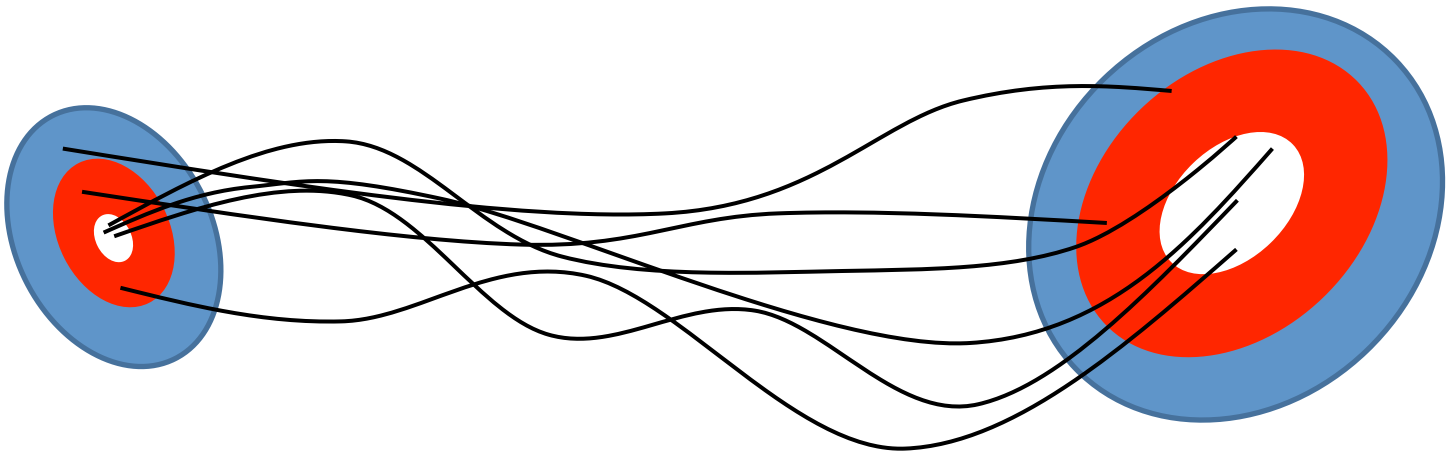
The Ensemble Kalman filter



Part I: Theory

Alison Fowler and Sanita Vetra-Carvalho

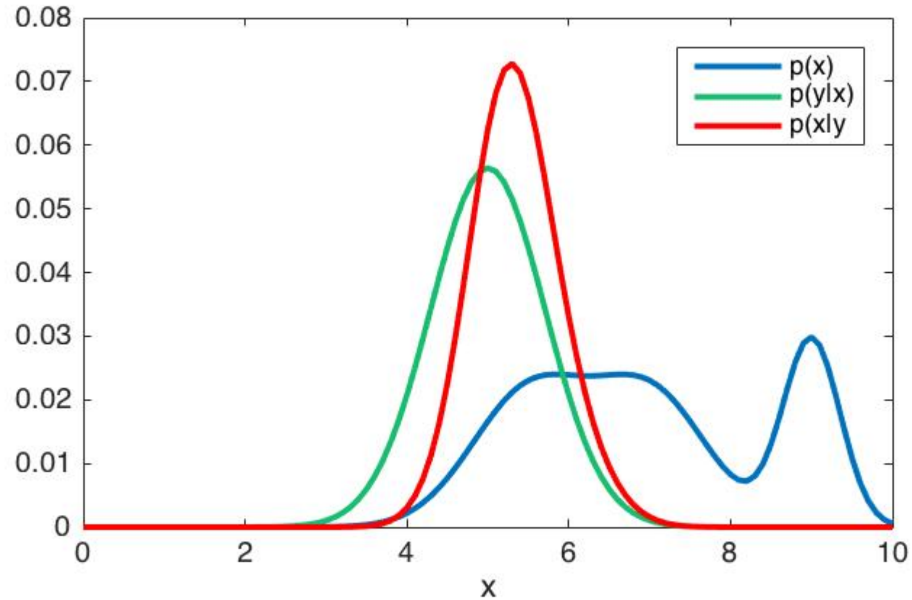
Data-assimilation training course. 5^{th} - 8^{th} March 2019, University of Reading

Recap of problem we wish to solve

- Given prior knowledge of the state of a system and a set of observations, we wish to estimate the state of the system at a given time. This is known as the posterior or analysis.
- Bayes' theorem states

$$p(\mathbf{x}|\mathbf{y}) \propto p(\mathbf{x})p(\mathbf{y}|\mathbf{x})$$

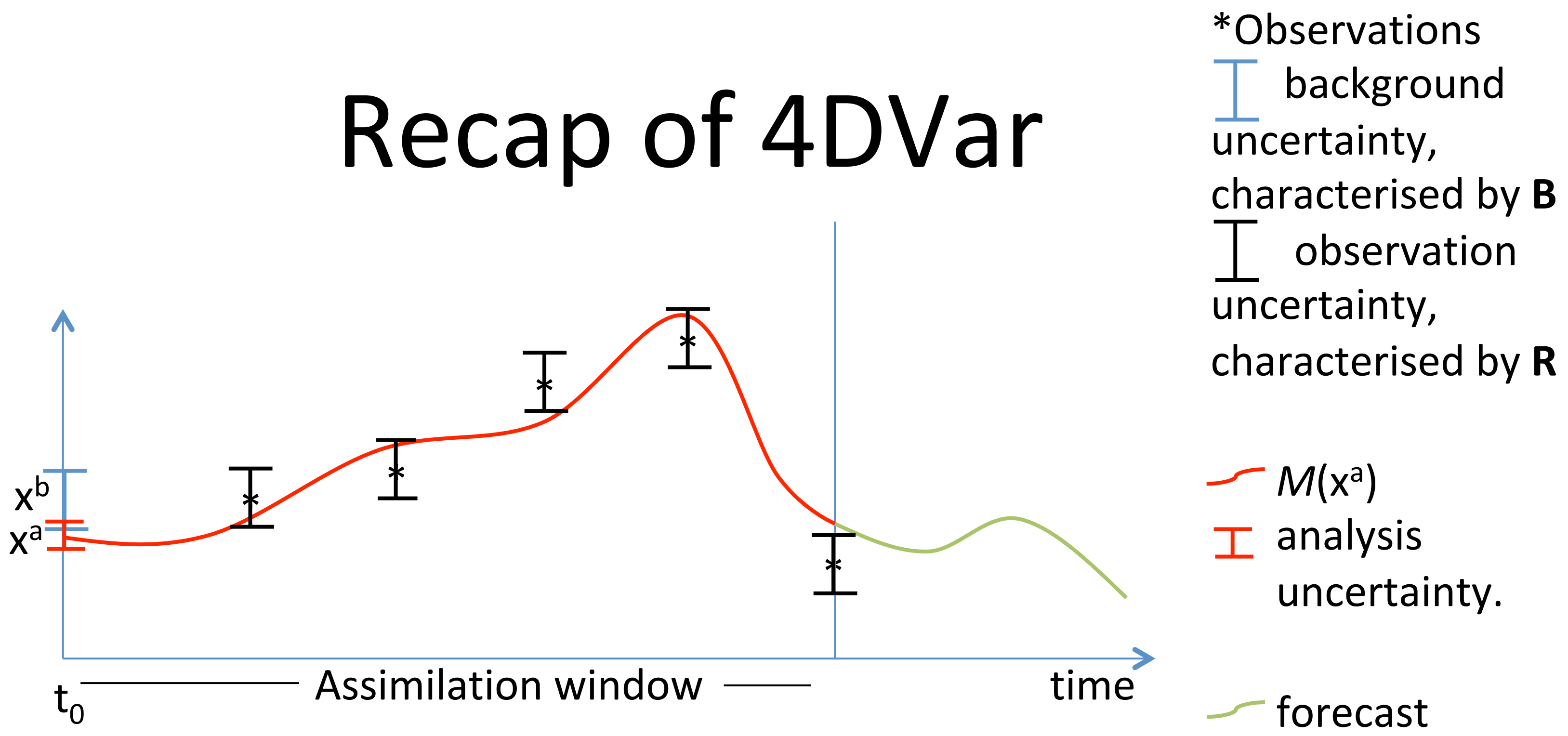
Figure: 1D example of Bayes' theorem.



No rain Moderate rain Heavy rain

For example this could be rainfall amount in a given grid box.
A-priori we are unsure if there will be moderate or heavy rainfall. The observation only gives probability to the rainfall being moderate.
Applying Bayes' theorem we can now be certain that the rainfall was moderate and the uncertainty is reduced compared to both the observations and our a-priori estimate.

Recap of 4DVar



- 4DVar aims to find the most likely state at time t_0, given an initial estimate, x_b, and a window of observations.

$$
\begin{aligned}
\mathbf{x}_0^a &= \arg\max_{\mathbf{x}_0}(p(\mathbf{x}_0|\mathbf{y}_1,\ldots,\mathbf{y}_p)) \\
&= \arg\min_{\mathbf{x}_0}(-log(p(\mathbf{x}_0|\mathbf{y}_1,\ldots,\mathbf{y}_p))) \\
&= \arg\min_{\mathbf{x}_0}(J(\mathbf{x}_0))
\end{aligned}
$$

Recap of 4DVar

- J (the cost function) is derived assuming Gaussian error distributions and a perfect model.

$$
\begin{aligned}
J(\mathbf{x}_0) = \quad & (\mathbf{x}_0 - \mathbf{x}_0^{\mathrm{b}})^T \mathbf{B}^{-1} (\mathbf{x}_0 - \mathbf{x}_0^{\mathrm{b}}) \\
& + \sum_{i=1}^{p} (\mathbf{y}_i - H(M(\mathbf{x}_0, t_0, t_i)))^T \mathbf{R}_i^{-1} (y_i - H(M(\mathbf{x}_0, t_0, t_i)))
\end{aligned}
$$

Recap of 4DVar: why do any different?

Advantages

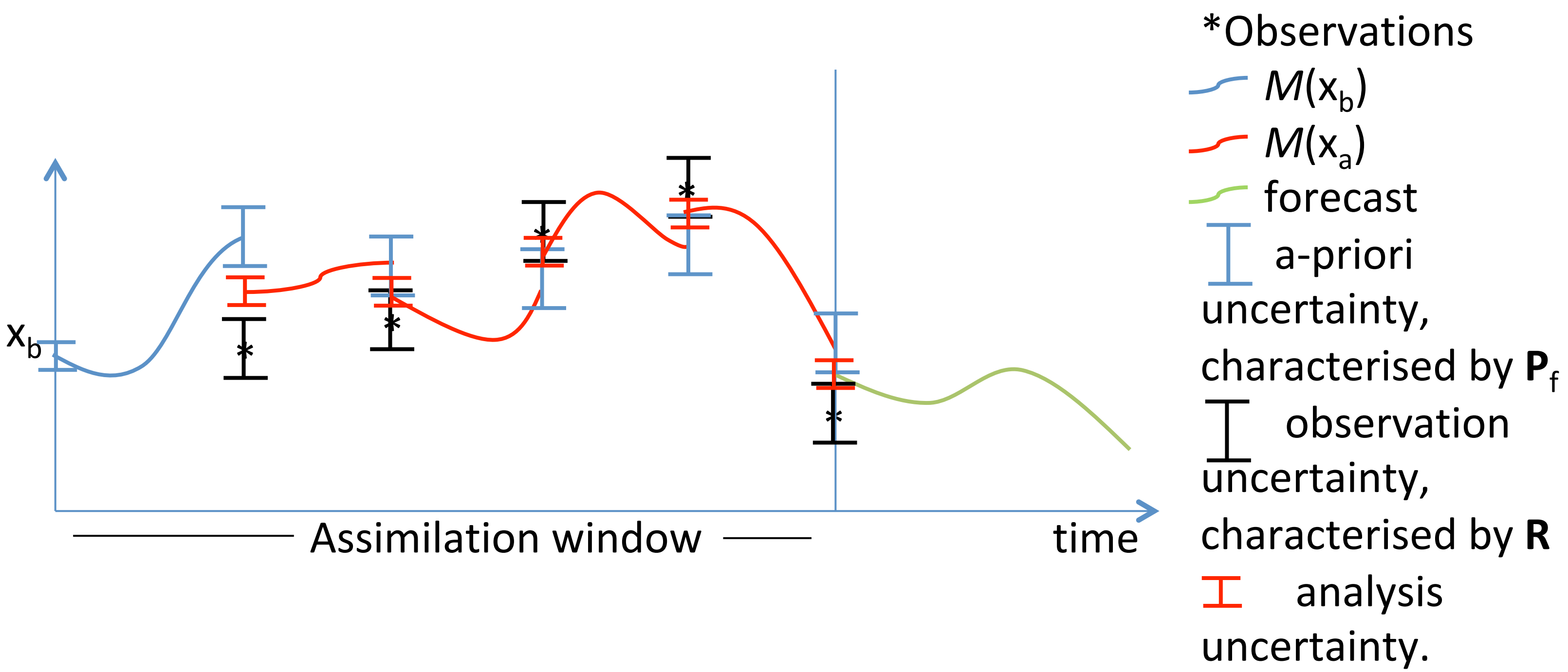
- Met Office and ECMWF both use methods based on 4DVar for their atmospheric assimilation.
- Gaussian and near-linear assumption makes this an efficient algorithm.
- Minimisation of the cost function is a well posed problem (the B-matrix is designed to be full rank).
- Analysis is consistent with the model (balanced).
- Lots of theory and techniques to modify the basic algorithm to make it a pragmatic method for various applications, e.g. incremental 4DVar, preconditioning, control variable transforms, weak constraint 4DVar...

Disadvantages

- Gaussian assumption is not always valid.
- Relies on the validity of TL and perfect model assumption. This tends to restrict the length of the assimilation window.
- Development of TL model, $\mathbf{M}$, and adjoint, $\mathbf{M}^T$, is very time consuming and difficult to update as the non-linear model is developed.
- B-matrix is predominately static.

This motivates a different approach…

Sequential DA



- Instead of assimilating all observations at one time, assimilate them sequentially in time.
- This can be shown to be equivalent to the variational problem, assuming a linear model and all error covariances are treated consistently.

Sequential DA- The Kalman equations

- At each observation time *k* the prior uncertainty, p(x_k), is updated to find the posterior, p(x_k|y_k).
- Recall Bayes' theorem: $p(\mathbf{x}_k|\mathbf{y}_k) \propto p(\mathbf{x}_k)p(\mathbf{y}_k|\mathbf{x}_k)$.

- Let us assume the distributions are Gaussian,

$$p(\mathbf{x}_k) \sim N(\mathbf{x}_k^f, \mathbf{P}_k^f), \qquad p(\mathbf{y}_k|\mathbf{x}_k) \sim N(\mathbf{y}_k, \mathbf{R}).$$

Note the change in notation! See final slide for a list of all notation used

- This implies that the posterior is also Gaussian $p(\mathbf{x}_k|\mathbf{y}_k) \sim N(\mathbf{x}_k^a, \mathbf{P}_k^a)$
- The mean of $p(\mathbf{x}_k|\mathbf{y}_k)$ is given by

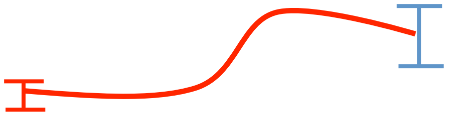
$$\mathbf{x}_k^a = \mathbf{x}_k^f + \mathbf{K}_k(\mathbf{y}_k - h(\mathbf{x}_k^f)) \quad \text{where } \mathbf{K} = \mathbf{P}_k^f\mathbf{H}^{\mathrm{T}}(\mathbf{H}\mathbf{P}_k^f\mathbf{H}^{\mathrm{T}} + \mathbf{R})^{-1}$$

- And the covariance is given by

$$\mathbf{P}_k^a = (\mathbf{I} - \mathbf{K}_k\mathbf{H})\mathbf{P}_k^f$$

The Kalman filter

- Need to be able to evolve the uncertainty in the state from one observation time to the next.



- That is to go from $p(x_{k-1}|y_{k-1})$ to $p(x_k)$, or

$$\mathbf{x}^{\mathrm{a}}_{k-1} \to \mathbf{x}^{\mathrm{f}}_{k} \text{ and } \mathbf{P}^{\mathrm{a}}_{k-1} \to \mathbf{P}^{\mathrm{f}}_{k}$$

- The Extended Kalman filter (EKF, Grewal and Andrews (2008)) does this using the non-linear model and its TL and adjoint.

The Kalman filter algorithm

Prediction step

- Evolve mean state from time *n-1* to time of observation, *n*.

$$\mathbf{x}_n^{\mathrm{f}} = M(\mathbf{x}_{n-1}^{\mathrm{a}}) + \eta_n$$

where $\eta \sim N(0, \mathbf{Q})$

- Evolve covariance to time of observation allowing for model error,

$$\mathbf{P}_n^{\mathrm{f}} = \mathbf{M}\mathbf{P}_{n-1}^{\mathrm{a}}\mathbf{M}^{\mathrm{T}} + \mathbf{Q}$$

Observation update step

- Update mean state given observation $\mathbf{x}_n^{\mathrm{a}} = \mathbf{x}_n^{\mathrm{f}} + \mathbf{K}_n(\mathbf{y}_n - \mathbf{H}\mathbf{x}_n^{\mathrm{f}})$

where $\mathbf{K}_n = \mathbf{P}_n^{\mathrm{f}}\mathbf{H}^{\mathrm{T}}(\mathbf{H}\mathbf{P}_n^{\mathrm{f}}\mathbf{H}^{\mathrm{T}} + \mathbf{R})^{-1}$

- Update error covariance given observation

$$\mathbf{P}_n^{\mathrm{a}} = (\mathbf{I} - \mathbf{K}_n\mathbf{H})\mathbf{P}_n^{\mathrm{f}}$$

Motivation for the ensemble Kalman filter (EnKF)

- The EKF still needs the TL and adjoint model to propagate the covariance matrix.

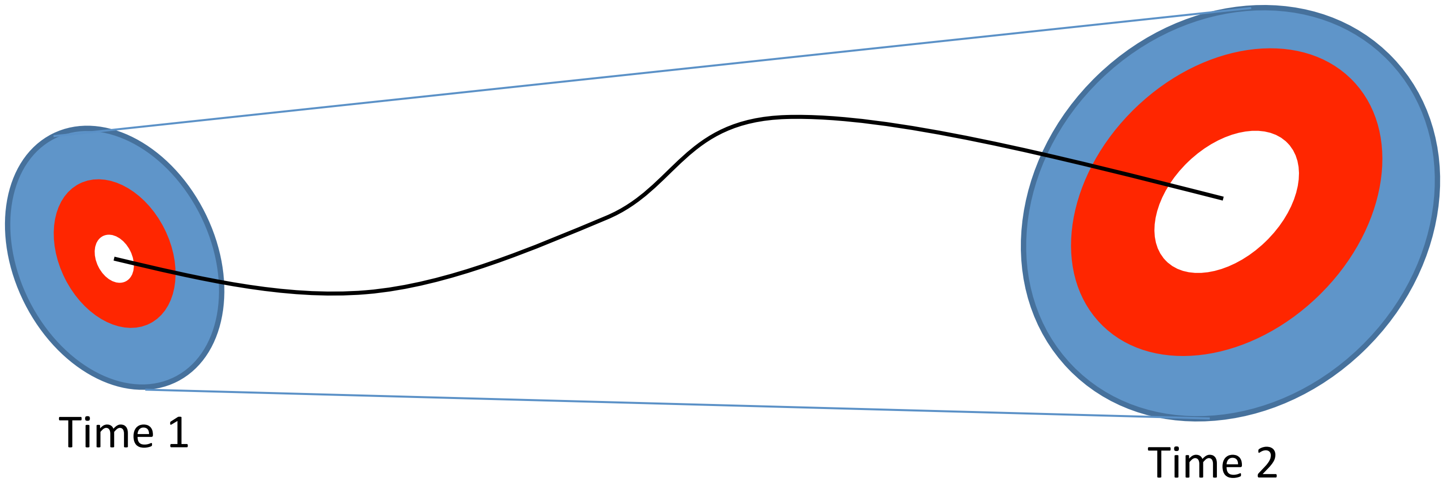
$$\mathbf{P}_n^{\mathrm{f}} = \mathbf{M}\mathbf{P}_{n-1}^{\mathrm{a}}\mathbf{M}^{\mathrm{T}} + \mathbf{Q}$$

- Due to the size of this matrix for most environmental applications, the EKF is not feasible in practice.

- An alternative approach to explicitly evolving the full covariance matrix is to instead estimate it using a sample of evolved states (known as the **ensemble**).

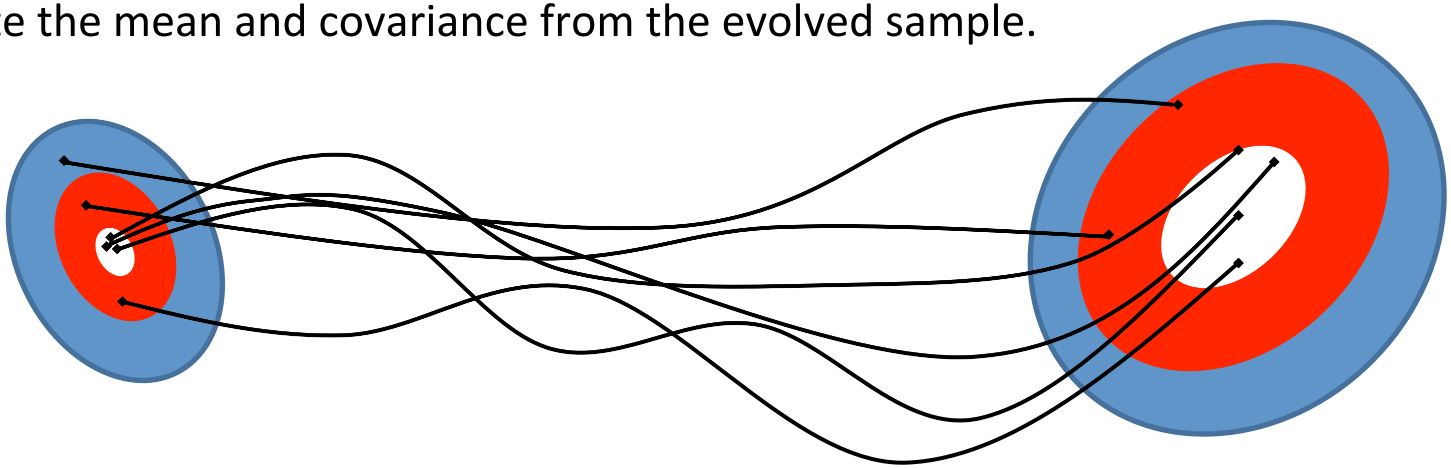
Extended Kalman filter approach

Explicitly evolve the mean and covariances forward in time using M, **M** and $\mathbf{M}^T$.



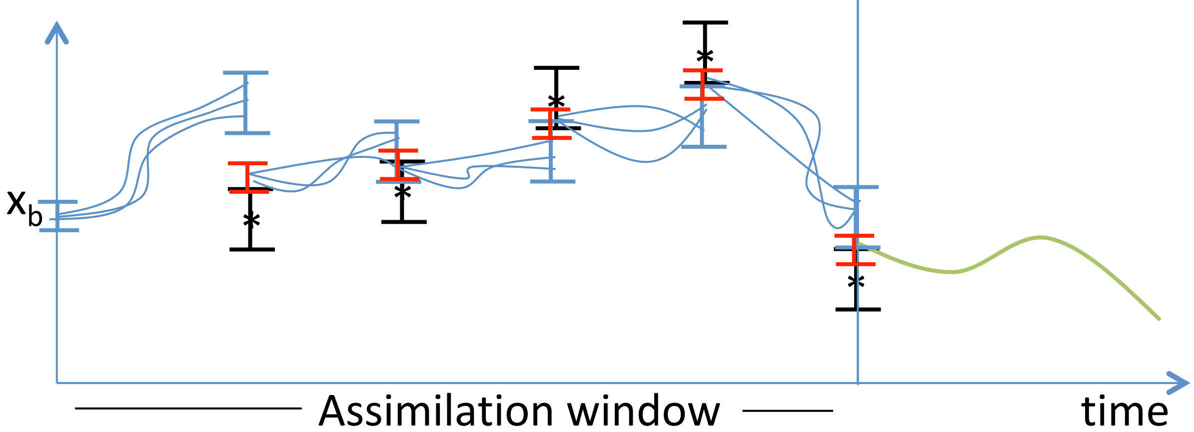
Ensemble Kalman filter approach

Sample from the initial time uncertainty, evolve each state forward in time using M, then estimate the mean and covariance from the evolved sample.



EnKF algorithms

- The EnKF (Envensen 1994) merges KF theory with Monte Carlo estimation methods.
- There are many many different flavours of EnKF.
- EnKF algorithms can be generalised into two main categories:
  - Stochastic algorithms (e.g. the perturbed observation Kalman filter)
  - Deterministic algorithms (e.g. the ensemble transform Kalman filter)
- All EnKF methods can be represented by the same basic schematic:



The perturbed observation Kalman Filter

Prediction step

- Evolve each ensemble member forward using the non-linear model with added noise.

$$\mathbf{x}_n^{(i),\mathrm{f}} = M(\mathbf{x}_{n-1}^{(i),\mathrm{a}}) + \eta^{(i)} \qquad \eta \sim N(0, \mathbf{Q})$$

- Reconstruct the ensemble mean

$$\bar{\mathbf{x}}_n^{\mathrm{f}} = \frac{1}{N}\sum_{i=1}^{N} \mathbf{x}_n^{(i),\mathrm{f}}$$

- And its covariance

$$\mathbf{P}_n^{\mathrm{f}} = \frac{1}{N-1}\mathbf{X}'^{\mathrm{f}}_n(\mathbf{X}'^{\mathrm{f}}_n)^{\mathrm{T}} = \frac{1}{N-1}\sum_{i=1}^{N}(\mathbf{x}_n^{(i),\mathrm{f}} - \bar{\mathbf{x}}_n^{\mathrm{f}})(\mathbf{x}_n^{(i),\mathrm{f}} - \bar{\mathbf{x}}_n^{\mathrm{f}})^{\mathrm{T}}$$

- Note there is no need to ever explicitly compute $\mathbf{P}^{\mathrm{f}} \in \mathbb{R}^{n\times n}$, just the perturbation matrix $\mathbf{X}'^{\mathrm{f}} \in \mathbb{R}^{n\times N}$, which is generally of a smaller dimension.

The perturbed observation Kalman Filter

Filtering step

– Update the ensemble using perturbed observations

$$\mathbf{x}_n^{(i),\mathrm{a}} = \mathbf{x}_n^{(i),\mathrm{f}} + \mathbf{K}_n(\mathbf{y}_n + \epsilon_{\mathrm{y}}^{(i)} - \mathbf{H}\mathbf{x}_n^{(i),\mathrm{f}})$$

Where $\epsilon_y \sim N(0, \mathbf{R})$ and $\mathbf{K}_n = \mathbf{P}_n^{\mathrm{f}}\mathbf{H}^{\mathrm{T}}(\mathbf{H}\mathbf{P}_n^{\mathrm{f}}\mathbf{H}^{\mathrm{T}} + \mathbf{R})^{-1}$

Derived from the ensemble

$$\frac{1}{N-1}\mathbf{X}'^{\mathrm{f}}_n(\mathbf{X}'^{\mathrm{f}}_n)^{\mathrm{T}}$$

The perturbed observation Kalman Filter

- The advantages of the perturbed observation KF is that it is very simple to implement and understand for toy models. However…
  - It is necessary to perturb the observations in order for the variance of the ensemble after the update step to correctly represent the uncertainty in the analysis.

$$\mathbf{P}_n^{\mathrm{a}} = (\mathbf{I} - \mathbf{K}_n \mathbf{H})\mathbf{P}_n^{\mathrm{f}}$$

  - This introduces additional sampling noise.

- This motivates the development of square-root or deterministic forms of the EnKF that do not need to perturb the observations.

Ensemble Square Root Filter

- The idea of ESRF is to create an updated ensemble with covariance consistent with

$$\mathbf{P}_n^{\mathrm{a}} = (\mathbf{I} - \mathbf{K}_n\mathbf{H})\mathbf{P}_n^{\mathrm{f}}$$

- Recall that the ensemble covariance matrix is given by

$$\mathbf{P}_n^{\mathrm{f}} = \frac{1}{N-1}\mathbf{X'}_n^{\mathrm{f}}(\mathbf{X'}_n^{\mathrm{f}})^{\mathrm{T}} = \frac{1}{N-1}\sum_{i=1}^{N}(\mathbf{x}_n^{(i),\mathrm{f}} - \bar{\mathbf{x}}_n^{\mathrm{f}})(\mathbf{x}_n^{(i),\mathrm{f}} - \bar{\mathbf{x}}_n^{\mathrm{f}})^{\mathrm{T}}$$

- We can write a similar expression for the analysis error covariance matrix in terms of the analysis perturbations

$$\mathbf{P}_n^{\mathrm{a}} = \frac{1}{N-1}\mathbf{X'}_n^{\mathrm{a}}(\mathbf{X'}_n^{\mathrm{a}})^{\mathrm{T}} = \frac{1}{N-1}\sum_{i=1}^{N}(\mathbf{x}_n^{(i),\mathrm{a}} - \bar{\mathbf{x}}_n^{\mathrm{a}})(\mathbf{x}_n^{(i),\mathrm{a}} - \bar{\mathbf{x}}_n^{\mathrm{a}})^{\mathrm{T}}$$

- Instead of updating each ensemble member separately, as in the perturbed observation KF, the ESRF generates the new ensemble simultaneously by updating $\mathbf{X}^{\mathrm{f}}$ instead of $\mathbf{x}^{(i),\mathrm{f}}$

Ensemble Square Root Filter

<u>Prediction step</u>

- This is the same as for the perturbed observation ensemble KF.
- Compute the evolved ensemble, $\mathbf{x}_n^{(i),\mathrm{f}}$, its mean, $\bar{\mathbf{x}}_n^{\mathrm{f}}$, and its perturbation matrix, $\mathbf{X}'^{\mathrm{f}}_n$.
- Compute the forecast-observation ensemble
  - Transform the forecast ensemble to observation space

$$\mathbf{y}_n^{(i),\mathrm{f}} = H(\mathbf{x}_n^{(i),\mathrm{f}})$$

  - from this can compute the mean $\bar{\mathbf{y}}_n^{\mathrm{f}} = \frac{1}{N}\sum_{i=1}^{N} \mathbf{y}_n^{(i),\mathrm{f}}$
  - and perturbation matrix

$$\mathbf{Y}'^{\mathrm{f}}_n = [(\mathbf{y}_n^{(1),\mathrm{f}} - \bar{\mathbf{y}}_n^{\mathrm{f}}), \ldots, (\mathbf{y}_n^{(i),\mathrm{f}} - \bar{\mathbf{y}}_n^{\mathrm{f}}), \ldots, (\mathbf{y}_n^{(N),\mathrm{f}} - \bar{\mathbf{y}}_n^{\mathrm{f}})]$$

Ensemble Square Root Filter

<u>Filtering step</u>

- Update ensemble mean $\bar{\mathbf{x}}_n^{\mathrm{a}} = \bar{\mathbf{x}}_n^{\mathrm{f}} + \mathbf{K}_n(\mathbf{y}_n - \bar{\mathbf{y}}_n^{\mathrm{f}})$
  - where $\mathbf{K}_n = \mathbf{X}'^{\mathrm{f}}_n(\mathbf{Y}'^{\mathrm{f}}_n)^{\mathrm{T}}(\mathbf{Y}'^{\mathrm{f}}_n(\mathbf{Y}'^{\mathrm{f}}_n)^{\mathrm{T}} + (N-1)\mathbf{R})^{-1}$

- Update perturbation matrix

$$\mathbf{X}'^{\mathrm{a}}_n = \mathbf{X}'^{\mathrm{f}}_n \mathbf{T}_n$$

  - Need to define the matrix **T**.

Ensemble Square Root Filter

The T matrix

- The matrix **T** is chosen such that

$$
\begin{aligned}
\mathbf{P}_n^{\mathrm{a}} &= \frac{1}{N-1}\mathbf{X'}_n^{\mathrm{a}}(\mathbf{X'}_n^{\mathrm{a}})^{\mathrm{T}} \\
&= \frac{1}{N-1}\mathbf{X'}_n^{\mathrm{f}}\mathbf{T}_n(\mathbf{X'}_n^{\mathrm{f}}\mathbf{T}_n)^{\mathrm{T}} \\
&\approx (\mathbf{I}-\mathbf{K}_n\mathbf{H})\mathbf{P}_n^{\mathrm{f}}
\end{aligned}
$$

- This does not uniquely define **T** which is why there are so many different variants of the ESRF, e.g. the Ensemble Adjustment Kalman Filter (Anderson (2001), and the Ensemble Transform Kalman Filter Bishop et al. (2001))
- Tippet et al. (2003) review several square root filters and compare their numerical efficiency. Show that although they lead to different ensembles they all span the same subspace.

Ensemble Square Root Filter

An expression for **T** can be found by rearranging $(\mathbf{I} - \mathbf{K}_n\mathbf{H})\mathbf{P}_n^{\mathrm{f}}$ using

$\mathbf{K}_n = \mathbf{P}_n^{\mathrm{f}}\mathbf{H}^{\mathrm{T}}(\mathbf{H}\mathbf{P}_n^{\mathrm{f}}\mathbf{H}^{\mathrm{T}} + \mathbf{R})^{-1}$, $\mathbf{P}_n^{\mathrm{f}} = \frac{1}{N-1}\mathbf{X}'^{\mathrm{f}}_n(\mathbf{X}'^{\mathrm{f}}_n)^{\mathrm{T}}$ and $\mathbf{Y}'^{\mathrm{f}} = \mathbf{H}\mathbf{X}'^{\mathrm{f}}$

$$
\begin{aligned}
&\tfrac{1}{N-1}\mathbf{X}'^{\mathrm{f}}\mathbf{T}(\mathbf{X}'^{\mathrm{f}}\mathbf{T})^{\mathrm{T}} \\
&= (\mathbf{I} - \mathbf{K}\mathbf{H})\mathbf{P}^{\mathrm{f}} \\
&= (\mathbf{I} - \mathbf{P}^{\mathrm{f}}\mathbf{H}^{\mathrm{T}}(\mathbf{H}\mathbf{P}^{\mathrm{f}}\mathbf{H}^{\mathrm{T}} + \mathbf{R})^{-1}\mathbf{H})\mathbf{P}^{\mathrm{f}} \\
&= \tfrac{1}{N-1}\left(\mathbf{I} - \tfrac{1}{N-1}\mathbf{X}'^{\mathrm{f}}(\mathbf{X}'^{\mathrm{f}})^{\mathrm{T}}\mathbf{H}^{\mathrm{T}}\left(\tfrac{1}{N-1}\mathbf{H}\mathbf{X}'^{\mathrm{f}}(\mathbf{X}'^{\mathrm{f}})^{\mathrm{T}}\mathbf{H}^{\mathrm{T}} + \mathbf{R}\right)^{-1}\mathbf{H}\right)\mathbf{X}'^{\mathrm{f}}(\mathbf{X}'^{\mathrm{f}}) \\
&= \tfrac{1}{N-1}\mathbf{X}'^{\mathrm{f}}\left(\mathbf{I} - (\mathbf{Y}'^{\mathrm{f}})^{\mathrm{T}}\left(\mathbf{Y}'^{\mathrm{f}}(\mathbf{Y}'^{\mathrm{f}})^{\mathrm{T}} + (N-1)\mathbf{R}\right)^{-1}\mathbf{Y}'^{\mathrm{f}}\right)(\mathbf{X}'^{\mathrm{f}})^{\mathrm{T}}
\end{aligned}
$$

$$
\implies \mathbf{T}\mathbf{T}^{\mathrm{T}} = \mathbf{I} - (\mathbf{Y}'^{\mathrm{f}})^{\mathrm{T}}\left(\mathbf{Y}'^{\mathrm{f}}(\mathbf{Y}'^{\mathrm{f}})^{\mathrm{T}} + (N-1)\mathbf{R}\right)^{-1}\mathbf{Y}'^{\mathrm{f}}
$$

The Ensemble Transform Kalman Filter

- First introduced by Bishop et al. (2001), later revised by Wang et al. (2004).
- **T** is computed using the Morrison-Woodbury identity to rewrite the previous expression for $\mathbf{TT}^{\mathrm{T}}$.

$$\mathbf{TT}^{\mathrm{T}} = \left(\mathbf{I} + \frac{1}{N-1}(\mathbf{Y}'^{\mathrm{f}})^{\mathrm{T}}\mathbf{R}^{-1}\mathbf{Y}'^{\mathrm{f}}\right)^{-1}$$
$$= (\mathbf{U\Sigma U}^{\mathrm{T}})^{-1}$$
$$\Longrightarrow \mathbf{T} = \mathbf{U\Sigma}^{-1/2}\mathbf{U}^{\mathrm{T}}$$

- The revision by Wang et al. highlighted that any **T** which satisfies the estimate of the analysis error covariance does not necessarily lead to an unbiased analysis ensemble, see Livings et al. (2008) for conditions that T must satisfy for the analysis ensemble to be centred on the mean.

Model error

- The ensemble Kalman filter allows for an imperfect model by adding noise at each time step of the model evolution.

$$\mathbf{x}_n^{(i)} = M(\mathbf{x}_{n-1}^{(i)}, t_{n-1}, t_n) + \eta_n^{(i)} \qquad \eta_n^{(i)} \sim N(\mathbf{0}, \mathbf{Q}_n)$$

- The matrix **Q** is not explicitly needed in the algorithm, only the effect of the model error in the evolution of the state.
- There have been many different strategies to including model error in the ensemble, based on where you think the source of the error lies. A few examples are
  - **Multiphysics-** different physical models are used in each ensemble member
  - **Stochastic kinetic energy backscatter-** replaces upscale kinetic energy loss due to unresolved processes and numerical integration.
  - **Stochastically perturbed physical tendencies**
  - **Perturbed parameters**
  - Or **combinations** of the above

Summary of the Ensemble Kalman Filter

Advantages

- The a-priori uncertainty is flow-dependent.
- The code can be developed separately from the dynamical model e.g. PDAF system which allows for any model to assimilate observations using ensemble techniques (see http://www.met.reading.ac.uk/~darc/empire).
- No need to linearise the model, only linear assumption is that statistics remain close to Gaussian.
- Easy to account for model error.
- Easy to parrallelise.

Disadvantages

- Sensitive to ensemble size. Undersampling can lead to filter divergence. Ideas to mitigate this include localisation and inflation (see next EnKF lecture).
- Assumes Gaussian statistics, for highly non-linear models this may not be a valid assumption (see Friday's lectures on particle filters)
- The updated ensemble mean may not be consistent with the model equations.

Summary of the Ensemble Kalman Filter

The different EnKF algorithms

– Many different algorithms exist.
– **Stochastic methods** update each ensemble member separately and then estimate the first two sample moments to give the ensemble mean and covariance.
– **Deterministic methods** update the ensemble simultaneously based on linear/Gaussian theory. May allow for a smaller ensemble than the stochastic methods as avoids some of the sampling error.

EnKF vs 4DVar

– Each method has its own advantages and disadvantages- there is no clear winner.
– Hybrid methods aim to combine the best bits of both (see tomorrow's lectures on hybrid methods).

Further reading

Kalman Filter: •Grewal and Andrews (2008) **Kalman Filtering: Theory and Practice using MATLAB**. *Wiley, New Jersey*. •Kalman (1960**) A new approach to linear filtering and prediction problems**. *J. Basic Engineering*, **82**, 32-45.

Stochastic Ensemble Kalman Filter:•Evensen (1994) **Sequential data assimilation with a nonlinear quasi-geostrophic model using Monte Carlo methods to forecast error statistics**. *J. Geophys. Res.,* **99(C5)**, 10143-10162.

Determanistic Ensemble Kalman filter:•Anderson (2001) **An ensemble adjustment filter for data assimilation**. *Mon. Weather Rev*., **129**, 2884-2903. •Bishop et al. (2001) **Adaptive sampling with the ensemble transform Kalman filter**. *Mon. Wea. Rev*., **126**, 1719-1724.•Tippet et al. (2003) **Ensemble Square Root Filters**. *Mon. Wea. Rev*., **131**, 1485-1490. •Livings et al. (2008) **Unbiased ensemble square root filters.** *Physica D*. **237**, 1021-1028. •Wang et al. (2004) **Which Is Better, an Ensemble of Positive–Negative Pairs or a Centered Spherical Simplex Ensemble?** *Mon. Wea. Rev*., **132**, 1590-1605

Model error:•Berner et al. (2011) **Model uncertainty in a mesoscale ensemble prediction system: Stochastic versus multiphysics representations**, *Mon. Weather Rev*., **139**, 1972–1995.

Reviews:•Bannister (2017) **A review of operational methods of variational and ensemble-variational data assimilation**. Q. J. R. Meteorol. Soc., 143: 607 – 633. •Houtekamer and Zhang (2016) **Review of the Ensemble Kalman Filter for Atmospheric Data Assimilation** . *Mon. Wea. Rev*., **144**, 4489–4532. •Vetra-Carvalho et al. (2018) **State-of-the-art stochastic data assimilation methods for high-dimensional non-Gaussian problems.** Tellus A, https://doi.org/10.1080/16000870.2018.1445364

Table 1: Notation for EnKF lectures

$\mathbf{B}$	Background error covariance matrix (4Dvar)	$p(\mathbf{x}\|\mathbf{y})$	posterior probability dist. fn
H, $\mathbf{H}$	observation operator, linearised observation operator	$p(\mathbf{x})$	prior probability dist. fn
$J(\mathbf{x})$	variational cost function	$p(\mathbf{y}\|\mathbf{x})$	likelihood probability dist. fn
$\mathbf{K}_n$	Kalman gain matrix valid at timestep n	$\mathbf{x}_n^{\mathrm{a}}$	analysis at time n
M, $\mathbf{M}$	time evolution model, linearised model	$\bar{\mathbf{x}}_n^{\mathrm{a}}$	mean analysis at time n
N	ensemble size	$\mathbf{x}^{\mathrm{b}}$	background (4Dvar)
$\mathbf{P}_n^{\mathrm{a}}$	Analysis error covariance matrix valid at timestep n	$\mathbf{x}_n^{\mathrm{f}}$	forecast at time n
$\mathbf{P}_n^{\mathrm{f}}$	Forecast error covariance matrix valid at timestep n	$\bar{\mathbf{x}}_n^{\mathrm{f}}$	mean forecast at time n
$\mathbf{R}$	Observation error covariance matrix	$\mathbf{x}_n^{(i),\mathrm{a}}$	i^{th} analysis ensemble member at time n
$\mathbf{R}_{\mathrm{e}}$	Reconstructed observation error covariance matrix (pert. ob EnKF)	$\mathbf{x}_n^{(i),\mathrm{f}}$	i^{th} forecast ensemble member at time n
$\mathbf{Q}$	Model error covariance matrix	$\mathbf{y}_n$	observation vector at time n (ESRF)
$\mathbf{T}$	T matrix (see ESRF)	$\bar{\mathbf{y}}_n^{\mathrm{f}}$	mean forecast-observation at time n (ESRF)
$\mathbf{X}_n^{\prime\mathrm{a}}$	Analysis perturbation matrix valid at timestep n	$\mathbf{y}_n^{(i),\mathrm{f}}$	i^{th} forecast-observation ensemble member at time n (ESRF)
$\mathbf{X}_n^{\prime\mathrm{f}}$	Forecast perturbation matrix valid at timestep n	$\epsilon_{\mathrm{y}}^{(i)}$	i^{th} realisation of the observation error (pert ob. EnKF)
$\mathbf{Y}_n^{\prime\mathrm{f}}$	Forecast-Observtaion perturbation matrix valid at timestep n (ESRF)	η	realisation of the model error
$\tilde{n}$	length state vector		
p	number of observations		