuming M is the identity operator, we see that, with $K = 0$, no weight is put on the observation y ; with $K = I$ (the identity operator), no weight is put on the forecast x^f . The Kalman gain is given by:

$$K = C^f M^T (M C^f M^T + R)^{-1} \quad (4)$$

where we assume that M is a linear operator (a matrix); M^T denotes its transpose. It is crucial that observations are treated as uncertain ($R > 0$), and, therefore, in the

ensemble Kalman filter, the observation probability density is represented by an ensemble; in other words, observations are perturbed (Burgers *et al.* 1998, pp. 1720-1721). It is convenient to let the number of ensemble members in the observation ensemble, denoted Y , equal the number of ensemble members in the state space ensemble, denoted X .

In the standard Kalman filter, both the forecast and analysis covariance (C^f and C^a) are in principle unknown; they are defined in terms of the unknown true state (see Evensen 2003, p. 347). In the ensemble Kalman filter, they are defined in terms of the ensemble means (\mathbf{E} denotes the mean or expected value):

$$C_e^f = \mathbf{E}[(X^f - \mathbf{E}[X^f])(X^f - \mathbf{E}[X^f])^T] \quad (5)$$

$$C_e^a = \mathbf{E}[(X^a - \mathbf{E}[X^a])(X^a - \mathbf{E}[X^a])^T] \quad (6)$$

That is, covariances are represented by the ensemble moments that carry the subscript e .

The observation covariance is also represented by the ensemble moment:

$$R_e = \mathbf{E}[(Y - y)(Y - y)^T] \quad (7)$$

The observation ensemble is defined such that it has the true (given) observation as its mean: $\mathbf{E}[Y] = y$. The ensemble Kalman gain is defined as

$$K_e = C_e^f M^T (M C_e^f M^T + R_e)^{-1} \quad (8)$$

We assume that the ensemble is large enough that $M C_e^f M^T$ and R_e are nonsingular; see Evensen (2003, p. 349). The analysis step (3) for ensemble member j is given by:

$$X^a(j) = X^f(j) + K_e(Y(j) - M X^f(j)) \quad (9)$$

It can be shown that, by updating the ensemble with the perturbed observations Y , the updated ensemble X^a has the correct error statistics (Evensen 2003, p. 349). The analysis covariance can be written as

$$C_e^a = (I - K_e M) C_e^f \quad (10)$$

which is equivalent to the standard Kalman filter expression for the covariance matrix.

See Evensen (2003) for derivations and further discussion.

1
2
3 The filter can estimate parameters by adding the parameters to the state vector –
4 in essence, by adding dimensions to the state space. Parameters are treated as
5 unobserved, constant model states, which implies they are assumed to have zero drift and
6 diffusion terms (Hansen and Penland 2007, Kivman 2003). With parameters in the state
7 space, involved operators must adapt to make them compatible with the extended state
8 vector. The distribution of the ensemble members in the relevant dimension of the state
9 space represents the conditional probability density function of the parameter. We
10 interpret the mean of the ensemble as the estimate and the spreading of the ensemble as a
11 measure of the estimate uncertainty.
12
13
14
15
16
17
18
19
20
21
22

23 The ensemble Kalman filter estimates state variables and parameters
24 simultaneously. As Evensen (2009, pp. 95-97) points out, the approach represents an
25 improvement to more traditional approaches that ignore model error and stochastic,
26 dynamic noise. The sequential nature of the approach yields, for each observation time t ,
27 parameter estimates conditional upon observations up until t ; estimates for the last
28 observation are conditional upon all observations and are usually the estimates of
29 interest. In situations where regime shifts or similar situations occur, one should inspect
30 the behavior of the sequential parameter estimates.
31
32
33
34
35
36
37
38
39
40

41 While the filter does not directly estimate the scaling of the diffusion term in
42 (1), the estimated C_e^a can be used to infer the appropriate noise scaling. C_e^a estimates
43 the second moment of the density of the state vector at a given moment in time (at, say, t).
44 C_e^a will vary with time (it is dynamic or flow-dependent; dynamic covariance is an
45 advantage with the ensemble Kalman filter over variational methods). The second
46 moment of the state vector density can be interpreted as the uncertainty in the
47 estimated state conditional upon the state at $t - 1$ and the uncertain observation at t .
48 The uncertainty in the state estimate accounts for parameter uncertainty, observational
49
50
51
52
53
54
55
56
57
58
59
60

1
2
3 uncertainty, and model inadequacy; the latter is what the diffusion term in (1)
4 represents. Thus, if the covariance is stable over time, or if it is stable after controlling
5 for some assumed functional form of the scaling term, such as $\sigma(x) = \sigma_0 \cdot x$, C_e^a can be
6 interpreted as an estimate of $\sigma(x)$ (or σ_0). How C_e^a varies over time maps out the
7 distribution of σ_0 ; that is, we essentially follow Hansen and Penland (2007).
8
9
10
11
12
13

14 The initial ensemble should reflect beliefs about the initial state of the system
15 (Evensen 2003, p. 350). The filter can be initialized by specifying means and standard
16 deviations that characterize the initial ensemble. In the case of unknown parameters,
17 initialization is not necessarily straightforward. Our experience is that, with large
18 enough standard deviations, such that the initial ensemble covers all eventualities, and
19 with enough ensemble members, it is possible to find reasonable traits of the initial
20 ensemble. Often, theory and earlier results are available to guide this process.
21
22
23
24
25
26
27
28
29

30 For a given time t , the ensemble Kalman filter provides an estimate of the state
31 of the system and its parameters conditional upon observations up until t . By
32 smoothing the filter estimates, we obtain estimates conditional upon all observations
33 (Evensen and van Leeuwen 2000). The filter and smoother estimates for the final
34 observation are identical, and the smoothed parameter estimates are constant through
35 time. The ensemble Kalman smoother can be formulated as a sequential method and in
36 terms of the filter analysis; see Evensen (2003, p.360) for details. That smoother
37 parameter estimates are constant and identical to the final filter estimates follows from
38 the explicit modeling of parameters as deterministic but unknown constants (see
39 Hansen and Penland 2007 and Kivman 2003) and is straightforward from the
40 formulation in terms of the filter estimates; see Evensen (2009) for details. The
41 ensemble Kalman smoother is particularly useful in problems involving unknown
42 parameters, as it provides estimates of the state variables, conditional upon all
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

1
2
3 observations, and upon parameter estimates, also conditional upon all observations. In
4
5 contrast, the filter provides, for a given t , state estimates conditional upon observations up
6
7 until t and upon parameter estimates conditional upon observations up until t ; these
8
9 estimates clearly are poor before the parameter estimates converge.
10

11 To summarize, the ensemble Kalman filter can be interpreted as a statistical Monte
12
13 Carlo method where the ensemble evolves in state space with the mean as the best
14
15 estimate and the spreading of the ensemble as the error variance (Burgers *et al.* 1998, p.
16
17 1720). For many problems, the sequential processing of observations proves to be a
18
19 better approach than the simultaneous processing that is typical in variational methods
20
21 (Evensen 2009, p. 101).
22
23
24
25
26

27 3 The Barents Sea Model

28
29 The Barents Sea is one of the most productive ocean areas in the world, and is subject to
30
31 extensive research (Gjøsæter *et al.* 2009, Huse *et al.* 2004, Durant *et al.* 2008; see also
32
33 further references therein). The commercially most important stocks are cod (*Gadus*
34
35 *morhua*) and capelin (*Mallotus villosus*); cod is highly valued as human food and capelin is
36
37 an important part of the cod diet. Capelin is also caught for fishmeal and oil production.
38
39 Juvenile herring (*Clupea harengus L.*) enters the Barents Sea when large year-classes arise
40
41 in the Norwegian Sea. Herring has an important influence on the ecosystem; it is preyed
42
43 on by cod while it preys on capelin larvae. We limit our model to these three fish stocks
44
45 for two main reasons. First, our model captures the dynamics of the cod stock to a high
46
47 degree, and the cod fishery, as the most important fishery in the region, is our main
48
49 interest. Second, for the model to be relevant for bioeconomic analysis, we have to limit
50
51 its complexity and dimensionality. We have in mind the type of analysis carried out in
52
53 Sandal and Steinshamn (2010) and Poudel *et al.* (2012); see also Kugarajh *et al.* (2006).
54
55
56
57
58
59
60

To limit complexity, we use simple growth functions and interaction terms common in traditional bioeconomic analysis. While dimensionality is based upon technical limitations, we find comfort in the view promoted by Holling and Meffe (1996, p. 333) that the driving forces of an ecosystem are confined to a relatively small subset of variables and relationships. While our choice of variables and relationships does not contain all driving forces of the Barents Sea ecosystem, we observe that our model captures much of the variation detected in stock assessments.

3.1 The State Space Model

The biomass of the three stocks are the state variables; cod is denoted x_1 , capelin is denoted x_2 , and herring is denoted x_3 . Both cod and capelin are harvested in the Barents Sea; h_1 and h_2 denote harvest rates of cod and capelin. Herring is not harvested in the Barents Sea, but eggs and larvae flow in from the Norwegian Sea. The best model would feature a lagged inflow variable (see discussion below), but a lagged variable would mean that the model is not autonomous; to accommodate dynamic analysis, the model should be autonomous. Thus, we use x_3 as a proxy for the herring inflow; again, see discussion below. Finally, we denote parameters c_i and vectors in boldface. The dynamic model for the system is written in differential form:

$$dx_1 = (f_1(x_1, c_1, c_2) + g_1(x_1, x_2, c_3) + g_2(x, c_4) - h_1)dt + \sigma_1(x)dB_1 \quad (11)$$

$$dx_2 = (f_2(x_2, c_5, c_6) - g_1(x_1, x_2, c_7) - g_1(x_2, x_3, c_8) - h_2)dt + \sigma_2(x)dB_2 \quad (12)$$

$$dx_3 = (f_2(x_3, c_9, c_{10}) + g_1(x_2, x_3, c_{11}) - g_2(x, c_{12}) + c_{13}x_3)dt + \sigma_3(x)dB_3 \quad (13)$$

where growth functions are denoted f_i and interaction terms are denoted g_i . Table 1 reports functional forms that we discuss further below. The stochastic increments dB_i are independent, with mean zero and variance dt . The scaling term $\sigma_i(x)$ reflect correlations in the noise processes. Two principal models of the scaling term were tried: white noise

1
2
3 $(\sigma_i(x) = \sigma_{0i})$ and state-dependent white noise $(\sigma_i(x) = \sigma_{0i} \cdot x = \sigma_{0i1}x_1 + \sigma_{0i2}x_2 +$
4
5 $\sigma_{0i3}x_3)$.

6
7 The first terms in each model equation are the growth functions. The growth functions
8 model the growth that does not happen through the modelled interactions. For cod (11),
9 we use the logistic growth function; for the pelagic stocks capelin (12) and herring (13), we
10 use the modified logistic growth function (see Table 1 for specifications). The related
11 parameters ($c_1, c_2, c_5, c_6, c_9,$ and c_{10}) are interpreted accordingly. (The idea of carrying
12 capacity, that is, the standard interpretation of the second parameter in the logistic and
13 modified logistic, becomes unclear in an ecosystem setting. The capacity of the ecosystem
14 to harbor any one species depends on the state of the entire system. Hence, intrinsic single
15 species notions such as carrying capacity must be treated with caution in our multispecies
16 approach.) While the logistic is the standard assumption, we found that the modified logistic
17 better captures the tendency of the pelagic stocks to remain at low stock levels for extended
18 periods before rapidly growing to high levels.
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33

34 All species interactions in the system are predator-prey relationships. Cod preys upon
35 both herring and capelin, while herring preys upon the capelin stock. (A competitive,
36 mutually destructive interaction between the pelagic species is an interesting alternative
37 model that was estimated, together with a few other specifications. The model reported here
38 provided the best overall fit to the data. Brief descriptions of, and results from, alternative
39 models are available from the authors.) The interaction terms are per definition positive,
40 and corresponding terms (cod-capelin corresponding to capelin-cod, for example) have
41 opposite signs. The capelin-cod and capelin-herring interaction terms ($g_1(\cdot)$) are inspired by
42 the crude Lotka-Volterra form of predator-prey interaction (May *et al.* 1979, p. 268), where
43 the product of the stock levels is adjusted by an intensity parameter. The functional form
44 of, for example, the capelin-cod interaction is $g_1(x_1, x_2, c_7) = c_7x_1x_2$, where c_7 is the
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

1
2
3 intensity parameter. We will discuss the interpretation of the interaction intensity
4
5 parameter further below.
6

7 While cod preys upon both capelin and herring, it has a tendency to prefer capelin when
8
9 both types of prey are available (Durant *et al.* 2008, Gjørseter *et al.* 2009). That is, the
10
11 predation pressure on herring from cod is reduced when capelin is abundant. Thus, we
12
13 modify the cod-herring interaction term to reflect cod's preference for capelin, and have
14
15 $g_2(\mathbf{x}, c_{12}) = c_{12}x_1x_3 \frac{x_3}{x_2+x_3}$. The fraction $\frac{x_3}{x_2+x_3}$ yields a model of preference such that, in
16
17 the extreme case with no capelin ($x_2 = 0$), the fraction equals one and the interaction term
18
19 is unmodified. When capelin is present ($x_2 > 0$), the fraction takes a value between zero
20
21 and one and the interaction (cod predation) is weaker.
22
23
24

25
26 As is evident from the model equations (11 - 13), the interaction terms g_1 and g_2
27
28 represent a biomass loss for the prey species and a biomass gain for the predator
29
30 species. The intensity parameters scale the product of biomass for each species in the
31
32 terms to account for the rate of biomass loss in the prey species. Biomass is not
33
34 conserved in the interactions, and the additional interaction parameters (c_3 , c_4 , and
35
36 c_{11}) reflect the loss of biomass in the interactions. The loss in the cod-capelin
37
38 interaction, for example, is given by $\frac{c_3}{c_7}$, and this and corresponding fractions are
39
40 expected to take values between zero and one. Because most of the biomass is lost,
41
42 these values are expected to lie closer to zero than one. We can think of these fractions
43
44 as biomass conversion rates between species. Presumably, regularities exist for
45
46 biomass conversion rates. While known or assumed interaction relationships would be
47
48 helpful in reducing the number of parameters in the model, biologists reject the idea of
49
50 stable relationships (S. Tjelmeland, personal communication). Thus, we refrain from
51
52 prescribing fixed conversion rates.
53
54
55
56
57
58
59
60

1
2
3 The final parameter, c_{13} , measures the influence of the inflow of herring on the herring
4 stock growth. Most of the time, the amount of herring biomass that enters the Barents Sea
5 is relatively small. Over a few years, however, the herring biomass grows substantially.
6
7 Thus, we could lag the inflow variable two years and multiply it with the scaling
8 parameter c_{13} . The idea is that three-year-old (and older) herring make up most of the
9 herring biomass in the Barents Sea, and the biomass influx two years earlier better
10 explains the change in the herring stock. As mentioned above, a lagged variable would
11 go against the need for an autonomous model. Thus, we use x_3 as a proxy for the inflow.
12
13 The correlation coefficient of x_3 and inflow lagged two years is 0.811 and is nonzero
14 with a p -value of $5.5 \cdot 10^{-14}$; therefore, we find x_3 to be a reasonable proxy. The
15 assimilation run with the lagged inflow yields only minor alterations to our results; these
16 results are available from the authors. (After three or four years in the Barents Sea, the
17 juvenile herring returns to its main habitat in the Norwegian Sea to mature and
18 eventually spawn; the herring growth rate in our model reflects the migration behavior.)
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33

34 To avoid negative parameters, parameters are all assumed to be log-normally
35 distributed. (Formally, they are treated as $c_i = \exp(\alpha_i)$, where each α_i is a stochastic
36 constant that is normally distributed.)
37
38
39
40

41 We treat estimates from stock assessments as measurements of the state variables, and
42 the measurement operator is thus the identity operator. Note that parameters are added
43 to the state vector as described above. We denote the extended state vector Ψ . The
44 measurement operator must thus be adjusted to be compatible with the state vector by
45 adding zeros. Parameters are treated as unobserved states. The observation equation
46 becomes
47
48
49
50
51
52

$$53 \quad \mathbf{y} = M\Psi + \mathbf{v} \quad (14)$$

54 where
55
56
57
58
59
60

$$\Psi = \begin{bmatrix} x_i \\ c_j \end{bmatrix}, \quad i = [1 \dots 3], \quad j = [1 \dots 13], \quad \text{and} \quad M = [\mathbf{I} \ \mathbf{0}] \quad (15)$$

\mathbf{I} is a three by three identity matrix and $\mathbf{0}$ is a three by thirteen null matrix. \mathbf{y} is a three-element vector of observations, and \mathbf{v} is the error term vector, which is normal, independent, and identically distributed with mean zero and variance \mathbf{R} .

3.2 An Alternative Model

For comparison, we set up a simple, alternative model in the spirit of much of the work in bioeconomic analysis. The crude idea is to fit the individual stock equations with ordinary least squares. The independent variables z_t are defined as observed changes in stock levels from year to year, controlling for harvest in the case of cod and capelin. We have $z_{i,t} = y_{i,t+1} - y_{i,t} + h_{i,t}$ for cod and capelin ($i = 1,2$); for herring (not harvested), we have $z_{3,t} = y_{3,t+1} - y_{3,t}$. The regression equations are as follows:

$$z_{1,t} = \gamma_1 + \beta_1 y_{1,t} + \beta_2 y_{1,t}^2 + \beta_3 y_{1,t} y_{2,t} + \beta_4 y_{1,t} y_{3,t} \frac{y_{3,t}}{y_{2,t} + y_{3,t}} + \epsilon_{1,t} \quad (16)$$

$$z_{2,t} = \gamma_2 + \beta_5 y_{2,t}^2 + \beta_6 y_{2,t}^3 + \beta_7 y_{1,t} y_{2,t} + \beta_8 y_{2,t} y_{3,t} + \epsilon_{2,t} \quad (17)$$

$$z_{3,t} = \gamma_3 + \beta_9 y_{3,t}^2 + \beta_{10} y_{3,t}^3 + \beta_{11} y_{2,t} y_{3,t} + \beta_{12} y_{1,t} y_{3,t} \frac{y_{3,t}}{y_{2,t} + y_{3,t}} + \beta_{13} x_{3,t} + \epsilon_{3,t} \quad (18)$$

The parameters denoted γ_i are the intercepts and are added to the otherwise unchanged equations for the main model. Thus, the parameters β_i refer to the same terms as do the parameters c_i above.

The model in equations (16-18) is formulated in discrete time, and the model in general and the parameters in particular are not directly comparable to those in the continuous time formulation (11-13). The comparison is nevertheless not completely irrelevant; as discussed above, fitting of continuous time model formulations in discrete time frameworks is widespread in the bioeconomic literature.

3.3 Data

The fish stocks in the Barents Sea cannot be observed directly. However, the Institute of Marine Research in Bergen and the Knipovich Polar Research Institute of Marine Fisheries and Oceanography in Murmansk carry out extensive, yearly ecosystem surveys. Based upon these surveys, they provide yearly estimates of the stock levels of all the important species in the Barents Sea. The stock estimates are published by the International Council for the Exploration of the Sea (ICES), and most of our data are collected from the ICES online database. We treat the stock estimates as observations. Notably, Ekerhovd and Gordon (2013) raise issues with stock estimates from virtual population models. We share their concern about the consistency in the stock estimates, but find it beyond our scope to apply the (Ekerhovd and Gordon 2013) adjustment here. Uncertainty in stock assessments is unfortunately not reported, and we are left to speculate. The herring inflow data was provided by S. Tjelmeland (personal communication).

We have stock estimates, catch data and herring inflow estimates from 1950 to 2007. However, the ICES database does not contain data on capelin prior to 1972. For the period prior to 1972, we collected catch data from Røttingen and Tjelmeland (2008, see Figure 2). Capelin stock estimates were collected from Marshall *et al.* (2000, see Figure 1, p. 2435). The early capelin stock estimates are more uncertain than later estimates, and we assume a 50% increased observation uncertainty on the capelin stock data prior to 1972.

All data are visually presented in Figure 1, with error bars showing assumed observation uncertainty. All numbers are given in tonnes.

3.4 Estimation Strategy and the Initial Ensemble

1
2
3 While the success of our approach hinges to some degree on reasonable characteristics of
4 the initial ensemble, it is not immediately clear what constitutes reasonable characteristics.
5
6 While we can rely on external, empirical evidence for a few of the parameters in the
7
8 interaction terms, we must produce reasonable initial ensemble characteristics for most
9
10 parameters in a heuristic fashion. The parameter subspace has thirteen dimensions (one
11
12 for each parameter), and, while it is not impossible to search, via trial and error, the
13
14 parameter subspace for an appropriate, initial ensemble, the high dimensionality makes
15
16 the approach unlikely to succeed. (Our main metrics of appropriateness are whether the
17
18 state estimates resemble the stock assessment data and to what degree the spread of the
19
20 ensemble in the parameter dimensions contracts over time. In addition, we have used the
21
22 Bayesian Information Criterion (BIC), but carefully, since the criterion is not unique
23
24 because of the Monte Carlo element of the filter (see Ekerhovd and Kvamsdal 2014). We
25
26 have also considered the distribution of the Kalman gain over time and stability of
27
28 parameter estimates.)
29
30
31
32
33

34 By first assimilating each equation individually, we reduce the dimensionality of the
35
36 relevant parameter subspace substantially. When we assimilate the cod equation (11),
37
38 for example, the state space consists of the cod stock level as the only state variable and
39
40 the four parameters in the equation ($c_1 - c_4$) as parameter variables. The variables x_2
41
42 and x_3 are treated as control variables.
43
44

45 We have good ideas about reasonable ensemble initializations of the biomass
46
47 conversion rates (limited support) and the interaction intensity parameters for the
48
49 cod-capelin and cod-herring interaction terms (empirical evidence). The capelin-
50
51 herring interaction intensity is assumed to be an order smaller than the cod-capelin
52
53 interaction intensity. Thus, when searching for reasonable initial ensemble
54
55 characteristics in the single equation assimilations, we need mostly to be concerned
56
57
58
59
60

1
2
3 with the parameters of the growth functions. What we have called the capacity
4
5 parameters are characterized by an ensemble mean higher than observed historic levels
6
7 (exploited fisheries usually have stock levels below their full capacity). To find reasonable
8
9 characteristics for the ensembles along the growth rate dimensions, we consider a range
10
11 of levels and compare, as mentioned above, model fit, ensemble contraction, the Bayesian
12
13 Information Criterion, and the distribution of the Kalman gain. To demonstrate, we
14
15 briefly discuss an example of the procedure in Appendix A.2. Means and spreads of the
16
17 initial ensemble for the parameter dimensions in the single equation assimilations are listed
18
19 in Table A2 in the appendix.
20
21

22
23 The estimates from the single equation assimilations are used to characterize the mean
24
25 of the normal distributions from which we draw the initial ensemble for assimilation of the
26
27 full model. Exceptions are those parameters for which we have empirical support for the
28
29 initial ensemble characteristics. Ensemble spreads (standard deviations of distributions
30
31 from which initial ensembles are drawn) are also inherited from the single equation
32
33 assimilations, with the same exceptions.
34
35

36
37 The initial ensemble is drawn randomly from a multivariate normal distribution. For the
38
39 three state variables, we use the first observations as the mean of the initial ensemble and
40
41 30% of the first observations as the standard deviation.
42

43
44 The initial ensemble for the interaction intensity parameters c_7 , c_8 , and c_{12} were
45
46 characterized based upon empirical evidence. The term $g_1(\mathbf{x}, \mathbf{c}) = c_7 x_1 x_2$ in (12) reflects
47
48 the loss of capelin biomass from the interaction with cod. Gjørseter *et al.* (2009, see figure
49
50 5, p. 45) estimated, from stomach content data, the amount of capelin consumed by the
51
52 Barents Sea cod for the years 1984-2006. The consumption varies over time, as do the cod
53
54 and capelin stock levels. To get a reasonable initial measure of c_7 , we regressed the total
55
56 consumption of capelin on the product $x_1 x_2$ (without intercept; we also ran the regression
57
58
59
60

with the intercept, and with and without two outliers; slope estimates varied somewhat but the distribution finally used for the initial ensemble spanned all estimates). Notably, Gjøsaeter *et al.* (2009) provided us with data for 1984-2007 (that is, one more year of data than what they based their original analysis upon). The estimated coefficient was $3.46 \cdot 10^{-10}$ (standard error $5.1 \cdot 10^{-11}$, R_{adj}^2 0.63). Similar data for the capelin-herring interaction are not available. Herring is, however, thought to have a smaller predation rate on capelin than does cod; we set the implied mean for c_8 at 10% of the implied mean of c_7 . For the herring-cod interaction intensity parameter c_{12} , data are available. Gjøsaeter *et al.* (2009) also estimated the amount of herring consumed by the Barents Sea cod. Regressing the consumed amount of herring on the term $x_1 x_3 \frac{x_3}{x_2 + x_3}$ yielded a coefficient of $3.03 \cdot 10^{-11}$ (standard error $6.46 \cdot 10^{-12}$, R_{adj}^2 0.44). As with c_7 , we set the mean of the initial shadow parameter (α_{12}) ensemble to correspond to the estimated coefficient. In comparison, regressing on the term $x_1 x_3$ produces the coefficient $2.49 \cdot 10^{-11}$ (standard error $3.76 \cdot 10^{-12}$, R_{adj}^2 0.61).

The additional interaction parameters c_3 , c_4 , and c_{11} (reflecting biomass conversion rates) cannot be larger than the corresponding parameters c_7 , c_8 , and c_{12} , as it is assumed that some biomass is lost in the interactions. The biomass loss assumption is not explicitly enforced, but initial implied ensemble means for the three additional interaction parameters were set to 25% of c_7 for c_3 , 10% of c_{12} for c_4 and 10% of c_8 for c_{11} . Typically, one assumes that 90% of the biomass is lost between trophic levels, but cod is particularly adapted to catching capelin and thus we specified a higher biomass conversion rate for the cod-capelin interaction.

We discuss further implementation details in Appendix A.1.

4 Results

1
2
3 Table 2 reports parameter estimates, with standard errors in parentheses, for the single
4 equation assimilations. Table A2 in the appendix reports the prior characterizations for
5 comparison. The third column ('Contraction') in Table 2 reports the standard error of the
6 estimates as a fraction of the standard deviation of the prior distribution. The ensemble
7 Kalman filter will mechanically contract parameter ensembles, but the amount of
8 contraction depends on the amount of information the filter retains. Assessing the
9 contraction is equivalent to comparing the width of the parameter confidence intervals at
10 the beginning and end of the assimilation. Both tables and also subsequent tables report
11 estimates of the shadow parameters α_i . However, our interest lies with the parameters
12 $c_i = \exp(\alpha_i)$, and Table 2 reports what we call the c -interval, which is the two standard
13 error interval around the mean estimate of the underlying parameter c_i .
14
15
16
17
18
19
20
21
22
23
24
25
26

27 We also calculate an estimate and standard error of the parameters in the diffusion
28 terms. We denote the parameters $\sigma_{i,i}$, where the subscripts denote the relevant state
29 variable. Table 2 reports the results.
30
31
32
33

34 Further, Table 2 reports the BIC-scores, the average root mean squared innovations,
35 the difference R-squared (R_D^2 , appropriate for time series data, see Harvey 1984), the
36 Lilliefors k -statistic and the Ljung-Box q -statistic (with 15 lags) for each equation. A
37 few comments on these statistics are in order. The BIC-scores, both here and later, are
38 evaluated with a bandwidth of 200.000 (tonnes); see Ekerhovd and Kvamsdal (2014) for
39 details. The innovation is the distance between the observations and the estimated state
40 variables. In our model, with the state-dependent noise scaling $\sigma \cdot x$, it is useful to
41 normalize the root mean squared innovations with the estimated state. Thus, what we
42 report as the average root mean squared innovation is the time-average of the following
43 expression
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

$$RMSI = \frac{\sqrt{E[(y - M\Psi_s^a)^2]}}{E[M\Psi_s^a]} \quad (19)$$

The subscript s denotes that it is the smoothed estimate which goes into the expression.

The lower the average root mean squared innovation, the better is the model fit. Note that, in the absence of the normalization issue, the average root mean squared innovation is the average distance between the ensemble members and the observation; if the observation and the ensemble mean are close, the average root mean squared innovation will be close to the estimate of the noise scaling term, which is derived from the second moment of the ensemble.

The Lilliefors k -statistic tests normality of error terms; the Ljung-Box q -statistic tests autocorrelation in errors. There are no signs of autocorrelation in any of the assimilations. Errors in the capelin equation deviate from normality to some degree (the p -value is slightly below the conventional 5% level). A density error plot shows that the deviation consists of an overrepresentation of large, negative errors. A scatterplot of the errors shows that large negative errors occur mainly at low stock levels and likely reflect that the model cannot predict exactly when the stock escapes near zero states.

To discuss the actual estimates in Table 2 is of limited interest; their main function is to serve as priors for the full model. We do note, however, that while the contraction rate is substantial for most other parameters, the interaction parameters (parameters 3, 4, 7, 8, 11, and 12) have not contracted much. As the full model results will show, contraction is somewhat better when we assimilate all equations simultaneously. The small contraction rates for the interaction parameters underline the need for informative priors.

The cod equation has the smallest average root mean squared innovation and has a similar BIC-score as that for the capelin equation. R_D^2 is much lower for the cod

1
2
3 equation than for the other equations, however. The high R_D^2 in the other equations likely
4
5 derives from relatively high gain values, which leads to state variable estimates close to the
6
7 observations. Comparing with results from the alternative (regression-based) model (see
8
9 Table 4), they all have lower R^2 -statistics. (We will discuss the results in Table 4 in more
10
11 detail below.) Finally, the estimated noise scaling parameters in Table 2 are highly significant
12
13 for all equations and underline the importance of the state-dependent noise processes.
14
15

16
17 Table 3 reports results for the full model assimilation. The BIC-score for the entire
18
19 model is 269.70, while R_D^2 is 0.9548. To properly test multivariate normality would require
20
21 more observations than are available here. The multivariate Ljung-Box q -statistic (Hosking
22
23 1980), with 15 lags, is 125.47 (p -value is 1.000; no signs of autocorrelation). Notably, the
24
25 prior for the full model assimilation is based upon the results reported in Table 2 for all
26
27 parameters, apart from the two parameters for which we have empirical evidence (c_7 and
28
29 c_{12}). For those parameters, we kept the original prior information as given in Table A2.
30
31

32
33 If we compare the contraction rates reported in Tables 2 and 3, we observe that,
34
35 overall, contraction is better in the full model assimilation for the capelin and herring
36
37 equation. In the cod equation, the interaction parameters have better contraction rates in
38
39 the full model assimilation, while the growth parameters contract better in the single
40
41 equation assimilation. That some growth parameters do not contract as much in the full
42
43 model assimilation is likely because most of the signal in the data about these parameters
44
45 is picked up in the single equation assimilation that was run prior to the full model
46
47 assimilation.
48
49

50
51 In Table 3, the average root mean squared innovations have improved considerably for all
52
53 state variables compared to the values in the single equation assimilations (Table 2). As
54
55 discussed above, the average root mean squared innovations can be close to the σ_0 estimate
56
57 if the ensemble mean is close to the observations. Further, significant cross-correlations
58
59
60

in σ_0 (the off-diagonal terms) may be challenging in model applications; as we report below, estimated cross-correlations are close to zero.

Below, we report estimates and standard errors of the noise scaling term σ_0 (20). All off-diagonal elements are statistically indifferent from zero, which suggests that there is little correlation in the different stochastic processes of the system. The diagonal elements are also relatively small, at least when compared to hypothetical scenarios studied in theoretical work (Poudel *et al.* 2012). The standard errors do not give a correct measure of the significance of the diagonal elements, as the elements are positive by definition, but nevertheless show that the system is stochastic. If one wishes to carry out studies of worst-case scenarios, it could be of interest to investigate whether high or low levels are correlated in time across equations.

$$\sigma_0 = \begin{bmatrix} 0.0802 & 0.0004 & 0.0008 \\ (0.0098) & (0.0034) & (0.0035) \\ 0 & 0.1113 & -0.0010 \\ & (0.0250) & (0.0030) \\ 0 & 0 & 0.0948 \\ & & (0.0047) \end{bmatrix} \quad (20)$$

Figure 1 shows the smoothed stock level estimates (solid curves) with two standard errors to each side (shaded areas) for all three state variables (top panel: cod; middle panel: capelin; bottom panel: herring). The figure also shows the observed stock levels (circles) with assumed observation uncertainty (the error bars show two standard deviations around the observations). Most observations lie within the four standard error band and the model captures most of the system dynamics. The smoothed parameter estimates are constant over time, and we interpret the smoothed estimates as model fit with uncertain but stable parameters (that is, as reported in Table 3).

Capelin stock data is more uncertain prior to 1972. As expected, the stock estimates have larger standard errors prior to 1972. Compare, for example, the width of the standard error band in Figure 1 (middle panel) in the years before and after 1972, or at

1
2
3 the peaks around 1970 and 1980, which are at roughly the same level. After 1972, the
4
5 capelin stock estimates, in addition to being more precise, lie closer to the measurements.
6

7
8 If we compare our estimates in the cod equation to those of Ussif *et al.* (2003),
9
10 there are considerable differences. (Our estimates are based on some ten years more of
11
12 data, but there is nothing spectacular about the last ten years of our data, perhaps apart
13
14 from considerable levels of herring.) Our central estimate for the carrying capacity is
15
16 6.5 million tonnes; their estimate is 5.2 million tonnes. As mentioned above, the
17
18 carrying capacity parameter in our model does not reflect the full capacity because
19
20 interaction terms may contribute. Thus, our estimate is a lower bound (perhaps the
21
22 lower end of the c -interval should be considered: 5.4 million tonnes). We estimate the
23
24 logistic growth rate in the cod equation at 0.55; Ussif *et al.* estimate it at 0.34. Our
25
26 estimate is larger and suggests that much larger harvest quotas can be sustainable. The
27
28 maximum of our estimated growth curve, interaction terms aside, is at 899 thousand
29
30 tonnes, while that of Ussif *et al.* (2013) is 460 thousand tonnes. In comparison, quotas
31
32 in recent years have exceeded our maximum. Thus, our estimates in the cod equation
33
34 seem to better align with reality. Published estimates of the remaining parameters were
35
36 not available to us.
37
38
39
40
41
42

43 4.1 Alternative Model Results

44
45 Ordinary least squares parameters estimates and related t -statistics, root mean squared errors,
46
47 R^2 -statistics, Lillifors k -statistics, and Ljung-Box q -statistics, for the individual equations (16
48
49 – 18), are reported in Table 4. The overall impression is a poor fit, with R^2 statistics barely
50
51 acceptable. There are no signs of autocorrelation, but errors are not normally distributed in the
52
53 cod (16) and herring (18) equations. Plots show that deviations are minor for the cod equation,
54
55 but severe deviations are present in the herring equation. Few parameter estimates are
56
57
58
59
60

1
2
3 statistically significant at the conventional 95% level. Most parameters are of the expected
4 sign in the cod and capelin equations (16 – 17), except for the insignificant capelin-cod
5 interaction parameter β_7 . The intercepts (γ_1 and γ_2) should be close to zero if the classical
6 growth model were to hold water, but they are not. The herring equation (18) seems entirely
7 driven by the inflow (β_{13}), with the rest of the equation seemingly representing noise. The
8 interaction parameters (β_{11} and β_{12}) are of the expected sign and somewhat removed from
9 zero; the growth parameters are essentially zero.

10
11
12
13
14
15
16
17
18
19 The root mean squared errors in Table 4 are reported in levels and do not, as such,
20 compare to the average root mean squared innovations reported in Table 3. More relevant is
21 the root mean squared error relative to the mean of the left-hand side variables (for equation i :
22 $RMSE_i/\bar{z}_{i,t}$), which evaluates to 0.4015, 2.673, and 52.44 in the three equations. Only the
23 first is similar to any of the average root mean squared innovations reported in Tables 2 and 3,
24 but is still higher than all of them. That $RMSE_i/\bar{z}_{i,t}$ evaluates as a static concept (all points
25 matter) while the average root mean squared innovations evaluates in a dynamic sense
26 (measured locally in time) reflect the larger point that, while ordinary least squares fits all
27 points simultaneously without regard for their order, the ensemble Kalman filter operates
28 sequentially and fits paths in the state space.

29
30
31
32
33
34
35
36
37
38
39
40
41 The carrying capacity in the cod equation in the alternative model (16) is estimated at
42 4.8 million tonnes; the growth rate at 0.35. The implied maximum of the growth curve, again
43 when interaction terms are ignored, is 436 thousand tonnes. As discussed above, the ensemble
44 Kalman filter estimates from the model (11 – 13) seem more realistic.

52 Conclusions

53
54 The ensemble Kalman filter relates structurally to the standard Kalman filter and the
55 extended Kalman filter in the sense that they minimize the variance of the state estimates.
56
57
58
59
60

1
2
3 However, the ensemble Kalman filter has some advantages. Unlike the extended Kalman
4 filter, it requires no linearization. It solves rank problems that may occur with large
5 numbers of observed variables. Unlike variational adjoint methods, it requires no adjoint
6 operator and is thereby simpler to implement, and it has flow-dependent (non-constant)
7 covariance. Further, the ensemble Kalman filter is well suited to large-scale problems, as it
8 extends to asynchronous and missing observations and other issues (Evensen 2009). On
9 the other hand, the ensemble integration (in the forecast step) can be computationally
10 costly and, with strongly nonlinear systems, iterative procedures called multiple data
11 assimilations hold better promise (Emerick and Reynolds 2012). As such, the ensemble
12 Kalman filter is just the tip of the iceberg of a range of related methods that apply to a
13 large range of problems (Evensen 2003).

14
15
16 In applying the ensemble Kalman filter, we have shown how relatively simple
17 aggregated biomass models, typical in bioeconomic analysis, can capture much of the
18 dynamics of ecosystems. When compared to earlier efforts of applying data assimilation
19 methods to bioeconomic models (Ussif *et al.* 2003), our model seems a better fit and
20 agrees better with stylized facts. The model presented here shows the most promise from
21 among a number of different specifications. (Among the specifications we tried were: capelin
22 and herring with common carrying capacity, that is, competitive, mutually destructive
23 interaction, notably with fewer parameters; pure, white – not level dependent – noise in the
24 error term; assumed perfect observations of the control variables; herring inflow as a state
25 variable; and herring inflow as white noise around a non-zero mean.) In the interest of space,
26 we cannot go into full details, but the alternative models have a number of undesirable
27 properties that, when added together, wipe out the advantage of, for example, fewer
28 parameters.

1
2
3 A prominent modeling possibility that could be explored is data timing. In our
4 current approach, we assume a constant harvest rate throughout each year. The
5 harvesting is more concentrated in winter and spring, however. Further, the stock
6 assessments are usually carried out in the fall. These nuances of timing could influence
7 the dynamics of the system were they taken into account. We have chosen not to go into
8 this in our current approach for two reasons. One is a need to limit the scope of our
9 present work. Second and more important, our current approach better serves the model
10 needs in a bioeconomic framework for decision and management analysis.

11
12 The main model does of course have room for other improvements. The c -intervals
13 for several of the parameters are not particularly tight, for example, and the estimates
14 of elements in the σ_0 matrix are not very precise. Based upon our experience, we
15 conclude that the best source of improvements would be more data. While some of the
16 series we use here extend further than what we utilize, herring inflow estimates are not
17 further available. Notwithstanding, estimates of parameters in chaotic systems are not
18 likely to be very precise, and management models should be flexible and adaptive
19 (Holling and Meffe 1996, p. 332). It is important that management models take the
20 uncertainty of the dynamics into account (Hill *et al.* 2007). Adaptive management
21 models such as feedback models are already well understood in the bioeconomic
22 literature (Sandal and Steinshamn 1997). The challenge is to solve models of higher
23 dimensionality that must underlie ecosystem-based management (Fulton *et al.* 2011).
24 We believe the ensemble Kalman filter has an important role to play in both
25 theoretical and operational management research, particularly in light of calls for
26 ecosystem-based management (Pew Oceans Commission 2003).

27
28 When we compare the ensemble Kalman filter to the standard approach that relies on
29 ordinary least squares, two major points emerge. First, the discrete time equivalent of a
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

1
2
3 model formulated in continuous time does not, in general, have the same functional form,
4
5 but this is nevertheless often assumed. If, as with certain discretization schemes, the discrete
6
7 time equivalent cannot be efficiently written in one equation, how to apply ordinary least
8
9 squares is unclear. If, on the other hand, one formulates the model in discrete time in order
10
11 to apply ordinary least squares, there is no unique representation of the discrete time model
12
13 in continuous time and, therefore, optimization schemes in continuous time, for example,
14
15 cannot be applied. Second, the ensemble Kalman filter fits the dynamics of the model in the
16
17 sense that it forces paths in the state space to track the observations. Ordinary least squares,
18
19 at least in the standard formulation, differentiates away the dynamics and is left with a cloud
20
21 of points where the order is no longer essential. Further, we have the flow-dependent
22
23 covariance (model error) that can be used to fit dynamic noise models and the explicit
24
25 incorporation of observational uncertainty; these features of the ensemble Kalman filter
26
27 have no equivalent in ordinary least squares, to our knowledge. In the final analysis, with
28
29 the ensemble Kalman filter, we obtain a model that represents the continuous time dynamics
30
31 of key species in the Barents Sea. Our adaptation of the standard approach is admittedly
32
33 naïve, but discrete time estimation methods based on ordinary least squares cannot provide
34
35 anything similar.
36
37
38
39

40
41 In a broader scope, we aim to answer calls for ‘flexible, adaptive, and experimental’
42
43 management models (Holling and Meffe 1996, p. 332), who further write that ‘effective
44
45 natural resource management that promotes long-term system viability must be based on an
46
47 understanding of the key processes that structure and drive ecosystems, and on acceptance of
48
49 both the natural ranges of ecosystems variation and the constraints of that variation for
50
51 long-term success and sustainability’ (p. 335). We think that, when models are simplified
52
53 and reduced down to the key driving phenomena, the ensemble Kalman filter can capture
54
55 both variation and stability in ecosystems and can serve in tractable management models.
56
57
58
59
60

Acknowledgements

We are grateful to Geir Evensen, Laurent Bertino, Sigurd Tjelmeland, Jonas Andersson, Arnt Ove Hopland, and two anonymous referees for help, comments, and suggestions. We acknowledge financial support from the Norwegian Research Council, project number 234238/E40.

References

- Burgers, Gerrit, Peter Jan van Leeuwen, Geir Evensen. 1998. Analysis scheme in the ensemble Kalman filter. *Monthly Weather Review* 126(6) 1719–1724.
- Durant, J. M., D. Ø. Hjermmann, P. S. Sabarros, N. C. Stenseth. 2008. Northeast Arctic cod population persistence in the Lofoten-Barents Sea system under fishing. *Ecological Applications* 18(3) 662–669.
- Ekerhovd, N.-A., D. V. Gordon. 2013. Catch, stock elasticity and an implicit index of fishing effort. *Marine Resource Economics* 28(4).
- Ekerhovd, N.-A., S. F. Kvamsdal. 2014. Up the Ante on Bioeconomic Submodels of Marine Foodwebs: A Data Assimilation-based Approach. SNF-Centre for Applied Research at NHH, Working Paper 13/14.
- Emerick, A. A., A. C. Reynolds. 2012. History matching time-lapse seismic data using the ensemble Kalman filter with multiple data assimilations. *Computational Geoscience* 16 639–659.
- Evensen, Geir. 2003. The Ensemble Kalman Filter: Theoretical formulation and practical implementation. *Ocean Dynamics* 53 343–367.
- Evensen, Geir. 2009. *Data Assimilation: The ensemble Kalman filter*. 2nd ed. Springer-Verlag, Berlin Heidelberg.

- 1
2
3 Evensen, Geir, Peter Jan van Leeuwen. 2000. An ensemble Kalman smoother for nonlinear
4
5 dynamics. *Monthly Weather Review* 128(6) 1852–1867.
6
- 7 Fulton, E. A., J. S. Link, I. C. Kaplan, M. Savina-Rolland, P. Johnson, C. Ainsworth, P.
8
9 Horne, R. Gorton, R. J. Gamble, A. D. M. Smith, D. C. Smith. 2011. Lessons in
10
11 modelling and management of marine ecosystems: the Atlantis experience. *Fish*
12
13 and *Fisheries* 12(2) 171–188.
14
15
- 16 Gjørseter, Harald, Bjarte Bogstad, Sigurd Tjelmeland. 2009. Ecosystem effects of the three
17
18 capelin stock collapses in the Barents Sea. *Marine Biology Research* 5 40–53.
19
- 20 Grønnevik, Rune, Geir Evensen. 2001. Application of ensemble-based techniques in fish
21
22 stock assessment. *Sarsia* 86 517–526.
23
24
- 25 Hansen, James A., Cécile Penland. 2007. On stochastic parameter estimation using data
26
27 assimilation. *Physica D* 230(1-2) 88–98.
28
29
- 30 Harvey, A.C. 1984. A Unified View of Statistical Forecasting Procedures. *Journal of*
31
32 *Forecasting* 3, 245-275.
33
- 34 Hill, S. L., G. M. Watters, A. E. Punt, M. K. McAllister, C. L. Quéré, J. Turner. 2007.
35
36 Model uncertainty in the ecosystem approach to fisheries. *Fish and Fisheries* 8(4)
37
38 315–336.
39
- 40 Holland, Daniel S., James N. Sanchirico, Robert J. Johnston, Deepak Joglekar. 2010.
41
42 Economic Analysis for Ecosystem-Based Management: Applications to Marine and
43
44 Coastal Environments. RFF Press, Washington. D.C
45
46
- 47 Holling, C. S., Gary K. Meffe. 1996. Command and control and the pathology of natural
48
49 resource management. *Conservation Biology* 10(2) 328–337.
50
51
- 52 Hosking, J. R. M. 1980. The Multivariate Portmanteau Statistic. *Journal of the American*
53
54 *Statistical Association* 75(371) 602-608.
55
56
57
58
59
60

- 1
2
3 Huse, Geir, Geir O. Johansen, Bjarte Bogstad, Harald Gjøsæter. 2004. Studying spatial
4 and thropic interactions between capelin and cod using individual- based
5 modelling. *ICES Journal of Marine Science* 61 1201–1213.
6
7
8
- 9
10 Kivman, G. A. 2003. Sequential parameter estimation for stochastic systems. *Nonlinear*
11 *Processes in Geophysics* 10 253–259
12
13
- 14 Kugarajh, Kanaganayagam, Leif K. Sandal, Gerhard Berge. 2006. Implementing a
15 stochastic bioeconomic model for the North-East Arctic cod fishery. *Journal of*
16 *Bioeconomics* 8(1) 35–53.
17
18
19
- 20 Marshall, C. Tara, Nathalia A. Yaragina, Bjørn Adlandsvik, Andrey V. Dolgov. 2000.
21 Reconstructing the stock-recruit relationship for Northeast Arctic cod using a
22 bioenergetic index of reproductive potential. *Canadian Journal of Fisheries and*
23 *Aquatic Sciences* 57(12) 2433–2442.
24
25
26
27
28
- 29 May, Robert M., John R. Beddington, Colin W. Clark, Sidney J. Holt, Richard M. Laws.
30 1979. Management of multispecies fisheries. *Science* 205 267–277.
31
32
33
- 34 Olsen, E. H. Gjøsæter, I. Røttingen, A. Dommasnes, P. Fossum, P. Sandberg. 2007. The
35 Norwegian ecosystem-based management plan for the Barents Sea. *ICES Journal of*
36 *Marine Science* 64(4) 599-602.
37
38
39
- 40 Pew Oceans Commission. 2003. America’s living oceans-charting a course for sea change.
41 recommendations for a new ocean policy. Tech. rep., Pew Foundation,
42 Washington, D.C.
43
44
45
46
- 47 Poudel, D., L. K. Sandal, S. I. Steinshamn, S. F. Kvamsdal. 2012. Do species interaction
48 and stochasticity matter to optimal management of multispecies fisheries? G. H.
49 Kruse, H. I. Browman, K. L. Cochrane, D. Evans, G. S. Jamieson, P. A. Livingston,
50 D. Woodby, C. I. Zhang, eds., *Global Progress in Ecosystem-Based Fisheries*
51 *Management*. Alaska Sea Grant, University of Alaska Fairbanks, 209–236.
52
53
54
55
56
57
58
59
60

- 1
2
3 Røttingen, Ingolf, Sigurd Tjelmeland. 2008. A quest for management objectives - case
4 study on the Barents Sea Capelin. ICES 2008 Annual Science Conference, Halifax,
5 Nova Scotia, Canada . O:08, ICES CM Documents 2008.
6
7
8
9 Sandal, Leif K., Stein I. Steinshamn. 1997. A stochastic feedback model for optimal
10 management of renewable resources. *Natural Resource Modeling* 10(1) 31–52.
11
12 Sandal, Leif K., Stein I. Steinshamn. 2010. Rescuing the prey by harvesting the
13 predator: Is it possible? Endre Bjørndal, Mette Bjørndal, Panos M. Pardalos,
14 Mikael Ronnqvist, eds., *Energy, Natural Resources and Environmental Economics .*
15 *Energy Systems*, Springer-Verlag Berlin Heidelberg, 359–378.
16
17
18 Skern-Mauritzen, Mette, Geir Ottersen, Nils O. Hagedard, Geir Huse, Gjert E. Dingsør, Nils
19 C. Stenseth, Olav S. Kjesbu. 2015. Ecosystem processes are rarely included in tactical
20 fisheries management. *Fish and Fisheries*, forthcoming. DOI: 10.1111/faf.12111
21
22
23 United Nations, The World Summit on Sustainable Development. 2002. Plan of
24 Implementation of the World Summit on Sustainable Development:
25 Johannesburg Declaration on Sustainable Development. UN Documents
26 A/Conf.199/20, United Nations.
27
28
29 Ussif, Al-Amin M., Leif K. Sandal, Stein I. Steinshamn. 2003. A new approach of fitting
30 biomass dynamics models to data. *Mathematical Biosciences* 182 67–79.
31
32
33 Vert-pre, Katyana A., Ricardo O. Amoroso, Olaf P. Jensen, Ray Hilborn. 2013. Frequency
34 and intensity of productivity regime shifts in marine fish stocks. *Proceedings of the*
35 *National Academy of Sciences* 110(5) 1779-1784.
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

Table 1: Functional forms used in the model equations.

Term	Functional Form
Logistic Growth	$f_1(x_i, c_p, c_q) = c_p x_i \left(1 - x_i/c_q\right)$
Modified Logistic Growth	$f_2(x_i, c_p, c_q) = c_p x_i^2 \left(1 - x_i/c_q\right)$
Modified Logistic Growth with Common Capacity	$f_3(x_i, x_j, c_p, c_q) = c_p x_i^2 \left(1 - x_i + x_j/c_q\right)$
Lotka-Volterra Interaction	$g_1(x_i, x_j, c_p) = c_p x_i x_j$
Modified Lotka-Volterra Interaction	$g_2(x_i, x_j, x_k, c_p) = c_p x_i x_k^{x_k/x_j + x_k}$

Table 2: Parameter estimates with standard errors in parentheses for the single equation assimilations (the horizontal lines separate the different assimilations). The table also reports contraction rates and the c -interval for each parameter, noise-scale estimates ($\sigma_{i,i}$), and BIC-scores, the average root mean squared innovation, R_D^2 , the Lilliefors k -statistic, and the Ljung-Box q -statistic (with 15 lags) for each equation (p -values in parentheses).

	Estimate	Contraction	c -interval
Cod, equation (12), BIC: 91.63, Avg. RMSI: 0.1518, R_D^2 : 0.2724, k : 0.1026 (0.1366), q : 22.60 (0.0928)			
α_1	-0.5694 (0.1159)	0.114	(0.5039, 0.6354)
α_2	15.64 (0.2672)	0.544	(4.780 e6, 8.158 e6)
α_3	-23.12 (0.4886)	0.986	(0.4540 e-10, 1.472 e-10)
α_4	-26.46 (0.4917)	0.982	(1.959 e-12, 5.240 e-12)
$\sigma_{1,1}$	0.1248 (0.0115)		
Capelin, equation (13), BIC: 177.63, Avg. RMSI: 0.2300, R_D^2 : 0.9938, k : 0.1182 (0.0450), q : 10.18 (0.8079)			
α_5	-12.58 (0.3831)	0.3780	(2.335 e-6, 5.024 e-6)
α_6	16.45 (0.2094)	0.710	(11.35 e6, 17.26 e6)
α_7	-21.74 (0.4845)	0.978	(2.206 e-10, 5.815 e-10)
α_8	-24.10 (0.485)	0.970	(2.095 e-11, 5.533 e-11)
$\sigma_{2,2}$	0.2289 (0.0430)		
Herring, equation (14), BIC: 89.13, Avg. RMSI: 0.1956, R_D^2 : 0.9972, k : 0.0929 (0.2488), q : 11.07 (0.7472)			
α_9	-11.58 (0.6948)	0.693	(0.4633 e-5, 1.859 e-5)
α_{10}	15.70 (0.3484)	0.690	(4.647 e6, 9.330 e6)
α_{11}	-26.36 (0.4757)	0.976	(2.206 e-12, 5.714 e-12)
α_{12}	-24.24 (0.9523)	0.978	(1.136 e-11, 7.631 e-11)
α_{13}	-0.1432 (0.4597)	0.929	(0.5471, 1.372)
$\sigma_{3,3}$	0.1953 (0.0092)		

Table 3: Parameter estimates with standard errors in parentheses for the full model assimilation. The table also reports contraction rates and the c -interval for each parameter, and the average root mean squared innovation for each state variable.

	Estimate	Contraction	c -interval
Cod, equation (12), Avg. RMSI: 0.09445			
α_1	-0.5829 (0.0735)	0.623	(0.5186, 0.6008)
α_2	15.69 (0.1922)	0.704	(5.403 e6, 7.937 e6)
α_3	-23.16 (0.4337)	0.914	(0.5660 e-10, 1.347 e-10)
α_4	-26.47 (0.4508)	0.929	(2.033 e-12, 5.009 e-12)
Capelin, equation (13), Avg. RMSI: 0.1202			
α_5	-12.81 (0.0643)	0.169	(2.548 e-6, 2.899 e-6)
α_6	16.56 (0.0788)	0.372	(1.445 e7, 1.692 e7)
α_7	-21.86 (0.4490)	0.931	(2.034 e-10, 4.994 e-10)
α_8	-24.31 (0.4415)	0.922	(1.779 e-11, 4.302 e-11)
Herring, equation (14), Avg. RMSI 0.1013			
α_9	-11.53 (0.1844)	0.270	(0.8099 e-5, 1.171 e-5)
α_{10}	15.18 (0.1634)	0.490	(3.327 e6, 4.614 e6)
α_{11}	-26.13 (0.4463)	0.937	(2.847 e-12, 6.951 e-12)
α_{12}	-24.36 (0.9020)	0.934	(1.065 e-11, 6.472 e-11)
α_{13}	0.1168 (0.3817)	0.879	(0.7673, 1.646)

Table 4: Parameter estimates with standard errors in parentheses for regressions of the individual equations in the alternative model. The table also reports t -statistics, the root mean squared errors R^2 , the Lilliefors k -statistic, and the Ljung-Box q -statistic (with 15 lags) for each equation (p -values in parentheses).

	Estimate	t -stat
Cod, equation (16), RMSE: 2.52 e5, R^2 : 0.258, k : 0.1583 (0.0013), q : 16.08 (0.3762)		
γ_1	1.147 e5 (2.02 e5)	0.56
β_1	0.357 (0.211)	1.69
β_2	-7.29 e-8 (4.57 e-8)	-1.59
β_3	1.49 e-8 (7.64 e-9)	1.95
β_4	1.88 e-8 (1.23 e-8)	1.52
Capelin, equation (17), RMSE: 1.43 e6, R^2 : 0.331, k : 0.0668 (>0.500), q : 13.96 (0.5278)		
γ_2	6.63 e5 (3.37 e5)	1.96
β_5	1.11 e-7 (5.96 e-8)	1.85
β_6	-1.44 e-5 (6.76 e-6)	-2.14
β_7	2.03 e-8 (6.21 e-8)	0.326
β_8	-1.53 e-7 (3.29 e-8)	-4.66
Herring, equation (18), RMSE: 1.17 e6, R^2 : 0.252, k : 0.1968 (<0.001), q : 21.76 (0.1141)		
γ_3	-1.01 e5 (3.15 e5)	-0.31
β_9	-3.07 e-7 (2.95 e-7)	-1.04
β_{10}	2.71 e-5 (2.98 e-5)	0.90
β_{11}	3.88 e-8 (3.05 e-8)	1.27
β_{12}	-2.13 e-7 (1.60 e-7)	-1.33
β_{13}	0.83 (0.65)	1.28

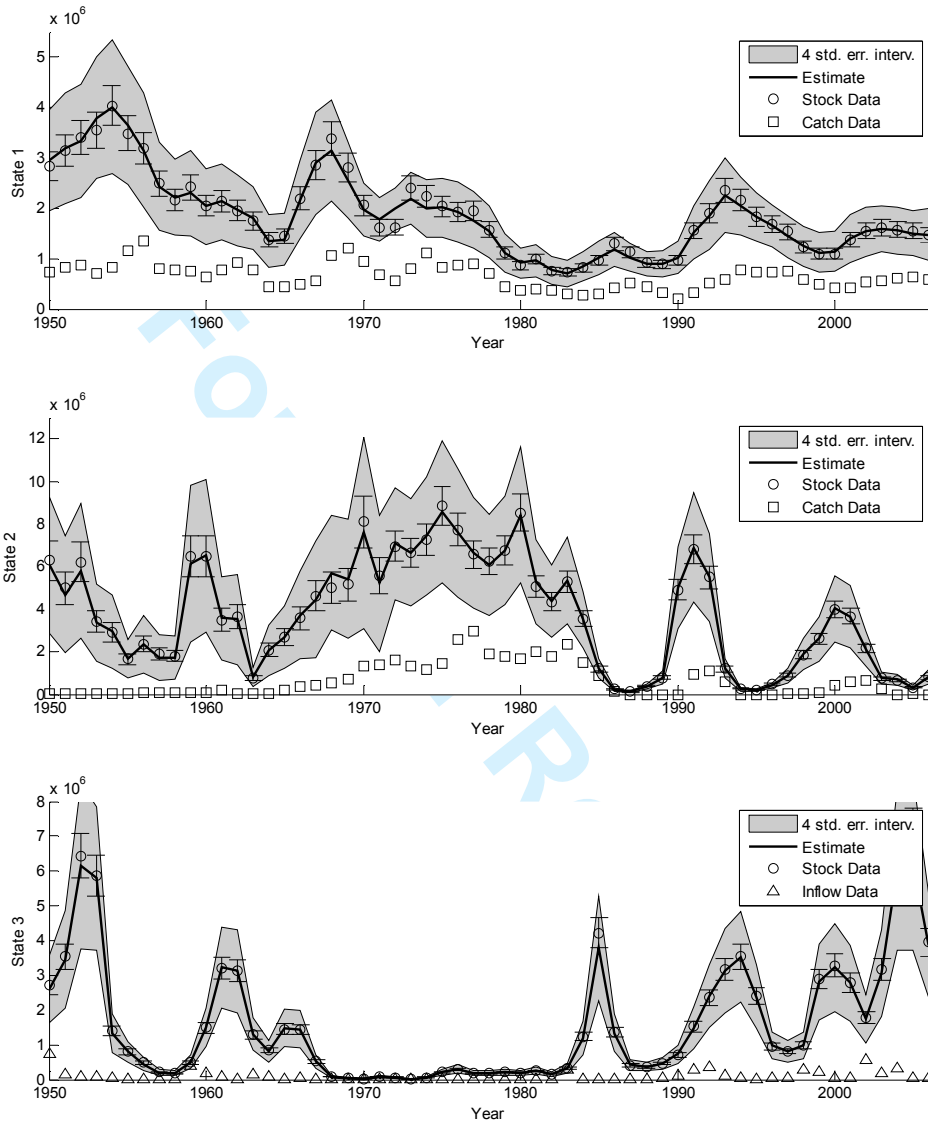


Figure 1: Smoothed stock level estimates (solid curves) with two standard errors to each side (shaded areas). Stock level observations with observation uncertainty (circles and error bars) and harvest (squares) and inflow (triangles) levels. Top panel: Cod. Middle panel: Capelin. Bottom panel: Herring.