The temporal cascade structure of reanalyses and Global Circulation models

J. Stolle¹, S. Lovejoy¹, D. Schertzer²,³

(1) Physics, McGill University, 3600 University St., Montreal, Que. H3A 2T8, Canada

(2) CEREVE, Université Paris Est, France

(3) Météo France, 1 Quai Branly, Paris 75005, France

Correspondence to: J. Stolle (stollej@physics.mcgill.ca), S. Lovejoy (lovejoy@physics.mcgill.ca)

Abstract

The spatial stochastic structure of deterministic models of the atmosphere has recently been shown to be well modelled by multiplicative cascade processes; in this paper we extend this to the time domain. Using data from reanalyses (ERA 40) and two meteorological models (GFS, GEM), we investigate the temporal cascade structures of the temperature, humidity, and zonal wind at various altitudes, latitudes, and forecast times. First, we estimate turbulent fluxes from the absolute second order time differences, showing that the fluxes are generally very close to those estimated in space at the model dissipation scale; thus validating the flux estimates. We then show that temporal cascades with outer scales typically in the range 5-20 days can accurately account for the statistical properties over the range from six hours up to 3-5 days. We quantify the (typically small) differences in the cascades from model to model, as functions of latitude, altitude and forecast time. By normalizing the moments by the theoretical predictions for universal multifractals, we investigate the “Levy collapse” of the statistics somewhat beyond the outer cascade limit into the low frequency weather regime.

Although the due to finite size effects and the small outer temporal scale, the temporal scaling range is narrow (12-24 hours up to 2-5 days), we compare the spatial and temporal statistics by constructing space-time (Stommel) diagrammes, finding space-time transformation velocities of 450 - 1000 km/day, comparable to those predicted based on the solar energy flux driving the system. This transition time scale corresponding to planetary size structures objectively defines the transition from usual weather to a low frequency weather regime with
much lower variability. Finally, we discuss the implications for ensemble forecasting systems, stochastic parametrization and stochastic forecasting.

1. Introduction

“Weather prediction by Numerical Process” (Richardson, 1922)) is widely celebrated as the pioneering work laying the basis of modern numerical weather forecasting. In it, Richardson eschewed approximations and attempted a brute force numerical integration of the dynamical fluid equations. While these equations are deterministic, numerical weather prediction has been increasingly transformed into sophisticated Ensemble Forecast Systems (EFS). This modern incarnation of “Richardson’s Dream” (Lynch, 2006) has the stochastic objectives of predicting the future states of the atmosphere as well as their probabilities of occurrence. EFS models involve stochastic subgrid parametrisations (Buizza et al., 1999; Palmer, 2001; Palmer and Williams, 2010) and therefore require knowledge of the stochastic properties of the deterministic models.

Interestingly, Richardson is not only the father of numerical weather forecasting, he is also – thanks to his iconic poem describing cascades – considered to be the “grandfather” of cascade approaches. Following the development of the classical turbulence laws of Richardson, Kolmogorov, Obhukhov, Corrsin and Bolgiano and starting in the 1960’s, these have spawned explicit stochastic cascade models. Today these are well understood and are known to be the generic multifractal process. Multifractal intermittency has been reported widely in laboratory and small scale atmospheric experiments from the mid-1980’s onwards (see the review (Anselmet, 2001)). Recently the deterministic and stochastic approaches have been shown to be surprisingly compatible. (Stolle et al., 2009) analyzed both meteorological reanalyses (the ERA40) and meteorological forecasting models (the GFS, GEM models) and shown that the spatial statistics of the turbulent fluxes associated with the temperature, zonal wind and humidity were indeed very close to those predicted by cascade models with deviations typically $\pm 1\%$ over the range of 5000 $km$ down to the model/reanalysis (hyper) viscous dissipation scales. Lovejoy and Schertzer (2011b) found similar results for the ECMWF interim reanalysis products extending these results to the geopotential height, and to the vertical and meridional winds the
fields as well as corresponding turbulent fluxes. At the same time, the increasing availability of global scale remote and in situ data has made it possible to directly confirm that these also have qualitatively and quantitatively very similar cascade structures also up to planetary scales. The similitude of the global scale data, models and reanalyses provides the basis for understanding atmospheric variability over huge spatial scales, see the review: (Lovejoy and Schertzer, 2010) and also (Lovejoy and Schertzer, 2011a). Wide scale range cascade processes are possible because it appears that the vertical structure of the atmosphere is scaling but with different exponents than in the horizontal. This implies that the cascades are anisotropic with structures getting progressively flatter at larger and larger scales but in a scaling manner ((Lovejoy et al., 2007), (Lovejoy et al., 2009a)).

These model, reanalysis, in situ and remote sensing studies have shown that the wind and other atmospheric fields are scaling over wide ranges (see (Lovejoy et al., 2009b) for the reinterpretation of aircraft measurements of wind). The wind is critical since physically it advects the fields and dimensionally it connects space and time. This suggests that the fields should be scaling in time as well as in space. It even seems that the corresponding space-time relations can even be quite accurately deduced from first principles, the argument (Lovejoy and Schertzer, 2010) goes as follows. Starting from the top of the atmosphere, there is roughly 1 kW/m² of incident solar energy flux. Of this, about 200 W/m² is absorbed by the earth (Monin, 1972). If we distribute this over the troposphere (thickness ≈ 10⁴ m), with an air density ≈ 1 Kg/m³, and assume a 2% conversion of energy into kinetic energy ((Palmén, 1959); (Monin, 1972)), then we obtain a value ε ≈ 4X10⁻⁴ m²/s³ which is indeed the typical value measured in small scale turbulence; see the review in (Lovejoy and Schertzer, 2010) where this and more modern data yield the global estimate ε ≈ 10⁻³ m²/s³; although with some variation with altitude and latitude (note, m²/s³ = W/kg). If we now use the Kolmogorov relation Δv(Δx) = ε¹/³ Δx¹/³ – which is apparently valid in the horizontal (but not vertical) up to near planetary scales – then we find that using Le = Δx ≈ 20000 km, Δv≈ 20 m/s and the corresponding lifetime of structures (“the eddy turn-over time”) is τw= Le²/³ε⁻¹/³ ≈ 10 days. While the predicted velocity difference is very close to the mean planetary antipodes differences (17.3±5.7 m/s), the lifetime is quite close to the transition from the high frequency weather regime to a lower frequency weather-climate “spectral plateau” with much flatter spectra (Lovejoy and Schertzer, 1986). Fig. 1 shows that for
instrumental surface temperatures, the implied “dimensional” transition” does indeed occur near this scale. Similarly – although the data is much sparser – the ocean appears to have global (near surface) energy flux $\varepsilon_o \approx 10^{-8} \text{ m}^2/\text{s}^3$ implying (again using $L_e$) a critical ocean time scale $\tau_o \approx 1$ year, confirmed in Fig. 1 using sea surface temperature measurements (this estimate is from ocean drifter data; see (Lovejoy and Schertzer, 2011a) for this and for a comprehensive review). For shorter scales, the spectra of “ocean weather” sea surface temperatures are very similar to atmospheric weather temperature spectra whereas at longer time scales they also display a marked flattening, a spectral plateau although with a larger (absolute) spectral exponent.

Fig. 1: This figure superposes the ocean and atmospheric spectral plateaus showing their great similarity. Left: A comparison of the monthly sea surface temperatures (SST) spectrum (bottom, blue) and monthly atmospheric temperatures over land (top, purple) for monthly temperature series from 1911-2010 on a 5°x5° grid; from NOAA NCDC data Smith 2008. Only those near complete series (missing less than 20 months out of 1200) were considered; 465 for the SST, 319 for the land series; the missing data were filled using interpolation. The reference slopes correspond to spectra of the form $E(\omega) \approx \omega^\beta$ with $\beta = 0.2$ (top), 0.6 bottom left and 1.8, bottom right. A transition at 1 year corresponds to a mean ocean $\varepsilon_o \approx 1 \times 10^{-8} \text{ m}^2/\text{s}^3$. Note the apparent beginning of a low frequency rise at around (30 yrs)$^{-1}$.

Right: The average of 5 spectra from a sections 6 years long of a thirty year series from daily temperatures at a station in France (black, taken from Lovejoy and Schertzer (1986)). The red reference line has a slope 1.8 (there is also a faint slope 0 reference line). The relative up-down placement of this daily spectrum with the monthly spectra
(corresponding to a constant factor) was determined by aligning the atmospheric spectral plateaus (i.e. the black and purple spectra).

The upshot of this is that we expect the meteorological models and reanalyses to display a scaling cascade structure for high “weather” frequencies, \( \omega > \omega_w = \tau_w^{-1} \) and that this will break down at scales somewhere in the vicinity of 10 days, with \( \omega < \omega_r \) “low frequency weather” following a relatively flat “spectral plateau” (Lovejoy and Schertzer 1986). Since the corresponding model and reanalysis outputs are only available every 6 hours at best - and for many (especially near surface) fields there is a technical analysis difficulty due to the diurnal cycle - one can anticipate that the scaling range will be rather limited. Nevertheless in this paper we attempt to quantitatively characterize it, compare it with the corresponding spatial cascades and to intercompare the various products.

This paper is structured as follows. In section 2 we present the numerical outputs which we analyse and explain the technique for estimating the fluxes in both space and in time. In section 3 we recapitulate pertinent spatial cascade results from (Stolle et al., 2009), we then systematically examine the temporal cascade structure for the different fields (horizontal wind, temperature, humidity), different models and as functions of altitude. In section 4 we compare the statistics of the fluxes estimated in the spatial domain at the dissipation scale with those estimated in the temporal domain in the scaling regime. We then discuss space-time (“Stommel”) diagrammes and show how to empirically construct them from the outputs using the spatial and temporal scale by scale flux statistics. We comment on the large space limit and the transition from weather to climate. Finally in section 5 we conclude.

2. Multiplicative cascades, turbulent fluxes and analysis techniques

2.1 Multiplicative cascades

2.1.1 Spatial cascades

During the 1960’s and early 1970’s, intermittency was increasingly acknowledged as an important phenomenon, but its effect was usually considered small, leading primarily to small “intermittency corrections” to the spectral exponents. The main statistical models (such as those used in statistical closures) assumed nonintermittent “quasi-Gaussian statistics”. The Gaussian
model leading to the classical Kolmogorov law \( \Delta v = \varepsilon^{1/3} \Delta x^{1/3} \) (for velocity fluctuations \( \Delta v \) over distances \( \Delta x \)) is obtained by taking a Gaussian white noise energy flux \( \varepsilon \) and giving \( \varepsilon^{1/3} \) a (fractional) integration of order 1/3 (i.e. a power law filter of order -1/3); the resulting \( \nu \) is a “fractional Brownian motion”.

In order to take into account intermittency, it suffices to replace the Gaussian \( \varepsilon \) by the result of a multiplicative cascade; this is the Fractionally Integrated Flux model (Schertzer and Lovejoy, 1987). In multiplicative cascades, large structures are broken up into smaller daughter structures which multiplicatively modulate the flux; this process is repeated to smaller and smaller scales (see Fig. 2 for a schematic). Normalized cascade processes generally lead to multifractal fields with statistics:

\[
\left\langle \phi_{\lambda}^q \right\rangle = \lambda^{K(q)}; \quad \lambda = L / l
\]  

(1a)

where “\( \langle \rceil \rangle \)” indicates ensemble (statistical) averaging, \( \phi \) is the turbulent flux (e.g. the energy flux \( \varepsilon \)) normalized such that \( \langle \phi_{\lambda} \rangle = 1 \), \( K(q) \) is a convex function describing the scaling behaviour the \( q \)th moment, \( \lambda \) is the ratio of the (large) scale \( L \) where the cascade starts to the scale of observation \( l \). Note that in the quasi-Gaussian (nonintermittent) classical model we have the trivial \( K(q) = 0 \).
Figure 2: The left of the figure shows a schematic of an isotropic cascade. Due to nonlinear interactions with other eddies or due to into instability, a large eddy/structure (indicated as a square) breaks up into daughter eddies (smaller squares). Following the left most arrow the energy flux is redistributed uniformly in space, the result is a homogeneous (non fractal) cascade. Following the right hand arrow, at each cascade step, we randomly allow one eddy in four to be “dead”, the result is that turbulence is only active on a fractal set. At the bottom, we see the average shape as a function of scale of more realistic (isotropic eddies). The right hand column is the same except that it shows an anisotropic cascade, a model of a vertical cross-section of the atmosphere (on the left a homogeneous, on the right, inhomogeneous, fractal model). The degree of stratification is characterized by an elliptical dimension $D_{el} = 1.5$ in the example. From Lovejoy and Schertzer (2010), adapted from Schertzer and Lovejoy (1987).
2.1.2 Universality classes

Since the cascade is multiplicative, its logarithm, the “generator” is additive. It is therefore not surprising that – due to the additive central limit theorem for the sums of identical independently distributed random variables – there are specific (stable, attractive) “universal” forms for the exponent $K(q)$:

$$K(q) = \frac{C_1}{(\alpha - 1)} (q^\alpha - q)$$  \hspace{1cm} (1b)

where $0 \leq C_1 \leq d$ is the “codimension of the mean”, which characterizes the sparseness of the set that gives the dominant contribution to the first order statistical moment (the mean), $d$ is the dimension of the space over which the cascade is observed (Schertzer and Lovejoy, 1987). The multifractal index $0 \leq \alpha \leq 2$ characterizes the degree of multifractality, i.e. the shape of the $K(q)$ function. It is also the Levy index of the generator. If the cascade is uni/mono-fractal, then $\alpha = 0$ whereas $\alpha = 2$ corresponds to the ‘lognormal’ multifractal. A “universal multifractal” is the basin of attraction for a wide variety of different multiplicative processes. In our analyses, we will see that the universal form (Eq. (1b)) fits the empirical $K(q)$ quite well so that irrespective of whether the numerical models are indeed universal multifractals, the parameters $C_1$, $\alpha$ give very convenient parameterizations of their forms and characterize the sparseness near the mean ($C_1$) and the curvature (and hence multifractality) near the mean ($\alpha$). In this paper we therefore often use $C_1$, $\alpha$ to reduce the characterization of the scaling (via the exponents $K(q)$) to manageable proportions. In any event, we show that over significant ranges of scale, the data display “Levy collapses” when the moments are normalized by the theoretical universal $K(q)$, and this even for scales somewhat outside the scaling range (see Stolle et al. (2009) and section 5 below). In other words, the universality relation Eq. (1b) may be respected even when the scaling relation (1a) is not.

2.1.3 Space-time cascades, causality

If the velocity field is scaling in space, we expect the entire model to be scaling in space-time – at least up to time scales corresponding to the external (planetary) scale. This is because physically, the wind transports the fields and dimensionally, it connects space and time.
However, it is not enough to simply relabel one of the axes in Fig. 2 as a time rather than as a space axis, and this, even if we make the cascade anisotropic by differentially stratifying it (the right hand side of Fig. 2). The problem is that these pedagogical discrete in scale cascades are fundamentally left-right symmetric, and hence the result of merely re-labelling would be an acausal process. However, it turns out to be fairly straightforward to make causal space-time cascades (Marsan et al., 1996), but the corresponding processes involve causal fractional integrations and cannot be as conveniently illustrated. The data analyses that follows is valid irrespective of the causal nature of the data; we will confine ourselves to pure temporal analyses (treating each pixel as a separate series), we will have the same formula as Eq. (1a) except that \( \lambda = \tau_{\text{ref}}/\tau \) where \( \tau \) is the time scale over which the flux is estimated and \( \tau_{\text{ref}} \) a convenient reference outer time scale, equal to the total length of the time series.

2.2 Estimating the turbulent fluxes

In order to test Eq. (1a), we must therefore use an approach that does not require \textit{a priori} assumptions about the physical nature of the relevant fluxes nor of their scale symmetries (isotropic or otherwise). If atmospheric dynamics are controlled by scale invariant turbulent cascades of various (scale by scale) conserved fluxes \( \varphi \) then in a scaling regime, the fluctuations \( \Delta f(\Delta t) \) in an observable \( f(t) \) (e.g. wind, temperature or radiance, \( t \) is time) over a duration \( \Delta t \) are related to the turbulent fluxes by a relation of the form \( \Delta f(\Delta t) = \varphi \Delta t^H \) (this is a generalization of the Kolmogorov law for velocity fluctuations, for temporal fluctuations in a Lagrangian frame, \( H = 1/2 \), and \( \varphi = \varepsilon^0 \), \( \eta = 1/2 \); for spatial fluctuations, we have \( H = 1/3 \) and \( \varphi = \varepsilon^{\eta} \), \( \eta = 1/3 \)). Without knowing \( \eta \) or \( H \) – nor even the physical nature of the flux – we can use this to estimate the normalized (nondimensional) flux \( \varphi' \) at the smallest resolution of our data:

\[
\varphi' = \varphi / \langle \varphi \rangle = \Delta f / \langle \Delta f \rangle
\]

(2).

Note that if the fluxes are realizations of pure multiplicative cascades then the normalized \( \eta \) powers, \( \varepsilon^\eta / \langle \varepsilon^\eta \rangle \), are also pure multiplicative cascades, so that \( \varphi' = \varepsilon^\eta / \langle \varepsilon^\eta \rangle \) is a normalized cascade quantity.
The fluctuation, $\Delta f(\Delta t)$ can be estimated in various ways; the simplest is to use either absolute first or second differences (respectively $\Delta f(\Delta t) = |f(t + \Delta t) - f(t)|$ or $\Delta f(\Delta t) = |f(t) - (f(t + \Delta r) + f(t - \Delta r))/2|$) with $\Delta t$ the smallest reliable resolution (we assume statistical translational invariance, no $t$ dependence). These “poor man’s wavelets” are usually adequate – when as is typically the case, $0 \leq H \leq 1$ (first) or $0 \leq H \leq 2$ (second) differences can be used. Alternatively other definitions of fluctuations (other wavelets) could be used. In the temporal analyses used here, we used absolute second differences which are the same order differences as the finite difference Laplacian used in the spatial analyses (Stolle et al., 2009, see also below). In section 5.1 we compare the two fluxes. Taking $\Delta t = t_i$ as the smallest available time scale and $\Lambda = t_{ref}/\tau_i$ as the corresponding ratio this yields: $\Delta f_\Lambda = |f(t) - (f(t + \tau_i) + f(x - \tau_i))/2|$). The resulting high resolution normalized flux estimates $\phi_\Lambda^\prime = \Delta f_\Lambda / \langle \Delta f_\Lambda \rangle$ can then be degraded (coarse-grained by averaging) to a lower resolution $\tau > \tau_i$ scale ratio $\lambda = t_{ref}/\tau < \Lambda$.

Since empirical data are nearly always sampled at scales much larger than the dissipation scales, the scaling range based technique described above has wide applicability. In numerical models however, where we have data down to the dissipation range, we find that the approach still works, but that the interpretation can be a little different. To see this, consider the example of the energy flux, $\varepsilon = \partial v^2 / \partial t$. Taking the scalar product of the velocity equation with velocity $v$ and recalling that at the dissipation scale the nonlinear terms are negligible, we obtain:

$$\varepsilon = \partial v^2 / \partial t \approx v \cdot \nabla^2 v$$

where $v$ is the viscosity.

Considering the spatial dependence, standard manipulations ((Landau and Lifschitz, 1959)) give:

$$\varepsilon \approx v \left( \frac{\partial v_i}{\partial x_j} + \frac{\partial v_j}{\partial x_i} \right)^2 = v \left( \frac{\Delta v}{\Delta x} \right)^2$$

so that if $\Delta x$ is in the dissipation range (e.g. the finest resolution of the model) then:
\[ \Delta v = \left( \frac{\varepsilon}{v} \right)^{1/2} \Delta x \]

(5)

The models actually use hyper-viscosities, which have the advantage of confining the dissipation to a small range of scales (about a factor of 3). This means that their dissipation is due to a Laplacian raised to the power \( h \) (typically \( h \) is either 2 or 3), this is discussed in Stolle et al. (2009); they still lead to \( \Delta v \propto \varepsilon^{1/2} \) i.e. to \( \varphi' = \varepsilon^\eta / \langle \varepsilon^\eta \rangle \) with \( \eta = 1/2 \). We therefore see that the spatial scaling in Eq. (5) – which is only valid for \( \Delta x \) in the dissipation range of the model – leads to the dissipation range estimate \( \varphi' = \varepsilon^\eta / \langle \varepsilon^\eta \rangle \) with \( \eta = 1/2 \) whereas the corresponding scaling range (i.e. the Kolmogorov law, essentially dimensional analysis but without the viscosity \( \nu \)), leads to the scaling regime formula \( \Delta v = \varepsilon^{1/3} \Delta x^{1/3} \) i.e. to \( \eta = 1/3 \). In Stolle et al. (2009), we show how to relate the statistics of the dissipation range normalized flux \( \varphi' = \varepsilon^{1/2} / \langle \varepsilon^{1/2} \rangle \) to the statistics of the scaling range normalized flux, \( \varphi' = \varepsilon^{1/3} / \langle \varepsilon^{1/3} \rangle \).

Consider now the temporal fluctuations at the dissipation scale. We see that since \( \varepsilon = \partial v^2 / \partial t \); the relation between \( \Delta v \), \( \Delta t \) and \( \varepsilon \) must be of the form \( \varepsilon = (\Delta v)^2 / \Delta t \) so that we obtain:

\[ \Delta v = \varepsilon^{1/2} \Delta t^{1/2} \]

(6)

Hence once again we have \( \varphi' = \varepsilon^{1/2} / \langle \varepsilon^{1/2} \rangle = \Delta v / \langle \Delta v \rangle \), i.e. \( \eta = 1/2 \); the same as the spatial dissipation range exponent. Interestingly, unlike the spatial case where the dissipation scale \( \Delta v(\Delta x) \) depends on the viscosity as well as \( \varepsilon \), in time \( \Delta v(\Delta t) \) does not explicitly involve \( v \) and the dissipation scale relation Eq. (6) is actually the same as the classical Lagrangian relation which holds for \( \Delta t \) in the scaling regime ((Inoue, 1951); (Landau and Lifschitz, 1959)). Alternatively, with a large scale space-time transformation velocity \( V_{\text{trans}} \), we can take \( \Delta x = V_{\text{trans}} \Delta t \) and use the Kolmogorov formula in space to obtain \( \eta = 1/3 \), see the discussion in section 5.2.

In the data analysis performed below, we did not estimate \( \Delta v(\Delta t) \) at the (model) temporal dissipation scale which is of the order of the integration time step (\( \tau_{\text{int}} \)) (see Table 1), instead the fluctuations were only available at the much longer scaling range \( \Delta t \) of 6 or 12 hours (ERA40, GFS, or GEM respectively, see Table 1). The only caveat is that there is a (nonscaling) diurnal
cycle which mostly affects the temperature at the 1000 mb level; this can lead to poor flux estimates. We avoided this either by taking the analysis inner scale \( \tau_i = 24 \) hours (at the price of losing some temporal resolution) or by detrending the fields by subtracting out the time-averaged value at each time for each longitude and latitude (see sections 3.2).

The fact that the spatial and the temporal analyses both involve the fluxes \( \varphi' = \varepsilon^n / \langle \varepsilon \rangle^n \); \( \eta = 1/2 \), shows that we are analyzing the same physical quantity: in one case degrading it in the spatial, and in the other case the temporal domain. In addition, if space-time is roughly isotropic, then when they are systematically degraded in resolution in either space or in time we expect to find a single linear space-time relation between them. In actual fact, the models do not satisfy all the assumptions of these classical turbulence arguments so that in section 5.1 we check the relations between the fluxes estimated by spatial and fluxes estimated by second order differences. However we find that the ratio of the empirical spatial and temporal exponents \( \eta_s/\eta_t \) is indeed close to unity.

It turns out that if we extend these arguments to passive scalars (which can used as simplified models of temperature and humidity) we obtain similar conclusions about the nature of the spatial versus temporal fluxes estimated from fluctuations (see also Stolle et al. (2009) for the more complex spatial scaling regime). Denoting by \( \rho \) the density of the passive scalar, and \( \chi = \partial \rho^2 / \partial t \) its variance flux, the dissipation range formula analogous to Eq. (3) is \( \chi = \rho \kappa V^2 \rho \) (\( \kappa \) is the molecular diffusivity) leading to \( \Delta \rho \approx (\chi / \kappa)^{1/2} \Delta x \) whereas the corresponding dissipation range temporal formula is \( \Delta \rho \approx \chi^{1/2} \Delta t^{1/2} \) which as for velocity has the same \( \chi^{1/2} \) behaviour (\( \eta = 1/2 \)) in both space and in time, the temporal relation also being valid in the scaling regime (at least for Lagrangian frames). For non-Lagrangian frames we again have use \( \Delta x = V_{trans} \Delta t \) which yields \( \eta = 1/3 \).

**2.3 The outer scales**

A final practical consideration is that in the analyses, the outer scale is not known a priori, but is an empirically estimated parameter. It is therefore convenient to define a reference scale to nondimensionalize the fluxes. With this convention, the basic prediction Eq. (1a) of
multiplicative cascades is that the normalized moments $M_q = \langle \phi^q \rangle / \langle \phi \rangle$ are expected to obey the generic multiscaling relation:

$$M_q = \left( \frac{\lambda}{\lambda_{\text{eff}}} \right)^{K(q)} ; \quad \lambda = \tau_{\text{ref}} / \tau; \quad \lambda_{\text{eff}} = \tau_{\text{ref}} / \tau_{\text{eff}}$$  \hspace{1cm} (7)

where "<,>" indicates statistical (ensemble) averaging $\tau$ is the temporal resolution of the flux, $\tau_{\text{ref}}$ is either take as the duration of the entire data set – or if space-time is quasi-isotropic so that the effective space-time transformation depends only on a “transformation” velocity $V_{\text{trans}}$ is known (see section 5.2) we could use $\tau_{\text{ref}} = L_{\text{ref}} / V_{\text{trans}}$. Finally, $\tau_{\text{eff}}$ is the effective outer temporal scale of the cascade (see section 3.4 and Fig. 7 for estimates). The scale ratio $\lambda / \lambda_{\text{eff}}$ is the overall ratio from the scale where the cascade started to the resolution scale $\tau$.

Before demonstrating this on the reanalysis and model outputs which have lower temporal resolutions and hence very short scaling ranges, we would like to present convincing evidence that the predictions of the cascade models are indeed accurately followed by real meteorological fields. Fig. 3a shows the temporal analysis of hourly infra red imagery from the MTSAT geostationary satellite. The flux was estimated using the finite difference Laplacian at a 30 km spatial resolution and then degraded in time over 1440 images (about 2 months). From about an one hour to about 2 days the evidence of converging straight lines (as predicted in equation 7) is quite convincing, (the corresponding spectrum starts to flatten at around 4 days), and one can see that the lines converge at somewhat larger scales ($\approx 48$ days). This figure demonstrates the challenge we face with even lower temporal resolution data: the lifetime of planetary scale structures may be $\approx 10$ days but the transition is not sharp so that the scaling of the moments of the fluxes is affected on scales as short as 2 days. That the outputs do indeed display cascade structures is shown in Fig. 3b which is a typical spatial analysis; deviations are <1% for scales < 5000 km.

Qualitatively, we can see that all the statistical moments have the same type of behaviour as a function of resolution $\lambda = \tau_{\text{ref}} / \tau$. In particular, starting at the small scale (large $\lambda$, far right) part of the graphs and moving to longer time scales, lower temporal resolutions (smaller $\lambda$, to the left), we notice that the moments initially respect power laws at least approximately emanating...
from a point. As q increases, the estimates are dominated by a smaller set of extreme values, eventually becoming spurious; (“multifractal phase transitions”); the results for the larger q values were therefore not given. This is the prediction of multiplicative cascades in the time domain, Eq. (1a, 7), the point marks the effective outer scale at which the cascade starts (λ = λeff). At larger scales (typically greater than a week or so, the statistics “flatten” out; below we argue that this corresponds to a transition from the weather to the climate regime.

Fig. 3a: Analysis of MTSAT hourly resolution thermal IR imagery over the Pacific. The temporal analysis of the spatial Laplacian (at 30 km resolution) geostationary MSTAT thermal IR imagery over the Pacific for two months. From Pinel et al (2011). The external scale is 48 days. The multifractal parameters are C1 ≈ 0.073, α ≈ 1.8.
3. The model products

3.1 The models

3.1.1 The Canadian Meteorological Centre (CMC) Global Environmental Multiscale (GEM) model

We chose two forecast models and one reanalysis, all recognized as being state-of-the-art: the Canadian Meteorological Centre (CMC) Global Environmental Multiscale (GEM), the NOAA Global Forecast System (GFS) model and the European Centre for Medium range Weather Forecasting’s (ECMWF) reanalysis (ERA40), see Table 1 for a summary of model and data set characteristics. For all three products, we analyzed the three most dynamically significant fields: temperature ($T$), east-west ($u$) wind fields and specific humidity ($h_s$; GEM, ERA 40) and relative humidity ($h_r$, GFS). For each, we investigated the cascade structure as a function of altitude, latitude and (for GEM, GFS) forecast horizon.

Fig. 3b – Trace Moment of the GEM $u$ field at 850mb at time step 00h for $q=0.0$ to 2.8 in increments of 0.2.
The GEM model is on a 0.25°x0.3° horizontal grid with 28 levels and our analysis used a 0.6°x0.6° resolution product (about 66 km resolution, the High-resolution CMC GRIB dataset) at 5 pressure levels (1000 mb, 850 mb, 700 mb, 500 mb, 200 mb). We used 505 realizations taken from September 20th, 2007 to June 2nd, 2008 which are initialized at either 12Z or 00Z and analysed the initial objective analyses and the 48 hour forecasts. In order to partially compensate for the smaller number of realizations and smaller spatial scale, $\tau_i$ was taken as 12 hours. 4DVAR 6-hourly assimilation is now used (CMC Global Data Assimilation System (DAS)). The model can be adapted, see Côté et al. (1998a, 1998b) for more details.
<table>
<thead>
<tr>
<th></th>
<th>GEM</th>
<th>GFS</th>
<th>ERA 40</th>
</tr>
</thead>
<tbody>
<tr>
<td>Time step (minutes) $\tau_{st}$</td>
<td>22.5</td>
<td>7.5</td>
<td>30</td>
</tr>
<tr>
<td>Number of realizations in sample</td>
<td>505</td>
<td>1340</td>
<td>4384</td>
</tr>
<tr>
<td>Analysis Time interval (hours) $\tau_i$</td>
<td>12</td>
<td>6, 24</td>
<td>6, 24</td>
</tr>
<tr>
<td>Assimilation cycle (hours)</td>
<td>12</td>
<td>6</td>
<td>6</td>
</tr>
<tr>
<td>External time scale of the data set $\tau_{ref}$</td>
<td>223 days</td>
<td>300 days</td>
<td>1000 days</td>
</tr>
<tr>
<td>Model Spatial resolution, $L_g$</td>
<td>$0.25^\circ \times 0.3^\circ$</td>
<td>$(0.47)^\circ \times (0.47)^\circ$</td>
<td>$(1.125)^\circ \times (1.125)^\circ$</td>
</tr>
<tr>
<td>Grid Speed (m/s), $V_{st} = L_g / \tau_{st}$</td>
<td>20.5</td>
<td>116,</td>
<td>69,</td>
</tr>
<tr>
<td>Number of vertical levels</td>
<td>28</td>
<td>64</td>
<td>60</td>
</tr>
<tr>
<td>Spatial resolution of analysis $L_i$</td>
<td>$0.6^\circ \times 0.6^\circ$</td>
<td>$1^\circ \times 1^\circ$</td>
<td>$1^\circ \times 1^\circ$</td>
</tr>
<tr>
<td>Analysis speed (m/s) $V_i = L_i / \tau_i$</td>
<td>0.77</td>
<td>1.3</td>
<td>1.3</td>
</tr>
<tr>
<td>Typical space-time transformation speed (700 mb) $V_{trans}$</td>
<td>500 km/day = 5.8 m/s</td>
<td>450 km/day = 5.2 m/s</td>
<td>1000 km/day = 11.6 m/s</td>
</tr>
</tbody>
</table>

Table 1: Comparison of various model parameters. The time step is the model integration time step, the “grid speed” is the ratio of finest model resolution to the time step. The time interval, $T_i$, is the smallest $\Delta t$ used in the analyses, the “analysis speed” is the ratio of the analysis spatial and temporal resolutions. For GFS, ERA40 the analyses were at 6 hours. The “typical transformation speed”, $V_{trans}$, is the mean speed obtained from the space-time analyses of section 5.2 for ERA40 and GFS. The value for GEM is for 700 mb used only $h_i$ and may not be representative of the other GEM fields.
3.1.2 The NOAA Global Forecast System (GFS) model

Like the CMC GEM model, the GFS is a global NWP model, which we analyzed at its analysis ($t = 0$) and 48h forecast ($t = 48$ hours). The GFS model uses T254 Spectral and 768x384 Gaussian grids on 64 vertical levels. The data is obtained at a 1°x1° resolution at 5 pressure levels (1000 mb, 850 mb, 700 mb, 500 mb, 200 mb) every 6 hours; each initialization starts at 00Z, 06Z, 12Z, or 18Z. The data were taken from August 1, 2007 to June 30, 2008 (with the exception of 700 mb u for which the first 61 days were corrupted). A total of 1340 realizations were analyzed. When needed, due to the diurnal variation, fields were averaged over 4 consecutive periods so that $T_i = \text{either 6 or 24 hours}$. The assimilation system used is 3DVAR (Okamoto and Derber, 2005) with an assimilation cycle of 6 hours. For more information, see Sela (1982), Sela (1988), and NCEP office note 442 (2003).

3.1.3 The European Centre for Medium range Weather Forecasting’s (ECMWF) reanalysis (ERA40) product

A reanalysis combines a model forecast with observational data using sophisticated data assimilation techniques, here a 6-hourly 3DVAR assimilation cycle was used. The product was made as uniform as possible over the 45 year period from September 1957 to August 2002 despite the changing observation network. The reanalyses were started every 6 hours for every day: 00Z, 06Z, 12Z, 18Z. Reanalyses are clearly dependent not only on the data, but also on the model being used for the assimilation, they are hybrid products.

This dataset was chosen because the data quality is relatively uniform over various multi-year periods, reanalyses are important tools in meteorological and climatology research, and this data is easily compared to the other datasets analysed in this paper (in particular the analysis of each of the models). We analyzed the most recent three years of the reanalysis: September 1999-August 2002 with a total of 4380 time steps analyzed at the following levels: 1000 mb, 850 mb, 700 mb, 500 mb, 200 mb. See Uppala et al. (2005) for more details.

In the reanalysis, the dynamic variables (that is, wind/vorticity/divergence, temperature, humidity) are calculated on the T159 (triangular truncation up to wavenumber 159) spherical harmonic grid, while the other variables are on an N80 Reduced Gaussian grid (80 latitudinal bands in each hemisphere with number of points varying from 18 near the poles to 320 around
the equator) on 60 vertical levels. The data were then interpolated onto a 1°x1° constant pressure level grid at 5 pressure levels indicated above. Only the dynamic variables temperature ($T$), east-west ($u$) wind fields and specific humidity ($h_s$) are analysed. For the cascade analyses, data between 30°N-30°S and 45°N-45°S were used. Although this interpolation causes the finest scales to be a little bit too smooth, problems such as aliasing are avoided.

![Figure 3 c: Schematic diagram of the space time scales used here. To make the time axis dimensionally commensurate with the spatial axis, it has been multiplied by the mean transformation velocity estimated in section 5.2 ($V_{trans}$, is denoted as $V$ in the figure). The raw model resolutions are the time step $V\tau_{st}$ and the inner scale $L_i$ (to simplify, we ignore the small difference between the inner analysis scale $L_i$ and raw grid scale $L_{st}$ see Table 1). $L_{ref}$ and $\tau_{ref}$ are the reference scales; $L_{ref} = 20,000 \text{ km}$ and $\tau_{ref}$ varies from 223 days to 1000 days depending on the dataset. The vertical rectangles show the space-time averaging regions used to define the spatial flux averages using “snapshots” whereas the horizontal rectangles show the space-time averaging regions used in the temporal analysis. The scales respect the inequalities $l \geq L \geq V\tau_{st}$ (spatial analyses) and $V\tau \geq V\tau_{st} \geq L_i$ (temporal analyses) as long as the $V_i < V < V_{st}$ (with $V_i = L_i/\tau_i$, $V_{st} = L_i/\tau_{st}$, see Table 1).]
3.2 Discussion

In attempting to determine the cascade structure, we must consider the space-time resolution of the data - how large a region (in space and time) does each analysis element represent? Fig. 3c shows the various scales on a schematic space-time diagramme; in order to make the space and time axes commensurate, we have multiplied the times by a large scale mean transformation velocity $V_{\text{trans}}$ (estimated in section 5.2 below); this is adequate if the (horizontal) spatial and temporal exponents are equal (space-time is only “trivially” anisotropic) so that the space-time relation is linear. This linearity is justified below on empirical grounds, in (Lovejoy et al., 2008)) it is given some theoretical justification (see also section 5). More generally, space-time can be scaling but anisotropic so that if the transformation is nonlinear (a power law), we need a generalized velocity; here we discuss the simplest linear case apparently relevant to the models.

Our analysis starts with the original raw model output which is at space-time resolution $L_iX(V_{\text{trans}} \tau_{\text{st}})$. We then estimate the turbulent fluxes either spatially by (absolute) finite difference Laplacian or temporally by (absolute) second order time differences. In the spatial case, the flux is a dissipation scale estimate and will have space-time resolution $L_iX(V_{\text{trans}} \tau_i)$ with (see Table 1), $L_i>(V_{\text{trans}} \tau_i)$, in the temporal case, since the available data are at either 6 or 12 hour intervals ($\tau_i \gg \tau_{\text{st}}$, see Table 1), and $\tau_i$ is in the scaling (not dissipation) range, we estimate temporal scaling fluxes (see section 2) at resolution $L_iX(V_{\text{trans}} \tau_i)$ with $(V_{\text{trans}} \tau_i)>L_i$ (see Fig. 3c). In both cases, the flux is only estimated by finite differences which involve neighbouring pixels, so that the fine resolution fluxes may suffer from “finite size effects” which become less and less important as their resolutions are degraded by subsequent spatial or temporal averaging. This presumably explains at least some of the deviations from scaling at the smallest space-time scales. In addition, we have mentioned that the assumption that $\tau_i$ is in the scaling range is not perfect since the scaling will be perturbed by the assimilation cycle (6 hours) as well as to varying degrees by the diurnal cycle. The latter effects can be minimized by first averaging four consecutive 6 hour or two consecutive 12 hour fields before the fluxes are estimated by differencing. However this has the unfortunate effect of losing a factor of 4 (or 2) in time scales which is quite a lot given that the temporal cascade begins to break down at scales of 1.5-5 days ($\tau_{\text{dev}}$). In order to
eliminate the effects of the diurnal cycle, we thus subtracted off the mean of the dataset at each
time of the day.

Now consider the effect of degrading the resolution of the fine scale flux estimates. The
flux at a space-time resolution \((l, \tau)\) is a space-time flux average over the corresponding
rectangles shown in the schematic. If \(l \gg V_{\text{trans}} \tau\), then the spatial resolution is dominant and the
overall scale ratio is \(\lambda = L_{\text{ref}}/l\). If on the contrary, \(V_{\text{trans}} \tau >> l\), then the temporal averaging
dominates and we obtain \(\lambda = L_{\text{ref}}/(V_{\text{trans}} \tau)\) (if the two are comparable, either can be used). From
this, it is clear that the optimum is to choose the inner analysis resolutions to be comparable so
that degrading purely in space or degrading purely in time gives us information about
respectively the spatial and temporal structures. If – as suggested in the diagram – the temporal
analysis resolution \(\tau\) is too large, then at first \((l \approx L_i)\) degrading in space will have little effect, it
is only when \(l>>V_{\text{trans}} \tau\) that the effect of spatial averaging will be felt.

A final comment about the temporal analysis is to signal the existence of a few missing
fields for GEM. For the spatial scaling behaviour, this omission had only a negligible effect
because each time step could be treated as a snapshot over which the spatial analysis technique
could be applied independently. When determining the temporal behaviour, the analysis
technique sees these holes. Because there were only a few missing time steps, it was decided
that the missing data would be filled in by interpolation.

4. The temporal analyses

4.1 Cascade analysis: forecast horizons, latitude and altitude dependence

4.1.1 \(t=0\)

Figs. 4 a-f, 5 a-f, and 6 a-f show the behaviour of the fluxes estimated using the absolute
second order difference at \(\tau = 6\) hours (ERA40 and GFS) and 12 hours (GEM). Each figure
shows the horizontal wind, temperature and humidity moments (graphs top to bottom, a-b, c-d, e-
f respectively) and at two pressure levels (1000 \(mb\), 700 \(mb\), left and right respectively) and for
the statistical moments of order \(q = 0.1, 0.2, \ldots, 2.9\) (for \(q > \approx 0.5\), the slopes monotonically
increase). We only consider up to \(q=2.9\) because the behaviour of larger moments is dominated
by increasingly rare events and so the moments become less reliable. Note that the fields at 1000 mb are influenced by the generally higher quality data near the ground, although this is somewhat offset by the interpolation and extrapolation techniques needed when the topography is significant (i.e. 1000 mb often corresponds to points below ground). In addition, the diurnal cycle which can lead to poor flux estimates (see above) it is much stronger at 1000 mb (especially for the temperature). As stated earlier, we infer that there is a temporal cascade at smaller timescales, of which we only see the longest timescale in our analysis, because of the limited temporal resolution because of the spatial cascade behaviour demonstrated in the spatial trace moments as shown in (Stolle et al, 2009).

The outer scale estimates for the various fields as functions of pressure level are given in Fig. 7; we see that they are typically of the order of 10-20 days which is in fact of the order of the lifetime of an atmospheric structure (see the discussion below and Fig. 7 for graphical results). We can also define a smaller scale at which the trace moments begin to deviate strongly from cascade behaviour. To quantify the deviations from Eq. (7) and to make this distinction more objective, we can define the deviations Δ from pure cascade behaviour using the formula:

\[
\Delta = \left\{ \log_{10}(M_q) - K(q) \log_{10}(\lambda / \lambda_{eff}) \right\}
\]

where the overbar represents averaging over all the moments \( q \leq 2 \) and over the scale ratios \( \lambda > \lambda_{dev} \). If the range of scales used to estimate \( \Delta \) is increased from the smallest scale (largest \( \lambda \)) to larger scales (smaller \( \lambda \)), we find that \( \Delta \) starts off being relatively stable (i.e. the deviations are of roughly constant size) but at some point \( \approx \tau_{dev} \) it rapidly increases when the flux moments begin to deviate systematically from the scaling predictions of Eq. (8). \( \tau_{dev} \) can then be somewhat objectively defined as the scale at which \( \Delta \) is 1.5 times the minimum value that it had in the smaller scale (cascade) region. Table 2 shows the variation of \( \tau_{dev} \) for the various fields; we see that it is in the range \( \approx 2-5 \) days, virtually independent of altitude. Recalling that \( \tau_{eff} \) is the effective outer cascade time scale, in between \( \tau_{eff} \) and \( \tau_{dev} \), the region is taken to be a transition region between the climate and meteorology (Fig. 4-6).

Considering the \( \tau_{eff} \) estimates for the various fields, there are few obvious trends (see Fig. 7.) The main ones are: a) the temperature tends to have a larger \( \tau_{eff} \) than the other fields, b)
ERA40 had generally smaller $\tau_{\text{eff}}$ than the other models, c) most of the estimates are on the order of 10 days, as expected based on the solar energy flux input and the horizontal scaling (see section 4.2), d) $\tau_{\text{eff}}$ for the lowest two altitudes of temperature for ERA40 and the highest three altitudes of temperature for GEM are much larger than the others.

Figure 4 a-f: Temporal flux estimated from the second order differences in time, temporal trace moments ($q = 0.0$ to 2.9 in increments of 0.1) for ERA40 analysis between 30°N and 30°S. a) $u$ 1000 mb, $\tau_{\text{eff}}$ 5.8 days, $\tau_{\text{dev}}$ 1.5 days; b)
$u$ 700 mb, $\tau_{\text{eff}}$ 5.4 days, $\tau_{\text{dev}}$ 1.5 days; c) $T$ 1000 mb, $\tau_{\text{eff}}$ 39.8 days, $\tau_{\text{dev}}$ 1.5 days; d) $T$ 700 mb, $\tau_{\text{eff}}$ 5.0 days, 1.5 days; e) $h_s$ 1000 mb, $\tau_{\text{eff}}$ 47.9 days, $\tau_{\text{dev}}$ 2.5 days; f) $h_s$ 700 mb, $\tau_{\text{eff}}$ 9.2 days, $\tau_{\text{dev}}$ 1.5 days. $\tau_{\text{ref}}$ = 1000 days.

Figure 5 a-f: Temporal flux estimated from absolute second order differences in time; trace moments ($q$ = 0.0 to 2.9 in increments of 0.1) for GEM analysis between 30°N and 30°S. a) $u$ 1000 mb, $\tau_{\text{eff}}$ 24.2 days, $\tau_{\text{dev}}$ 2.0 days; b) $u$ 700 mb, $\tau_{\text{eff}}$ 13.6, days, $\tau_{\text{dev}}$ 2.5 days; c) $T$ 1000 mb, $\tau_{\text{eff}}$ 12.4 days, $\tau_{\text{dev}}$ 1.5 days; d) $T$ 700 mb, $\tau_{\text{eff}}$ 30.4 days, $\tau_{\text{dev}}$ 2.5 days; e) $h_s$ 1000 mb, $\tau_{\text{eff}}$ 10.3 days, $\tau_{\text{dev}}$ 1.5 days; f) $h_s$ 700 mb, $\tau_{\text{eff}}$ 13.0 days, $\tau_{\text{dev}}$ 2.0 days. $\tau_{\text{ref}}$ = 223 days.
Figure 6 a-f: Temporal flux estimated from the absolute second order differences in time; trace moments for GFS analysis between 30°N and 30°S a) $u$ 1000 mbar $\tau_{\text{eff}}$ - 5.8 days, $\tau_{\text{dev}}$ - 1.25 days; b) $u$ 700 mbar $\tau_{\text{eff}}$ - 23.4 days, $\tau_{\text{dev}}$ - 1.25 days; c) $T$ 1000 mbar $\tau_{\text{eff}}$ - 141.2 days, $\tau_{\text{dev}}$ - 1.25 days; d) $T$ 700 mbar $\tau_{\text{eff}}$ - 4.5 days, $\tau_{\text{eff}}$ -
1.25 days; e) $h_r$ 1000 mbar, $\tau_{\text{eff}}$ - 35.5 days, $\tau_{\text{dev}}$ - 1.25 days; f) $h_r$ 700 mbar; $\tau_{\text{eff}}$ - 13.5 days, $\tau_{\text{dev}}$ - 1.25 days. $\tau_{\text{ref}}$ = 300 days.
Figure 7: Estimates of $\tau_{\text{eff}}$ for GFS (green lines, square points), GEM (blue lines circular points), ERA40 (red lines, diamond points) for $t=00h$. The lines indicate which points belong to the same dataset/field $h$, $T$, $u$ (solid lines), $h$, $T$, $u$ (short dashed lines), $T$ (solid lines), $u$ (long dashed lines).

<table>
<thead>
<tr>
<th>$\tau_{\text{dev}}$ (days)</th>
<th>$h$</th>
<th>$T$</th>
<th>$u$</th>
</tr>
</thead>
<tbody>
<tr>
<td>GFS</td>
<td>1.0-1.5</td>
<td>1.25</td>
<td>1.25</td>
</tr>
<tr>
<td>GEM</td>
<td>1.5-2.0</td>
<td>2.5-3.0</td>
<td>2.0-2.5</td>
</tr>
<tr>
<td>ERA40</td>
<td>1.5-2.5</td>
<td>1.5</td>
<td>1.5</td>
</tr>
</tbody>
</table>

Table 2: Estimate of $\tau_{\text{dev}}$ for GFS, GEM, ERA40 (roughly the same for all altitudes for each field).
4.1.2 Differences in Temporal Cascade Behaviour for 00 h and 48 h Fields

Fig. 8 a-f shows the effect of model integration, we compare analysis ($t = 0$) and 48 hour forecast for the two forecast models. We see that there is very little difference between them. Apparently, the model integration has little effect on the cascades. Since the 48 hour forecast is less affected by the initial data than the $t = 0$ fields, this gives some support to the idea that the long term “climate” of the model also involves cascades possibly with very similar statistics. Before continuing, we should mention that the regional dependence was also investigated by comparing the cascade structures for data between ±30° and ±45°. The differences in the cascade parameters were mostly small. For more systematic comparisons of this type (but on ECMWF interim reanalysis products), where it was found that the main effect was on the outer time scale, not the exponents, and that this variation followed fairly closely that predicted using the mean energy flux at the given latitude to predict the corresponding eddy turn-over time; see Lovejoy and Schertzer (2011b) for more details.
Figure 8 a-f: Temporal flux, temporal trace moments (q=0.0 to 2.0 in steps of 0.1) for GEM ($T_{ref} = 223$ days, $T_{ref} = 300$ days) and GFS between 30°N and 30°S at $t = 0$ (cyan) and $t = 48$ h (red). a) GEM U 1000 mb; b) GFS U 1000 mb; c) GEM T 700 mb; d) GFS T 700 mb; e) GEM $h_s$ 700 mb; f) GFS $h_s$ 700 mb.
4.3 Levy collapse and universality

We have primarily been interested in establishing the fundamental prediction of multiplicative cascade models, Eq. (1a, 7). However, we also argued that there are basic physical, mathematical reasons (essentially the existence of a kind of multiplicative central limit theorem) that make it plausible that the model fields fall into special universality classes in which the basic scale invariant exponent \( K(q) \) is given by Eq. (1b) characterized by just two parameters \( C_1, \alpha \). In the analysis of spatial cascades (Stolle et al., 2009), it was directly shown that universality works well – at least up to a critical moment \( q_c \) beyond which there is a “multifractal phase transition” where \( K \) becomes asymptotically linear (a sample size dependent effect corresponding to the domination of the statistics by the largest flux values present (Schertzer et al., 1993)). If the flux follows Eq. (1b), it implies that the generator of the cascade (log flux) is a Levy variable, index \( \alpha \). In that case, one can attempt to “collapse” the moments \( M_q \) to a unique curve by dividing \( \log M_q \) by the theoretical \( K(q) \) for (say) \( C_1 = 1 \), i.e. by dividing it by \( (q^\alpha-q)/(\alpha-1) \); if \( M_q \) does indeed follow Eqs. (1a) and (1b) with parameters \( C_1, \alpha \), then we obtain \(( (\alpha-1) \log M_q ) / (q^\alpha-q) = C_1 \log \lambda \) which is independent of the moment \( q \) so that all the curves with different \( q \) values “collapse” onto a single curve whose slope is given by \( C_1 \). The interest of such plots goes beyond just testing for the universal cascade behaviour: on a single plot we can independently evaluate both the scaling (the straightness of the collapsed lines) as well as the log-Levy nature of the generator – by the thinness of the collection of lines i.e. how well at a given scale the different moments collapse, how well they follow the form \( (q^\alpha-q) \).

Before continuing, we should mention that the stochastic cascade model that generates the “weather” regime up to \( \tau_w \approx 10 \) days (see Fig. 1) can be extended to much longer time scales, i.e. to scales where the space-time weather cascade process has been reduced to an essentially purely temporal process. As argued in (Lovejoy and Schertzer, 2010), the collapse can be considered as a “dimensional” transition” the model energy flux \( \varepsilon \) can be written as a product of high frequency space-time weather regime flux \( \varepsilon_w(r,t) \) multiplied by a low frequency weather process \( \varepsilon_{lw}(t) \) which is only a function of the time scales longer than \( \tau_w \). Theory shows that \( \varepsilon_{lw}(t) \) has a asymptotically at low enough scales – a singular autocorrelation function \( \langle \varepsilon_{lw}(r)\varepsilon_{lw}(t-\Delta t) \rangle \approx \Delta t^{-1} \); the corresponding spectrum has a low frequency divergence. In
practice this means that the spectrum $E(\omega)$ for frequencies lower than $\tau_w^{-1}$ will have spectral exponent $\beta$ which depends only the outer scale (and weakly on $\alpha$) with realistic values in the range $\beta \approx 0.2 - 0.4$ (with $E(\omega) \approx \omega^\beta$; this reproduces the observed spectral plateau, see (Lovejoy and Schertzer, 2011), Fig. 1).

Further consequences of the dimensional transition are that although the “bare” low frequency weather process $\mathcal{E}_{lw}(t)$ process (i.e. constructed down to a given scale and stopped) has roughly log-Levy distributions, that the empirically measured “dressed” process i.e. continued down to much smaller scales and then averaged as in our data analysis, will eventually have quasi Gaussian statistics, therefore the Levy collapse would not continue to time scales much longer than $\tau_w$. However, as mentioned in the introduction, this purely atmospheric model doesn’t take into account interactions with the oceans which apparently have similar turbulence and transitions, only with $\tau_o$ at about $\approx 1$ year (as mentioned in the introduction $\mathcal{E}_o \approx 10^{-8} \text{m}^2/\text{s}^3$ which is is much lower than $\mathcal{E}_w \approx 10^{-3} \text{m}^2/\text{s}^3$). The interaction between the ocean and the atmosphere may therefore provide the extra correlations needed to allow the log-Levy distributions to continue to longer time scales of the order of one year.

With this theoretical motivation, let us turn to the analyses. In Fig. 9 we see that the collapse for the different models and different fields is fairly convincing and this out to scales well into the “climatological” regime; typically the spread is 2-10 percent of the mean, comparable to that in space, see Stolle et al. (2009), which we illustrate in Fig. 10. Interestingly, the temperature has the largest variability (highest curves), and the humidity the lowest (except for the forecast models (GEM, GFS) at 1000 mb where the most variable field is the wind). Note that the largest deviations are at the very smallest scales (largest $\lambda$) due to the “finite size” effects mentioned earlier, and at the very largest scales (smallest $\lambda$) where the statistics are poor.
Figure 9 a-f: Levy collapse \( \log_{10} M' = (\alpha - 1) \log_{10} M / (q^\alpha - q) \) of temporal flux, of the temporal trace moments at \( t=0 \) \( (q=0.1 \) to \( 2.0 \) excluding \( q=1.0 \)) for ERA40 (\( \tau_{ref}=1 \) year), GEM (\( T_{ref}=223 \) days), and GFS (\( \tau_{ref}=300 \) days, except \( u \) 700mb, \( \tau_{ref}=250 \) days) analysis between 30°N and 30°S of \( u \) (red), \( T \) (green), \( h_r \) or \( h_s \) (blue). a) ERA40 1000 mb, b) ERA40 700 mb, c) GEM 1000 mb, d) GEM 700 mb, e) GFS 1000 mb, f) GFS 700 mb. The collapse is visible by the bunching, overlap of the curves.
Figure 10 a-f: The relative root mean square spread of Levy collapse of fig. 9 (the relative standard deviation of $\frac{\text{Log}_{10} M_q}{(\alpha - 1) \text{Log}_{10} M_q / (q^\alpha - q)}$) of temporal flux for $q=0.1$ to $2.0$ excluding $q=1.0$) for ERA40 ($\tau_{\text{ref}}=1\text{year}$), GEM ($T_{\text{ref}}=223\text{ days}$), and GFS ($T_{\text{ref}}=300\text{ days}$, except $u$ 700mb, $\tau_{\text{ref}}=250\text{ days}$) analysis between $30^\circ\text{N}$ and $30^\circ\text{S}$ of $u$ (red), $T$ (green), $h_r$ or $h_s$ (blue). a) ERA40 1000 mb, b) ERA40 700 mb, c) GEM 1000 mb, d) GEM 700 mb, e) GFS 1000 mb, f) GFS 700 mb. The collapse is visible by the bunching, overlap of the curves.
5. Comparing the spatial and temporal cascades

5.1 Spatial analyses using spatial and temporal flux estimates

We noted in section 2 that turbulent fluxes could be estimated by using either (horizontal) spatial Laplacians (as in the analysis Stolle et al. (2009)) or by second order differences in time (the discrete Laplacian analogue). However, the fluxes estimated in space and in time are not necessarily the same notably because the spatial estimates are at the (model) dissipation scale whereas the temporal estimates are in the scaling regime (with the caveats about the diurnal and assimilation cycles mentioned earlier). In section 2, we argued that if the normalized spatial and temporal fluxes are different powers of the underlying flux (say $\phi_x = \varepsilon^{\eta_x}/<\varepsilon^{\eta_x}>$, $\phi_t = \varepsilon^{\eta_t}/<\varepsilon^{\eta_t}>$ in space and time respectively), then if $\varepsilon$ are also universal multifractals with index $\alpha$, then the ratio $\log<\phi_x^q>/\log<\phi_t^q> = (\eta_x/\eta_t)^\alpha$; i.e. they are in a constant ratio. In Fig. 11 we rescaled the $\phi_t$ moments by taking the optimum powers of the moments of the temporal fluxes so that they optimally overlap the $\phi_x$ moments. We see that there is generally good agreement between the rescaled temporal fluxes with the spatial fluxes. The fluxes estimated in the time domain assume that the latter is in a scaling regime at the 6 hour (or 12 for GEM) time scale used for estimating the second order differences. However, the flux estimate may be poor due to the nonscaling perturbations such as the assimilation or the diurnal cycle. The fact that generally the degree of overlap is very high gives us confidence that the flux estimates are quite trustworthy. Indeed, we see that the main case of poor overlap is precisely the 1000 mb $T$ fluxes which are particularly affected by the diurnal cycle. In Table 3, we see that the ratio is close to unity (with the exception of $u$ at 1000 mb and GFS $h_v$) with GEM having a little smaller values (~0.7), GFS averaging to a little above 1 between the different fields (though the ratio for the 1000 mb $h_v$ field is about 2), and the average of the values for ERA40 being slightly larger than 1. Since $\alpha \approx 1.8$ for all the fields, for GEM this corresponds to $\eta_x/\eta_t \approx 0.85$. This is close enough to the value 1 that it is possible that all the results are compatible with 1 with the small differences due to some other cause. Recall that $\eta_x = \eta_t = 1/2$ (hence $\eta_x/\eta_t = 1$) in the case of the energy flux estimated at the dissipation scale in space and in the Lagrangian
scaling regime in time. Note that we use the 6 h fluxes as opposed to the 1 day fluxes for GFS and ERA40 because it allows for a larger range of analysis; this procedure is validated in Fig. 11.

Figure 11 a-f: Overlapping Mean Levy Collapse \((\log_{10} M' = (\alpha - 1) \log_{10} M_y / (q^\alpha - q))\) of spatial Laplacian flux (red) and temporal second order difference flux (blue) spatial trace moments at \(t = 0 \ h \ (q=0.1 \text{ to } 2.0 \text{ excluding})\)
\( q = 1.0 \) for ERA40 \((T_{_{\text{ref}}} = 1\, \text{year})\), GEM \((T_{_{\text{ref}}} = 223\, \text{days})\), and GFS \((T_{_{\text{ref}}} = 300\, \text{days}, \text{except } u\, 700\, \text{mb}; T_{_{\text{ref}}} = 250\, \text{days})\) analysis between 30°N and 30°S. The time trace moment is multiplied by the factor \((\eta_x/\eta_t)^\alpha\) given in table 3. a) ERA40 \( u\, 1000\, \text{mb} \); b) ERA40 \( T\, 700\, \text{mb} \); c) GEM \( u\, 1000\, \text{mb} \); d) GEM \( h\, 700\); e) GFS \( u\, 1000\, \text{mb} \); f) GFS \( h\, 700\, \text{m} \). The slopes are \( C_1 = K(q)/(\eta_x - q) \) of the spatial trace moments are the same as in Stolle et al (2009).

| Dataset          | \( \log \langle \varphi_x^q \rangle / \log \langle \varphi_t^q \rangle \) |
|------------------|------------------|------------------|------------------|
| \( u\, 1000\, \text{mb} \) | 1.6              | 1.1              | 1.0              |
| \( T\, 700\, \text{mb} \) | 1.4              | 1.4              | 1.0              |
| \( u\, 1000\, \text{mb} \) | 0.7              | 0.6              | 0.8              |
| \( h\, 700\, \text{mb} \) | 0.7              | 0.7              | 1.0              |
| \( u\, 1000\, \text{mb} \) | 1.6              | 1.6              | 2.0              |
| \( h\, 700\, \text{mb} \) | 1.0              | 1.0              | 1.0              |

Table 3: ratio, \( \log \langle \varphi_x^q \rangle / \log \langle \varphi_t^q \rangle \) connecting space and time flux according to theory, this equals \((\eta_x/\eta_t)^\alpha\).

### 5.2 Space-time diagrams

In order to systematically explore the space-time relations, we can compare the statistics of the flux moments as functions of time scale and as functions of space scale; it is important that we degrade the same flux in both space and in time. The previous subsection shows that both the spatial finite difference Laplacian or the second order temporal difference give essentially equivalent results (related by a factor close to unity to go from one to the other due to the different fluxes that are estimated). We therefore chose to use the absolute second order temporal differences and estimate the space-time ("Stommel") diagrammes for the various fields using the following method. The basic idea is that if for a spatial scale \( L \) and a temporal scale \( \tau \) and for all \( q \), we have

\[
\langle \varepsilon_x^q / \tau \rangle = \langle \varepsilon_t^q / L \rangle \tag{9}
\]
then the scales \( L \) and \( \tau \) are implicitly equivalent in the sense that their statistics are the same; for example, if Eq. (9) holds for all \( q \), then at the level of random variables it implies: \( \varepsilon_{\tau_{\text{ref}}/\tau} \overset{d}{=} \varepsilon_{L_{\text{ref}}/L} \), where \( \overset{d}{=} \) means equality of probability distributions.

If the horizontal scaling of the flux is the same as the temporal scaling the same flux, then to determine the relation between \( \tau \) and \( L \), it suffices to superpose the spatial and temporal \( M_q \) and adjust them left-right on the log-log plot to find the optimum space-time ratio (i.e. velocity) connecting them. If the scaling between space and time is anisotropic, i.e. \( \tau_{\text{ref}} / \tau = (L_{\text{ref}} / L)^{H_t} \) with \( H_t \neq 1 \) then the space-time relation is nonlinear; see (Radkevitch et al., 2008) for a discussion and space-time analysis of lidar backscatter data). Theoretically, there are two classical possibilities for \( H_t \): if we ignore advection, then for the horizontal wind field dominated by the turbulent energy flux, \( \varepsilon \), in space the (Kolmogorov) scaling is \( \Delta v \approx \varepsilon^{1/3} \Delta x^{1/3} \) whereas the corresponding (Lagrangian frame) formula in time is \( \Delta v \approx \varepsilon^{1/2} \Delta t^{1/2} \). This leads to \( H_t = (1/2)/(1/3) = 2/3 \). However, even if there is no imposed large scale wind, the (random) largest planetary scale eddies will advect all the smaller scale eddies, and this up to the outer scale of the scaling regime (here, planetary in extent). This advection will lead to \( H_t = 1 \) with the overall advection velocity being the large scale turbulent velocity mentioned earlier: \( V \approx \varepsilon^{1/3} L_{\text{eff}}^{1/3} \). This argument is an extension of one first invoked by Tennekes (1975); see the discussion in Lovejoy et al. (2008). Fig. 12a shows a typical result when the technique is applied to the hourly resolution infra radiances (see Fig. 1). We see that the space-time relation is roughly linear up to 2 days in the north-south direction and nearly 10 days in the east-west direction.
Fig. 12 a: The horizontal space-time diagram constructed by comparing the temporal MTSAT cascades (Fig. 1) and north-south and east-west analyses of the corresponding spatial moments. From Lovejoy and Schertzer 2010, (we thank J. Pinel). Over the scales where the data are linear, spatial and temporal statistics (for the q=2 moment) are related by a constant velocity of \( \approx 900 \text{ km/day} \).
Fig. 12b: Space-time plots for the ECMWF interim reanalysis for 700 mb fields analysed for 2006 using the $q = 2$ moments and using $\lambda = T_{\text{ref}}/\Delta t$ and $\lambda = L_{\text{ref}}/\Delta x$ for time and space respectively (east-west and time). Yellow = $h_i$, green = $T$, cyan = $u$, blue = $v$, purple = $w$, red = $z$. The black reference line has a slope 1; it corresponds to a speed of $\approx 225 \ km/day$; the spread in the lines indicates a variation over a factor of about 1.6 in speed. Reproduced from Lovejoy and Schertzer (2011).

Once again, the difficulty with calculating space-time diagrams for the outputs is that the temporal resolution is barely adequate. Figs. 12b-f show the results for 1000 mb and 700 mb for all of the datasets. In addition, the single smallest scale point, corresponding to 6 hours (or 12 hours for GEM) in time (for which the fluxes are not necessarily well estimated due to “finite size effects”), is also a little too low. The space-time relation is not as linear as in Fig. 12a or as was found in the ECMWF interim reanalyses in Lovejoy and Schertzer (2011b), reproduced for comparison in fig. 12b. We seem to have deviations at small $\Delta t$ due to finite size effects, a narrow scaling regime ($\approx 12$hrs- 2 days), and then the beginning of the low frequency weather spectral plateau. The reference lines in Fig. 13 are plausible given “finite size effects” affecting the smallest time steps (6 and 12 hours) and the small range of the temporal scaling.

A transition to the low frequency weather regime is easily seen on the space-time diagrams by the flattening of the curves. Overall, GEM shows, a clearer transition at 1000 mb than at 700 mb despite the noisier space-time diagram. In addition, the corresponding advection velocities are around 500-1000 km/day for GFS and ERA40 which are close to that found in the analysis of geostationary infra red satellite imagery (Fig. 12a) and close to those estimated from the solar energy injection as discussed in section 4.2. GEM, on the other hand, has a characteristic velocity of 250 km/day, which is somewhat smaller though perhaps not inconsistent. With the exception of the GFS 1000 mb and GEM 700 mb fields, the wind, humidity and temperature have very similar space-time transformation velocities ($V_{\text{trans}}$).

As seen in the space-time diagrams (ignoring the exceptions in GFS 1000 mb and GEM 700 mb), they start to level off at around 5000-10000 km corresponding to $\approx 10$ days. This is a little larger than the end of the accurate temporal scaling regime, $\tau_{\text{dev}}$ (1.5-6.0 days) – the scaling value where the trace moments were found to deviate from cascade behaviour. The difference between $\tau_{\text{dev}}$ and the value at which the space-time diagrams flatten differ because the both the spatial and temporal trace moments tend to curve at the largest scales, so that the space-time diagrams could have a linear part.
Figure 12 a-d: Space-time (Stommel) diagrams based on q=2.0 moments using absolute second temporal difference flux estimate at t=0h (q=0.1 to 2.0 excluding q=1.0) for ERA40 and GFS analysis between 30°N and 30°S of u (red), T (green), h_r or h_s (blue). The black line with slope $H_t = 1$ represents the magnitude of the “typical” transformation velocity. a) ERA40 1000 mb, $V_{trans}=1000$ km/day; b) ERA40 700 mb, $V_{trans}=1000$ km/day; c) GEM 1000 mb, $V_{trans}=500$ km/day; d) GEM 700 mb, $V_{trans}=500$ km/day; e) GFS 1000 mb, $V_{trans}=450$ km/day; f) GFS 700 mb, $V_{trans}=450$ km/day.
over a wider range than the scaling in either space or in time. When the space-time diagramme curves become flat, the climate regime begins. To determine whether or not cascade behaviour is exhibited in the climate regime, the analysis would need to be extended to longer length scales, for which the datasets used were not sufficient.

6. Conclusions

Cascades and brute force numerical approaches to modelling and understanding the atmosphere have co-existed ever since Richardson in the 1920’s. Today, the divorce between these approaches is such that atmospheric scientists only use rather vague cascade ideas based largely on dimensional analysis coupled with strong assumptions about isotropy (in either two or three dimensions), while turbulence scientists focus on verifying the predictions of multiplicative cascades in laboratory experiments, usually – unfortunately – with similarly strong isotropy assumptions. However, due to gravity, the inertial, Coriolis and other forces, the atmosphere is anisotropic, yet scaling over wide ranges of scale, which has been supported by massive satellite-based studies, and reanalyses of classical aircraft studies. Since multiplicative cascades are scaling and can readily be anisotropic, this anisotropy allows the atmosphere to have cascades acting over huge ranges of scale. This leads to the hypothesis that the deterministic numerical models have underlying multiplicative cascade structures. Since the corresponding turbulent fluxes are anisotropic, they will not be classical (e.g. energy, enstrophy); it is important to estimate the fluxes without a priori assumptions.

In a recent paper (Stolle et al., 2009), we studied the spatial cascade structure of the temperature, humidity and horizontal wind for ERA40 reanalyses and for GFS and GEM forecast models is satisfied to typically ±1% up to scales of about 5000 km (for moments up to second order), and for various altitudes, latitudes, and forecast times. In this paper, we extend these analyses to the time domain and then compare space and time cascade structures by constructing space-time diagrammes. The temporal cascades have effective outer scales $\tau_{\text{eff}} \approx$ of around 15 days, although the scaling starts to worsen for scales greater than $\tau_{\text{dev}} \approx$1.25-5 days. We also displayed “Levy collapses” of the temporal moments, which indicate that the generators (logs) of the fluxes are of the predicted “universal” Levy form even outside the scaling regime. We can use the flux statistics to relate the space and time domains. If there is scaling in space and in
time, then the resulting space-time diagrams will be scaling. Unfortunately, the scaling regime is quite narrow in time so that the scaling of the space-time relations are not so good, although we obtain space-time transformation velocities are in the range 250 – 1000 km/day which is of the same order as that predicted by assuming the solar flux drives the dynamics and that the horizontal wind field is dominated by the resulting turbulent energy flux up to planetary scales.

The close equivalence between deterministic models and stochastic cascades opens up new avenues in atmospheric science; this is particularly true since the goal of weather modelling today is increasingly a stochastic one with ensemble forecasting systems becoming the norm. At the moment these systems have difficulty going from deterministic to stochastic descriptions (for example in choosing the ensemble initializations as in the ad hoc methods of Corazza et al. (2003)) as well as in “stochastic parametrisation” of the subgrid effects, which are currently fairly ad hoc (Buizza et al., 1999; Palmer, 2001). Direct exploitation of the known cascade structures would presumably already benefit the latter. However, looking further into the future, rather than making cumbersome attempts to add randomness to a fundamentally deterministic system, it may be possible to use the cascades as a basis of direct stochastic forecasting in which conditional probabilities of the state of the weather at $t = 0$ are used to directly estimate the conditional probabilities (or conditional expectations) of the weather over various forecast horizons. Cascade-based pure stochastic approaches would have the advantage of being able to (statistically) account for huge ranges of scales as well as to be able to make forecasts at any desired resolution.

7. Acknowledgements

We would like to acknowledge NSERC for Jonathan Stolle’s scholarship support as well as Dr. Michel Bourqui, Dr. Andrew Ryzkhov, and Dr. Charles Lin for their helpful discussions and technical expertise. The analysis of GEM data was based on Environment Canada data and the ERA40 data developed by the ECMWF was obtained through UCAR. This research did not enjoy any specific funding.

References


