Consortium

for

Small-Scale Modelling

Technical Report No. 29

A Stochastic Pattern Generator for ensemble applications

July 2016

DOI: 10.5676/DWD_pub/nwv/cosmo-tr_29

Deutscher Wetterdienst
MeteoSwiss
Ufficio Generale Spazio Aereo e Meteorologia
ΕΘΝΙΚΗ ΜΕΤΕΩΡΟΛΟΓΙΚΗ ΥΠΗΡΕΣΙΑ
Instytucie Meteorogii i Gospodarki Wodnej
Administratia Nationala de Meteorologie
ROSHYDROMET
Agenzia Regionale per la Protezione Ambientale del Piemonte
Agenzia Regionale per la Protezione Ambientale dell'Emilia-Romagna
Centro Italiano Ricerche Aerospaziali
Amt für GeoInformationswesen der Bundeswehr

www.cosmo-model.org

Editor: Massimo Milelli, ARPA Piemonte

A Stochastic Pattern Generator

for ensemble applications

M. Tsyrulnikov and D. Gayfulin

Hydrometeorological Center of Russia
11-13 B. Predtechensky Lane
123242 Moscow
Russia

Contents

1 Abstract

A generator of spatio-temporal pseudo-random Gaussian fields that satisfy the "proportionality of scales" property (Tsyroulnikov 2001) is presented. The generator is based on a third-order in time stochastic differential equation with a pseudo-differential spatial operator defined on a limited area 2D or 3D domain in the Cartesian coordinate system. The generated pseudo-random fields are homogeneous and isotropic in space-time. The correlation functions in any spatio-temporal direction belong to the Matérn class. The spatio-temporal correlations are non-separable. A spectral-space numerical solver is implemented and accelerated exploiting properties of real-world geophysical fields, in particular, smoothness of their spatial spectra. The generator is designed to simulate additive or multiplicative, or other spatio-temporal perturbations that represent uncertainties in numerical prediction models in geophysics. The generator is tested with the COSMO model as a source of additive spatio-temporal perturbations to the forecast model fields.

2 Introduction

2.1 Stochastic dynamic prediction

Since the works of [1] and [2], we know that accounting for the uncertainty in the initial forecast fields can improve weather (and other geophysical) predictions. Assigning a probability distribution for the truth at the start of the forecast (instead of using deterministic initial data) and attempting to advance this distribution in time according to the dynamic (forecast) model is called stochastic dynamic prediction.

The advantage of the stochastic dynamic prediction paradigm is twofold. First, the resulting forecast probability distribution provides a valuable measure of the *uncertainty* in the prediction, leading to probabilistic forecasting and flow-dependent background-error statistics in data assimilation. Second, for a nonlinear physical model, switching from the deterministic forecast to the mean of the forecast probability distribution improves the mean-square accuracy of the prediction, i.e. can improve the deterministic forecasting.

2.2 Model errors

Since [3], we realize that not only uncertainties in the initial data (analysis errors) matter, forecast model (including boundary conditions) imperfections also play an important role. Simulation of model errors is the subject of this study, so we define them now. Let the forecast model be of the form

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

where t is time, $\mathbf{x}$ is the vector that represents the (discretized) state of the system, and $\mathbf{F}$ is the model (forecast) operator. The imperfection of the model Eq.(1) means that the (appropriately discretized) truth does *not* exactly satisfy this equation. The discrepancy is called the model error [e.g. 4]:

$$\boldsymbol{\xi}^t = \mathbf{F}(\mathbf{x}^t) - \frac{\mathrm{d}\mathbf{x}^t}{\mathrm{d}t}. \tag{2}$$

The true model error $\boldsymbol{\xi}^t$ is normally unknown. In order to include model errors in the stochastic dynamic prediction paradigm, one *models* $\boldsymbol{\xi}^t(t)$ as a *random process*, $\boldsymbol{\xi}(t)$, or, in

other words, as a spatio-temporal random field $\xi(t, \mathbf{s})$ (where $\mathbf{s}$ is the spatial vector). The probability distribution of $\boldsymbol{\xi}(t)$ is assumed to be known.

Rearranging the terms in Eq.(2), and replacing the unknown $\boldsymbol{\xi}^t$ with its stochastic counterpart $\boldsymbol{\xi}$, we realize that the resulting model of truth is the stochastic dynamic equation

$$\frac{\mathrm{d}\mathbf{x}}{\mathrm{d}t} = \mathbf{F}(\mathbf{x}) - \boldsymbol{\xi}. \tag{3}$$

Thus, the extended stochastic dynamic prediction (or modeling) paradigm requires two input probability distributions (that of initial errors and that of model errors) and aims to transform them to the output (forecast) probability distribution.

2.3 Ensemble prediction

Stochastic dynamic modeling of complex geophysical systems is hampered by their high dimensionality and non-linearity. The output probability distribution appears to be in most cases analytically intractable. An affordable approximate solution is provided by the Monte-Carlo method called in geophysics *ensemble prediction*.

In ensemble prediction, the input uncertainties (i.e. initial and model errors) are represented by simulated *pseudo-random draws* from the respective probability distributions. A relatively small affordable number of these draws are fed to the forecast model giving rise to an *ensemble* of predictions (forecasts). Members of this ensemble (called ensemble forecasts) are solutions to Eq.(3) with $\boldsymbol{\xi}$ replaced with simulated pseudo-random draws from the model-error probability distribution. The ensemble forecasts start from the initial data perturbed according to the probability distribution of initial errors.

If initial and model errors are sampled from the correct respective distributions, then the forecast ensemble members are draws from the correct probability distribution of the truth given all available external data (initial and boundary conditions). This mathematically justifies the ensemble prediction principle. From the practical perspective, members of the forecast ensemble can be interpreted as "potential truths" consistent with all available information.

In what follows, we concentrate on the model error field $\xi(t, \mathbf{s})$. We briefly review existing models for $\xi(t, \mathbf{s})$ and then present our stochastic pattern generator, whose goal is to simulate pseudo-random draws of $\xi(t, \mathbf{s})$ from a meaningful and flexible distribution.

2.4 Practical model error modeling

In meteorology, our knowledge of the actual model error probability distribution is scarce. Justified stochastic model-error models are still to be devised and verified. In the authors' opinion, the best way to stochastically represent spatio-temporal forecast-model-error fields is to treat each error source separately, so that, say, each physical parametrization is accompanied with a spatio-temporal stochastic model of its uncertainty. Or, even better, to completely switch from deterministic physical parameterizations to stochastic ones. There is a growing number of such developments [see 5, for a review], but the problem is so complex that we cannot expect it to be solved in the near future. Its solution is further hampered by the fact that the existing meteorological observations are too scarce and too inaccurate for model errors to be objectively identified by comparison with measurement data with satisfactory accuracy [6].

As a result, in meteorology *ad-hoc* model-error models are in wide use. The existing approaches include generating pseudo-random additive or multiplicative perturbations of the right-hand sides of the model equations [e.g. 7, 8] in the course of forecast. These two model-error modeling techniques as well as Stochastic Kinetic Energy Backscatter schemes [9] require a pseudo-random spatio-temporal field as a stochastic input. Stochastic parameterization schemes can also demand such fields [see e.g. 10].

The simplest non-constant pseudo-random field is the white noise, i.e. the uncorrelated in space and time random field. The white noise is the default forcing in stochastic differential equations [e.g. 11, 12]. Its advantage is the complete absence of any spatio-temporal structure, it is a pristine source of stochasticity. But in model-error modeling, this lack of structure precludes its direct use as an additive or multiplicative perturbation field because model errors are related to the weather pattern and so should be correlated (dependent) both in space and time. Tsyrulnikov [13] showed in a simulation study that model errors can exhibit complicated spatio-temporal behavior.

A correlated pseudo-random spatio-temporal field can be easily computed by generating independent random numbers at points of a *coarse* spatio-temporal grid and then assigning each of them to all model grid points within the respective coarse-grid cell [8]. As a result, the model-grid field becomes correlated in space and time. The decorrelation space and time scales are, obviously, defined by the respective coarse-grid spacings. This determines the choice of the coarse grid, e.g. in [8] the spatial grid spacing was about 1000 km and the temporal one 6 hours). This technique is extremely simple but it suffers from two flaws.

First, the resulting model-grid field appears to be discontinuous and inhomogeneous. Second, the spatio-temporal structure of the field is not scale dependent, that is, the resulting temporal length scales do not depend on the respective spatial scales. In reality, longer spatial scales "live longer" than shorter spatial scales, which "die out" quicker. This 'proportionality of scales' is widespread in geophysical fields [see 14, and references therein] and other media, [e.g. 15, p.129], so we believe this property should be represented by model-error models. Note also that the "proportionality of scales" is a special case of the *non-separability* of spatio-temporal covariances. For a critique of simplistic separable space-time covariance models, see [16, 17, 18] and Appendix A in this report.

Another popular pseudo-random field generation technique in space and time employs a spectral transform in space and then imposes independent temporal auto-regressions for the coefficients of the spectral expansion [19, 9, 20, 21, 22]. This technique is more general and produces homogeneous fields, but the above implementations use *the same time scale* for all spatial wavenumbers so that there are still no space-time interactions in the generated spatio-temporal fields.

In this report, we propose and test a spatio-temporal Stochastic (pseudo-random) Pattern Generator (SPG) that accounts for the above "proportionality of scales" and imposes meaningful space-time interactions. The SPG operates on a limited-area domain. It is based on a (spectral-space) solution to a stochastic partial differential equation, more precisely, to a stochastic differential equation in time with a pseudo-differential spatial operator. In what follows, we present the technique, examine properties of the resulting spatio-temporal fields in 2D and 3D spatial domains, and describe the numerical scheme. We start with a first-order in time SPG model. Then we show that this model needs to be modified in order to meet all the criteria we impose. Eventually, we end up with a third-order in time model. The technique is implemented as a Fortran program freely available from https://github.com/gayfulin/SPG.

3 SPG: Requirements

The general requirements are:

- The SPG should produce stationary in time and homogeneous and isotropic in space Gaussian pseudo-random fields $\xi(t, \mathbf{s})$ in 3D and 2D spatial domains.
- The SPG should be fast enough so that it does not significantly slow down the forecast model computations.
- Variance as well as spatial and temporal length scales of $\xi(t, \mathbf{s})$ are to be tunable.

We also impose more specific requirements:

1. The random field $\xi(t, \mathbf{s})$ should have finite variance and continuous realizations (sample paths).
2. The spatio-temporal covariances should obey the "proportionality of scales" principle: larger (shorter) spatial scales should be associated with larger (shorter) temporal scales [14].
3. The SPG *ansatz* should be flexible enough to allow for practicable solutions in both physical space and spectral space.

4 The proposed solution

We select the general class of *linear evolutionary stochastic partial differential equations* (SPDE) as a starting point in the development of the SPG. This choice is motivated by the flexibility of this class of spatio-temporal models [e.g. 23]. In particular, for an SPDE, it is relatively easy to introduce inhomogeneity in space and time as well as local anisotropy—either by changing coefficients of the spatial operator or by changing local properties of the driving noise. One can also produce non-Gaussian fields by making the random forcing non-Gaussian [e.g. 24, 25]. Physical-space discretizations of SPDEs lead to *sparse* matrices, which give rise to fast numerical algorithms. If an SPDE has constant coefficients, then it can be efficiently solved using spatial spectral-space expansions.

In this study, we develop the SPG that relies on a spatio-temporal stochastic model with *constant coefficients* so that both physical-space and spectral-space solvers can be employed. To facilitate the spectral-space solution, the general strategy is to define the SPG model on a standardized spatial domain. The operational pseudo-random fields are then produced by mapping the generated fields from the standardized domain to the forecast-model domain. In 3D, the standardized spatial domain is chosen to be triply periodic: the three-dimensional (3D) unit torus. In 2D, the standardized domain is the 2D unit torus. The 3D and 2D cases are distinguished by the dimensionality $d = 2$ or $d = 3$ in what follows. To simplify the presentation, the default dimensionality will be $d = 3$.

5 Tentative first-order SPG model

5.1 Physical-space model

The random field in question $\xi(t, \mathbf{s})$ is a function of the time coordinate t and the space vector $\mathbf{s} := (x, y, z)$, where (x, y, z) are the three spatial coordinates. Each of the spatial coordinates belongs to the the unit circle $\mathbb{S}^1$, so that $\mathbf{s}$ is on the unit torus $\mathbb{T}^3 \equiv \mathbb{S}^1 \times \mathbb{S}^1 \times \mathbb{S}^1$ ($\mathbb{T}^2$ in the 2D case).

We start with the simplest general form of a first-order Markov model:

$$\frac{\partial \xi(t, \mathbf{s})}{\partial t} + A\,\xi(t, \mathbf{s}) = \alpha(t, \mathbf{s}), \tag{4}$$

where A is the spatial linear operator to be specified and α is the driving noise postulated to be homogeneous in space and white in time.

The SPG is required to be fast, so we choose A to be a *differential* operator (because, as we noted, in this case a physical-space discretization of A gives rise to a very sparse matrix).

Further, since we wish $\xi(t, \mathbf{s})$ to be homogeneous and isotropic in space, we define A to be a polynomial of the negated spatial Laplacian:

$$A := P(-\Delta) := \sum_{j=0}^{q} c_j (-\Delta)^j, \tag{5}$$

where $P(x)$ is the polynomial and q its degree (a positive integer). We will refer to q as the spatial order of the SPG model. Note that the negation of the Laplacian is convenient because $(-\Delta)$ is a non-negative definite operator.

The coefficients c_j are selected to ensure that the spatial operator $\sum c_j(-\Delta)^j$ has only positive eigen-values not close enough to 0, which can be achieved if all $c_j \geq 0$ and $c_0 > 0$; this guarantees *stability* of the SPG.

The model Eq.(5) appears to be too rich for the purposes of the SPG at the moment, so in what follows we employ an even more reduced (but still quite flexible) form

$$A = P(-\Delta) := \mu(1 - \lambda^2\Delta)^q, \tag{6}$$

where μ and λ are positive real parameters.

So, we start with the following SPG equation:

$$\frac{\partial \xi(t, \mathbf{s})}{\partial t} + \mu(1 - \lambda^2\Delta)^q\,\xi(t, \mathbf{s}) = \alpha(t, \mathbf{s}). \tag{7}$$

5.2 Spectral-space model

On the torus $\mathbb{T}^d$, a Fourier series is an expansion in the basis functions $\mathrm{e}^{\mathrm{i}(\mathbf{k},\mathbf{s})} \equiv \mathrm{e}^{\mathrm{i}(\mathrm{mx}+\mathrm{ny}+\mathrm{lz})}$, where the wavevector $\mathbf{k}$ is the triple of integer wavenumbers, $\mathbf{k} := (m, n, l)$. We perform the Fourier decomposition for both $\alpha(t, \mathbf{s})$ and $\xi(t, \mathbf{s})$,

$$\alpha(t, \mathbf{s}) = \sum_{\mathbf{k} \in \mathbb{Z}^d} \tilde{\alpha}_{\mathbf{k}}(t) \mathrm{e}^{\mathrm{i}(\mathbf{k},\mathbf{s})} \tag{8}$$

and

$$\xi(t,\mathbf{s}) = \sum_{\mathbf{k}\in\mathbb{Z}^d} \tilde{\xi}_{\mathbf{k}}(t)\mathrm{e}^{\mathrm{i}(\mathbf{k},\mathbf{s})} \tag{9}$$

(where $\mathbb{Z}$ denotes the set of integer numbers) and substitute these expansions into Eq.(7). Noting that the application of $P(-\Delta)$ to $\mathrm{e}^{\mathrm{i}(\mathbf{k},\mathbf{s})}$ returns $P(\mathbf{k}^2)\,\mathrm{e}^{\mathrm{i}(\mathbf{k},\mathbf{s})}$, recalling that $P(-\Delta)$ is defined by Eq.(6), and using orthogonality of the basis functions, we obtain that Eq.(7) decouples into the set of ordinary stochastic differential equations [OSDE, e.g. 11, 12] in time:

$$\frac{\mathrm{d}\tilde{\xi}_{\mathbf{k}}}{\mathrm{d}t} + a_{\mathbf{k}}\,\tilde{\xi}_{\mathbf{k}}(t) = \tilde{\alpha}_{\mathbf{k}}(t), \tag{10}$$

where

$$a_{\mathbf{k}} := \mu(1+\lambda^2\mathbf{k}^2)^q. \tag{11}$$

From the postulated homogeneity of α in space, its spectral-space coefficients $\tilde{\alpha}_{\mathbf{k}}(t)$ are mutually independent. This is well known for random fields on $\mathbb{R}^d$ (where spectra are continuous), see e.g. Chapter 2 in [26] or section 8 in [27], and can be directly verified in our case of the fields on the torus (where spectra are discrete). Therefore, for different wavevectors $\mathbf{k}$, the resulting spectral-space equations, Eqs.(10)–(11), are probabilistically completely *independent* from each other. This greatly simplifies the solution of the SPG equations because instead of handling the complicated SPDE Eq.(7) we have to solve a number of independent simple OSDEs Eq.(10).

Further, from the postulated whiteness of α in time, all $\tilde{\alpha}_{\mathbf{k}}(t)$ are white in time random processes (see Appendix B.4). So, we may write

$$\tilde{\alpha}_{\mathbf{k}}(t) = \sigma_{\mathbf{k}}\,\Omega_{\mathbf{k}}(t), \tag{12}$$

where $\Omega_{\mathbf{k}}(t)$ are independent *standard* white noises, i.e. derivatives of independent standard Wiener processes $W_{\mathbf{k}}(t)$ such that

$$\Omega_{\mathbf{k}}(t)\mathrm{d}t = \mathrm{d}W_{\mathbf{k}}(t) \tag{13}$$

and $\sigma_{\mathbf{k}}$ are *intensities* of the white-noise processes. In space, $\sigma_{\mathbf{k}}^2$ is proportional to the spatial spectrum of the driving noise $\alpha(t,\mathbf{s})$, see Eq.(8), Eq.(12), and Appendix B. If $\sigma_{\mathbf{k}} = \mathsf{const}$, then $\alpha(t,\mathbf{s})$ is white both in space and time, otherwise $\alpha(t,\mathbf{s})$ is a colored in space and white in time noise.

Thus, the first-order SPG model reduces to a series of OSDEs

$$\mathrm{d}\tilde{\xi}_{\mathbf{k}} + \mu(1+\lambda^2\mathbf{k}^2)^q\,\tilde{\xi}_{\mathbf{k}}\,\mathrm{d}t = \sigma_{\mathbf{k}}\,\mathrm{d}W_{\mathbf{k}}. \tag{14}$$

For practical purposes the series is truncated, so that $\mathbf{k} \equiv (m,n,l)$ is limited: $|m| < m_{max}$, $|n| < n_{max}$, and $|l| < l_{max}$, where m_{max}, n_{max}, and l_{max} are the truncation limits. If not otherwise stated, all the truncation limits are the same and denoted by n_{max}.

5.3 Stationary spectral-space statistics

Equation (14) is a first-order OSDE with constant coefficients sometimes called the Langevin equation, see e.g. [12] or Example 4.12 in [11]. Its generic form is

$$\mathrm{d}\eta + a\eta\,\mathrm{d}t = \sigma\mathrm{d}W, \tag{15}$$

where $\eta(t)$ is the random process in question, a and σ are constants, and $W(t)$ is the standard Wiener process. The solution to Eq.(15) is known as the Ornstein-Uhlenbeck random process, whose stationary (steady-state) temporal covariance function is

$$B_\eta(t) = \frac{\sigma^2}{2a}\,\mathrm{e}^{-\mathrm{a}|\mathrm{t}|} \tag{16}$$

[e.g. 11, Example 4.12]. From Eq.(16), it is clear that a has the meaning of the inverse temporal length scale $\tau := 1/a$.

Now, consider the stationary covariance function of the elementary random process $\tilde{\xi}_{\mathbf{k}}(t)$,

$$\mathsf{E}\,\tilde{\xi}_{\mathbf{k}}(t_0)\cdot\overline{\tilde{\xi}_{\mathbf{k}}(t_0+t)} = b_{\mathbf{k}}\cdot C_{\mathbf{k}}(t), \tag{17}$$

where $b_{\mathbf{k}}$ is the variance and $C_{\mathbf{k}}(t)$ the correlation function. According to Eq.(9), $\tilde{\xi}_{\mathbf{k}}$ is the spatial spectral component of the random field in question $\xi(t,\mathbf{s})$. Therefore $b_{\mathbf{k}} = \mathsf{Var}\,\tilde{\xi}_{\mathbf{k}}$ is called the *spatial spectrum* of $\xi(t,\mathbf{s})$. From Eqs.(14) and (16), we have

$$b_{\mathbf{k}} = \frac{\sigma_{\mathbf{k}}^2}{2\mu(1+\lambda^2\mathbf{k}^2)^q} \tag{18}$$

and

$$C_{\mathbf{k}}(t) = \mathrm{e}^{-\frac{|\mathrm{t}|}{\tau_{\mathbf{k}}}}, \tag{19}$$

where

$$\tau_{\mathbf{k}} := \frac{1}{a_{\mathbf{k}}} = \frac{1}{\mu(1+\lambda^2\mathbf{k}^2)^q} \tag{20}$$

is the temporal length scale associated with the spatial wavevector $\mathbf{k}$.

Note that by the spectrum (e.g. $b_{\mathbf{k}}$), we always mean the *modal* spectrum, i.e. the variance associated with a single basis function (a single wavevector $\mathbf{k}$); the modal spectrum is not to be confused with the variance (or energy) spectrum.

5.4 Physical-space statistics

In the stationary regime (i.e. after an initial transient period has passed), the above independence of the spectral random processes $\tilde{\xi}_{\mathbf{k}}(t)$ (see section 5.2) implies that the random field $\xi(t,\mathbf{s})$ is spatio-temporally *homogeneous*, i.e. invariant under shifts in space and time:

$$\mathsf{E}\,\xi(t,\mathbf{s})\cdot\overline{\xi(t+\Delta t,\mathbf{s}+\Delta\mathbf{s})} = B(\Delta t,\Delta\mathbf{s}), \tag{21}$$

where

$$B(t,\mathbf{s}) = \sum_{\mathbf{k}} b_{\mathbf{k}}\,C_{\mathbf{k}}(t)\,\mathrm{e}^{\mathrm{i}(\mathbf{k},\mathbf{s})}. \tag{22}$$

In particular, the spatial covariance function is

$$B(\mathbf{s}) = B(t=0,\mathbf{s}) = \sum_{\mathbf{k}} b_{\mathbf{k}}\,\mathrm{e}^{\mathrm{i}(\mathbf{k},\mathbf{s})}, \tag{23}$$

where it is seen that the spatial spectrum $b_{\mathbf{k}}$ is the Fourier transform of the spatial covariance function $B(\mathbf{s})$. Finally, the variance is

$$\mathsf{Var}\,\xi = B(t=0,\mathbf{s}=\mathbf{0}) = \sum_{\mathbf{k}} b_{\mathbf{k}}. \tag{24}$$

5.5 Imposing the SPG requirements

In section 3, we have formulated requirements 1–3 the SPG model should satisfy. In more specific terms, they imply the following three conditions.

Requirement 1 states, in particular, that

$$\mathsf{Var}\,\xi < \infty. \tag{25}$$

Next, the more precise formulation of requirement 2 states that that for large k, the temporal length scale $\tau_{\mathbf{k}}$ should be inversely proportional to k:

$$\tau_{\mathbf{k}} \sim \frac{1}{k} \quad \text{as} \quad k \to \infty. \tag{26}$$

Finally, requirement 3 entails that the driving noise $\alpha(t, \mathbf{s})$ should be white not only in time but also in space:

$$\sigma_{\mathbf{k}} = \sigma = \mathsf{const}. \tag{27}$$

This is because the physical-space simulation of the discretized white noise is cheap (since its grid-point values are just independent zero-mean Gaussian pseudo-random variables), whereas to simulate a non-white noise requires building a model for the noise and solving its equations, which is normally expensive (and complicates the design of the SPG).

In the rest of this section we show that these three conditions cannot be simultaneously satisfied for the first-order SPG model Eq.(7).

5.5.1 "Proportionality of scales" implies that $q = \frac{1}{2}$

We start with condition Eq.(26). Substituting Eq.(20) into Eq.(26) yields

$$(1 + \lambda^2 k^2)^q \sim k \quad \text{as k} \to \infty \quad \Leftrightarrow \quad \boxed{q = \frac{1}{2}}. \tag{28}$$

Note that here and elsewhere, boxed equations are the ones that present the key aspects of our final SPG model.

With $q = \frac{1}{2}$, the model's spatial operator A becomes (see Eq.(6))

$$\boxed{A = \mu(1 - \lambda^2 \Delta)^{\frac{1}{2}} \equiv \mu\sqrt{1 - \lambda^2 \Delta}}. \tag{29}$$

This is a pseudo-differential operator [e.g. 28] with the *symbol*

$$\boxed{a(k) := \mu\sqrt{1 + \lambda^2 k^2}}, \tag{30}$$

so that the action of A on the test function $\varphi(\mathbf{s})$ is defined as follows. First, we Fourier transform $\varphi(\mathbf{s})$ getting $\{\tilde{\varphi}_{\mathbf{k}}\}$. Then, $\forall \mathbf{k} \in \mathbb{Z}^d$, we multiply $\tilde{\varphi}_{\mathbf{k}}$ by the symbol $a(k)$. Finally, we perform the backward Fourier transform of $\{a(k)\tilde{\varphi}_k\}$ retrieving the result, the function $(A\varphi)(\mathbf{s})$.

Obviously, there is no problem with the above fractional negated and shifted Laplacian in spectral space (as its action on test functions is well defined, see the previous paragraph). Importantly, the pseudo-differential operator A appears to have nice properties in physical space, too. Specifically, A can be approximated by a discrete-in-space operator which is

a *very sparse matrix*, see Appendix C. So, in both spectral space and physical space, the resulting operator A with the fractional degree $q = \frac{1}{2}$ is numerically tractable.

The spectral-space SPG model Eq.(14) becomes

$$\mathrm{d}\tilde{\xi}_{\mathbf{k}} + \mu\sqrt{1+\lambda^2\mathbf{k}^2}\,\tilde{\xi}_{\mathbf{k}}\,\mathrm{d}t = \sigma_{\mathbf{k}}\,\mathrm{d}W_{\mathbf{k}}. \tag{31}$$

The resulting spatial spectrum is

$$b_k = \frac{\sigma_{\mathbf{k}}^2}{2\mu\sqrt{1+\lambda^2 k^2}}. \tag{32}$$

The SPG model becomes

$$\frac{\partial\xi(t,\mathbf{s})}{\partial t} + \mu\sqrt{1-\lambda^2\Delta}\cdot\xi(t,\mathbf{s}) = \alpha(t,\mathbf{s}). \tag{33}$$

Below, we will see that the choice $q = \frac{1}{2}$ implies that besides the "proportionality of scales", the generated spatio-temporal field has other nice properties.

5.5.2 For $\mathsf{Var}\,\xi(t,\mathbf{s})$ to be finite, $\alpha(t,\mathbf{s})$ needs to be a red noise in space

Next, we turn to condition Eq.(25) in section 5.5. Substituting Eq.(32) into Eq.(24) yields

$$\mathsf{Var}\,\xi = \frac{1}{2\mu}\sum\frac{\sigma_k^2}{\sqrt{1+\lambda^2k^2}}. \tag{34}$$

Assuming that σ_k is smoothly varying, we may approximate the sum in this equation with the integral (where $\sigma(k) = \sigma(\mathbf{k})$ for integer wavenumbers), getting

$$\mathsf{Var}\,\xi \approx \int_{\mathbb{R}^d}\frac{\sigma_k^2}{\sqrt{1+\lambda^2k^2}}\,\mathrm{d}\mathbf{k} \propto \int_{\mathbb{R}}\frac{\sigma_k^2}{\sqrt{1+\lambda^2k^2}}\,k^{d-1}\,\mathrm{d}k. \tag{35}$$

To check the convergence of the integral in Eq.(35), we examine the $k\to\infty$ limit (where $\sqrt{1+\lambda^2k^2}\sim k$). As we know, the integral of this kind converges if the integrand decays faster than $\frac{1}{k^{1+\epsilon}}$ with some $\epsilon > 0$. This is the case whenever

$$\sigma_k^2 \sim \frac{1}{k^{d-1+\epsilon}} \quad \text{as} \quad k\to\infty. \tag{36}$$

So, to satisfy the requirement $\mathsf{Var}\,\xi < \infty$, the spectrum σ_k^2 needs to be a *decaying* function of its argument k. This clearly contradicts to white-in-space-driving-noise condition Eq.(27).

5.5.3 Implications for the SPG design

As we have just seen, the conditions Eqs.(25)–(27) cannot be met by the first-order SPG model. So, the model Eq.(7) is to be somehow changed. The solution is to increase the temporal order of the stochastic model.

6 Higher-order in time model

6.1 Motivation and formulation

Equation (36) shows that with the first order SPG model, we could met both the the finite-variance condition Eq.(25) and the proportionality-of-scales condition Eq.(26) if we specified

a *red* in space (i.e. with the decaying spectrum) driving noise α. The red noise can be obtained by the application of a linear integral operator *with the decaying symbol* to the white noise. The problem here is that with the quite rapidly decaying red-noise spectrum Eq.(36), the support of the physical-space integral operator will be rather large, resulting in a computationally expensive numerical scheme.

The idea is, instead of introducing of an expensive *integral* operator with a decaying symbol to the *right-hand side* of the model, Eqs.(4) or (7) or (33), to introduce a *differential* operator (with a growing symbol) to the *left-hand side* of the SPG equation. The simplest way of doing so is to raise the operator $\partial/\partial t + A$, which already acts in our model on ξ in the l.h.s. of Eq.(4), to a power. This implies that the *temporal order* of the SPG model increases:

$$\boxed{\left(\frac{\partial}{\partial t} + \mu\sqrt{1-\lambda^2\Delta}\right)^p \xi(t,\mathbf{s}) = \alpha(t,\mathbf{s})}, \tag{37}$$

where p is the temporal order of the modified SPG model (a positive integer) and α is white both in time and in space. In spectral space, the model Eq.(37) becomes, obviously,

$$\boxed{\left(\frac{\mathrm{d}}{\mathrm{d}t} + \mu\sqrt{1+\lambda^2 k^2}\right)^p \tilde{\xi}_{\mathbf{k}}(t) = \sigma\,\Omega_{\mathbf{k}}(t)}, \tag{38}$$

where, we recall, $\Omega_{\mathbf{k}}(t)$ are mutually independent standard white noises.

The resulting higher-order SPG model satisfies condition Eq.(27) by construction. Now, we show that the model Eq.(37) can be defined to satisfy the two remaining conditions, Eqs.(25)–(26).

6.2 Stationary spectral-space statistics

For each $\mathbf{k}$, Eq.(38) is a pth-order in time OSDE. In Appendix D, we examine properties the generic pth-order OSDE, specifically, the steady-state statistics of its solution. Using Table 2 in Appendix D, we can write down the stationary variance $b_{\mathbf{k}}$ and the temporal correlation function $C_{\mathbf{k}}(t)$ of the solution to Eq.(38), the process $\tilde{\xi}_{\mathbf{k}}(t)$:

$$b_{\mathbf{k}} \propto \frac{\sigma^2}{\mu^{2p-1}(1+\lambda^2 k^2)^{p-\frac{1}{2}}} \tag{39}$$

and

$$C_{\mathbf{k}}(t) = \left(1 + \frac{|t|}{\tau_{\mathbf{k}}} + r_2\frac{|t|^2}{\tau_{\mathbf{k}}^2} + \cdots + r_{p-1}\frac{|t|^{p-1}}{\tau_{\mathbf{k}}^{p-1}}\right)\mathrm{e}^{-\frac{|t|}{\tau_{\mathbf{k}}}}. \tag{40}$$

Here $r_2, \ldots, r_{p-1}$ are real numbers (given for $p = 1, 2, 3$ in Table 2) and $\tau_{\mathbf{k}}$ is the temporal length scale associated with the spatial wavevector $\mathbf{k}$:

$$\boxed{\tau_{\mathbf{k}} = \frac{1}{\mu\sqrt{1+\lambda^2 k^2}}}. \tag{41}$$

Specifically, for the temporal order $p = 2$, we have

$$b_{\mathbf{k}}|_{p=2} = \frac{\sigma^2}{4\mu^3(1+\lambda^2 k^2)^{\frac{3}{2}}} \tag{42}$$

and

$$C_{\mathbf{k}}(t)|_{p=2} = \left(1 + \frac{|t|}{\tau_{\mathbf{k}}}\right)\mathrm{e}^{-\frac{|t|}{\tau_{\mathbf{k}}}}. \tag{43}$$

For the temporal order $p = 3$, we have

$$\boxed{b_{\mathbf{k}}|_{p=3} = \frac{3\sigma^2}{16\mu^5(1+\lambda^2k^2)^{\frac{5}{2}}}} \tag{44}$$

and

$$\boxed{C_{\mathbf{k}}(t)|_{p=3} = \left(1 + \frac{|t|}{\tau_{\mathbf{k}}} + \frac{1}{3}\,\frac{|t|^2}{\tau_{\mathbf{k}}^2}\right) \mathrm{e}^{-\frac{|t|}{\tau_{\mathbf{k}}}}.} \tag{45}$$

From Eq.(41), it is seen that the "proportionality of scales" condition Eq.(26) is indeed satisfied because $\tau_{\mathbf{k}}$ is indeed inversely proportional to k for large k. In order to achieve the desired dependency of $\tau_{\mathbf{k}}$ not only on k (which we already have from Eq.(41)), but also on λ (the greater is λ the greater should be $\tau_{\mathbf{k}}$), we parameterize μ as

$$\mu = \frac{U}{\lambda}, \tag{46}$$

where $U > 0$ is the velocity-dimensioned tuning parameter. With this parameterization, λ affects both the spatial length scale of ξ (due to Eq.(39)) and the temporal length scale (thanks to Eq.(41)). In contrast, U affects only the temporal length scale.

6.3 Finite-variance criterion

Substituting b_k from Eq.(39) into Eq.(24), approximating the sum over the wavevectors by the integral (as in Eq.(35)), and exploiting the isotropy of the integrand yields

$$\mathsf{Var}\,\xi \approx \mathsf{const} \cdot \int_0^\infty \frac{\sigma^2}{(1+\lambda^2k^2)^{p-\frac{1}{2}}}\, k^{d-1}\,\mathrm{d}k, \tag{47}$$

so that $\mathsf{Var}\,\xi < \infty$, i.e. the finite-variance condition Eq.(25) is met, whenever

$$\boxed{p > \frac{d+1}{2}}. \tag{48}$$

Thus, in the higher-order-in-time model Eq.(37) we can rely on the white in space and time driving noise $\alpha(t, \mathbf{s})$ *provided that the temporal order is large enough*: in 2D, it is required that $p \geq 2$ whilst in 3D, we have to set up $p \geq 3$.

6.4 Isotropy in space-time

In this subsection, we show that, remarkably, $q = \frac{1}{2}$ is the unique spatial order for which the field $\xi(t, \mathbf{s})$ appears to be isotropic is space-time. In particular, the shape of the correlation function is the same in any spatial or temporal or any other direction in the spatio-temporal domain $\mathbb{T}^d \times \mathbb{R}$.

6.4.1 Spatial isotropy

First, we note that spatial isotropy of the random field ξ is invariance of its covariance function $B(\mathbf{s})$ under rotations. If we were in $\mathbb{R}^d$ rather than on $\mathbb{T}^d$, isotropy of $B(\mathbf{s}) = B(s)$, where $s := |\mathbf{s}|$, would be equivalent to isotropy of its Fourier transform (spectrum) $b(\mathbf{k})$, so that the latter would be dependent only on the modulus of the wavevector $\mathbf{k}$, i.e. the

total wavenumber $k := |\mathbf{k}| = \sqrt{m^2 + n^2 + l^2}$. On the torus, spectra are discrete, i.e. m, n, l take only *integer* values, so, strictly speaking, $b(\mathbf{k})$ cannot be isotropic there. To avoid this technical difficulty, we resort to the device used in sections 5.5.2 and 6.3, the approximation of a sum over the wavevectors by the integral.

Specifically, we assume that $b(\mathbf{k})$ is smooth enough (which is tantamount to the assumption that $B(\mathbf{s})$ decays on length scales much smaller than the domain's extents) for the validity of the approximation

$$B(\mathbf{s}) = \sum_{\mathbf{k}\in\mathbb{Z}^{\mathbf{d}}} b_{\mathbf{k}}\, \mathrm{e}^{\mathrm{i}(\mathbf{k},\mathbf{s})} \approx \int_{\mathbb{R}^{\mathrm{d}}} \mathrm{b}(\mathbf{k})\, \mathrm{e}^{\mathrm{i}(\mathbf{k},\mathbf{s})} \mathrm{d}\mathbf{k}, \tag{49}$$

where $b(\mathbf{k})$ is a smooth function of the real vector argument $\mathbf{k}$ such that $\forall \mathbf{k} \in \mathbb{Z}^{\mathbf{d}}, \mathbf{b}(\mathbf{k}) = \mathbf{b_k}$. From Eq.(49), it is obvious that $B(\mathbf{s})$ is indeed approximately invariant under rotations because so is $b(\mathbf{k})$, see Eq.(39).

In the theoretical analysis in this section, we will rely on the approximation Eq.(49) and thus assume that the "spectral grid" is dense enough for the spatial spectra to be treated as continuous ones.

6.4.2 Spatio-temporal spectra

Consider the OSDE Eq.(38) in the stationary regime. Following [27, section 8], the stationary random process can be spectrally represented as the stochastic integral

$$\tilde{\xi}_{\mathbf{k}}(t) = \int_{\mathbb{R}} \mathrm{e}^{\mathrm{i}\omega \mathrm{t}}\, \mathrm{Z}_{\mathbf{k}}(\mathrm{d}\omega), \tag{50}$$

where ω is the angular frequency (temporal wavenumber) and Z is the orthogonal stochastic measure such that

$$\mathsf{E}\,|Z_{\mathbf{k}}(\mathrm{d}\omega)|^2 = b_{\mathbf{k}}(\omega)\,\mathrm{d}\omega, \tag{51}$$

where $b_{\mathbf{k}}(\omega)$ is the spectral density of the process $\tilde{\xi}_{\mathbf{k}}(t)$ and, at the same time, the spatio-temporal spectrum of the field ξ. In the spectral expansion of the driving white noise $\Omega_{\mathbf{k}}(t)$ (see Eq.(38)),

$$\Omega_{\mathbf{k}}(t) = \int_{\mathbb{R}} \mathrm{e}^{\mathrm{i}\omega \mathrm{t}}\, \mathrm{Z}_{\Omega_{\mathbf{k}}}(\mathrm{d}\omega), \tag{52}$$

we have $\mathsf{E}\,|Z_{\Omega_{\mathbf{k}}}(\mathrm{d}\omega)|^2 = \mathsf{const} \cdot \mathrm{d}\omega$ because the white noise has constant spectral density. Next, we substitute Eqs.(50) and (52) into Eq.(38), getting

$$(\mathrm{i}\omega + \mu\sqrt{1 + \lambda^2 k^2})^p Z_{\mathbf{k}}(\mathrm{d}\omega) = Z_{\Omega_{\mathbf{k}}}(\mathrm{d}\omega). \tag{53}$$

In this equation, taking expectation of the squared modulus of both sides yields

$$[(\omega^2 + \mu^2(1 + \lambda^2 k^2)]^p\, b_{\mathbf{k}}(\omega) = \mathsf{const}, \tag{54}$$

whence, recalling that $\mu = U/\lambda$ and introducing the scaled angular frequency $\omega' := \omega/U$, we finally obtain

$$b_{\mathbf{k}}(\omega') \equiv b_{\mathbf{K}} \propto \frac{1}{(\lambda^{-2} + (\omega')^2 + k^2)^p} = \frac{1}{(\lambda^{-2} + \mathbf{K}^2)^p}, \tag{55}$$

where

$$\mathbf{K} = (\ \frac{\omega}{U}, \mathbf{k}) \equiv (\frac{\omega}{U}, m, n, l) \tag{56}$$

is the spatio-temporal wavevector.

From Eq.(55), one can see that with the scaled frequency (note that the change $\omega \to \omega/U$ corresponds to the change of the time coordinate $t \to t \cdot U$), the spatio-temporal spectrum $b_{\mathbf{k}}(\omega') = b_{\mathbf{K}}$ becomes *isotropic in space-time*. This implies that the correlation function of ξ in isotropic in space-time as well (with the scaled time coordinate). Note that this remarkable property can be achieved only with the spatial order $q = \frac{1}{2}$.

6.4.3 Continuity of realizations of ξ in space-time

The functional form of the spatio-temporal spectrum Eq.(55) together with the constraint Eq.(48) imply that the conditions of Theorem 3.4.3 in [26] are satisfied, so that spatio-temporal sample paths of the random field ξ are almost surely continuous, as we demanded in section 3, see requirement 1.

6.5 Spatio-temporal covariances: the Matérn class

The spatio-temporal field satisfying the p-th order SPG model Eq.(37) has the spatio-temporal correlation function belonging to the so-called *Matérn* class of covariance functions [e.g. 29, 30]. To see this, we denote

$$\nu := p - \frac{d+1}{2} > 0, \tag{57}$$

where positivity follows from Eq.(48). Then Eq.(55) rewrites as

$$b_{\mathbf{K}} \propto \frac{1}{(\lambda^{-2} + \mathbf{K}^2)^{\nu + \frac{d+1}{2}}}. \tag{58}$$

Note that here $d+1$ is the dimensionality of space-time. Equation (58) indeed presents the spectrum of the Matérn family of correlation functions, see e.g. Eq.(32) in [29]. The respective isotropic correlation function is given by the equation that precedes Eq.(32) in [29] or by Eq.(1) in [30]:

$$B(r) \propto (r/\lambda)^{\nu} K_{\nu}(r/\lambda), \tag{59}$$

where r is the distance (in our case, the spatio-temporal distance $r = \sqrt{(Ut)^2 + s^2}$, with s being the spatial distance) and K_{ν} is the MacDonald function (the modified Bessel function of the second kind).

The Matérn family is often recommended for use in spatial analysis due to its notable flexibility with only two free parameters: ν and λ, see e.g. [29] and [30]. Specifically, λ controls the length scale, whereas $\nu > 0$ determines the degree of smoothness (the higher ν the smoother sample paths of the random field, for illustration see Appendix E).

Table 1 lists the resulting spatial correlation functions for several combinations of d and p [see 30, for details].

With the fixed d, the larger p corresponds, according to Eq.(57), to the larger ν and so to the smoother in space and time field ξ. This allows us to change the degree of smoothness of the generated field by changing the temporal order of the SPG model.

From the constraint Eq.(48), the minimal temporal order p that can be used in both 2D and 3D is equal to 3. This value $\boxed{p = 3}$ will be used by default in what follows and in the current SPG computer program.

Table 1: **The spatial correlation functions $B(s)$ for some plausible combinations of the dimensionality d and the temporal order p**

d	p	$\nu = p - \frac{d+1}{2}$	$B(s)$
2	2	$\frac{1}{2}$	$\mathrm{e}^{-\frac{s}{\lambda}}$
2	3	$\frac{3}{2}$	$(1+\frac{s}{\lambda})\,\mathrm{e}^{-\frac{s}{\lambda}}$
2	4	$\frac{5}{2}$	$(1+\frac{s}{\lambda}+\frac{1}{3}\left(\frac{s}{\lambda}\right)^2)\,\mathrm{e}^{-\frac{s}{\lambda}}$
3	3	1	$\frac{s}{\lambda}K_1(\frac{s}{\lambda})$

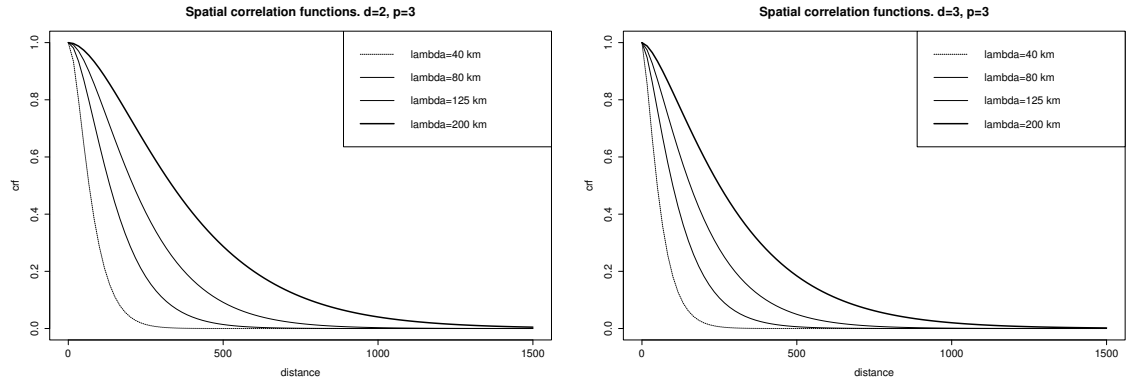

Figure 1: The spatial correlation functions for $p = 3$ in 2D (the left panel) and 3D (the right panel)—for four spatial length scales indicated in the legend.

6.5.1 Spatial correlation functions

The above spatio-temporal isotropy (see section 6.4.2 and Eq.(59)) means, in particular, that the spatial correlations are also isotropic. Figure 1 presents the *spatial* correlation functions calculated for different length scales in 2D and 3D following Eq.(59). To make the plots more accessible, it is arbitrarily assumed that the extent of the standardized spatial domain (the torus) in each dimension equals 3000 km, so that the distance is measured in kilometers. From Fig.1, one can notice, first, that the actual length scale is indeed well controlled by the parameter λ. Second, it is seen that in 2D (the left panel), where, according to Eq.(57), $\nu = \frac{3}{2}$, the correlation functions are somewhat *smoother* at the origin than in 3D (the right panel), where $\nu = 1$. This is consistent with the above statement that the greater ν the smoother the field. But in general, the 2D and 3D spatial correlation functions are quite similar.

6.5.2 Temporal correlation functions

Equation (59) shows that the spatial and temporal correlations have the same *shape*. The latter feature is very nice because atmospheric spectra are known to be similar in the spatial and in the temporal domain, e.g. the well-known "-5/3" spectral slope law is observed both in space and time, see e.g. [31, section 23]. So, our SPG does reproduce this observed in the nature similarity of spatial and temporal correlations.

Figure 2 shows the temporal correlation functions for the truncated (with $n_{max} = 90$) spatial spectral series Eq.(22). It is evident that the temporal length scale is well controlled by

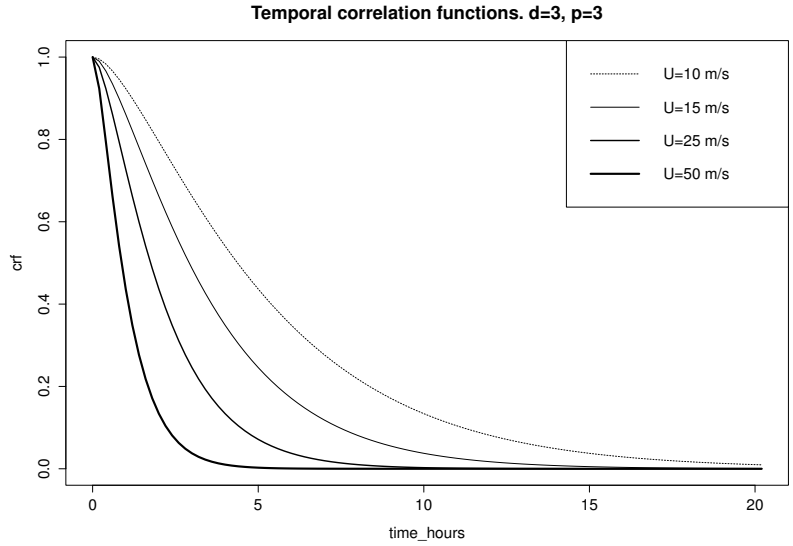

Figure 2: The temporal correlation functions in 3D for the four values of U indicated in the legend and $\lambda = 125$ km.

the parameter U. Comparing Fig.2 with Fig.1(right), one can observe that the spatial and temporal correlations indeed have the same shape; the effect of the spatial spectral truncation, which can cause a difference in the shapes, is barely visible.

6.5.3 Spatio-temporal correlations

Here, we explore the 3D spatio-temporal correlations calculated using Eq.(22) with the maximal wavenumbers in all three dimensions equal to $n_{max} = 90$, $\lambda = 125$ km, and $U = 20$ m/s. Figure 3 presents the spatial correlation functions for four time lags. Figure 4 displays the spatio-temporal correlation function.

From both Fig.3 and Fig.4, one can see the noticeable *non-separability* of the spatio-temporal covariances. The larger the time lag, the broader the spatial correlations. Note that this is consistent with the behavior of the spatio-temporal covariances found by [their Fig.8 16] in real-world wind speed data.

6.6 The final formulation of the SPG model

1. The temporal order of the SPG model is $\boxed{p = 3}$.
2. The driving noise $\alpha(t, \mathbf{s})$ is white both in time and space, so that the intensities of the spectral-space driving noises $\tilde{\alpha}_{\mathbf{k}}(t)$ are constant (and equal to σ). The intensity of the spatio-temporal white noise α is $(2\pi)^{d/2}\,\sigma$.

The resulting SPG model is

$$\boxed{\left(\frac{\partial}{\partial t} + \frac{U}{\lambda}\sqrt{1 - \lambda^2\Delta}\right)^3 \xi(t, \mathbf{s}) = \alpha(t, \mathbf{s})}. \tag{60}$$

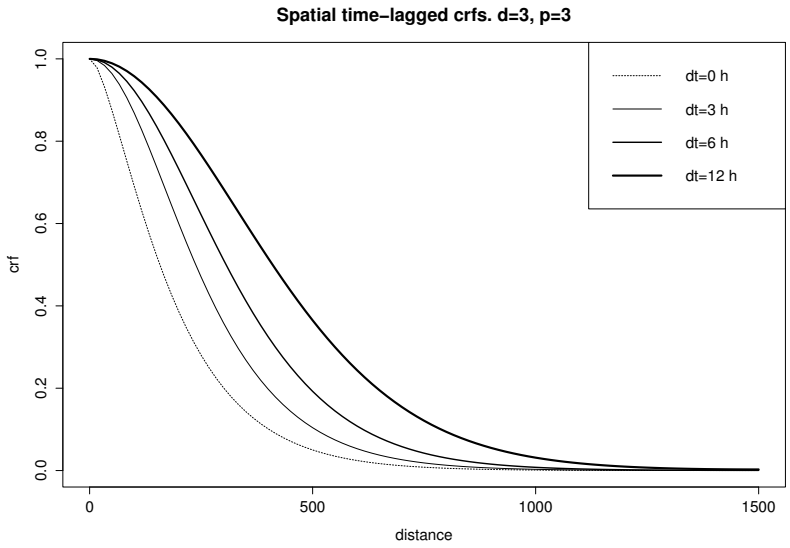

Figure 3: The spatial correlations in 3D for the four time lags indicated in the legend and $U = 20$ m/s.

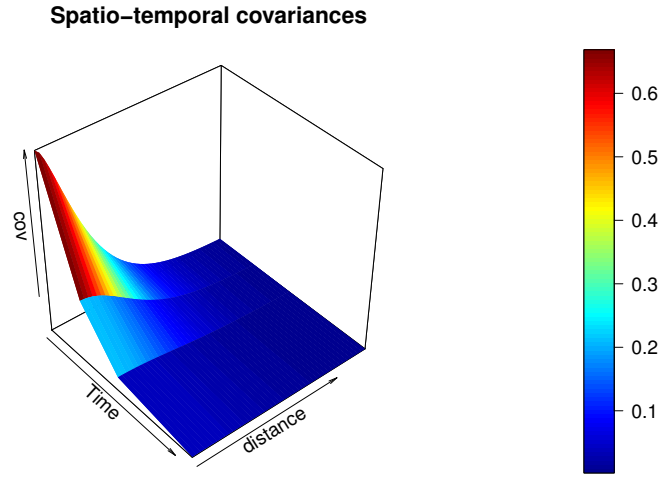

Figure 4: The spatio-temporal covariances.

In spectral space, each spectral coefficient $\tilde{\xi}_{\mathbf{k}}(t)$ satisfies the equation

$$\boxed{\left(\frac{\mathrm{d}}{\mathrm{d}t}+\frac{U}{\lambda}\sqrt{1+\lambda^2k^2}\right)^3\tilde{\xi}_{\mathbf{k}}(t)=\tilde{\alpha}_{\mathbf{k}}(t)=\sigma\,\Omega_{\mathbf{k}}(t)}, \tag{61}$$

where $\Omega_{\mathbf{k}}(t)$ are mutually independent complex standard white noise processes.

7 Time discrete solver for the third-order in time SPG model

In *physical space*, our final evolutionary model Eq.(60) can be discretized using the approximation of the operator $\sqrt{1-\lambda^2\Delta}$ proposed in Appendix C. The respective physical-space solver looks feasible but we do *not* examine it in this study. Below, we present our basic *spectral-space* technique. From this point on, we will consider only the *spectral* SPG.

7.1 The spectral solver

To numerically integrate the SPG model equations in spectral space, we discretize Eq.(61) using an implicit scheme. The model operator $(\frac{\mathrm{d}}{\mathrm{d}t}+a_{\mathbf{k}})^3$ (where, we recall, $a_{\mathbf{k}}:=\frac{U}{\lambda}\sqrt{1+\lambda^2k^2}$) is discretized by replacing the time derivative $\frac{\mathrm{d}}{\mathrm{d}t}$ with the backward finite difference $\frac{\mathcal{I}-\mathcal{B}}{\Delta t}$, where Δt is the time step, $\mathcal{I}$ is the identity operator, and $\mathcal{B}$ is the backshift operator. The r.h.s. of Eq.(61) (the white noise) is discretized following Appendix B, Eq.(89). As a result, we obtain the time discrete evolution equation

$$\hat{\xi}_{\mathbf{k}}(i)=\frac{1}{\varkappa^3}\left[3\varkappa^2\hat{\xi}_{\mathbf{k}}(i-1)-3\varkappa\hat{\xi}_{\mathbf{k}}(i-2)+\hat{\xi}_{\mathbf{k}}(i-3)+\sigma\Delta t^{\frac{5}{2}}\zeta_{\mathbf{k}t}\right], \tag{62}$$

where $i=0,1,2,\ldots$ denotes the discrete time instance, $\varkappa:=1+a_{\mathbf{k}}\Delta t$, and $\zeta_{\mathbf{k}t}\sim CN(0,1)$ are independent complex standard Gaussian pseudo-random variables (for their definition, see Appendix B.6). Note that the solution of the time-discrete Eq.(62) is denoted by the hat, $\hat{\xi}_{\mathbf{k}}(i)$, in order to distinguish it from the solution of the time-continuous Eq.(61), which is denoted by the tilde, $\tilde{\xi}_{\mathbf{k}}(t)$.

It can be shown that the numerical stability of the scheme Eq.(62) is guaranteed whenever $\varkappa>1$, which is always the case because $a_{\mathbf{k}}>0$, see Eq.(11).

Note that the derivation of the numerical scheme for a higher-order (i.e. with $p>3$) SPG model is straightforward: one should just raise the difference operator $\frac{\mathcal{I}-\mathcal{B}}{\Delta t}$ to a power higher than 3.

7.2 Correction of spectral variances

Because of discretization errors, the time discrete scheme Eq.(62) gives rise to the steady-state spectral variances $\hat{b}_{\mathbf{k}}:=\mathsf{Var}\,\hat{\xi}_{\mathbf{k}}(i)$, which are different from the "theoretical" ones, $b_{\mathbf{k}}$, given in Eq.(42) or Eq.(44). The idea is to correct $\hat{\xi}_{\mathbf{k}}(i)$ so that their steady-state variances coincide with $b_{\mathbf{k}}$. To this end, we derive $\hat{b}_{\mathbf{k}}$ from Eq.(62), see Eq.(104) in Appendix F, and then, knowing the "theoretical" $b_{\mathbf{k}}$, we introduce the correction coefficients, $\sqrt{b_{\mathbf{k}}/\hat{b}_{\mathbf{k}}}$, to be applied to $\hat{\xi}_{\mathbf{k}}(i)$. As a result of this correction, $\mathsf{Var}\,\hat{\xi}_{\mathbf{k}}(i)$ becomes, obviously, equal to the desired spectral variances $b_{\mathbf{k}}$. This simple device ensures that for any time step, the spatial spectrum and thus the *spatial* covariances are perfect. But the *temporal* correlations do depend on the time step, this aspect is discussed below in section 7.4.1.

7.3 "Warm start": ensuring stationarity from the beginning of time integration

To start the numerical integration of the third-order scheme Eq.(62) (for any wavevector $\mathbf{k}$), we obviously need three initial conditions. If the integration is the continuation of a previous run, then we just take values of $\hat{\xi}_{\mathbf{k}}(i)$ at the three last time instances i from that previous run; this ensures "continuity" of the resulting trajectory. If we start a new integration, we have to somehow generate values of $\hat{\xi}_{\mathbf{k}}(i)$ at $i = 1, 2, 3$, let us denote them here as the vector $\boldsymbol{\xi}^{ini} := (\hat{\xi}_{\mathbf{k}}(1), \hat{\xi}_{\mathbf{k}}(2), \hat{\xi}_{\mathbf{k}}(3))^{\top}$. Simplistic choices like specifying zero initial conditions give rise to a substantial initial transient period, which distorts the statistics of the generated field in the short time range.

In order to have the steady-state regime right from the beginning of the time integration and thus avoid the initial transient period altogether, we simulate $\boldsymbol{\xi}^{ini}$ as a pseudo-random draw from the multivariate Gaussian distribution with zero mean and the steady-state covariance matrix of $\hat{\xi}_{\mathbf{k}}(i)$. In Appendix F, we derive the components of this 3×3 matrix, namely, its diagonal elements (all equal to the steady-state variance), see Eq.(104), and the lag-1 and lag-2 covariances, see Eq.(105).

7.4 Computational efficiency

In this subsection, we describe two techniques that allow us to significantly decrease the computational cost of running the spectral SPG.

7.4.1 Making the time step Δt dependent on the spatial wavenumber k

For an ordinary differential equation, the accuracy of a finite-difference scheme depends on the time step. More precisely, it depends on the ratio of the time step Δt to the temporal length scale τ of the process in question. For high accuracy, $\Delta t \ll \tau$ is needed.

In our problem, $\tau_{\mathbf{k}}$ decays with the total wavenumber k, see Eq.(41). This implies that for higher k, smaller time steps are needed. To maintain the accuracy across the wavenumber spectrum, we choose the time step to be a portion of the time scale:

$$(\Delta t)_k := \gamma \tau_k. \tag{63}$$

The less γ, the more accurate and, at the same time, more time consuming the numerical integration scheme.

We note that in the atmospheric spectra, small scales have, normally, much less variance (energy) than large scales. But with the constant γ, the computational time would be, on the contrary, spent predominantly on high wavenumbers (because the latter require a smaller time step on the one hand and are much more abundant in 3D or 2D on the other hand). So, to save computer time whilst ensuring reasonable overall (i.e. for the whole range of wavenumbers) accuracy, we specify γ to be wavenumber dependent (growing with the wavenumber) in the following ad-hoc way:

$$\gamma_k := \gamma_{min} + (\gamma_{max} - \gamma_{min}) \left(\frac{k}{k_{max}} \right)^2, \tag{64}$$

where γ_{min} and γ_{max} are tunable parameters, $k = \sqrt{m^2 + n^2 + l^2}$, and $k_{max} := \max k$.

Note that the choice of γ depends on the shape of the temporal correlation function. For an OSDE of the type defined by Eq.(91) and τ defined to be equal to a^{-1}, the higher is the order p, the slower is the decay of the temporal correlation function for the same τ and thus the larger is to be γ.

The choice of the "optimal" γ_{min} and γ_{max} is discussed just below in section 7.4.2.

7.4.2 Introduction of a coarse grid in spectral space

Here we propose another technique to reduce the computational cost of the spectral solver. The technique exploits the *smoothness* of the SPG spectrum $b_{\mathbf{k}}$ Eq.(44). This smoothness allows us to introduce a *coarse grid in spectral space* and perform the integration of the time discrete spectral OSDEs Eq.(62) only for those wavevectors that belong to the coarse grid. The spectral coefficients $\hat{\xi}_{\mathbf{k}}(i)$ are then interpolated from the coarse grid to the dense (full) grid in spectral space.

The latter interpolation would introduce correlations between different spectral coefficients $\hat{\xi}_{\mathbf{k}}(i)$, which would destroy the spatial homogeneity. In order to avoid this, we employ a device used to generate so-called surrogate time series [32, section 2.4.1]. At each t, we multiply the interpolated $\hat{\xi}_{\mathbf{k}}(i)$ by $\mathrm{e}^{\mathrm{i}\theta_{\mathbf{k}}}$, where $\theta_{\mathbf{k}}$ are independent *random phases*, i.e. independent for different $\mathbf{k}$ random variables uniformly distributed on the segment $[0, 2\pi]$. It can be easily seen that this multiplication removes any correlation between the spectral coefficients.

Note also that the random phase rotation does not destroy the Gaussianity because $\hat{\xi}_{\mathbf{k}}(i)$ are complex circularly-symmetric random variables with uniformly distributed and independent of $|\hat{\xi}_{\mathbf{k}}(i)|$ arguments (phases), see e.g. [33, section A.1.3].

To preserve the temporal correlations of the field $\xi(t, \mathbf{s})$, we keep the set of $\theta_{\mathbf{k}}$ constant during the SPG-model time integration.

The exact spectrum $b_{\mathbf{k}}$ after the trilinear (bilinear in 2D) interpolation of $\hat{\xi}_{\mathbf{k}}(i)$ from the coarse to the full spectral grid is imposed in a way similar to that described in section 7.2 as follows. At any time instance when we wish to compute the physical space field, for each $\mathbf{k}$ on the full spectral grid, the linearly interpolated value $\check{\xi}_{\mathbf{k}}$ is a linear combination of the closest coarse-grid points $\mathbf{k}_j$:

$$\check{\xi}_{\mathbf{k}} = \sum_{j=1}^{2^d} w_j \, \hat{\xi}_{\mathbf{k}_j} \tag{65}$$

where $\check{}$ denotes the interpolated value and w_j is the interpolation weight (note that the set of the closest coarse-grid points $\mathbf{k}_j$ depends, obviously, on $\mathbf{k}$). In Eq.(65), the coarse-grid variances $\mathsf{Var}\,\hat{\xi}_{\mathbf{k}_j} = b_{\mathbf{k}_j}$ are known for all $\mathbf{k}_j$ from the spectrum $\{b_{\mathbf{k}}\}$, see Eqs.(42) or (44). Therefore, we can find $\mathsf{Var}\,\check{\xi}_{\mathbf{k}} = \sum_j w_j^2 \, b_{\mathbf{k}_j}$. Besides, we know which variance $\check{\xi}_{\mathbf{k}}$ *should* have on the fine grid, namely $b_{\mathbf{k}}$. So, we normalize $\check{\xi}_{\mathbf{k}}$ by multiplying it by $\sqrt{b_{\mathbf{k}}/(\mathsf{Var}\,\check{\xi}_{\mathbf{k}})}$, thus imposing the exact spatial spectrum for all $\mathbf{k}$.

Technically, the 3D coarse spectral grid is the direct product of three 1D grids. Any of the (non-uniform) 1D coarse grids is specified as follows. The jth coarse grid point is located at the fine-grid wavenumber n_j, which equals j for $|j| \le n_0$ (where n_0 is an integer) and the closest integer to $n_0(1+\varepsilon)^{|n|-n_0}$ for $|j| > n_0$. Here, ε is a tunable small positive number. In the below numerical experiments, the coarse-grid parameters were $n_0 = 20$ and $\varepsilon = 0.2$, which resulted in the following positive 1D coarse-grid points: 1 2 3 ... 19 20 24 29 35 42 50 60 72 86 103 124 150 (if not otherwise stated, the 1D grid extent was 300 points and, correspondingly, the maximal wavenumber was 150).

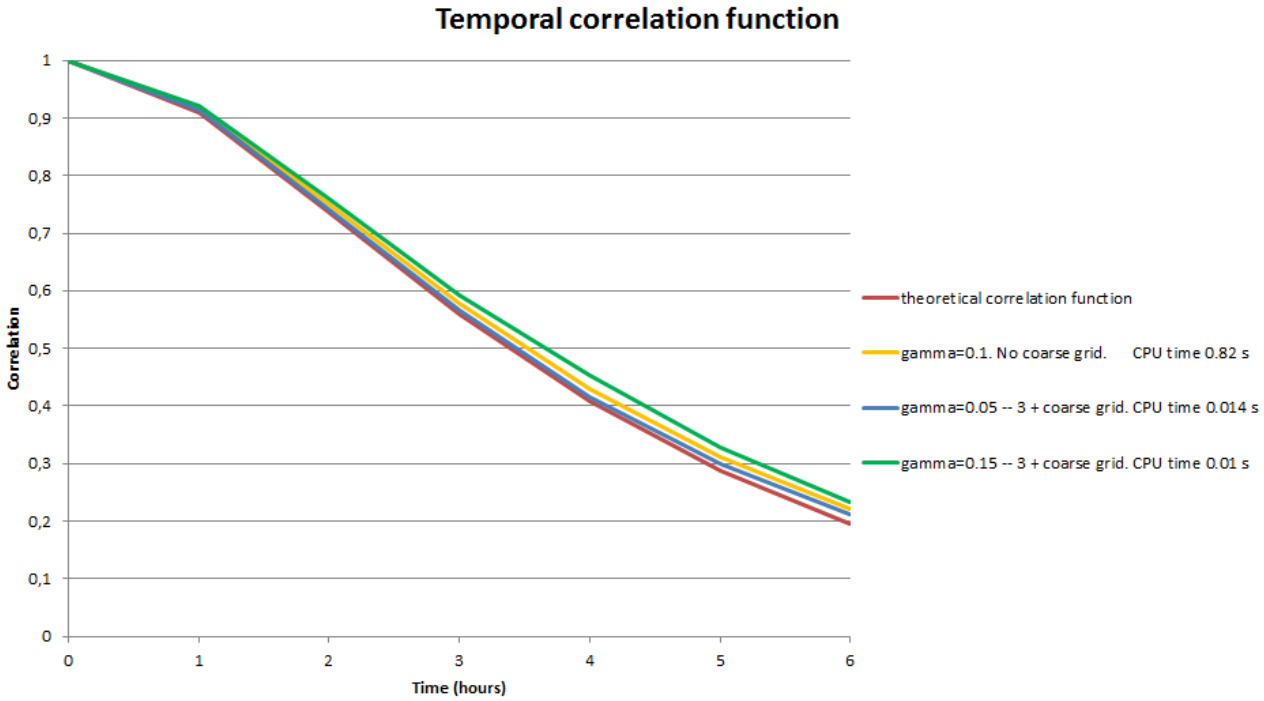

Figure 5: The theoretical and estimated temporal correlations and CPU times for the 2D SPG. The legend indicates the range γ_{min}–γ_{max} and whether the coarse grid was used. The respective spectral-space computation time (per one hour of time integration on one CPU) is also indicated in the legend. The SPG setup: $\lambda = 85$ km, $U = 12$ m/s.

7.4.3 Numerical acceleration: results

As the two above acceleration techniques guarantee that the *spatial* spectrum is always precise, we test how these techniques impact the *temporal* correlations (and what is the speedup).

In the numerical experiments, the introduction of the coarse spectral grid impacted the temporal correlations to a lesser extent than an increase in the time-step parameter γ. So, we examined the role of the two parameters γ_{min} and γ_{max} of the acceleration technique described in section 7.4.1, and the impact of the presence or absence of the coarse spectral grid introduced in section 7.4.2.

Figure 5 shows the temporal correlations functions for the different setups (indicated in the legend) of the 2D SPG. The respective CPU times for the spectral-space computations (on one CPU per one hour of model integration) are also shown in the legend. The discrete backward Fourier transforms were performed every hour of lead time.

From Fig.5, one can see that the combined effect of both numerical acceleration techniques (the green curve) was dramatic: the speedup was about 80 times as compared to the non-accelerated scheme (i.e. without the coarse grid and with constant $\gamma = 0.1$, the yellowish curve). The contributions of the two above numerical acceleration techniques to this speedup were comparable in magnitude (not shown). Most importantly, this big speedup was achieved at the very little cost: the temporal correlation length scale was distorted, as a result of approximations caused by the two acceleration techniques, by only some 5 percent w.r.t. the non-accelerated scheme and some 10 percent w.r.t. the theoretical model (the red curve).

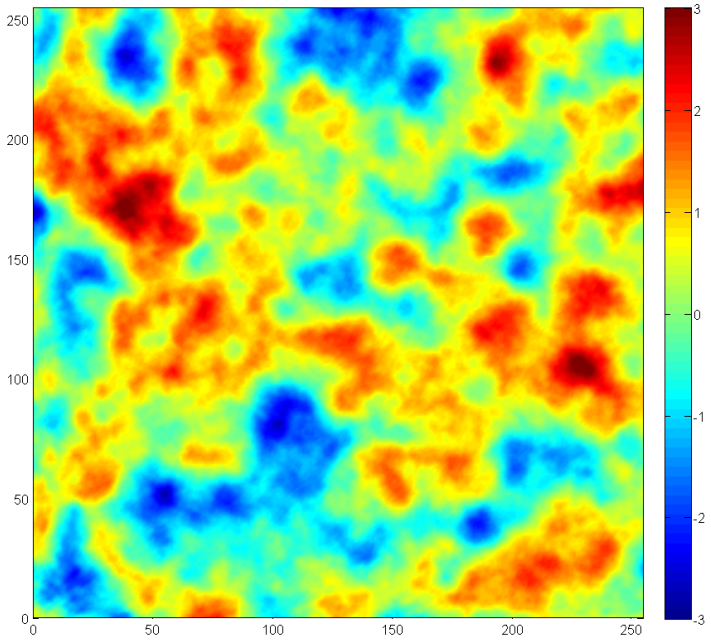

Figure 6: 2D (x-y) cross-section of a spatial SPG field.

Note that the speedup was for the spectral-space computations only, i.e. it did *not* include the cost of the discrete backward Fourier transform.

In summary, the 2D-in-space SPG took only 1 second on one CPU to perform the spectral-space model integration for as long as 100 hours of lead time. The respective cost of the backward Fourier transform performed every hour was about 4 seconds. The total cost was thus 5 seconds.

The same computations performed for the 3D-in-space grid with $300 \times 300 \times 64$ points took 60 seconds on one CPU for 100 hours of spectral-space model integration and 110 seconds for the hourly backward Fourier transform. The speedup of the spectral-space computations for the 3D scheme due to the two acceleration techniques was about 50 times.

7.5 Examples of the SPG fields

Figure 6 shows a "horizontal" x-y cross-section and Fig.7 a spatio-temporal x-t cross-section of a simulated pseudo-random field $\xi(t, x, y)$. Note that in each spatial direction, there were 300 grid points, whilst only 256 contiguous points are shown in the Figures. This is done in the SPG for practical purposes in order to avoid correlations between the opposite sides of the spatial domain, which would be spurious in real-world applications.

8 Application to the COSMO model

The SPG was embedded into the Fortran code of the limited-area meteorological non-hydrostatic model COSMO [34]. Within COSMO, the SPG code was parallelized using

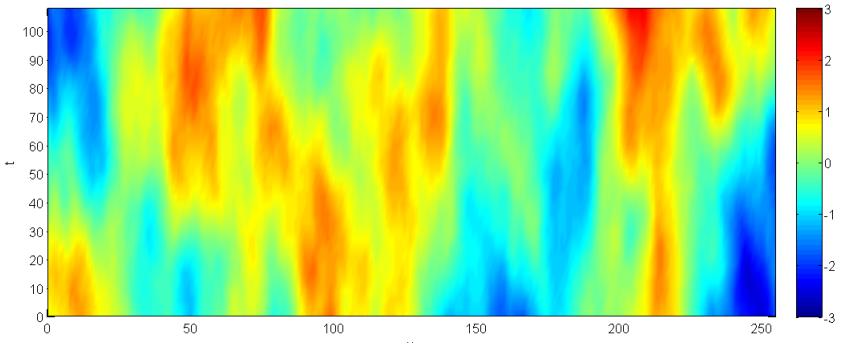

Figure 7: 2D (abscissa is x, ordinata is t) spatio-temporal cross-section of an SPG field.

MPI. The SPG was used to generate additive perturbations of the model's right-hand sides.

Figure 8 displays results of a numerical experiment, in which independent SPG perturbation fields for temperature and horizontal wind were simulated and added to the model fields every 15 minutes of lead time. The standard deviations of the SPG perturbations were specified as 0.005 K for temperature and 0.01 m/s for each of the two mutually orthogonal wind components. Besides, pressure perturbations were computed (and added to the model pressure 3D field) from the SPG temperature perturbation using the hydrostatic equation with the zero pressure perturbation at the model's top level.

The first generated by the SPG meridional wind perturbation field on the model level 25 (about 3 km above the ground) is displayed in Fig.8(top). The forecast (i.e. accumulated over time) meridional wind perturbation field after 3 hours of model integration is shown in Fig.8(bottom). Comparing the top and bottom panels of Fig.8 suggests that the perturbation field becomes more large-scale and, at the same time, less smooth over time. Specifically, the large-scale structure of the perturbation field becomes more large scale. This is to be expected because the time integration reduces (filters out) small-scale-in-time components of the spatio-temporal field and so, from the "proportionality of scales", filters out smaller scales also in space.

Besides, an opposite effect, namely appearance of a kind of fine structure can be seen in the forecast perturbation fields, see Fig.8(bottom). In particular, in the forecast perturbation field, there are many very localized features, which presumably develop due to a non-linear interaction with convection and effects related to steep orography. Finally, we note that in contrast to the "input" SPG model-error perturbation field, the "output" forecast perturbation field is clearly seen to be flow dependent (as desired in ensemble prediction and ensemble data assimilation schemes).

Figure 9 displays the growth of the meridional wind RMS forecast perturbation at the same model level 25 in response to the above SPG model-error perturbations. One can see that the magnitude of the forecast perturbation grows monotonically and "regularly" (as expected).

Finally, we note that at the time of writing, the numerical acceleration techniques described in section 7.4 were implemented only in the stand-alone version of the SPG. Without those accelerators, the cost of running the SPG within COSMO was about 0.8% of the total COSMO model wall-clock time per generated SPG field.

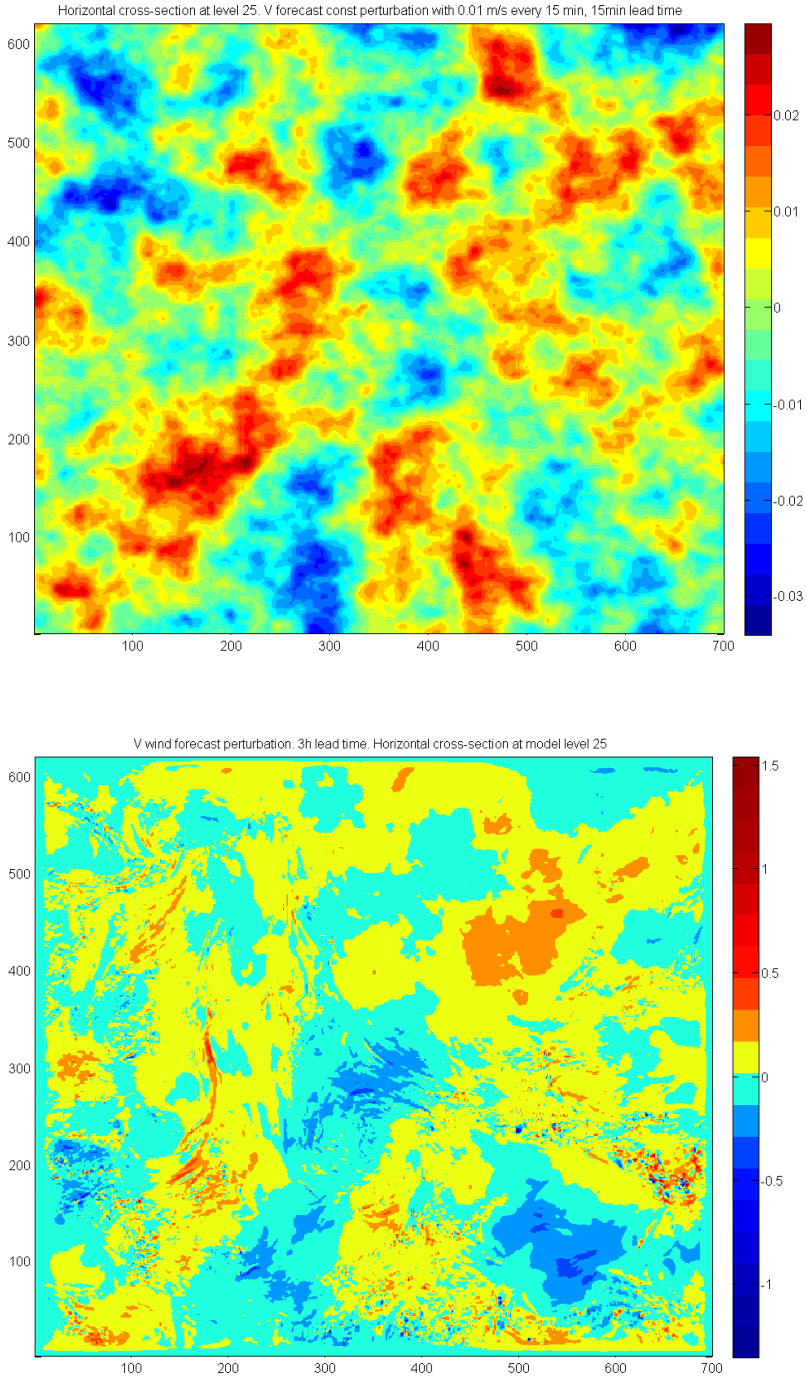

Figure 8: *The top panel*: The first SPG V-wind perturbation added to the COSMO field at the lead time 15 minutes. *The bottom panel*: the COSMO 3-h forecast V-wind perturbation field in response to the additive SPG perturbations of temperature, pressure, and both wind components added every 15 minutes of lead time.

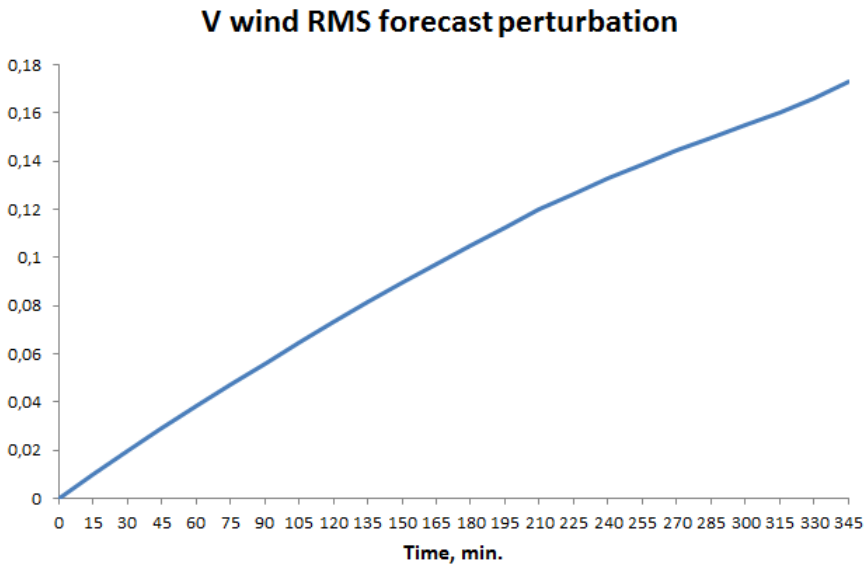

Figure 9: The COSMO forecast RMS V-wind perturbation in response to the SPG forcing.

9 Discussion

9.1 Physical-space or spectral-space SPG solver?

In this study, we have investigated both the spectral-space and the physical-space approximations of the SPG spatio-temporal model. We have found that both approaches can be used to build a practical SPG scheme. We have selected the spectral-space technique. Here, we briefly compare both approaches.

Advantages of the spectral-space technique are the following.

- Simplicity of realization. If the SPG model has constant coefficients, then the complicated SPG equation decouples into a series of simple OSDEs.
- Straightforward accommodation of non-local-in-physical-space spatial operators.

Advantages of the physical-space approach are:

- The relative ease of introduction of inhomogeneous and anisotropic capabilities to the SPG.
- The SPG solver can be implemented in domains with complex boundaries.
- Better suitability for an efficient implementation on massively parallel computers.

9.2 Extensions of the SPG

The proposed SPG technique can be extended in the future along the following lines.

- Development of a physical-space solver.

- Introduction of advection to the SPG model.
- Introduction of spatial inhomogeneity/anisotropy and non-stationarity.
- Introduction of non-Gaussianity. This can be done either by applying a nonlinear transform to the output SPG fields, or by introducing a non-Gaussian driving noise [as in 24, 25]. The former approach is simpler but the latter allows for much richer deviations from Gaussianity, including the multi-dimensional aspect.
- Going beyond additive and multiplicative perturbations for highly non-Gaussian variables like humidity and cloud fields.
- Simulation of several mutually correlated pseudo-random fields.
- Making the temporal order p a user defined variable. As noted above, the larger p the smoother the generated field.

10 Conclusions

10.1 Summary

- The proposed Stochastic Pattern Generator (SPG) produces pseudo-random spatio-temporal Gaussian fields on 2D and 3D spatial domains.
- The SPG model is defined on a standardized domain in space, specifically, on the unit 2D or 3D torus. Fields on a limited-area geophysical domain in question are obtained by mapping from the standardized domain.
- The SPG is based on a linear third-order in time stochastic model driven by the white in space and time Gaussian noise.
- The spatial operator of the stochastic model is built to ensure that solutions to the SPG model, i.e. the generated pseudo-random fields satisfy the "proportionality of scales" property: large-scale (small-scale) in space field components have large (small) temporal length scales.
- Besides the "proportionality of scales", the generated fields possess a number of other nice properties:
  - The spatio-temporal realizations are (almost surely) continuous.
  - With the appropriately scaled time coordinate, the spatio-temporal fields are isotropic in space-time.
  - The spatial and temporal correlation functions belong to the Matérn class.
  - The spatial and temporal correlations have the same shape.
- It is shown that the spatial operator of the SPG model can be effectively discretized both in physical space and spectral space.
- The basic SPG solver is spectral-space based.
- Two techniques to accelerate the spectral-space computations are proposed and implemented. The first technique selects the time step of the spectral-space numerical integration scheme to be dependent on the wavenumber, so that the discretization error is smaller for more energetic larger spatial scales and is allowed to be larger for

less energetic smaller scales. The second technique introduces a coarse grid in spectral space. The combined speedup for spectral-space computations from both techniques is as large as 50–80 times.

- The SPG is embedded in the COSMO model as a source of additive model-error perturbations. The numerical experiments show that the forecast perturbations in response to the SPG model-error perturbations behave as expected.
- For a peer reviewed publication on the SPG, see [35].

10.2 Applications

Potential applications of the SPG include ensemble prediction and ensemble data assimilation in meteorology, oceanography, hydrology, and other areas.

The SPG can be used to generate spatio-temporal perturbations of the model fields (in the additive or multiplicative or other mode), and of the boundary conditions.

Acknowledgements

The SPG has been developed as part of the Priority Project KENDA (Kilometer scale Ensemble Data Assimilation) of COSMO. We have used the discrete fast Fourier package `fft991` developed by C. Temperton at ECMWF in 1978. Christoph Schraff and Michael Baldauf kindly helped us to interpret the appearance of the small-scale noise in the COSMO forecast perturbation fields.

A Illustration of the "proportionality of scales" property

Figure 10 shows a realization of the spatio-temporal field with *non-separable* correlations that satisfy the "proportionality of scales" property (the top panel) and a realization of the field with *separable* spatio-temporal correlations (the bottom panel).

The non-separable spatio-temporal correlation function is defined as $B(x,t) = \exp(-r/L)$, where $r := \sqrt{x^2 + (Ut)^2}$, $U = 10$ m/s, $L = 200$ km, and the domain size in the x direction is 3000 km. The separable correlation function is $B(x,t) = \exp(-|x|/L) \cdot \exp(-|Ut|/L)$. So, both separable and non-separable fields have exactly the same spatial correlation functions and the same temporal correlation functions.

Note that both the separability and the exponential temporal correlation function is what the scale-independent first-order auto-regression used in [19, 9, 20, 21, 22] implies.

Comparing the two panels of Fig.10, one can see that the two fields are quite different. In the non-separable case, Fig.10(top), large spatial structures indeed tend to live longer than small structures, as it is expected from the "proportionality of scales" principle. In contrast, in the case with separable space-time correlations, Fig.10(bottom), the "longevity" of a spatial pattern is rather independent of its size (which is unphysical). Besides, in the non-separable case, a kind of spatio-temporal "organization" is evident, which is absent in the separable case. Finally, the field with non-separable correlations exhibits a sort of spatio-temporal isotropy, again, not visible in the separable case.

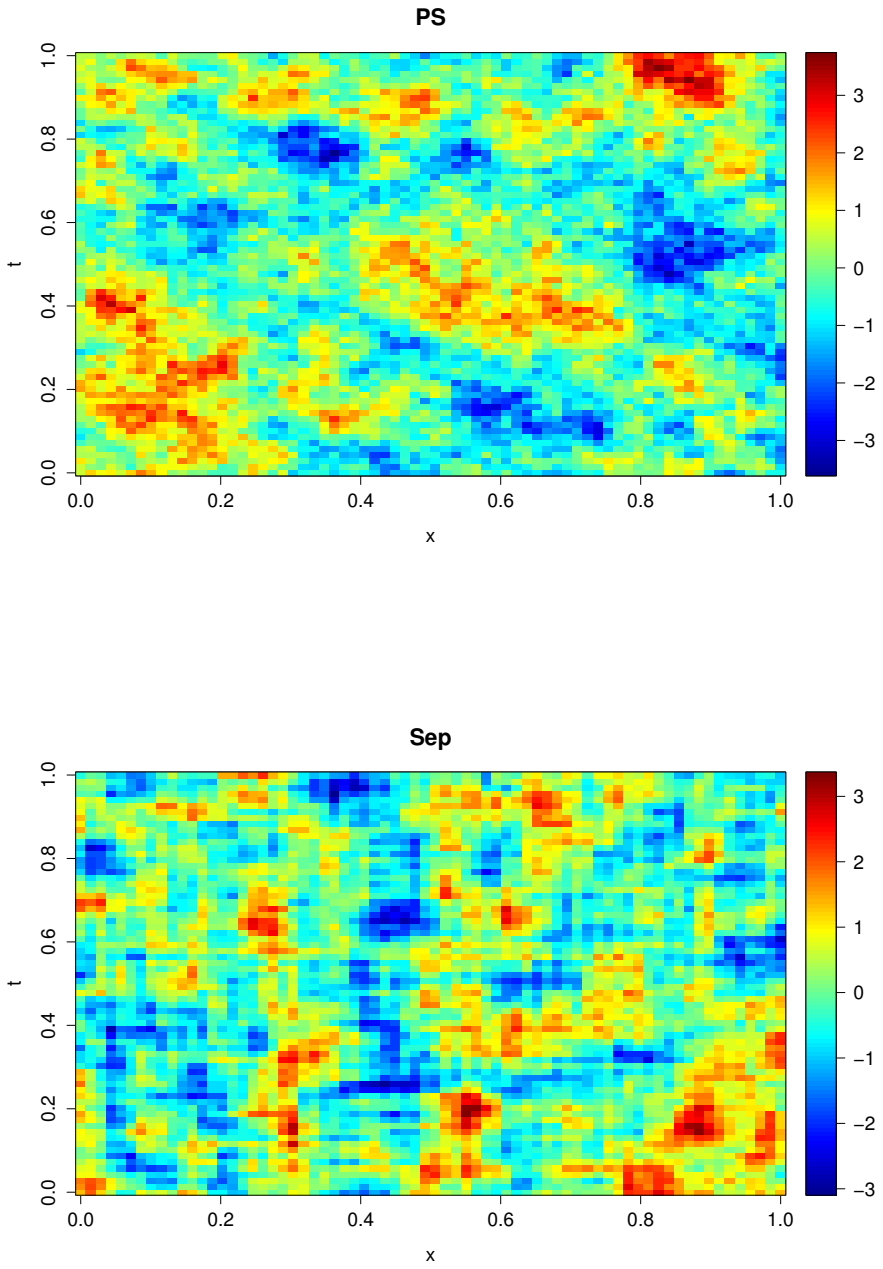

Figure 10: Simulated spatio-temporal fields. *Top*: With non-separable space-time correlations satisfying the "proportionality of scales" principle. *Bottom*: With separable space-time correlations.

B Spatio-temporal structure of the driving 4-D noise

Here, we recall the general definition of the white noise, define the spatial spectrum of the white noise on the d-dimensional unit torus, and find its spatial spectral decomposition in the spatio-temporal case. Then we introduce a colored in space and white in time noise, and find it spatial spectrum. Finally, we define the time discrete complex-valued white-noise process.

B.1 White noise

By definition, see e.g. [36, section 1.1.3] or [37, section 3.1.4], the *standard white noise* $\Omega(\mathbf{x})$ defined on a manifold $\mathbb{D}$ is a *generalized random field* that acts on a test function $\varphi(\mathbf{x})$ (where $x \in \mathbb{D}$) as follows:

$$(\Omega, \varphi) := \int \varphi(\mathbf{x})\, \Psi(\mathrm{d}\mathbf{x}), \tag{66}$$

where Ψ is the Gaussian orthogonal stochastic measure such that for any Borel set A, $\Psi(A)$ is a (complex, in general) Gaussian random variable with $\mathsf{E}\,\Psi(A) = 0$ and $\mathsf{E}\,|\Psi(A)|^2 = |A|$, where $|.|$ denotes the Lebesgue measure.

The equivalent definitions of the standard white noise are

$$\mathsf{E}\,|(\Omega, \varphi)|^2 := \int |\varphi(\mathbf{x})|^2\, \mathrm{d}\mathbf{x}, \tag{67}$$

and

$$\mathsf{E}\,(\Omega, \varphi) \cdot \overline{(\Omega, \psi)} := \int \varphi(\mathbf{x}) \overline{\psi(\mathbf{x})}\, \mathrm{d}\mathbf{x}, \tag{68}$$

where ψ is another test function. Thus, we have defined of the *standard* white noise. By the general Gaussian white noise, we mean a multiple of the *standard* white noise.

B.2 Spectrum of the white noise on $\mathbb{T}^d$

The formal Fourier transform of the spatial white noise $\Omega(\mathbf{s})$ (where $\mathbf{s} \in \mathbb{T}^3$),

$$\tilde{\Omega}_{\mathbf{k}} = \frac{1}{(2\pi)^d} \int_{\mathbb{T}^d} \Omega(\mathbf{s})\, \mathrm{e}^{-\mathrm{i}(\mathbf{k},\mathbf{s})}\, \mathrm{d}\mathbf{s}, \tag{69}$$

can be rigorously justified as the action of the white noise Ω on the test function

$$\chi(\mathbf{s}) := \frac{1}{(2\pi)^d}\, \mathrm{e}^{-\mathrm{i}(\mathbf{k},\mathbf{s})}. \tag{70}$$

Then, the spatial spectrum of $\Omega(\mathbf{s})$ is

$$b_{\mathbf{k}} := \mathsf{E}\,|\tilde{\Omega}_{\mathbf{k}}|^2 \equiv \mathsf{E}\,|(\Omega, \chi)|^2 = \int_{\mathbb{T}^d} |\chi(\mathbf{s})|^2\, \mathrm{d}\mathbf{s} = \frac{1}{(2\pi)^d}. \tag{71}$$

Here, the third equality is due to Eq.(67). We stress that it is the *modal* spectrum that is constant for the white noise (not the variance spectrum), see also remark at the end of section 5.3.

B.3 Space-integrated spatio-temporal white noise on $\mathbb{T}^d \times \mathbb{R}$

Let us consider the spatio-temporal white noise $\Omega = \Omega(t, \mathbf{s})$, where $t \in \mathbb{R}$ is time and $\mathbf{s} \in \mathbb{T}^d$ the spatial coordinate vector. Take a spatial test function $c(\mathbf{s})$ and define the temporal process $\Omega_1(t)$ formally as

$$\Omega_1(t) := \int_{\mathbb{T}^d} \Omega(t, \mathbf{s}) c(\mathbf{s}) \, \mathrm{d}\mathbf{s}, \tag{72}$$

so that it acts on a test function in the temporal domain, $\varphi(t)$, as

$$(\Omega_1, \varphi) := \int_{\mathbb{R}} \Omega_1(t) \varphi(t) \, \mathrm{d}t = \int_{\mathbb{R}} \int_{\mathbb{T}^d} \Omega(t, \mathbf{s}) c(\mathbf{s}) \varphi(t) \, \mathrm{d}\mathbf{s} \, \mathrm{d}t. \tag{73}$$

Here, we note that the latter double integral is nothing other than the result of action of the original white noise $\Omega(t, \mathbf{s})$ on the spatio-temporal test function $c(\mathbf{s}) \cdot \varphi(t)$. This enables us to mathematically rigorously *define* $\Omega_1(t)$ as the generalized random process that, with the fixed $c(\mathbf{s})$, acts on the test function $\varphi(t)$ as follows:

$$(\Omega_1(t), \varphi(t)) := (\Omega(t, \mathbf{s}),\, c(\mathbf{s})\varphi(t)). \tag{74}$$

Now, using the definition Eq.(67) of the white noise $\Omega(t, \mathbf{s})$, we have

$$\mathsf{E} \, |(\Omega(t, \mathbf{s}),\, c(\mathbf{s})\varphi(t))|^2 = \int_{\mathbb{T}^d} \int_{\mathbb{R}} |c(\mathbf{s})|^2 |\varphi(t)|^2 \, \mathrm{d}\mathbf{s} \, \mathrm{d}t = \int_{\mathbb{T}^d} |c(\mathbf{s})|^2 \, \mathrm{d}\mathbf{s} \int_{\mathbb{R}} |\varphi(t)|^2 \, \mathrm{d}t. \tag{75}$$

Since we have fixed $c(\mathbf{s})$, we observe that

$$\sigma^2 := \int |c(\mathbf{s})|^2 \, \mathrm{d}\mathbf{s} \tag{76}$$

is a constant such that

$$\mathsf{E} \, |(\Omega_1(t), \varphi(t))|^2 = \sigma^2 \int_{\mathbb{R}} |\varphi(t)|^2 \, \mathrm{d}t. \tag{77}$$

Comparing this equation with one of the definitions of the standard white noise, Eq.(67), we recognize $\Omega_1(t)$ as a general Gaussian white noise in time, i.e. the standard temporal white noise multiplied by σ. We call σ the *intensity* of the white noise.

B.4 Spatial spectrum of a spatio-temporal white noise

Now, we are in a position to derive the spatial spectrum of the standard spatio-temporal white noise $\Omega(t, \mathbf{s})$. In the formal Fourier decomposition

$$\Omega(t, \mathbf{s}) = \sum_{\mathbf{k}} \tilde{\Omega}_{\mathbf{k}}(t) \, \mathrm{e}^{\mathrm{i}(\mathbf{k}, \mathbf{s})}, \tag{78}$$

the elementary temporal processes $\tilde{\Omega}_{\mathbf{k}}(t)$ can be shown to be white noises in time. Indeed, again formally, we have

$$\tilde{\Omega}_{\mathbf{k}}(t) = \frac{1}{(2\pi)^d} \int_{\mathbb{T}^d} \Omega(t, \mathbf{s}) \, \mathrm{e}^{-\mathrm{i}(\mathbf{k}, \mathbf{s})} \, \mathrm{d}\mathbf{s}. \tag{79}$$

Here, we recognize an expression of the kind given by Eq.(72) with $c(\mathbf{s}) := \mathrm{e}^{-\mathrm{i}(\mathbf{k}, \mathbf{s})}/(2\pi)^{\mathrm{d}}$. Therefore, from Eq.(77), $\tilde{\Omega}_{\mathbf{k}}(t)$ is a temporal white noise with the intensity $\sigma^{\Omega}_{\mathbf{k}}$ squared equal to

$$(\sigma^{\Omega}_{\mathbf{k}})^2 = \int |c(\mathbf{s})|^2 \, \mathrm{d}\mathbf{s} = \frac{1}{(2\pi)^{2d}} \int_{\mathbb{T}^d} |\mathrm{e}^{-\mathrm{i}(\mathbf{k}, \mathbf{s})}|^2 \, \mathrm{d}\mathbf{s} = \frac{1}{(2\pi)^{\mathrm{d}}}. \tag{80}$$

In addition, using Eq.(68), it is easy to show that $\tilde{\Omega}_{\mathbf{k}}(t)$ and $\tilde{\Omega}_{\mathbf{k}'}(t)$ are mutually orthogonal for $\mathbf{k} \neq \mathbf{k}'$.

To summarize, $\tilde{\Omega}_{\mathbf{k}}(t)$ are mutually orthogonal white-in-time noises, all with equal intensities $\sigma_{\mathbf{k}}^{\Omega} = (2\pi)^{-d/2}$:

$$\tilde{\Omega}_{\mathbf{k}}(t) = \frac{1}{(2\pi)^{d/2}}\,\Omega_{\mathbf{k}}(t), \tag{81}$$

where $\Omega_{\mathbf{k}}(t)$ are the standard white noises.

B.5 Spectral decomposition of a white in time and colored in space noise

In order to introduce a white in time and colored in space noise, let us *convolve* the spatio-temporal white noise $\Omega(t, \mathbf{s})$ with a *smoothing kernel* in space $u(\mathbf{s})$, getting

$$\alpha(t, \mathbf{s}) := \int_{\mathbb{T}^d} u(\mathbf{s} - \mathbf{r})\,\Omega(t, \mathbf{r})\,\mathrm{d}\mathbf{r}. \tag{82}$$

In this equation, the stochastic integral is defined, for any t and $\mathbf{s}$, following Eq.(74) with $c(\mathbf{r}) := u(\mathbf{s} - \mathbf{r})$. Fourier transforming $u(\mathbf{s})$,

$$u(\mathbf{s}) = \sum_{\mathbf{k}} \tilde{u}_{\mathbf{k}}\,\mathrm{e}^{\mathrm{i}(\mathbf{k},\mathbf{s})}, \tag{83}$$

and, in space, $\alpha(t, \mathbf{s})$,

$$\alpha(t, \mathbf{s}) = \sum_{\mathbf{k}} \tilde{\alpha}_{\mathbf{k}}(t)\,\mathrm{e}^{\mathrm{i}(\mathbf{k},\mathbf{s})}, \tag{84}$$

we easily obtain that the elementary spectral processes $\tilde{\alpha}_{\mathbf{k}}(t)$ are independent white noises in time with the intensities squared

$$\sigma_{\mathbf{k}}^2 = (2\pi)^d |\tilde{u}_{\mathbf{k}}|^2, \tag{85}$$

so that the stochastic differential $\tilde{\alpha}_{\mathbf{k}}(t)\mathrm{d}t$ is

$$\tilde{\alpha}_{\mathbf{k}}(t)\,\mathrm{d}t = \sigma_{\mathbf{k}}\,\mathrm{d}W_{\mathbf{k}}(t). \tag{86}$$

Equivalently,

$$\tilde{\alpha}_{\mathbf{k}}(t) = \sigma_{\mathbf{k}}\,\Omega_{\mathbf{k}}(t). \tag{87}$$

B.6 Discretization of the spectral processes $\tilde{\alpha}_{\mathbf{k}}(t)$ in time

Being white noises, $\tilde{\alpha}_{\mathbf{k}}(t)$ have infinite variances. They become ordinary random processes if, e.g., we discretize them in time. With the time step Δt, we define the discretized process $\hat{\alpha}_{\mathbf{k}}(t_j)$ at the time instance t_j by replacing, in Eq.(86), $\mathrm{d}t$ with Δt and $\mathrm{d}W_{\mathbf{k}}$ with $\Delta W_{\mathbf{k}}$:

$$\hat{\alpha}_{\mathbf{k}}(t_j)\,\Delta t := \sigma_{\mathbf{k}}\,\Delta W_{\mathbf{k}}(t). \tag{88}$$

As $\mathsf{E}\,|\Delta W_{\mathbf{k}}(t)|^2 = \Delta t$, we obtain

$$\hat{\alpha}_{\mathbf{k}}(t_j) = \frac{\sigma_{\mathbf{k}}}{\sqrt{\Delta t}} \cdot \zeta_{\mathbf{k}j}, \tag{89}$$

where $\zeta_{\mathbf{k}j}$ are independent complex standard Gaussian random variables $CN(0, 1)$. The latter is defined as a complex random variable whose real and imaginary parts are mutually uncorrelated zero-mean random variables with variances equal to $1/2$. $CN(0, 1)$ is sometimes referred to as circularly symmetric complex Gaussian (normal) random variable [e.g. 33].

Equation (89) shows that the spatial spectrum of the time discrete driving noise is $\sigma_{\mathbf{k}}^2/\Delta t$.

C Physical-space approximation of the operator $\sqrt{1-\lambda^2\Delta}$

As we have discussed in section 5.5.1, the fractional power (square root) of the negated and shifted Laplacian operator, $\mathcal{L} := \sqrt{1-\lambda^2\Delta}$, is defined as the pseudo-differential operator with the symbol $\tilde{l}(\mathbf{k}) := \sqrt{1+\lambda^2\mathbf{k}^2}$. In the literature, one can find approaches to discretization of fractional powers of elliptic operators, e.g. finite elements were used in the spatial context in [38].

Here, we propose a simple technique to build a spatial discretization scheme that approximates the operator $\sqrt{1-\lambda^2\Delta}$ in the sense that the symbol of the approximating operator is close to $\sqrt{1+\lambda^2\mathbf{k}^2}$.

To this end, we do the following.

1. Perform the backward Fourier transform of the symbol $\tilde{l}(\mathbf{k})$, getting the function $l(\mathbf{s})$. As multiplication in Fourier space by $\tilde{l}(\mathbf{k})$ is equivalent to convolution in physical space with $l(\mathbf{s})$, we obtain that for any test function $\varphi(\mathbf{s})$,

$$(\mathcal{L}\varphi)(\mathbf{s}) = \int_{\mathbb{T}^3} l(\mathbf{s}-\mathbf{r})\,\varphi(\mathbf{r})\,\mathrm{d}\mathbf{r}. \tag{90}$$

   The crucial moment here is that the kernel function $l(\mathbf{s})$ appears to be oscillating while rapidly decreasing in modulus as $|\mathbf{s}|$ increases (see below). This enables its efficient approximation with a compact-support (truncated) function.

2. With the discretization on the grid with n points in each of the d dimensions on the torus $\mathbb{T}^d$, the kernel function $l(\mathbf{s})$ is represented by the set of its grid-point values $l(\mathbf{s_i})$, where $\mathbf{s} = (s_1, \ldots, s_d)$, $\mathbf{i} = (i_1, \ldots, i_d)$, and $\mathbf{s_i} = (s_1(i_1), \ldots, s_d(i_d))$. If $l(\mathbf{s})$ appears to be rapidly decreasing away from $\mathbf{s} = \mathbf{0}$, we truncate the $l(\mathbf{s_i})$ function by limiting its support near the origin, thus getting the function $l_{trunc}(\mathbf{s_i})$. E.g. in 3D, the support of $l_{trunc}(\mathbf{s_i})$ consists of the grid points $\mathbf{i} = (i_1, i_2, i_3)$ that simultaneously satisfy the following constraints: $|i_1| \le J$, $|i_2| \le J$, and $|i_3| \le J$, where J is the spatial order of the scheme. Below, we present results with $J = 1$ (3 grid points in the support of the truncated kernel function in each dimension) and $J = 3$ (7 grid points in the support in each dimension).

3. Fourier transform $l_{trunc}(\mathbf{s})$ back to the spectral space, getting the approximated symbol $\tilde{l}_{trunc}(\mathbf{k})$.

4. Compare $\tilde{l}(\mathbf{k})$ with $\tilde{l}_{trunc}(\mathbf{k})$ and conclude whether a parsimonious (that is, with a very small J) approximation is viable.

Now, we present the results. We found that for $d = 1$, $d = 2$, and $d = 3$, the goodness of fit was similar, so we examine the 3D case below.

We selected the grid of $n = 2 \cdot n_{max} = 256$ points in each of the three dimensions. We specified the spatial non-dimensional length scale λ to be much greater than the mesh size $h = 2\pi/n$ and much less than the domain's extents, 2π. Specifically, we chose $\lambda = 1/n_1$, where $n_1 := \sqrt{n_{max}}$. (The results were not much sensitive to changes in n_1 within the whole wavenumber range on the grid.)

Figure 11 displays the resulting kernel function $l(s)$ for positive s (note that $l(s)$ is an even function of the scalar distance s). One can see the remarkably fast decay of $|l(s)|$ with the growing s. Consequently, a stencil with just a few points in each dimension can be expected to work well.

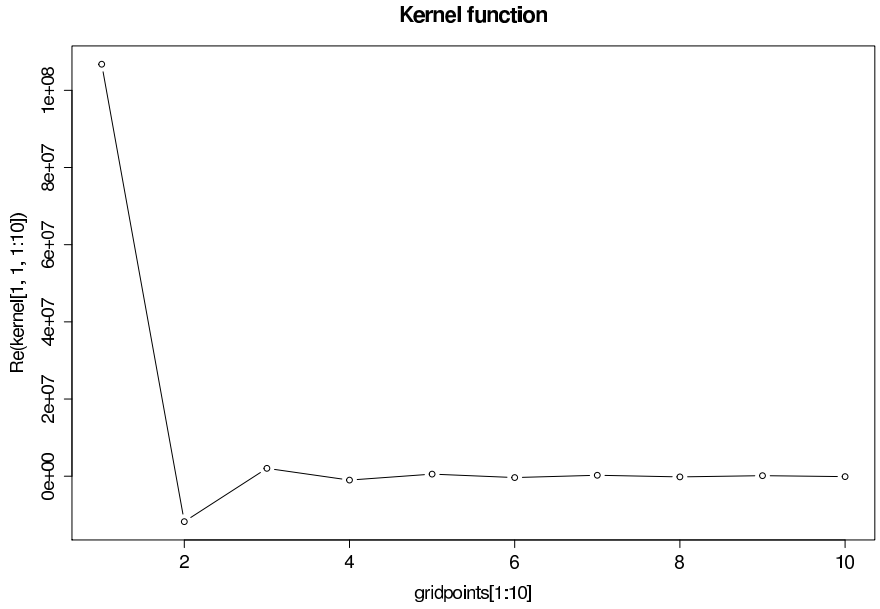

Figure 11: The kernel function

Figure 12 shows the exact and approximated symbols for the stencil that contains 3 grid points in each dimension (the left panel) and the stencil that contains 7 grid points in each dimension (the right panel). (The 5-point scheme worked not much better than the 3-point one and so its performance is not shown.)

From Fig.12, one can see that the 3-point scheme's performance is rather mediocre, whereas the 7-point scheme works very well (in terms of the reproduction of the operator's symbol).

Finally, we verified that the symbol $\tilde{l}_{trunc}(\mathbf{k})$ of the discrete operator for $J = 3, 5, 7$ was everywhere positive, which guarantees that the operator is positive definite and so the discretized SPG model should be stable.

To summarize, the operator $\sqrt{1-\lambda^2\Delta}$ can be approximated with parsimonious physical-space discretization schemes. For simulation of uncertainty in meteorology, where precise

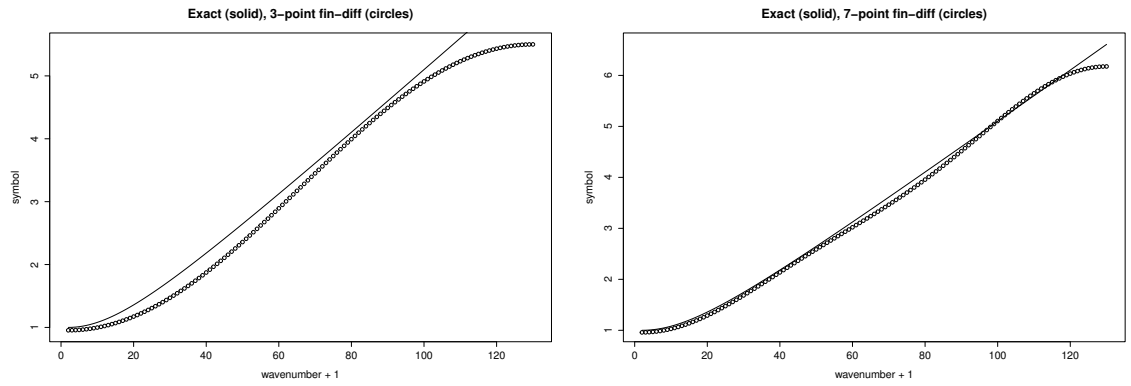

Figure 12: Goodness of fit of the symbol $\tilde{l}(\mathbf{k})$ (solid curve) by $\tilde{l}_{trunc}(\mathbf{k})$ (circles). Left: the 3-point stencil in each dimension. Right: the 7-point stencil

error statistics is not available, the simplest 3-point (in each direction) scheme seems most appropriate and computationally attractive. For more demanding applications, the 7-point scheme can be more appropriate.

D Stationary statistics of a higher-order OSDE

We examine the OSDE Eq.(38) in its generic form:

$$\boxed{\left(\frac{\mathrm{d}}{\mathrm{d}t}+a\right)^p \eta(t)=\sigma\Omega(t),} \tag{91}$$

where σ and a are the positive numbers, p is the positive integer, and $\Omega(t)$ is the standard white noise (see Appendix B).

The goal here is to find the variance and the correlation function of $\eta(t)$ in the stationary regime. The technique is to reduce the p-th order OSDE to a system of first-order OSDEs.

To simplify the exposition, we consider the third-order OSDE ($p=3$) and rewrite Eq.(91) as

$$\left(\frac{\mathrm{d}}{\mathrm{d}t}+a\right)\ \left\{\left(\frac{\mathrm{d}}{\mathrm{d}t}+a\right)\ \left[\left(\frac{\mathrm{d}}{\mathrm{d}t}+a\right)\eta(t)\right]\right\}=\sigma\Omega(t). \tag{92}$$

Here, by η_1 we denote the term in brackets,

$$\left(\frac{\mathrm{d}}{\mathrm{d}t}+a\right)\eta =: \eta_1 \tag{93}$$

and by η_2 the term in braces,

$$\left(\frac{\mathrm{d}}{\mathrm{d}t}+a\right)\eta_1 =: \eta_2, \tag{94}$$

so that the original equation Eq.(91) can be rewritten as

$$\left(\frac{\mathrm{d}}{\mathrm{d}t}+a\right)\eta_2=\sigma\Omega. \tag{95}$$

In Eqs.(93)–(95), the last equation is the familiar first-order OSDE forced by the white noise, whereas the other equations are not forced by the white noise. Generalizing the above construction, Eqs.(92)–(95), to the arbitrary $p>0$, we form the following first-order vector-matrix OSDE (a system of first-order OSDEs):

$$\mathrm{d}\boldsymbol{\eta}+\mathbf{A}\boldsymbol{\eta}\mathrm{d}t=\boldsymbol{\Sigma}\boldsymbol{\Omega}\mathrm{d}t, \tag{96}$$

where $\boldsymbol{\eta}:=(\eta,\eta_1,\ldots,\eta_{p-2},\eta_{p-1})$, $\boldsymbol{\Omega}:=(0,0,\ldots,0,\Omega)$, and the design of the matrices $\mathbf{A}$ and $\boldsymbol{\Sigma}$ is obvious (not shown).

With Eq.(96) in hand, we derive a differential equation for the covariance matrix $\mathbf{P}:=\mathsf{E}\,\boldsymbol{\eta}\boldsymbol{\eta}^*$, where * denotes transpose complex conjugate [see e.g. 11, example 4.16]. First, we compute the increment of $\mathbf{P}$:

$$\Delta\mathbf{P}=\mathsf{E}\,(\boldsymbol{\eta}+\mathrm{d}\boldsymbol{\eta})(\boldsymbol{\eta}+\mathrm{d}\boldsymbol{\eta})^*-\mathsf{E}\,\boldsymbol{\eta}\boldsymbol{\eta}^*=\mathsf{E}\,\boldsymbol{\eta}\mathrm{d}\boldsymbol{\eta}^*+\mathsf{E}\,\mathrm{d}\boldsymbol{\eta}\boldsymbol{\eta}^*+\mathsf{E}\,\mathrm{d}\boldsymbol{\eta}\mathrm{d}\boldsymbol{\eta}^*. \tag{97}$$

Then, using Eq.(96) and the fact that $\mathsf{E}\,|\Omega\mathrm{d}t|^2=\mathsf{E}\,|\mathrm{d}W|^2=\mathrm{d}t$, we obtain the differential of $\mathbf{P}$ from Eq.(97):

$$\mathrm{d}\mathbf{P}=-\mathbf{A}\mathbf{P}\mathrm{d}t-\mathbf{P}\mathbf{A}^*\mathrm{d}t+\boldsymbol{\Sigma}\boldsymbol{\Sigma}^*\mathrm{d}t. \tag{98}$$

In the stationary regime $\mathrm{d}\mathbf{P}=0$, so the equation for the stationary covariance matrix is

$$\boxed{\mathbf{A}\mathbf{P}+\mathbf{P}\mathbf{A}^*=\boldsymbol{\Sigma}\boldsymbol{\Sigma}^*.} \tag{99}$$

Next, we look at the *first* diagonal entry of the resulting stationary covariance matrix $\mathbf{P}$, which represents the required $\mathsf{Var}\,\eta$ (because η is defined above to be the first entry of the

vector $\boldsymbol{\eta}$). Dropping tedious derivations, we present in Table 2 (second row) the formulas for the temporal orders $p = 1$, $p = 2$, $p = 3$, and for the general p.

Finally, we derive the temporal correlation function for the pth-order OSDE. To this end, we multiply Eq.(91) by $\eta(s)$ with $s < t$ and take expectation. Since a is non-stochastic, we may interchange the expectation and the differential operator $\left(\frac{\mathrm{d}}{\mathrm{d}t} + a\right)^p$, getting the pth-order ordinary differential equation for the temporal covariance function, whose solutions for different p are presented in row 3 of Table 2.

Table 2: **The variances $\mathsf{Var}\,\eta$ and correlation functions $C_\eta(t)$ of the stationary solution to Eq.(91) for different temporal orders p**

p	1	2	3	Arbitrary p
$\mathsf{Var}\,\eta$	$\frac{\sigma^2}{2a}$	$\frac{\sigma^2}{4a^3}$	$\frac{3\sigma^2}{16a^5}$	$\propto \frac{\sigma^2}{a^{2p-1}}$
$C_\eta(t)$	$\mathrm{e}^{-\mathrm{a\|t\|}}$	$(1 + a\|t\|)\,\mathrm{e}^{-\mathrm{a\|t\|}}$	$(1 + a\|t\| + \frac{a^2t^2}{3})\,\mathrm{e}^{-\mathrm{a\|t\|}}$	$\propto R_{p-1}(a\|t\|) \cdot \mathrm{e}^{-\mathrm{a\|t\|}}$

Here $R_{p-1}(x)$ is a polynomial of order $p - 1$.

E Smoothness of sample paths of the spatial Matérn random field for different ν

Here, we show how sample paths (realizations) of the Matérn random field with the smoothness parameter ν look. Specifically, in Fig.13, we present three plots with 1D cross-sections of randomly chosen realizations of the Matérn random field for the following three values of ν: 1/2, 3/2, and 5/2. The spatial length scale parameter λ is selected in each of the three cases in such a way that the spatial correlation function intersects the 0.7 level at approximately the same distance (we denote this distance by $L_{0.7}$): $L_{0.7} = 500$ km. Again, as in section 6.5 and Appendix A, we assume, for convenience, that the extent of the spatial domain in each coordinate direction is 3000 km (rather than 2π). For comparison, we also display a realization with $L_{0.7} = 1500$ km for $\nu = 1/2$ (the bottom panel of Fig.13).

One can see that, indeed, the larger ν, the smoother the realizations—in the sense that they have less small-scale "noise". By contrast, increasing the length scale λ (compare the top and bottom panels of Fig.13) makes the *large-scale* pattern smoother but does *not* remove the smallest scales. So, the *large-scale* behavior is determined by the length scale λ, whereas the degree of *small-scale* smoothness/roughness depends predominantly on the smoothness parameter ν.

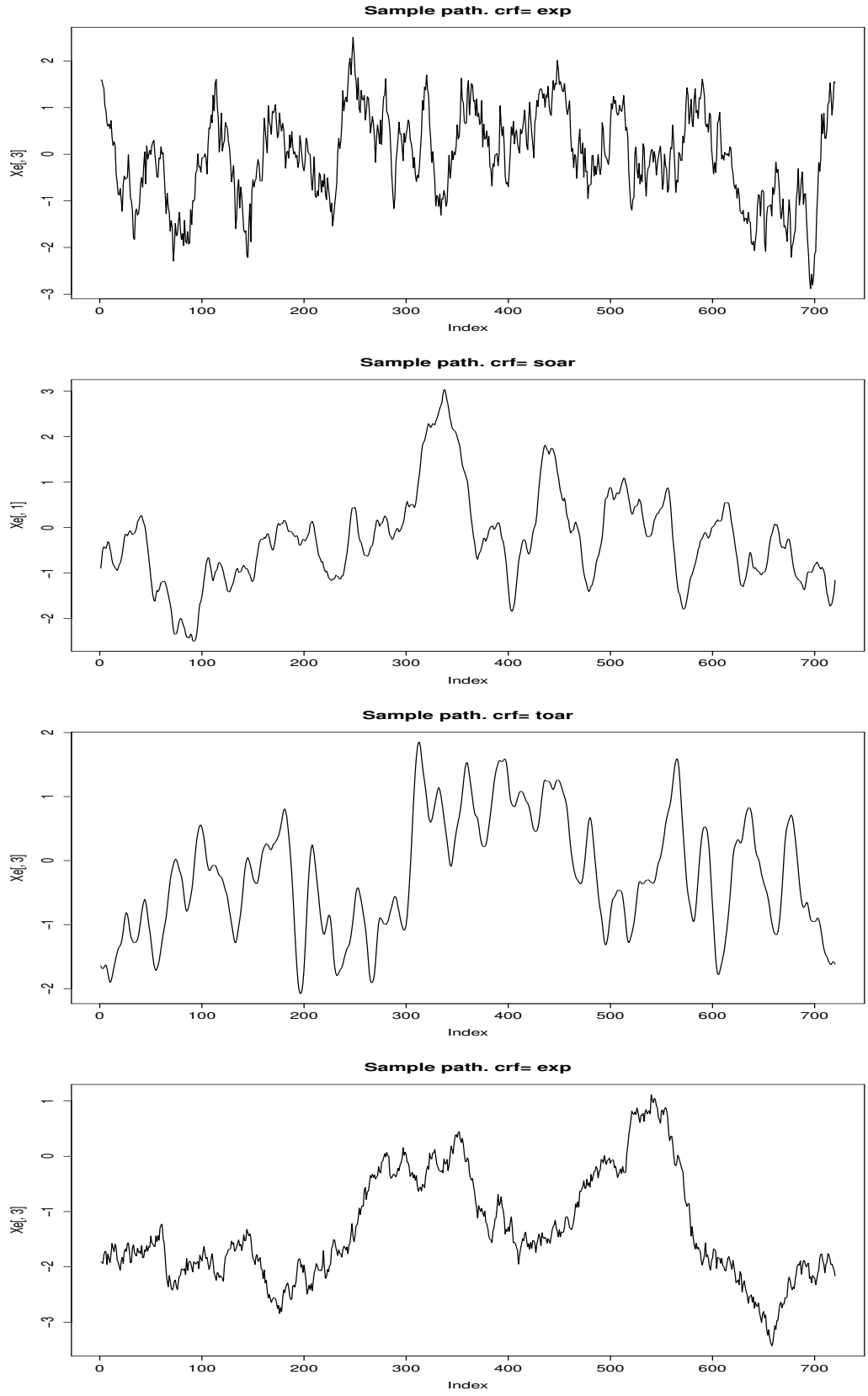

Figure 13: Sample paths for various ν and $L_{0.7}$. From the *top* to the *bottom*: $[\nu = \frac{1}{2}, L_{0.7} = 500]$, $[\nu = \frac{3}{2}, L_{0.7} = 500]$, $[\nu = \frac{5}{2}, L_{0.7} = 500]$, $[\nu = \frac{1}{2}, L_{0.7} = 1500]$

F Stationary statistics of a time discrete higher-order OSDE

Here, to simplify the exposition, we first examine the simplest first-order (i.e. with $p = 1$) OSDE and then give the results for the third-order OSDE used in the current version of the SPG.

F.1 First-order numerical scheme

Discretization of the Langevin Eq.(15) by an implicit scheme yields

$$\eta_i - \eta_{i-1} + a\eta_i\,\Delta t = \sigma\Delta W_i, \tag{100}$$

so that

$$\eta_i = \frac{\eta_{i-1} + \sigma\Delta W_i}{1 + a\Delta t}. \tag{101}$$

In the stationary regime, $\mathsf{Var}\,\eta_i = \mathsf{Var}\,\eta_{i-1}$, whence, bearing in mind that $\mathsf{Var}\,\Delta W_i = \Delta t$ and ΔW_i is independent on the values of η for all time moments up to and including the moment $i-1$, we apply the variance operator to both sides of Eq.(101) and obtain the stationary variance

$$V(\Delta t) := \lim_{i\to\infty} \mathsf{Var}\,\eta_i = \frac{\sigma^2}{2a + (g\Delta t)^2}. \tag{102}$$

Note that, as $\Delta t \to 0$, $V(\Delta t)$ tends to the continuous-time variance $\frac{\sigma^2}{2a}$, see Eq.(16).

F.2 Third-order numerical scheme

Consider the continuous-time OSDE, Eq.(91), with $p = 3$. The implicit scheme Eq.(62) we use to numerically solve it is reproduced here as

$$\eta_i = \frac{1}{\varkappa^3}\left[3\varkappa^2\eta_{i-1} - 3\varkappa\eta_{i-2} + \eta_{i-3} + \sigma(\Delta t)^2\,\Delta W_i\right], \tag{103}$$

where $\varkappa := 1 + a\Delta t$. Here, the goal is to find the stationary variance $V := \lim_{i\to\infty} \mathsf{Var}\,\eta_i$ along with lag-1 and lag-2 stationary covariances, $c_1 := \lim_{i\to\infty} \mathsf{E}\,\eta_i\eta_{i-1}$ and $c_2 := \lim_{i\to\infty} \mathsf{E}\,\eta_i\eta_{i-2}$, respectively. To reach this goal, we build three linear algebraic equations for the three unknowns, V, c_1, and c_2. The first equation is obtained by applying the variance operator to both sides of Eq.(103). The second and third equations are obtained by multiplying Eq.(103) by η_{i-1} and η_{i-2}, respectively, and applying the expectation operator to both sides of the resulting equations. Omitting the derivations, we write down the results:

$$V = \frac{\varkappa^4 + 4\varkappa^2 + 1}{(\varkappa^2 - 1)^5}\,(\Delta t)^5\sigma^2. \tag{104}$$

$$c_1 = \frac{3\varkappa(\varkappa^2 + 1)}{(\varkappa^2 - 1)^5}\,(\Delta t)^5\sigma^2, \qquad c_2 = \frac{6\varkappa^2}{(\varkappa^2 - 1)^5}\,(\Delta t)^5\sigma^2. \tag{105}$$

As in the first-order case, one can see that as $\Delta t \to 0$, V tends to the continuous-time variance $\frac{3}{16}\frac{\sigma^2}{a^5}$, see Table 2 in Appendix D.

References

[1] E.S. Epstein. Stochastic dynamic prediction. *Tellus*, 21(6):739–759, 1969.

[2] B.V. Tatarsky. The use of dynamic equations for a probabilistic forecasting of the pressure fields. *Izv. Akad. Nauk SSSR, FAO*, 5:293–297, 1969.

[3] E.J. Pitcher. Application of stochastic dynamic prediction to real data. *J. Atmos. Sci.*, 34(1):3–21, 1977.

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

[5] J. Berner, U. Achatz, L. Batte, A. De La Camara, D. Crommelin, H. Christensen, M. Colangeli, S. Dolaptchiev, C.L. Franzke, P. Friederichs, et al. Stochastic parameterization: Towards a new view of weather and climate models. *arXiv preprint arXiv:1510.08682*, 2015.

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

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

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

[9] J. Berner, G.J. Shutts, M. Leutbecher, and T.N. Palmer. A spectral stochastic kinetic energy backscatter scheme and its impact on flow-dependent predictability in the ECMWF ensemble prediction system. *J. Atmos. Sci.*, 66(3):603–626, 2009.

[10] L. Bengtsson, M. Steinheimer, P. Bechtold, and J.F. Geleyn. A stochastic parametrization for deep convection using cellular automata. *Q. J. R. Meterol. Soc.*, 139(675): 1533–1543, 2013.

[11] A.H. Jazwinski. *Stochastic processes and filtering theory*. Academic Press, 1970.

[12] L. Arnold. *Stochastic differential equations*. Wiley, 1974.

[13] M. Tsyrulnikov. Stochastic modelling of model errors: A simulation study. *Q. J. R. Meterol. Soc.*, 131(613):3345–3371, 2005.

[14] M. Tsyroulnikov. Proportionality of scales: An isotropy-like property of geophysical fields. *Q. J. R. Meterol. Soc.*, 127(578):2741–2760, 2001.

[15] N. Meunier and J. Zhao. Observations of photospheric dynamics and magnetic fields: from large-scale to small-scale flows. *Space Sci. Rev.*, 144(1-4):127–149, 2009.

[16] N. Cressie and H.C. Huang. Classes of nonseparable, spatio-temporal stationary covariance functions. *J. Amer. Statist. Assoc.*, 94(448):1330–1339, 1999.

[17] M.L. Stein. Space–time covariance functions. *J. Amer. Statist. Assoc.*, 100(469):310–321, 2005.

[18] T. Gneiting, M.G. Genton, and P. Guttorp. Geostatistical space-time models, stationarity, separability, and full symmetry. *Monographs Statist. Appl. Probab.*, 107:151, 2006.

[19] X. Li, M. Charron, L. Spacek, and G. Candille. A regional ensemble prediction system based on moist targeted singular vectors and stochastic parameter perturbations. *Mon. Wea. Rev.*, 136(2):443–462, 2008.

[20] N.E. Bowler, A. Arribas, S.E. Beare, K.R. Mylne, and G.J. Shutts. The local ETKF and SKEB: Upgrades to the MOGREPS short-range ensemble prediction system. *Q. J. R. Meterol. Soc.*, 135(640):767–776, 2009.

[21] T.N. Palmer, R. Buizza, F. Doblas-Reyes, T. Jung, M. Leutbecher, G.J. Shutts, M. Steinheimer, and A. Weisheimer. Stochastic parametrization and model uncertainty. *ECMWF Tech. Memo. n. 598, ECMWF, Shinfield Park, 42 pp.*, 2009.

[22] Bouttier F., Vié B., O. Nuissier, and L. Raynaud. Impact of stochastic physics in a convection-permitting ensemble. *Mon. Wea. Rev.*, 140(11):3706–3721, 2012.

[23] F. Lindgren, H. Rue, and J. Lindström. An explicit link between Gaussian fields and Gaussian Markov random fields: the stochastic partial differential equation approach. *J. Roy. Statist. Soc. B*, 73(4):423–498, 2011.

[24] S. Åberg and K. Podgórski. A class of non-Gaussian second order random fields. *Extremes*, 14(2):187–222, 2011.

[25] J. Wallin and D. Bolin. Geostatistical modelling using non-Gaussian Matérn fields. *Scand. J. Statist.*, 42(3):872–890, 2015.

[26] R.J. Adler. *The geometry of random fields*. Wiley, 1981.

[27] A.M. Yaglom. *Correlation theory of stationary and related random functions, Volume 1: Basic results*. Springer Verlag, 1987.

[28] M.A. Shubin. *Pseudodifferential operators and spectral theory*. Springer, 1987.

[29] M.L. Stein. *Interpolation of spatial data: some theory for kriging*. Springer, New York, 1999.

[30] P. Guttorp and T. Gneiting. Studies in the history of probability and statistics XLIX. On the Matern correlation family. *Biometrika*, 93(4):989–995, 2006.

[31] A.S. Monin and A.M. Yaglom. *Statistical fluid mechanics, Volume II: Mechanics of turbulence*. Courier Corp., 2013.

[32] J. Theiler, S. Eubank, A. Longtin, B. Galdrikian, and J.D. Farmer. Testing for nonlinearity in time series: the method of surrogate data. *Physica D*, 58(1-4):77–94, 1992.

[33] D. Tse and P. Viswanath. *Fundamentals of wireless communication*. Cambridge University Press, 2005.

[34] M. Baldauf, A. Seifert, J. Förstner, D. Majewski, M. Raschendorfer, and T. Reinhardt. Operational convective-scale numerical weather prediction with the COSMO model: description and sensitivities. *Mon. Wea. Rev.*, 139(12):3887–3905, 2011.

[35] M. Tsyrulnikov and D. Gayfulin. A limited-area spatio-temporal stochastic pattern generator for ensemble prediction and ensemble data assimilation. *Meteorol. Zeitschrift (under review)*, 2016.

[36] Y.A. Rozanov. *Markov random fields*. Springer, 1982.

[37] H.H. Kuo. White noise theory. *In "Handbook of stochastic analysis and applications" Kannan D. and Lakshmikantham V. (Eds.)*, pages 107–158, 2001.

[38] D. Simpson, F. Lindgren, and H. Rue. Think continuous: Markovian Gaussian models in spatial statistics. *Spatial Statist.*, 1:16–29, 2012.

List of COSMO Newsletters and Technical Reports

(available for download from the COSMO Website: www.cosmo-model.org)

COSMO Newsletters

No. 1: February 2001.

No. 2: February 2002.

No. 3: February 2003.

No. 4: February 2004.

No. 5: April 2005.

No. 6: July 2006.

No. 7: April 2008; Proceedings from the 8th COSMO General Meeting in Bucharest, 2006.

No. 8: September 2008; Proceedings from the 9th COSMO General Meeting in Athens, 2007.

No. 9: December 2008.

No. 10: March 2010.

No. 11: April 2011.

No. 12: April 2012.

No. 13: April 2013.

No. 14: April 2014.

No. 15: July 2015.

No. 16: July 2016.

COSMO Technical Reports

No. 1: Dmitrii Mironov and Matthias Raschendorfer (2001):
Evaluation of Empirical Parameters of the New LM Surface-Layer Parameterization Scheme. Results from Numerical Experiments Including the Soil Moisture Analysis.

No. 2: Reinhold Schrodin and Erdmann Heise (2001):
The Multi-Layer Version of the DWD Soil Model TERRA_LM.

No. 3: Günther Doms (2001):
A Scheme for Monotonic Numerical Diffusion in the LM.

No. 4: Hans-Joachim Herzog, Ursula Schubert, Gerd Vogel, Adelheid Fiedler and Roswitha Kirchner (2002):
LLM ¯ the High-Resolving Nonhydrostatic Simulation Model in the DWD-Project LITFASS.
Part I: Modelling Technique and Simulation Method.

No. 5: Jean-Marie Bettems (2002):
EUCOS Impact Study Using the Limited-Area Non-Hydrostatic NWP Model in Operational Use at MeteoSwiss.

No. 6: Heinz-Werner Bitzer and Jürgen Steppeler (2004):
Documentation of the Z-Coordinate Dynamical Core of LM.

No. 7: Hans-Joachim Herzog, Almut Gassmann (2005):
Lorenz- and Charney-Phillips vertical grid experimentation using a compressible nonhydrostatic toy-model relevant to the fast-mode part of the 'Lokal-Modell'.

No. 8: Chiara Marsigli, Andrea Montani, Tiziana Paccagnella, Davide Sacchetti, André Walser, Marco Arpagaus, Thomas Schumann (2005):
Evaluation of the Performance of the COSMO-LEPS System.

No. 9: Erdmann Heise, Bodo Ritter, Reinhold Schrodin (2006):
Operational Implementation of the Multilayer Soil Model.

No. 10: M.D. Tsyrulnikov (2007):
Is the particle filtering approach appropriate for meso-scale data assimilation ?

No. 11: Dmitrii V. Mironov (2008):
Parameterization of Lakes in Numerical Weather Prediction. Description of a Lake Model.

No. 12: Adriano Raspanti (2009):
COSMO Priority Project "VERification System Unified Survey" (VERSUS): Final Report.

No. 13: Chiara Marsigli (2009):
COSMO Priority Project "Short Range Ensemble Prediction System" (SREPS): Final Report.

No. 14: Michael Baldauf (2009):
COSMO Priority Project "Further Developments of the Runge-Kutta Time Integration Scheme" (RK): Final Report.

No. 15: Silke Dierer (2009):
COSMO Priority Project "Tackle deficiencies in quantitative precipitation forecast" (QPF): Final Report.

No. 16: Pierre Eckert (2009):
COSMO Priority Project "INTERP": Final Report.

No. 17: D. Leuenberger, M. Stoll and A. Roches (2010):
Description of some convective indices implemented in the COSMO model.

No. 18: Daniel Leuenberger (2010):
Statistical analysis of high-resolution COSMO Ensemble forecasts in view of Data Assimilation.

No. 19: A. Montani, D. Cesari, C. Marsigli, T. Paccagnella (2010):
Seven years of activity in the field of mesoscale ensemble forecasting by the COSMO–LEPS system: main achievements and open challenges.

No. 20: A. Roches, O. Fuhrer (2012):
Tracer module in the COSMO model.

No. 21: Michael Baldauf (2013):
A new fast-waves solver for the Runge-Kutta dynamical core.

No. 22: C. Marsigli, T. Diomede, A. Montani, T. Paccagnella, P. Louka, F. Gofa, A. Corigliano (2013):
The CONSENS Priority Project.

No. 23: M. Baldauf, O. Fuhrer, M. J. Kurowski, G. de Morsier, M. Müllner, Z. P. Piotrowski, B. Rosa, P. L. Vitagliano, D. Wójcik, M. Ziemiański (2013):
The COSMO Priority Project 'Conservative Dynamical Core' Final Report.

No. 24: A. K. Miltenberger, A. Roches, S. Pfahl, H. Wernli (2014):
Online Trajectory Module in COSMO: a short user guide.

No. 25: P. Khain, I. Carmona, A. Voudouri, E. Avgoustoglou, J.-M. Bettems, F. Grazzini (2015):
The Proof of the Parameters Calibration Method: CALMO Progress Report.

No. 26: D. Mironov, E. Machulskaya, B. Szintai, M. Raschendorfer, V. Perov, M. Chumakov, E. Avgoustoglou (2015):
The COSMO Priority Project 'UTCS' Final Report.

No. 27: J-M. Bettems (2015):
The COSMO Priority Project 'COLOBOC': Final Report.

No. 28: Ulrich Blahak (2016):
RADAR_MIE_LM and RADAR_MIELIB - Calculation of Radar Reflectivity from Model Output.

COSMO Technical Reports

Issues of the COSMO Technical Reports series are published by the *COnsortium for Small-scale MOdelling* at non-regular intervals. COSMO is a European group for numerical weather prediction with participating meteorological services from Germany (DWD, AWGeophys), Greece (HNMS), Italy (USAM, ARPA-SIMC, ARPA Piemonte), Switzerland (MeteoSwiss), Poland (IMGW), Romania (NMA) and Russia (RHM). The general goal is to develop, improve and maintain a non-hydrostatic limited area modelling system to be used for both operational and research applications by the members of COSMO. This system is initially based on the COSMO-Model (previously known as LM) of DWD with its corresponding data assimilation system.

The Technical Reports are intended

- for scientific contributions and a documentation of research activities,
- to present and discuss results obtained from the model system,
- to present and discuss verification results and interpretation methods,
- for a documentation of technical changes to the model system,
- to give an overview of new components of the model system.

The purpose of these reports is to communicate results, changes and progress related to the LM model system relatively fast within the COSMO consortium, and also to inform other NWP groups on our current research activities. In this way the discussion on a specific topic can be stimulated at an early stage. In order to publish a report very soon after the completion of the manuscript, we have decided to omit a thorough reviewing procedure and only a rough check is done by the editors and a third reviewer. We apologize for typographical and other errors or inconsistencies which may still be present.

At present, the Technical Reports are available for download from the COSMO web site (www.cosmo-model.org). If required, the member meteorological centres can produce hard-copies by their own for distribution within their service. All members of the consortium will be informed about new issues by email.

For any comments and questions, please contact the editor:

Massimo Milelli
Massimo.Milelli@arpa.piemonte.it