GCM Parameter estimation based on ETKF

Juan J. Ruiz[1] and Manuel Pulido[2]

[1]Centro de Investigaciones del Mar y de la Atmósfera (CONICET-UBA), Atmospheric Sciences Department, Universidad Nacional de Buenos Aires, Buenos Aires, Argentina [2] Universidad Nacional del Nordeste and CONICET, Corrientes, Argentina.

Motivation: General Circulation Models use a set of parametrizations to represent the effect of those processes that cannot be fully resolved at the model spatial resolution (i.e. cloud micro physics, turbulence, radiative transfer, etc.). Some of the parameters used in these parameterizations are not exactly known so they have to be estimated in order to adequately represent the atmospheric circulation. This work explores the performance of a parameter estimation algorithm based on the Ensemble Transform Kalman Filter technique that can be easily coupled with a conventional data assimilation cycle at almost no extra computational cost.

Data assimilation and parameter estimation approach:

Initial conditions estimation

Local ensemble transform data assimilation is generally used to estimate the optimal initial conditions. The implementation is as in Hunt et al. 2007 and Miyoshi et al. 2007. This method is extended to obtain also the optimal value for certain model parameters.

The analysis ensemble mean is obtained as a linear combination of the ensemble members.

The method include also a way to update the ensemble perturbations for the next data assimilation cycle.

For the computation of the analysis the algorithm is applied locally (localization is implemented as in Miyoshi et al. 2007)

Parameter estimation

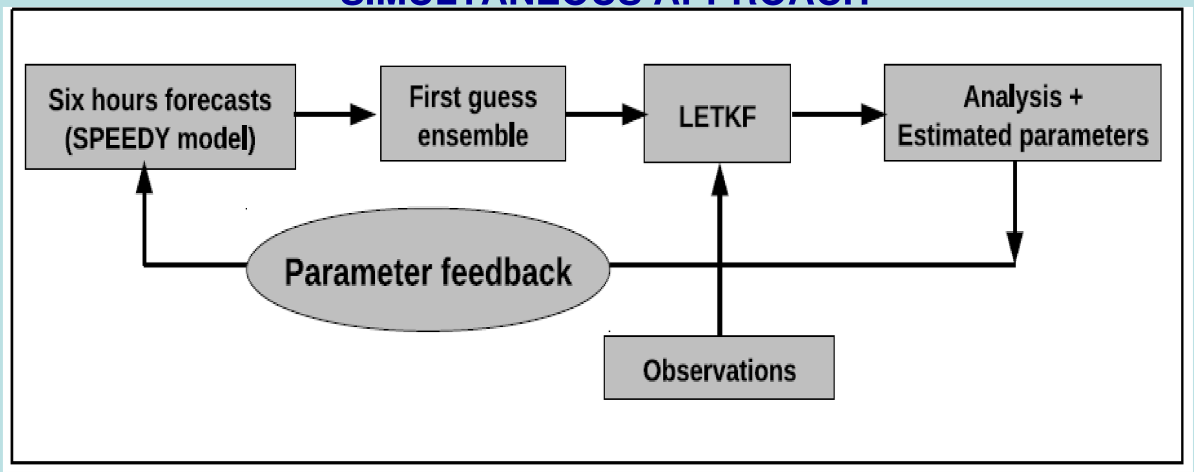
Parameters are estimated using a similar approach to that used in the estimation of the state variables. However the behavior of the parameters differ from that of the state variables in some key aspects:

-Parameter ensemble spread does not increases during the forecast step as for the state variables. It is assumed to be constant.

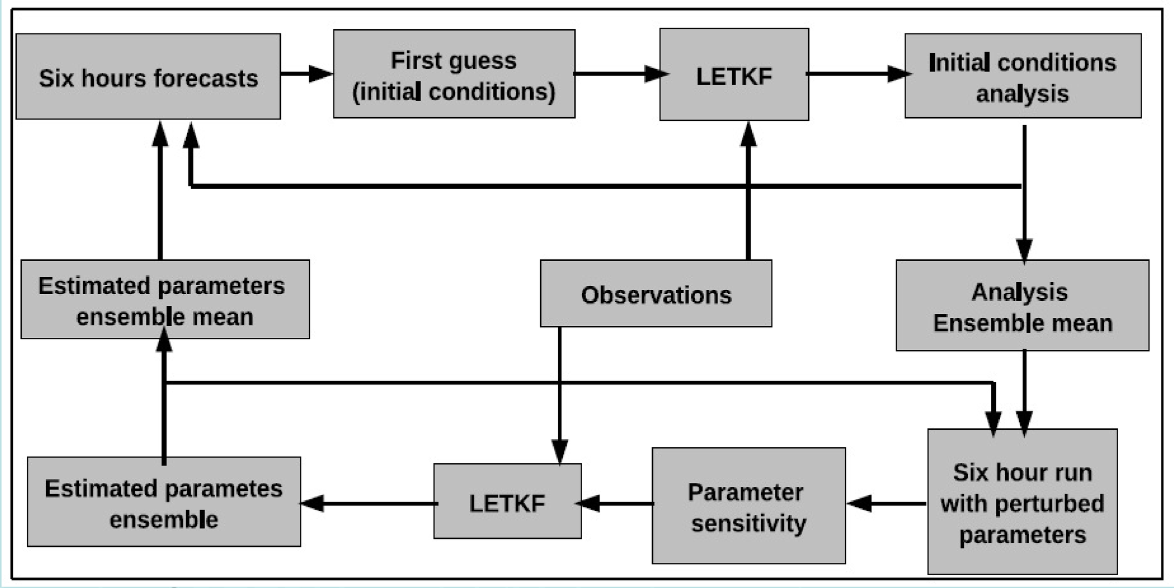
-Global parameters might be correlated with state variables all over the model domain so that no localization is used for the parameter estimation.

Two different approaches to include parameter estimation within a data assimilation cycle

SIMULTANEOUS APPROACH



PARALLEL APPROACH



Experimental settings.

Data assimilation is conducted using twin experiments based on the SPEEDY model (Molteni, 2003).

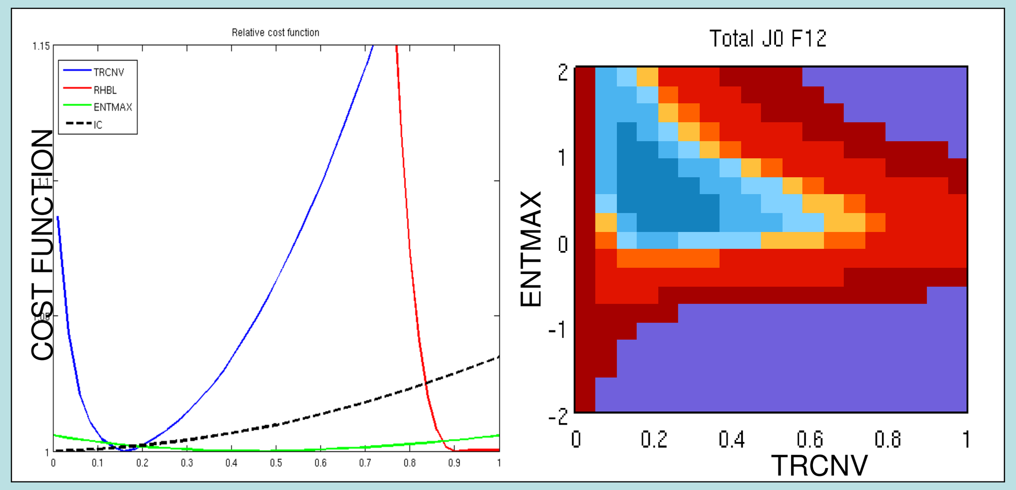
The initial ensemble of state variables is constructed using random perturbations, the parameter initial ensemble is generated as a random set of parameters.

A long continuous simulation is performed with the SPEEDY model using standard paramters which is cosidered as the true state of the system. Observations are generated from this true state adding a normally distributed random noise at some selected grid points.

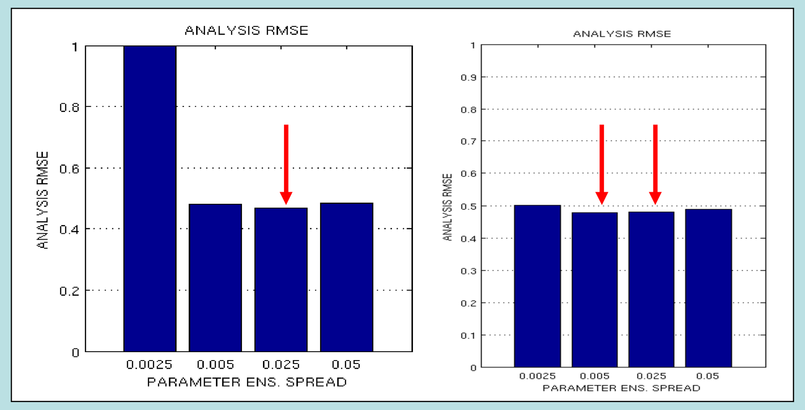
The selected parameters to be estimated are 3 parameters associated with the convective scheme in the SPEEDY model. The TRCNV (convective relaxation time), RHBL (threshold for convection initiation) and ENTMAX (associated with the rate of cumulus entrainment).

Model sensitivity to the selected parameters.

The sensitivity of the model to the different parameters has been computed (cost function). The model is more sensitive to changes in TRCNV and RHBL parameters. The sensitivity to RHBL is quite asymmetric, since the model shows almost no sensitivity for values over 0.9 which is the parameter value taken as the truth. The joint sensitivity to TRCNV and ENTMAX and for other parameters combinations has also been computed.



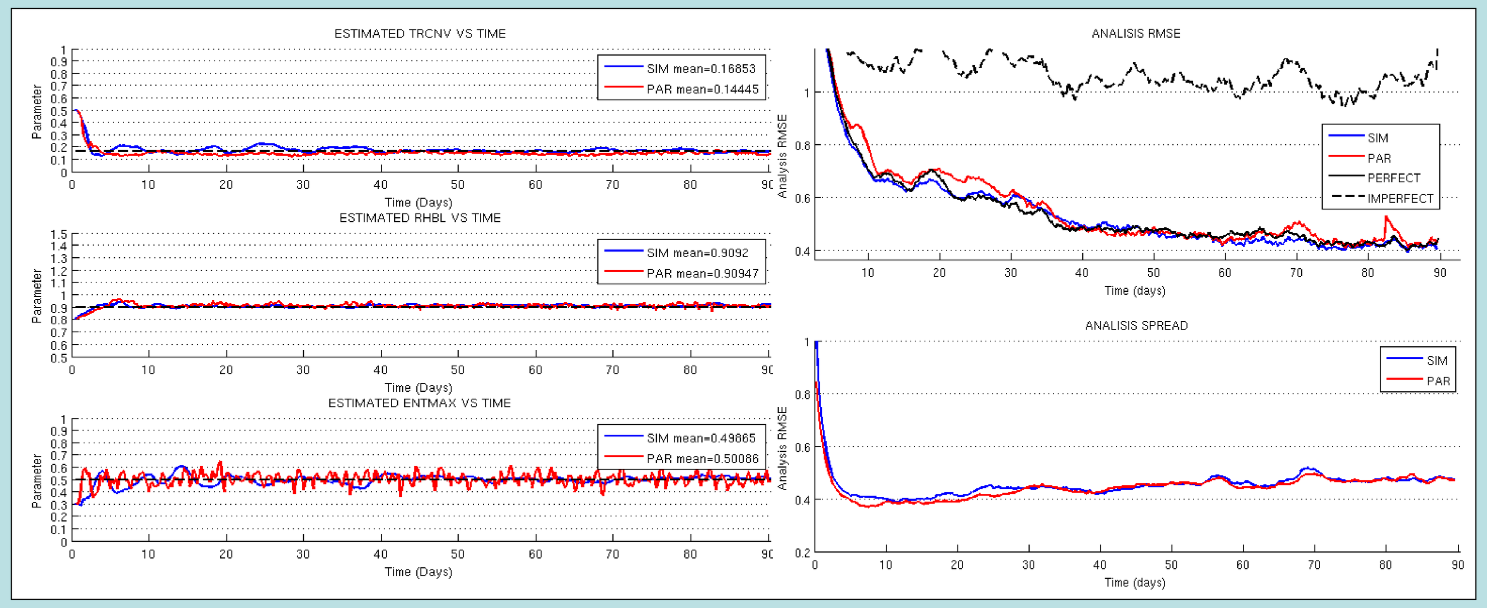
Results: 20 member sim. estimation vs 20+5 member par. estimation



Sensitivity of analysis error to the parameter ensemble spread.

-In the simultaneous experiment, when the parameter ensemble spread is too small the parameters does not converge to their true values. This might be because in this case the sensitivity of the model to the perturbations in the parameters is to small compared to the sensitivity to the perturbations in the initial conditions.

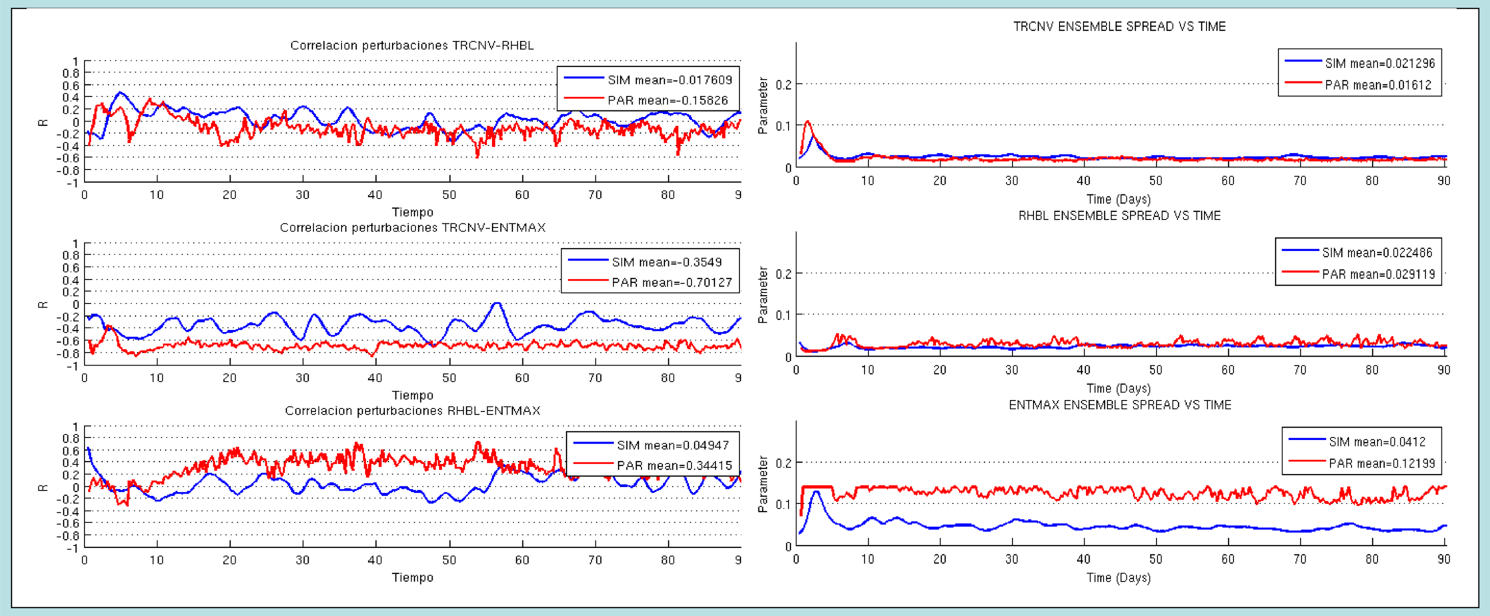
-In both cases an optimal value for the parameter ensemble dispersion is found.



Both the simultaneous and the parallel approach succeed in the estimation of the 3 parameters, although the parallel approach exhibit a small bias in the estimation of the TRCNV parameter.

In both cases when parameter estimation is included within the data assimilation cycle the analysis error is reduced and is close to the analysis error obtained using a perfect model.

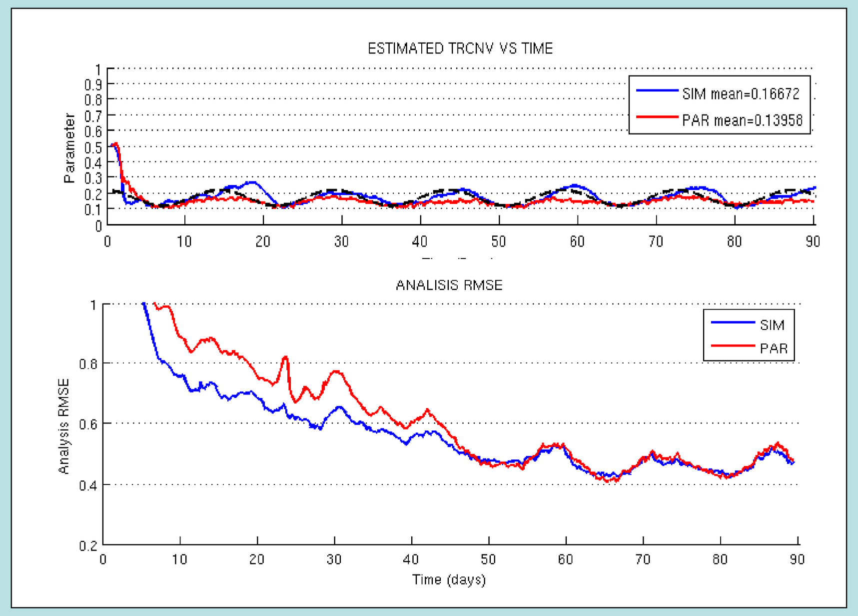
The analysis error obtained with the simultaneous and parallel approaches are similar. However the simultaneous approach is computationally cheaper.



The ETKF method can efficiently find the adequate spread relationship among different parameters.

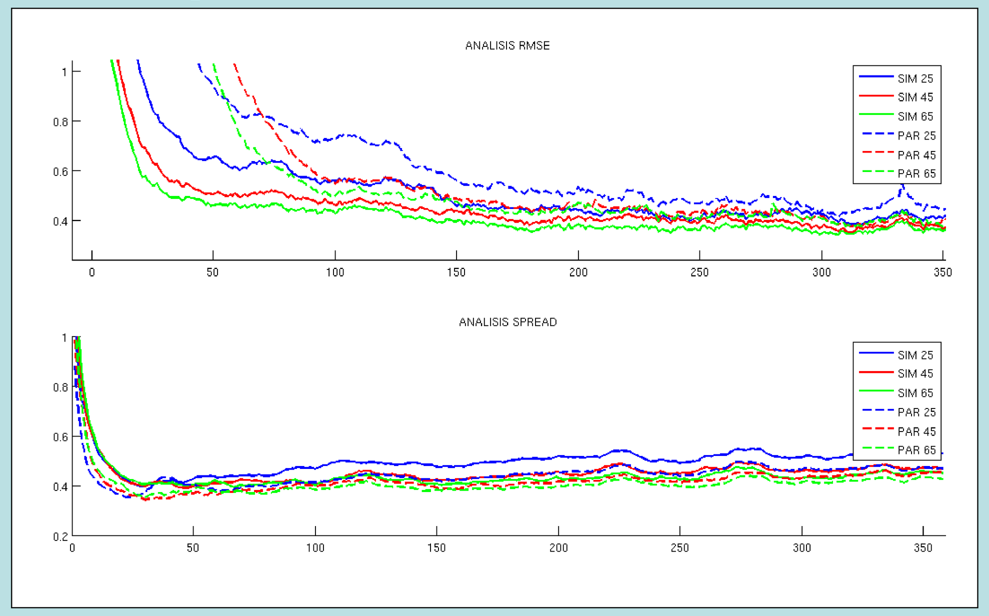
It can also represent the correlations among the parameters perturbations in order to increase the parameter ensemble spread in those directions where the sensitivity of the model to changes in the parameters is weaker.

Time dependent parameters



If the estimated parameters are time dependent, the method can still provide a good estimate of the parameters. However some temporal lag between the estimated and the true parameter arises.

Computational cost



If the simultaneous and parallel approaches are compared using the same total number of ensemble members in both cases and for different ensemble sizes, then the simultaneous approach is always better (smaller analysis error).

Conclusions

The inclusion of parameter estimation within a data assimilation cycle efficiently reduces model error associated with parameter uncertainty.

The simultaneous approach where initial conditions and parameters are simultaneously perturbed produces better results and is almost computationally free.

The ETKF provide a good estimation of the components of the error covariance matrix corresponding to the parameters.

Aknowledgements: This research was partly financed by ANPICT 2007-411. Thanks to Takemasa Miyoshi for providing the LETKF code for the SPEEDY model.

Thanks to CNRS for travel and loddging support.