Results in Applied Mathematics 8 (2020) 100114

Contents lists available at ScienceDirect

Results in Applied Mathematics

journal homepage: www.elsevier.com/locate/results-in-applied-mathematics

Stochastic parameterization of the time-relaxation model of turbulence

Eric Olson

Department of Mathematics and Statistics, University of Nevada, Reno, NV 89557, USA

ARTICLE INFO

Article history:
Received 18 January 2020
Received in revised form 5 April 2020
Accepted 3 May 2020
Available online 20 May 2020

ABSTRACT

We parameterize the error in the time-relaxation model of turbulence by means of a deterministic corrector with stochastic noise term that represents the unresolved effects of the small scales on the large scales. Our method uses the large scales from a reference solution of the Navier–Stokes equations to compute the corrections needed to ensure the large scales from the turbulence model are the same as those in the reference solution over time. In the absence of model error, the theory of determining modes implies the size of the corrections needed over time decrease to zero provided enough Fourier modes were observed from the reference solution. In the presence of model error, the corrections do not decrease to zero. In this case, we call them the unresolved tendencies. After a theoretical analysis on the size of the these tendencies, our main result parameterizes them using a least-squares fit and characterizes the resulting residual as a time-correlated Gaussian noise processes. In doing so, we not only obtain an improved turbulence model, but a concrete representation of the error in the time-relaxation model of turbulence.

1. Introduction

The time-relaxation model of turbulence was introduced in 2001 by Stolz, Adams and Kleiser [1] as a way to approximate evolution of incompressible fluid flow with less computational effort than a direct numerical simulation of Navier–Stokes equations. The truncation of scales in the time-relaxation model were analyzed by Layton and Neda in [2] and a parameter sensitivity analysis performed by Neda, Pahlevani and Waters [3]. Additional studies appear in Dunca and Neda [4], in [5] and references therein. Advantages of the time-relaxation model include simplicity of numerical implementation using well-developed pressure solvers already present in existing subroutine libraries as well as suitability for wall-bounded flows.

In this paper the time-relaxation model represents a typical approximation of the incompressible Navier–Stokes needed for a definite context in which to develop and test the method of stochastic parameterization that will be described shortly. A similar parameterization could be performed for any dynamical system designed to approximate another. Doing so not only has the potential to improve the accuracy of the approximate dynamics, but the resulting characterization of errors further provides a quantitative framework for comparing different turbulence models. Before embarking on such a comparative analysis, it is important to first develop the theoretical and computational framework needed to effectively perform the stochastic parameterization. This, then, is our goal.

The method of stochastic parameterization presented here is inspired by the work of Wilks [6] for the two-layer Lorenz 96 model. In that paper the effects of the small scales in the second layer on the large scales in the first was modeled

E-mail address: ejolson@unr.edu.

https://doi.org/10.1016/j.rinam.2020.100114

by a fourth-degree polynomial plus a time-correlated noise term given by a discrete Ornstein–Uhlenbeck process, also called a red-noise Markov process. Lu, Lin and Chorin [7] perform a similar study of the effects of the small scales in the Kuramoto–Sivashinsky equation on a low-dimensional Fourier truncation. In that work a second-degree polynomial based on an approximate inertial manifold was used plus a discrete q-step noise process designed to more accurately characterize the time-correlations present in the errors of the low-dimensional truncation. Other research along these lines include Mohebujjaman, Rebholz, Iliescu [8], among others. See also the references in each of the above papers.

We note that the reduced dynamics considered in all the above studies started out as a low-dimensional projection which did not initially contain any terms to model the effects of the small scales on the large ones. As a result, significant improvements could be made by introducing terms deduced from the data to represent those effects. One distinguishing feature of the time-relaxation model (and other reasonable turbulence models) is that add-hoc and theoretical averaging techniques have already been used to construct terms which represent the effects of the missing scales. As a result, we are working in a context where improvements are more difficult to come by and where simply characterizing existing model errors becomes increasingly important. Turbulence models may also include variables and degrees of freedom that do not obviously correspond to any of the available data when using data-driven correction techniques. This is another difficulty that has to be overcome.

To motivate our research, recall that a natural task in the experimental sciences is to use observational measurements of a physical system taken over time to improve existing mathematical models of those physical systems. Related is the task of characterizing the model error in a model in order to obtain estimates on the uncertainties in predictions made using those approximations. Both these activities are directly connected to one of the primary uses of mathematics in science: to predict the future. With this in mind, one of our tasks is to improve an existing turbulence model to obtain greater predictive ability.

Consider, therefore, two dynamical systems

$$\frac{du}{dt} = F(u) \qquad \text{and} \qquad \frac{d\tilde{u}}{dt} = \widetilde{F}(\tilde{u})$$

where F represents the exact dynamics we wish to compute and $\widetilde{F}$ represents an approximation of those dynamics that we are able to compute. Our goal is to use observational data of the solution u to construct a more accurate dynamical system

$$\frac{dv}{dt} = G(v) \tag{1}$$

such that $v(t)$ approximates $u(t)$ for longer into the future than $\tilde{u}(t)$ and then characterize the errors in that model.

In [9] the residual model error was defined as

$$\frac{dR}{dt} = F(u) - \widetilde{F}(u) \qquad \text{where} \qquad \frac{du}{dt} = F(u)$$

and subsequently computed for the LANS-α and NS-α deconvolution models of turbulence. It was discovered in all cases that the model error was correlated over short distances in both space and time. However, for certain choices of parameters $|R|$ grew as $\sqrt{t}$ while for other values it grew linearly. As a good turbulence model should be free of the systematic biases that lead to linear error growth, one of the purposes of the corrector we shall obtain through data-driven stochastic parameterization is to remove such systematic biases. The remaining model error can then be characterized by a mean-zero time-correlated noise process.

Although the techniques presented in this paper could also be applied to the LANS-α and NS-α deconvolution models of turbulence, we have chosen to study only the time-relaxation model. We remark that each of these models employ the same spatial filter to obtain average velocity fields; however, in the time-relaxation model that filter is applied as a relaxation term while in the others it appears as a smoothing term in the nonlinear velocity transport. It would be interesting to use stochastic parameterization to compare these and other turbulence models. As this is outside the scope of the present research, such comparisons are left for future work.

Returning to the task at hand, note that obtaining an improved dynamical model from observational measurements is different but similar to the goal of data-assimilation, whereby incomplete observational measurements are used over time to obtain an improved estimate of the current state for subsequent forecasting. We therefore couple the solutions u and $\tilde{u}$ in a way similar to the direct-insertion method of Charney, Halem and Jastrow [10]. Given observational measurements of u represented by Pu where P is a low-dimensional orthogonal projection, compute q such that $q_0 = 0$ and

$$\frac{dq}{dt} = (I - P)\widetilde{F}(q + Pu) \qquad \text{where} \qquad \frac{du}{dt} = F(u) \tag{2}$$

to obtain an approximation of u given by $q + Pu$. An alternative way to describe this same algorithm may be obtained by writing $U = q + Pu$ and $Q = I - P$ to obtain that

$$\frac{dU}{dt} = Q\widetilde{F}(U) + PF(u) \qquad \text{where} \qquad \frac{du}{dt} = F(u). \tag{3}$$

In real-world applications the projection P will likely include filtering to regularize and remove measurement errors from the observational data. As data assimilation is not our focus, we avoid such difficulties by assuming Pu is given by

a noise-free Fourier projection onto the large scales of the flow. Such an idealized setting was first considered for the two-dimensional incompressible Navier–Stokes equations in a periodic domain by Browning, Henshaw and Kreiss [11], see also Henshaw, Kreiss and Ystrom [12]. In [13] and [14] it was theoretically proved and verified numerically that

$$\|u - U\|_{L^2} \to 0 \quad \text{exponentially as} \quad t \to \infty \tag{4}$$

for projections P of sufficient rank when $\widetilde{F} = F$ and there is no model error.

When $\widetilde{F} \neq F$ and there is model error, the most one can hope for is a bound such as

$$\limsup_{t\to\infty} \|u - U\|_{L^2} \leq E \tag{5}$$

where E is a function of the model error. Ideally, we would like that

$$E \to 0 \qquad \text{as} \qquad \widetilde{F} \to F \tag{6}$$

for projections P such that (4) holds. This is, in fact, the result of Theorem 2 as proved in the present paper and further verified through numerical simulation.

We shall relate the continuous-in-time coupling described by (3) with an easier to compute near-continuous coupling defined as follows. Let $t_n = hn$ for $h > 0$ small and consider

$$\frac{d\tilde{u}}{dt} = \widetilde{F}(\tilde{u}) \qquad \text{for} \qquad t \in [t_n, t_{n+1}) \tag{7}$$

with initial conditions on each interval given by

$$\tilde{u}(t_n) = \begin{cases} Pu_0 & \text{for } n = 0 \\ Pu(t_n) + (I - P)\lim_{t \nearrow t_n} \tilde{u}(t) & \text{for } n > 0. \end{cases} \tag{8}$$

The above algorithm under the assumption $\widetilde{F} = F$ of no model error was studied in [15], see also [16], for the two-dimensional incompressible Navier–Stokes equations. Analysis in those works prove for any $h > 0$ that there is P of sufficient rank such that

$$\|u - \tilde{u}\|_{L^2} \to 0 \quad \text{exponentially as} \quad t \to \infty.$$

When $\widetilde{F} \neq F$ and there is model error, a result similar to (5) should hold. That is, given suitable choices for h and P we expect there to be a bound of the form

$$\limsup_{t\to\infty} \|u - \tilde{u}\|_{L^2} \leq E \tag{9}$$

where E, as before, is a function of the model error.

Upon defining the tendencies

$$\tau^n = (\nu h)^{-1}\big(Pu(t_n) - P\tilde{u}^n\big) \qquad \text{where} \qquad \tilde{u}^n = \lim_{t \nearrow t_n} \tilde{u}(t) \tag{10}$$

one can express the coupling in (8) as $\tilde{u}(t_n) = \nu h \tau^n + \tilde{u}^n$. Here ν is a positive constant – the viscosity in the Navier–Stokes equation – included to scale the values of τ^n in a way that is convenient later. We now make a critical remark:

If the tendencies τ^n could be predicted for t_n in the future, we could then use those tendencies to compute (7) for as far into the future as we like while maintaining the bound described in (9).

Of course, knowing the τ^n arbitrarily far into the future is essentially as difficult as knowing the solution u itself. However, such a rephrasing of the problem indicates how $\widetilde{F}$ needs to be modified to obtain a uniform-in-time approximation of the exact solution far into the future. In particular, we now have a framework in which to construct a statistical model based on the observations $Pu(t_n)$ obtained at times t_n in the past to predict τ^n in the future. The resulting characterization of τ^n can then be used to construct the dynamical model denoted above by G which we shall call the corrected model as well as a corresponding system of stochastic differential equations which make the model error explicit.

This paper is organized as follows. In Section 2 we recall the two-dimensional Navier–Stokes equations and describe the time-relaxation model in detail. We then prove Theorem 2 and finish by testing the sharpness of that theorem using numerical simulation. Section 3 performs the data-driven stochastic parameterization. After a linearly-constrained least-squares fit is performed to find the corrector the remaining mean-zero residuals are then modeled by fitting a time-correlated red-noise process. We also state and prove two propositions relating the near-continuous dynamics used to numerically fit our parameters to the continuous dynamics used to describe the final system of stochastic equations. The paper finishes with some numerical tests in Section 4 comparing the original time-relaxation model to our improved deterministic and stochastic models. This is followed by concluding remarks and directions for future research.

2. Theory and data assimilation

The two-dimensional incompressible Navier–Stokes equations are given by

$$\frac{\partial u}{\partial t} - \nu \Delta u + (u \cdot \nabla)u + \nabla p = f \qquad \text{and} \qquad \nabla \cdot u = 0, \tag{11}$$

where $u = u(x,t)$ is the Eulerian velocity field, ν is the kinematic viscosity, p is the physical pressure and f a time-independent body forcing.

We begin by recalling the Fourier representation of these equations on periodic domains and then describe the time-relaxation model of turbulence. Throughout, we shall employ the functional notation of Constantin and Foias [17], see also Robinson [18]. In particular, define the norms

$$\|u\|_\gamma^2 = 4\pi^2 \sum_{k\in\mathbf{Z}^2} |k|^{2\gamma}|u_k|^2 \qquad \text{for} \qquad u = \sum_{k\in\mathbf{Z}^2} u_k e^{ik\cdot x}$$

where $u_k \in \mathbf{C}^2$ and the corresponding spaces

$$V_\gamma = \left\{ u : k\cdot u_k = 0, \quad u_{-k} = \overline{u_k}, \quad u_0 = 0 \quad \text{and} \quad \|u\|_\gamma < \infty \right\}.$$

The condition $k\cdot u_k = 0$ implies the elements of each space are divergence free, $u_{-k} = \overline{u_k}$ implies they are real valued and $u_0 = 0$ indicates the average velocity is zero.

For simplicity denote

$$|u| = \|u\|_0, \qquad \|u\| = \|u\|_1 \qquad \text{and} \qquad \|u\|_* = \|u\|_{-1}$$

as well as

$$H = V_0, \qquad V = V_1 \qquad \text{and} \qquad V^* = V_{-1}.$$

Note that H is a Hilbert space and V^* is the dual of V with respect to the pairing

$$(u,v) = 4\pi^2 \sum_{k\in\mathbf{Z}^2} \overline{u}_k v_k.$$

Thus, $V \subseteq H \subseteq V^*$ forms a Gelfand triple with continuous, compact and dense injections.

Now, write the Leray–Helmholtz projector and Stokes operator as

$$\Pi u = \sum_{k\in\mathbf{Z}^2} \frac{u_k\cdot(k_2,-k_1)}{|k|^2}\begin{bmatrix} k_2 \\ -k_1\end{bmatrix} e^{ik\cdot x} \quad \text{and} \quad Au = \Pi \sum_{k\in\mathbf{Z}^2} |k|^2 u_k e^{ik\cdot x},$$

respectively, and the bilinear term as

$$B(u,v) = \Pi \sum_{k\in\mathbf{Z}^2} B_k e^{ik\cdot x} \qquad \text{where} \qquad B_k = \sum_{\ell\in\mathbf{Z}^2} i(u_{k-\ell}\cdot\ell)v_\ell.$$

Note that including the projection Π in the definitions of A and B ensures the range of each operator is divergence free. This, along with the assumption that $f \in H$, allows us to express the incompressible two-dimensional Navier–Stokes equations (11) as

$$\frac{du}{dt} + \nu Au + B(u,u) = f. \tag{12}$$

It is known, see [17] or [18] for details, that the incompressible two-dimensional Navier–Stokes equations (12) possess unique strong solutions depending continuously on the initial condition u_0. Namely, we have

Theorem 1. *Let $u_0 \in V$ and $f \in H$. Then (12) has unique strong solutions that satisfy*

$$u \in L^\infty\big((0,T);V\big) \cap L^2\Big((0,T);V_2\Big) \quad \text{and} \quad \frac{du}{dt} \in L^2\big((0,T);H\big)$$

for any $T > 0$. Furthermore, this solution is in $C\big([0,T];V\big)$ and depends continuously on the initial data u_0 in the V norm.

Moreover, these equations possess a unique global attractor $\mathcal{A}$ such that

$$u \in \mathcal{A} \qquad \text{implies} \qquad \|u\| \le \rho_V = \nu\lambda_1^{1/2} G$$

where $G = |f|/(\nu^2\lambda_1)$ is the Grashof number. Here λ_1 has units length^{-2} and is a dimensional constant – the smallest eigenvalue of the Stokes operator – which happens to have the numerical value of 1 since we are working with 2π-periodic functions. We further recall the orthogonality condition for the bilinear term and Ladyzhenskaya bound given by

$$\big(B(u,v),v\big) = 0 \qquad \text{and} \qquad \big|\big(B(u,v),u\big)\big| \le c_1|u|\|u\|\|v\|, \tag{13}$$

respectively, which hold for functions $u, v \in V$. Here c_1 is an absolute constant depending only on the shape of the domain and in the present case satisfies $c_1 \le 2 + (2\pi)^{-1}$.

The time-relaxation model may be obtained from the incompressible two-dimensional Navier–Stokes equations by adding a relaxation term that contains a spatial filter designed to attenuate small scales. We denote that filter by

$$\mathcal{H}\tilde{u} = \sum_{k\in\mathbf{Z}^2} \mathcal{H}_k \tilde{u}_k e^{ik\cdot x} \qquad \text{where} \qquad \mathcal{H}_k = 1 - \frac{\alpha^2|k|^2}{1+\alpha^2|k|^2}$$

and express the time-relaxation model as

$$\frac{d\tilde{u}}{dt} + \nu A\tilde{u} + B(\tilde{u},\tilde{u}) = f + \chi\big(\mathcal{H}\tilde{u} - \tilde{u}\big). \tag{14}$$

The time-relaxation model contains two parameters, α which determines the length-scales upon which the filter acts as well as the relaxation parameter χ. We shall write $\widetilde{F} \to F$ to indicate that the time-relaxation model converges to the two-dimensional incompressible Navier–Stokes equations as either $\alpha \to 0$ or $\chi \to 0$.

Define the orthogonal projection P onto the large-scale Fourier modes appearing in (2) and its complement Q by

$$Pu = \sum_{0<|k|^2\le\lambda} u_k e^{ik.x} \qquad \text{and} \qquad Q = I - P.$$

Here $\lambda > 0$ is a parameter that determines the resolution of observational measurements represented by Pu.

Before turning to the task of proving the theoretical bounds given in (5), we note, after checking that the additional term $\chi\big(\mathcal{H}(\tilde{u}) - \tilde{u}\big)$ does no harm to the estimates, that the well-posedness of (2) follows directly from the proof of Theorem 3.1 in [13] covering the case when $\widetilde{F} = F$. We now state and prove Theorem 2, which not only implies $E \to 0$ as $\widetilde{F} \to F$ but suggests the rate of convergence in terms of α and χ.

Theorem 2. *Suppose the reference solution $u \in \mathcal{A}$ lies on the global attractor of the incompressible two-dimensional Navier–Stokes equations. Let U be the solution to (3) and $\lambda \ge 4c_1^2\nu^{-2}\rho_V^2$. Then*

$$\limsup_{t\to\infty} |u - U| \le \alpha^2\chi C_1 \qquad \textit{where} \qquad C_1 = \frac{2\rho_V}{\nu^{1/2}}.$$

Proof. Setting $w = u - U$ it follows that $Pw = 0$ and $Qw = w$. Moreover,

$$\begin{aligned}\frac{dw}{dt} &= QF(u) - Q\widetilde{F}(U) \\ &= -\nu Aw - QB(u,u) + QB(U,U) - \chi Q(\mathcal{H}U - U).\end{aligned} \tag{15}$$

Differencing yields

$$-B(u,u) + B(U,U) = -B(w,u) + B(w,w) - B(u,w)$$

and

$$-\chi(\mathcal{H}U - U) = \chi(\mathcal{H}w - w) - \chi(\mathcal{H}u - u).$$

In light of the fact that $Qw = w$, the orthogonality (13) implies

$$\big(QB(w,w),w\big) = 0 \qquad \text{and} \qquad \big(QB(u,w),w\big) = 0.$$

Now, multiplying (15) by w and integrating over space yields

$$\frac{1}{2}\frac{d|w|^2}{dt} + \nu\|w\|^2 = -\big(B(w,u),w\big) + \chi(\mathcal{H}w - w, w) - \chi(\mathcal{H}u - u, w).$$

Estimate using Ladyzhenskaya's inequality (13) to obtain

$$\big(B(w,u),w\big) \le |w|_{L^4}^2\|u\| \le \frac{c_1^2}{\nu}|w|^2\rho_V^2 + \frac{\nu}{4}\|w\|^2.$$

Estimate

$$(\mathcal{H}w - w, w) = -4\pi^2 \sum_{|k|^2>\lambda} \frac{\alpha^2|k|^2}{1+\alpha^2|k|^2}|w_k|^2 \le 0 \tag{16}$$

and using similar techniques estimate

$$\begin{aligned}\big\|Q(\mathcal{H}u - u)\big\|_*^2 &= 4\pi^2 \sum_{|k|^2>\lambda} \Big(\frac{\alpha^2|k|^2}{1+\alpha^2|k|^2}\Big)^2 |k|^{-2}|u_k|^2 \\ &= 4\pi^2\alpha^4 \sum_{|k|^2>\lambda} \Big(\frac{1}{1+\alpha^2|k|^2}\Big)^2 |k|^2|u_k|^2 \le \alpha^4\|u\|^2.\end{aligned}$$

It follows that

$$\chi|(\mathcal{H}u-u,w)| \le \|\mathcal{H}u-u\|_*\|w\| \le \frac{\chi^2}{\nu}\|\mathcal{H}u-u\|_0^2+\frac{\nu}{4}\|w\|^2$$
$$\le \frac{\alpha^4\chi^2}{\nu}\|u\|^2+\frac{\nu}{4}\|w\|^2 \le \frac{\alpha^4\chi^2}{\nu}\rho_V^2+\frac{\nu}{4}\|w\|^2.$$

Since $Pw=0$ the improved Poincaré inequality

$$|w|^2=4\pi^2\sum_{|k|^2>\lambda}|w_k|^2 \le 4\pi^2\sum_{|k|^2>\lambda}\frac{|k|^2}{\lambda}|w_k|^2=\lambda^{-1}\|w\|^2$$

now implies

$$\frac{d|w|^2}{dt}+\Big(\nu\lambda-\frac{2c_1^2}{\nu}\rho_V^2\Big)|w|^2 \le 2\alpha^4\chi^2\rho_V^2.$$

By hypothesis $\nu^2\lambda \ge 4c_1^2\rho_V^2$; consequently,

$$\frac{d|w|^2}{dt}+\frac{\nu\lambda}{2}|w|^2 \le 2\alpha^4\chi^2\rho_V^2.$$

Now, multiplying by the integrating factor $e^{\nu t/2}$ yields

$$\frac{d|w|^2e^{\nu\lambda t/2}}{dt} \le 2\alpha^4\chi^2\rho_V^2e^{\nu\lambda t/2}$$

from which it follows that

$$|w|^2 \le |w_0|^2e^{-\nu\lambda t/2}+2\alpha^4\chi^2\rho_V^2e^{-\nu\lambda t/2}\int_0^t e^{\nu\lambda s/2}ds$$
$$=|w_0|^2e^{-\nu\lambda t/2}+\alpha^4\chi^2\Big(\frac{4\rho_V^2}{\nu\lambda}\Big)\big(1-e^{-\nu\lambda t/2}\big). \tag{17}$$

Therefore,

$$\limsup_{t\to\infty}|w|^2 \le \alpha^4\chi^2C_1 \qquad \text{where} \qquad C_1=\frac{4\rho_V^2}{\nu\lambda}.$$

This finishes the proof of the theorem. ■

Note that $\widetilde{F}\to F$ as either $\alpha\to 0$ or $\chi\to 0$ and in both these cases we have that the upper bound $E\to 0$ as mentioned in (6). The theorem, however, also suggests a dependency of E as a function of α and χ. As it will be necessary to compute U for different values of α and χ to find the corrector and resulting stochastic parameterization which form the focus of our research, we pause momentarily and use those same computations to check the sharpness of the bounds found in Theorem 2.

To obtain results related to previous computations we consider the flow conditions studied in [13], see also [14]. Since, it is convenient for the corrector and resulting parameterization to involve fewer parameters, we specifically select a forcing function f and viscosity ν for which the size of λ is modest and for which the resulting fluid is energetic and undergoes complicated time-dependent motion. Thus, we set $\nu=0.0001$ and take the body force to be supported on an annulus in Fourier space as

$$f=\sum_{16\le|k|^2\le 34}f_ke^{ik\cdot x}$$

with the time-independent modes $f_k\in\mathbf{C}^2$ chosen such that

$$k\cdot f_k=0, \qquad f_{-k}=\overline{f_k}, \qquad \text{and} \qquad |f|=0.0025.$$

Note in this case that the Grashof number $G=250\,000$ and that $\lambda=5$ is sufficient in the case $\widetilde{F}=F$ to ensure (4) holds.

All computations were performed with 256×256 Fourier transforms using a spectral Galerkin method consisting of 170×170 dealiased modes, only half of which need to be computed because of the conjugate symmetry. Time was integrated using a 3rd-order Adams–Bashforth method with a step size of 0.0078125. A detailed description of the numerical code appears in [13] and [14]. It is interesting that the dual-processor Pentium III computer on which the original code was developed had an aggregate parallel performance of about 11 time steps per second while the 32-core EPYC server used for the present computations achieves an aggregate parallel performance of 13 184 time steps per second. Although computer performance has increased 1000 fold, the scope of practical problems being solved in applications has increased at an even greater rate. Therefore, methods for constructing approximate dynamics that efficiently yield accurate results are as relevant today as ever.

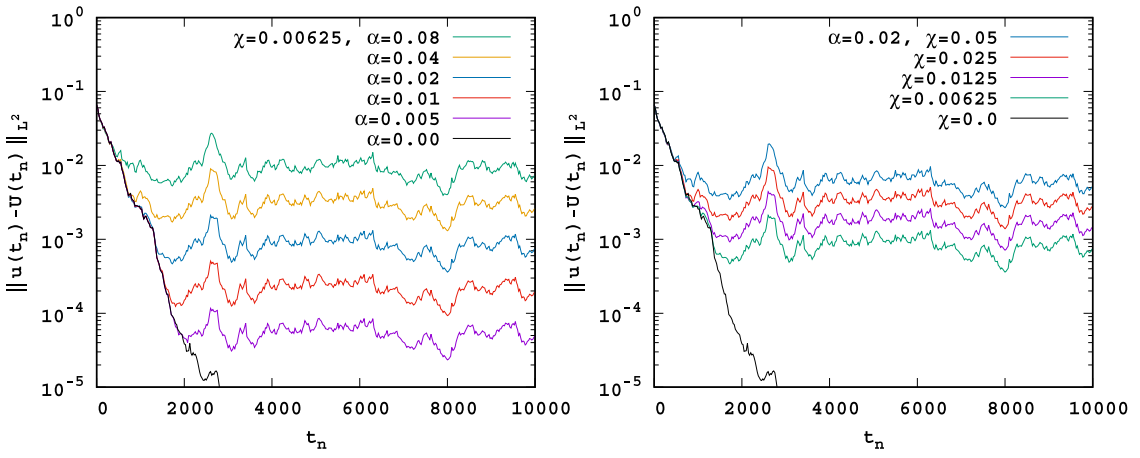

Fig. 1. The degree to which the time-relaxation model synchronizes to the Navier–Stokes equations when coupled on Fourier modes with $|k|^2 \leq 5$. The left shows the dependency on the filter length-scale α; the right considers the relaxation parameter χ.

To obtain an initial condition $u_0 \in \mathcal{A}$ for the reference solution u used in our numerical experiments, we perform a preliminary long-time simulation of only the two-dimensional incompressible Navier–Stokes equations starting at time $t = -25\,000$ in the past. By $t = 0$ the energetics of the flow have long since appeared to enter into a time-varying state with stationary statistical properties. The solution u obtained by moving forward in time from u_0 is consequently presumed to lie close to a point on the global attractor $\mathcal{A}$ and to within the limits of rounding error exactly on the global attractor corresponding to the discretized dynamics. For reference, following the definition appearing in [19], the large-eddy turnover time T_ε of the resulting flow averaged for $t \in [0, 25\,000]$ is $T_\varepsilon = 90.5775$.

Numerical computations to check the sharpness of the bounds in Theorem 2 were performed for values of α and χ such that

$$\alpha \in \{0.0, 0.005, 0.01, 0.02, 0.04, 0.08\} \tag{18}$$

and

$$\chi \in \{0.0, 0.00625, 0.0125, 0.025, 0.05\}. \tag{19}$$

Fig. 1 illustrates the behavior of $|u - U|$ for $U_0 = Pu_0$ where $u_0 \in \mathcal{A}$ is the initial condition described above. For all choices of parameters tested, the norm $|u - U|$ initially decreased at the same exponential rate as the case when $\widetilde{F} = F$. Though not shown in the graphs, when $\widetilde{F} = F$ the difference between $\tilde{u}$ and u continued to decrease at the same exponential rate until reaching the limits of rounding error, that is, until $|u - U| \leq 10^{-15}$. While the theoretical estimate given by (17) also describes an initial decay whose exponential rate is independent of the choice of parameters, it should be noted that the rate $e^{-\nu\lambda t/2}$ predicted by the theory is much slower than observed numerically. The transfer of energy between scales due to the bilinear term B is likely responsible for the rapid convergence of the approximating solution U to the reference solution u; however, the effects of these physics did not play a role in the analysis.

When $\widetilde{F} \neq F$ the evolution of $|u - U|$ levels out after time $t = 2000$ and begins to fluctuate around some average value. Interestingly, the pattern of fluctuation looks quite similar for each of the different values of α and χ tested. Note also that the separation between each of the curves is noticeably greater when varying α compared to varying χ. In particular, the theoretical bound given in Theorem 2 where α appears with a higher power than χ seems to reflect a computational property shown by the graphs.

Before turning to the task of stochastic parameterization, we further examine the theoretical bound in Theorem 2 in relation to the numerical bounds

$$E(\alpha, \chi) = \max\big\{|u - U| : t \in [10\,000, 25\,000]\big\}.$$

Using the values summarized in Table 1, we then suppose that

$$\log E(\alpha, \chi) \approx m_0 + m_1 \log \alpha + m_2 \log \chi$$

and perform a least-squares fit to obtain

$$m_0 \approx 4.46007, \qquad m_1 \approx 1.63203 \qquad \text{and} \qquad m_2 \approx 0.83383.$$

Although, the values of m_1 and m_2 are not exactly equal to 2 and 1, respectively, it is interesting that m_1 is almost exactly the double of m_2.

Table 1
Upper bounds on $\|u - U\|_{L^2}$ where $t \in [10\,000, 25\,000]$ as a function of α and χ for the time-relaxation model.

α	$\chi = 0.00625$	$\chi = 0.0125$	$\chi = 0.25$	$\chi = 0.05$
0.005	1.64938e−04	3.27937e−04	6.48201e−04	1.26636e−03
0.010	6.46036e−04	1.26227e−03	2.41116e−03	4.41967e−03
0.020	2.38330e−03	4.37428e−03	7.76947e−03	1.42511e−02
0.040	7.45131e−03	1.37361e−02	2.32833e−02	3.58971e−02
0.080	2.11713e−02	3.26244e−02	5.40707e−02	8.02256e−02

On the other hand, since $m_1 < 2$ and $m_2 < 1$ then the bounds of Theorem 2 are guaranteed to dominate in the limit. This implies that the values we have estimated for m_1 and m_2 from Table 1 could not hold possibly hold as $\alpha \to 0$ or $\chi \to 0$. Additional computations for smaller values of α and χ resolve this problem but would constitute a distraction too far from our ultimate goal. It is worth noting, however, that a quick reexamination of Fig. 1 indicates the curves for smaller values of α and χ are just perceptibly farther apart.

3. Stochastic parameterization

Having studied what happens when the large-scale Fourier modes of an approximating solution governed by the time-relaxation model are replaced by the corresponding Fourier modes of a direct numerical simulation of the two-dimensional incompressible Navier–Stokes equations, we now look into the energetics needed to realize that coupling as an additional force. To this end, consider the tendencies τ^n defined by (10). We begin by making a connection between the discrete-in-time coupling given in (7) and the continuous-in-time coupling (2) studied in the previous section.

From a numerical point of view, the only thing necessary to identify (7) with (2) is to take h equal to the size of the time steps used by the time integrator. As all our simulations have been performed using a fixed time step of size 0.0078125, we consequently take $h = 0.0078125$. From a physical point of view, this means we are sampling τ^n about 11 600 times per large-eddy turnover. Mathematically, on the other hand, there is a significant difference between near-continuous and continuous coupling.

To provide a mathematical connection between the two consider the limit $h \to 0$. We begin by studying the convergence of the low modes with

Proposition 1. *Suppose for any $T > 0$ fixed that there is M independent of h and large enough such that*

$$|P\widetilde{F}(\tilde{u})| \le M \qquad \text{for all} \qquad t \in [0, T].$$

Then $|P\tilde{u}(t) - PU(t)| \to 0$ as $h \to 0$.

Proof. Since $u \in \mathcal{A}$, then by taking M larger if necessary, we may further assume

$$|PF(u)| \le M \qquad \text{for all} \qquad t \in [0, T].$$

By definition

$$u(t_n) = u(t_{n-1}) + \int_{t_{n-1}}^{t_n} F(u) \quad \text{and} \quad \tilde{u}^n = \tilde{u}(t_{n-1}) + \int_{t_{n-1}}^{t_n} \widetilde{F}(\tilde{u}).$$

Therefore, since (8) implies $P\tilde{u}(t_{n-1}) = Pu(t_{n-1})$, it follows that

$$\nu h \tau^n = P\big(u(t_n) - \tilde{u}^n\big) = \int_{t_{n-1}}^{t_n} P\big(F(u) - \widetilde{F}(\tilde{u})\big). \tag{20}$$

For notational convenience denote by $\delta(t)$ the Dirac impulse and formally rewrite Eqs. (7) governing the evolution of $\tilde{u}$ as

$$\frac{d\tilde{u}}{dt} = \widetilde{F}(\tilde{u}) + \nu \sum_{n=1}^{\infty} \delta(t - t_n) h \tau^n \qquad \text{where} \qquad \tilde{u}_0 = Pu_0. \tag{21}$$

Consequently, for any $t \in [0, T]$ such that $t \ne t_n$ for all n, we have that

$$\begin{aligned} \tilde{u}(t) &= \tilde{u}(0) + \int_0^t \widetilde{F}(\tilde{u}) + \nu \int_0^t \sum_{n=1}^{\infty} \delta(t - t_n) h \tau^n \\ &= Pu_0 + \int_0^t \widetilde{F}(\tilde{u}) + \nu \sum_{t_n < t} \delta(t - t_n) h \tau^n \end{aligned}$$

$$= Pu_0 + \int_0^t \widetilde{F}(\tilde{u}) + \int_0^{t_n} P\big(F(u) - \widetilde{F}(\tilde{u})\big)$$
$$= Pu_0 + \int_0^{t_n} Q\widetilde{F}(\tilde{u}) + \int_0^{t_n} PF(u) + \int_{t_n}^t \widetilde{F}(\tilde{u}).$$

Integrating (3) in the case of continuous coupling yields

$$U(t) = Pu_0 + \int_0^t Q\widetilde{F}(U) + \int_0^t PF(u).$$

Therefore,

$$|P\tilde{u}(t) - PU(t)| \le \int_{t_n}^t |PF(u)| + \int_{t_n}^t |P\widetilde{F}(\tilde{u})|. \tag{22}$$

At the same time, if $t = t_n$ for some n, then $P\tilde{u}(t) = PU(t)$. Thus, the above bound (22) actually holds for all $t \in [0, T]$ and $h > 0$.

Given the hypothesis that both $|PF(u)|$ and $|P\widetilde{F}(\tilde{u})|$ are bounded by M and the fact that $|t - t_n| \le h$ we conclude

$$|P\tilde{u}(t) - PU(t)| \le 2hM \to 0 \qquad \text{as} \qquad h \to 0.$$

This verifies the low modes of $\tilde{u}$ and U agree in the limit. ■

We remark the hypothesis in Proposition 1 on the existence of an M independent of h is a strong one. It is likely such an assumption can be justified when λ is sufficiently large using arguments similar to those appearing in [15]. We leave such things as future research while noting for completeness that when $\lambda = 5$ our numerics fully support the claim that $|\widetilde{F}\tilde{u}|$ is bounded independent of h.

Proposition 2. *Suppose $\lambda \ge 4c_1^2\nu^{-2}\rho_V^2$ as in Theorem 2. Then, there exists a non-decreasing function g where $g(\varepsilon) \to 0$ as $\varepsilon \to 0$ such that*

$$\sup\big\{\, |P\tilde{u}(t) - PU(t)| : t \in [0, T] \,\big\} \le \varepsilon$$

implies $|Q\tilde{u} - QU(t)| \le e^{g(\varepsilon)T} - 1$ for all $t \in [0, T]$ and any $\varepsilon > 0$.

Proof. Since the high modes are governed identically as

$$\frac{dQ\tilde{u}}{dt} = Q\widetilde{F}(\tilde{u}) \qquad \text{and} \qquad \frac{dQU}{dt} = Q\widetilde{F}(U),$$

then setting $W = Q\tilde{u} - QU$ and $\tilde{w} = \tilde{u} - U$ yields

$$\frac{dW}{dt} = -\nu AW - QB(\tilde{u}, \tilde{u}) + QB(U, U) + \chi(\mathcal{H}(W) - W).$$

Thus,

$$\frac{dW}{dt} = -\nu AW - QB(\tilde{w}, \tilde{u}) + QB(\tilde{w}, \tilde{w}) - QB(\tilde{u}, \tilde{w}) + \chi(\mathcal{H}(W) - W).$$

Taking inner product with W yields

$$\frac{1}{2}\frac{d|W|^2}{dt} + \nu\|W\|^2 \le -(B(\tilde{w}, \tilde{u}), W) + (B(\tilde{w}, \tilde{w}), W) - (B(\tilde{u}, \tilde{w}), W).$$

Using

$$|\tilde{w}|^{1/2} \le |W|^{1/2} + |P\tilde{w}|^{1/2} \qquad \text{and} \qquad \|\tilde{w}\|^{1/2} \le \|W\|^{1/2} + \|P\tilde{w}\|^{1/2}$$

estimate as

$$\begin{aligned}
\big|(B(\tilde{w}, \tilde{u}), W)\big| &\le c_1|\tilde{w}|^{1/2}\|\tilde{w}\|^{1/2}|W|^{1/2}\|W\|^{1/2}\|u\| \\
&\le c_1\rho_V|W|\|W\| + c_1\lambda^{-1/4}\varepsilon^{1/2}\rho_V|W|\|W\|^{1/2} \\
&\quad + c_1\varepsilon^{1/2}\rho_V|W|^{1/2}\|W\| + c_1\lambda^{1/4}\varepsilon\rho_V|W|^{1/2}\|W\|^{1/2}, \\
\big|(B(\tilde{w}, \tilde{w}), W)\big| &= |(B(\tilde{w}, P\tilde{w}), W)| \\
&\le c_1|\tilde{w}|^{1/2}\|\tilde{w}\|^{1/2}|W|^{1/2}\|W\|^{1/2}\|P\tilde{w}\| \\
&\le c_1\varepsilon|W|\|W\| + c_1\lambda^{-1/4}\varepsilon^{3/2}|W|\|W\|^{1/2} \\
&\quad + c_1\varepsilon^{3/2}|W|^{1/2}\|W\| + c_1\lambda^{1/4}\varepsilon^2|W|^{1/2}\|W\|^{1/2}
\end{aligned}$$

and

$$\begin{aligned}\big|(B(\tilde{u},\tilde{w}),W)\big| &= |(B(\tilde{u},P\tilde{w}),W)|\\ &\le c_1|\tilde{u}|^{1/2}\|\tilde{u}\|^{1/2}|W|^{1/2}\|W\|^{1/2}\|P\tilde{w}\|\\ &\le c_1\varepsilon\lambda^{1/2}\lambda_1^{1/4}\rho_V|W|^{1/2}\|W\|^{1/2}.\end{aligned}$$

After repeatedly applying the inequality

$$ab \le \frac{|a|^p}{p}+\frac{|b|^q}{q} \qquad \text{where} \qquad \frac{1}{p}+\frac{1}{q}=1$$

for various powers of p and q, we finally obtain that

$$\begin{aligned}\big|(B(\tilde{w},\tilde{u}),W)\big| + \big|(B(\tilde{w},\tilde{w}),W)\big| + \big|(B(\tilde{u},\tilde{w}),W)\big|\\ \le \frac{\nu}{2}\|W\|^2+\frac{c_1^2\rho^2}{\nu}|W|^2+\frac{g(\varepsilon)}{2}\big(1+|W|^2\big)\end{aligned}$$

where $g(\varepsilon)\to 0$ as $\varepsilon\to 0$. It follows that

$$\frac{d|W|^2}{dt}+\Big(\nu\lambda-\frac{2c_1^2\rho^2}{\nu}\Big)|W|^2 \le g(\varepsilon)\big(1+|W|^2\big).$$

This implies

$$\frac{d(1+|W|^2)}{dt} \le g(\varepsilon)\big(1+|W|^2\big).$$

Therefore

$$1+|W(t)|^2 \le \big(1+|W_0|^2\big)e^{g(\varepsilon)t},$$

and since $W_0=0$, then $|W(t)|^2 \le e^{g(\varepsilon)T}-1$ for all $t\in[0,T]$. ■

Since Proposition 1 implies $\varepsilon\to 0$ as $h\to 0$ and hence $e^{g(\varepsilon)T}-1\to 0$, then combining Propositions 1 and 2 together yields that

$$|\tilde{u}(t)-U(t)|\to 0 \qquad \text{uniformly for } t\in[0,T] \text{ as} \qquad h\to 0. \tag{23}$$

Having finished the mathematical discussion connecting near-continuous coupling to the continuous coupling in the limit $h\to 0$, we return to our main focus: Developing a statistical model for predicting the values of τ^n into the future.

Philosophically, our method of prediction is based on the idea that the exact dynamics denoted by F governing the evolution of u are either unknown or computationally too expensive to implement. In the present case, of course, this is not true. The two-dimensional incompressible Navier–Stokes equations are well known and computationally tractable. Though we focus on two-dimensional turbulence, our purpose, in reality, is to better understand how to improve any type of approximate dynamics using data-driven stochastic parameterization based on observational measurements. Thus, we pretend for a moment that the only thing we can compute is $\tilde{u}$ using the time-relaxation model and that we can also make observations Pu of a solution u mysteriously computed by someone else. Now, since it is known that the low Fourier modes given by Pu are sufficient to ensure $\tilde{u}$ stays uniformly close to u, then all we need is to find a way of predicting the values of Pu, or equivalently τ^n, for times in the future. Switching things around, since τ^n depends on Pu and Pu is approximated by $P\tilde{u}$, we suppose that it may be possible to approximate τ^n as a function of $P\tilde{u}$. Thus, our first task is to find a corrector given as a functional relationship between τ^n and $P\tilde{u}^n$.

In our simulations $\lambda=5$ and the rank of P is 20. However, due to the fact that the velocity field is real valued, it holds that $\tilde{u}_{-k}=\overline{\tilde{u}_k}$. Therefore, only 10 complex degrees of freedom are represented by the tendencies τ^n. For definiteness we postulate a functional dependency of τ^n on $P\tilde{u}^n$ such that

$$\tau^n \approx \tilde{\tau}^n \qquad \text{where} \qquad \tilde{\tau}^n=\Omega P\tilde{u}^n+\beta$$

and $\Omega\in\mathbf{C}^{10\times 10}$ and $\beta\in\mathbf{C}^{10}$. This form for $\tilde{\tau}^n$ is not only convenient but it can correct for excess dissipation – a linear phenomenon – and the biases that affect some turbulence models. To maintain a real-valued and divergence-free flow we further impose the constrains that

$$\tilde{\tau}^n_{-k}=\overline{\tilde{\tau}^n_k} \quad \text{and} \quad k\cdot\tilde{\tau}^n_k=0 \qquad \text{for every} \qquad n\in\mathbf{N}.$$

Note that the flow is automatically mean zero since $k=(0,0)$ is not included among the modes onto which P is a projection.

Upon counting the entries in Ω and β, one sees there are 110 complex parameters that need to be found in order construct our corrector. Even though this may seem like a large number of parameters, there is little problem over fitting because of how large we take the time series in n. In particular, we employ $N=3\,200\,000$ values for τ^n and $\tilde{u}^n$ in our

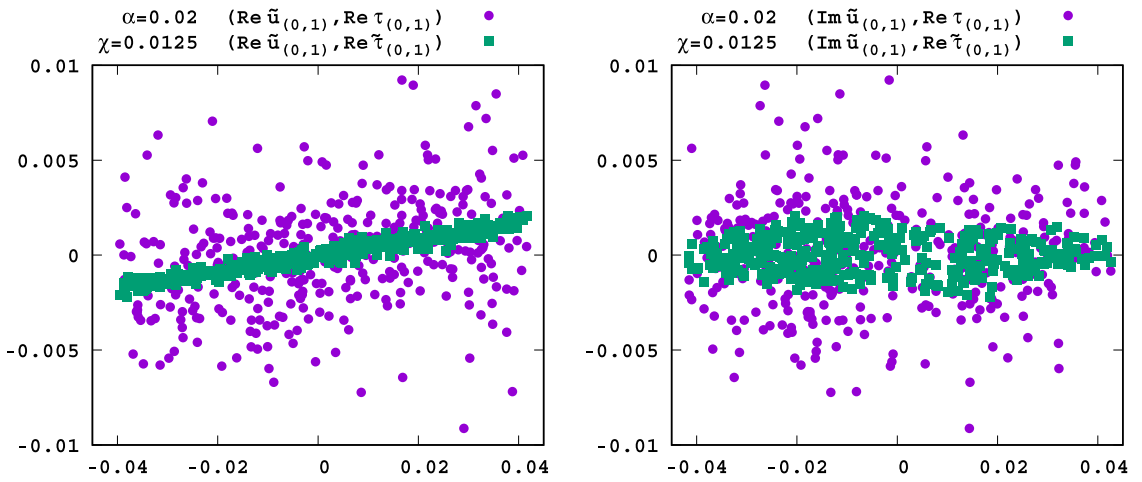

Fig. 2. Comparison of predicted tendencies $\tilde{\tau}_k$ to actual tendencies τ_k viewed as a function of $\tilde{u}_k$ for $k = (0, 1)$. To show detail only data taken at the times t_n for $n = 0, 8000, 16\,000, \ldots, 3\,192\,000$ have been plotted.

calculations. Since $h = 0.0078125$, this means the resulting corrector was obtained by fitting values over the time interval $[0, T]$ where $T = 25\,000$.

In order to model the errors in the predicted tendencies $\tilde{\tau}^n$ with a mean-zero noise process, we impose a further linear constraint on the choice of Ω and β to ensure

$$\frac{1}{N}\sum_{n=1}^{N}\tau^n = \frac{1}{N}\sum_{n=1}^{N}\tilde{\tau}^n. \tag{24}$$

This reduces the number of complex parameters to 100. Forming the matrix needed to perform a least-squares fit results in a $10N \times 100$ complex-valued matrix taking 24 GB of RAM which can be processed on a well-equipped desktop computer. For computational efficiency, similar estimates of Ω and β can be obtained by subsampling the tendency information as

$$\tilde{u}^{n\mu} \quad \text{and} \quad \tau^{n\mu} \qquad \text{for a fixed} \qquad \mu \in \mathbf{N} \tag{25}$$

chosen (for example $\mu = 32$) to ensure the resulting matrices are easy to process. It is, therefore, possible to quickly find approximations for Ω and β corresponding to each of the combinations of α and χ specified by (18) and (19) in the previous section.

To visualize how the tendencies $\tilde{\tau}^n$ predicted by $P\tilde{u}^n$ are related to the actual tendencies τ^n, we create scatter plots made out of the real slices

$$(\operatorname{Re}\tilde{u}_k^n, \operatorname{Re}\tau_\ell^n) \qquad \text{compared with} \qquad (\operatorname{Re}\tilde{u}_k^n, \operatorname{Re}\tilde{\tau}_\ell^n)$$

for values of k and ℓ such that $0 < |k|^2 \le 5$ and $0 < |\ell|^2 \le 5$ as well as

$$(\operatorname{Re}\tilde{u}_k^n, \operatorname{Im}\tau_\ell^n) \qquad \text{compared with} \qquad (\operatorname{Re}\tilde{u}_k^n, \operatorname{Im}\tilde{\tau}_\ell^n)$$
$$(\operatorname{Im}\tilde{u}_k^n, \operatorname{Re}\tau_\ell^n) \qquad \text{compared with} \qquad (\operatorname{Im}\tilde{u}_k^n, \operatorname{Re}\tilde{\tau}_\ell^n)$$

and

$$(\operatorname{Im}\tilde{u}_k^n, \operatorname{Im}\tau_\ell^n) \qquad \text{compared with} \qquad (\operatorname{Im}\tilde{u}_k^n, \operatorname{Im}\tilde{\tau}_\ell^n).$$

As there are 400 such plots, we only reproduce a few representative examples.

Fig. 2 illustrates a pronounced correlation between the reals parts of $\tilde{u}_{(0,1)}$ and $\tau_{(0,1)}$ but little correlation when the real and imaginary parts are mixed. Note that a graph similar to the one on the left reveals a similar correlation between imaginary and imaginary parts; a graph similar to the one of the right indicates little correlation when the real and imaginary parts are mixed the other way around. These observations also hold when $k = \ell$ for other Fourier modes. On the other hand, when $k \neq \ell$ the graphs in Fig. 3 along with other representative examples suggest there is very little correlation between τ_ℓ and $\tilde{u}_k$.

In general, the diagonal terms in the matrix Ω are noticeably larger than the off-diagonal terms. Moreover, as those diagonal terms do not represent significant correlations between the real and imaginary parts, they are also approximately real. Since the filter denoted by $\mathcal{H}$ in the time-relaxation model (14) acts in Fourier space only along the diagonal, it is natural that corrector found by our least-squares regression is concentrated along the diagonal. Moreover, as shown by the analysis in (16) among other places, the relaxation term acts in a purely dissipative way. Therefore, as confirmed by

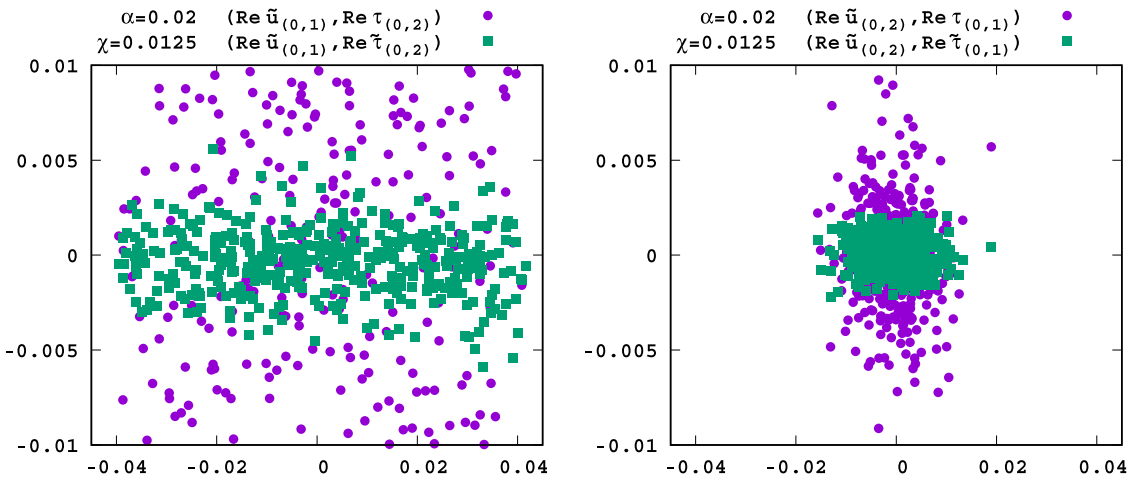

Fig. 3. Comparison of predicted tendencies $\tilde{\tau}_\ell$ to actual tendencies τ_ℓ viewed as a function of $\tilde{u}_k$ for $k, \ell \in \{(0,1),(0,2)\}$. To show detail only data taken at the times t_n for $n = 0, 8000, 16\,000, \ldots, 3\,192\,000$ have been plotted.

the positive slope of the correlation shown on the left in Fig. 2, it seems reasonable that the diagonal of Ω be positive to correct for the extra dissipation.

Before characterizing the errors in our predicted values of τ^n, we explicitly write down the function G which provides the more accurate dynamics described in (1). We first argue, in a somewhat heuristic fashion, that Ω and β, which depend on h, have limits when $h \to \infty$. To see this, first set $h = t/n$ so that $t = t_n$ is fixed and rewrite (20) as

$$\tau(t) = \frac{1}{\nu h}\int_{t-h}^{t} P\big(F(u) - \widetilde{F}(\tilde{u})\big).$$

Under the assumptions where Proposition 1 holds Eq. (23) implies the convergence of $\tilde{u} \to U$ is uniform. For similar reasons $P\widetilde{F}(\tilde{u}) \to P\widetilde{F}(U)$ uniformly. Therefore,

$$\frac{1}{\nu h}\int_{t-h}^{t} P\big(\widetilde{F}(U) - \widetilde{F}(\tilde{u})\big) \to 0 \qquad \text{as} \qquad h \to 0.$$

Also, since both $F(u)$ and $\widetilde{F}(U)$ are continuous functions of t, then

$$\frac{1}{\nu h}\int_{t-h}^{t} P\big(F(u) - \widetilde{F}(U)\big) \to \nu^{-1} P\big(F(u(t)) - \widetilde{F}(U(t))\big) \quad \text{as} \quad h \to 0.$$

Thus, $\tau(t) \to \tau^\circ(t)$ as $h \to 0$ where $\tau^\circ(t) = \nu^{-1} P\big(F(u(t)) - \widetilde{F}(U(t))\big)$.

Let Ω° and β° be chosen to minimize

$$\int_0^T |\tau^\circ(s) - \tilde{\tau}^\circ(s)|^2 ds \qquad \text{where} \qquad \tilde{\tau}^\circ(s) = \Omega^\circ P u(s) + \beta^\circ \tag{26}$$

subject to the constrains that

$$\tilde{\tau}^\circ_{-k}(t) = \overline{\tilde{\tau}^\circ_k(t)} \quad \text{and} \quad k \cdot \tilde{\tau}^\circ_k(t) = 0 \qquad \text{for all} \qquad t \in [0, T]$$

as well as

$$\frac{1}{T}\int_0^T \tau^\circ(s) ds = \frac{1}{T}\int_0^T \tilde{\tau}^\circ(s) ds.$$

Since the definition of τ^n includes h in the denominator, then the least-squares problems used to define Ω and β are exactly the Riemann sums whose limits correspond to (26). Because of this, we further claim without proof that $\Omega \to \Omega^\circ$ and $\beta \to \beta^\circ$ as $h \to 0$.

Under this assumption we now insert $\tilde{\tau}^n$ into (21) to obtain

$$\frac{d\tilde{v}}{dt} = \widetilde{F}(\tilde{v}) + \nu \sum_{n=1}^{\infty} \delta(t - t_n) h(\Omega P \tilde{v}^n + \beta) \quad \text{where} \quad \tilde{v}^n = \lim_{t \nearrow t_n} \tilde{v}(t)$$

and then take the limit as $h \to 0$ to obtain the corrected model

$$\frac{dv}{dt} = G(v) \qquad \text{where} \qquad G(v) = \widetilde{F}(v) + \nu(\Omega^\circ P v + \beta^\circ). \tag{27}$$

Since the dynamics of $\widetilde{F}$ have been modified only in the low modes, the problem of solving for v is still well posed. Whether the system is still dissipative is less clear as indeed it appears the numerical value of Ω found above acts as a source of energy.

We turn now to the task of using a red-noise process of the form

$$\begin{aligned} X_k^n &= \phi_{x,k}^h X_k^{n-1} + \sigma_{x,k}(1-\phi_{x,k}^{2h})^{1/2}\xi_k^n \\ Y_k^n &= \phi_{y,k}^h Y_k^{n-1} + \sigma_{y,k}(1-\phi_{y,k}^{2h})^{1/2}\eta_k^n \end{aligned} \tag{28}$$

where ξ_k^n and η_k^n are independent standard Gaussian random variables to model the errors x_k^n and y_k^n in the predicted tendencies defined by

$$x_k^n + iy_k^n = \tau_k^n - \tilde{\tau}_k^n \qquad \text{where} \qquad 0 < |k|^2 \le \lambda.$$

There are two natural ways to initialize the noise process given in (28): One could simply set $X_k^0 = \sigma_{x,k}\xi_k^0$ and $Y_k^0 = \sigma_{y,k}\eta_k^0$ where ξ_k^0 and η_k^0 are again independent standard Gaussian random variables; alternatively, one could set $X_k^0 = x_k^0$ and $Y_k^0 = y_k^0$ so that $\tilde{\tau}_k^0 + X_k^0 + iY_k^0$ exactly agrees with the last observed tendency τ_k^0 prior to making predictions. The latter makes more sense in the context of data assimilation and consequently the strategy taken in Section 4. We now seek to find $\sigma_{x,k}$, $\sigma_{y,k}$, $\phi_{x,k}$ and $\phi_{y,k}$ for which the statistical properties of X_k and Y_k reflect those of the actual errors x_k and y_k.

Since the noise-processes used for the real and imaginary parts have an identical structure, we describe only how to fit the parameters $\sigma_{x,k} > 0$ and $\phi_{x,k} \in [0, 1]$. Moreover, for simplicity, we omit x from the subscripts and write only σ_k and ϕ_k in what follows. Now, as in [6], we estimate σ_k by computing the variance as

$$\sigma_k^2 = \frac{1}{N}\sum_{n=1}^{N} |x_k^n|^2 \qquad \text{since} \qquad \frac{1}{N}\sum_{n=1}^{N} x_k^n = 0$$

is already ensured by the constraint (24) used to determine Ω and β.

Next, estimate ϕ_k. Since the autocorrelation of the noise process X_k^n is given by

$$\phi_k^{h\ell} = \frac{\mathrm{Cov}(X_k^{n+\ell}, X_k^n)}{\mathbf{V}[X_k^n]},$$

we first estimate the corresponding quantities for the actual errors in the tendencies as

$$\begin{aligned} \gamma_k^\ell &= \frac{1}{N-\ell}\sum_{j=1}^{N-\ell} \big(x_k^j - \langle x_k^0\rangle\big)\big(x_k^{j+\ell} - \langle x_k^\ell\rangle\big) \\ &= \frac{1}{N-\ell}\sum_{j=1}^{N-\ell} \big(x_k^j x_k^{j+\ell} - \langle x_k^0\rangle\langle x_k^\ell\rangle\big) \end{aligned}$$

where

$$\langle x_k^m\rangle = \frac{1}{N-\ell}\sum_{j=1}^{N-\ell} x_k^{j+m}.$$

Although Ω and β were chosen so the average over all values of x_k^n is zero, it is unlikely that every shifted average $\langle x_k^m\rangle$ is also zero. It is important, therefore, to subtract these terms when estimating the autocorrelation.

The next step is to choose ϕ_k such that

$$\phi_k^{h\ell} \approx \gamma_k^\ell/\gamma_k^0 \qquad \text{for all} \qquad \ell \in \mathbf{N}.$$

Unfortunately, the combination of non-linear dynamics and continuous dependence on initial conditions that lead to the time correlations represented by γ_k^ℓ/γ_k^0 are quite different in nature than the exponentially decaying correlations represented by ϕ_k. As a result, it is impossible to obtain a truly convincing fit no matter what value of ϕ_k is chosen. While a more complicated noise process as considered in [7] could lead to a better fit, the noise given by (28) is easier to handle in the limit $h \to 0$ and appears sufficient to quantify the model error remaining in the corrected dynamics given by G as well as predict the rate at which solutions obtained from that model deviate from the exact solution u.

Therefore, to find ϕ_k we first choose ℓ so that $\gamma_k^\ell/\gamma_k^0 \approx \varepsilon$ where ε is a fixed value in $(0, 1)$, and then choose ϕ_k so that $\phi_k^{h\ell}$ has the same value. Specifically, set $\varepsilon = 0.5$ and then interpolate ϕ_k as $\phi_k = \varepsilon^{1/(h\ell+\theta)}$ where ℓ and θ satisfy

$$\frac{\gamma_k^{\ell+1}}{\gamma_k^0} \le \varepsilon < \frac{\gamma_k^\ell}{\gamma_k^0} \qquad \text{and} \qquad \theta = h\log\Big(\frac{\varepsilon\gamma_k^0}{\gamma_k^\ell}\Big)\Big/\log\Big(\frac{\gamma_k^{\ell+1}}{\gamma_k^\ell}\Big).$$

Before carrying out the above procedure, we say a few words about why our estimates for σ_k and ϕ_k are already scaled in such a way as to have limits when $h \to 0$. First, since we already know $\tau^n \to \tau^\circ(t)$, the definition of σ_k can then be

Table 2
Values for the noise in the stochastic parameterization of the time-relaxation model when $\alpha = 0.02$ and $\chi = 0.0125$.

k_1	k_2	$\sigma^2_{x,k}$	$\sigma^2_{y,k}$	$\phi_{x,k}$	$\phi_{y,k}$
0	1	1.05831e−06	9.83669e−07	0.92203	0.92224
0	2	2.26852e−05	2.07332e−05	0.92575	0.92721
1	0	1.01637e−06	1.10255e−06	0.92271	0.92134
1	1	3.69466e−06	3.84067e−06	0.92258	0.91853
1	2	2.23293e−04	2.20599e−04	0.94458	0.94387
2	0	2.16467e−05	2.36145e−05	0.92643	0.92423
2	1	2.25934e−04	2.56874e−04	0.93935	0.94655
−2	1	2.59807e−04	2.28193e−04	0.94238	0.93992
−1	1	4.24431e−06	3.83753e−06	0.91897	0.92331
−1	2	2.13486e−04	2.23961e−04	0.93927	0.94268

interpreted as a Riemann sum which converges to

$$(s^\circ)^2 = \int_0^T |\tau^\circ(s) - \tilde{\tau}^\circ(s)|^2 \qquad \text{as} \qquad h \to 0.$$

For ϕ_k the argument is somewhat different. Express the red-noise process in terms of h as

$$X_k(t+h) = \phi_k^h X_k(t) + \sigma_k(1-\phi_k^{2h})^{1/2}\xi_k(t+h)$$

and then iterate to obtain

$$\begin{aligned} X_k(t+2h) &= \phi_k^h X_k(t+h) + \sigma_k(1-\phi_k^h)^{1/2} + \xi_k(t+h) \\ &= \phi_k^{2h} X_k(t) + \sigma_k(1-\phi_k^{4h})^{1/2}\zeta_k(t+2h) \end{aligned}$$

where

$$\zeta_k(t+2h) = \frac{\xi_k(t+h) + \xi_k(t)}{(1+\phi_k^{2h})^{1/2}}$$

is again a standard Gaussian random variable. Notice, since the exponent on ϕ_k scales exactly as the increment in time with no further adjustments needed to the constants, that this suggests – again without rigorous proof – there is some ϕ_k° such that $\phi_k \to \phi_k^\circ$ as $h \to 0$.

In summary, having verified that σ_k and ϕ_k have no obvious dimensional dependencies on h, we assume their limits exist as $h \to 0$ and for reference write down the corresponding continuous in time Ornstein–Uhlenbeck process

$$dX_k = (\log \phi_k^\circ) X_k\, dt + \sigma_k^\circ(-2\log\phi_k^\circ)^{1/2} dW_k$$

where the W_k are independent standard Brownian motions. Upon defining $dZ_k = dX_k + i dY_k$ we obtain the full stochastic model

$$dV = \tilde{F}(V)dt + \nu(\Omega^\circ PV + \beta^\circ)dt + \nu dZ. \tag{29}$$

While it would be nice to make the above arguments rigorous, we leave such things for future work and return to the numerics of the present paper.

Numerical values of $\sigma^2_{x,k}$, $\sigma^2_{y,k}$, $\phi_{x,k}$ and $\phi_{y,k}$ are tabulated in Table 2. Note that the time correlations in the noise processes denoted by $\phi_{x,k}$ and $\phi_{y,k}$ are similar for each of the Fourier modes and have an average values of about 0.9307. Thus, the estimate

$$\phi_k^{T_\varepsilon} \approx (0.9307)^{90.5775} \approx 0.0014958$$

shows the tendencies before and after one large-eddy turnover are essentially uncorrelated. This is strikingly similar to the result reported in [9], that after one large-eddy turnover the residual error in the LANS-α model is also uncorrelated.

Table 2 indicates that the variances $\sigma^2_{x,k}$ and $\sigma^2_{y,k}$ vary over two decimal orders of magnitude. To measure the overall effectiveness of our corrector in predicting the tendencies we compute the total variance

$$\Sigma^2(\alpha,\chi) = \sum_{0<|k|^2\le 5} \left(\sigma^2_{x,k} + \sigma^2_{y,k}\right)$$

for the values of α and χ given by (18) and (19). Table 3 lists values for $\Sigma^2(\alpha,\chi)$ computed with respect to the least squares fit of Ω and β obtained by subsampling the tendencies as suggested in (25) by the factor $\mu = 32$. A few of these entries were compared to the results obtained when Ω and β were found using the complete time series with no appreciable differences. We remark that the values of $\Sigma^2(\alpha,\chi)$ depend more strongly on α than they do as a function of χ. This is reminiscent to how $E(\alpha,\chi)$ in Table 1 depended on α and χ.

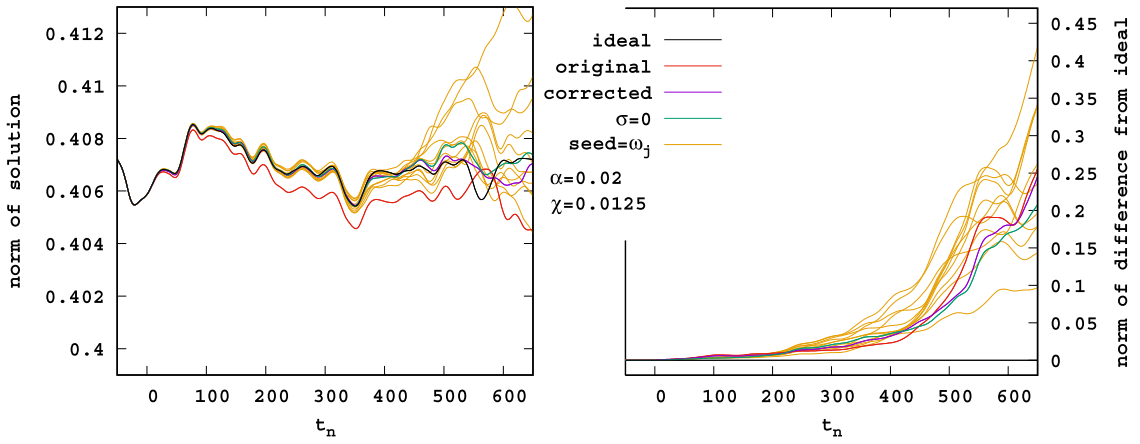

Fig. 4. The ideal solution constructed using the actual tendencies compared with the original time-relaxation model, the corrected model, the stochastic model with $\sigma = 0$ and eleven different realizations of the full stochastic model.

Table 3
Values of the total variance $\Sigma^2(\alpha, \chi)$ of the error in the predicted tendencies of the time-relaxation model.

α	$\chi = 0.00625$	$\chi = 0.0125$	$\chi = 0.25$	$\chi = 0.05$
0.005	2.08738e−06	8.33508e−06	3.32248e−05	1.31978e−04
0.010	3.29975e−05	1.31081e−04	5.17078e−04	2.01083e−03
0.020	5.03846e−04	1.96061e−03	7.41123e−03	2.63869e−02
0.040	6.79747e−03	2.43741e−02	7.87789e−02	2.18037e−01
0.080	6.32765e−02	1.82659e−01	4.38065e−01	8.69494e−01

4. Improved predictions

In this section we compare the energetics of the corrected model (27) to the original time-relaxation model and then check whether the corrected model can be used to better predict the future. We also check whether the full stochastic equations (29) well-characterize the model errors in our approximate dynamics by computing spaghetti plots made using different realizations of the underlying noise process.

To do this we first compute a solution using (7) based on the exact tendencies τ^n which couple the low Fourier modes in two-dimensional incompressible Navier–Stokes equations to the corresponding modes in the time-relaxation model at every time step. Such a computation was performed using every choice of α and χ satisfying (18) and (19) for 25 000 units of time—approximately 275 large-eddy turnovers. Then we turn off the coupling and make a prediction into the future based on the current state in the following ways: By computing

- A free-running solution using the original time-relaxation model (14).
- An improved solution using the corrected dynamics (27).
- A deterministic version of (29) obtained by taking $\sigma = 0$.
- An ensemble of stochastic solutions to (29) computed with respect to different realizations of the noise process.

To use as a reference for the others, we also to compute the coupled solution (7) forward in time and call this the ideal solution.

The graphs on the left in Fig. 4 and in Fig. 5 depict the time evolutions of the L^2 norms of the solutions described above. When $\alpha = 0.02$ and $\chi = 0.125$ the norm of the solution computed by the original dynamics follows that of the ideal solution for about one large-eddy turnover and then proceeds at a reduced level still reflecting the same pattern of fluctuations until about $t = 450$. On the other hand, the norm of the solutions produced by the corrected dynamics and stochastic models do not exhibit any increased levels of dissipation and track the norm of the ideal solution quite closely to time $t = 450$. After that the norms diverge, but the average levels of the L^2 energy remain comparable. When $\alpha = 0.04$ and $\chi = 0.025$ the excess dissipation in the original time-relaxation model becomes evident almost immediately and norm of that solution runs off the bottom of the graph at $t = 350$. It continues to decrease until the end of the computation at $t = 700$. As in the previous case, the corrected and stochastic solutions maintain the same level of energetics on average as the ideal solution throughout most of the interval.

Before examining the graphs on the right, we also notice the ensemble of stochastic solutions includes both the ideal, the corrected and the deterministic solution with $\sigma = 0$ in their envelope. Moreover, the different members of the ensemble start to diverge from each other at a time and rate similar to the way they diverge from the ideal solutions.

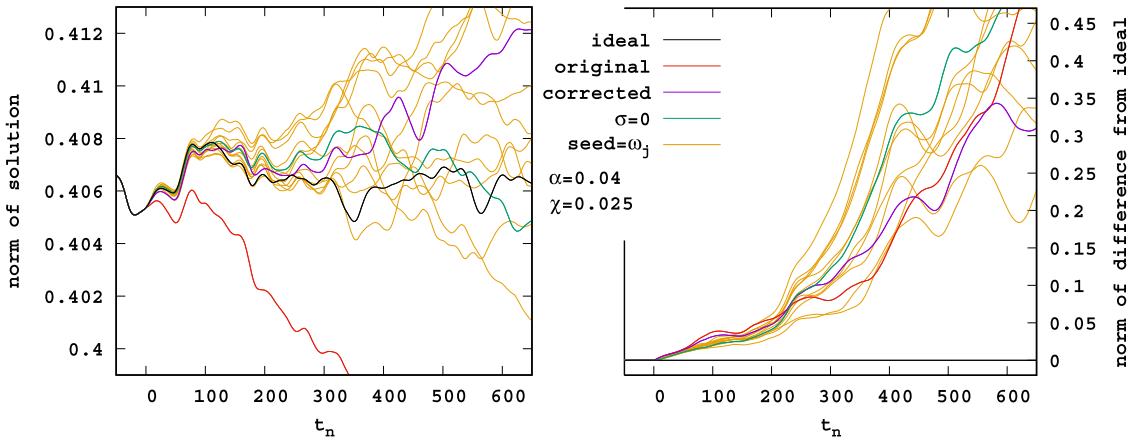

Fig. 5. The same comparison as Fig. 4 with different values for the parameters α and χ in the time-relaxation model.

Thus, one does not need the ideal solution to estimate the errors in the prediction. This is important, because for practical problems the ideal solution is known only in retrospect through data-assimilation.

The graphs on the right in Figs. 4 and 5 depict the L^2 norm of differences between the ideal solution and the original, the corrected, the case when $\sigma = 0$ and the fully stochastic solutions. Since the trajectories obtained by the deterministic solutions again fall within the envelope of stochastic solutions, one can again estimate the error in our predictions from the stochastic solutions, for example, by computing the norms of the differences between members of the stochastic ensemble.

When $\alpha = 0.02$ and $\chi = 0.0125$ we are happy to note that the solution obtained from the corrected dynamics as well as the deterministic solution when $\sigma = 0$ both remain closer to the idea solution than the solution obtained from the original time-relaxation model over the whole interval shown by the graph. However, such happiness is short lived, because when $\alpha = 0.04$ and $\chi = 0.025$ the situation is reversed: the solution computed using the original dynamics stay closer to the ideal solution over much of the graph up to time $t = 450$. The reason for this warrants further study. On one hand this may be an effect of deterministic chaos resolved positively by repeatedly performing the same experiment using different initial conditions; on the other, it could be a limitation of the methods used.

5. Conclusions

This paper improved the approximate dynamics of the time-relaxation model of turbulence by creating a corrector and stochastic parameterization obtained by coupling the time-relaxation model through observational measurements to a direct-numerical simulation of the two-dimensional incompressible Navier–Stokes equations. The final result was a system of stochastic differential equations representing an improved dynamical model with more accurate energetics, less bias and the ability to estimate errors in the resulting predictions. While an improved ability to predict the exact flow was also observed in some cases, such improvements were not universal.

We have attempted to be as rigorous as possible; however, the present paper is of a numerical nature with rigorous justification of theoretical results planned for the future. In particular, there is still significant work to be done, for example, to fully characterize the circumstances under which the hypothesis of Proposition 1 hold. Moreover, we have argued in a heuristic way when passing to the limit as $h \to 0$. It would appear that providing fully rigorous analysis connecting the continuous and near-continuous dynamical systems discussed in this paper is an important direction for further research.

We remarked in the beginning that the error in turbulence models is correlated in time as well as space. The present paper characterized the time correlations in terms of a red-noise process but did not address correlations over short distances in space. We note, however, that the noise process appearing in our parameterization is concentrated only on the Fourier modes such that $0 < |k|^2 \leq 5$. Because of this, we automatically obtain spatial correlations. While there is reason to believe these spatial correlations are on the right scales, this has not been verified. Moreover, it is not clear such spatial correlations remain invariant if a larger number of Fourier modes, for example those with $0 < |k|^2 \leq 10$, were used for the parameterization instead.

Note also, since the parameterization considered in this study involves only the large-scale Fourier modes where $0 < |k|^2 \leq 5$, that this parameterization does not change the dynamics of the inertial-range. In particular, this implies the energy and enstrophy cascades of the original and corrected time-relaxation turbulence models should be the same, as well as the scaling of the corresponding energy spectra.

We close with a speculation that average properties such as the norm may better reflect the physical questions people are interested in than the exact velocity field. In meteorology, for example, people want to know what the temperature

will be and how much it will snow, rather than the exact position of the molecules which make up the air. Therefore, in light of the mixed success in obtaining a better prediction of the velocity field itself, it is comforting to reflect on how much improved were the predictions of the L^2 energetics.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgments

The author would like to thank the organizers of the Conference on Computational Mathematics and Applications held in Las Vegas on October 2019. Particular thanks goes Monika Neda, who not only invited me to talk on the material which later became this paper, but who also introduced me to the time-relaxation model of turbulence. The author would also like to thank the anonymous referees for their helpful comments.

References

[1] Stolz S, Adams NA, Kleiser L. An approximate deconvolution model for large eddy simulation with application to wall-bounded flows. Phys Fluilds 2001;13(4):997–1015.
[2] Layton W, Neda M. Truncation of scales by time relaxation. J Math Anal Appl 2007;325:788–807.
[3] Neda M, Pahlevani F, Waters J. Sensitivity analysis and computations of the time relaxation model. J Adv Appl Math Mech 2015;7:89–115.
[4] Dunca AA, Neda M. Numerical analysis of a nonlinear model of fluids. J Math Anal Appl 2005;4:1095–115.
[5] Hill T, Neda M, Olson E, Pahlevani F. Energy and Enstrophy Study of the Time Relaxation Model, submitted for publication.
[6] Wilks DS. Effects of stochastic parametrizations in the Lorenz 96 system. Q J R Meteorol Soc 2005;131:389–407.
[7] Lu F, Lin KK, Chorin AJ. Data-based stochastic model reduction for the Kuramoto–Sivashinsky Equation. Physica D 2017;340:46–57.
[8] Mohebujjaman M, Rebholz LG, Iliescu T. Physically constrained data-driven correction for reduced-order modeling of fluid flows. Internat J Numer Methods Fluids 2019;89(3):103–22.
[9] Olson E. Stochastic model error in the LANS-alpha and NS-alpha deconvolution models of turbulence. Int J Numer Anal Model 2018;(6):811–33.
[10] Charney J, Halem M, Jastrow R. Use of incomplete historical data to infer the present state of the atmosphere. J Atmos Sci 1969;26:1160–3.
[11] Browning G, Henshaw W, Kreiss H. A numerical investigation of the interaction between the large and small scale of the two-dimensional Navier–Stokes Equations. Technical report 98–23, Los Angeles, CA: Department of Mathematics, University of California; 1998.
[12] Henshaw WD, Kreiss H, Ystrom J. Numerical experiments on the interaction between the large and small-scale motions of the Navier–Stokes Equations. Multiscale Model Simul 2003;1(1):119–49.
[13] Olson E, Titi ES. Determining modes for continuous data assimilation in 2D turbulence. J Stat Phys 2003;113(5–6):799–840.
[14] Olson E, Titi ES. Determining modes and grashof number in 2D Turbulence—A numerical case study. Theor Comput Fluid Dyn 2008;22(5):327–39.
[15] Hayden K, Olson E, Titi ES. Discrete data assimilation in the Lorenz and 2D Navier–Stokes Equations. Physica D 2011;240(18):1416–25.
[16] Celik E, Olson E, Titi ES. Spectral filtering of interpolant observables for discrete-in-time data assimilation. SIAM J Appl Dyn Syst 2019;18(2):1118–42.
[17] Constantin P, Foias C. Navier–Stokes Equations. University of Chicago Press; 1988.
[18] Robinson J. Infinite-dimensional dynamical systems. Cambridge texts in applied mathematics, 2001.
[19] Gesho M, Olson E, Titi ES. A computational study of a data assimilation algorithm for the two-dimensional Navier–Stokes Equations. Commun Comput Phys 2016;19(4):1094–110.