Progress with model error evaluation and modelling: truncation errors

Michael Tsyrulnikov and Dmitry Gayfulin

HydroMetCenter of Russia

Genève from Moscow, 3 Sep 2020

Outline

1. What is model error and how is it evaluated?
2. Why truncation error?
3. Results of truncation error evaluation
4. Building a model for truncation errors
5. Ensemble prediction experiments with the new truncation-model-error model

What is model error and how is it evaluated?

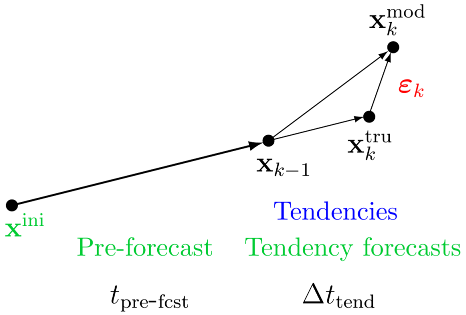
Model error = error in the model's instantaneous tendency

•For a low-resolution model in question, we compute the model error with respect to a significantly higher-resolution model.

•We start the two models from "the same initial data", compute the two short-term tendencies and claim that their difference is the model error.

•The hiRes fields are upscaled (coarse grained) when projected on the lowRes grid.

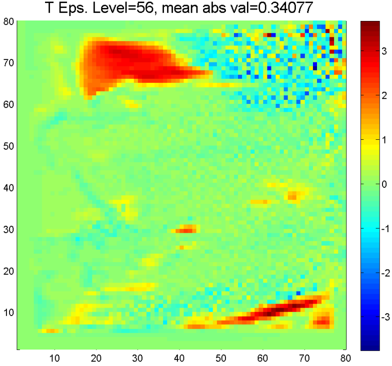
•COSMO-2.2km is the forecast model, COSMO-0.22km is the truth (both L65).



Why truncation error?

Failure to capture errors in physical parameterizations by a stochastic model

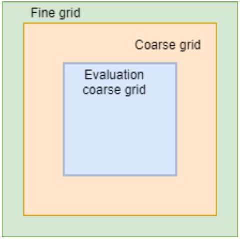
Only 3D turbulence parameterization is on:



The red spots preclude random-field modelling (no predictors found).

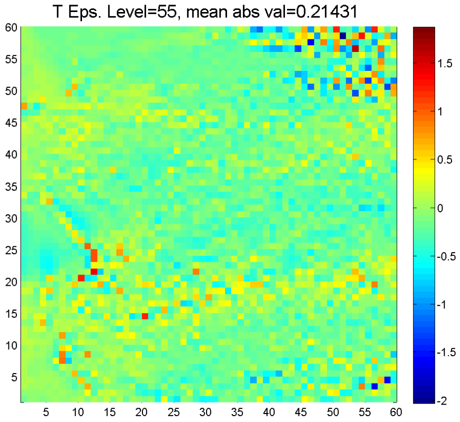
Results of truncation error evaluation

Domain and cases



- Domain over sea: North Sea.
- Fine-grid-model's domain: 851*851 points (187*187 km, greenish).
- Coarse-grid-model's domain: 80*80 points (176*176 km, pinkish).
- Evaluation-coarse-grid's domain (where model errors are computed): 60*60 points (132*132 km, bluish).
- 3 cases were studied :
  17 June, 30 June, and 5 July 2020.
- Time of the day: morning (6 UTC).

Truncation model-error field ε (T, K per hour, level 55)



$\Rightarrow$ Looks like a kind of white noise,
whose magnitude is, likely, determined by meteorology.

Building a model for truncation errors

Truncation model-error: search for predictors

Two predictors sensitive to smaller space and time scales were found.

Reason: we hypothesize that the amount of small resolvable (i.e. detectable by the LowRes model) scales is closely related to the amount of subgrid scales (unavailable to the LowRes model and responsible for its model uncertainty).

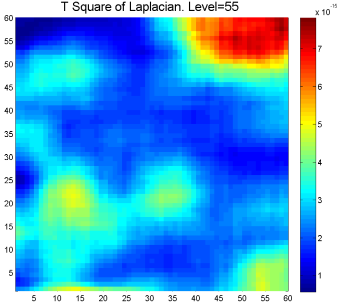
$$\boxed{F(x,y,z,t) = \mathcal{S}_{\mathrm{xy}} \cdot |\Delta_{\perp}^{2} \psi(x,y,z,t)|}$$

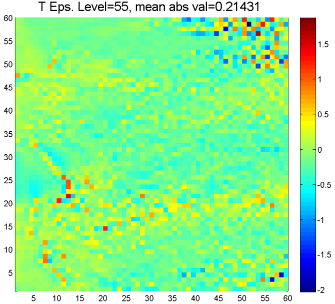
($\mathcal{S}_{\mathrm{xy}}$ is a horizontal smoothing operator) and

$$F_1(x,y,z,t) = \mathcal{S}_{\mathrm{xy}} \cdot \left| \frac{\mathrm{d}^4 \psi(x,y,z,t)}{\mathrm{d}t^4} \right|.$$

The 4th time derivative appeared to be noisy.
We tried to switch to the 2nd time derivative – but with less effect, so eventually we abandoned F_1 and **used only** F.

Performance of the predictor: F (left) vs. ε (right)





$\Rightarrow$ F is seen to be a good predictor for the truncation-model-error magnitude.

Model for ε

1. The linear homogeneous regression was fitted:

$$\mathbb{E}\left(\mathrm{abs}\left(\varepsilon\right)|F\right) = a \cdot F$$

2. The regression coefficient $a \sim 10^{14}$ appeared to only weakly depend on the level and even on the variable (T, u, v). This greatly simplified the model.
3. The normalized model error $\chi = \varepsilon / F$ appeared to be not too non-Gaussian (kurtosis = 4–6, with the Gaussian kurtosis 3).

The truncation-error model:

$$\boxed{\varepsilon(x, y, z, t) = a \cdot F(x, y, z, t) \cdot \chi(x, y, z, t)}$$

where $\chi(x, y, z, t)$ is the dimensionless zero-mean unit-variance 4D stationary Gaussian random field.

More on model for ε

With our initial error model $\varepsilon(x, y, z, t) = a \cdot F(x, y, z, t) \cdot \chi(x, y, z, t)$ the COSMO-model equations are changed from their native form

$$\frac{\mathrm{d}\psi}{\mathrm{d}t} = \Phi(\psi)$$

to the *stochastic* form

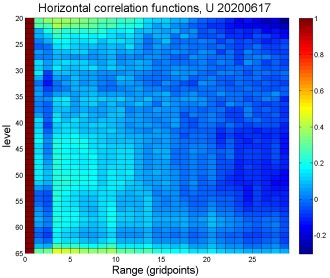
$$\frac{\mathrm{d}\psi}{\mathrm{d}t} = \Phi(\psi) + a\,\chi(x, y, z, t)\,\mathcal{S}_{\mathrm{xy}}\,|\Delta_\perp^2 \psi(x, y, z, t)|$$

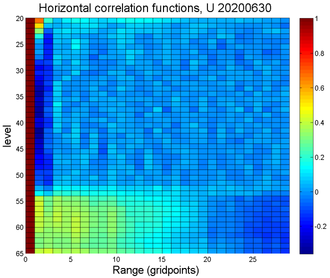
which, **in places where both χ and $\Delta_\perp^2$ are positive**, means a kind of negative diffusion (which pumps energy, especially small scales, into the system!).

In our initial experiments, the perturbed model indeed exploded..

To break the positive feedback loop (small-scale field $\rightarrow$ large $\Delta_\perp^2$ $\rightarrow$ large negative diffusion $\rightarrow$ more small scales $\rightarrow$ larger $\Delta_\perp^2$ etc.), we update the predictor field F with a longer time step (10 minutes). This fixed the problem.

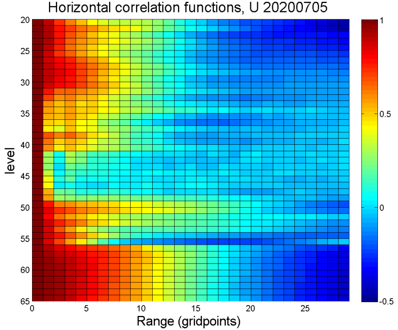
A model for $\chi = \varepsilon/F$. Horizontal correlations. U Cases 1 and 2





$\Rightarrow$ The white-noise-in-the-horizontal hypothesis looks quite good.

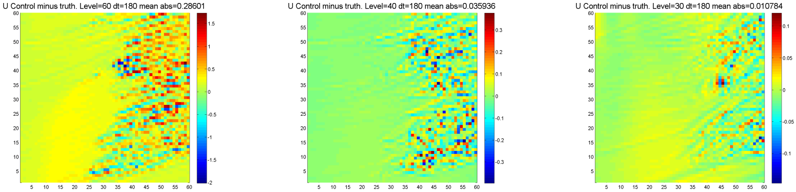
Horizontal correlations of $\chi = \varepsilon / F$. U. Case 3



$\Rightarrow$ The white-noise-in-the-horizontal hypothesis looks problematic here.

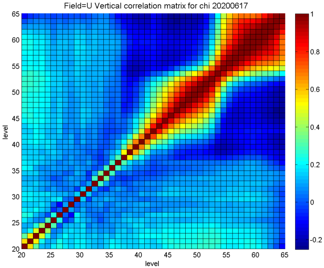
Why are there correlations in case 3?

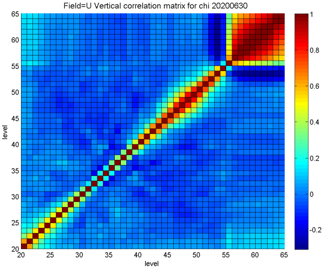
Model error. Levels 60, 40, 30. Case 3



It is the continuous (yellow) areas with biases that cause the correlations.

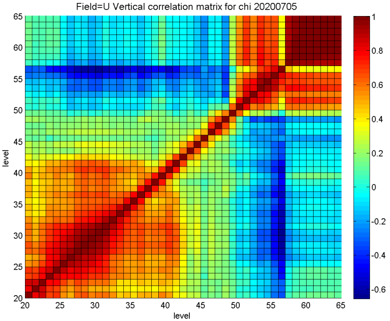
Vertical correlations of $\chi = \varepsilon / F$. U. Cases 1 and 2





$\Rightarrow$ The white-noise-in-the-vertical hypothesis is OK only outside the boundary layer.

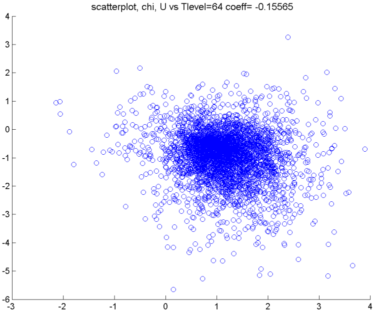
Vertical correlations of $\chi = \varepsilon / F$. U. Case 3



$\Rightarrow$ The white-noise-in-the-vertical hypothesis is not good here even above the boundary layer (due to the biases).

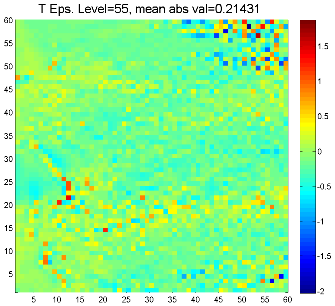
Time and inter-variable correlations of the model error

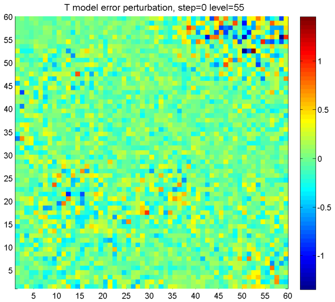
- Time correlations were measured only with 15 min lag and found very small.
- Inter-variable dependencies are also weak:



To start with, we adopted the simplest possible hypothesis that χ has **no correlations in space, and between the variables**. Knowing only that the time scale is less than 15 min, we updated the driving white noise every 2–10 minutes in our ensemble prediction experiments.

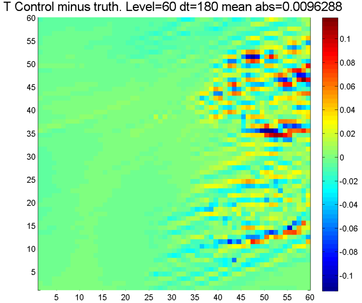
True model error (left) vs. *simulated model error* (right). Case 2

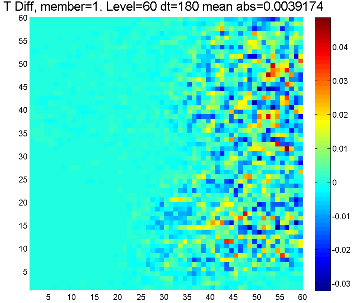




⇒ A very nice similarity!

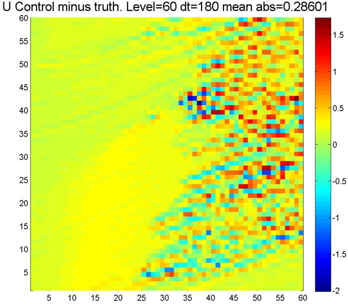
True model error (left) vs. *simulated model error* (right). (Bad) case 3. *T*

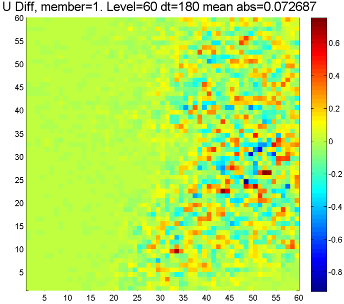




$\Rightarrow$ Very similar even in the worst case 3!

True model error (left) vs. *simulated model error* (right). (Bad) case 3. *U*



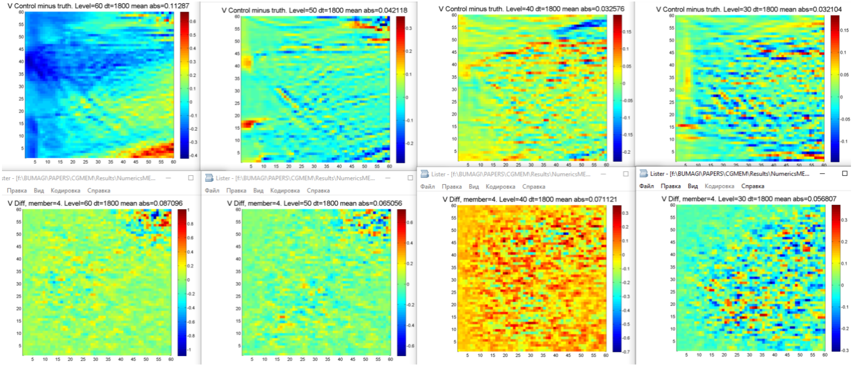


⇒ Still very similar, but the bias in the big yellow spot (the left panel, center at (23,20)) is not modeled...

Ensemble prediction experiments with the new truncation-model-error model

Ensemble prediction experiments. *V*. Case 2

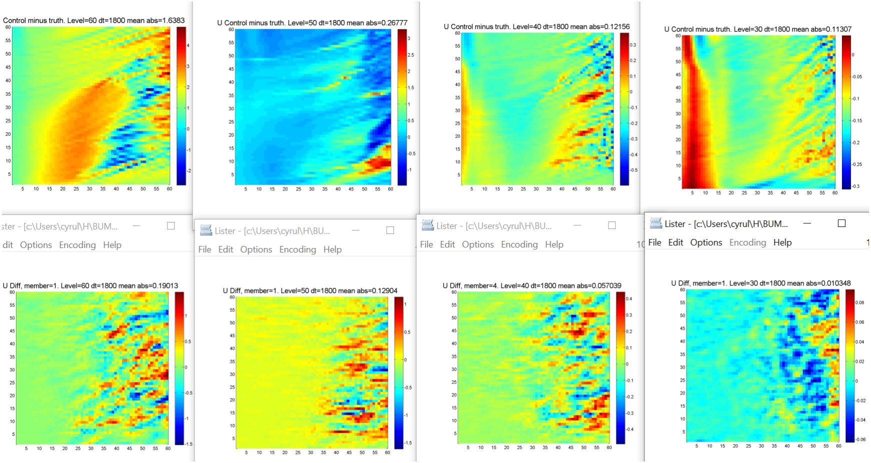
30-min Forecast error (upper panel) vs.
30-min Ensemble perturbation (lower panel)



Ensemble prediction experiments. *U*. (Bad) case 3

30-min Forecast error (upper panel) vs.

30-min Ensemble perturbation (lower panel)



Conclusions

- Truncation model-tendency-errors were evaluated using the coarse-graining approach.
- A (fast) stochastic model for truncation errors was identified and estimated.
- The truncation-error model was used in ensemble prediction experiments with known truth.
- In ensemble prediction experiments, the truncation-error model produced reasonably good forecast perturbation *patterns* but often misrepresented their magnitude and failed to predict occasional biases in the forecast.

The truncation-error model used in ensemble prediction experiments was **objectively** specified.

Further steps

- Origin of occasional biases in the truncation error fields needs to be discovered and, hopefully, modeled.
- More specialized predictors of the subgrid-scale field variance (other than Laplacian and its square) will be examined.
- Spatio-temporal error correlations in the boundary layer will be examined in more detail.
- The assumptions underlying the truncation-error model (Gaussianity, whiteness, independence of variables) are to be examined on a greater number of cases, possibly, with ICON.

Thank you!

Many thanks to D. Blinov for his help with the COSMO model.

Strong waves from the boundaries (March 2020 case)

"Model uncertainty" estimated with the 0.22km model.