Exploratory workshop DADA – October 15-18 2012 Buenos Aires, Argentina

Parameter estimation using EnKF

Juan Ruiz

In colaboration with Manuel Pulido and Takemasa Miyoshi

jruiz@cima.fcen.uba.ar

UMI IFAECI (CNRS-CONICET-UBA)

Department of Atmospheric and Oceanographic Sciences, University of Buenos Aires

What is parameter estimation?

- ✓ Numerical models have a dynamic core and parameterizations of the ”physics”.
- ✓ In both components there are parameters that have some impact upon the model performance.
- ✓ Most of this parameters arise from the formulation of numerical schemes and the simplifying assumptions made in parameterizations. So these parameters are intrinsically unknown and have to be tuned in order to get a good performance.
- ✓ Some external forcings may also be treated as parameters in the model equations.

$$\frac{dx}{dt} = \underbrace{f(x,t,p_1)}_{\text{Dynamic core}} + \underbrace{par(x,t,p_2)}_{\text{Parameterizations (model physics)}} + \underbrace{f}_{\text{Forcing}}$$

Why do we want to estimate model parameters?

- Climate change detection and attribution. (??)

- Optimization of climate model for improved representation of the current climatology. Climate is independent of initial conditions, so that parameters plays in important role. (e.g. Annan et al. 2005)

- To optimize model performance in operational weather forecasting. Optimal parameters should depend on the region, season, etc. (e.g. Koyama and Watanabe 2010, Aksoy et al 2006)

- To estimate biases produced by model error. (e.g. Dee and Da Silva 2005, Baek et al 2006)

- Estimate parameter uncertainty in order to introduce a better representation of model error associated with the uncertain parameters in the forecast (i.e. ensemble forecasts with perturbations in the parameters, stochastic parameterizations, etc). (e.g. Hansen and Penland 2007)

Parameter estimation for the earth system

✓ A state of the art numerical model of the atmosphere or the ocean can have a large number of parameters. The number is substantially larger if we consider the 2D or 3D variability of the parameters (e.g. Bocquet 2011, Kang et al. 2011).

✓ Traditional optimization techniques that require several model evaluations to describe the model sensitivity to the parameter and to optimize their value, are too expensive to optimize time and space varying parameters.

✓ Data assimilation methods can provide an efficient way to estimate these model parameters and they can take into account the temporal and spatial dependence of the parameters.

Data assimilation and parameter estimation

Data assimilation is a group of methodologies that optimally combine information provided by observations with information coming from numerical models (previous observations + system dynamics) to obtain an accurate representation of the state of a system.

We can consider the model parameters as state variables in a data assimilation cycle. This approach is known as state augmentation and has been widely used.

State variables (state vector)

$$x = \begin{bmatrix} x_1 \\ \dots \\ x_i \\ \dots \\ x_n \end{bmatrix}$$

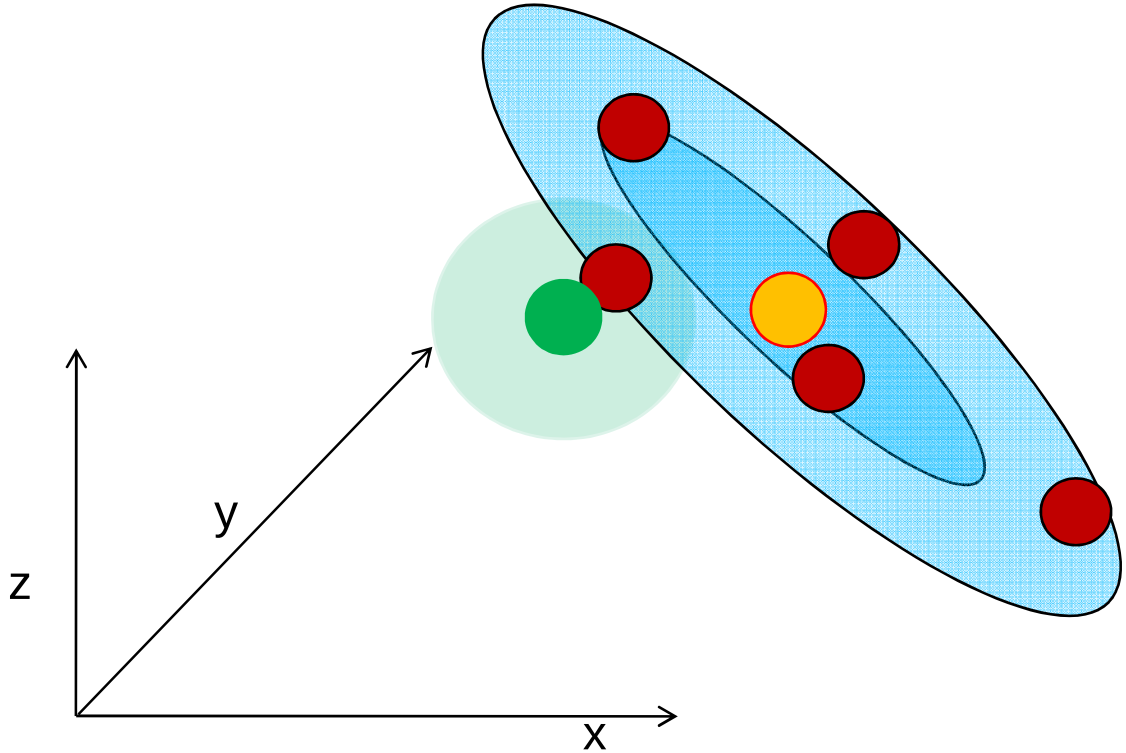
$$s = \begin{bmatrix} x_1 \\ \dots \\ x_n \\ p_1 \\ \dots \\ p_M \end{bmatrix}$$

State variables + parameters (augmented state vector)

Observations are now providing information about the optimal state of the system and also about the optimal value of the model parameters.

Data assimilation cycle with parameter estimation:

Ensemble Kalman filter



Data assimilation based on the ensemble Kalman filter: assimilation cycle and parameter estimation

$$\boxed{\begin{gathered} x_k^{f,i} = m(x_{k-1}^{a,i}, p_{k-1}^{a,i}) \\ p_k^{f,i} = p_{k-1}^{a,i} \end{gathered}}$$

Run the model
Perturb the initial conditions and the model parameters.
Parameters are assumed to be constant during the integration time.

$$P_k^f \approx \frac{1}{N-1}\sum_{j=1}^{N}(s_k^{f,i} - \overline{s^f})(s_k^{f,i} - \overline{s^f})^T$$

$$\overline{s_k^a} = \overline{s_k^b} + K_g\left(y_k^o - h\left(\overline{x_k^b}\right)\right)$$

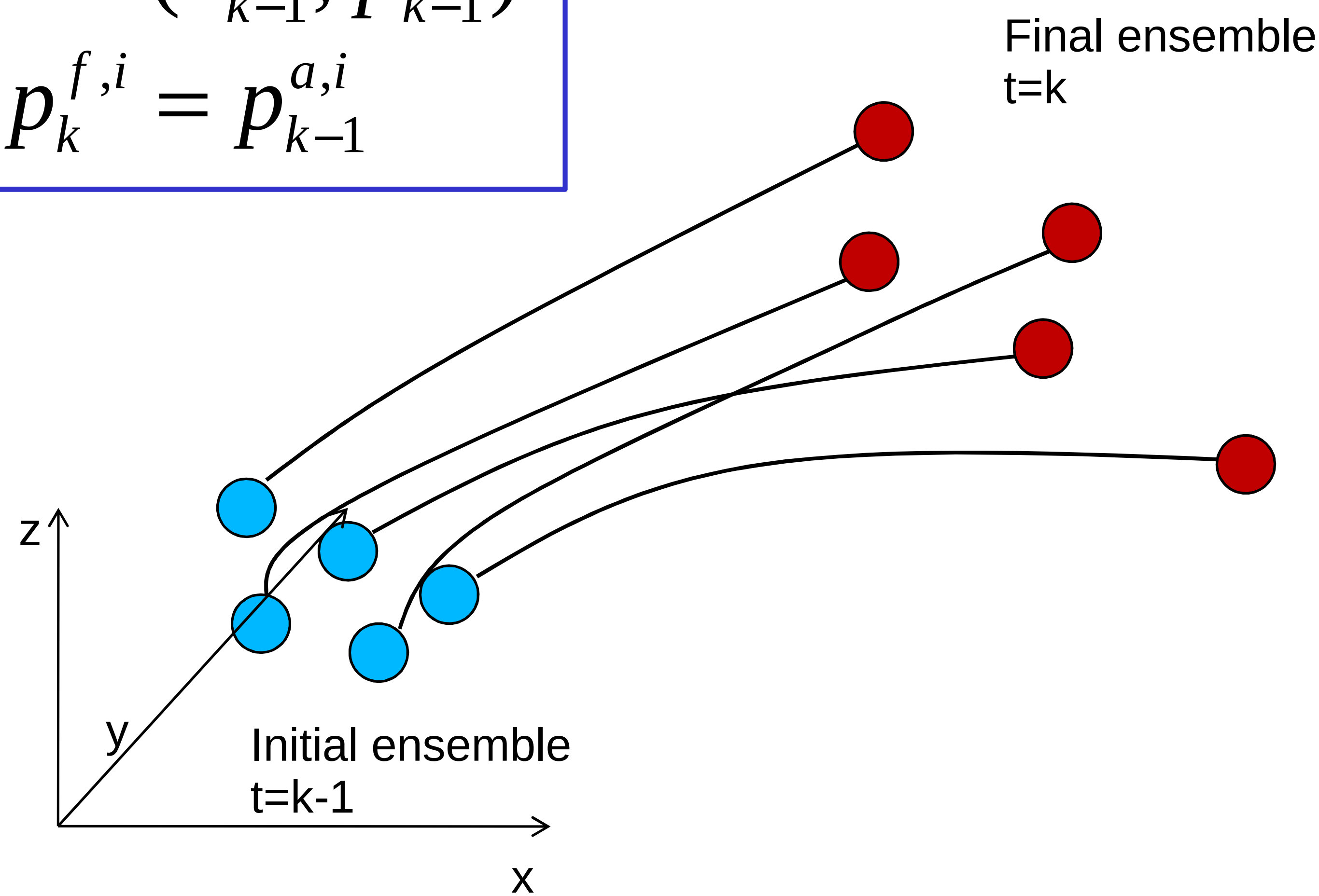
$$K_g = P_k^f H^T\left(HP_k^f H^T + R_k\right)^{-1}$$

$$P_k^a = (I - K_g)HP_k^f$$

$$P_k^a = \frac{1}{N-1}\sum_{j=1}^{N}(x_k^{a,i} - \overline{x^a})(x_k^{a,i} - \overline{x^a})^T$$

k=k+1

$$x_k^{f,i} = m(x_{k-1}^{a,i}, p_{k-1}^{a,i})$$
$$p_k^{f,i} = p_{k-1}^{a,i}$$



- ✓Integration time depends on the application (synoptic scale data assimilation ~ 6 hr). Longer integrations increases the influence of non-linear effects.

- ✓Due to chaotic behavior of the system, different model trajectories usually diverge (i.e. uncertainty about the system state grows with time).

- ✓This step is computationally intensive since the model has to be run several times. (~30-50 times)

Data assimilation based on the ensemble Kalman filter: assimilation cycle and parameter estimation

$$x_k^{f,i} = m(x_{k-1}^{a,i}, p_{k-1}^{a,i})$$

$$p_k^{f,i} = p_{k-1}^{a,i}$$

$$\boxed{P_k^f \approx \frac{1}{N-1}\sum_{j=1}^{N}(s_k^{f,i} - \overline{s^f})(s_k^{f,i} - \overline{s^f})^T}$$

Estimate the PDF of the forecast from the ensemble.
Compute the first and the second moments from the ensemble.

$$\overline{s_k^a} = \overline{s_k^b} + K_g\left(y_k^o - h\left(\overline{x_k^b}\right)\right)$$

$$K_g = P_k^f H^T \left(H P_k^f H^T + R_k\right)^{-1}$$

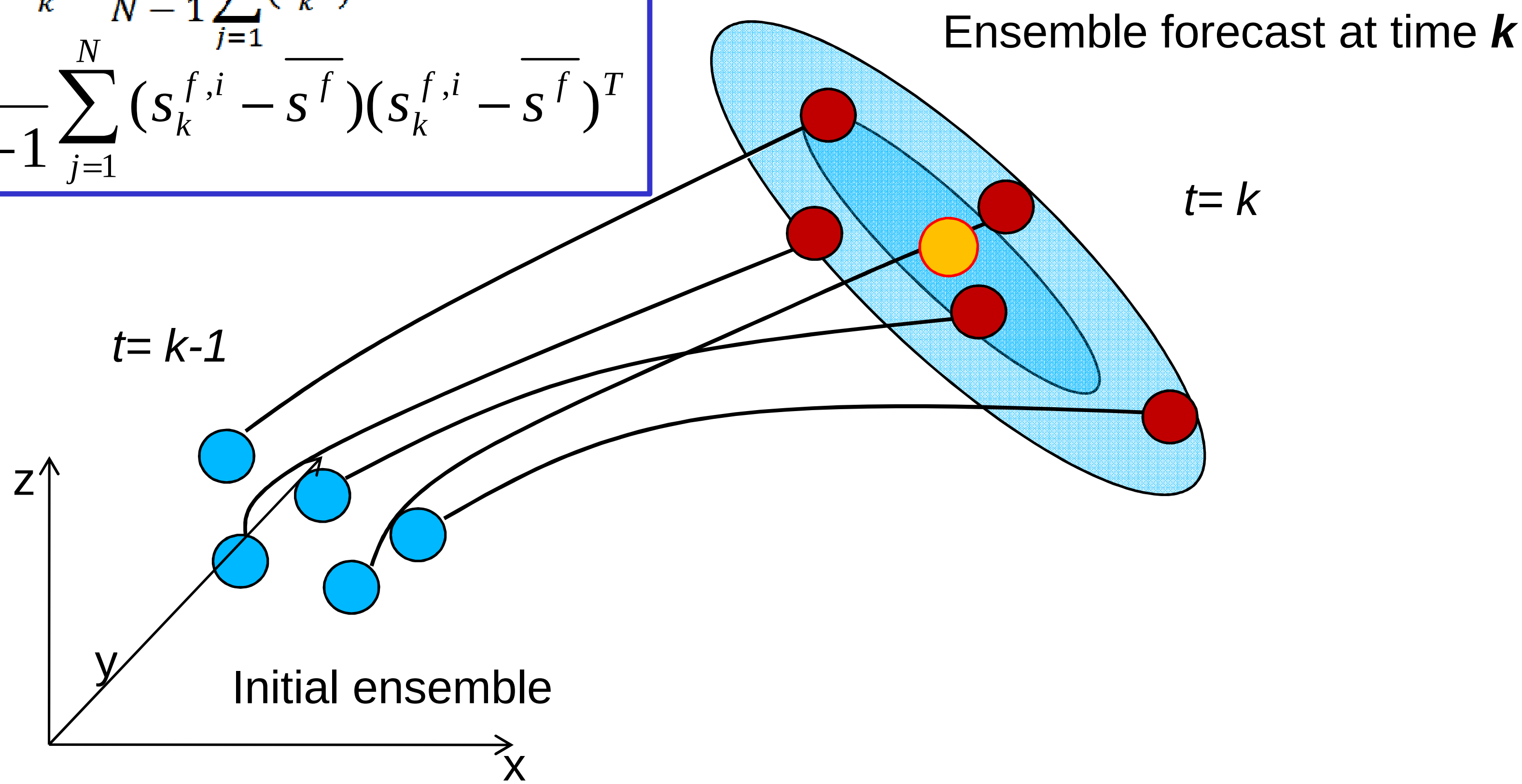
$$P_k^a = (I - K_g) H P_k^f$$

$$P_k^a = \frac{1}{N-1}\sum_{j=1}^{N}(x_k^{a,i} - \overline{x^a})(x_k^{a,i} - \overline{x^a})^T$$

k=k+1

$$\overline{s_k^f} = \frac{1}{N-1}\sum_{j=1}^{N}(s_k^{f,i})$$

$$P_k^f \approx \frac{1}{N-1}\sum_{j=1}^{N}(s_k^{f,i}-\overline{s^f})(s_k^{f,i}-\overline{s^f})^T$$



- Estimate the first and second moments of the PDF of the forecast.
- Pf contains in this case covariances between the state variables and the parameters.
- Pf evolves in time according to the non-linear dynamics of the system.
- The PDF of the forecast at time ***k*** is assumed to be Gaussian.
- When non-linear effects are important the actual PDF may be far from Gaussianity.

Data assimilation based on the ensemble Kalman filter: assimilation cycle and parameter estimation

$$x_k^{f,i} = m(x_{k-1}^{a,i}, p_{k-1}^{a,i})$$

$$p_k^{f,i} = p_{k-1}^{a,i}$$

$$P_k^f \approx \frac{1}{N-1}\sum_{j=1}^{N}(s_k^{f,i}-\overline{s^f})(s_k^{f,i}-\overline{s^f})^T$$

$$\overline{s_k^a} = \overline{s_k^b} + K_g\left(y_k^o - h\left(\overline{x_k^b}\right)\right)$$

$$K_g = P_k^f H^T \left(HP_k^f H^T + R_k\right)^{-1}$$

$$P_k^a = (I-K_g)HP_k^f$$

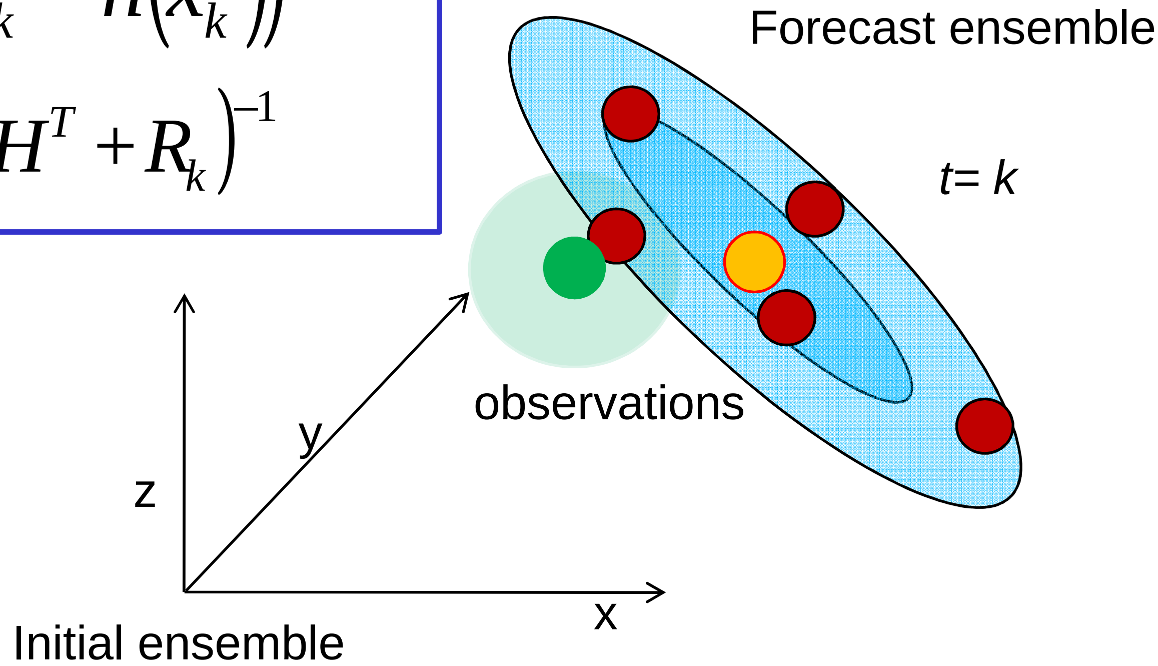
Compute the analysis and the optimal parameters
Get the analysis and its uncertainty.

$$P_k^a = \frac{1}{N-1}\sum_{j=1}^{N}(x_k^{a,i}-\overline{x^a})(x_k^{a,i}-\overline{x^a})^T$$

k=k+1

$$\overline{s_k^a} = \overline{s_k^b} + K_g\left(y_k^o - h\left(\overline{x_k^b}\right)\right)$$

$$K_g = P_k^f H^T \left(H P_k^f H^T + R_k\right)^{-1}$$



✓Get the observations for time ***k .***

✓Errors associated to the observations are assumed to be Gaussian. Errors in the observations, are often assumed to be uncorrelated among each other.

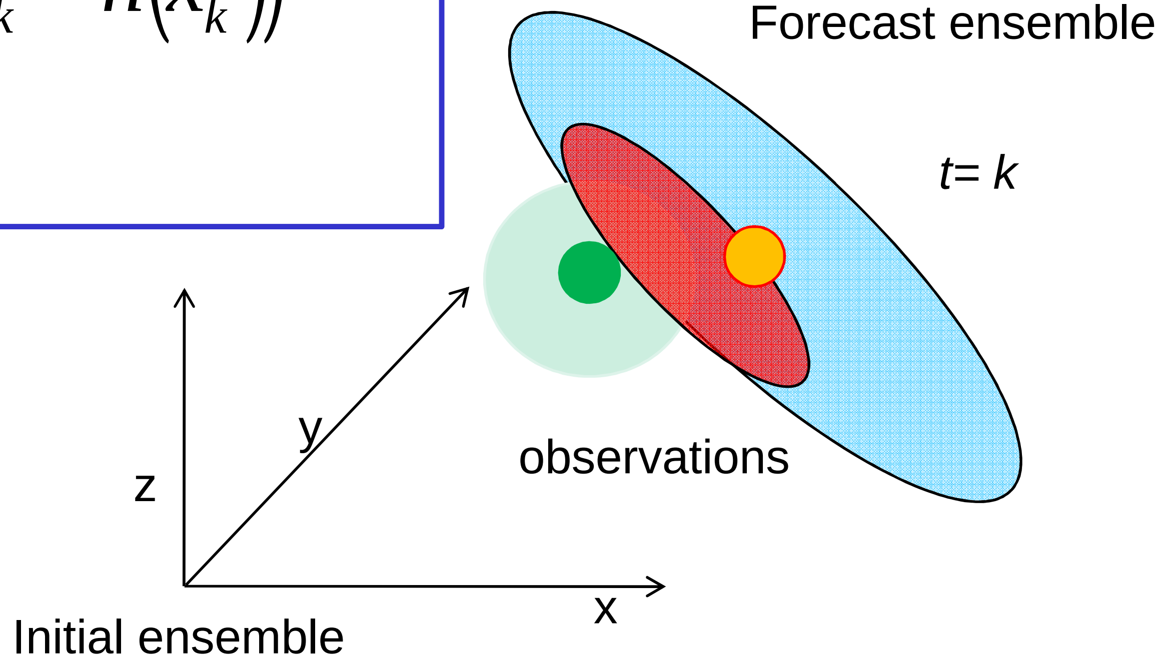
✓Errors in the forecast and in the observations are assumed to be unbiased.

✓Compare the ensemble mean against the observations. (note that parameters are not directly observed)

✓The analysis mean is obtained for the state variables and the parameters. Parameters are estimated based on the covariances between errors in the state variables and in the parameters.

$$\overline{s_k^a} = \overline{s_k^b} + K_g\left(y_k^o - h\left(\overline{x_k^b}\right)\right)$$

$$P_k^a = (I - K_g)HP_k^f$$



✓Using the Kalman filter equations obtain the analysis PDF of model states at time ***k***.

✓Obtain the PDF of the analysis. Under the afore mentioned assumptions, the shape of the analysis PDF is also Gaussian.

✓The variance of the state variables and parameters is reduced due to the information provided by the observations. This is a consequence of assuming that the distribution of the errors in the model and in the observations is Gaussian.

Data assimilation based on the ensemble Kalman filter: assimilation cycle and parameter estimation

$$x_k^{f,i} = m(x_{k-1}^{a,i}, p_{k-1}^{a,i})$$

$$p_k^{f,i} = p_{k-1}^{a,i}$$

$$P_k^f \approx \frac{1}{N-1}\sum_{j=1}^{N}(s_k^{f,i} - \overline{s^f})(s_k^{f,i} - \overline{s^f})^T$$

$$\overline{s_k^a} = \overline{s_k^b} + K_g\left(y_k^o - h\left(\overline{x_k^b}\right)\right)$$

$$K_g = P_k^f H^T \left(H P_k^f H^T + R_k\right)^{-1}$$

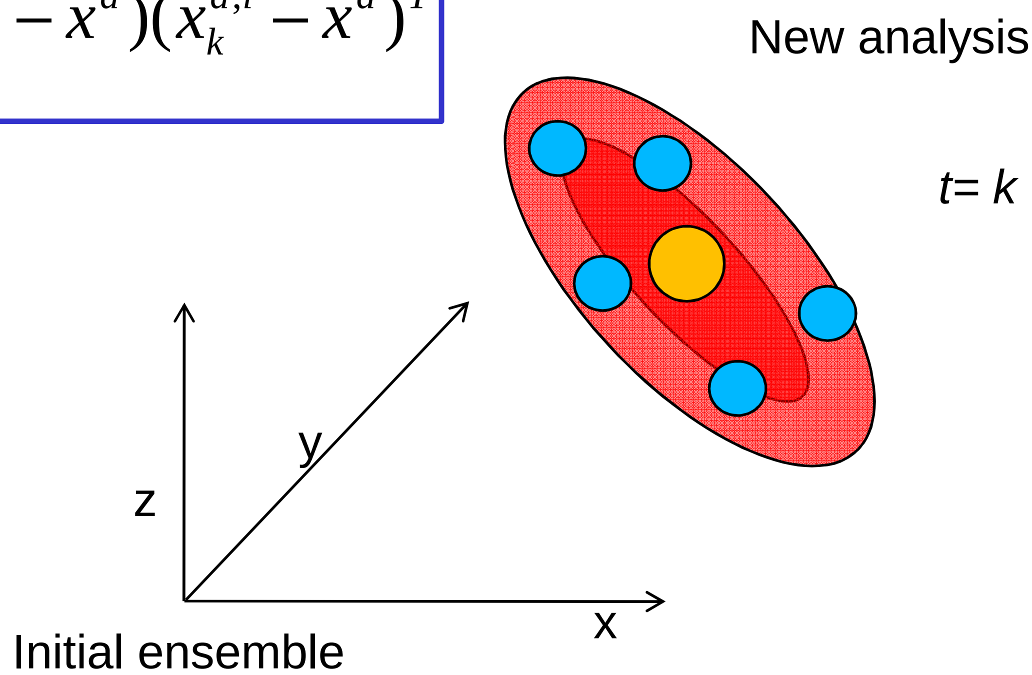
$$P_k^a = (I - K_g) H P_k^f$$

$$\boxed{P_k^a = \frac{1}{N-1}\sum_{j=1}^{N}(x_k^{a,i} - \overline{x^a})(x_k^{a,i} - \overline{x^a})^T}$$

Obtain the new ensemble members
Sample them from the analysis PDF.

k=k+1

$$P_k^a = \frac{1}{N-1}\sum_{j=1}^{N}(x_k^{a,i} - \overline{x^a})(x_k^{a,i} - \overline{x^a})^T$$



✓ Generate new ensemble members consistent with the analysis PDF.

✓ This new ensemble members for the model variables and for the parameters will be used to produce a new short range ensemble forecast and the cycle is closed!

A simple model example: Parameter estimation in the Lorenz's three variable model. (Lorenz 1963).

$$\frac{dx}{dt} = a(y - x)$$

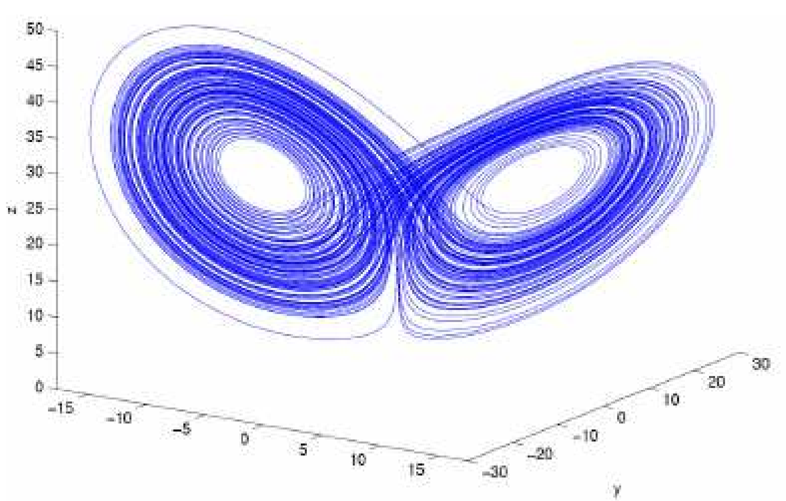
$$\frac{dy}{dt} = bx - y - xz$$

$$\frac{dz}{dt} = xy - cz$$

x, y, z are model variables

a, b, c are model parameters

strange….



…yet attractive.

Twin experiments

Twin experiment and perfect model

- The model is integrated without assimilating any information for 1000 time units. This run is assumed to be the true system evolution.

- Parameters are constant during the integration.

- Synthetic observations are generated every 1 time unit adding a random Gaussian noise to the true state.

- A 30 members ensemble is used to estimate the PDF of the model states. (in this simple experiment the ensemble size is larger than the dimension of the augmented state).

$$P^f = \begin{bmatrix} \sigma^2_{xx} & \sigma^2_{xy} & \sigma^2_{xz} & \sigma^2_{xa} & \sigma^2_{xb} & \sigma^2_{xc} \\ \sigma^2_{yx} & \sigma^2_{yy} & \sigma^2_{yz} & \sigma^2_{ya} & \sigma^2_{yb} & \sigma^2_{yc} \\ \sigma^2_{zx} & \sigma^2_{zy} & \sigma^2_{zz} & \sigma^2_{za} & \sigma^2_{zb} & \sigma^2_{zc} \\ \sigma^2_{ax} & \sigma^2_{ay} & \sigma^2_{az} & \sigma^2_{aa} & \sigma^2_{ab} & \sigma^2_{ac} \\ \sigma^2_{bx} & \sigma^2_{by} & \sigma^2_{bz} & \sigma^2_{ba} & \sigma^2_{bb} & \sigma^2_{bc} \\ \sigma^2_{cx} & \sigma^2_{cy} & \sigma^2_{cz} & \sigma^2_{ca} & \sigma^2_{cb} & \sigma^2_{cc} \end{bmatrix}$$

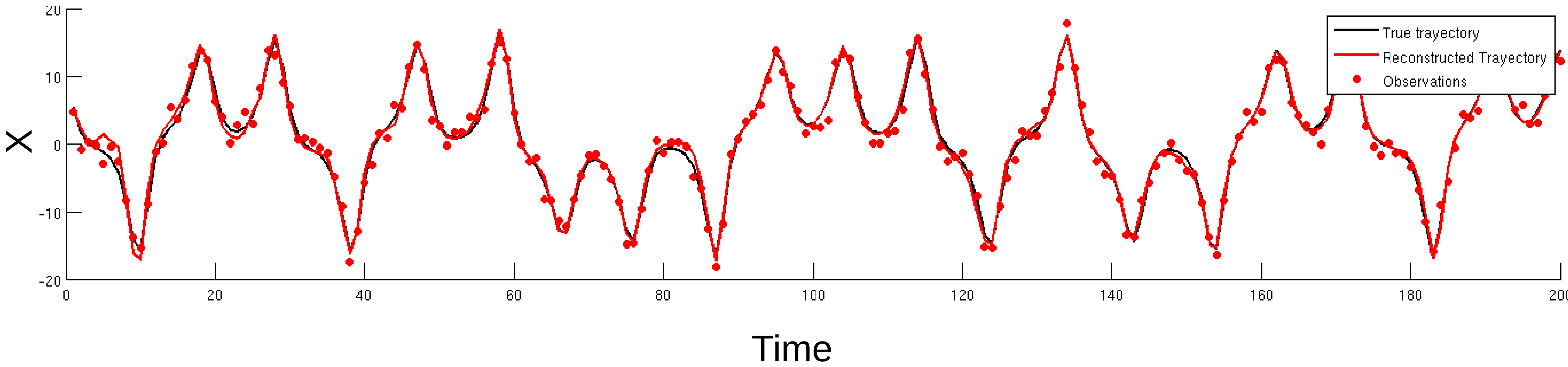
$$h = H = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

Assimilation experiment 1: constant parameters

PDF mean

A data assimilation cycle is started, at the beginning the model parameters and the model state are imperfectly known.

True X evolution, estimated x evolution and observations



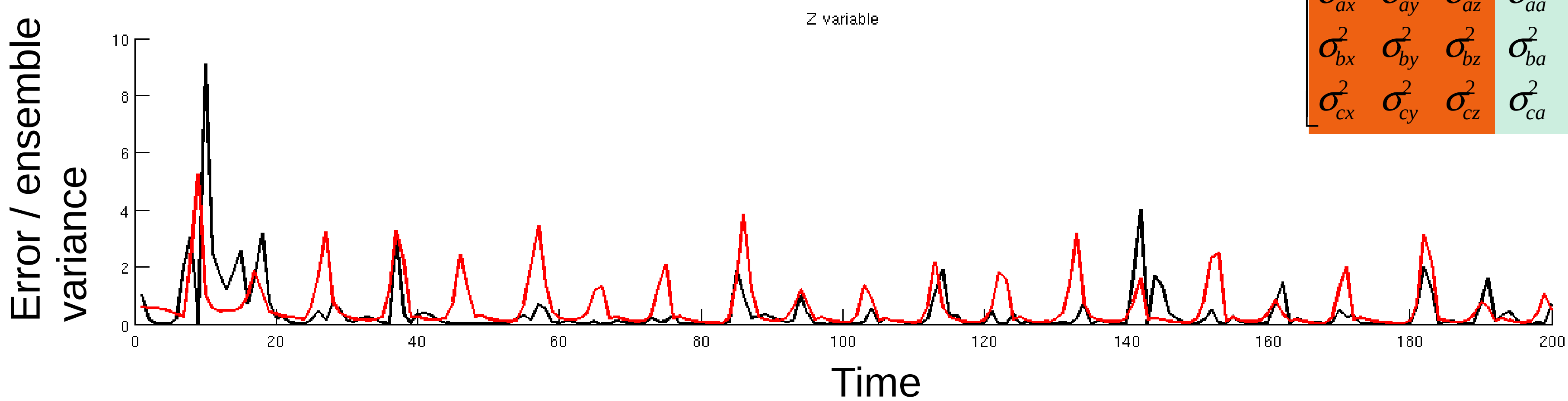
In this example we can see how the data assimilation scheme is able to reconstruct the true evolution using the information provided by the observations and by the model.

Assimilation experiment 1: constant parameters

PDF covariance

$$P^f = \begin{bmatrix} \sigma_{xx}^2 & \sigma_{xy}^2 & \sigma_{xz}^2 & \sigma_{xa}^2 & \sigma_{xb}^2 & \sigma_{xc}^2 \\ \sigma_{yx}^2 & \sigma_{yy}^2 & \sigma_{yz}^2 & \sigma_{ya}^2 & \sigma_{yb}^2 & \sigma_{yc}^2 \\ \sigma_{zx}^2 & \sigma_{zy}^2 & \sigma_{zz}^2 & \sigma_{za}^2 & \sigma_{zb}^2 & \sigma_{zc}^2 \\ \sigma_{ax}^2 & \sigma_{ay}^2 & \sigma_{az}^2 & \sigma_{aa}^2 & \sigma_{ab}^2 & \sigma_{ac}^2 \\ \sigma_{bx}^2 & \sigma_{by}^2 & \sigma_{bz}^2 & \sigma_{ba}^2 & \sigma_{bb}^2 & \sigma_{bc}^2 \\ \sigma_{cx}^2 & \sigma_{cy}^2 & \sigma_{cz}^2 & \sigma_{ca}^2 & \sigma_{cb}^2 & \sigma_{cc}^2 \end{bmatrix}$$

Z variance and error in the estimation as a function of time



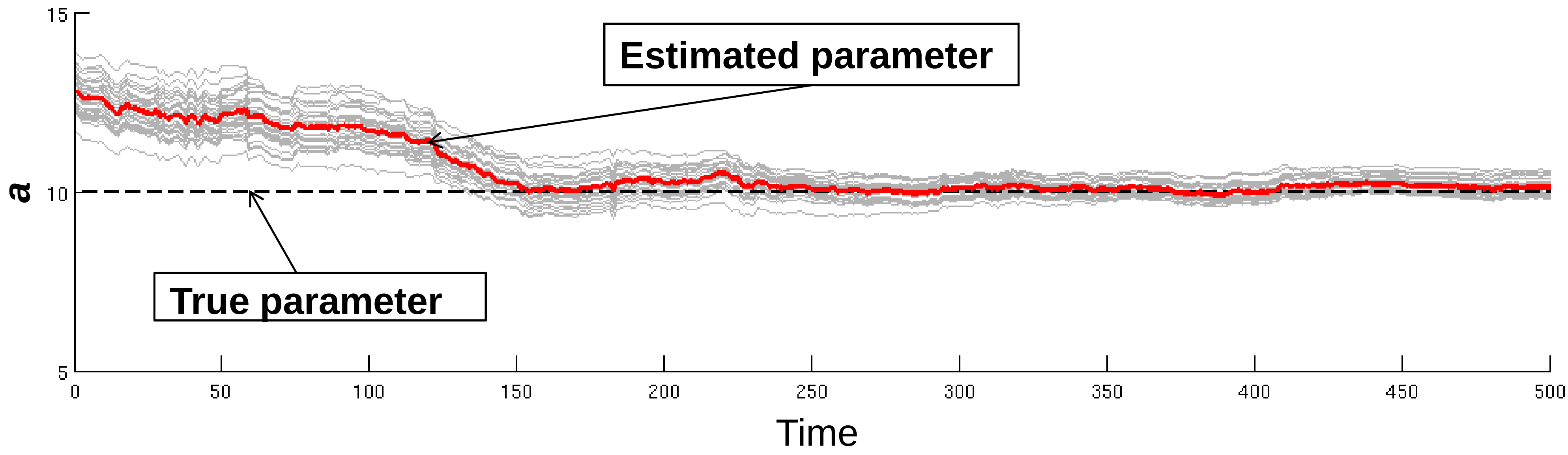
- ✓ The method captures also changes in the PDF of the model variables. Usually a large error in the estimated value of a variable is associated with a larger variance in the estimated PDF.

- ✓ These changes in the PDF are caused by the system dynamics. The PDF of the forecast errors is “state dependent”

Assimilation experiment 1: constant parameters

Estimated parameters

Time evolution of the estimated parameter *a*

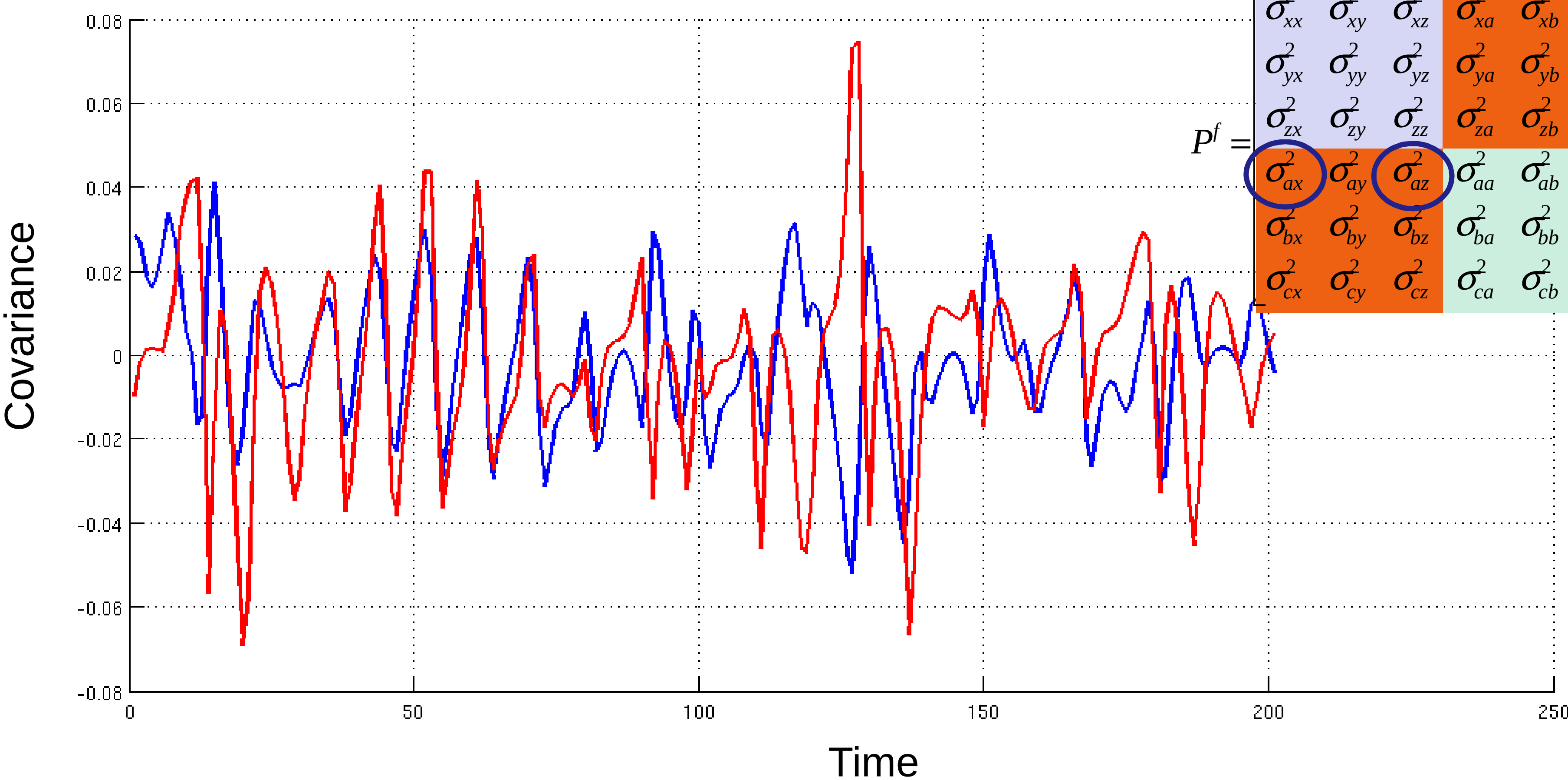


- At the beginning of the experiment a value of 13 is assumed for ***a*.** After several data assimilation cycles the value of the parameter ensemble mean (read line) converges towards the true value (black line).

- Observations provide enough information to find the optimal value for the parameters (similar results are obtained for the other two parameters).

Assimilation experiment 1: constant parameters

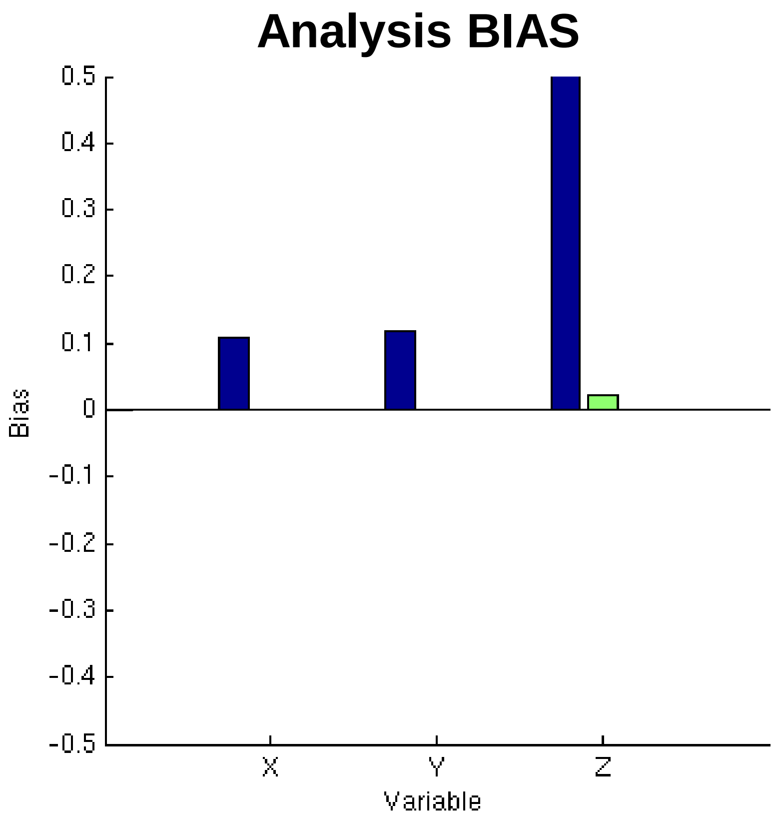
Covariance between *a* and x, and between *a* and z as a function of time

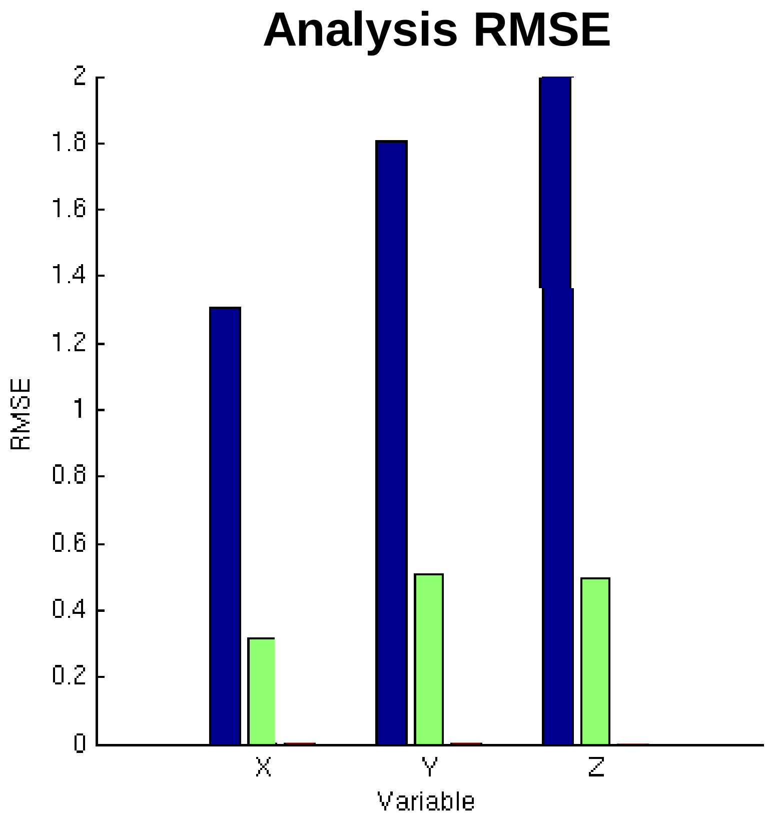


$$P^f = \begin{bmatrix} \sigma^2_{xx} & \sigma^2_{xy} & \sigma^2_{xz} & \sigma^2_{xa} & \sigma^2_{xb} & \sigma^2_{xc} \\ \sigma^2_{yx} & \sigma^2_{yy} & \sigma^2_{yz} & \sigma^2_{ya} & \sigma^2_{yb} & \sigma^2_{yc} \\ \sigma^2_{zx} & \sigma^2_{zy} & \sigma^2_{zz} & \sigma^2_{za} & \sigma^2_{zb} & \sigma^2_{zc} \\ \sigma^2_{ax} & \sigma^2_{ay} & \sigma^2_{az} & \sigma^2_{aa} & \sigma^2_{ab} & \sigma^2_{ac} \\ \sigma^2_{bx} & \sigma^2_{by} & \sigma^2_{bz} & \sigma^2_{ba} & \sigma^2_{bb} & \sigma^2_{bc} \\ \sigma^2_{cx} & \sigma^2_{cy} & \sigma^2_{cz} & \sigma^2_{ca} & \sigma^2_{cb} & \sigma^2_{cc} \end{bmatrix}$$

- ✓ Covariance between the model variables and the parameters are also state dependent. In this case the sign of the covariance can change depending on the state variables.
- ✓ The EnKF estimate these covariances from the ensemble.

Assimilation experiment 1: constant parameters

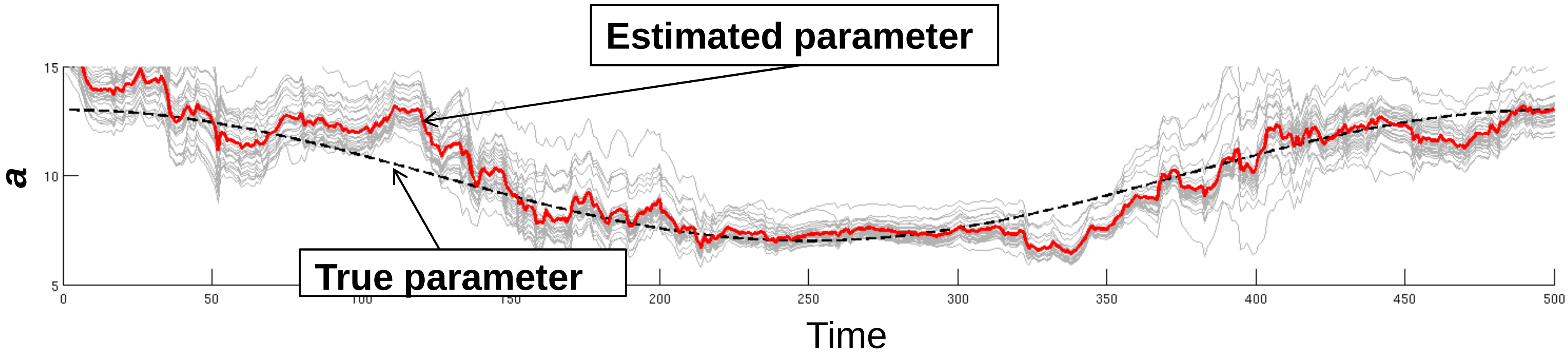




The optimization of model parameters improves the quality in the estimation of the model variables (x,y and z in this case).

Assimilation experiment 2: time-varying parameters

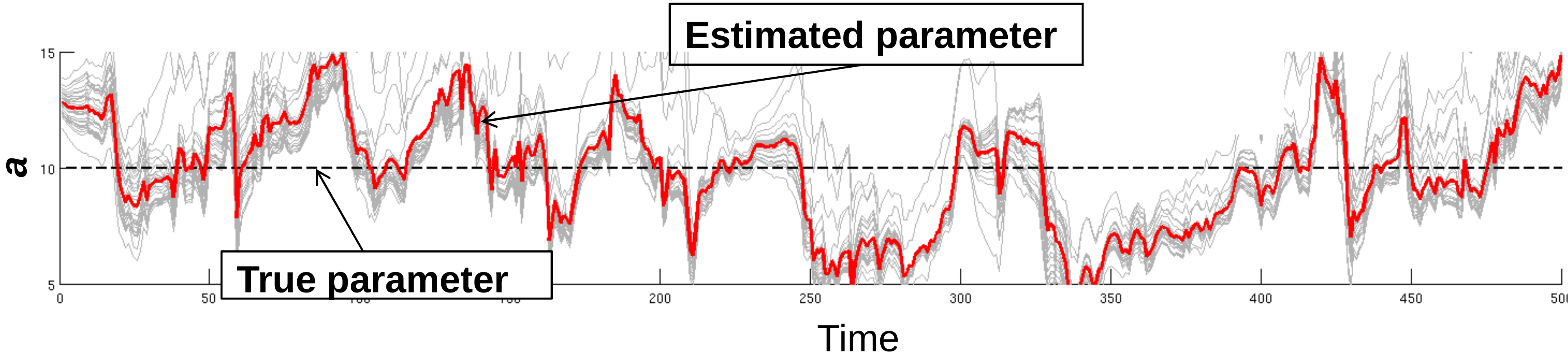
A new true evolution is generated assuming that the model parameters changes slowly with time.



- ✓ Estimated parameters can capture the time changes in the optimal parameters.
- ✓ In this case results are more noisy. We are still assuming that the parameter is constant with time during the model integration, but this is not a good assumption for this case.
- ✓ The uncertainty in the estimated parameter increases in this case. The adaptive inflation of Miyoshi (2011) is used in this case to estimate the uncertainty of the parameters.
- ✓ In most application of parameter estimation the parameters being estimated changes in space and time.

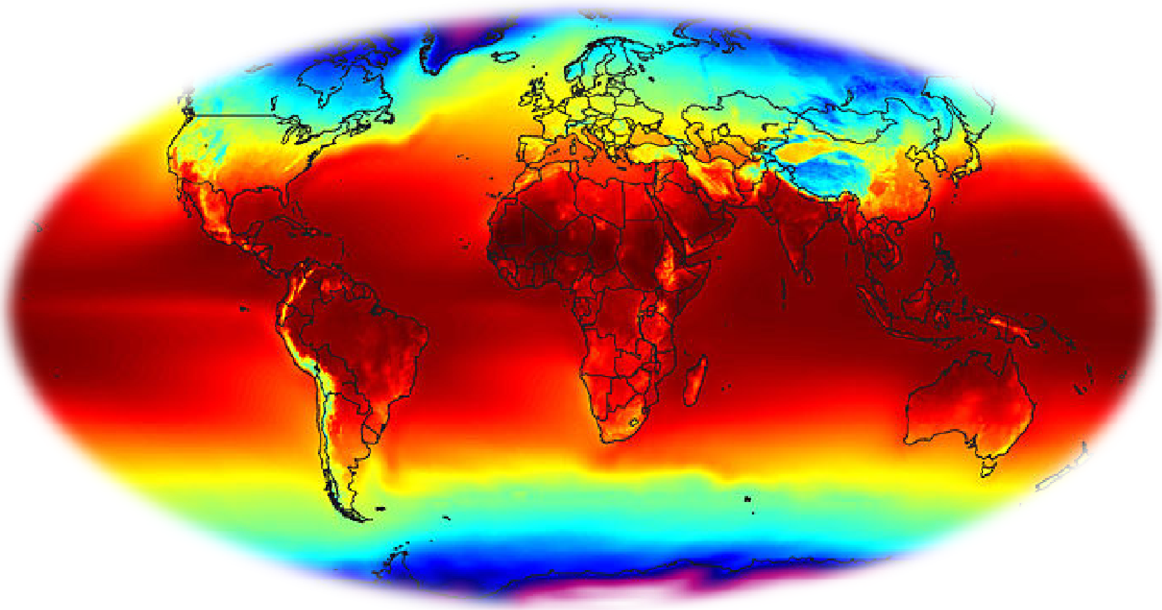
Assimilation experiment 3: Imperfect model

Only “***a***” parameter is estimated. ***b*** and ***c*** are fixed and they have errors.



- In this case the estimated parameter shows large oscillations around the true parameter.
- The estimated parameter is compensating errors associated with the other two parameters that are not being estimated.
- It is not possible for the method (as implemented in this case) to distinguish between errors associated with the estimated parameters and other sources of model error.
- The estimation of the model variables are still improved by the estimation of ***a***, even when it does not converge to its true value.

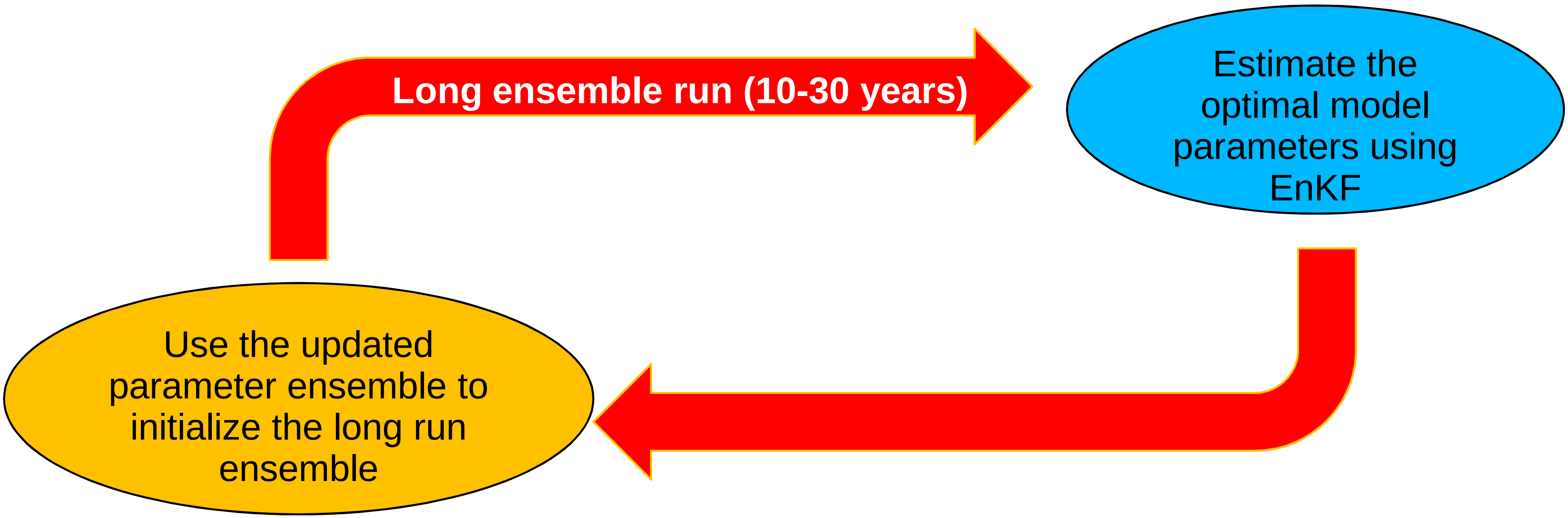
Parameter estimation based on the model climatology



Climate model optimization
(Annan and Hargreaves 2004)

Instead of assimilating observations every 1 time unit, build an ensemble of long runs.
Study the sensitivity of the climatology of the model to the model parameters.

Use the climatology of the observations instead of observations at particular times.



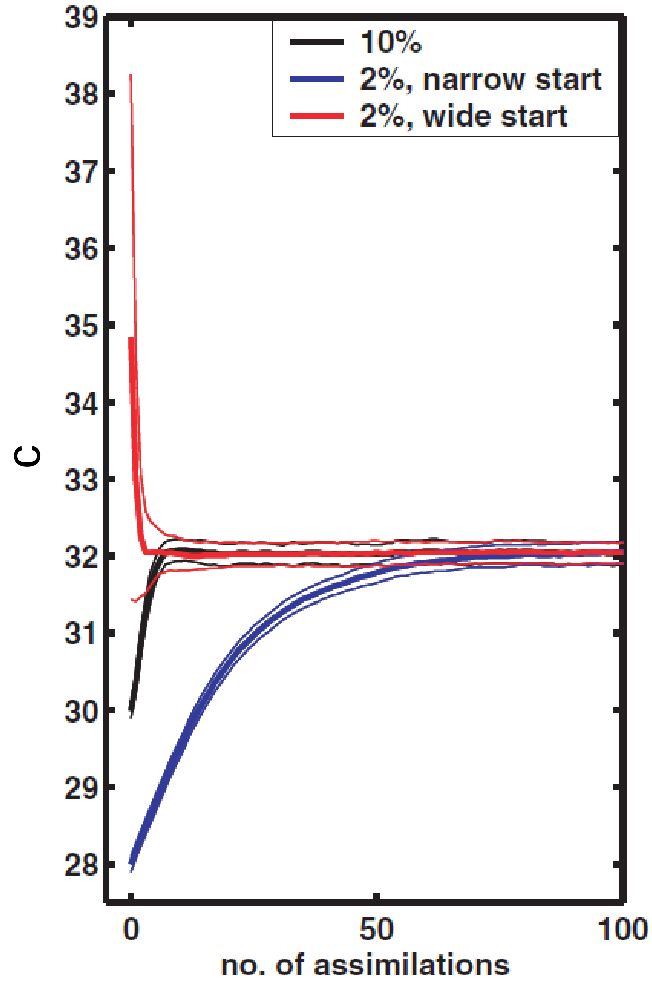
Several assimilation cycles are performed but in each cycle the same observations are used.
This may help to deal with non-linearities in the system?

Simple experiments using the Lorenz model are performed.

Climate model optimization
(Annan and Hargreaves 2004)

Estimated parameters as a function of the number of assimilation cycles.

Each assimilation cycle requires obtaining an ensemble of model climatologies with the updated values of the parameter ensemble.



The convergence rate is a function of the initial variance of the parameter ensemble. (red vs blue lines)

Climate model optimization

(Annan and Hargreaves 2004)

Non-linear sensitivity of the model to the parameters.
Joint cost function of *b* and *c* parameters.

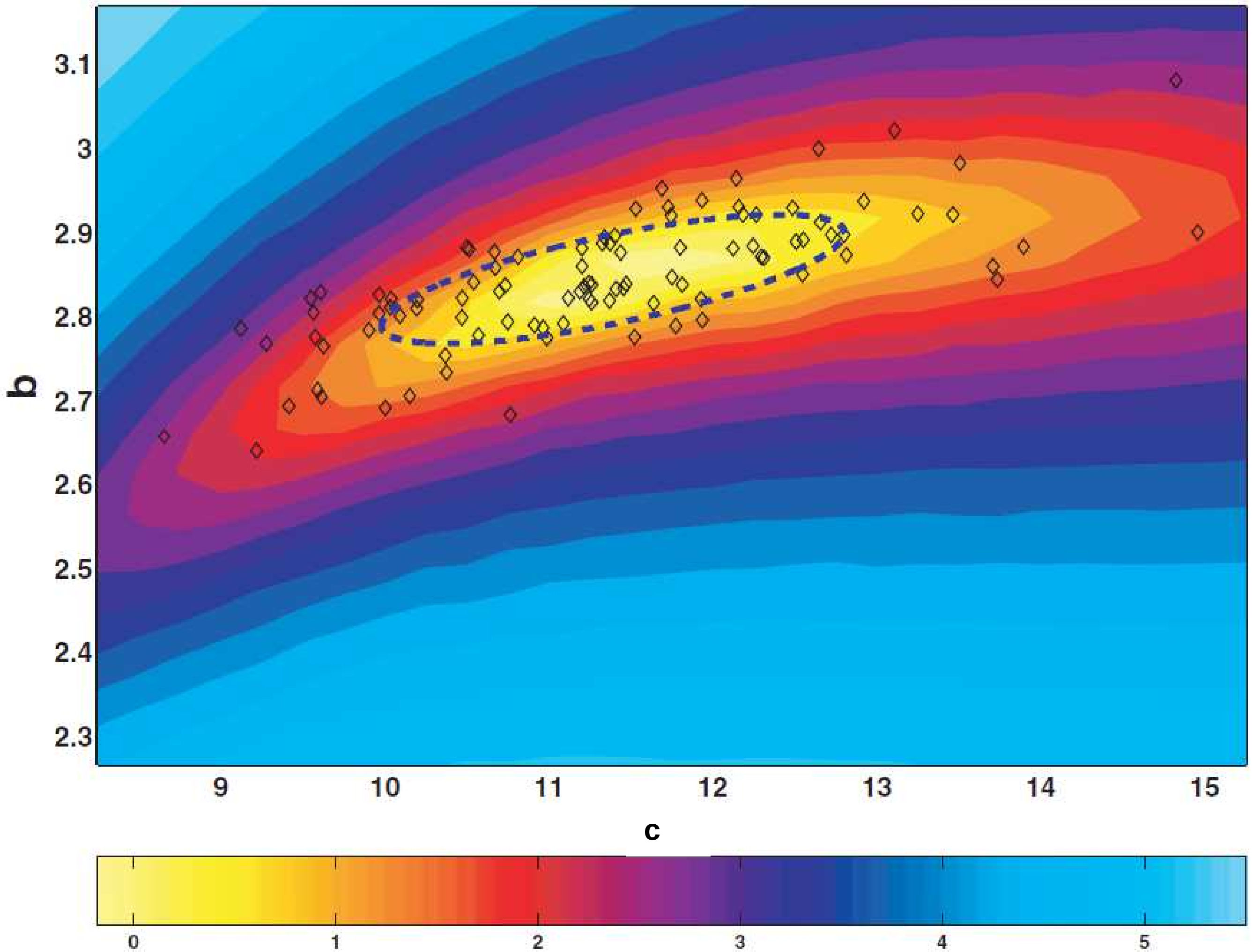


Fig 4. Cost function and ensemble results in the plane $r = 32$. The contours indicate the logarithm of the cost function, and the diamonds show the projection of the results of an EnKF experiment projected on to this plane. The ellipse indicates the one standard deviation width of the ensemble.

In this case the Gaussian approximation for the shape of the PDF is not so bad. But as can be seen in the figure the PDF of the uncertain parameters is not Gaussian.

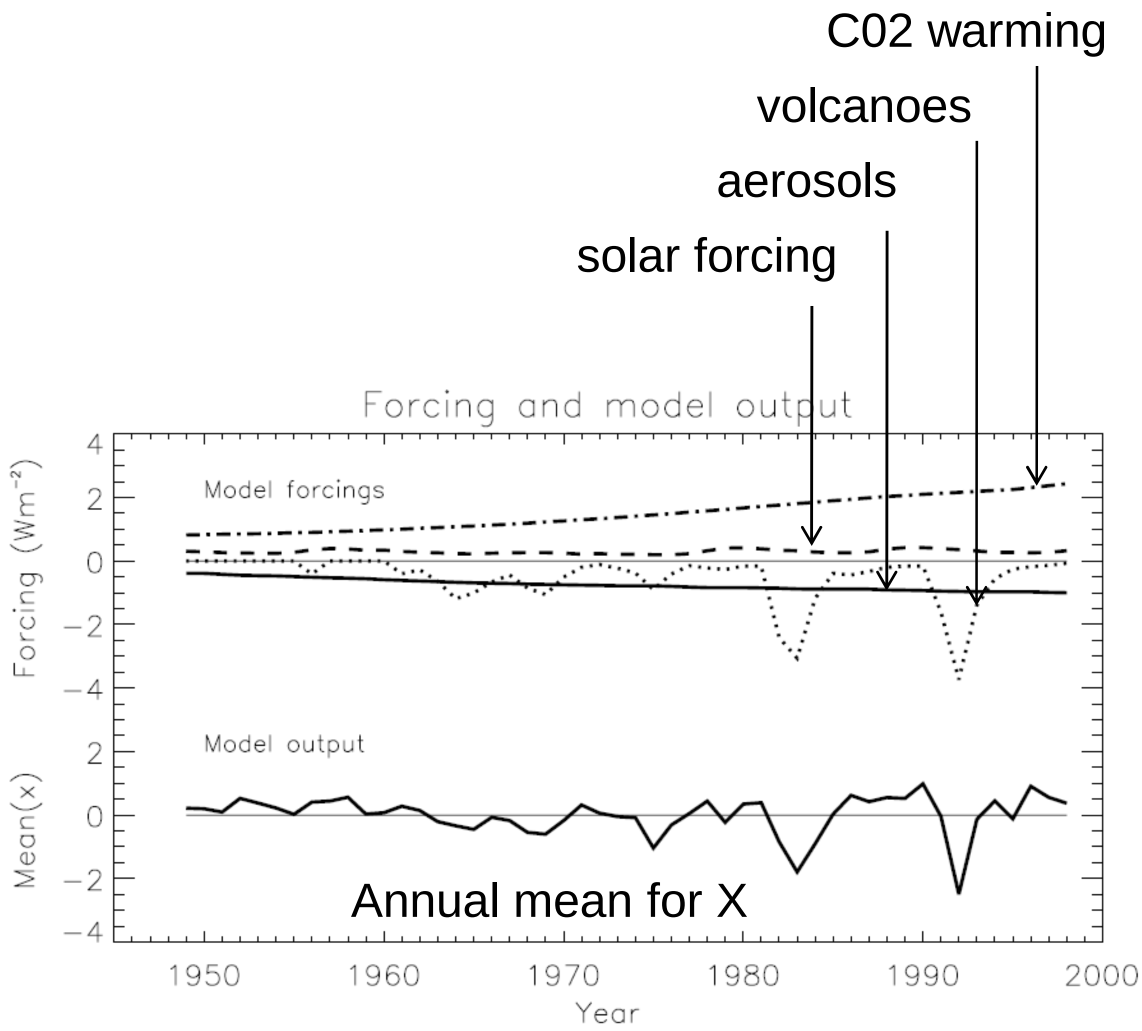
DADA kind experiment with the Lorenz model
(Annan 2005)

$$\frac{dx}{dt} = a(y - x) + f$$

$$\frac{dy}{dt} = bx - y - xz + f$$

$$\frac{dz}{dt} = xy - cz$$

Forcing terms were added to the Lorenz equations.

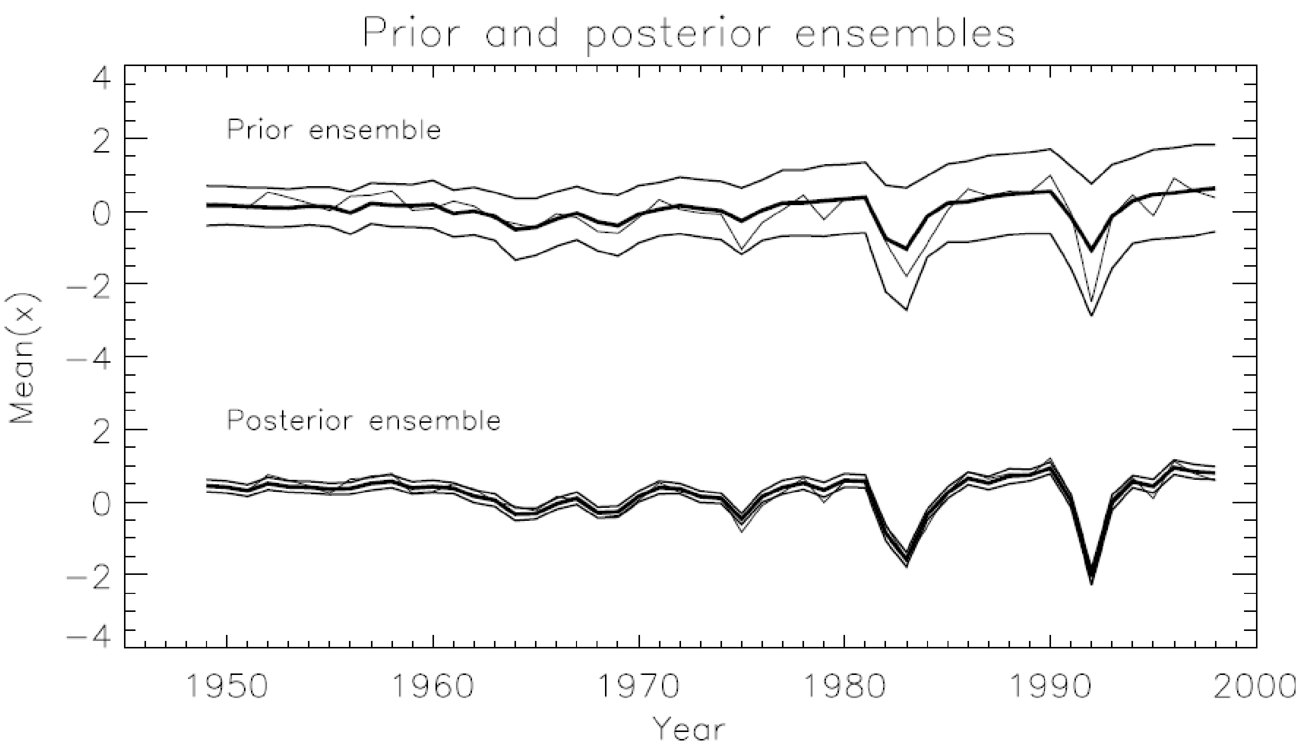


DADA kind experiment with the Lorenz model
Annan 2005

Time dependence of the forcing is assumed to be known

$$f = \alpha_{solar} f_{solar} + \alpha_{ghg} f_{ghg} + \alpha_{volcanoes} f_{volcanoes} + \alpha_{aerosols} f_{aerosols}$$

In this case the parameters α are estimated using the observations, but the shape of the forcing is assumed to be known a priori.



X evolution before the assimilation

X evolution after the assimilation

✓The total forcing is accurately reconstructed here.

✓In this simple model, the sensitivity of the climatology to the parameters is almost linear.

✓The issue of model error is not taken into account in this simple experiments.

Experiments with a simple GCM

- ✓ SPEEDY has a 48x92x7 grid and simple parameterizations (convection, pbl, soil model, radiation, large scale condensation)
- ✓ The EnKF method (LETKF Hunt et al. 2007) is used for the simultaneous estimation of model variables and parameters in a simple GCM (SPEEDY).
- ✓ Three parameters associated with the convective scheme are estimated (P1, P2 and P3). (Ruiz et al. 2012)

Twin experiments with the SPEEDY model

Ruiz et al 2012, Kang et al 2009, Fertig et al 2007, Miyoshi et al 2005.

Nature run.

- ✓ A three month run with the SPEEDY model using the original set of parameters is used as the nature run.
- ✓ Observations are generated taking the nature run values at every other grid point, at all vertical levels and every six hours.

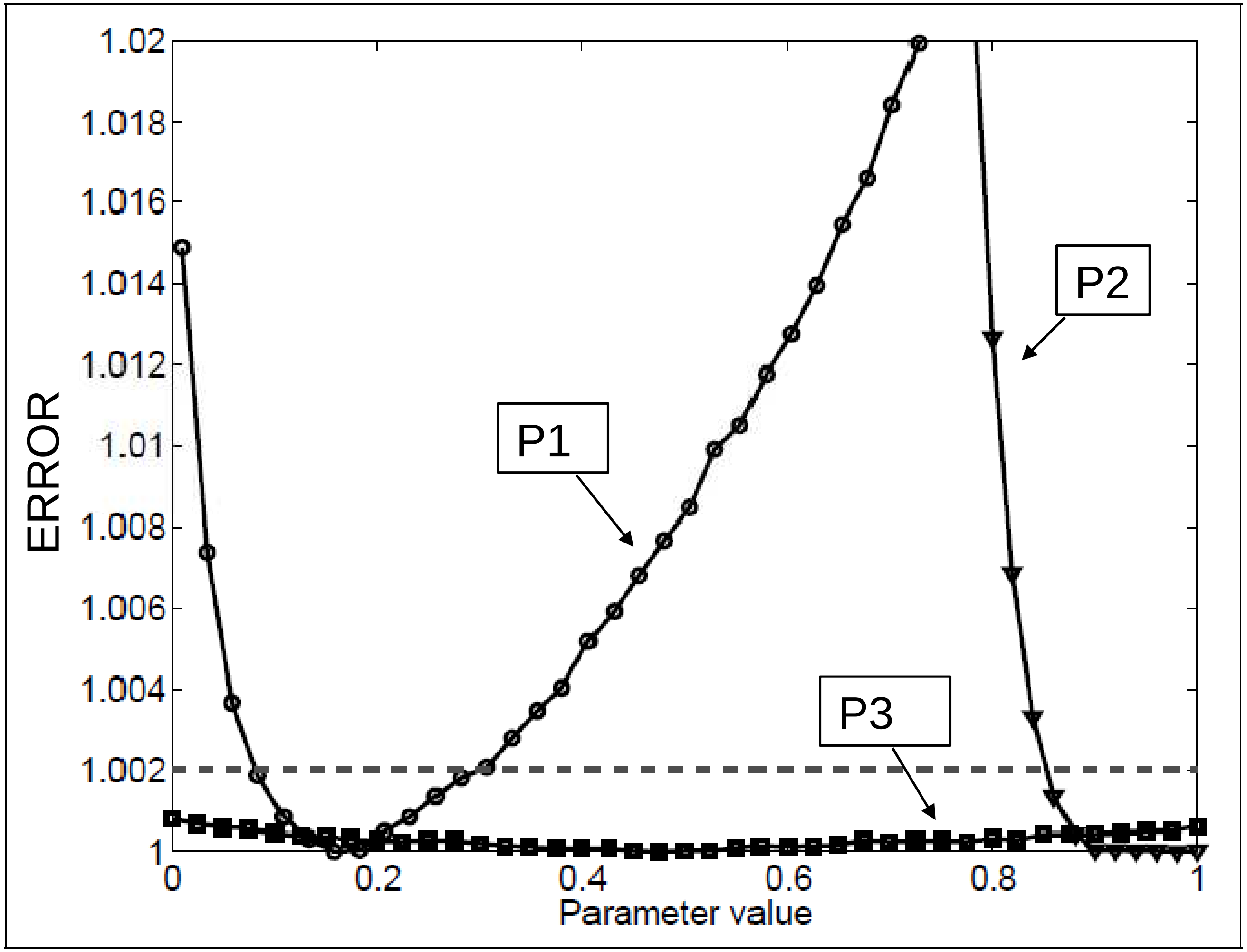
Data assimilation

- ✓ Observations are assimilated every six hours.
- ✓ The model used in the assimilation experiment may only differ from the true model in the value of the convective parameter scheme.

Experiments with a simple GCM

Ruiz et al. 2012

The success of parameter estimation depends on the sensitivity of the model to the estimated parameters. If the model is not sensitive to changes in these parameters, then parameter estimation will fail.

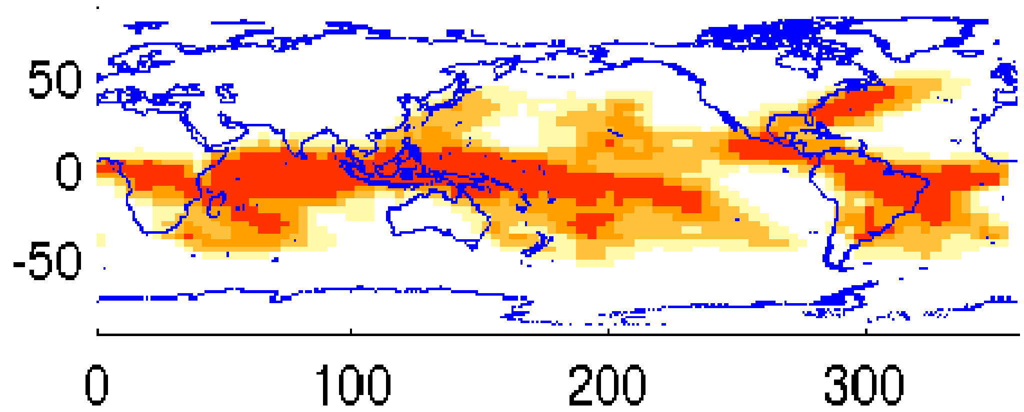


- 6 hr Forecast ERROR
- The larger the response, the smaller the uncertainty in the estimation of the parameter.

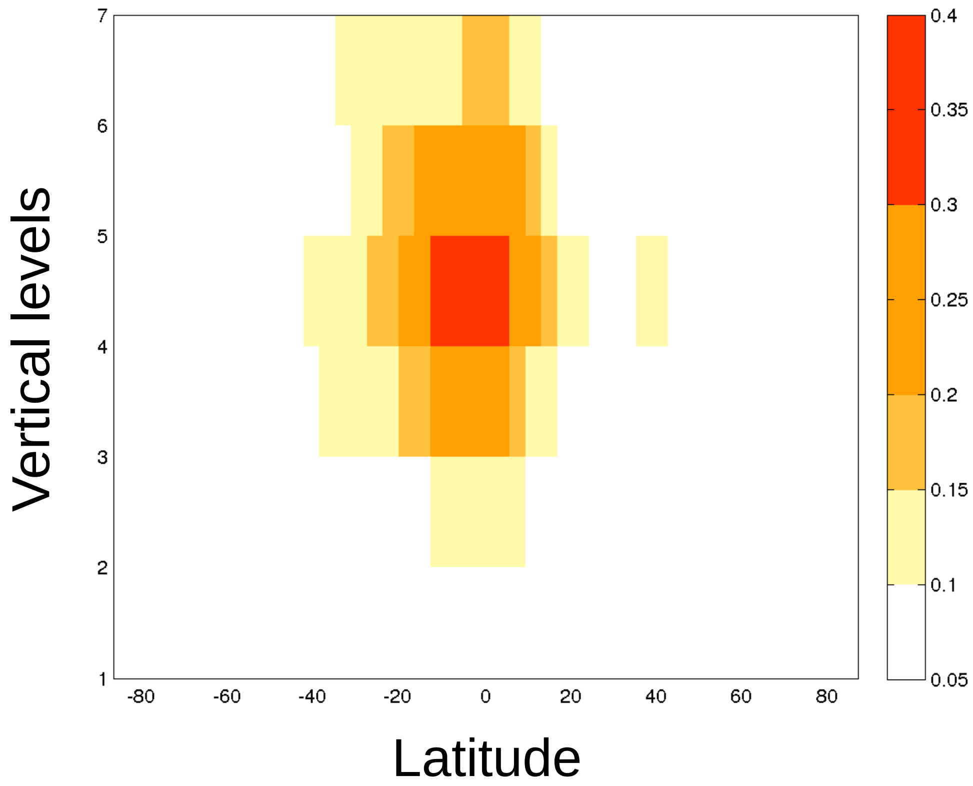
Experiments with a simple GCM

Ruiz et al. 2012

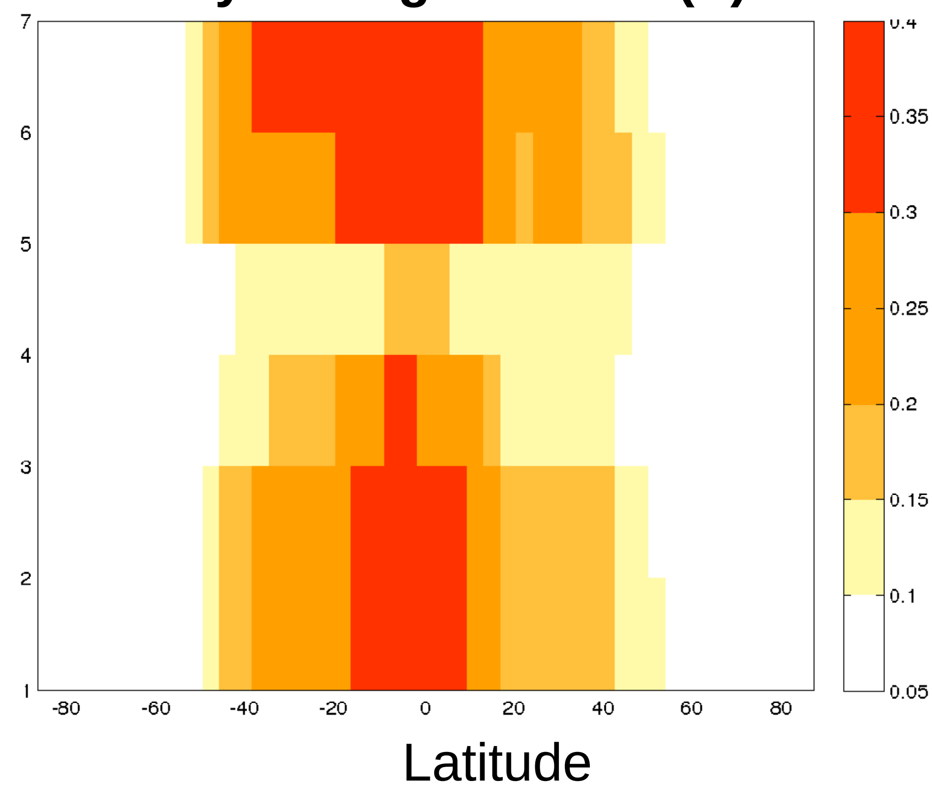
RMSE U at 200 hPa, P1



Zonally averaged RMSE (temp) P1



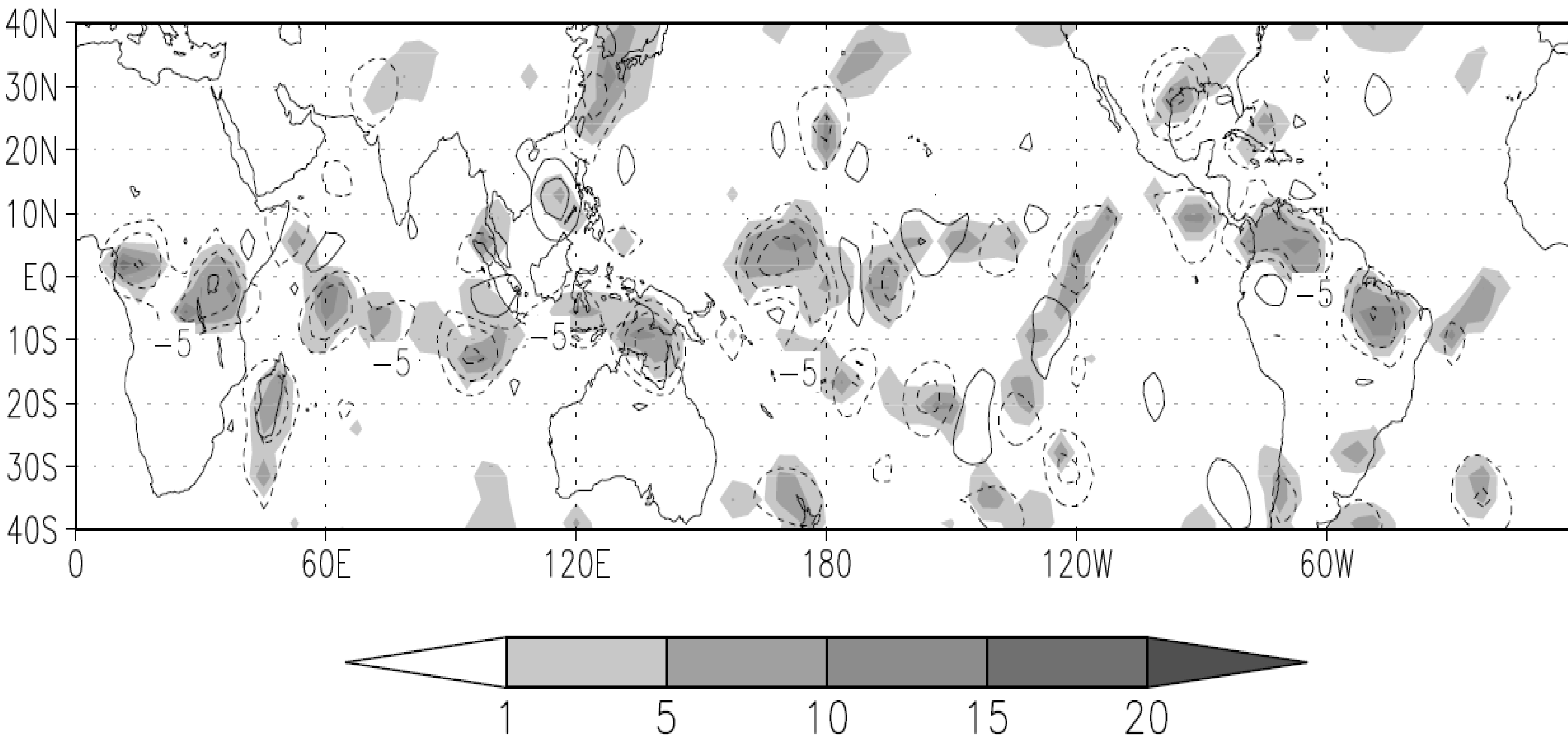
Zonally averaged RMSE (V) P1



Experiments with a simple GCM

Ruiz et al. 2012

Covariance between the **temperature** at mid levels and **P1** (contours) and forecasted precipitation (shaded).
Lower P1 means weaker convection which produces less warming and lower temperatures at mid levels.

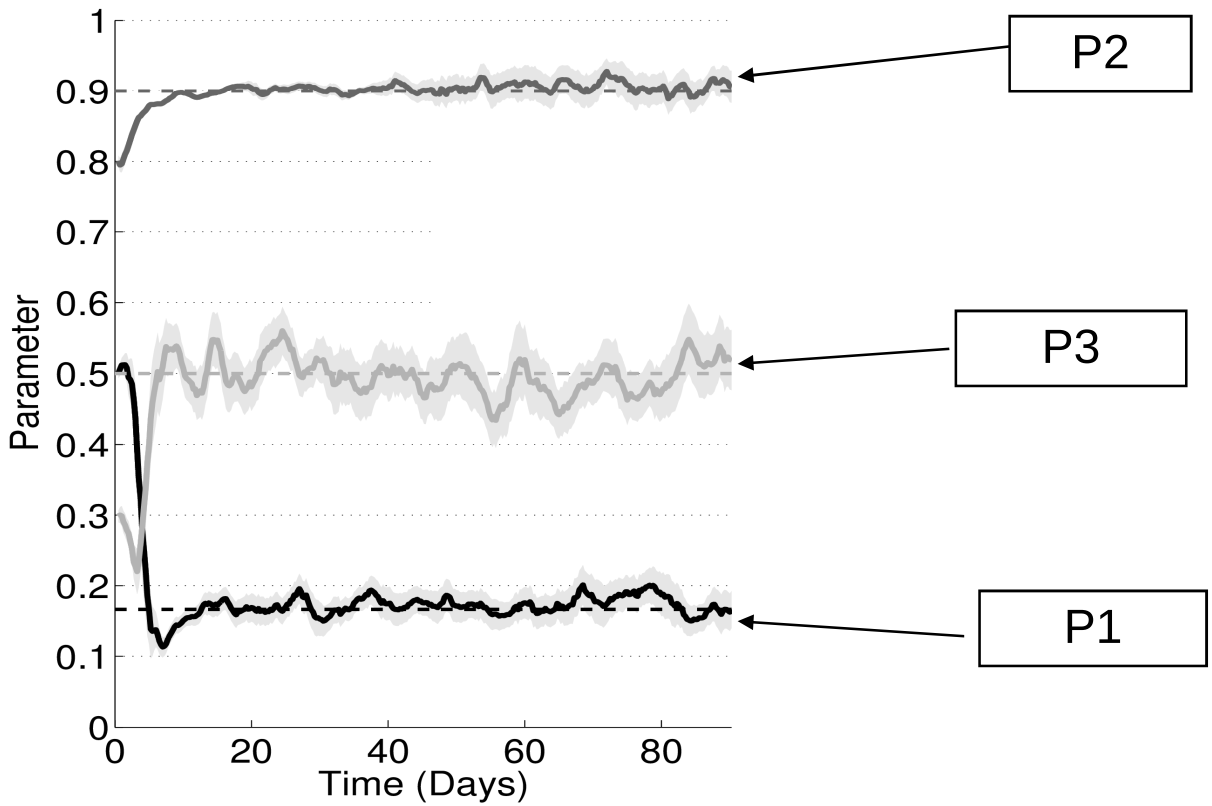


Covariance structure is strongly flow dependent because convection is intermitent in space and time.

Experiments with a simple GCM

Ruiz et al. 2012

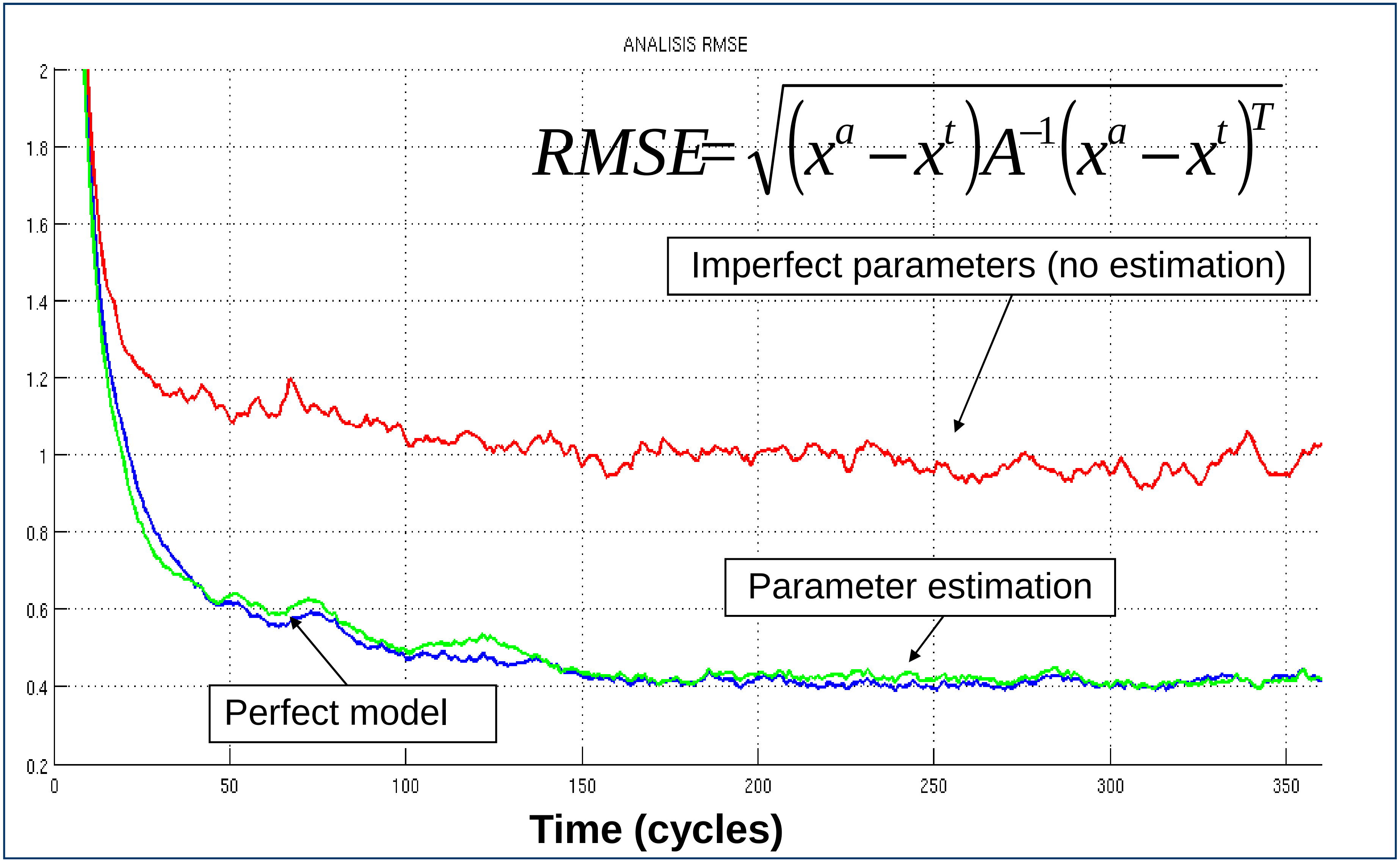
Estimated parameters as a function of time, and parameter uncertainty (gray shaded)



Convective schemes parameter are accurately estimated and the spin-up time is around 15 days (including the spin-up of the initial conditions).

Experiments with a simple GCM

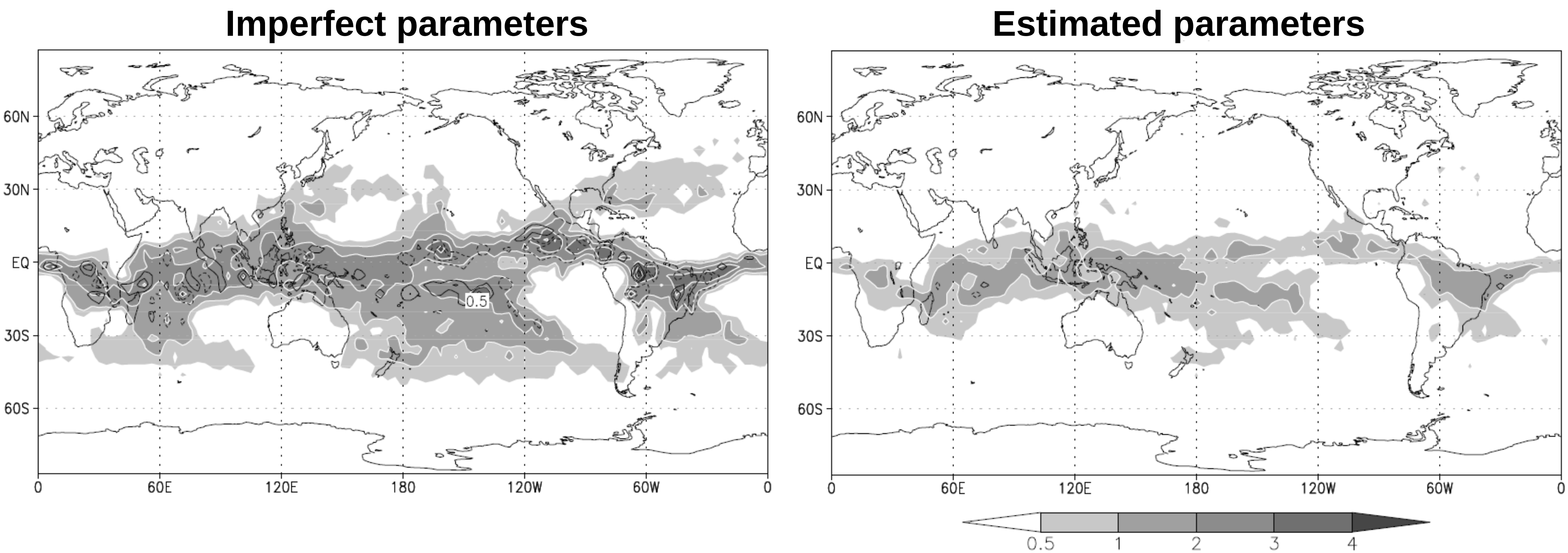
Ruiz et al. 2012



Parameter estimation produces a strong impact upon analysis error. In this twin experiment the parameter estimation experiment is as good as the perfect model.

Experiments with a simple GCM
Ruiz et al. 2012

RMSE of short range precipitation forecast (shaded) and its BIAS (contours)



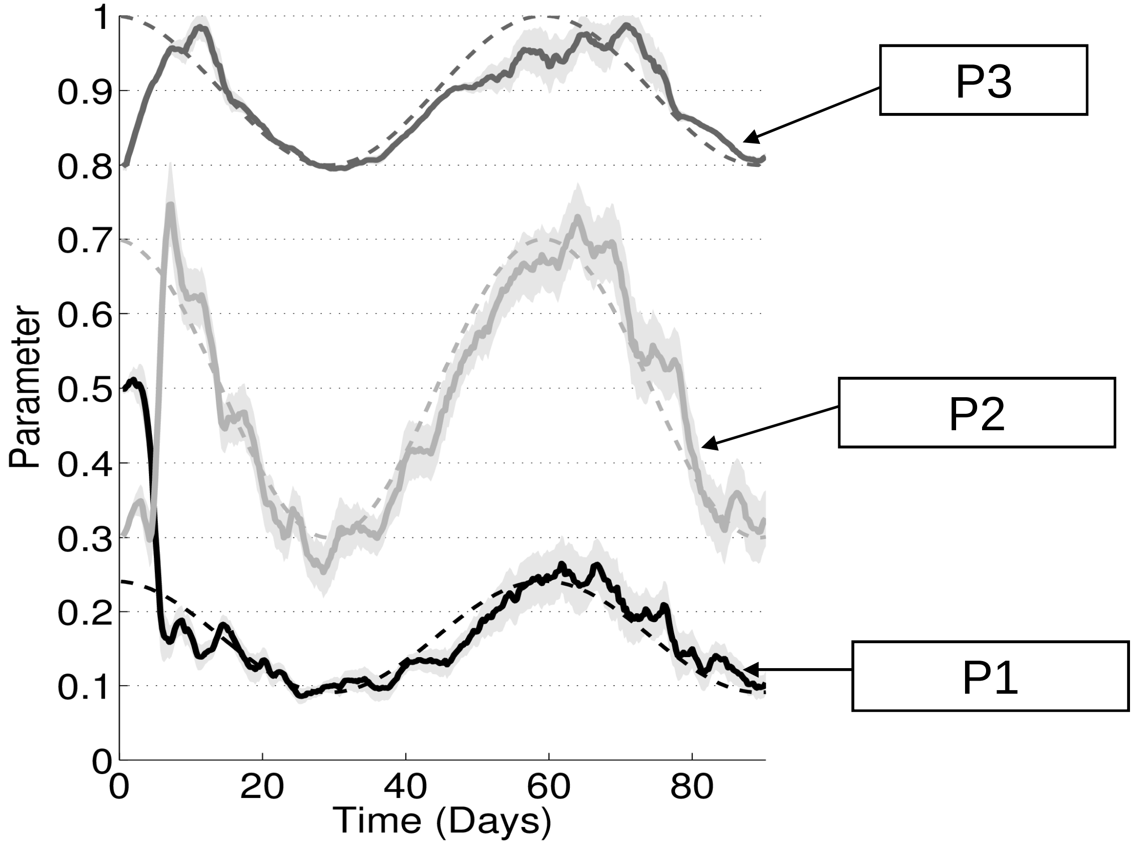
- ✓ Estimated parameters produce a positive impact upon the precipitation forecast. Most forecasted precipitation in the tropics is produced by the convective scheme.

- ✓ The increase in the accuracy of the estimated model variables (initial conditions for the forecast) also explains the increase in the skill of the precipitation forecast.

Experiments with a simple GCM

Ruiz et al. 2012

See also, Kang et al. (2012), Koyama and Watanabe (2010).
Time dependent parameters are included in the nature run.



- ✓ **Time dependent parameters are well estimated. However a small lag is present in the estimated parameters.**
- ✓ **Parameter evolution is not known a priori.**
- ✓ **This lag is due to the persistance model used in the parameter estimation.**

Parameter estimation in the presence of other sources of model error.

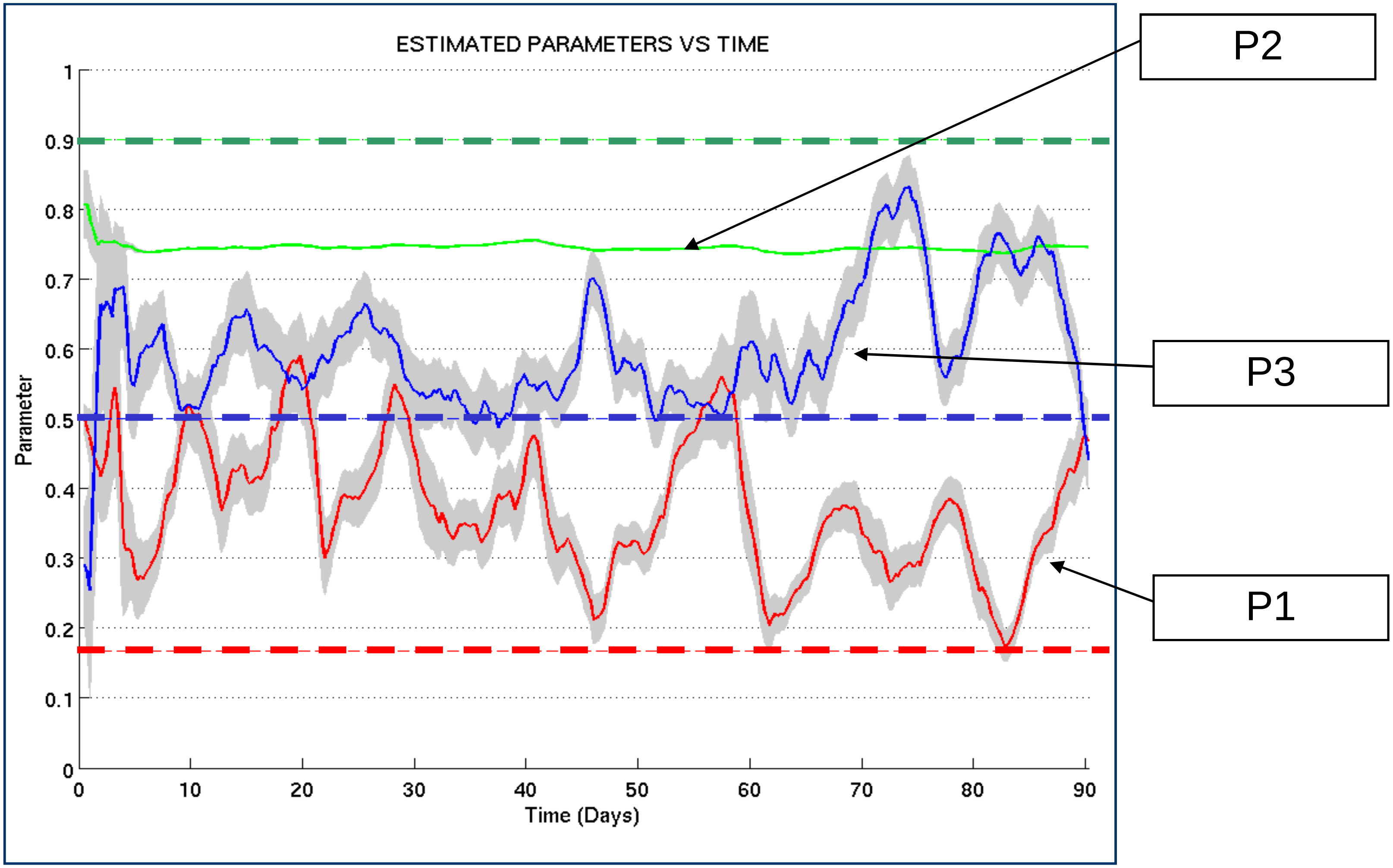
- In the experiments presented so far, the model was (almost) perfect.
- All model error is due to the uncertainty in the convective scheme parameters.
- What happens when those are not the only source of errors and when the model error cannot be completely corrected by tuning some model parameters (as in real applications)?
- How can parameter estimation be combined with other methods that include a representation of model error in the ensemble Kalman filter?

Imperfect model experiments:

Other sources of model imperfection are introduced.

- ✓ To simulate other sources of model imperfection the value of some parameters (numerical diffusion, surface exchange coefficients) are modified.
- ✓ Convective scheme parameter are the only ones being estimated in the parameter estimation experiments.
- ✓ All other settings are as in the perfect model experiment.

Parameter estimation in the presence of other model error.



In the presence of other sources of model error, the estimated parameters do not converge to the true parameter value.

Estimated parameters shows an increased variability in time.

Conclusions.

Parameter estimation based on data assimilation is a very promising tool for objectively tuning the model. It can partially correct model error.
Implementation of parameter estimation in most state-of-the-art data assimilation systems is straightforward and produces almost no extra computational cost.

Some issues:

- ✓How to estimate the parameter uncertainty? This is very important to have good results in parameter estimation.

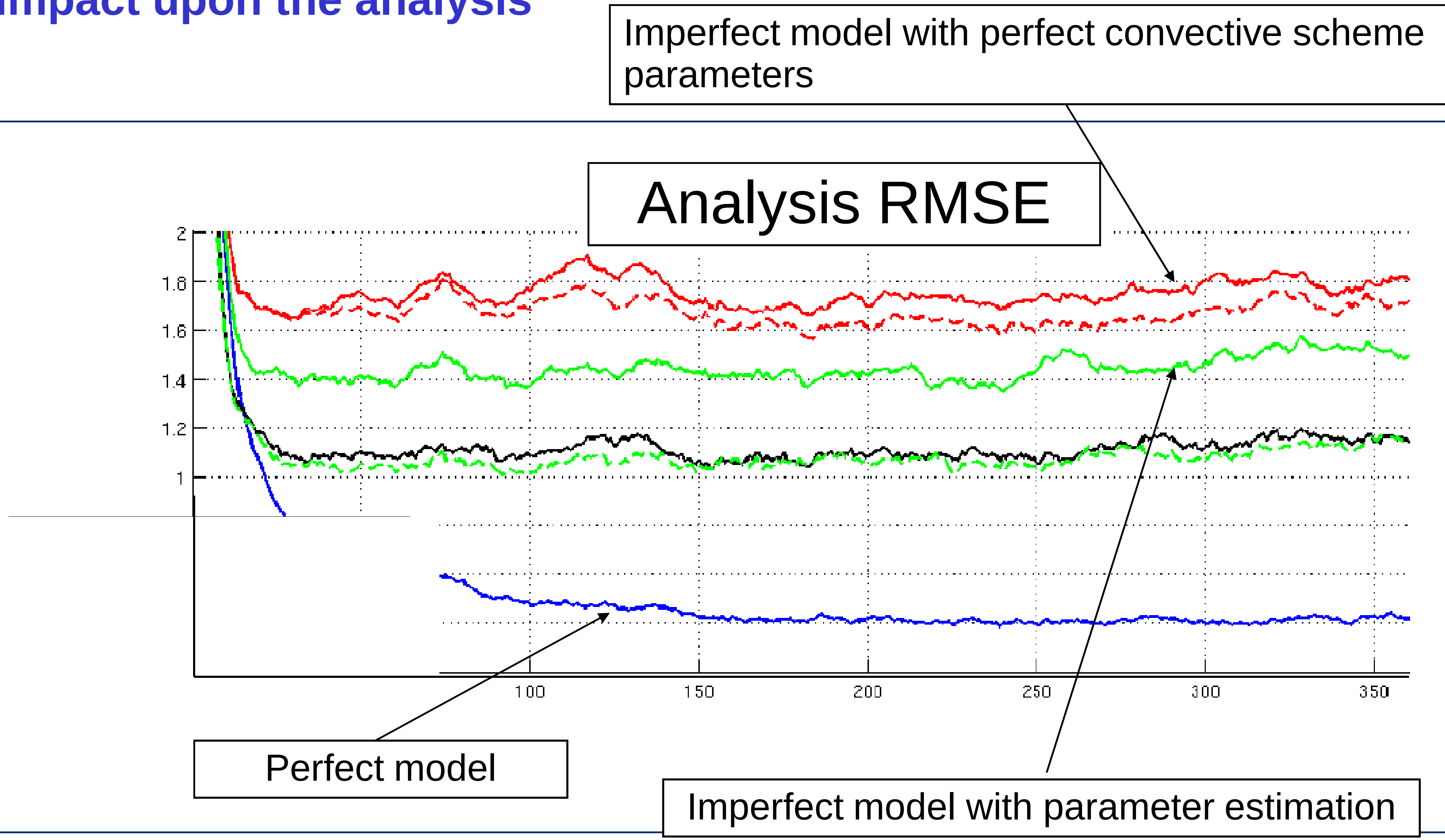
- ✓How to deal with parameters that are not constant in time.

- ✓Non-linear sensitivity of the model to the parameters. KF assumes Gaussian PDF, multimodal or asymmetric PDF that results from non-linear sensitivities are not well represented in this framework. (Posselt and Bishop 2012). Parameter estimation usually increases the non-linearity of the model.

- ✓Physical interpretation of the estimated parameters in the presence of other sources of model error. Can this effect be quantified?

- ✓Some applications will suffer from computational issues. Climate model tuning may be too expensive. (Tune the climate model as a short range numerical weather prediction model?).

Impact upon the analysis



Estimated parameters produced a lower analysis RMSE than the true convective scheme parameters.

Parameter estimation can improve the analysis even in the presence of other sources of model error.

Combine parameter estimation with adaptive inflation and additive inflation.

•Additive inflation (applied to the state variables only):

Random samples of “true” model error are used as perturbations for the additive inflation approach. This can compensate for low ensemble spread associated with model error and can modify the structure of the initial conditions perturbations.

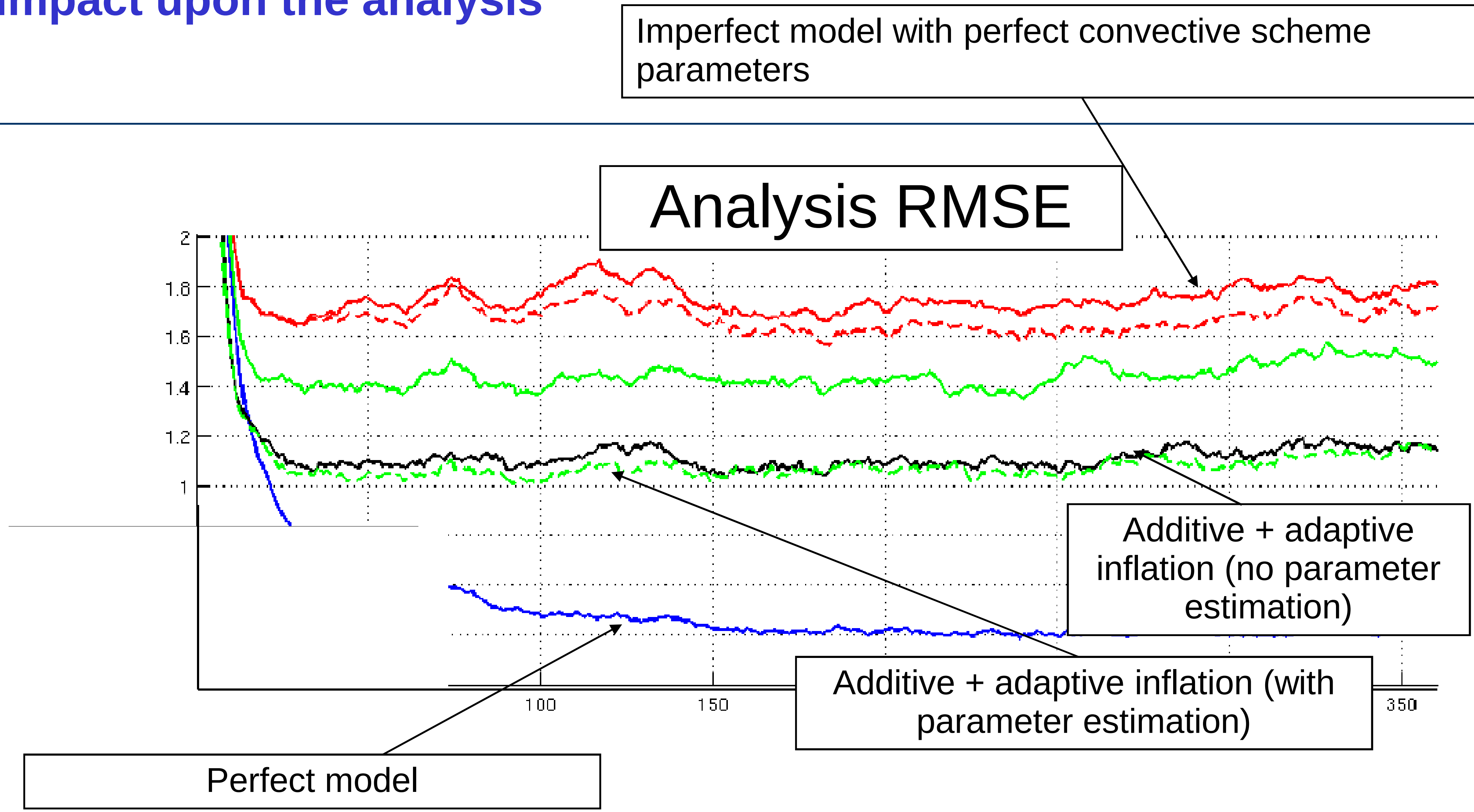
(Li et al 2009)

•Adaptive multplicative inflation (applied to the state variables only):

Adaptive multiplicative inflation uses the observations to find the optimal inflation level (Li et al, 2009, Miyoshi 2011). This can compensate low ensemble spread due to error sources not considered in the ensemble formulation.

Kang 2009 showed that multiplicative inflation applied to the state variables has a possitive impact upon the estimation of the parameters.

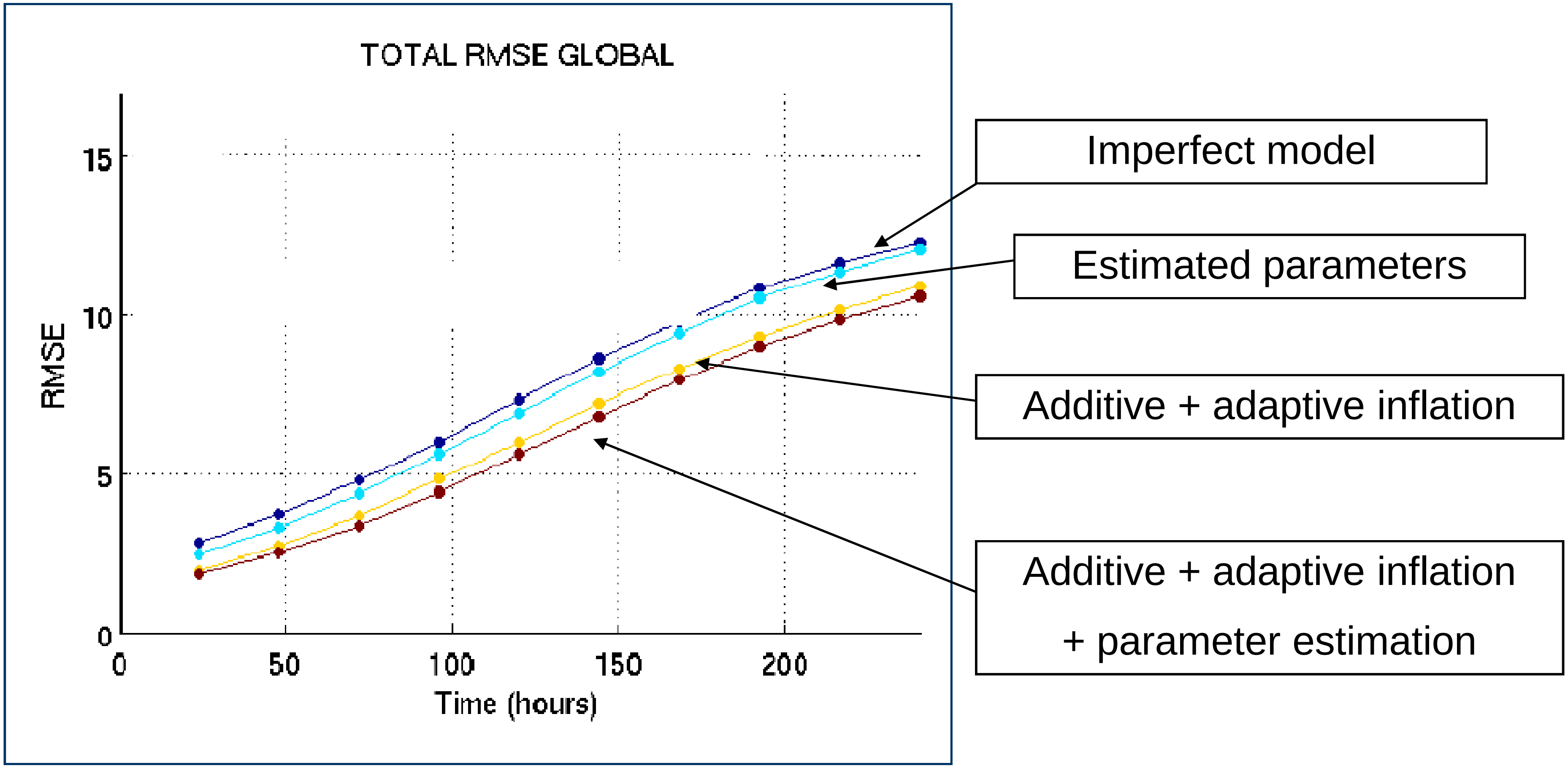
Impact upon the analysis



Additive and adaptive inflation produces a larger impact than parameter estimation alone.

Combining parameter estimation with these techniques produces an small improvement of the analysis.

Impact upon the short to medium range forecast: Imperfect model.



Eventhough the estimated parameters do not converge to the true parameters, they produce an improvement of the forecast skill.