The Ensemble Kalman Filter:
Theoretical formulation and practical implementation

Geir Evensen

Norsk Hydro Research Centre, Bergen, Norway

Based on Evensen 2003, Ocean Dynamics, Vol 53, No 3

The Ensemble Kalman Filter (EnKF)

- Represents error statistics using an ensemble of model states
- Evolves error statistics by ensemble integrations
- "Variance minimizing" analysis scheme operating on the ensemble

$$\Downarrow$$

- Monte Carlo, low rank, error subspace method
- Converges to the Kalman Filter with increasing ensemble size
- Fully nonlinear error evolution, contrary to EKF
- Assumption of Gaussian statistics in analysis scheme

The error covariance matrix

Define ensemble covariances around the ensemble mean

$$\boldsymbol{P}^{\mathrm{f}} \simeq \boldsymbol{P}_{\mathrm{e}}^{\mathrm{f}} = \overline{(\boldsymbol{\psi}^{\mathrm{f}} - \overline{\boldsymbol{\psi}^{\mathrm{f}}})(\boldsymbol{\psi}^{\mathrm{f}} - \overline{\boldsymbol{\psi}^{\mathrm{f}}})^{\mathrm{T}}}$$

$$\boldsymbol{P}^{\mathrm{a}} \simeq \boldsymbol{P}_{\mathrm{e}}^{\mathrm{a}} = \overline{(\boldsymbol{\psi}^{\mathrm{a}} - \overline{\boldsymbol{\psi}^{\mathrm{a}}})(\boldsymbol{\psi}^{\mathrm{a}} - \overline{\boldsymbol{\psi}^{\mathrm{a}}})^{\mathrm{T}}}$$

- The ensemble mean $\overline{\psi}$ is the best-guess.
- The ensemble spread defines the error variance.
- The covariance is determined by the smoothness of the ensemble members.
- A covariance matrix can be represented by an ensemble of model states (not unique).

Dynamical evolution of error statistics

- Each ensemble member evolve according to the model dynamics which is expressed by a stochastic differential equation

$$d\boldsymbol{\psi} = \boldsymbol{f}(\boldsymbol{\psi})dt + \boldsymbol{g}(\boldsymbol{\psi})d\boldsymbol{q}.$$

- The probability density then evolve according to Kolmogorov's equation

$$\frac{\partial \phi}{\partial t} + \sum_i \frac{\partial (f_i \phi)}{\partial \psi_i} = \frac{1}{2} \sum_{i,j} \frac{\partial^2 \phi (\boldsymbol{g}\boldsymbol{Q}\boldsymbol{g}^T)_{ij}}{\partial \psi_i \partial \psi_j},$$

- This is the fundamental equation for evolution of error statistics and can be solved using Monte Carlo methods.

Analysis scheme (1)

- Given an ensemble of model forcasts ψ_j^{f}.
- Create an ensemble of observations

$$\boldsymbol{d}_j = \boldsymbol{d} + \boldsymbol{\epsilon}_j,$$

  with
  - $\boldsymbol{d}$ the first guess observations,
  - $\boldsymbol{\epsilon}_j$ a vector of observation noise.
- The measurement error covariance matrix is

$$\boldsymbol{R} \simeq \boldsymbol{R}_e = \overline{\boldsymbol{\epsilon}\boldsymbol{\epsilon}^{\mathrm{T}}}.$$

Analysis scheme (2)

- Update each ensemble member according to

$$\boldsymbol{\psi}_j^{\mathrm{a}} = \boldsymbol{\psi}_j^{\mathrm{f}} + K_e(\boldsymbol{d}_j - H\boldsymbol{\psi}_j^{\mathrm{f}}).$$

where

$$K_e = P_e^f H^{\mathrm{T}}(HP_e^f H^{\mathrm{T}} + R_e)^{-1}$$

- This is equivalent to updating the mean

$$\overline{\boldsymbol{\psi}^{\mathrm{a}}} = \overline{\boldsymbol{\psi}^{\mathrm{f}}} + K_e(\boldsymbol{d} - H\overline{\boldsymbol{\psi}^{\mathrm{f}}}).$$

- The posterior error covariance becomes

$$\boldsymbol{P}_{\mathrm{e}}^{\mathrm{a}} = (\boldsymbol{I} - \boldsymbol{K}_{\mathrm{e}}\boldsymbol{H})\boldsymbol{P}_{\mathrm{e}}^{\mathrm{f}}.$$

Analysis of the Analysis scheme (1)

- Define the ensemble matrix

$$\boldsymbol{A} = (\boldsymbol{\psi}_1, \boldsymbol{\psi}_2, \dots, \boldsymbol{\psi}_N) \in \Re^{n \times N}.$$

- The ensemble mean is (defining $\mathbf{1}_N \in \Re^{N \times N} \equiv 1/N$)

$$\overline{\boldsymbol{A}} = \boldsymbol{A}\mathbf{1}_N.$$

- The ensemble perturbations becomes

$$\boldsymbol{A}' = \boldsymbol{A} - \overline{\boldsymbol{A}} = \boldsymbol{A}(\boldsymbol{I} - \mathbf{1}_N).$$

- The ensemble covariance matrix $\boldsymbol{P}_\mathrm{e} \in \Re^{n \times n}$ becomes

$$\boldsymbol{P}_\mathrm{e} = \frac{\boldsymbol{A}'(\boldsymbol{A}')^\mathrm{T}}{N-1}$$

Analysis equation (2)

- Given a vector of measurements $\boldsymbol{d} \in \Re^m$, define

$$\boldsymbol{d}_j = \boldsymbol{d} + \boldsymbol{\epsilon}_j, \quad j = 1, \ldots, N,$$

stored in

$$\boldsymbol{D} = (\boldsymbol{d}_1, \boldsymbol{d}_2, \ldots, \boldsymbol{d}_N) \in \Re^{m \times N},$$

- The ensemble perturbations are stored in

$$\boldsymbol{\Upsilon} = (\boldsymbol{\epsilon}_1, \boldsymbol{\epsilon}_2, \ldots, \boldsymbol{\epsilon}_N) \in \Re^{m \times N},$$

thus, the measurement error covariance matrix becomes

$$\boldsymbol{R}_{\mathrm{e}} = \frac{\boldsymbol{\Upsilon}\boldsymbol{\Upsilon}^{\mathrm{T}}}{N-1}.$$

Analysis equation (3)

- The analysis equation can now be written

$$\boldsymbol{A}^{\mathrm{a}} = \boldsymbol{A} + \boldsymbol{P}_{\mathrm{e}}\boldsymbol{H}^{\mathrm{T}}(\boldsymbol{H}\boldsymbol{P}_{\mathrm{e}}\boldsymbol{H}^{\mathrm{T}} + \boldsymbol{R}_{\mathrm{e}})^{-1}(\boldsymbol{D} - \boldsymbol{H}\boldsymbol{A}).$$

- Defining the innovations $\boldsymbol{D}' = \boldsymbol{D} - \boldsymbol{H}\boldsymbol{A}$ and using previous definitions:

$$\boldsymbol{A}^{\mathrm{a}} = \boldsymbol{A} + \boldsymbol{A}'\boldsymbol{A}'^{\mathrm{T}}\boldsymbol{H}^{\mathrm{T}}\left(\boldsymbol{H}\boldsymbol{A}'\boldsymbol{A}'^{\mathrm{T}}\boldsymbol{H}^{\mathrm{T}} + \boldsymbol{\Upsilon}\boldsymbol{\Upsilon}^{\mathrm{T}}\right)^{-1}\boldsymbol{D}'.$$

i.e., analysis expressed entirely in terms of the ensemble

Practical computation of analysis (1)

- Use pseudo inverse

$$\boldsymbol{H}\boldsymbol{A}'\boldsymbol{A}'^{\mathrm{T}}\boldsymbol{H}^{\mathrm{T}} + \boldsymbol{\Upsilon}\boldsymbol{\Upsilon}^{\mathrm{T}} = \boldsymbol{Z}\boldsymbol{\Lambda}\boldsymbol{Z}^{\mathrm{T}}$$

$$(\boldsymbol{H}\boldsymbol{A}'\boldsymbol{A}'^{\mathrm{T}}\boldsymbol{H}^{\mathrm{T}} + \boldsymbol{\Upsilon}\boldsymbol{\Upsilon}^{\mathrm{T}})^{-1} = \boldsymbol{Z}\boldsymbol{\Lambda}^{-1}\boldsymbol{Z}^{\mathrm{T}}.$$

- Computational cost is:
  - m^2N to form $\boldsymbol{H}\boldsymbol{A}'\boldsymbol{A}'^{\mathrm{T}}\boldsymbol{H}^{\mathrm{T}}$,
  - $\mathcal{O}(m^2)$ for eigenvalue decomposition.
- Unafordable for large m!

Practical computation of analysis (2)

- Note that $\boldsymbol{H}\boldsymbol{A}'\boldsymbol{\Upsilon}^{\mathrm{T}} \equiv \mathbf{0}$, thus

$$\boldsymbol{H}\boldsymbol{A}'\boldsymbol{A}'^{\mathrm{T}}\boldsymbol{H}^{\mathrm{T}} + \boldsymbol{\Upsilon}\boldsymbol{\Upsilon}^{\mathrm{T}} = (\boldsymbol{H}\boldsymbol{A}' + \boldsymbol{\Upsilon})(\boldsymbol{H}\boldsymbol{A}' + \boldsymbol{\Upsilon})^{\mathrm{T}}.$$

- Compute SVD, $\boldsymbol{H}\boldsymbol{A}' + \boldsymbol{\Upsilon} = \boldsymbol{U}\boldsymbol{\Sigma}\boldsymbol{V}^{\mathrm{T}}$, giving

$$\boldsymbol{H}\boldsymbol{A}'\boldsymbol{A}'^{\mathrm{T}}\boldsymbol{H}^{\mathrm{T}} + \boldsymbol{\Upsilon}\boldsymbol{\Upsilon}^{\mathrm{T}} = \boldsymbol{U}\boldsymbol{\Sigma}\boldsymbol{V}^{\mathrm{T}}\boldsymbol{V}\boldsymbol{\Sigma}^{\mathrm{T}}\boldsymbol{U}^{\mathrm{T}} = \boldsymbol{U}\boldsymbol{\Sigma}\boldsymbol{\Sigma}^{\mathrm{T}}\boldsymbol{U}^{\mathrm{T}}.$$

- Computational cost is $\mathcal{O}(mN)$ for SVD.

Practical computation of analysis (3)

- The analysis equation can now be written

$$\boldsymbol{A}^{\mathrm{a}} = \boldsymbol{A} + \boldsymbol{A}'(\boldsymbol{H}\boldsymbol{A}')^{\mathrm{T}}\boldsymbol{U}\boldsymbol{\Lambda}^{-1}\boldsymbol{U}^{\mathrm{T}}\boldsymbol{D}'.$$

- The computation goes as follows with $p \leq N$

$$\begin{aligned}
\boldsymbol{X}_1 &= \boldsymbol{\Lambda}^{-1}\boldsymbol{U}^{\mathrm{T}} && \in \Re^{N\times m} && mp,\\
\boldsymbol{X}_2 &= \boldsymbol{X}_1\boldsymbol{D}' && \in \Re^{N\times N} && mNp,\\
\boldsymbol{X}_3 &= \boldsymbol{U}\boldsymbol{X}_2 && \in \Re^{m\times N} && mNp,\\
\boldsymbol{X}_4 &= (\boldsymbol{H}\boldsymbol{A}')^{\mathrm{T}}\boldsymbol{X}_3 && \in \Re^{N\times N} && mNN,\\
\boldsymbol{A}^{\mathrm{a}} &= \boldsymbol{A} + \boldsymbol{A}'\boldsymbol{X}_4 && \in \Re^{n\times N} && nNN,
\end{aligned}$$

- All m^2N computations reduced to mN^2.

Practical computation of analysis (4)

- The final update can be written as

$$\begin{aligned} \boldsymbol{A}^{\mathrm{a}} &= \boldsymbol{A} + (\boldsymbol{A} - \overline{\boldsymbol{A}})\boldsymbol{X}_4 \\ &= \boldsymbol{A} + \boldsymbol{A}(\boldsymbol{I} - \mathbf{1}_N)\boldsymbol{X}_4 \\ &= \boldsymbol{A}(\boldsymbol{I} + \boldsymbol{X}_4) \\ &= \boldsymbol{A}\boldsymbol{X}_5, \end{aligned}$$

  thus, the analysis is a “weakly nonlinear combination” of the forecast ensemble.

- Note also

$$\begin{aligned} \boldsymbol{A}^{\mathrm{a}} &= \boldsymbol{A} + \boldsymbol{P}_{\mathrm{e}}\boldsymbol{H}^{\mathrm{T}}(N-1)\boldsymbol{X}_3 \\ &\equiv \boldsymbol{A}\boldsymbol{X}_5 \end{aligned}$$

Remarks on the analysis equation (1)

- Covariances only needed between observed variables at measurement locations.
- Covariances never computed but indirectly used to determine $\boldsymbol{HPH}^{\mathrm{T}}$.
- Analysis may be interpreted as:
  - combination of ensemble members, or,
  - forecast pluss combination of covariance functions.
- Covariances only needed to compute $\boldsymbol{X}_5$.
- Accuracy of analysis is determined by
  - the accuracy of $\boldsymbol{X}_5$
  - the properties of the ensemble error space

Remarks on the analysis equation (2)

- For a linear model, any choice of $\boldsymbol{X}_5$ will result in an analysis which is also a solution of the model.
- Filtering of covariance functions introduces nondynamical modes in the analysis.

Local analysis

- A local analysis is computed grid point by grid point using only nearby measurements.
- Introduces nondynamical modes in the analysis.
- Different $\boldsymbol{X}_5$ for each grid point.
- Allows us to reach a larger class of solutions.

Nonlinear measurements

- Measurement equation

$$\boldsymbol{d} = \boldsymbol{h}(\boldsymbol{\psi}) + \boldsymbol{\epsilon}.$$

- Define ensemble of model prediction of the measurements

$$\widehat{\boldsymbol{A}} = (\boldsymbol{h}(\boldsymbol{\psi}_1), \ldots, \boldsymbol{h}(\boldsymbol{\psi}_N)), \in \Re^{\hat{m}\times N}.$$

- The analysis then becomes

$$\boldsymbol{A}^{\mathrm{a}} = \boldsymbol{A} + \boldsymbol{A}'\widehat{\boldsymbol{A}}'^{\mathrm{T}} \left(\widehat{\boldsymbol{A}}'\widehat{\boldsymbol{A}}'^{\mathrm{T}} + \boldsymbol{\Upsilon}\boldsymbol{\Upsilon}^{\mathrm{T}}\right)^{-1} (\boldsymbol{D} - \widehat{\boldsymbol{A}}),$$

- Analysis based on covariances between $\boldsymbol{h}(\boldsymbol{\psi})$ and $\boldsymbol{\psi}$.

Ensemble Kalman Smoother (EnKS)

- Starts with EnKF solution.
- Computes updates backward in time;
  - sequentially for each measurement time,
  - using covariances in time,
  - no backward integrations.
- The analysis becomes for $t_{i-1} \leq t' < t_i \leq t_k$:

$$\boldsymbol{A}^{\mathrm{a}}_{\mathrm{EnKS}}(t') = \boldsymbol{A}_{\mathrm{EnKF}}(t') \prod_{j=i}^{k} \boldsymbol{X}_5(t_j)$$

Some recent applications of the EnKF

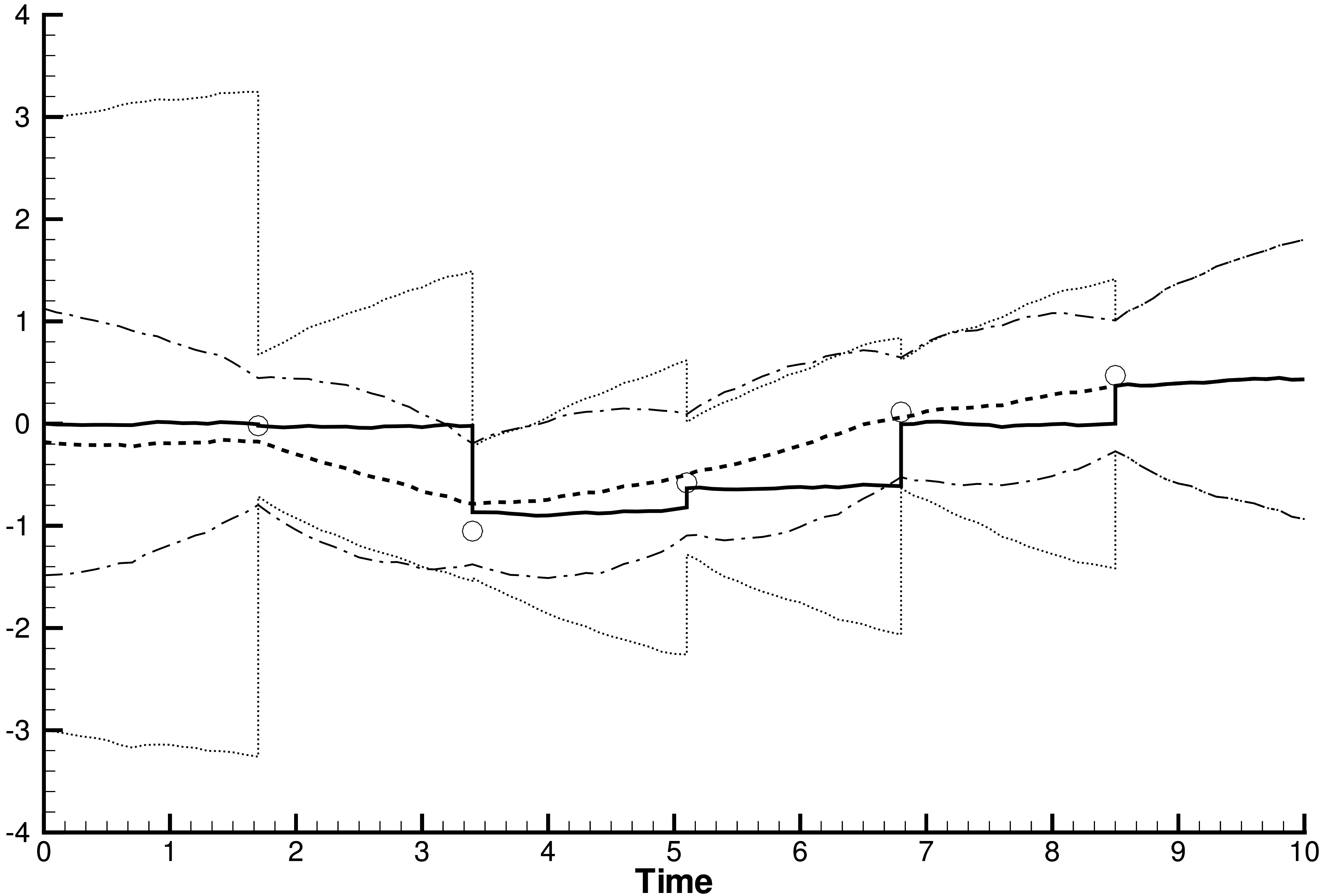
- Haugan and Evensen (2002), Ocean Dynamics.
- Mitchell et al. (2002), MWR.
- Brusdal et al (2003), JMS.
- Natvik and Evensen (2003a,b), JMS.
- Keppenne and Rienecker (2003), JMS.
- DIADEM project
- TOPAZ project (topaz.nersc.no)
- MERSEA project

Time correlated model noise

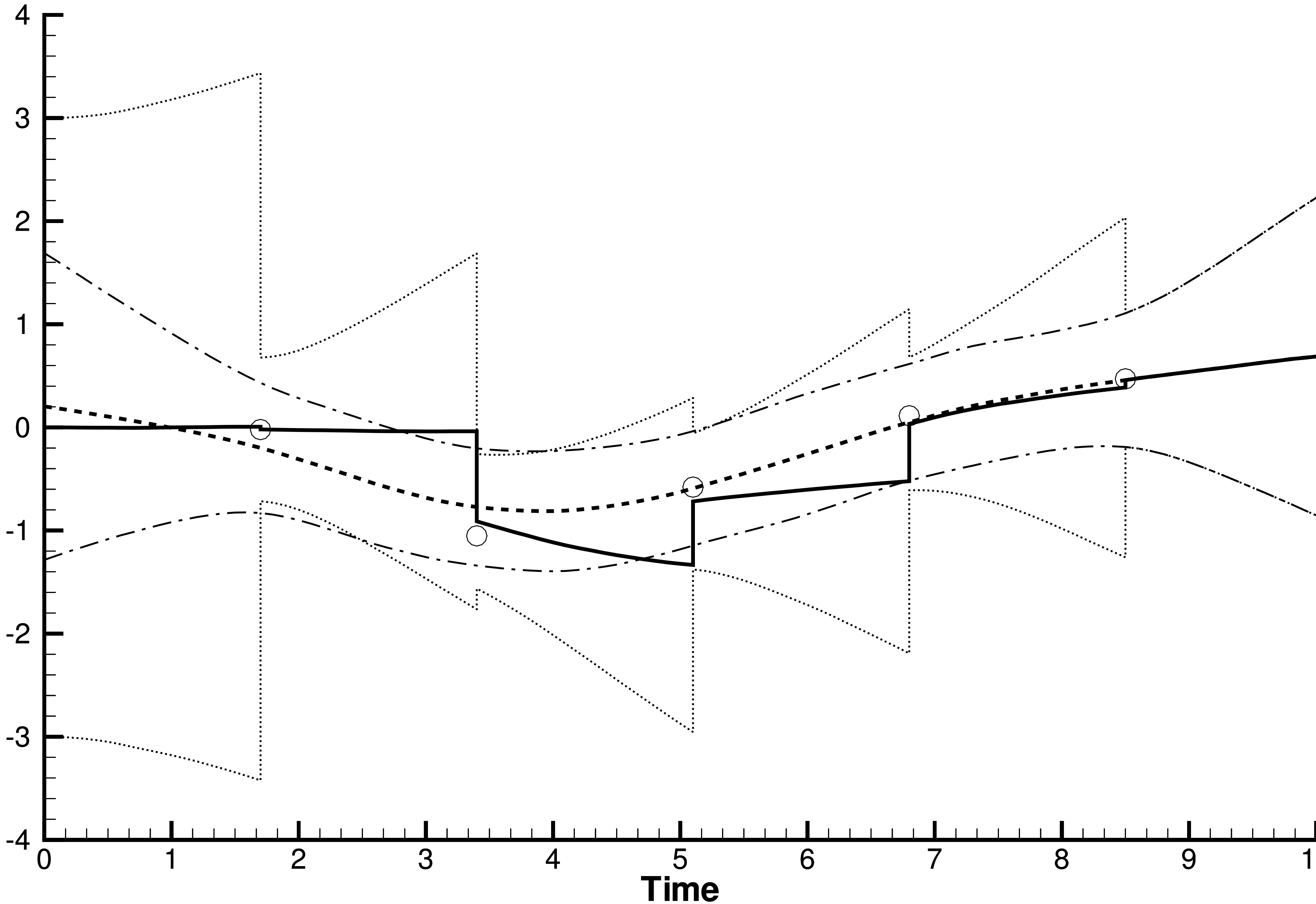
- Scalar model

$$\begin{pmatrix} q_k \\ \psi_k \end{pmatrix} = \begin{pmatrix} \alpha q_{k-1} \\ \psi_{k-1} + \sqrt{\Delta t}\sigma\rho q_k \end{pmatrix} + \begin{pmatrix} \sqrt{1-\alpha^2} w_{k-1} \\ 0 \end{pmatrix}.$$

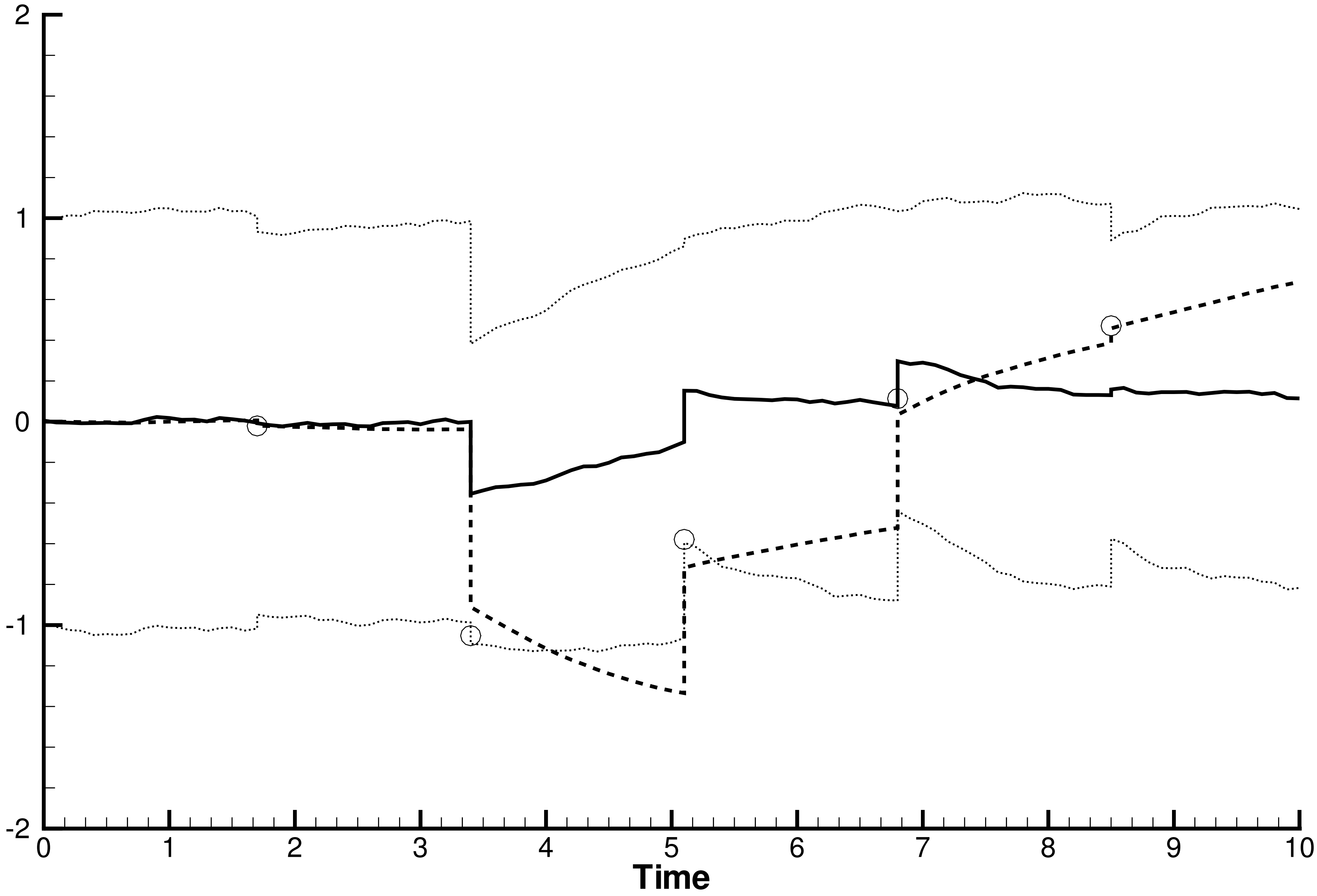
Results ($\alpha = 0$)



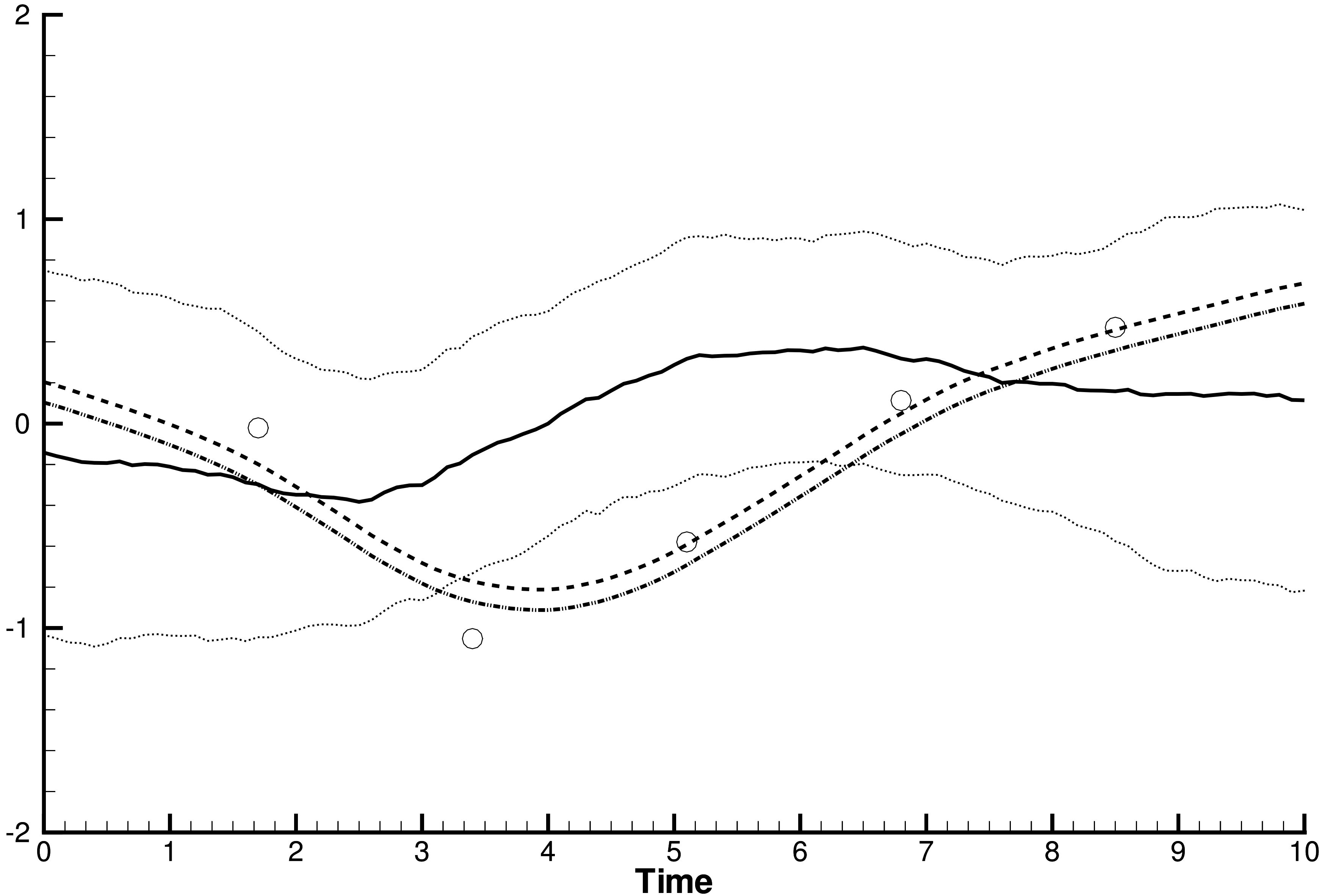
Results ($\alpha = 0.95$)



Estimate of model noise, EnKF



Estimate of model noise, EnKS



Model forced by estimated model error

$$\psi_k = \psi_{k-1} + \sqrt{\Delta t}\sigma\rho\hat{q}_k$$

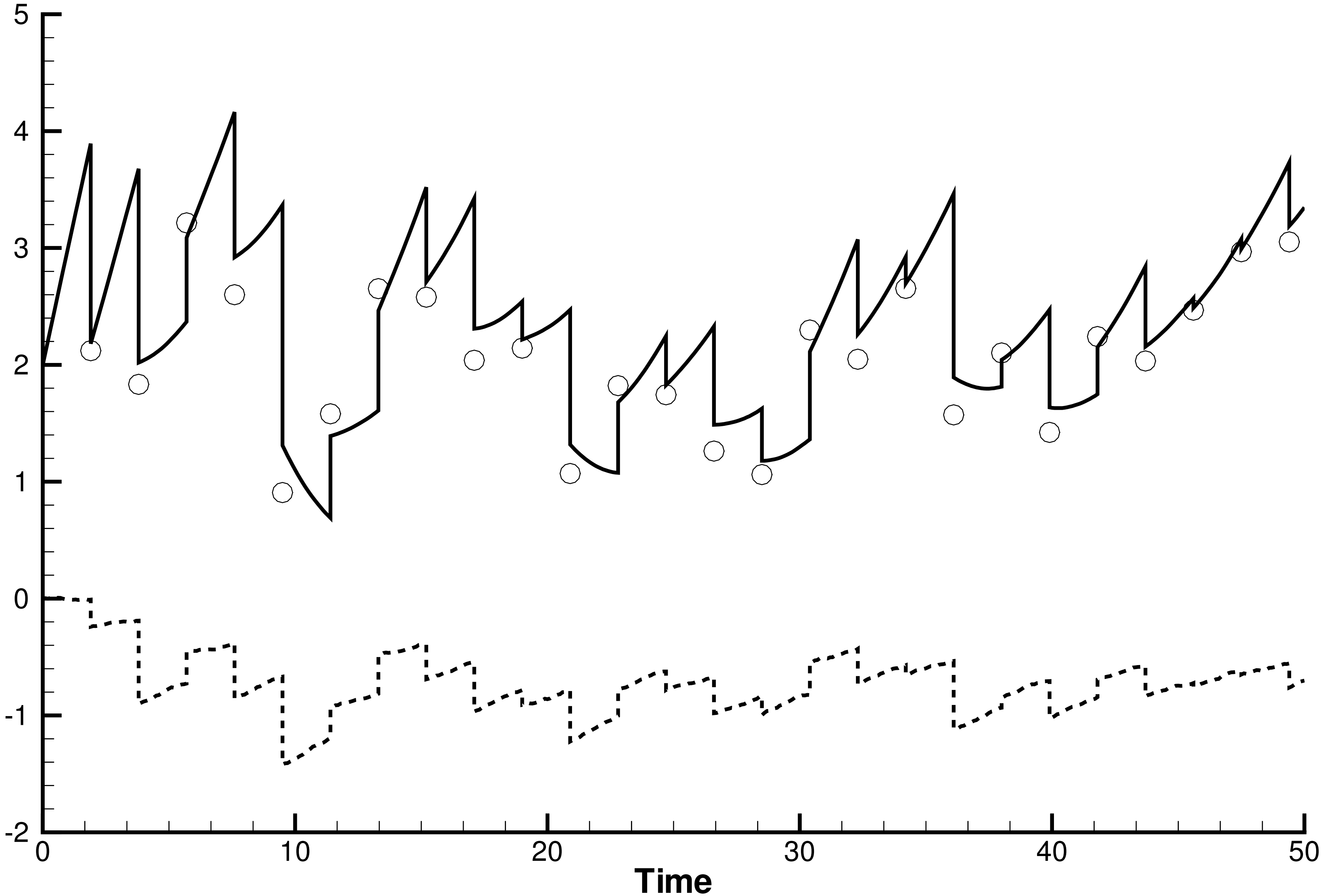
$$\psi_0 = \hat{\psi}_0$$

Parameter and bias estimation

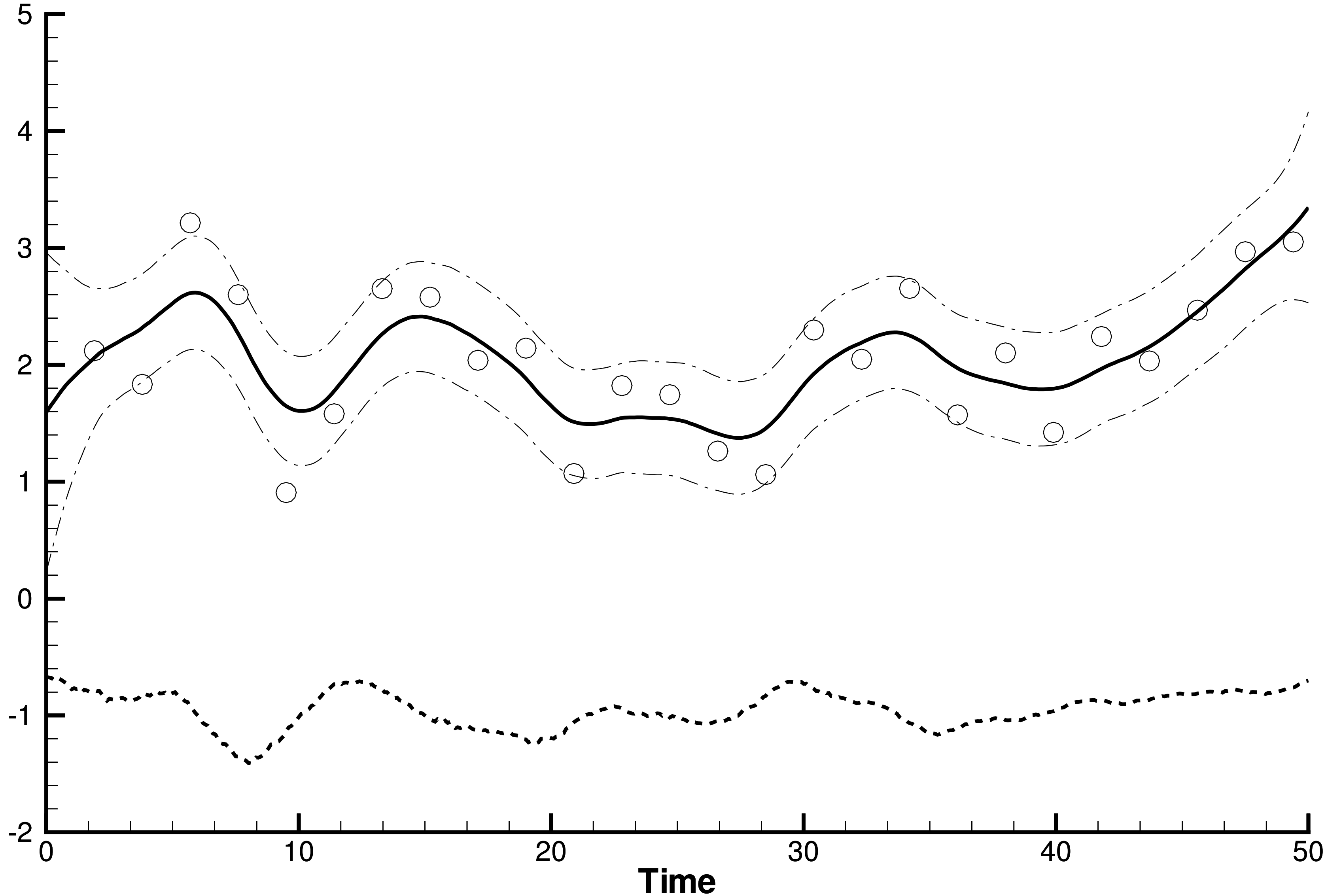
- Introduces poorly known parameter β_k in model

$$\begin{pmatrix} q_k \\ \beta_k \\ \psi_k \end{pmatrix} = \begin{pmatrix} \alpha q_{k-1} \\ \beta_{k-1} \\ \psi_{k-1} + (\eta + \beta_k)\Delta t + \sqrt{\Delta t}\sigma \rho q_k \end{pmatrix} + \begin{pmatrix} \sqrt{1-\alpha^2} w_{k-1} \\ 0 \\ 0 \end{pmatrix}.$$

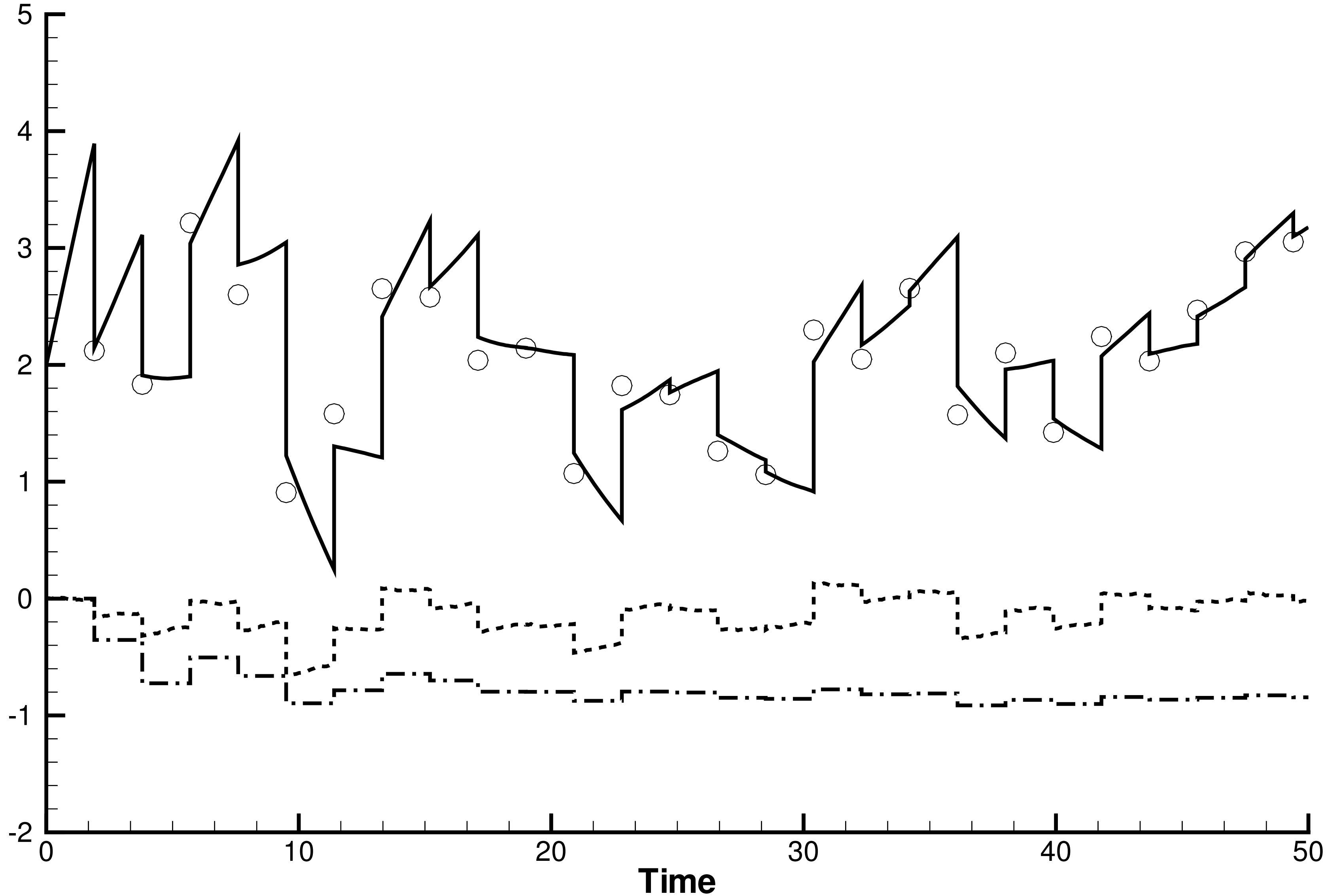
Estimate and model error, EnKF



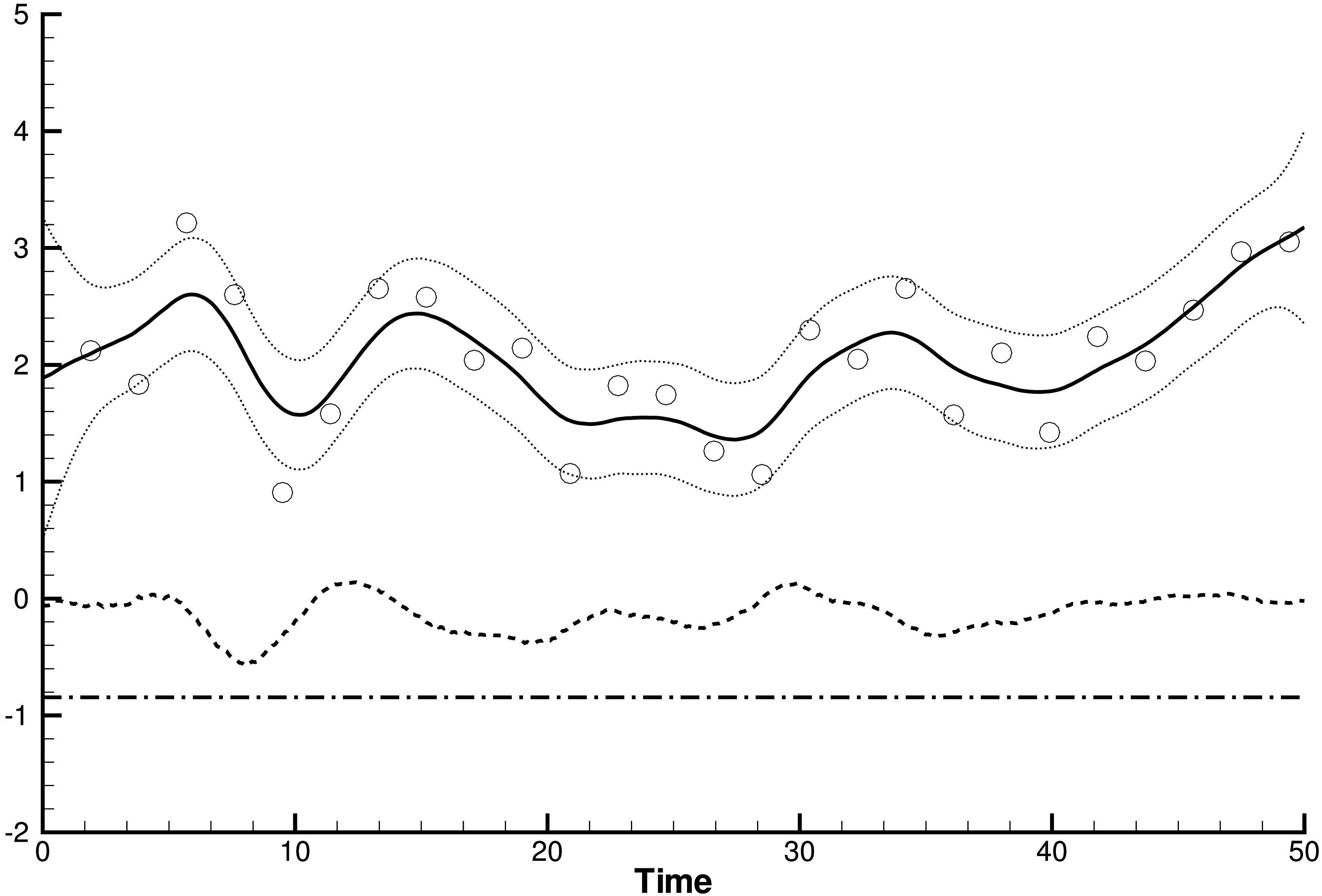
Estimate and model error, EnKS



Estimate, model error and bias, EnKF



Estimate, model error and bias, EnKS



Estimated bias and std dev