Parameter uncertainty of chaotic systems

H. Haario[1,2], with people from UH, LUT, FMI, ECMWF,
University of Montana, US

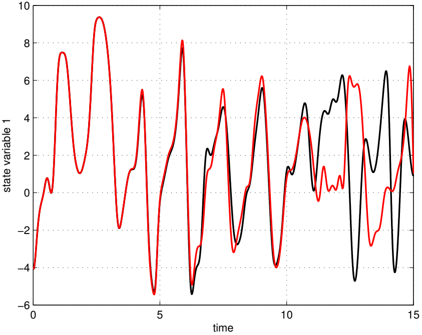
April 13, 2016

[1] Lappeenranta University of Technology (LUT), Finland
[2] Finnish Meteorological Institute (FMI)

Contents

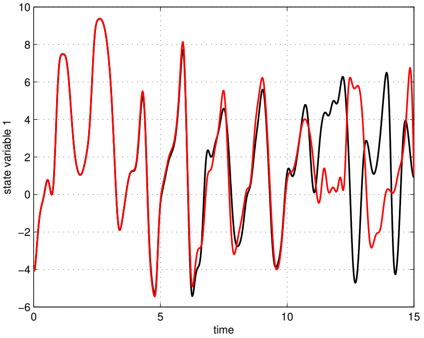
- Parameter estimation for chaotic dynamic systems:
  - Short-time integrations: from ensemble UQ to parameter optimization
  - Long-time integrations: naive summary statistics vs fractal concepts
- Examples, from 3D chaos to SWE, operational weather and climate models

Chaotic Systems



After a predictable interval, any changes (of initial values, model parameters, solver settings) lead to unpredictable deviations.
Options:

Chaotic Systems



After a predictable interval, any changes (of initial values, model parameters, solver settings) lead to unpredictable deviations.
Options:

- **Avoid chaos**: deal with predictable time intervals only (Weather)
- **Face it**: deal with behaviour after predictability (Climate)

Short time simulations: parameter estimation by monitoring operational ensemble systems

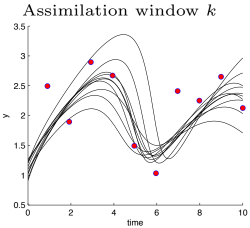
- Estimate numerical weather prediction (NWP) model parameters on-line, employing operational EPS runs.
- EPS (ensemble prediction system): an ensemble of predictions is run, to estimate the uncertainty of prediction.
- Combine parameter optimization and EPS: only monitor the results of EPS, no new model simulations added, **no additional CPU needed!**

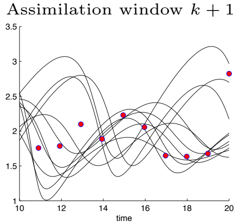
- Laine M, Solonen A, Haario H, Järvinen H. *Ensemble prediction and parameter estimation system: the method*. Q. J. R. Meteorol. Soc. Vol. 138, nro. 663, pp. 289-297, 2012.

The EPPES concept

Assume some assimilation method used, to get initial values for each assimilation window.

- EPS: Ensembles of simulations by initial/model perturbations.
- EPPES: model parameters θ are additionally perturbed. Sampled from an adapted Gaussian proposal distribution.
- The parameters are weighted by importance sampling by a cost function that depends on forecast skill.





Cost function

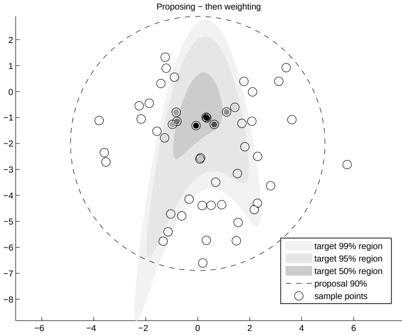
Cost function for model F, observations y_k, initial values x_k, and parameter θ_k for forecast time window k:

$$(y_k - F(x_k;\theta_k))' \, \Sigma_{\text{obs}}^{-1} \, (y_k - F(x_k;\theta_k)) \\ + (\theta_k - \mu)' \, \Sigma^{-1} \, (\theta_k - \mu)$$

We weight local parameters θ_k by respective performance, and update the global hyper parameters μ and Σ that give the proposal distribution for θ_{k+1}

Proposal covariance updates by importance sampling

1. Sample $\widetilde{\theta}_k^{(j)}$ from a Gaussian distribution q, $j = 1, \ldots, n_{\text{ens}}$
2. Resample $\widetilde{\theta}_k^{(j)}$ with weights $w_j \propto \frac{p(\widetilde{\theta}_k^{(j)} | y_k)}{q(\widetilde{\theta}_k^{(j)})}$ to produce $\theta_k^{(j)}$.
3. For the next stage $k+1$, update the proposal distribution $N(\mu_k, \Sigma_k)$. Go back to step 1.

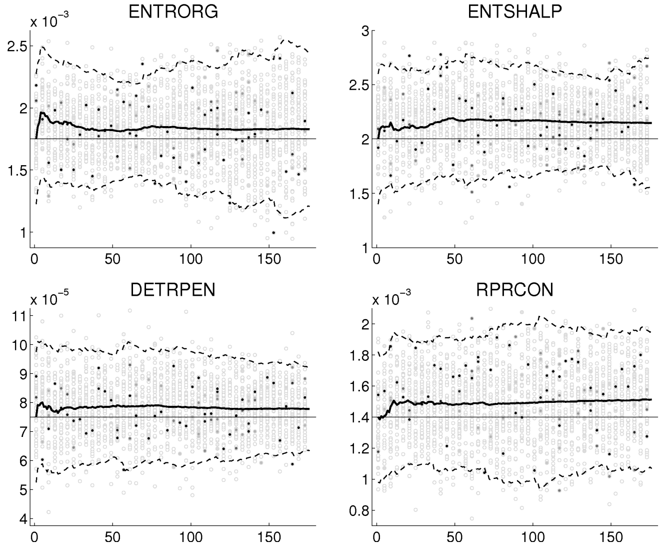


Large scale NWP models

- After toy model experiments (Lorenz 95) the method tested by running ECHAM 5 climate model in NWP mode utilizing ECMWF initial values.
- ECMWF implementation (IFS, code update CY38R1), by FMI and ECMWF people.
- Tune model parameters by training data and validate by independent forecast runs, using the RMSE and ACC as cost function. SPPT used in addition to initial value perturbation.
- Selected model fields used as criteria;
  - z500, the 500 hPa geopotential height
  - Total energy norm

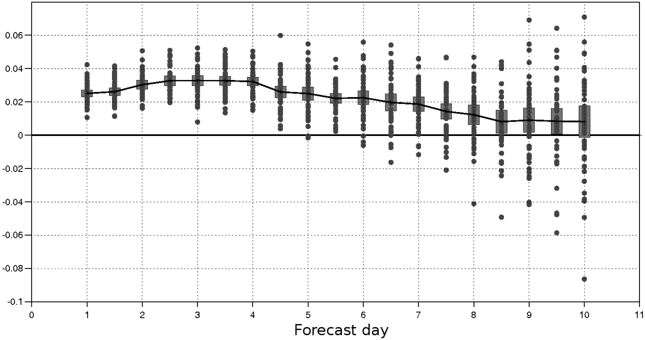
- Ollinaho, P., Bechtold, P., Leutbecher, M., Laine, M., Solonen, A., Haario, H., and Järvinen, H.: *Parameter variations in prediction skill optimization at ECMWF*, Nonlin. Processes Geophys., 20, 6,1001-1010, 2013.
- Ollinaho, P., Järvinen, H., Bauer, P., Laine, M., Bechtold, P., Susiluoto, J., and Haario, H.: *Optimization of NWP model closure parameters using total energy norm of forecast error as a target*, Geosci. Model Dev., 7, 1889-1900, doi:10.5194/gmd-7-1889-2014, 2014.

Results: IFS performance



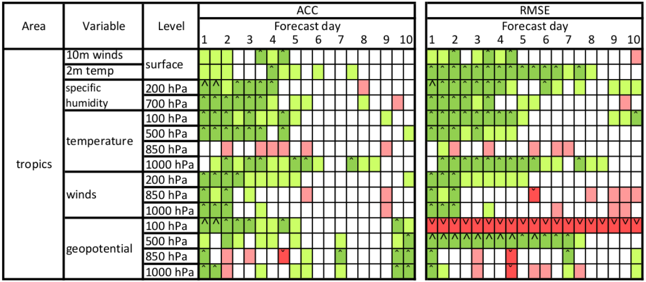
Parameter evolutions as functions of assimilation steps

Results: IFS performance



Performance of the default (already highly tuned IFS model !) and optimised model, in relative RMSE values.

Results: IFS performance



Score Cards of various other performance criteria. Green: improved, **Red: deteriorated.**

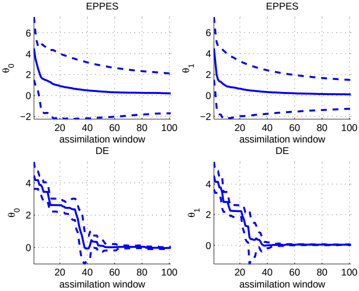
Optimizing by evolution algorithms

Need for multicriteria optimization, for higher number of unknowns, with faster convergence.

- Interprete each ensemble in an assimilation window as a 'generation' of a population for a genetic optimisation algorithm.
- We employ the Differential Evolution algorithm
  - Mutation: add scaled differences of ensemble vectors to the present ones
  - Crossover: random survival of mutations
  - Selection: survival of improvements
- The EPPES framework leads to a multicriteria stochastic optimization problem, certain modifications needed for DE to maintain population diversity.

Optimizing by DE

- Faster convergence than basic EPPES
- Multicriteria by products of the EPPES importance weights: all selected criteria required to improve.



Optimizing by DE/EPPES

Next:

- Multicriteria test runs with SWE, openIFS
- Implement for limited area models (HARMONIE ?)
- Combine EPPES-style optimization with stochastically perturbed parametrizations discussed here, to include multivariate correlations between parameters?

Long time simulations: summary statistics

The aim: characterise the distribution of model parameters that produce the 'same' known long-time behaviour. Ideally, by Monte Carlo (MCMC) sampling of a statistical likelihood cost function.

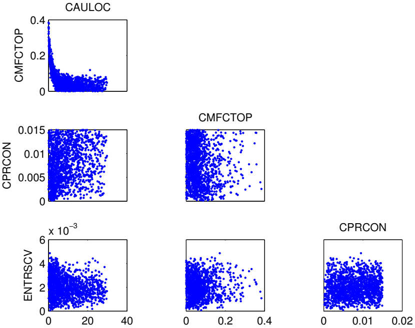
- Observations and simulations are averaged in space and time to create 'summary statistics'.
- If the statistics of the summary expression is known, a likelihood is formulated which yields the posterior for the model parameters.

Long time simulations: summary statistics

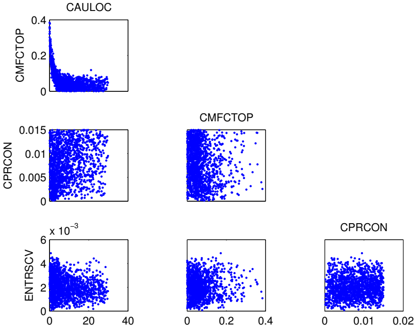
The aim: characterise the distribution of model parameters that produce the 'same' known long-time behaviour. Ideally, by Monte Carlo (MCMC) sampling of a statistical likelihood cost function.

- Observations and simulations are averaged in space and time to create 'summary statistics'.
- If the statistics of the summary expression is known, a likelihood is formulated which yields the posterior for the model parameters.
- Example: the approach was implemented for the ECHAM5 climate model, using likelihoods based on monthly global and zonal net radiation averages.
- MCMC was used to estimate four parameters related to cloud formation and precipitation. Technically possibly but...

Climate model MCMC results



Climate model MCMC results



- Direct, naive summary statistics (projections) do not identify the parameters, i.e., characterise the simulated trajectories.

- Järvinen, H., Räisänen, P., Laine, M., Tamminen, J., Ilin, A., Oja, E., Solonen, A., and Haario, H.: *Estimation of ECHAM5 climate model closure parameters with adaptive MCMC*, Atmos. Chem. Phys., Vol. 10, nro. 2, 9993-10002, 2010.

Distance between attractors, based on fractal concepts

In chaotic dynamics a fixed model parameter corresponds to different trajectories, depending on slightly different initial conditions, solver settings etc. But they all give samples of the same underlying attractor.

- We want to separate the 'internal' model variability due to initial values etc, but with fixed model parameters, from that due to different model parameters.
- We modify the concept of Correlation Dimension: from fractal dimension estimation to a statistical distance concept between attractors.
- Get a Gaussian likelihood ('by CLT') for the 'internal' variability.

Heikki Haario, Leonid Kalachev, Janne Hakkarainen *Generalized Correlation integral vectors: A distance concept for chaotic dynamical systems*. Chaos, 25, 2015

Numerical estimation of Correlation Dimension

In numerical practice, we have a finite time interval $[0, T]$ the trajectory vector s_i is evaluated on a finite number of time instants $t_i, i = 1, 2, ..., N$. For $R > 0$ set

$$C(R, N) = 1/N^2 \sum_{i,j} \#(||s_i - s_j|| < R)$$

and define then the correlation integral as the limit $C(R) = \lim_{N\to\infty} C(R, N)$. So we take the total number of points closer than R, normalize by the number of pairs N^2 and take the limit. Note that for each N we have $1/N \leq C(R, N) \leq 1$.

If ν is the dimension of the trajectory, we should have

$$C(R) \sim R^\nu$$

and the Correlation Dimension ν is defined as the limit

$$\nu = \lim_{R\to 0} \log C(R)/\log(R).$$

Distance via a generalized correlation sums

Modify the definition to get a measure for the distance between **two** model trajectories, as given, e.g., with different model parameters:

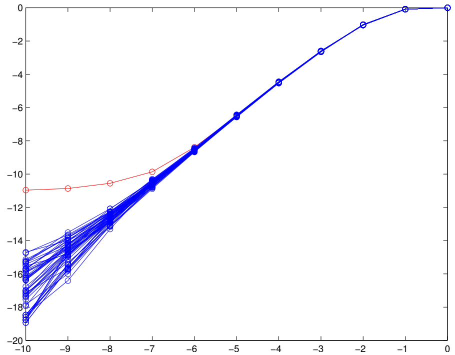
$$C(R, N, \theta, x, \tilde{\theta}, \tilde{x}) = 1/N^2 \sum_{i,j} \#(||s_i - \tilde{s}_j|| < R), \qquad (1)$$

where $\theta, \tilde{\theta}$ denote the respective model parameters and $x, \tilde{x}$ the initial values. For $\tilde{\theta} = \theta, \tilde{x} = x$ the formula reduced to the original definition of the correlation sum.

Correlation Curve variability with fixed model parameter

First, characterize the 'within variability' of a chaotic dynamical system with fixed model parameter vector:

1. Repeatedly simulate the trajectory, with varying initial values (and solver tolerances), but fixed model parameter θ_0.
2. Compute the distance matrix between (all) different trajectory pairs, to get the values $C(R, N, \theta_0, x, \theta_0, \tilde{x})$. An example for Lorenz3, with a log-scale for R:



Cost function for parameter estimation

We treat the above vectors $y = C(R_k, N, \theta_0, x, \theta_0, \tilde{x}), k = 1, ..., M$ as 'measurements' of the variability of a chaotic trajectory with a given fixed model parameter. Construct the respective likelihood:

1. Empirically estimate the statistics of $y = C(R, N, \theta_0, x, \theta_0, \tilde{x})$ from repeated simulations.
2. Create the empirical likelihood function.
3. For any trajectory $s(\theta)$ compute the distance matrix from the reference trajectory, and the respective $C(R_k, N, \theta, x, \tilde{\theta}, \tilde{x})$. Evaluate the likelihood.

Example: Likelihood for 3D Lorenz

Fix an integration time interval $[0, T]$ and the time points where the state vector is observed.

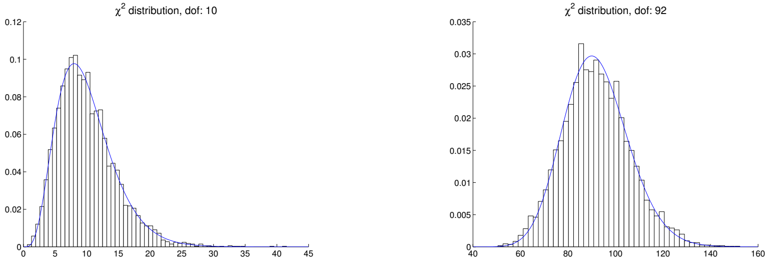
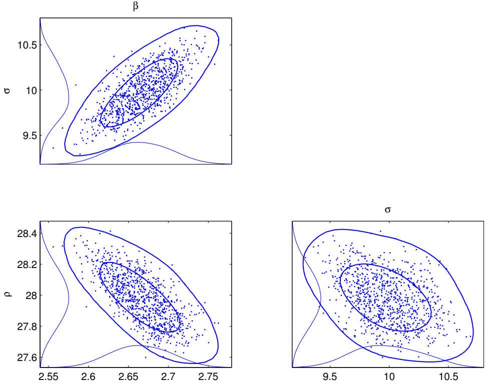


Figure: Normality check of the correlation integral vector by the χ^2 test for the Lorenz 63 system. Left: with 10 radius values used. Right: with 92 radius values

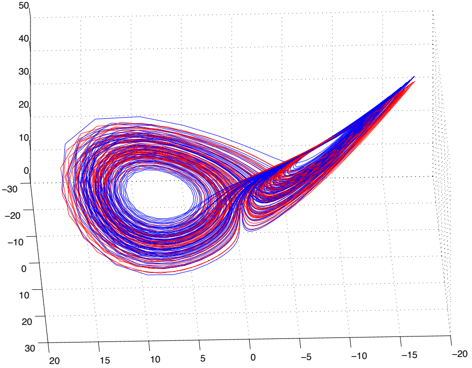
Example: 3D Lorenz

Find the distribution of model parameters that generate the 'same' trajectories, by MCMC.



Example: 3D Lorenz

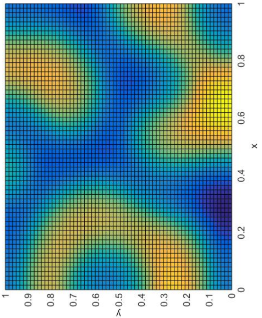
Verification: a trajectory created by a model parameter slightly outside the sampled posterior, vs the reference.



Other examples

Similar results for

- Rössler equation
- Chua circuits (3D state but more complex attractors)
- Lorenz95 (dimension 42 or 210)
- Shallow water (high dimensional, GPU implementation)
- FitzHugh-Nagumo pattern formation



Where is the real data ?

- No measured data is directly used for parameter estimation. Instead, assume "basic" model parameters given, and want to determine the posterior of parameters that would produce essentially the same chaotic dynamics.
- An example: reanalysis studies of weather and climate models (e.g., the ERA-40 data and ECHAM5), that combine past real data and model predictions to achieve the best understanding of the systems.
- The aim here: characterize the parameter distributions of the reanalyzed models, that fit the "climatology" of long time runs of a given climate model. Further use them to quantify the uncertainty of model predictions with respect to the given parameters, by parameter ensemble simulations under various scenarios, such as increased CO_2 levels.

Conclusion, Next

The summary statistics

- Need long integration times to cover the underlying attractor.
- Direct projection approaches have problems in properly identifying the parameter.
- Fractal dimension-based approaches promising.
- Ongoing work: High dimension. Shallow water, openIFS ?

References, thanks to collaborators:

- Järvinen, H., Räisänen, P., Laine, M., Tamminen, J., Ilin, A., Oja, E., Solonen, A., and Haario, H.: *Estimation of ECHAM5 climate model closure parameters with adaptive MCMC*, Atmos. Chem. Phys., Vol. 10, nro. 2, 9993-10002, 2010.
- Heikki Haario, Leonid Kalachev, Janne Hakkarainen *Generalized Correlation integral vectors: A distance concept for chaotic dynamical systems*. Chaos, 25, 2015

- Järvinen H, Laine M, Solonen A, Haario H. *Ensemble prediction and parameter estimation system: the concept*. Q. J. R. Meteorol. Soc. Vol. 138, nro. 663, pp. 281-288, 2012.
- Laine M, Solonen A, Haario H, Järvinen H. *Ensemble prediction and parameter estimation system: the method*. Q. J. R. Meteorol. Soc. Vol. 138, nro. 663, pp. 289-297, 2012.
- Ollinaho P, Laine M, Solonen A,Haario H, Järvinen H.:*NWP model forecast skill optimization via closure parameter variations*. Q. J. R. Meteorol. Soc., 139, 675, pp. 1520-1532, 2013.
- Ollinaho, P., Bechtold, P., Leutbecher, M., Laine, M., Solonen, A., Haario, H., and Järvinen, H.: *Parameter variations in prediction skill optimization at ECMWF*, Nonlin. Processes Geophys., 20, 6,1001-1010, 2013.
- Ollinaho, P., Järvinen, H., Bauer, P., Laine, M., Bechtold, P., Susiluoto, J., and Haario, H.: *Optimization of NWP model closure parameters using total energy norm of forecast error as a target*, Geosci. Model Dev., 7, 1889-1900, doi:10.5194/gmd-7-1889-2014, 2014.
- Vladimir Shemyakin, Heikki Haario: *Optimizing Ensemble Prediction System parameters with Differential Evolution* Submitted.