Estimation of model errors on convective scales: a coarse-graining study (preliminary stage results)

Michael Tsyrulnikov and Dmitry Gayfulin

mik.tsyrulnikov@gmail.com

Abstract

An attempt to objectively estimate model tendency errors using a "true model" is described. The model in question (the "model") is COSMO with the horizontal resolution 2.2 km. The "true model" is COSMO with the horizontal resolution 0.55 km. The model error is evaluated as the difference between 1-minute "model" and upscaled (coarse-grained) "true-model" tendencies started from the same initial conditions. Preliminary results show, that, first, convection is not to be treated with this approach. Second, non-convective model errors contain both additive and multiplicative components. The additive component and the multiplier (applied to the physical tendency) appear to be approximately Gaussian. Third, the model-error field is too complex (especially in the planetary boundary layer) to be modeled with a reasonably simple stochastic model, so a process-level model error treatment is to be employed.

1 Introduction

To perform numerical weather prediction, three components are needed: initial conditions, boundary conditions, and a forecast model. The classical paradigm is deterministic: we (naively) assume that all these three components are perfect (however they are prepared), and come up with a deterministic forecast. But in reality, the three components needed to compute a forecast are, of course, imperfect and subject to uncertainty, so that the forecast inevitably contains an error. The (expected) magnitude of the error is of great interest to any user of the forecast and thus should be quantified.

1.1 Ensemble prediction

The most widely used paradigm to account for the uncertainties is stochastic: the forecast-error is assumed to be a random field with the probability distribution to be computed/estimated/specified. Correspondingly, all data used to prepare the forecast and the forecast-model itself are assumed to be random. Initial and boundary conditions are treated as multidimensional random fields. Forecast model equations are assumed to be subject of error traditionally represented by the *model error*, the difference of the model's right-hand sides from the hypothetical true right-hand sides.The model-error field is, then, also modeled as a spatiotemporal random field.The randomness of the three "input" random fields, that is, initial, boundary, and model-error fields, results in a randomness of the forecast. Therefore, to quantify the forecast probability distribution, we need, first, to adequately model the three input probability distributions and second, to map these input distributions to the output (i.e. forecast) distribution.
These two tasks are performed nowadays using Monte-Carlo, that is, the input probability distributions are represented by pseudo-random *samples* from an initial ensemble, a boundary ensemble, and a model-error ensemble. These samples are fed to the forecast model (in other words, the initial and boundary data as

doi:10.5676/dwd_pub/nwv/cosmo-nl_19_04

well as the model equations are "perturbed")[1]. The perturbed input fields give rise to a perturbed forecast. Multiple realizations of the input fields give rise to a *forecast ensemble*. The forecast ensemble is then, by construction, a sample from the forecast probability distribution. Otherwise stated, the forecast ensemble is obtained, ideally, by replacing deterministic initial and boundary conditions and deterministic right-hand sides of the model equations by random samples from the respective distributions. This approach is known in geosciences as ensemble modeling and prediction. From the forecast ensemble, a *probabilistic forecast* can be computed, replacing the deterministic forecast and quantifying the forecast uncertainty.

For the ensemble prediction to adequately describe the forecast uncertainty, the "input" uncertainties need to be adequately represented. The initial ensemble is, normally, generated by a data assimilation scheme. The lateral-boundary ensemble for a limited-area model is generated from a parent-model forecast ensemble. The lower-boundary ensemble requires a forecast ensemble in the soil/sea/lakes/rivers etc. In global models, the upper boundary condition is normally not perturbed. In limited-area models, the upper-boundary ensemble is generated in the same way as the lateral-boundary ensemble. It remains to generate samples from the distribution of model errors. Our focus in this study is on model errors.

1.2 Model errors

Existing approaches to model-error modeling can be characterized as follows. Non-stochastic approaches include multi-model and multi-physics techniques e.g., Berner et al. [1]. The most popular stochastic approach is the scheme called Stochastic Perturbations of Physical Tendencies (SPPT) e.g., Buizza et al. [2], see also its more flexible version Christensen et al. [4]. Another stochastic technique involves *additive* perturbations (called additive inflation in data assimilation) e.g., Houtekamer et al. [6], Tsyrulnikov and Gayfulin [17]. The following techniques are also widely used: parameter perturbation schemes Ollinaho et al. [9], Christensen et al. [3], the Stochastic Kinetic Energy Backscatter scheme (SKEB) Shutts [13], the stochastic convection scheme Kober and Craig [7], and the Stochastic Convection Backscatter Shutts [14]. See also the review by Leutbecher et al. [8] and other references therein. The common shortcoming of *all* the numerous above-mentioned schemes is their lack of objective justification and objective parameter estimation.

The goal of this research is to objectively estimate and stochastically model the multivariate spatiotemporal model-error field using the coarse-graining approach Shutts and Palmer [16], Shutts and Pallarès [15]. According to this approach, a higher-resolution "true model" is introduced and used (after upscaling or coarse-graining to the model resolution — hence the name of the approach) to explicitly evaluate model errors with respect to the true model. The limitation of this approach is its reliance on a *"true model"*, which of course is also approximate (as any model). The advantage of the coarse-graining approach is that it offers an opportunity to carefully and rigorously identify and estimate a model for *proxy model errors*.

Our approach is as follows.

1. As a "model", take the COSMO model with a relatively high, convection-permitting resolution (2.2 km).
2. As a "true model", take the same model but with a significantly higher resolution and more sophisticated physical parameterizations.
3. Start the "model" and the "true model" from the same initial conditions and compute the two short-time tendencies.
4. Upscale (coarse-grain) the "true-model" tendencies to the "model" resolution and explicitly compute the model-error field.
5. Accumulate a sample of model-error fields.
6. Use this sample to build a multivariate spatiotemporal stochastic model for model errors.

[1] When forecast models become inherently stochastic, there will be no need for the additive model error field, and then the stochastic ("perturbed") forecast model will be directly applied to the initial and boundary ensembles.

Note that Tsyrulnikov [19] used a similar approach that involved a "true model" but for a much simpler pair of "model" and "true-model" (vorticity equation vs. the shallow-water model).

2 Model error definition

2.1 Standard definition

To define the model error we follow Orrell et al. [10] but convert their definition from the time continuous to the time discrete form. Let the time continuous forecast model equation be

$$\frac{\mathrm{d}\mathbf{x}^{\mathrm{mod}}}{\mathrm{d}t} = \mathbf{F}(\mathbf{x}^{\mathrm{mod}}), \tag{1}$$

$\mathbf{x}^{\mathrm{mod}} \in \mathcal{X}^{\mathrm{mod}}$ is the model state vector, $\mathcal{X}^{\mathrm{mod}}$ is the model state space, and $\mathbf{F}(\mathbf{x}^{\mathrm{mod}})$ is the model operator. Numerically integrating this equation yields its time discrete solution

$$\mathbf{x}^{\mathrm{mod}}_k = \mathcal{M}(\mathbf{x}^{\mathrm{mod}}_{k-1}), \tag{2}$$

where k denotes the time step and $\mathcal{M}(\mathbf{x}^{\mathrm{mod}})$ is the time discrete model operator.

Next, let the truth be denoted by $\mathbf{x}^{\mathrm{tru}}_k \in \mathcal{X}^{\mathrm{tru}}$, where $\mathcal{X}^{\mathrm{tru}}$ is the true state space.

Finally, postulate that there is a *projection* (see, e.g., Tsyrulnikov [19]) $\mathbf{\Pi} : \mathcal{X}^{\mathrm{tru}} \rightarrow \mathcal{X}^{\mathrm{mod}}$ such that any true state is mapped to the model space, getting a model state (denoted by the tilde):

$$\widetilde{\mathbf{x}}^{\mathrm{tru}} = \mathbf{\Pi}\mathbf{x}^{\mathrm{tru}}. \tag{3}$$

Normally, elements of $\mathcal{X}^{\mathrm{tru}}$ are *continuous* fields in physical space (as opposed to space-discrete fields in $\mathcal{X}^{\mathrm{mod}}$), in this case the action of $\mathbf{\Pi}$ amounts to *upscaling* the true field to the model-space fields resolution[2]. Now we are in a position to define the model error. To this end, we start the "model" from the true initial conditions $\mathbf{x}^{\mathrm{tru}}_{k-1}$ at time $k-1$ (that is, from the upscaled truth $\widetilde{\mathbf{x}}^{\mathrm{tru}}_{k-1}$), compute the one-step model forecast, and compare it with the (upscaled) truth at time k.

The difference is solely due to the inability of the model to predict the true evolution of the system (atmosphere) and therefore it is called the model error:

$$\boxed{\boldsymbol{\varepsilon}_k = \mathcal{M}(\widetilde{\mathbf{x}}^{\mathrm{tru}}_{k-1}) - \widetilde{\mathbf{x}}^{\mathrm{tru}}_k.} \tag{4}$$

Adding and subtracting $\widetilde{\mathbf{x}}^{\mathrm{tru}}_{k-1}$ from the right-hand side of Eq.(4) shows that the model error is the model *tendency* error as well:

$$\boxed{\boldsymbol{\varepsilon}_k = \left[\mathcal{M}(\widetilde{\mathbf{x}}^{\mathrm{tru}}_{k-1}) - \widetilde{\mathbf{x}}^{\mathrm{tru}}_{k-1}\right] - \left[\widetilde{\mathbf{x}}^{\mathrm{tru}}_k - \widetilde{\mathbf{x}}^{\mathrm{tru}}_{k-1}\right] \equiv \mathbf{T}^{\mathrm{mod}}_k - \widetilde{\mathbf{T}}^{\mathrm{tru}}_k,} \tag{5}$$

where $\mathbf{T}$ stands for the one-step (total) tendency.

2.2 Definition of model error that assumes that there is a true model

It is essential that the model operator is applied in Eq.(4) to the (upscaled) **truth**. To understand why this is required, let us hypothesize that there is a *true model*:

$$\mathbf{x}^{\mathrm{tru}}_k = \mathcal{M}^{\mathrm{tru}}(\mathbf{x}^{\mathrm{tru}}_{k-1}), \tag{6}$$

where $\mathcal{M}^{\mathrm{tru}}$ is the operator of the true model. Note that from now on the superscript "tru" denotes the true model (not the truth as in section 2.1).

[2] It is also plausible that the truth or the "true model" involves more fields (e.g. additional air constituents) than the model. In that case, the projector $\mathbf{\Pi}$ just ignores the fields present in the truth or the "true model" but absent in the "model". This situation is not considered in this study.

Substituting $\mathbf{x}_k^{\mathrm{tru}}$ from Eq.(6) into Eq.(4) yields

$$\boldsymbol{\varepsilon}_k = \mathcal{M}(\widetilde{\mathbf{x}}_{k-1}^{\mathrm{tru}}) - \widetilde{\mathcal{M}}^{\mathrm{tru}}(\mathbf{x}_{k-1}^{\mathrm{tru}}), \tag{7}$$

where $\widetilde{\mathcal{M}}^{\mathrm{tru}}(\mathbf{x}_{k-1}^{\mathrm{tru}}) \equiv \mathbf{\Pi}\mathcal{M}^{\mathrm{tru}}(\mathbf{x}_{k-1}^{\mathrm{tru}})$. From Eq.(7), we see that the model error is the difference of the one-time-step *model* solution and the one-time-step *true-model* solution **provided that the two models are initialized from the same state** (up to the subgrid-scale field components). We call this requirement the *same start condition*. Informally, the same start condition is very meaningful: in order to compare the two models we specify the same *inputs* and look at the *outputs* so that difference is only due to the difference in the model's operators.

Remark. The above model error definition can also be viewed as follows. If we take the upscaled truth at the two consecutive time steps $k-1$ and k and substitute them into the forecast-model Eq.(2), then Eq.(2) will not be exactly satisfied because the truth is not governed by the (inevitably approximate) forecast model equation. The discrepancy is, by Eq.(4), the model error. Formulated this way, the model error definition appears to exactly correspond to the definition of the *truncation error* of a numerical scheme in solving a differential equation. Indeed, following eg Richtmyer and Morton [12], section 1.6, we substitute the exact solution to the differential problem (i.e. the truth we seek to approximate) into the finite-difference scheme (the "approximating model") and call the residual the truncation (model) error.

Equation (7) can be rewritten in terms of *tendencies* by subtracting and adding $\widetilde{\mathbf{x}}_{k-1}^{\mathrm{tru}}$ in its right-hand side and rearranging the terms:

$$\boxed{\boldsymbol{\varepsilon}_k = \left[\mathcal{M}(\widetilde{\mathbf{x}}_{k-1}^{\mathrm{tru}}) - \widetilde{\mathbf{x}}_{k-1}^{\mathrm{tru}}\right] - \mathbf{\Pi}\left[\mathcal{M}^{\mathrm{tru}}(\mathbf{x}_{k-1}^{\mathrm{tru}}) - \mathbf{x}_{k-1}^{\mathrm{tru}}\right] \equiv \mathbf{T}_k^{\mathrm{mod}} - \widetilde{\mathbf{T}}_k^{\mathrm{tru}}.} \tag{8}$$

We reiterate that here and for the remainder of the article $\widetilde{\mathbf{T}}_k^{\mathrm{tru}}$ denotes the tendency of the true model, not the true tendency as in Eq.(5). Technically, since the true model may have a shorter time step, we allow the tendencies to be computed for several time steps so that the *tendency forecasts* used to compute $\mathbf{T}_k^{\mathrm{mod}}$ and $\mathbf{T}_k^{\mathrm{tru}}$ have the same lead time.

This generalization can also be used to check if there is an initial transient process due to possible imbalances in the starting points (fields) from which the tendency forecasts are initialized (by comparing the tendency fields for different tendency-forecast lengths, see section).

To summarize, the standard definition of model error, Eq.(4), assumes that the truth at the time instants in question, $k-1$ and k, that is, $\mathbf{x}_{k-1}^{\mathrm{tru}}$ and $\mathbf{x}_k^{\mathrm{tru}}$, is the *actual truth*, that is, the truth actually observed in the nature by real-world observations. By contrast, the definition that makes use of a true model, Eqs.(7)–(8), allows us to lift this restriction and assume that the $\mathbf{x}_{k-1}^{\mathrm{tru}}$ is **can be *any* point** on the true-model "attractor" (in practice, not far from the true-model "attractor", that is, with reasonably balanced initial conditions for the true model).

3 Evaluation of model error

3.1 Motivation

Tsyrulnikov and Gorin [18] tried to use the *standard* definition of model error, Eq.(4), to evaluate model errors through comparing finite time model tendencies with finite-time observed tendencies. They found in a simulation study with the COSMO model (with 20 km horizontal resolution) that the main obstacle was the requirement to start the model from the truth. Starting from analyses led to too high estimation errors. In order to make the analyses accurate enough to reliably estimate even the simplest constant-in-space and piecewise-constant-in-time model error of realistic magnitude, the assimilated observations of temperature and wind needed to have currently unreachable accuracy (0.1 K for temperature and 0.02 m/s for winds) and be available at each model grid point. Hence, reliable estimation of realistic model errors by comparing finite-time model tendencies with finite-time observed tendencies is **not possible** with existing observational

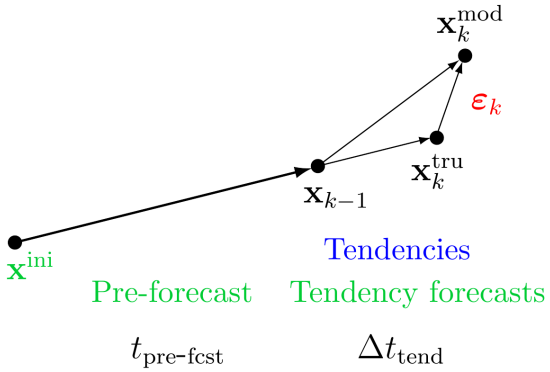

Figure 1: Model error evaluation schematic. Dots denote points in state space. Differences between the true state space and model state space are not highlighted for simplicity here.

networks. This has motivated the present research, in which we explicitly specify a **"true model"** and use the extended definition of the model error introduced in section to evaluate the model error.

3.2 Notation

The "model" in question is referred to as the coarse-grid model (cgm, also abbreviated as mod). The coarse grid is abbreviated as cg. Fields on cg are denoted by the tilde.

The "true model" is referred to as the fine-grid model (fgm, also abbreviated as tru). The fine grid is abbreviated as fg. By "grid-scales" (abbreviated as gs) we mean scales resolved on cg. By "subgrid-scales" (abbreviated as sgs) we mean scales *not* resolved on cg (but resolved by fgm).

3.3 The proposed approach

In order to generate the common starting point for the two tendency forecasts and apply the model-error evaluation methodology described in section , we run a cgm *pre-forecast* (to "spin" the model up). The pre-forecast starts from the initial point $\mathbf{x}^{\mathrm{ini}}$ generated from a global model. The alternative approach in which the pre-forecast is performed using fgm will be investigated on a later stage of the project.

At the end of the pre-forecast, we obtain the cgm forecast fields denoted by $\mathbf{x}_{k-1}$ in Fig.1 and by $\widetilde{\mathbf{x}}_{\mathrm{start}}$ in what follows. The cgm *tendency forecast* starts immediately from $\widetilde{\mathbf{x}}_{\mathrm{start}}$. The fgm *tendency forecast* starts from a *downscaled* version of $\widetilde{\mathbf{x}}_{\mathrm{start}}$ denoted by $\mathbf{x}_{\mathrm{start}}$.

Then, we run two very-short-term tendency forecasts *of the same lead time* $\mathbf{x}_k^{\mathrm{mod}}$ and $\mathbf{x}_k^{\mathrm{tru}}$, compute the two tendencies, downscale the fgm tendency, and finally compute the model error field ε_k following Eq.(8). A more precise and detailed description of our approach is given in section 4.2 .

3.4 Upscaling

As noted in section , an upscaling (coarse-graining, aggregation) is needed to properly project a high-resolution fgm field onto the cgm phase space. The upscaling removes the scales not represented on the coarse model grid and thus makes the true field *comparable* with its model counterpart. Note that not performing upscaling (i.e. simply *restricting* the high-resolution fields on a coarse grid) would give rise to the phenomenon known as *aliasing* so that the sgs field components would be superimposed on the gs field components, irreversibly distorting them.
Normally, the resolution of fgm is higher than cgm not only in space but also in time (shorter time steps),

which implies that the upscaling must involve the dime dimension as well as the (three) spatial dimensions. There are two common approaches to upscaling: spectral and physical-space based. A *spectral* upscaling is more common in mathematics. It is performed by Fourier transforming the spatial field $\mathbf{x}^{\mathrm{tru}}$, truncating the resulting expansion at the model-grid resolution, and computing $\widetilde{\mathbf{x}}^{\mathrm{tru}}$ as the inverse Fourier transform of the truncated expansion. This approach exactly removes all sgs spectral components.

In meteorology, a *physical-space* upscaling is more common, it consists in averaging the high-resolution fg field $\mathbf{x}^{\mathrm{tru}}$ over cells of lower-resolution cg, see e.g. Shutts and Palmer [16]. This technique is simpler (as it is local, in contrast to the spectral approach) and more physically appealing, albeit not precise mathematically. We will adopt this physical-space definition to simplify our analysis.

Note that in the context of coarse-graining studies, Shutts and Pallarès [15] used a spectral low-pass filter in the horizontal with a squared exponential (i.e. not rectangular as in the ideal low-pass filter) transfer function. In time, they performed low-pass filtering by ad-hoc averaging. Both filters were applied both to the model (they called it "target") and to the true model ("truth").

4 Numerical experiments

4 cases were studied (all 12 UTC): 1 July and 29 July 2017 ("convective" days) and 17 July and 1 December 2017("non-convective" days).

4.1 Models'setup

The cgm was COSMO with 50 vertical levels and horizontal resolution 2.2 km. The fgm was the same COSMO with the following differences from cgm:

1. The horizontal resolution was 0.55 km.
2. The time step was 5 s in fgm vs. 20 s in cgm.
3. The shallow convection parameterization (Tiedtke) was switched off in fgm whilst switched on in cgm.
4. A 3D turbulence scheme was used in fgm vs. a 1D scheme in cgm.
5. Some more sophisticated options were used in the fgm's cloud and precipitation scheme as well as in the radiation scheme as compared to cgm.

The models domains were centered at 52N 35E, see Fig.2. The outermost grid is cg (250*250 points, 550*550 km, marked in greenish). The second-largest grid in Fig.2 is fg (600*600 points, 330*330 km, marked in pinkish). The innermost grid is the model-error evaluation grid (mesh size 2.2 km, 110*110 points, 220*220 km, marked in bluish). The three domains/grids were nested one in another with the intention to reduce any impact of lateral boundaries in a 3h forecast. This will be useful at a later stage of the project when the developed model-error model is verified in an ensemble prediction system.

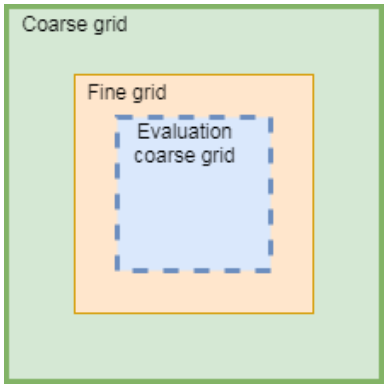

Figure 2: The horizontal grids: fg (the outermost grid), cg (the intermediate grid), and the innermost model-error evaluation grid.

4.2 Computation of the model error

The model-error evaluation technique is summarized as follows.

1. Pre-forecast: run cgm for 1 h lead time. The pre-forecast is used as the cgm *starting point* $\widetilde{\mathbf{x}}_{\text{start}}$.
2. Downscale $\widetilde{\mathbf{x}}_{\text{start}}$ to fg (using the COSMO tool INT2LM). This is $\mathbf{x}_{\text{start}}$. This procedure is meant to approximately satisfy the "same start" condition.
3. Run cgm for 3 time steps (60 s in total) starting from $\widetilde{\mathbf{x}}_{\text{start}}$. Calculate the total tendency $\mathbf{T}_3^{\text{mod}}$.
4. Run fgm for 12 time steps (60 s in total) starting from $\mathbf{x}_{\text{start}}$. Calculate the total tendency $\mathbf{T}_{12}^{\text{tru}}$.

5. Upscale $\mathbf{T}_{12}^{\mathrm{tru}}$ to the coarse grid, getting $\widetilde{\mathbf{T}}_{12}^{\mathrm{tru}}$.
6. Compute the model error as $\boxed{\boldsymbol{\varepsilon} = \mathbf{T}_{3}^{\mathrm{mod}} - \widetilde{\mathbf{T}}_{12}^{\mathrm{tru}}}$.

The length of the tendency forecasts (60 s) was selected by trial and error. We also tried the tendency-forecast lengths 20 s and 5 min and found that the results presented below were quite stable and not critically dependent on the tendency-forecast length within the above range. With 20 s, there were some indications of an initial imbalance (not shown). The model-error fields for 1 min and 5 min were similar in terms of their spatial scales and variability.

4.3 Results

To give an impression of how model errors are related to model tendencies, we show two (i.e. cgm and fgm) total tendencies for the zonal wind component at an arbitrarily selected model level 41 (about 700 m above ground). One can see that the two tendencies are very similar, implying that the model error (which is their difference) is quite small, as expected (recall that COSMO is an operational-class model used in many countries).

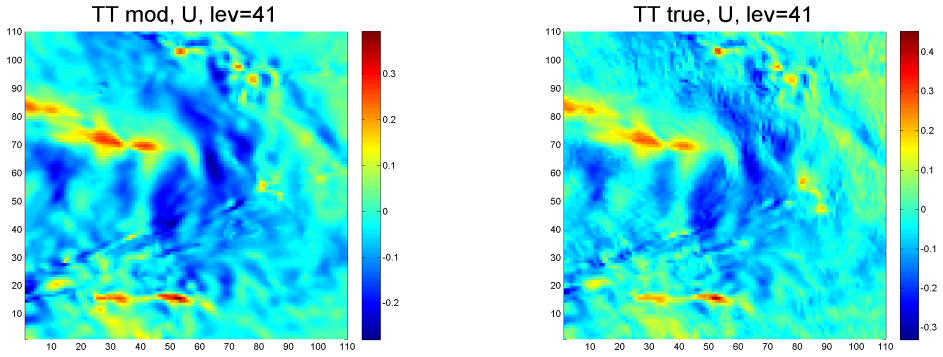

Figure 3: Total tendencies: cgm (left) and fgm (right)

4.3.1 Role of convection

Figure 4 shows the model-error field ε (left) along with the convective physical-tendency field P_{conv} (right) at the model level 32 (about 2.5 km above ground). It is striking that the model-error field ε is dominated by a relatively small number of *outliers*, with the rest of the field being relative close to zero. Comparing the left and the right panels of Fig.4 suggests that it is the convective parameterization that produces those large model errors. At some grid points where the large cgm's P_{conv} is matched with a large tendency produced by the fgm-explicitly-resolved convection — at those points, ε is small. At other points where the cgm's P_{conv} is large, it is *not* matched with a similarly large fgm's tendency — at those points, ε is big (and seen in Fig.4(left) as outliers).

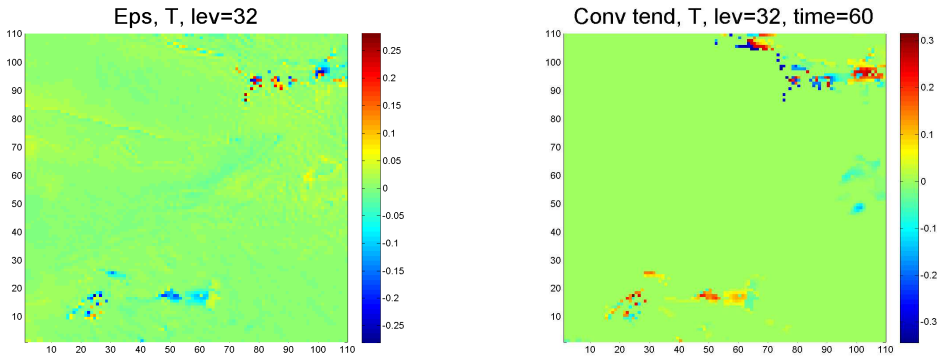

Figure 4: Model errors (left) and convective tendency (right) at the model level 32. The units are K/min.

This outstanding impact of the shallow convection parameterization (a similar effect of the *deep* convection parameterization in a 7-km-resolution COSMO was even much bigger, not shown) suggests that errors in the convective parameterization need a special treatment (like any outliers in general). We attempted to use predictors like CAPE and the vertical lapse rate to spot those large convective model errors,but those attempts failed dramatically (not shown). We also realized that, given the *complexity* of the convection phenomenon, a purely stochastic approach looks unsuitable to model convective model errors.A physical model is needed. Besides, convection is a *fast* and strong phenomenon so that the convective model errors we can measure are the *outcome* of convection, not its *source*. And it is a "convective source" that we would like to isolate, study, and model in this study (and then perturb in an ensemble forecast).

To verify the conjecture that it the invisible "convective source" that needs to be perturbed, we introduced tiny and constant-in-space-and-time model-error perturbations at all model levels and looked at the resulting forecast perturbation in a 15-min cgm forecast. We imposed model-error perturbations with the magnitude $5 \cdot 10^{-5}$ K per time step in T and 10^{-4} m/s in U, V. The resulting forecast-error perturbation of temperature at the model level 30 is shown in Fig.5(right). The respective model-error field is shown in Fig.5(left). It is seen that both fields look quite similar, which means that realistic convection-related forecast errors can be obtained by just "any" perturbation of temperature and other fields. This can be compared with findings by Flack et al. [5], who introduced temperature model-error perturbations of magnitude 0.1 K with the spatial length scale of $6\Delta x$ (where Δx is the horizontal mesh size) every 15 minutes during the model integration at a single model level at the model hybrid height coordinate 261.6 m.

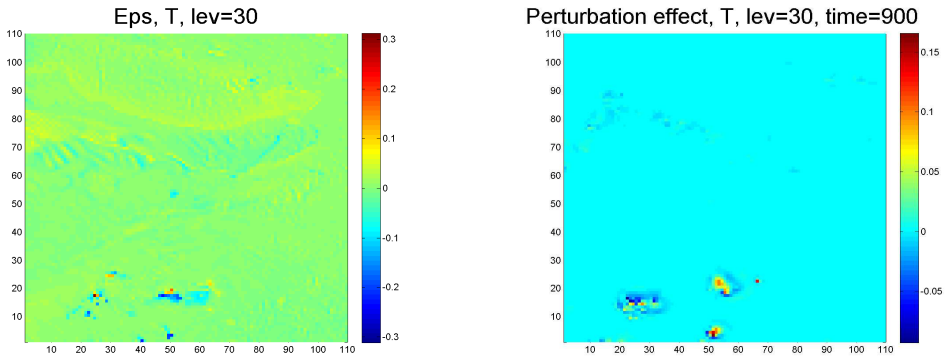

Figure 5: Model errors, K/min, (left) and forecast perturbation, K (right).

4.3.2 Conclusions on convection

The above results indicate that errors in parameterized convection **cannot be treated using the "true-model" approach** (which we pursue in this study) for the following reasons.

1. Evaluating model errors as differences between cgm and fgm short-time tendencies implies, in case of convection, that differences in the *outcomes* of convection are actually measured.

   However, initiation of convection plumes is a sporadic process "modulated" by a hypothetical "convective source" (like CAPE). And it is this source whose uncertainty needs to be modeled, not the outcome. Objectively sensing the uncertainty in this "convective source" is not possible with our approach and, maybe, impossible in principle.

2. A *deterministic* convective parameterization (currently in use in the standard configuration of COSMO adopted in this research) attempts to represent the contribution of subgrid convective elements to the grid-scale fields. However, with the cg mesh as small as 1–2 km, the number of those convective elements (modeled by fgm) in a cg cell is not large enough for their combined effect to be considered as deterministic Shutts and Pallarès [15]. It is inherently random with high variance. And this high variability is the major contributor to the cgm-minus-fgm difference we can measure — because the deterministic cgm tendency produced by the convective scheme is inevitably almost always far from the highly random upscaled fgm tendency. But this random difference is, actually, not the error, it is the *uncertainty* related to the manifestation of the stochastic nature of convection and should not be regarded as model error. This kind of error requires a *stochastic* convective parameterization scheme like Plant and Craig [11]), whose development, estimation, and calibration is beyond the scope of this research.

So, in this work, we do not treat model errors due to parameterized convection. The convective parameterization is switched off both in cgm and fgm, and predominantly non-convective cases are studied.

4.3.3 Non-convective model errors

First, we show the temperature model-error field (ε) at a high enough model level 21 (about 7 km above ground)such that there is, likely, no convection there, see Fig.6. The ε field looks here like a Gaussian random field (in contrast to the above convection-contaminated model-error fields).

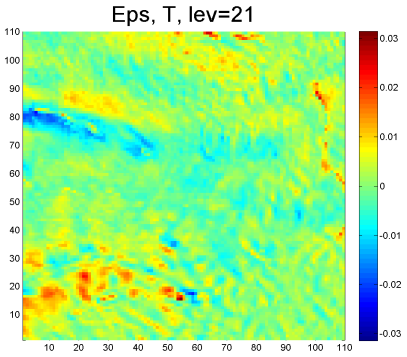

Figure 6: Non-convective model error at level 21, temperature, K/min

Next, Fig.(7) displays the temperature model-error within the planetary boundary layer (level 43 about 500 m above ground). The field looks like a random field with very complicated structure, with multiple scales, and, likely, with multiple components. Building a stochastic model for such field is a very challenging task.

4.3.4 Physical tendency as a predictor for model error

We start with looking at the physical tendency as a potential predictor for model errors. Figure 8 shows the meridional-wind (V) model-error field (left) along with the physical-tendency field P (right) at the model level 31 (about 3 km above ground). It is clearly visible that if P is large, then ε is also, usually, large, so that the physical tendency is indeed a useful indicator of the magnitude of model error. However, there are several areas where the model-error field ε is large while the physical-tendency field P is small. This implies that physical tendency is, actually, of limited use as a model-error predictor.

To get a deeper understanding of the relationship between the model error and the physical tendency, we estimated the conditional probability density of ε given the absolute value of P, i.e. $p(\varepsilon \,|\, \mathsf{abs(P)})$. In the estimator, 2 percent of largest $|\varepsilon|$ and $|\mathsf{P}| \equiv \mathsf{abs(P)}$ were taken down with the intention to filter out grid points with convection. Values of $|\mathsf{P}|$ were then grouped in 10 equipopulated bins for which histograms of ε were plotted. As an example, the histogram of ε for the 4-th bin of $|\mathsf{P}|$ (temperature, level 30) is displayed in Fig.9.

Remarkably, this conditional distribution is seen to be not too far from Gaussian. More qualitatively, its kurtosis is 4.0 (the Gaussian kurtosis is 3).For other variables and levels, kurtosis remained, mainly, between 3 and 4,thus indicating that Gaussianity is, perhaps, a reasonable hypothesis for the probability distribution of non-convective model errors given the physical tendency, $\varepsilon \,|\, \mathsf{abs(P)}$.

Next, having estimated $p(\varepsilon \,|\, \mathsf{abs(P)})$, we used it to examine the conditional variance of non-convective model errors: $\mathsf{Var}\,(\varepsilon \,|\, \mathsf{abs(P)})$. We computed scatterplots of ε^2 vs. P^2 and then smoothed them using a kernel smoother with the Epanechnikov kernel and an empirically selected bandwidth. The resulting dependencies are exem-

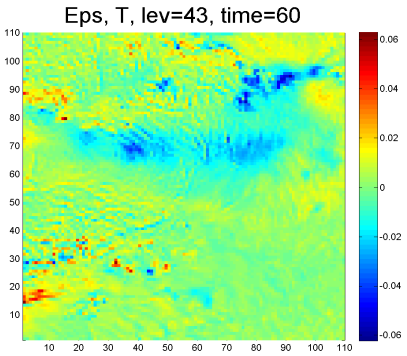

Figure 7: Non-convective model error at level 43, temperature, K/min

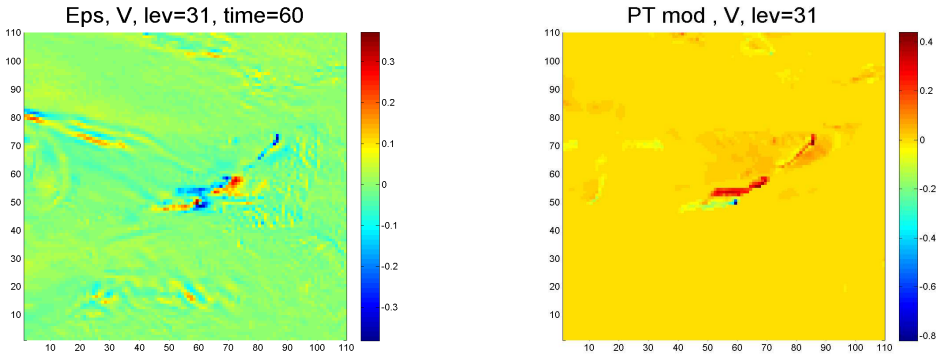

Figure 8: Model error (left) and physical tendency (right), meridional wind, $\mathrm{m}s^{-1}\mathrm{min}^{-1}$

plified in Fig.10 for temperature at the model level 28 (about 4 km above ground), where the x-axis is P^2 and the y-axis is ε^2.

The resulting conditional variance curves were somewhat noisy even after smoothing but two their salient features were standing out. First, there always was a significantly non-zero offset (the value of $\mathbb{E}\varepsilon^2$ for $\mathsf{P} = 0$), which can be interpreted as the variance of the *additive* (physical-tendency independent) model-error component.Second, the model-error variance, by and large, grew with the increasing physical tendency. The growth was, in a first approximation, linear, thus suggesting that it can be interpreted as the *multiplicative* (physical-tendency dependent) model-error variance. As a result, the following **additive-multiplicative** model-error model is our first (and preliminary) finding:

$$\boxed{\varepsilon(\mathbf{s}) = \alpha(\mathbf{s}) + \mu(\mathbf{s}) \cdot \mathsf{P}(\mathbf{s}),} \tag{9}$$

where $\alpha(\mathbf{s})$ is the additive model-error component and $\mu(\mathbf{s})$ is the random multiplier field.

To a first approximation, $\alpha(\mathbf{s})$ and $\mu(\mathbf{s})$ can be assumed to be Gaussian random fields with their horizontal, vertical, temporal, and multivariate structure to be identified.

Finally, in Table 1 we show the relationship between the magnitudes of the additive and multiplicative model-error components. One can see that the magnitudes of the additive error components were somewhat larger than the magnitudes of the multiplicative error components. Only in the boundary layer (where turbulence dominates the physical tendency), the multiplicative errors were comparable to additive errors or even larger (not shown).

The difference between the values for U and for V is, perhaps, due to insufficient statistics.

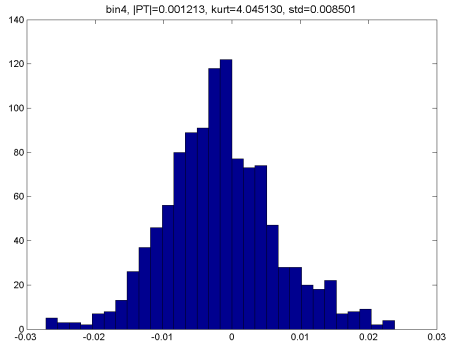

Figure 9: Histogram of ε for the 4th bin of P. Temperature, level 30.

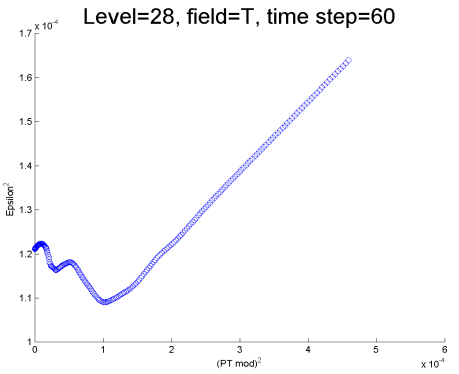

Figure 10: Conditional variance $\mathsf{Var}\,(\varepsilon \,|\, \mathsf{abs}(\mathsf{P}))$ for non-convective model errors. Temperature, level 28.

Table 1: Vertically averaged ratio of multiplicative to additive error st.dev.

	T	U	V
$\frac{\text{s.d. (mult)}}{\text{s.d. (add)}}$	0.5	0.5	0.8

5 Conclusions

Approach and preliminary stage results of a model-error objective estimation and modeling study are presented. The approach is based on the comparison of model's instantaneous tendencies with those of a higher-resolution "true model". The "true-model" tendencies are upscaled (coarse-grained) to the resolution of the model-in-question and subtracted from the (total) tendencies of the latter, yielding a proxy to the model-error fields. The goal of the study is a multivariate spatiotemporal stochastic model-error model. The model-error model is to be identified and estimated from a sample of the proxy model-error fields.

The model in question (the coarse-grid model) is COSMO with the horizontal resolution 2.2 km and 50 levels in the vertical.The high-resolution ("true") model is COSMO with the horizontal resolution 0.55 km, the same vertical grid, and a number of differences in the setup of the physical parameterizations. Preliminary results show that, first, errors in convective parameterization cannot and should not be tackled with the coarse-graining approach. Second, we found that non-convective model errors have both additive and multiplicative components. The additive model-error component is independent of the physical tendency and approximately Gaussian. The multiplicative model-error component is proportional to the physical tendency, with the multiplier being, again, approximately Gaussian. Third, the spatial structure of the non-convective model-error field is too complex to be modeled with a reasonably simple model-error model, especially in the lower troposphere and in the planetary boundary layer. This suggests that process-level model-error treatment is to be attempted.

Next steps (which are underway) are the following. Technically, (i) an even higher-resolution true model (fgm) is to be used, (ii) a more careful treatment of static fields (including orography), soil fields, and all initial fields in the two models is to be employed (so that their coarse-grid-resolved components be the same for the two models).Conceptually, (i) the starting point of the tendency-forecasts is to be computed using the true model (fgm) instead of the model (cgm) — to better represent the role of subgrid scales in the formation of model errors,(ii) errors due to different physical parameterizations are to be treated separately whenever possible, (iii) a spatial (horizontal and vertical), spatiotemporal, and multivariate (mutual dependencies between temperature, winds, etc.) aspects are to be addressed in the model-error stochastic modeling.

References

[1] J. Berner, S.-Y. Ha, J. Hacker, A. Fournier, and C. Snyder. Model uncertainty in a mesoscale ensemble prediction system: Stochastic versus multiphysics representations. *Mon. Wea. Rev.*, 139(6):1972–1995, 2011.

[2] R. Buizza, M. Miller, and T. Palmer. Stochastic representation of model uncertainties in the ECMWF ensemble prediction system. *Quart. J. Roy. Meteor. Soc.*, 125(560):2887–2908, 1999.

[3] H. Christensen, I. Moroz, and T. Palmer. Stochastic and perturbed parameter representations of model uncertainty in convection parameterization. *J. Atmos. Sci.*, 72(6):2525–2544, 2015.

[4] H. Christensen, S.-J. Lock, I. Moroz, and T. Palmer. Introducing independent patterns into the stochas-

tically perturbed parametrization tendencies (SPPT) scheme. *Quart. J. Roy. Meteor. Soc.*, 143(706): 2168–2181, 2017.

[5] D. L. Flack, S. L. Gray, R. S. Plant, H. W. Lean, and G. C. Craig. Convective-scale perturbation growth across the spectrum of convective regimes. *Monthly Weather Review*, 146(1):387–405, 2018.

[6] P. Houtekamer, H. L. Mitchell, and X. Deng. Model error representation in an operational ensemble Kalman filter. *Mon. Wea. Rev.*, 137(7):2126–2143, 2009.

[7] K. Kober and G. C. Craig. Physically based stochastic perturbations (PSP) in the boundary layer to represent uncertainty in convective initiation. *Journal of the Atmospheric Sciences*, 73(7):2893–2911, 2016.

[8] M. Leutbecher, S.-J. Lock, P. Ollinaho, S. T. Lang, G. Balsamo, P. Bechtold, M. Bonavita, H. M. Christensen, M. Diamantakis, E. Dutra, et al. Stochastic representations of model uncertainties at ECMWF: State of the art and future vision. *Quarterly Journal of the Royal Meteorological Society*, 143 (707):2315–2339, 2017.

[9] P. Ollinaho, S.-J. Lock, M. Leutbecher, P. Bechtold, A. Beljaars, A. Bozzo, R. M. Forbes, T. Haiden, R. J. Hogan, and I. Sandu. Towards process-level representation of model uncertainties: Stochastically perturbed parametrisations in the ECMWF ensemble. *Quart. J. Roy. Meteor. Soc.*, 143(702):408–422, 2017.

[10] D. Orrell, L. Smith, J. Barkmeijer, and T. Palmer. Model error in weather forecasting. *Nonlin. Proc. Geophys.*, 8(6):357–371, 2001.

[11] R. Plant and G. C. Craig. A stochastic parameterization for deep convection based on equilibrium statistics. *Journal of the Atmospheric Sciences*, 65(1):87–105, 2008.

[12] R. D. Richtmyer and K. W. Morton. Difference methods for initial-value problems. *Wiley*, 1967.

[13] G. Shutts. A kinetic energy backscatter algorithm for use in ensemble prediction systems. *Quart. J. Roy. Meteor. Soc.*, 131(612):3079–3102, 2005.

[14] G. Shutts. A stochastic convective backscatter scheme for use in ensemble prediction systems. *Quart. J. Roy. Meteor. Soc.*, 141(692):2602–2616, 2015.

[15] G. Shutts and A. C. Pallarès. Assessing parametrization uncertainty associated with horizontal resolution in numerical weather prediction models. *Phil. Trans. R. Soc. A*, 372(2018):20130284, 2014.

[16] G. Shutts and T. Palmer. Convective forcing fluctuations in a cloud-resolving model: Relevance to the stochastic parameterization problem. *Journal of Climate*, 20(2):187–202, 2007.

[17] M. Tsyrulnikov and D. Gayfulin. A limited-area spatio-temporal stochastic pattern generator for simulation of uncertainties in ensemble applications. *Meteorol. Zeitschrift*, 26(5):549–566, 2017.

[18] M. Tsyrulnikov and V. Gorin. Are atmospheric-model tendency errors perceivable from routine observations? *COSMO Newsletter No. 13*, pages 3–18, 2013.

[19] M. D. Tsyrulnikov. Stochastic modelling of model errors: A simulation study. *Quart. J. Roy. Meteor. Soc.*, 131(613):3345–3371, 2005.