Fractional relaxation noises, motions and the fractional energy balance equation

Shaun Lovejoy
Physics, McGill University,
3600 University st.
Montreal, Que. H3A 2T8
Canada

Abstract:

We consider the statistical properties of solutions of the stochastic fractional relaxation equation (a fractional Langevin equation) that has been proposed as a model for the Earth's energy balance. In this equation, the (scaling) fractional derivative term models the energy storage processes that occur over a wide range of scales. Up until now, stochastic fractional relaxation processes have been considered in the context of random walk processes where it yields highly nonstationary behaviour. Instead, we consider the stationary solutions of the Weyl fractional relaxation equations whose domain is $-\infty$ to t rather than 0 to t.

We follow a framework developped for handling the simplest fractional equation driven by Gaussian white noise forcings: fractional Gaussian noise (fGn) and fractional Brownian motion (fBm). These more familiar processes are the high frequency limits of the resulting fractional relaxation motions (fRm) and fractional relaxation noises (fRn). Since these processes are Gaussian, their properties are determined by their second order statistics; using Fourier and Laplace techniques, we analytically develop power series as well as asymptotic expansions. We show extensive analytic and numerical results on the autocorrelation functions, Haar fluctuations and spectra. We display sample realizations.

Finally, we discuss the prediction of fRn, fRm which – due to long memories is a *past* value problem, not an *initial* value problem (used for example in monthly and seasonal temperature forecasts). We develop an analytic formula for the fRn forecast skill and compare it to fGn. The large scale white noise limit is attained in a slow power law manner so that when the temporal resolution of the series is small compared to the relaxation time (of the order of a few years in the Earth), fRn can mimic a long memory process with a range of exponents wider than possible with fGn or fBm. We discuss the implications for monthly, seasonal, annual forecasts of the Earth's temperature as well as for projecting the temperature to 2050 and 2100.

1. Introduction:

Over the last decades, stochastic approaches have rapidly developed and have spread throughout the geosciences. From early beginnings in hydrology and turbulence, stochasticity has made inroads in many traditionally deterministic areas. This is notably illustrated by stochastic parametrisations of Numerical Weather Prediction models, e.g.

[*Buizza et al.*, 1999], and the "random" extensions of dynamical systems theory, e.g. [*Chekroun et al.*, 2010].

Pure stochastic approaches have developed primarily along two distinct lines. One is the classical (integer ordered) stochastic differential equation approach based on the Itô or Stratonivch calculii that goes back to the 1950's (see the useful review [*Dijkstra*, 2013]). The other is the scaling strand that encompasses both linear (monofractal, [*Mandelbrot*, 1982]) and nonlinear (multifractal) models (see the review [*Lovejoy and Schertzer*, 2013]) that are based on phenomenological scaling models, notably cascade processes. These and other stochastic approaches have played important roles in nonlinear Geoscience.

Up until now, the scaling and differential equation strands of stochasticity have had surprisingly little overlap. This is at least partly for technical reasons: integer ordered stochastic differential equations have exponential Green's functions that are incompatible with wide range scaling. However, this shortcoming can – at least in principle - be easily overcome by introducing at least some derivatives of fractional order. Once the (typically) ad hoc restriction to integer orders is dropped, the Green's functions are based on "generalized exponentials" that are in turn are based on fractional powers (see the review [*Podlubny*, 1999]). The integer-ordered stochastic equations that have received most attention are thus the exceptional, nonscaling special cases. In physics they correspond to classical Langevin equations; in geophysics and climate modelling, they correspond to the Linear Inverse Modelling (LIM) approach that goes back to [*Hasselmann*, 1976] later elaborated notably by [*Penland and Magorian*, 1993], [*Penland*, 1996], [*Sardeshmukh et al.*, 2000], [*Sardeshmukh and Sura*, 2009] and [*Newman*, 2013]. Although LIM is not the only stochastic approach to climate, in two recent representative multi-author collections ([*Palmer and Williams*, 2010] and [*Franzke and O'Kane*, 2017]), all 32 papers shared the integer ordered assumption (the single exception being [*Watkins*, 2017], see also [*Watkins et al.*, 2020]).

Under the title "Fractal operators" [*West et al.*, 2003], reviews and emphasizes that in order to yield scaling behaviours, it suffices that stochastic differential equations contain fractional derivatives. However, when it is the time derivatives of stochastic variables that are fractional - fractional Langevin equations (FLE) - then the relevant processes are generally non-Markovian [*Jumarie*, 1993], so that there is no Fokker-Planck (FP) equation describing the corresponding probabilities. Furthermore, we expect that - as with the simplest scaling stochastic model – fractional Brownian motion (fBm, [*Mandelbrot and Van Ness*, 1968]) - that the solutions will not be semi-martingales and hence that the Itô calculus used for integer ordered equations will not be applicable (see [*Biagini et al.*, 2008]). This may explain the relative paucity of mathematical literature on stochastic fractional equations (see however [*Karczewska and Lizama*, 2009]). In statistical physics, starting with [*Mainardi and Pironi*, 1996], [*Metzler and Klafter*, 2000], [*Lutz*, 2001] and helped with numerics, the FLE (and a more general "Generalized Langevin Equation" [*Kou and Sunney Xie*, 2004], [*Watkins et al.*, 2019]) has received a little more attention as a model for (nonstationary) particle diffusion (see [*West et al.*, 2003] for an introduction, or [*Vojta et al.*, 2019] for a more recent example).

These technical difficulties explain the apparent paradox of Continuous Time Random Walks (CTRW) and other approaches to anomalous diffusion that involve fractional equations. While CTRW probabilities are governed by the deterministic fractional ordered Generalized Fractional Diffusion equation (e.g. [*Hilfer*, 2000], [*Coffey* *et al.*, 2012]), the walks themselves are based on specific particle jump models rather than (stochastic) Langevin equations. Alternatively, a (spatially) fractional ordered Fokker-Planck equation may be derived from an integer-ordered but nonlinear Langevin equation for a diffusing particle driven by an (infinite variance) Levy motion [*Schertzer et al.*, 2001].

In nonlinear geoscience, it is all too common for mathematical models and techniques developed primarily for mathematical reasons, to be subsequently applied to the real world. This approach - effectively starting with a solution and then looking for a problem - occasionally succeeds, yet historically the converse has generally proved more fruitful. The proposal that an understanding of the Earth's energy balance requires the Fractional Energy Balance Equation (FEBE, [*Lovejoy et al.*, 2020], announced in [*Lovejoy*, 2019]) is an example of the latter. First, the scaling exponent of macroweather (monthly, seasonal, interannual) temperature stochastic variability was determined ($H_I \approx -0.085\pm0.02$) and shown to permit skillful global temperature predictions, [*Lovejoy*, 2015], [*Lovejoy et al.*, 2015], [*Del Rio Amador and Lovejoy*, 2019], [*Del Rio Amador and Lovejoy*, 2020]. Then, the multidecadal deterministic response to external (anthropogenic) forcing was shown to also obey a scaling law but with a different exponent [*Hebert*, 2017], [*Lovejoy et al.*, 2017], [*Procyk et al.*, 2020] ($H_F \approx -0.5\pm0.2$). It was only later that it was realized that the FEBE naturally accounts for both the high and low frequency exponents with $H = H_I + 1/2$ and $H_F = -H$ with the empirical exponents recovered with a FEBE of order $H \approx 0.42\pm0.02$. The realization that the FEBE fit the basic empirical facts motivated the present research into its statistical properties.

The FEBE is a stochastic fractional relaxation equation, it is the FLE for the Earth's temperature treated as a stochastic variable. The FEBE determines the Earth's global temperature when the energy storage processes are scaling and modelled by a fractional time derivative term. Whereas earlier approaches ([*van Hateren*, 2013], [*Rypdal*, 2012], [*Hebert*, 2017], [*Lovejoy et al.*, 2017]) postulated that the climate response function itself is scaling, the FEBE instead situates the scaling in the energy storage processes.

The FEBE differs from the classical energy balance equation (EBE) in several ways. Whereas the EBE is integer ordered and describes the deterministic, exponential relaxation of the Earth's temperature to equilibrium, the FEBE is both stochastic and of fractional order. The FEBE unites the forcing due to internal and external variabilities. Whereas the former represents the forcing and response to the unresolved degrees of freedom - the "internal variability" - and is treated as a zero mean Gaussian noise, the latter represents the external (e.g. anthropogenic) forcing and the forced response modelled by the (deterministic) total external forcing. Complementary work [*Procyk et al.*, 2020] focuses on the deterministic FEBE equation and its application to projecting the Earth's temperature to 2100.

An important but subtle EBE - FEBE difference is that whereas the former is an *initial* value problem whose initial condition is the Earth's temperature at $t = 0$, the FEBE is effectively a *past* value problem whose prediction skill improves with the amount of available past data and - depending on the parameters - it can have an enormous memory. To understand this, recall that an important aspect of fractional derivatives is that they are defined as convolutions over various domains. To date, the main one that has been applied to physical problems is the Riemann-Liouville (RL and the related Caputo) fractional derivative in which the domain of the convolution is the interval between an initial time = 0 and a later time t. This is the domain considered in Podlubny's mathematical monograph on deterministic fractional differential equations [*Podlubny*, 1999] as well as in the stochastic fractional physics discussed in [*West et al.*, 2003], [*Herrmann*, 2011], [*Atanackovic et al.*, 2014], and most of the papers in [*Hilfer*, 2000] (with the partial exceptions of [*Schiessel et al.*, 2000], and [*Nonnenmacher and Metzler*, 2000]). A key point of the FEBE is that it is instead based on Weyl fractional derivatives i.e. derivatives defined over semi-infinite domains, here from $-\infty$ to t. This is the natural range to consider for the Earth's energy balance and it is needed to obtain statistically stationary responses. Although in some respects this semi-infinite domain is easy to handle the statistics of the resulting processes are not available in the literature.

In the EBE, energy storage is modelled by a uniform slab of material implying that when perturbed, the temperature exponentially relaxes to a new thermodynamic equilibrium. However, the actual energy storage involves a hierarchy of mechanisms and the assumption that this storage is scaling is justified by the observed spatial scaling of atmospheric, oceanic and surface (e.g. topographic) structures (reviewed in [*Lovejoy and Schertzer*, 2013]). A consequence is that the temperature relaxes to equilibrium in a power law manner.

This is the phenomenological justification for the FEBE developped in [*Lovejoy et al.*, 2020] where the fractional derivative of order H is an empirically determined parameter with $H = 1$ corresponding to the classical (exponential) exception. Alternatively, [*Lovejoy*, 2020a; b] used Babenko's operator method to show that the special $H = 1/2$ FEBE - the Half-ordered Energy Balance Equation (HEBE) - could be derived analytically from the classical Budyko-Sellers energy balance models ([*Budyko*, 1969], [*Sellers*, 1969]). To obtain the HEBE, it is only necessary to improve the mathematical treatment of the radiative boundary conditions in the classical energy transport equation. In other words, the $H = 1/2$ process discussed below is completely classical.

The purpose of this paper is to understand various statistical properties of the statistically stationary solutions of noise driven fractional relaxation - oscillation equations that underpin the FEBE: "fractional Relaxation noise" (fRn) - and its integral "fractional Relaxation motion" (fRm) with stationary increments. fRn, fRm are direct extensions of the widely studied fractional Gaussian noise (fGn) and fractional Brownian motion (fBm) processes. We derive the main statistical properties of both fRn and fRm including spectra, correlation functions and (stochastic) predictability limits needed for forecasting the Earth temperature ([*Lovejoy et al.*, 2015], [*Del Rio Amador and Lovejoy*, 2019], [*Del Rio Amador and Lovejoy*, 2020]) or projecting it to 2050 or 2100 [*Hébert et al.*, 2020], [*Procyk et al.*, 2020].

The choice of a Gaussian white noise forcing was made both for theoretical simplicity but also for physical realism. While the temperature forcings in the (nonlinear) weather regime are highly intermittent, multifractal, in the lower frequency macroweather regime over which the FEBE applies it quite exceptional inasmuch as its intermittency is low so that the temperature anomalies are not far from Gaussian ([*Lovejoy*, 2018]). Responses to multifractal or Levy process FEBE forcings are likely however to be of interest elsewhere.

This paper is structured as follows. In section 2 we present the classical models of fractional Brownian motion and fractional Gaussian noise as solutions to fractional Langevin equations and define the corresponding fractional Relaxation motions (fRm) and fractional Relaxation noises (fRn) as generalizations. We develop a general framework for handling Gaussian noise driven linear fractional Weyl equations taking care of both high and low frequency divergence issues. Applying this to fBm, fRm we show that they both have stationary increments. Similarly, application of the framework to fGn and fRn shows that they are stationary noises (i.e. with small scale divergences). In section 3 we discuss analytic formulae for the second order statistics including autocorrelations, structure functions, Haar fluctuations and spectra that determine all the corresponding statistical properties (with many details in appendix A). In section 4 we discuss the problem of prediction – important for macrowether forecasting - deriving expressions for the theoretical prediction skill as a function of forecast lead time. In section 5 we conclude and in appendix B, we derive the properties of the HEBE special case.

2. Unified treatment of fBm and fRm:

2.1 fRn, fRm, fGn and fBm

In the introduction, we outlined physical arguments that the Earth's global energy balance could be well modelled by the (linearized) fractional energy balance equation, more details will be published elsewhere. Taking T as the globally averaged temperature, τ_r as the characteristic time scale for energy storage/relaxation processes, F as the (stochastic) forcing (energy flux; power per area), and λ the climate sensitivity (temperature increase per unit flux of forcing) the FEBE can be written in Langevin form as:

$$\tau_r^H\left({}_aD_t^H T\right)+T=\lambda F \qquad , \qquad (1)$$

where the Riemann-Liouville fractional derivative symbol ${}_aD_t^H$ is defined as:

$${}_aD_t^H T=\frac{1}{\Gamma(1-H)}\frac{d}{dt}\int_a^t (t-s)^{-H}\,T(s)\,ds; \quad 0<H<1 \quad , \qquad (2)$$

Where Γ is the standard gamma function. Derivatives of order $\nu>1$ can be obtained using $\nu = H+m$ where m is the integer part of ν, and then applying this formula to the m^{th} ordinary derivative. The main case studied in applications (e.g. random walks) is $a = 0$ so that Laplace transform techniques are often used (alternatively, the somewhat different Caputo fractional derivative is used). However, here we will be interested in $a=-\infty$: the Weyl fractional derivative ${}_{-\infty}D_t^H$ which is naturally handled by Fourier techniques (section 3.5 and appendix A), and in this case, the distinction is unimportant. Although it is in many respects simpler, the statistical characterizations and prediction properties are not available in the literature justifying the following developments.

Since equation 1 is linear, by taking ensemble averages, it can be decomposed into deterministic and random components with the former driven by the mean forcing external to system $\langle F\rangle$, and the latter by the fluctuating stochastic component $F - \langle F\rangle$ representing the internal forcing driving the internal variability. Elsewhere we will consider the deterministic part, in the following, we consider the simplest purely stochastic model in which $\langle F\rangle = 0$ and $F = \gamma$ where γ is a Gaussian "delta correlated" white noise:

$$\langle\gamma(s)\rangle=0; \quad \langle\gamma(s)\gamma(u)\rangle=\delta(s-u) \qquad . \qquad (3)$$

In [*Hebert*, 2017], [*Lovejoy et al.*, 2017], [*Hébert et al.*, 2020] it was argued on the basis of an empirical study of ocean- atmosphere coupling that $\tau_r \approx 2$ years (recent work indicates a value somewhat higher, (≈ 5 years, [*Procyk et al.*, 2020]) and in [*Lovejoy et al.*, 2015] and [*Del Rio Amador and Lovejoy*, 2019] that the value $H \approx 0.4$ reproduced both the Earth's temperature both at scales $< \tau_r$ as well as for macroweather scales (longer than the weather regime scales of about 10 days) but still $< \tau_r$.

When $0 < H < 1$, eq. 1 with $\gamma(t)$ replaced by a deterministic forcing is a fractional generalization of the usual ($H = 1$) relaxation equation; when $1 < H < 2$, it is a generalization of the usual ($H = 2$) oscillation equation, the "fractional oscillation equation", see e.g. [*Podlubny*, 1999]. This classification is based on the deterministic equations; for the noise driven equations, we find that there are two critical exponents $H = 1/2$ and $H = 3/2$ and hence three ranges. Although we focus on the range $0 < H < 3/2$ (especially $0 < H < 1/2$), we also give results for the full range $0 < H < 2$ that includes the strong oscillation range.

To simplify the development, we use the relaxation time τ to nondimensionalize time i.e. to replace time by t/τ_r to obtain the canonical Weyl fractional relaxation equation:

$$\left({}_{-\infty}D_t^H + 1\right)U_H = \gamma; \qquad U_H = \frac{dQ_H}{dt} \tag{4}$$

for the nondimensional process U_H. The dimensional solution of eq. 1 with nondimensional $\gamma = \lambda F$ is simply $T(t) = \tau_r^{-1}\, U_H(t/\tau_r)$ so that in the nondimensional eq. 4, the characteristic transition "relaxation" time between dominance by the high frequency (differential) and the low frequency (U_H term) is $t = 1$. Although we give results for the full range $0 < H < 2$ - i.e. both the "relaxation" and "oscillation" ranges – for simplicity, we refer to the solution $U_H(t)$ as "fractional Relaxation noise" (fRn) and to $Q_H(t)$ as "fractional Relaxation motion" (fRm). Note that we take $Q_H(0) = 0$ so that Q_H is related to U_H via an ordinary integral from time = 0 to t and that fRn is only strictly a noise when $H \leq 1/2$.

In dealing with fRn and fRm, we must be careful of various small and large t divergences. For example, eqs. 1 and 4 are the fractional Langevin equations corresponding to generalizations of integer ordered stochastic diffusion equations: the solution with the classical $H = 1$ value is the Ohrenstein-Uhlenbeck process. Since $\gamma(t)$ is a "generalized function" - a "noise" - it does not converge at a mathematical instant in time, it is only strictly meaningful under an integral sign. Therefore, a standard form of eq. 4 is obtained by integrating both sides by order H (i.e. by differentiating by $-H$ and assuming that differentiation and integration of order H commute):

$$U_H(t) = -{}_{-\infty}D_t^{-H}U_H + {}_{-\infty}D_t^{-H}\gamma = -\frac{1}{\Gamma(H)}\int_{-\infty}^{t}(t-s)^{H-1}U_H(s)\,ds + \frac{1}{\Gamma(H)}\int_{-\infty}^{t}(t-s)^{H-1}\gamma(s)\,ds, \tag{5}$$

(see e.g. in [*Karczewska and Lizama*, 2009]). The white noise forcing in the above is statistically stationary; we show below that the solution for $U_H(t)$ is also statistically stationary. It is tempting to obtain an equation for the motion $Q_H(t)$ by integrating eq. 4 from $-\infty$ to t to obtain the fractional Langevin equation: ${}_{-\infty}D_t^H Q_H + Q_H = W$ where W is Wiener process (a standard Brownian motion) satisfying $dW = \gamma(t)dt$. Unfortunately the Wiener process integrated $-\infty$ to t almost surely diverges, hence we relate Q_H to U_H by an integral from 0 to t.

In the high frequency limit, the derivative dominates and we obtain the simpler fractional Langevin equation:

$$ {}_{-\infty}D_t^H F_H = \gamma(t); \quad F_H = \frac{dB_H}{dt} \tag{6} $$

Whose solution F_H is the fractional Gaussian noise process (fGn), and whose integral B_H is fractional Brownian motion (fBm). We thus anticipate that F_H and B_H are the high frequency limits of fRn, fRm.

Before continuing, a comment is necessary on the use of the symbol H that Mandelbrot introduced for fBm in honour of E. Hurst who pioneered the study of long memory processes in Nile flooding [*Hurst*, 1951]. First, note that eq. 6 implies that the root mean square (RMS) increments of B_H over intervals Δt grow as $\left\langle \Delta B_H(\Delta t)^2 \right\rangle^{1/2} \propto \Delta t^{H+1/2}$ (see below). Since fBm is often defined by this scaling property, it is usual to use the fBm exponent $H_B = H+1/2$. In terms of H_B, from eq. 6, we see that fGn (F_H) is a fractional integration of a white noise of order $H = H_B - 1/2$, whereas fBm is an integral of order $H_B + 1/2$, the 1/2 being a consequence of the fundamental scaling of the Wiener measure whose density is $\gamma(t)$. While the parametrization in terms of H_B is convenient for fGn and fBm, in this paper, we follow [*Schertzer and Lovejoy*, 1987] who more generally used H to denote an order of fractional integration. This more general usage includes the use of H as a general order of fractional integration in the Fractionally Integrated Flux (FIF) model [*Schertzer and Lovejoy*, 1987] which is the basis of space-time multifractal modelling (see the monograph [*Lovejoy and Schertzer*, 2013]). In the FIF generalization, the density of a Wiener measure (i.e. the white noise forcing in eq. 6) is replaced by the density of a (conservative) multifractal measure. The scaling of this multifractal measure is different from that of the Wiener measure so that the extra 1/2 term does not appear. A consequence is that in multifractal processes, H simultaneously characterizes the order of fractional differentiation/integration ($H < 0$ or $H > 0$), and has a straightforward empirical interpretation as the "fluctuation exponent" that characterizes the rate at which fluctuations grow ($H > 0$) or decay ($H < 0$) with scale. In comparison, for fBm, the critical H distinguishing integration and differentiation is still zero, but $H > 0$ or $H < 0$ corresponds to fluctuation exponents $H_B > 1/2$ or $H_B < 1/2$; which for these Gaussian processes is termed "persistence" and "antiperistence". There are therefore several H's in the literature and below, we continue to denote the order of the fractional integration by H but we relate it to other exponents as needed.

2.2 Green's functions

As usual, we can solve inhomogeneous linear differential equations by using appropriate Green's functions:

$$ \begin{aligned} F_H(t) &= \int_{-\infty}^{t} G_{0,H}^{(fGn)}(t-s)\gamma(s)\,ds \\ U_H(t) &= \int_{-\infty}^{t} G_{0,H}^{(fRn)}(t-s)\gamma(s)\,ds \end{aligned}, \tag{7} $$

where $G_{0,H}^{(fGn)}$ and $G_{0,H}^{(fRn)}$ are Green's functions for the differential operators corresponding respectively to ${}_{-\infty}D_t^H$ and ${}_{-\infty}D_t^H+1$.

$G_{0,H}^{(fGn)}$ and $G_{0,H}^{(fRn)}$ are the usual "impulse" (Dirac) response Green's functions (hence the subscript "0"). For the differential operator Ξ they satisfy:

$$\Xi G_{0,H}(t)=\delta(t). \tag{8}$$

Integrating this equation we find an equation for their integrals $G_{1,H}$ which are thus "step" (Heaviside, subscript "1") response Green's functions satisfying:

$$\Xi G_{1,H}(t)=\Theta(t);\quad \Theta(t)=\int_{-\infty}^{t}\delta(s)\,ds;$$

$$\frac{dG_{1,H}}{dt}=G_{0,H}, \tag{9}$$

where Θ is the Heaviside (step) function. The inhomogeneous equation:

$$\Xi f(t)=F(t) \tag{10}$$

has a solution in terms of either an impulse or a step Green's function:

$$f(t)=\int_{-\infty}^{t}G_{0,H}(t-s)F(s)\,ds=\int_{-\infty}^{t}G_{1,H}(t-s)F'(s)\,ds;\quad F'(s)=\frac{dF}{ds}, \tag{11}$$

the equivalence being established by integration by parts with the conditions $F(-\infty)=0$ and $G_{1,H}(0) = 0$. The use of the step rather than impulse response is standard in the Energy Balance Equation literature since it gives direct information on energy balance and the approach to equilibrium (see e.g. [*Lovejoy et al.*, 2020]). The step response for the noise is also the basic impulse response function for the motion (although care is needed for the convergence, see below).

For fGn, the Green's functions are simply the kernels of Weyl fractional integrals:

$$F_H(t)=\frac{1}{\Gamma(H)}\int_{-\infty}^{t}(t-s)^{H-1}\gamma(s)\,ds, \tag{12}$$

obtained by integrating both sides of eq. 6 by order H. We conclude:

$$G_{0,H}^{(fGn)}=\frac{t^{H-1}}{\Gamma(H)};\quad G_{1,H}^{(fGn)}=\frac{t^{H}}{\Gamma(H+1)};\qquad -\frac{1}{2}\le H<\frac{1}{2}. \tag{13}$$

Due to the statistical stationarity of the white noise forcing $\gamma(t)$, that the Riemann-Liouville Green's functions can be used:

$$U_H(t) = \int_{-\infty}^{t} G_{0,H}^{(fRn)}(t-s)\gamma(s)ds \qquad (14)$$

with:

$$G_{0,H}^{(fRn)}(t) = \sum_{n=1}^{\infty}(-1)^{n+1}\frac{t^{nH-1}}{\Gamma(nH)} \qquad\qquad 0 < H \le 2 \ , \qquad (15)$$
$$G_{1,H}^{(fRn)}(t) = \sum_{n=1}^{\infty}(-1)^{n+1}\frac{t^{nH}}{\Gamma(nH+1)}$$

so that $G_{0,H}^{(fGn)}$, $G_{1,H}^{(fGn)}$ are simply the first terms in the power series expansions of the corresponding fRn, fRm Green's functions.

We now recall some classical results useful in geophysical applications. First, these Green's functions are often equivalently written in terms of Mittag-Leffler functions ("generalized exponentials"), $E_{\alpha,\beta}$:

$$G_{0,H}^{(fRn)}(t) = t^{H-1}E_{H,H}\left(-t^H\right) \qquad E_{\alpha,\beta}(z) = \sum_{n=0}^{\infty}\frac{z^n}{\Gamma(\alpha n+\beta)}$$

$$G_{1,H}^{(fRn)}(t) = t^H E_{H,H+1}\left(-t^H\right) \qquad H \ge 0 \quad . \qquad (16)$$

Second, we note that at the origin, for $0 < H < 1$, $G_{0,H}$ is singular whereas $G_{1,H}$ is regular so that it is may be advantageous to use the latter (step) response function (for example in the numerical simulations in section 4). These Green's function responses are shown in figure 1. When $0 < H \le 1$, the step response is monotonic; in an energy balance model, this would correspond to relaxation to equilibrium. When $1 < H < 2$, we see that there is overshoot and oscillations around the long term value; it is therefore (presumably) outside the physical range of an equilibrium process.

In order to understand the relaxation process – i.e. the approach to the asymptotic value 1 in fig. 1 for the step response $G_{1,H}$ - we need the asymptotic expansion:

$$G_{\zeta,H}^{(fRn)}(t) = \sum_{n=0}^{\infty}\frac{(-1)^n}{\Gamma(\zeta-nH)}t^{\zeta-1-nH}; \quad t >> 1 \ , \qquad (17)$$

Where $G_{\zeta,H}(t)$ is the ζ order (fractionally) integrated impulse response $G_{0,H}$. Specifically, for $\zeta = 0, 1$ we obtain the special cases corresponding to impulse and step responses:

$$G_{0,H}^{(fRn)}(t) = \sum_{n=0}^{\infty}(-1)^n\frac{t^{-1-nH}}{\Gamma(-nH)}; \quad t >> 1$$

$$G_{1,H}^{(fRn)}(t) = \sum_{n=0}^{\infty}(-1)^n\frac{t^{-nH}}{\Gamma(1-nH)}; \quad t >> 1 \ , \qquad (18)$$

($0 < H < 1$, $1 < H < 2$; note that the $n = 0$ terms are 0, 1 for $G_{0,H}^{(fRn)}$, $G_{1,H}^{(fRn)}$ respectively) [*Podlubny*, 1999], i.e. power laws in t^{-H} rather than t^{H}. According to this, the asymptotic approach to the step function response (bottom row in fig. 1) is a slow, power law process. In the FEBE, this implies for example that the classical CO_2 doubling experiment would yield a power law rather than exponential approach to a new thermodynamic equilibrium. Comparing this to the EBE, i.e. the special case $H = 1$, we have:

$$G_{0,1}(t) = e^{-t}; \quad G_{1,1}(t) = 1 - e^{-t} \quad , \tag{19}$$

so that when $H = 1$, the asymptotic step response is instead approached exponentially fast. There are also analytic formulae for fRn when $H = 1/2$ (the HEBE) discussed in appendix B notably involving logarithmic corrections.

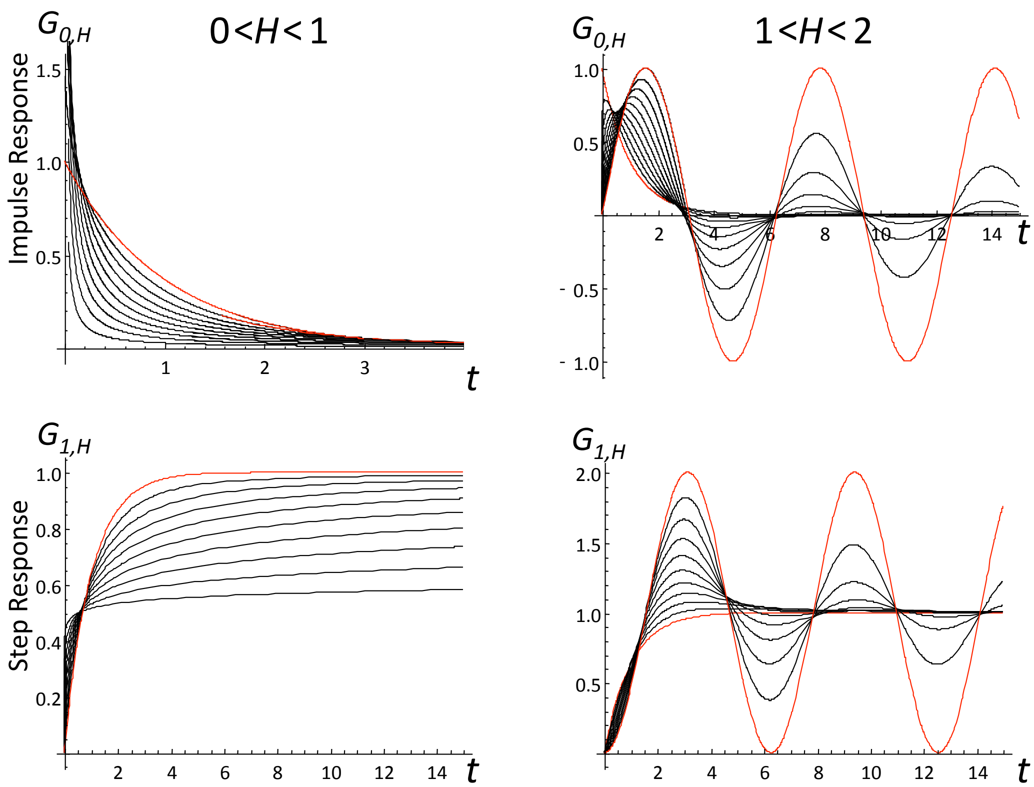

Fig. 1: The impulse (top) and step response functions (bottom) for the fractional relaxation range ($0 < H < 1$, left, red is $H = 1$, the exponential), the black curves, bottom to top are for H = 1/10, 2/10, ..9/10) and the fractional oscillation range ($1 < H < 2$, red are the integer values $H = 1$, bottom, the exponential, and top, $H = 2$, the sine function, the black curves, bottom to top are for H = 11/10, 12/10, ..19/10.

2.3 A family of Gaussian noises and motions:

In the above, we discussed fGn, fRn and their integrals fBm, fRm, but these are simply special cases; a wide variety of Green's functions could be used, for example we expect our approach to applies to the stochastic Basset's equation which could be regarded as a natural extension of the stochastic relaxation equation (see [*Karczewska and Lizama*, 2009] for the more general case of finite and complex vector-valued processes).

With the motivation outlined in the previous sections, and following [*Mandelbrot and Van Ness*, 1968] (see also [*Biagini et al.*, 2008]), the simplest way to proceed is to start by defining the general motion $Z_H(t)$ as:

$$Z_H(t) = N_H \int_{-\infty}^{t} G_{1,H}(t-s)\gamma(s)ds - N_H \int_{-\infty}^{0} G_{1,H}(-s)\gamma(s)ds, \tag{20}$$

where N_H is a normalization constant and H is an index. It is advantageous to rewrite this in standard notation (e.g. [*Biagini et al.*, 2008]) as:

$$Z_H(t) = N_H \int_{\mathbb{R}} \left(G_{1,H}(t-s)_+ - G_{1,H}(-s)_+\right)\gamma(s)ds \quad , \tag{21}$$

where the "+" subscript indicates that the argument is > 0, and the range of integration is over all the real axis $\mathbb{R}$. Here and throughout, the Green's functions need only be specified for $t>0$ corresponding to their causal range.

The advantage of starting with the motion Z_H is that it is based on the step response $G_{1,H}$ which is finite at small t; the disadvantage is that integrals may diverge at large scales. The second (constant) term in eq. 20 was introduced by [*Mandelbrot and Van Ness*, 1968] for fBm precisely in order to avoid large scale divergences in fBm. The introduction of this constant physically corresponds to considering the long-time behaviour of the fractional random walks discussed in [*Kobelev and Romanov*, 2000] and [*West et al.*, 2003]. The physical setting of the random walk applications is a walker starting at the origin corresponding to a fractionally diffusing particle whose velocity obeys the fractional Riemann-Liouville relaxation equation.

From the definition (eq. 20 or 21), we have:

$$Z_H(0) = 0; \quad \langle Z_H(0) \rangle = 0 \quad . \tag{22}$$

Hence, the origin plays a special role, so that the $Z_H(t)$ process is nonstationary.

The variance $V_H(t)$ of Z_H (not to be confused with the velocity of a random walker) is:

$$V_H(t) = \langle Z_H^2(t) \rangle = N_H^2 \int_{\mathbb{R}} \left(G_{1,H}(t-s)_+ - G_{1,H}(-s)_+\right)^2 ds. \tag{23}$$

Equivalently, with an obvious change of variable:

$$V_H(t) = N_H^2 \int_0^{\infty} \left(G_{1,H}(s+t) - G_{1,H}(s)\right)^2 ds + N_H^2 \int_0^{t} G_{1,H}(s)^2 ds, \tag{24}$$

so that $V_H(0) = 0$. Z_H will converge in a root mean square sense if V_H converges. If at large scales $G_{1,H} \propto t^{H_l};\quad t >> 1$, then $H_l < 1/2$ is required for convergence. Similarly, if at small scales $G_{1,H} \propto t^{H_h};\quad t << 1$, then convergence of V_H requires $H_h > -1/2$. We see that for fBm (eq. 13), $H_l = H_h = H$ so that this restriction implies $-1/2 < H < 1/2$ which is equivalent to the usual range $0 < H_B < 1$ with $H_B = H + 1/2$. Similarly, for fRm, using $G^{(fRn)}_{1,H}(t)$, we have $H_h = H$, (eq. 15) and $H_l = -H$, (eq. 18) so that fRm converges for $H > -1/2$, i.e. over the entire range $0 < H < 2$ discussed in this paper. Since the small scale limit of fRm is fBm, we see that the range $0 < H < 2$ overlaps with the range of fBm and extends it at large H.

From eq. 20 we can consider the statistics of the increments:

$$\begin{aligned} Z_H(t)-Z_H(u) &= N_H \int_{\mathbb{R}} \left(G_{1,H}(t-s)_+ - G_{1,H}(u-s)_+\right)\gamma(s)ds \\ &\overset{d}{=} N_H \int_{\mathbb{R}} \left(G_{1,H}(t-u-s')_+ - G_{1,H}(-s')_+\right)\gamma(s')ds'; \qquad s' = s-u \end{aligned}, \quad (25)$$

where we have used the fact that $\gamma(s')\overset{d}{=}\gamma(s)$ where $\overset{d}{=}$ means equality in a probability sense. This shows that:

$$Z_H(t)-Z_H(u)\overset{d}{=}Z_H(t-u)-Z_H(0)=Z_H(t-u), \quad (26)$$

so that the increments $Z_H(t)$ are stationary. From this, we obtain the variance of the increments $\Delta Z_H(\Delta t) = Z_H(t)-Z_H(t-\Delta t)$:

$$\left\langle \Delta Z_H(\Delta t)^2 \right\rangle = V_H(\Delta t); \quad \Delta t = t-u. \quad (27)$$

Since $Z_H(t)$ is a mean zero Gaussian process, its statistics are determined by the covariance function:

$$C_H(t,u) = \left\langle Z_H(t) Z_H(u)\right\rangle = \frac{1}{2}\left(V_H(t)+V_H(u)-V_H(t-u)\right). \quad (28)$$

The noises are the derivatives of the motions and as we mentioned, depending on H, we only expect their finite integrals to converge. Let us therefore define the resolution τ noise $Y_{H,\tau}$ corresponding to the mean increments of the motions:

$$Y_{H,\tau}(t) = \frac{Z_H(t)-Z_H(t-\tau)}{\tau}. \quad (29)$$

The noise, $Y_H(t)$ can now be obtained as the limit $\tau \to 0$:

$$Y_H(t) = \frac{dZ_H(t)}{dt}. \quad (30)$$

Applying eq. 27, we obtain the variance:

$$\left\langle Y_{H,\tau}(t)^2\right\rangle = \left\langle Y_{H,\tau}{}^2\right\rangle = \tau^{-2}V_H(\tau), \quad (31)$$

since $\left\langle Y_{H,t}(0)\right\rangle = 0$, $Y_{H,\tau}(t)$ could be considered as the anomaly fluctuation of Y_H, so that $\tau^{-2}V_H(\tau)$ is the anomaly variance at resolution τ.

From the covariance of Z_H (eq. 28) we obtain the correlation function:

$$\begin{aligned} R_{H,\tau}(\Delta t) &= \left\langle Y_{H,\tau}(t) Y_{H,\tau}(t-\Delta t)\right\rangle = \tau^{-2}\left\langle \left(Z_H(t)-Z_H(t-\tau)\right)\left(Z_H(t-\Delta t)-Z_H(t-\Delta t-\tau)\right)\right\rangle \\ &= \tau^{-2}\frac{1}{2}\left(V_H(\Delta t-\tau)+V_H(\Delta t+\tau)-2V_H(\Delta t)\right) \end{aligned} \qquad \Delta t \geq \tau$$

$$R_{H,\tau}(0)=\left\langle Y_{H,\tau}(t)^2\right\rangle=\tau^{-2}V_H(\tau);\quad \Delta t=0 \quad . \tag{32}$$

Alternatively, taking time in units of the resolution $\lambda = \Delta t/\tau$:

$$\begin{aligned} R_{H,\tau}(\lambda\tau)&=\left\langle Y_{H,\tau}(t)Y_{H,\tau}(t-\lambda\tau)\right\rangle=\tau^{-2}\left\langle\left(Z_H(t)-Z_H(t-\tau)\right)\left(Z_H(t-\lambda\tau)-Z_H(t-\lambda\tau-\tau)\right)\right\rangle \\ &=\tau^{-2}\frac{1}{2}\left(V_H\left((\lambda-1)\tau\right)+V_H\left((\lambda+1)\tau\right)-2V_H(\lambda\tau)\right) \end{aligned} \qquad \lambda\geq 1$$

$$R_{H,\tau}(0)=\left\langle Y_{H,\tau}(t)^2\right\rangle=\tau^{-2}V_H(\tau);\quad \lambda=0 \quad . \tag{33}$$

$R_{H,\tau}$ can be conveniently written in terms of centred finite differences:

$$R_{H,\tau}(\lambda\tau)=\frac{1}{2}\Delta_\tau^2 V_H(\lambda\tau)\approx\frac{1}{2}V_H''(\Delta t);\quad \Delta_\tau f(t)=\frac{f(t+\tau/2)-f(t-\tau/2)}{\tau} \quad . \tag{34}$$

The finite difference formula is valid for $\Delta t \geq \tau$. For finite τ, it allows us to obtain the correlation behaviour by replacing the second difference by a second derivative, an approximation that is very good except when Δt is close to τ.

Taking the limit $\tau \to 0$ in eq. 34 to obtain the second derivative of V_H, and after some manipulations, we obtain the following simple formula for the limiting function $R_H(\Delta t)$:

$$R_H(\Delta t)=\frac{1}{2}\frac{d^2V_H(\Delta t)}{d\Delta t^2}=N_H^2\int_0^\infty G_{0,H}(s+\Delta t)G_{0,H}(s)\,ds;\quad G_{0,H}=\frac{dG_{1,H}}{ds} \quad . \tag{35}$$

If the integral for V_H converges, this integral for $R_H(\Delta t)$ will also converge except possibly at $\Delta t = 0$ (in the examples below, when $H \leq 1/2$).

Eq. 35 shows that R_H is the correlation function of the noise:

$$Y_H(t)=N_H\int_{-\infty}^{t} G_{0,H}(t-s)\gamma(s)\,ds. \tag{36}$$

This result could have been derived formally from:

$$\begin{aligned} Y_H(t)=Z_H'(t)&=\frac{dZ_H(t)}{dt}=N_H\frac{d}{dt}\int_{-\infty}^{t}G_{1,H}(t-s)\gamma(s)\,ds \\ &=N_H\int_{-\infty}^{t}G_{0,H}(t-s)\gamma(s)\,ds \end{aligned} \tag{37}$$

but the above derivation explicitly handles the convergence issues.

A useful statistical characterization of the processes is by the statistics of their Haar fluctuations over an interval Δt. For an interval Δt, Haar fluctuations are the differences between the averages of the first and second halves of an interval. For the noise Y_H, the Haar fluctuation is:

$$\Delta Y_H(\Delta t)_{Haar}=\frac{2}{\Delta t}\int_{t-\Delta t/2}^{t}Y_H(s)\,ds-\frac{2}{\Delta t}\int_{t-\Delta t}^{t-\Delta t/2}Y_H(s)\,ds \quad . \tag{38}$$

In terms of $Z_H(t)$:

$$\Delta Y_H(\Delta t)_{Haar} = \frac{2}{\Delta t}\left(Z_H(t) - 2Z_H(t-\Delta t/2) + Z_H(t-\Delta t)\right). \tag{39}$$

Therefore:

$$\begin{aligned}\left\langle \Delta Y_H(\Delta t)^2_{Haar}\right\rangle &= \left(\frac{2}{\Delta t}\right)^2\left(2\left\langle \Delta Z_H(\Delta t/2)^2\right\rangle - 2\left\langle Y_{H,\Delta t/2}(t)Y_{H,\Delta t/2}(t-\Delta t/2)\right\rangle\right)\\ &= \left(\frac{2}{\Delta t}\right)^2\left(4V_H(\Delta t/2) - V_H(\Delta t)\right)\end{aligned}. \tag{40}$$

This formula will be useful below.

3 Application to fBm, fGn, fRm, fRn:

3.1 fBM, fGn:

The above derivations were for noises and motions derived from differential operators whose impulse and step Green's functions had convergent $V_H(t)$. Before applying them to fRn, fRm, we illustrate this by applying them first to fBm and fGn.

The fBm results are obtained by using the fGn step Green's function (eq. 13) in eq. 24 to obtain:

$$V_H^{(fBm)}(t) = N_H^2\left(\frac{2\sin(\pi H)\Gamma(-1-2H)}{\pi}\right)t^{2H+1}; \quad -\frac{1}{2} \le H < \frac{1}{2}. \tag{41}$$

The standard normalization and parametrisation is:

$$\begin{aligned}N_H = K_H &= \left(\frac{\pi}{2\sin(\pi H)\Gamma(-1-2H)}\right)^{1/2}\\ &= \left(-\frac{\pi}{2\cos(\pi H_B)\Gamma(-2H_B)}\right)^{1/2};\end{aligned} \qquad H_B = H + \frac{1}{2}; \quad 0 \le H_B < 1. \tag{42}$$

This normalization turns out to be convenient not only for fBm but also for fRm so that we use it below to obtain:

$$V_{H_B}^{(fBm)}(t) = t^{2H+1} = t^{2H_B}; \quad 0 \le H_B < 1, \tag{43}$$

so that:

$$\left\langle \Delta B_H(\Delta t)^2\right\rangle^{1/2} = \Delta t^{H_B}; \quad \Delta B_H(\Delta t) = B_H(t) - B_H(t-\Delta t), \tag{44}$$

so – as mentioned earlier - H_B is the fluctuation exponent for fBm. Note that fBm is usually *defined* as the Gaussian process with V_H given by eq. 43 i.e. with this normalization (e.g. [*Biagini et al.*, 2008]).

We can now calculate the correlation function relevant for the fGn statistics. With the normalization:

$$R_{H,\tau}^{(fGn)}(\lambda\tau)=\frac{1}{2}\tau^{2H-1}\left((\lambda+1)^{2H+1}+(\lambda-1)^{2H+1}-2\lambda^{2H+1}\right);\qquad \lambda\geq 1;\quad -\frac{1}{2}<H<\frac{1}{2}$$

$$R_{H,\tau}^{(fGn)}(0)=\tau^{2H-1}$$

$$R_{H_B,\tau}^{(fGn)}(\lambda\tau)\approx H(2H+1)(\lambda\tau)^{2H-1}=H_B(2H_B-1)(\lambda\tau)^{2(H_B-1)};\qquad \begin{matrix}-\frac{1}{2}<H<\frac{1}{2}\\ \lambda>>1\end{matrix}\quad ,\qquad (45)$$

the bottom approximations are valid for large scale ratios λ. We note the difference in sign for $H_B > ½$ ("persistence"), and for $H_B < ½$ ("antipersistence"). When $H_B = ½$, the noise corresponds to standard Brownian motion, it is uncorrelated.

3.2 fRm, fRn

3.2.1 $V_H(t)$

Since fRm, fRn are Gaussian, their properties are determined by their second order statistics, by $V_H(t)$, $R_H(t)$. These statistics are second order in $G_{0,H}(t)$ and can most easily be determined using the Fourier representation of $G_{0,H}(t)$, (section 3.5, appendix A). The development is challenging because unlike the $G_{0,H}(t)$ functions that are entirely expressed in series of fractional powers of t, $V_H(t)$ and $R_H(t)$ involve mixed fractional and integer power expansions, the details are given in appendix A, here we summarize the main results. To lighten the notation, we drop the superscripts "fRn", "fRm" and use the unnormalized functions ($N_H = 1$).

First, for the motions, we have:

$$V_H(t)=2\sum_{n=2}^{\infty}D_n\Gamma(-1-Hn)t^{1+Hn}+2\sum_{j=1,odd}^{\infty}F_j\frac{t^{j+1}}{\Gamma(j+2)};\quad 0<H<2 \qquad (46)$$

$$D_n=(-1)^n\frac{\sin\left(nH\frac{\pi}{2}\right)\sin\left((n-1)H\frac{\pi}{2}\right)}{2\pi\sin\left(H\frac{\pi}{2}\right)}$$

$$F_j=-\frac{1}{\pi H}\cot\left(\frac{\pi H}{2}\right)\left(\Phi\left(-1,1,1-\frac{j}{H}\right)+\Phi\left(-1,1,\frac{j}{H}\right)\right)$$

where Φ is the Hurwitz-Lerch phi function $\Phi(z,s,a)=\sum_{n=0}^{\infty}z^n(n+a)^{-s}$. When $0<H<1/2$, then the leading term is t^{1+Hn} with $n = 2$, so that the coefficient can be used for normalization: $N_H^{-2}=K_H^{-2}=2D_2\Gamma(-1-2H)$ (the fBm normalization). When $½<H<2$, then this becomes negative, so that it cannot be used, however in this case, the leading term is t^2 and its coefficient may be used for normalization:

$$N_H^{-2} = F_1 = -\frac{1}{\pi H}\cot\left(\frac{\pi H}{2}\right)\left(\Phi\left(-1,1,1-\frac{1}{H}\right)+\Phi\left(-1,1,\frac{1}{H}\right)\right) = \int_0^\infty G_{0,H}(s)^2\,ds; \quad 1/2 < H < 2 \tag{47}$$

(see section 3.5 A for the relation with $G_{0,H}$). Since $\Phi\left(-1,1,1-\frac{j}{H}\right)$ diverges for all integer j/H and since we sum over odd integer j, the expansion only converges for irrational H. Therefore, the convergence properties are not clear, but due to the presence of the Γ functions they appear to converge for all t although the convergence is slow (see the numerical results in appendix A, and also for a slightly different expansion that converges more rapidly, useful in applications).

For multidecadal global climate projections, the relaxation time has been estimated at ≈ 5 years ([*Procyk et al.*, 2020]), so that we are interested in the long time behaviour (exploited for example in [*Hébert et al.*, 2020]). For this, asymptotic expansions are useful, in appendix A we show that:

$$V_H(t) = t + a_H - 2\sum_{n=1}^{\infty} D_{-n}\Gamma(-1+nH)t^{1-nH} + 2P_{H,-}(t); \quad t >> 1 \tag{48}$$

where we have included the term:

$$\begin{aligned} P_{H,\pm}(t) &= 0; & 0 < H < 1 \\ P_{H,\pm}(t) &= -e^{t\cos\left(\frac{\pi}{H}\right)} \frac{\sin\left(\pm\frac{\pi}{H}+\frac{H\pi}{2}+t\sin\left(\frac{\pi}{H}\right)\right)}{H\sin\left(\frac{\pi H}{2}\right)}; & 1 < H < 2 \end{aligned} \tag{49}$$

for $1<H<2$, $\cos(\pi/H)<0$ so that at large t, $P_H(t)$ is subdominant, however it explains the oscillations visible in fig. 2. The constant a_H can be determined numerically if needed.

For convenience, the leading terms of the normalized V_H are:

$$V_H^{(norm)}(t) = t^{1+2H} + O\left(t^{1+3H}\right) + O\left(t^2\right); \qquad N_H = K_H = \left(2D_2\Gamma(-1-2H)\right)^{-1/2}; \quad 0 < H < 1/2 \tag{50}$$

and for $½<H<2$, using $N_H = \left(F_1\right)^{-1/2}$:

$$V_H^{(norm)}(t) = t^2 - \frac{2\Gamma(-1-2H)\sin(\pi H)}{\pi F_1}t^{1+2H} + O\left(t^{1+3H}\right); \quad 1/2 < H < 3/2 \tag{51}$$

$$V_H^{(norm)}(t) = t^2 + \frac{F_3}{12F_1}t^4 + O\left(t^{2H+1}\right); \quad 3/2 < H < 2$$

Note that for 3/2<H<2, $F_3 = -\int_0^\infty G'_{0,H}(s)^2 ds$ (appendix A). The change in normalization for $H > 1/2$ is necessary since $K_H^2<0$ for this range. Fig. 2 shows plots of $V^{(norm)}{}_H(t)$, the small t^2 behaviour for H > 1/2 corresponds to fRm increments $\left\langle \Delta Q_H^2(\Delta t)\right\rangle^{1/2} = \left(V_H^{(norm)}(\Delta t)\right)^{1/2} \approx \Delta t$ i.e. to a smooth process, differentiable of order 1; see section 3.4.

Since $\left\langle \Delta Q_H(\Delta t)^2\right\rangle = V_H(\Delta t)$, the corrections imply that at large scales $\left\langle \Delta Q_H(\Delta t)^2\right\rangle^{1/2} < \Delta t^{1/2}$ so that the fRm process Q_H appears to be anti-persistent at large scales.

3.2.2 $R_H(t)$

The formulae for R_H can be obtained from the above using $R_H(t) = (1/2)d^2V_H(t)/dt^2$ (eq. 35, appendix A):

$$R_H(t) = \sum_{n=2}^{\infty} D_n \Gamma(1-Hn)t^{-1+Hn} + \sum_{j=1,odd}^{\infty} F_j \frac{t^{j-1}}{\Gamma(j)} \tag{52}$$

The normalized autocorrelation functions are thus:

$$R_H^{(norm)}(t) = H(1+2H)t^{-1+2H} + O\left(t^{-1+3H}\right); \quad \tau << t << 1; \quad 0 < H < 1/2$$

$$R_H^{(norm)}(t) = 1 - \frac{\left|\Gamma(1-2H)\right|\sin(\pi H)}{\pi F_1} t^{-1+2H} + O\left(t^{-1+3H}\right); \quad t << 1; \quad 1/2 < H < 3/2$$

$$R_H^{(norm)}(t) = 1 + \frac{t^2}{2F_1} F_3 + O\left(t^{-1+2H}\right)...; \quad t << 1; \quad 3/2 < H < 2 \tag{53}$$

(note F_3<0 for 3/2<H<2).

The asymptotic expansions are:

$$R_H(t) = -\sum_{n=1}^{\infty} D_{-n}\Gamma(1+nH)t^{-(1+nH)} + P_{H,+}(t); \quad t >> 1 \tag{54}$$

(when 0 < H < 1/2, for $t \approx \tau$ we must use the exact resolution τ fGn formula, eq. 45, top, note the absolute value sign for ½<H<3/2). For large t:

$$R_H(t) = -\frac{1}{\Gamma(-H)} t^{-1-H} + O\left(t^{-1-2H}\right): \quad 0 < H < 2 \quad ; t>>1. \tag{55}$$

Note that for 0<H<1, Γ(-H)<0 so that R>0 over this range (fig. 3). Formulae 53 shows that there are three qualitatively different regimes: 0 < H < 1/2, 1/2 < H < 3/2, 3/2 < H < 2; this is in contrast with the deterministic relaxation and oscillation regimes (0 < H <1 and 1 < H <2). We return to this in section 3.4.

Now that we have worked out the behaviour of the correlation function, we can comment on the issue of the memory of the process. Starting in turbulence, there is the notion of "integral scale" that is conventionally defined as the long time integral of the correlation function. When the integral scale diverges, the process is conventionally termed a "long memory process". With this definition, if the long time exponent of R_H is > -1, then the process has a long memory. Eq. 55 shows that the long time exponent = -(1+H) so that for all H considered here, the integral scale converges. However, it is of the order of the relaxation time which may be much larger than the length of the available sample series. For example, eq. 55 shows that when H <1/2, the effective exponent $2H$ -1 implies (in the absence of a cut-off), a divergence at long times, so that up to the relaxation scale, fRn mimics a long memory process.

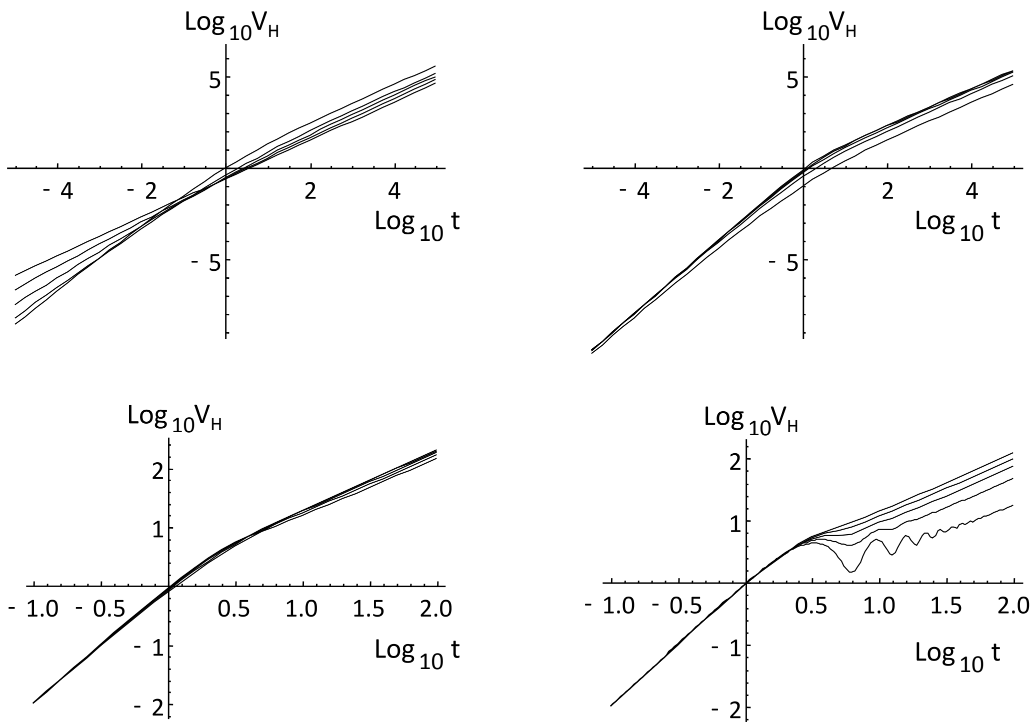

Fig. 2: The normalized V_H functions for the various ranges of H for fRm. The plots from left to right, top to bottom are for the ranges 0 < H<1/2, 1/2 < H < 1, 1 < H < 3/2, 3/2 < H < 2. Within each plot, the lines are for H increasing in units of 1/10 starting at a value 1/20 above the plot minimum; overall, H increases in units of 1/10 starting at a value 1/20, upper left to 39/20, bottom right (ex. for the upper left, the lines are for H = 1/20, 3/10, 5/20, 7/20, 9/20). For all H's the large t behaviour is linear (slope = 1, although note the oscillations for the lower right hand plot for 3/2<H<2). For small t, the slopes are 1+2H (0<H≤1/2) and 2 (1/2≤H<2).

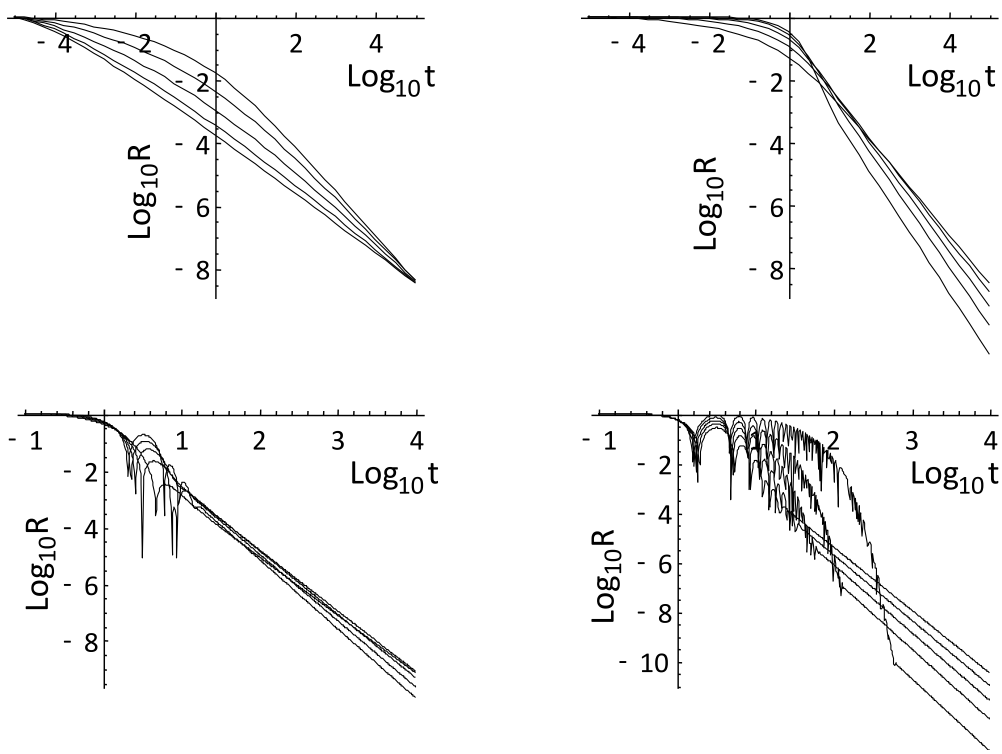

Fig. 3: The normalized correlation functions R_H for fRn corresponding to the V_H function in fig. 2 $0 < H$ 1/2 (upper left) $1/2 < H < 1$ (upper right), $1 < H < 3/2$) lower left, $3/2 < H < 2$ lower right. In each plot, the curves correspond to H increasing from bottom to top in units of 1/10 starting from 1/20 (upper left) to 39/20 (bottom right). For $H<1/2$, the resolution is important since $R_{H,\tau}$ diverges at small τ. In the upper left figure, $R_{H,\tau}$ is shown with $\tau = 10^{-5}$; they were normalized to the value at resolution $\tau = 10^{-5}$, for $H > 1/2$, the curves are normalized with $N_H = F_3^{-1/2}$. In all cases, the large t slope is $-1-H$.

3.3 Haar fluctuations

Using eq. 40 we can determine the behaviour of the RMS Haar fluctuations. Applying this equation to fGn we obtain $\left\langle \Delta F_H\left(\Delta t\right)_{Haar}^2 \right\rangle^{1/2} \propto \Delta t^{H_{Haar}}$ with $H_{Haar} = H - 1/2$ (the subscript "Haar" indicates that this is not a difference/increment fluctuation but rather a Haar fluctuation). For the motion, the Haar exponent is equal to the exponents of the increments (eq. 44) so that $\left\langle \Delta B_H\left(\Delta t\right)_{Haar}^2 \right\rangle^{1/2} \propto \Delta t^{H_{Haar}}$ with $H_{Haar} = H_B = H + 1/2$ ([*Lovejoy et al.*, 2015]). Therefore, from an empirical viewpoint if we have a scaling Gaussian process and (up to the relaxation time scale) when $-1/2 < H_{Haar} < 0$, it has the scaling of an fGn and when $0 < H_{Haar} < 1/2$, it scales as an fBm.

Using eq. 40, we can determine the Haar fluctuations for fRn $\left\langle \Delta U_H(\Delta t)^2_{Haar} \right\rangle^{1/2}$. With the small and large t approximations for $V_H(t)$, we can obtain the small and large Δt behaviour of the Haar fluctuations. Therefore, the leading terms for small Δt are:

$$\left\langle \Delta U_H(\Delta t)^2_{Haar} \right\rangle^{1/2} = \Delta t^{H_{Haar}} \qquad \begin{matrix} H_{Haar} = H - 1/2; & 0 < H < 3/2 \\ H_{Haar} = 1; & 3/2 < H < 2 \end{matrix} \quad ; \quad \Delta t << 1 \quad , \qquad (56)$$

where the $\Delta t^{H\text{-}1/2}$ behaviour comes from terms in $V_H \approx t^{1+2H}$. Note (eq. 40) that $\left\langle \Delta U_H(\Delta t)^2_{Haar} \right\rangle^{1/2}$ depends on $4V_H(\Delta t/2)$ - $V_H(\Delta t)$ so that quadratic terms in $V_H(t)$ cancel. The H_{Haar} = 1 behaviour from the $V_H \approx t^4$ terms that arise when $H > 3/2$.

As H increases past the critical value $H = 1/2$, the sign of H_{Haar} changes so that when $1/2 < H < 3/2$, we have $0 < H_{Haar} < 1$ so that over this range, the small Δt behaviour mimics that of fBm rather than fGn (discussed in the next section).

For large Δt, the corresponding formula is:

$$\left\langle \Delta U^2_{Haar}(\Delta t)^2 \right\rangle^{1/2} \propto \quad \Delta t^{-1/2}; \quad \Delta t >> 1; \quad 0 < H < 2 \quad . \qquad (57)$$

This white noise scaling is due to the leading behavior $V_H(t) \approx t$ over the full range of H (eq. 48), see fig. 4a.

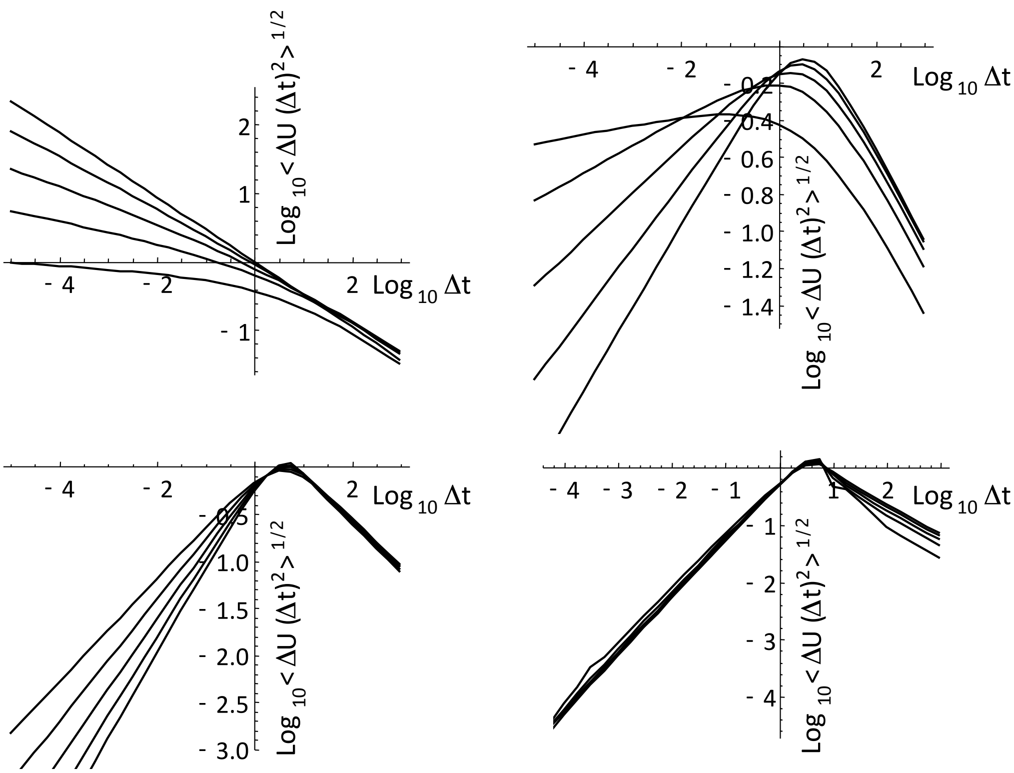

Fig. 4a: The RMS Haar fluctuation plots for the fRn process for $0 < H < 1/2$ (upper left), $1/2 < H < 1$ (upper right), $1 < H < 3/2$ (lower left), $3/2 < H < 2$ (lower right). The individual curves correspond to those of fig. 2, 3. The small Δt slopes follow the theoretical values H - 1/2 up to H = 3/2 (slope= 1); for larger H, the small t slopes all = 1. Also, at large t due to dominant $V \approx t$ terms, in all cases we obtain slopes $t^{-1/2}$.

3.4 fBm, fRm or fGn?

Our analysis has shown that there are three regimes with qualitatively different small scale behaviour, let us compare them in more detail. The easiest way to compare the different regimes is to consider their increments. Since fRn is stationary, we can use:

$$\left\langle \Delta U_H(\Delta t)^2 \right\rangle = \left\langle \left(U_H(t) - U_H(t-\Delta t) \right)^2 \right\rangle = 2\left(R_H^{(fRn)}(0) - R_H^{(fRn)}(\Delta t) \right). \quad (58)$$

Over the various ranges for small Δt, ($\tau \ll 1$ is the resolution) recall that we have:

$$\begin{aligned} &\left\langle \Delta U_{H,\tau}(\Delta t)^2 \right\rangle \approx 2\tau^{-1+2H} - 2H(2H+1)\Delta t^{-1+2H}; \quad 1 >> \Delta t >> \tau; \quad 0 < H < 1/2 \\ &\left\langle \Delta U_H(\Delta t)^2 \right\rangle \approx \Delta t^{-1+2H}; \qquad 1/2 < H < 3/2 \\ &\left\langle \Delta U_H(\Delta t)^2 \right\rangle \approx \Delta t^2; \qquad 3/2 < H < 2 \end{aligned} , \quad (59)$$

(when H>1/2 the resolution is not important, the index is dropped). We see that in the small H range, the increments are dominated by the resolution τ, the process is a noise that does not converge point-wise, hence the τ dependence. In the middle ($1/2 < H < 3/2$) regime, the process is point-wise convergent (take the limit τ->0) although it cannot be differentiated by any positive integer order. Finally, the largest H regime $3/2<H<2$), the process is smoother: $\lim_{\Delta t \to 0} \left\langle \left(\Delta U_H(\Delta t) / \Delta t \right)^2 \right\rangle = 1$, so that it is almost surely differentiable of order 1. Since the fRm are simply order one integrals of fRn, their orders of differentiability are simply augmented by one.

Considering the first two ranges i.e. $0 < H < 3/2$, we therefore have several processes with the same small scale statistics and this may lead to difficulties in interpreting empirical data that cover ranges of time scales smaller than the relaxation time. For example, we already saw that over the range $0 < H < 1/2$ that at small scales we could not distinguish fRn from the corresponding fGn; they both have anomalies (averages after the removal of the mean) or Haar fluctuations that decrease with time scale with exponent H-1/2, (eq. 56). This similitude was not surprising since they both were generated by Green's functions with the same high frequency term. From an empirical point of view, with data only available over scales much smaller than the relaxation time, it might be impossible to distinguish the two; their statistics can be very close.

The problem is compounded when we turn to increments or fluctuations that increase with scale. To see this, note that in the middle range ($1/2 < H < 3/2$), the exponent $-1+2H$ spans the range 0 to 2. This overlaps the range 1 to 2 spanned by fRm (Q_H) with $0 < H < 1/2$:

$$\left\langle \Delta Q_H(\Delta t)^2 \right\rangle = V_H^{(fRm)}(\Delta t) \propto \Delta t^{1+2H}; \quad \Delta t << 1; \quad 0 < H < 1/2 \qquad , \tag{60}$$

and with fBm (B_H) over the same H range (but for all Δt):

$$\left\langle \Delta B_H(\Delta t)^2 \right\rangle = V_H^{(fBm)}(\Delta t) = \Delta t^{1+2H}; \qquad 0 < H < 1/2 \qquad . \tag{61}$$

If we use the usual fBm exponent $H_B = H + 1/2$, then, over the range $0 < H < 1/2$ we may not only compare fBm with fRm with the same H_B, but also with an fRn process with an H larger by unity, i.e. with $H_B = H - 1/2$ in the range $1/2 < H < 3/2$. In this case, we have:

$$\begin{aligned} \left\langle \Delta U_H(\Delta t)^2 \right\rangle &\propto \Delta t^{2H_B}; \quad \Delta t << 1; \quad 0 < H_B < 1 \\ &\propto 2\left(1 - a\Delta t^{-H_B - 3/2}\right); \quad \Delta t >> 1 \end{aligned} \tag{62}$$

$$\begin{aligned} \left\langle \Delta Q_H(\Delta t)^2 \right\rangle &\propto \Delta t^{2H_B}; \quad \Delta t << 1; \quad 1/2 < H_B < 1 \\ &\propto \Delta t - b\Delta t^{3/2 - H_B}; \quad \Delta t >> 1 \end{aligned} \qquad ,$$

$$\left\langle \Delta B_H(\Delta t)^2 \right\rangle = \Delta t^{2H_B}; \qquad 0 < H_B < 1$$

where a, b are constants (section 3.2). Over the entire range $0 < H_B < 1$, we see that the only difference between fBm, and fRn, fRm is their different large scale corrections to the small scale Δt^{2H_B} behaviour. Therefore, if we found a process that over a finite range was scaling with exponent $1/2 < H_B < 1$, then over that range, we could not tell the difference between fRn, fRm, fBm, see fig. 4b for an example with $H_B = 0.95$.

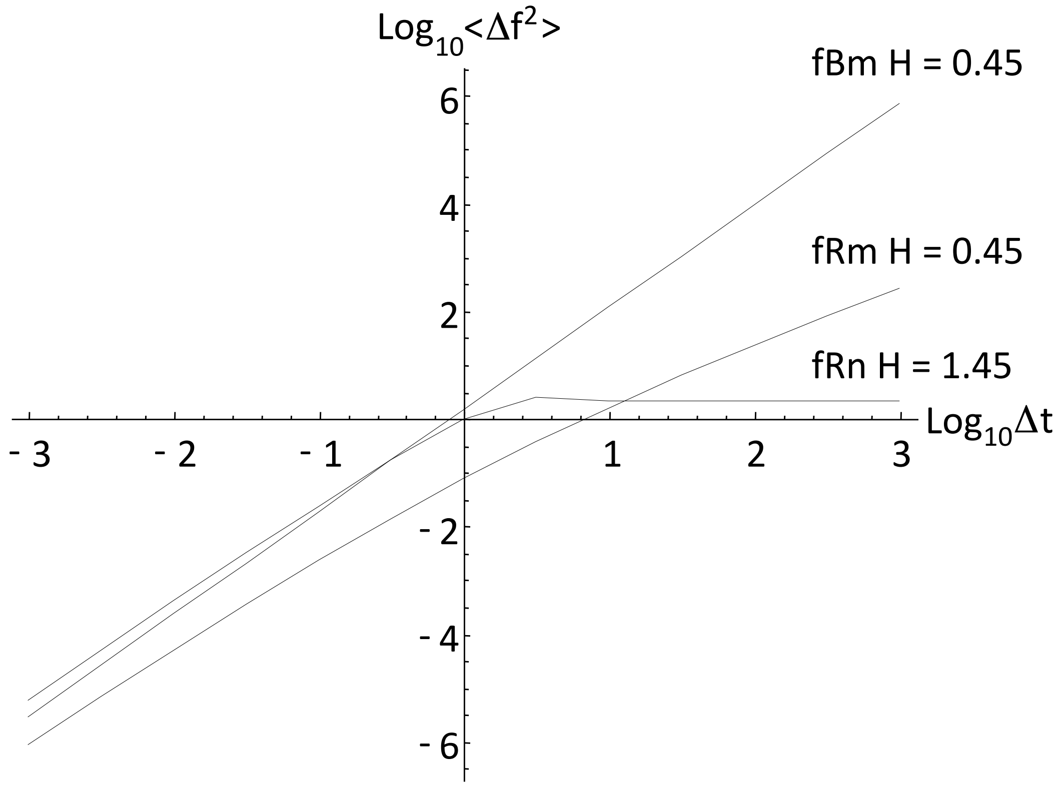

Fig. 4b: A comparison of fRn with $H = 1.45$, fRm with $H = 0.45$ and fBm with $H = 0.45$. For small Δt, they all have RMS increments with exponent $H_B = 0.95$ and can only be distinguished by their behaviours at Δt larger than the relaxation time ($\log_{10}\Delta t = 0$ in this plot).

3.5 Spectra:

Since $Y_H(t)$ is stationary process, its spectrum is the Fourier transform of the correlation function $R_H(t)$ (the Wiener-Khintchin theorem). However, it is easier to determine it directly from the fractional relaxation equation using the fact that the Fourier transform (F.T., indicated by the tilda) of the Weyl fractional derivative is simply $F.T.\left[{}_{-\infty}D_t^H Y_H\right] = \left(i\omega\right)^H \widetilde{Y_H}\left(\omega\right)$ (e.g. [*Podlubny*, 1999], this is simply the extension of the usual rule for the F.T. of integer-ordered derivatives). Therefore take the F.T. of eq. 4 (the fRn), to obtain:

$$\left(\left(i\omega\right)^H + 1\right)\widetilde{U_H} = \tilde{\gamma}, \qquad (63)$$

so that the Fourier transform of $G_{0,H}$ is:

$$\widetilde{G_{0,H}}(\omega)=\frac{1}{1+(i\omega)^{H}} \qquad (64)$$

And the spectrum of Y_H is:

$$\begin{aligned} E_U(\omega)&=\left\langle\left|\widetilde{U_H}(\omega)\right|^2\right\rangle=\left|\widetilde{G_{0,H}}(\omega)\right|^2\left\langle\left|\tilde{\gamma}(\omega)\right|^2\right\rangle=\frac{1}{\left(1+(-i\omega)^H\right)\left(1+(i\omega)^H\right)}.\\ &=\left(1+2\mathrm{Cos}\left(\frac{\pi H}{2}\right)\omega^H+\omega^{2H}\right)^{-1} \end{aligned} \qquad (65)$$

(since the Gaussian white noise was normalized such that $\left\langle\left|\tilde{\gamma}(\omega)\right|^2\right\rangle=1$). Due to the Wiener-Khintchin theorem, the spectrum is the Fourier transform of the autocorrelation function, hence:

$$R_H(t)=\frac{1}{2\pi}\int_{-\infty}^{\infty}e^{i\omega t}E_U(\omega)d\omega=\frac{1}{2\pi}\int_{-\infty}^{\infty}\frac{e^{i\omega t}}{\left(1+(i\omega)^H\right)\left(1+(-i\omega)^H\right)}d\omega \qquad (66)$$

We use this relationship extensively in appendix A in order to derive the main fRn, fRm statistical properties that were discussed above.

From eq. 66 we already can immediately obtain some basic results. First, due to Parseval's theorem:

$$R_H(0)=\frac{1}{2\pi}\int_{-\infty}^{\infty}\left|\widetilde{G_{0,H}}(\omega)\right|^2 d\omega=\int_{0}^{\infty}G_{0,H}(s)^2\,ds \qquad (67)$$

When H<1/2 this is divergent, but when H>1/2, this can be used to normalize R_H.

We may easily obtain the asymptotic high and low frequency behaviours:

$$E_U(\omega)=\begin{array}{ll}\omega^{-2H}+O\left(\omega^{-3H}\right); & \omega>>1\\ 1-2\cos\left(\frac{\pi H}{2}\right)\omega^H+O\left(\omega^{2H}\right) & \omega<<1\end{array}. \qquad (68)$$

This corresponds to the scaling regimes determined by direct calculation above:

$$R_H(t)\propto\begin{array}{ll}t^{-1+2H}+.. & t<<1\\ t^{-1-H}+.. & t>>1\end{array}. \qquad (69)$$

($H\neq1$). Note that the usual (Orenstein-Uhlenbeck) result for $H=1$ has no ω^H term, hence no t^{-1-H} term; it has an exponential rather than power law decay at large t.

3.6 Sample processes

It is instructive to view some samples of fRn, fRm processes. For simulations, both the small and large scale divergences must be considered. Starting with the approximate methods developed by [*Mandelbrot and Wallis*, 1969], it took some time for exact fBm, and fGn simulation techniques to be developed [*Hipel and McLeod*, 1994], [*Palma*, 2007]. Fortunately, for fRm, fRn, the low frequency situation is easier since the long time memory is much smaller than for fBm, fGn. Therefore, as long as we are careful to always simulate series a few times the relaxation time and then to throw away the earliest 2/3 or 3/4 of the simulation, the remainder will have accurate correlations. With this procedure to take care of low frequency issues, we can therefore use the solution for fRn in the form of a convolution (eqs. 19, 35, 36), and use standard numerical convolution algorithms.

However, we still must be careful about the high frequencies since the impulse response Green's functions $G_{0,H}$ are singular for H<1. In order to avoid singularities, simulations of fRn are best made by first simulating the motions Q_H using $Q_H \propto G_{1,H} * \gamma$ (* denotes a Weyl convolution) and obtain the resolution τ fRn, using $U_{H,\tau}(t) = \left(Q_H(t+\tau) - Q_H(t)\right)/\tau$. Numerically, this allows us to use the smoother (nonsingular) G_1 in the convolution rather than the singular G_0. The simulations shown in figs. 5, 6 follow this procedure and the Haar fluctuation statistics were analyzed verifying the statistical accuracy of the simulations.

In order to clearly display the behaviours, recall that when t>>1, we showed that all the fRn converge to Gaussian white noises and the fRm to Brownian motions (albeit in a slow power law manner). At the other extreme, for $t << 1$, we obtain the fGn and fBm limits (when $0 < H < 1/2$) and their generalizations for $1/2 < H < 2$.

Fig. 5a shows three simulations, each of length 2^{19}, pixels, with each pixel corresponding to a temporal resolution of $\tau = 2^{-10}$ so that the unit (relaxation) scale is 2^{10} elementary pixels. Each simulation uses the same random seed but they have H's increasing from $H = 1/10$ (top set) to $H = 5/10$ (bottom set). The fRm at the right is from the running sum of the fRn at the left. Each series has been rescaled so that the range (maximum - minimum) is the same for each. Starting at the top line of each group, we show 2^{10} points of the original series degraded by a factor 2^9. The second line shows a blow-up by a factor of 8 of the part of the upper line to the right of the dashed vertical line. The line below is a further blown up by factor of 8, until the bottom line shows 1/512 part of the full simulation, but at full resolution. The unit scale indicating the transition from small to large is shown by the horizontal red line in the middle right figure. At the top (degraded by a factor 2^9), the unit (relaxation) scale is 2 pixels so that the top line degraded view of the simulation is nearly a white noise (left), (ordinary) Brownian motion (right). In contrast, the bottom series is exactly of length unity so that it is close to the fGn limit with the standard exponent $H_B = H+1/2$. Moving from bottom to top in fig. 5a, one effectively transitions from fGn to fRn (left column) and fBm to fRm (right).

If we take the empirical relaxation scale for the global temperature to be 2^7 months (≈10 years, [*Lovejoy et al.*, 2017]) and we use monthly resolution temperature anomaly data, then the nondimensional resolution is 2^{-7} corresponding to the second series from the top (which is thus 2^{10} months ≈ 80 years long). Since $H \approx 0.42\pm0.02$ ([*Del Rio Amador and Lovejoy*, 2019]), the second series from the top in the bottom set is the most realistic, we can make out the low frequency ondulutions that are mostly present at scales 1/8 of the series (or less).

Fig. 5b shows realizations constructed from the same random seed but for the extended range $1/2 < H < 2$ (i.e. beyond fGn). Over this range, the top (large scale, degraded resolution) series is close to a white noise (left) and Brownian motion (right). For the bottom series, there is no equivalent fGn or fBm process, the curves become smoother although the rescaling may hide this somewhat (see for example the H = 13/20 set, the blow-up of the far right 1/8 of the second series from the top shown in the third line. For $1 < H < 2$, also note the oscillations with frequency $2\pi / \sin(\pi / H)$ (eq. 49), this is the fractional oscillation range.

Fig. 6a shows simulations similar to fig. 5a (fRn on the left, fRm on the right) except that instead of making a large simulation and then degrading and zooming, all the simulations were of equal length (2^{10} points), but the relaxation scale was changed from 2^{15} pixels (bottom) to 2^{10}, 2^5 and 1 pixel (top). Again the top is white noise (left), Brownian motion (right), and the bottom is (nearly) fGn (left) and fBm (right), fig. 6b shows the extensions to $1/2 < H < 2$.

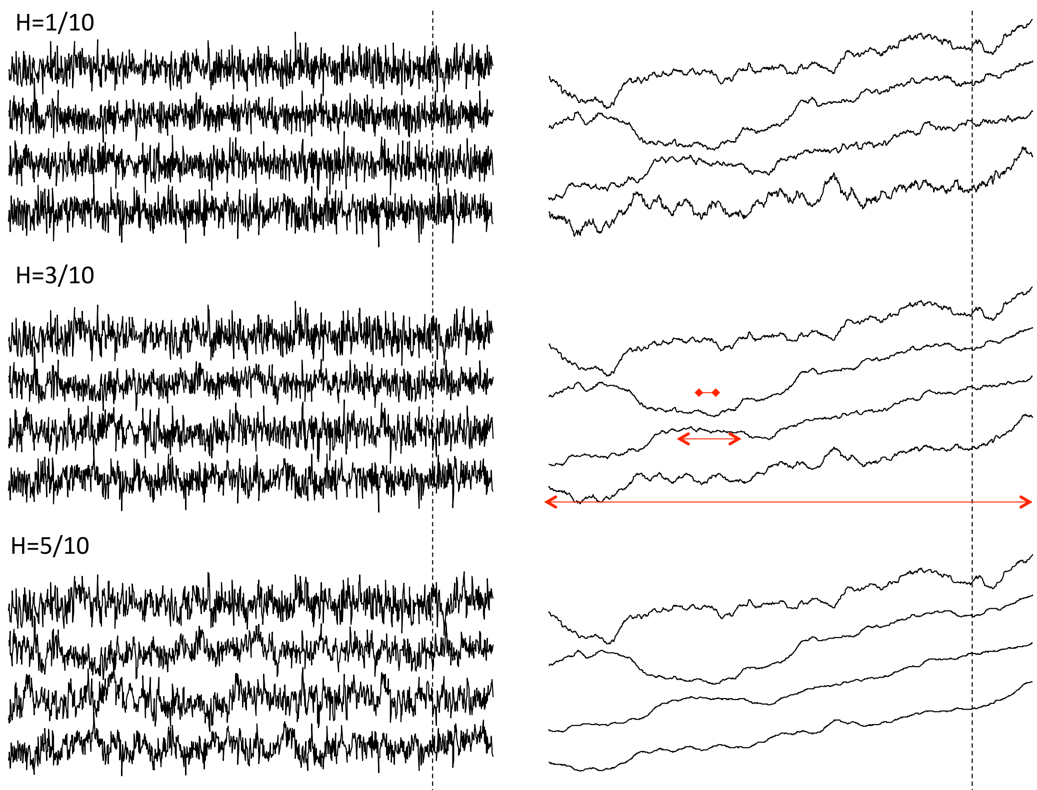

Fig. 5a: fRn and fRm simulations (left and right columns respectively) for H = 1/10, 3/10, 5/10 (top to bottom sets) i.e. the exponent range that overlaps with fGn and fBm. There are three simulations, each of length 2^{19} pixels, each use the same random seed with the unit scale equal to 2^{10} pixels (i.e. a resolution of $\tau = 2^{-10}$). The entire simulation therefore covers the range of scale 1/1024 to 512 units. The fRm at the right is from the running sum of the fRn at the left.

Starting at the top line of each set, we show 2^{10} points of the original series degraded in resolution by a factor 2^9 . Since the length is $t = 2^9$ units long, each pixel has resolution $\tau = 1/2$). The second line of each set takes the segment of the upper line lying to the right of the dashed vertical line, 1/8 of its length. It therefore spans t=0 to $t = 2^9/8 = 2^6$ but resolution was taken as $\tau = 2^{-4}$, hence it is still 2^{10} pixels long. Since each pixel has a resolution of 2^{-4}, the unit scale is 2^4 pixels long, this is shown in red in the second series from the top (middle set). The process of taking 1/8 and blowing up by a factor of 8 continues to the third line (length $t = 2^3$, resolution $\tau = 2^{-7}$), unit scale $=2^7$ pixels (shown by the red arrows in the third series) until the bottom series which spans the range $t = 0$ to $t = 1$ and a resolution $\tau = 2^{-10}$ with unit scale 2^{10} pixels (the whole series displayed). Each series was rescaled in the vertical so that its range between maximum and minimum was the same.

The unit relaxation scales indicated by the red arrows mark the transition from small to large scale. Since the top series in each set has a unit scale of 2 (degraded) it is nearly a white noise (left), or (ordinary) Brownian motion (right). In contrast, the bottom series is exactly of length $t = 1$ so that it is close to the fGn and fBm limits (left and right) with the standard exponent $H_B = H+1/2$. As indicated in the text, the second series from the top in the bottom set is most realistic for monthly temperature anomalies.

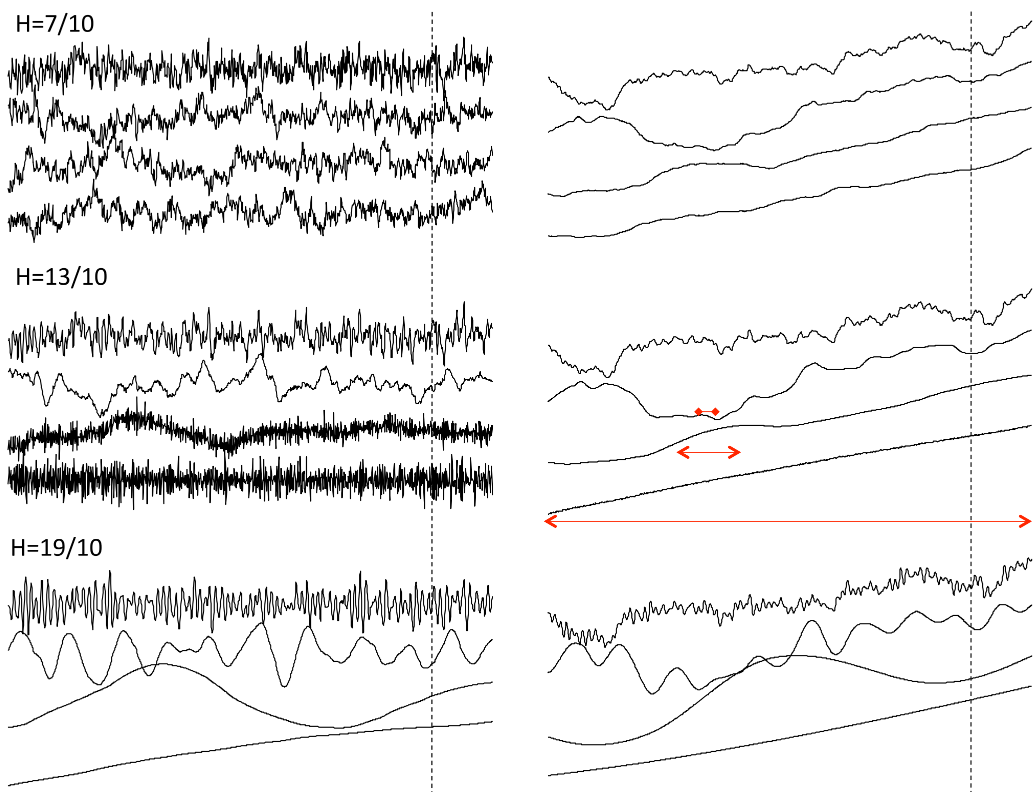

Fig. 5b: The same as fig. 5a but for $H = 7/10$, $13/10$ and $19/10$ (top to bottom). Over this range, the top (large scale, degraded resolution) series is close to a white noise (left) and Brownian motion (right). For the bottom series, there is no equivalent fGn or fBm process, the curves become smoother although the rescaling may hide this somewhat (see for example the middle $H = 13/20$ set, the blow-up of the far right 1/8 of the second series from the top shown in the third line). Also note for the bottom two sets with $1 < H < 2$, the oscillations that have frequency $2\pi / \sin(\pi / H)$, this is the fractional oscillation range.

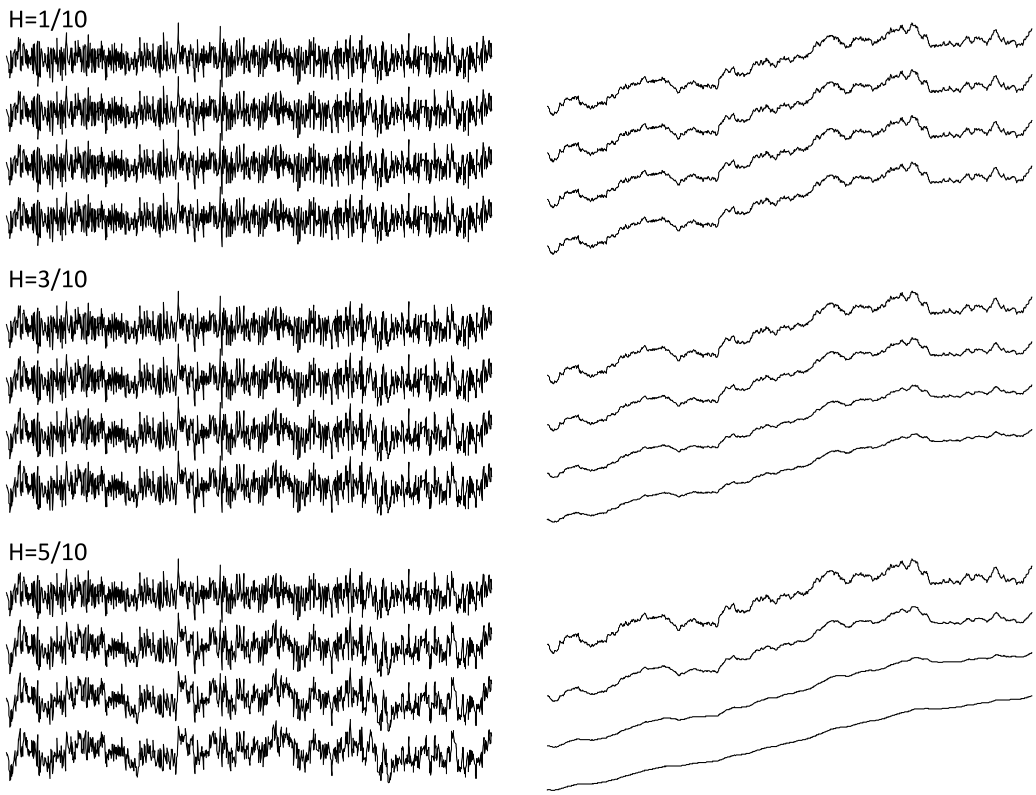

Fig. 6a: This set of simulations is similar to fig. 5a (fRn on the left, fRm on the right) except that instead of making a large simulation and then degrading and zooming, all the simulations were of equal length (2^{10} points), but resolutions $\tau = 2^{-15}, 2^{-10}, 2^{-5}, 1$ (bottom to top). The simulations therefore spanned the ranges of scale 2^{-15} to 2^{-5} ; 2^{-10} to 1 ; 2^{-5} to 2^{5} ; 1 to 2^{10} and the same random seed was used in each so that we can see how the structures slowly change when the relaxation scale changes. The bottom fRn, $H = 5/10$ set is the closest to that observed for the Earth's temperature, and since the relaxation scale is of the order of a few years, the second series from the top of this set (with one pixel = one month) is close to that of monthly global temperature anomaly series. In that case the relaxation scale would be 32 months and the entire series would be $2^{10}/12 \approx$ 85 years long.

The top series (of total length 2^{10} relaxation times) is (nearly) a white noise (left), and Brownian motion (right), and the bottom is (nearly) an fGn (left) and fBm (right). The total range of scales covered here (2^{10}x2^{15}) is larger than in fig. 5a and allows one to more clearly distinguish the high and low frequency regimes.

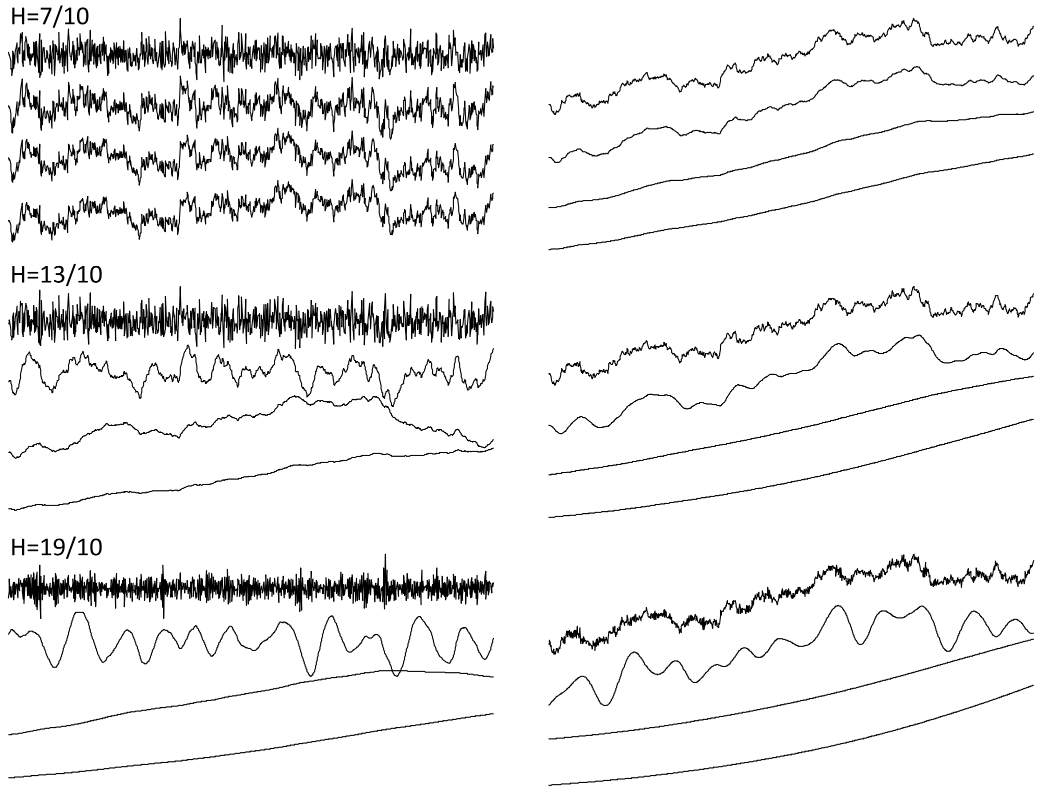

Fig. 6b: The same fig. 6a but for larger H values; see also fig. 5b.

4. Prediction

The initial value for Weyl fractional differential equations is effectively at $t=-\infty$, so that for fRn it is not directly relevant at finite times (although the ensemble mean is assumed = 0; for fRm, $Q_H(0)=0$ is important). The prediction problem is thus to use past data (say, for $t < 0$) in order to make the most skilful prediction of the future noises and motions at $t > 0$. We are therefore dealing with a *past value* rather than a usual *initial value* problem. The emphasis on past values is particularly appropriate since in the fGn limit, the memory is so large that values of the series in the distant past are important. Indeed, prediction of fGn with a finite length of past data involves placing strong (mathematically singular) weights on the most ancient data available (see [*Gripenberg and Norros*, 1996], [*Del Rio Amador and Lovejoy*, 2019], [*Del Rio Amador and Lovejoy*, 2020]). This is quite different from standard stochastic predictions that are based on short memory (exponential) auto-regressive or moving average type processes that are not much different from initial value problems.

In general, there will be small scale divergences (for fRn, when $0 < H \leq 1/2$) so that it is important to predict the finite resolution fRn: $Y_{H,\tau}(t)$. Using eq. 28 for $Y_{H,\tau}(t)$, we have:

$$
\begin{aligned}
Y_{H,\tau}(t) &= \frac{N_H}{\tau}\left[\int_{-\infty}^{t} G_{1,H}(t-s)\gamma(s)ds - \int_{-\infty}^{0} G_{1,H}(-s)\gamma(s)ds\right] - \\
&\quad \frac{N_H}{\tau}\left[\int_{-\infty}^{t-\tau} G_{1,H}(t-\tau-s)\gamma(s)ds - \int_{-\infty}^{0} G_{1,H}(-s)\gamma(s)ds\right] \quad . \qquad (70)\\
&= \frac{N_H}{\tau}\left[\int_{-\infty}^{t} G_{1,H}(t-s)\gamma(s)ds - \int_{-\infty}^{t-\tau} G_{1,H}(t-\tau-s)\gamma(s)ds\right]
\end{aligned}
$$

Let us define the predictor for $t \geq 0$ (indicated by a circonflex):

$$\widehat{Y}_\tau(t) = \frac{N_H}{\tau}\left[\int_{-\infty}^{0} G_{1,H}(t-s)\gamma(s)ds - \int_{-\infty}^{0} G_{1,H}(t-\tau-s)\gamma(s)ds\right]. \qquad (71)$$

To show that it is indeed the optimal predictor, consider the error $E_\tau(t)$ in the predictor:

$$
\begin{aligned}
E_\tau(t) = Y_\tau(t) - \widehat{Y}_\tau(t) &= N_H\tau^{-1}\left[\int_{-\infty}^{t} G_{1,H}(t-s)\gamma(s)ds - \int_{-\infty}^{t-\tau} G_{1,H}(t-\tau-s)\gamma(s)ds\right] \\
&\quad - N_H\tau^{-1}\left[\int_{-\infty}^{0} G_{1,H}(t-s)\gamma(s)ds - \int_{-\infty}^{0} G_{1,H}(t-\tau-s)\gamma(s)ds\right] \quad . \qquad (72)\\
&= N_H\tau^{-1}\left[\int_{0}^{t} G_{1,H}(t-s)\gamma(s)ds - \int_{0}^{t-\tau} G_{1,H}(t-\tau-s)\gamma(s)ds\right]
\end{aligned}
$$

Eq. 72 shows that the error depends only on $\gamma(s)$ for $s>0$ whereas the predictor (eq. 71) only depends on $\gamma(s)$ for $s<0$, hence they are orthogonal:

$$\left\langle E_\tau(t)\widehat{Y}_\tau(t)\right\rangle = 0, \qquad (73)$$

this is a sufficient condition for $\widehat{Y}_\tau(t)$ to be the minimum square predictor which is the optimal predictor for Gaussian processes, (e.g. [*Papoulis*, 1965]). The prediction error variance is:

$$\left\langle E_\tau(t)^2\right\rangle = N_H^2\tau^{-2}\left[\int_{0}^{t-\tau}\left(G_{1,H}(t-s) - G_{1,H}(t-\tau-s)\right)^2 ds + \int_{t-\tau}^{t} G_{1,H}(t-s)^2 ds\right], \qquad (74)$$

or with a change of variables:

$$\left\langle E_\tau(t)^2\right\rangle = \tau^{-2}V_H(\tau) - N_H^2\tau^{-2}\left[\int_{t-\tau}^{\infty}\left(G_{1,H}(u+\tau) - G_{1,H}(u)\right)^2 du\right], \qquad (75)$$

where we have used $\left\langle Y_\tau^{\,2}\right\rangle = \tau^{-2}N_H^{-2}V_H(\tau)$ (the unconditional variance).

Using the usual definition of forecast skill (also called the Minimum Square Skill Score or MSSS) we obtain:

$$S_{k,\tau}(t)=1-\frac{\left\langle E_\tau(t)^2\right\rangle}{\left\langle E_\tau(\infty)^2\right\rangle}=\frac{N_H^2\int_{t-\tau}^{\infty}\left(G_{1,H}(u+\tau)-G_{1,H}(u)\right)^2 du}{V_H(\tau)}$$
$$=\frac{\int_{t-\tau}^{\infty}\left(G_{1,H}(u+\tau)-G_{1,H}(u)\right)^2 du}{\int_0^{\infty}\left(G_{1,H}(u+\tau)-G_{1,H}(u)\right)^2 du+\int_0^{\tau}G_{1,H}(u)^2 du}. \quad (76)$$

When $H < 1/2$ and $G_{1,H}(t)=G_{1,H}^{(fGn)}(t)=\frac{t^H}{\Gamma(1+H)}$, we can check that we obtain the fGn result:

$$\int_{t-\tau}^{\infty}\left(G_{1,H}(u+\tau)-G_{1,H}(u)\right)^2 du\approx\frac{\tau^{1+2H}}{\Gamma(1+H)^2}\int_{\lambda-1}^{\infty}\left((v+1)^H-v^H\right)^2 dv; \quad v=u/\tau; \quad \lambda=t/\tau \quad (77)$$

[*Lovejoy et al.*, 2015]. This can be expressed in terms of the function:

$$\xi_H(\lambda)=\int_0^{\lambda-1}\left((u+1)^H-u^H\right)^2 du, \quad (78)$$

so that the usual fGn result (independent of τ) is:

$$S_k=\frac{\xi_H(\infty)-\xi_H(\lambda)}{\xi_H(\infty)+\frac{1}{2H+1}}. \quad (79)$$

To survey the implications, let's start by showing the τ independent results for fGn, shown in fig. 7 which is a variant on a plot published in [*Lovejoy et al.*, 2015]. We see that when $H \approx 1/2$ ($H_B \approx 1$) that the skill is very high, indeed, in the limit $H \to 1/2$, we have perfect skill for fGn forecasts (this would of course require an infinite amount of past data to attain).

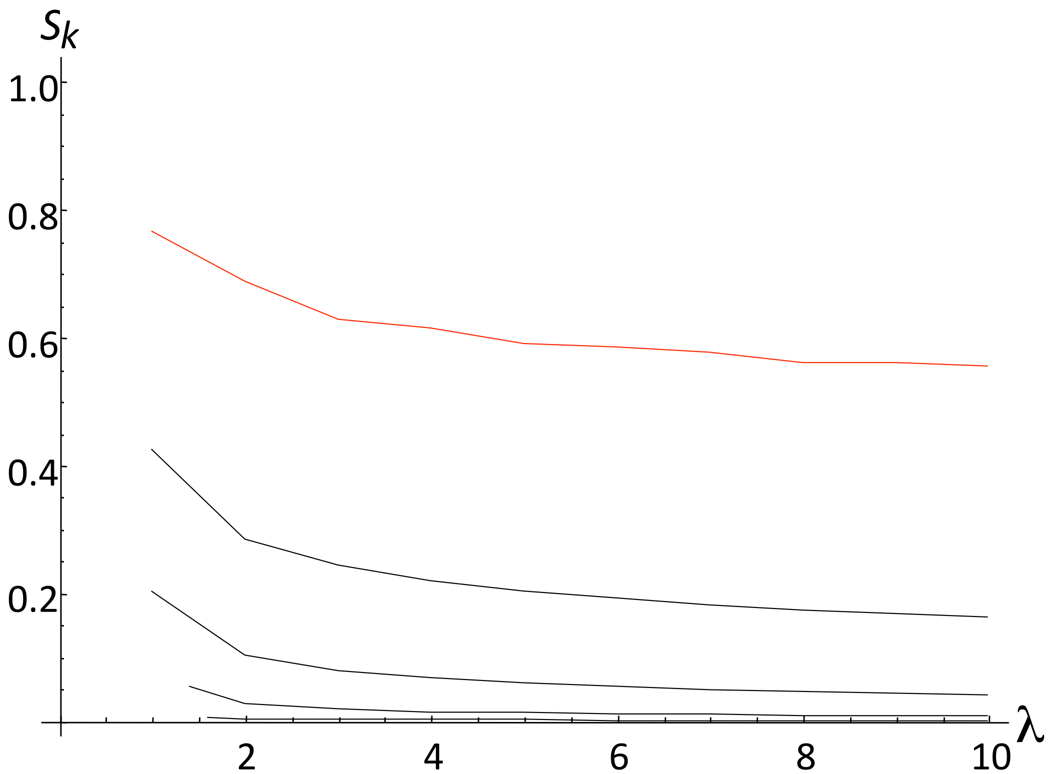

Fig. 7: The prediction skill (S_k) for pure fGn processes for forecast horizons up to $\lambda = 10$ steps (ten times the resolution). This plot is non-dimensional, it is valid for time steps of any duration. From bottom to top, the curves correspond to $H = 1/20$, $3/10$, …$9/20$ (red, top, close to the empirical H).

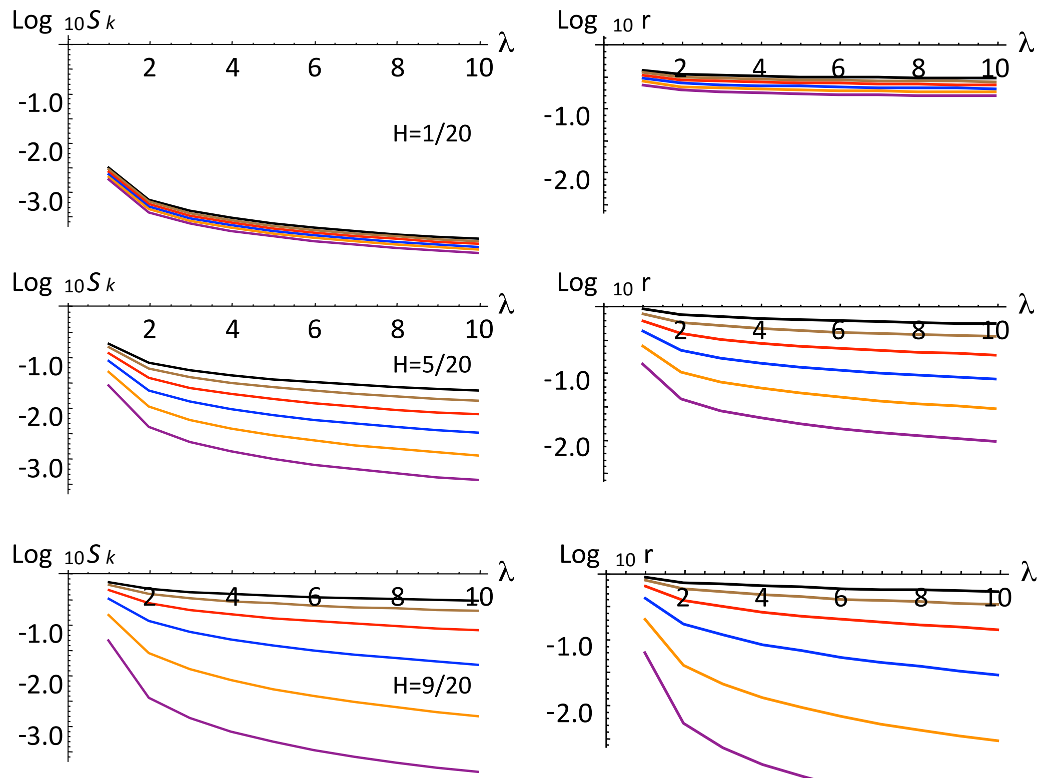

Fig. 8: The left column shows the skill (S_k) of fRn forecasts (as in fig. 7 for fGn) for fRn skill with H = 1/20, 5/20, 9/20 (top to bottom set); λ is the forecast horizon, the number of steps of resolution τ forecast into the future. The right hand column shows the ratio (r) of the fRn to corresponding fGn skill.

Here the result depends on τ; each curve is for different values increasing from 10^{-4} (top, black) to 10 (bottom, purple) increasing by factors of 10 (the red set in the bottom plots with $\tau = 10^{-2}$, H = 9/20 are closest to the empirical values).

Now consider the fRn skill. In this case, there is an extra parameter, the resolution of the data, τ. Figure 8 shows curves corresponding to fig. 7 for fRn with forecast horizons integer multiples (λ) of τ i.e. for times $t = \lambda\tau$ in the future, but with separate curves, one for each of five τ values increasing from 10^{-4} to 10 by factors of ten. When τ is small, the results should be close to those of fGn, i.e. with potentially high skill, and in all cases, the skill is expected to vanish quite rapidly for $\tau>1$ since in this limit, fRn becomes an (unpredictable) white noise (although there are scaling corrections to this).

To better understand the fGn limit, it is helpful to plot the ratio of the fRn to fGn skill (fig. 8, right column). We see that even with quite small values $\tau = 10^{-4}$ (top, black curves), that some skill has already been lost. Fig. 9 shows this more clearly, it shows one time step and ten time step skill ratios. To put this in perspective, it is helpful to compare this using some of the parameters relevant to macroweather forecasting. According to [*Lovejoy et al.*, 2015] and [*Del Rio Amador and Lovejoy*, 2019], the relevant empirical Haar exponent is ≈ -0.08 for the global temperature so that H = 1/2 - 0.08 ≈ 0.42. Although direct empirical estimates of the relaxation time, are difficult since the responses to anthropogenic forcing begin to dominate over the internal variability after ≈10 years [*Procyk et al.*, 2020] have used the determistic response to estiamte a global relaxation time of ≈ 5 years. For monthly resolution forecasts, the non-dimensional resolution is $\tau \approx 1/100$. With these values, we see (red curves) that we may have lost ≈ 30% of the fGn skill for one month forecasts and ≈ 85% for ten month forecasts. Comparing this with fig. 7 we see that this implies about 60% and 10% skill (see also the red curve in fig. 8, bottom set).

Going beyond the $0 < H < 1/2$ region that overlaps fGn, fig. 10 clearly shows that the skill continues to increase with H. We already saw (fig. 4) that the range $1/2 < H < 3/2$ has RMS Haar fluctuations that for $\Delta t < 0$ mimic fBm and these do indeed have higher skill, approaching unity for H near 1 corresponding to a Haar exponent ≈ 1/2, i.e. close to an fBm with $H_B = 1/2$, i.e. a regular Brownian motion. Recall that for Brownian motion, the increments are unpredictable, but the process itself is predictable (persistence).

Finally, in figure 11a, b, we show the skill for various H's as a function of resolution τ. Fig. 11a for the $H < 3/2$ shows that for all H, the skill decreases rapidly for $\tau > 1$. Fig. 12b in the fractional oscillation equation regime shows that the skill also oscillates.

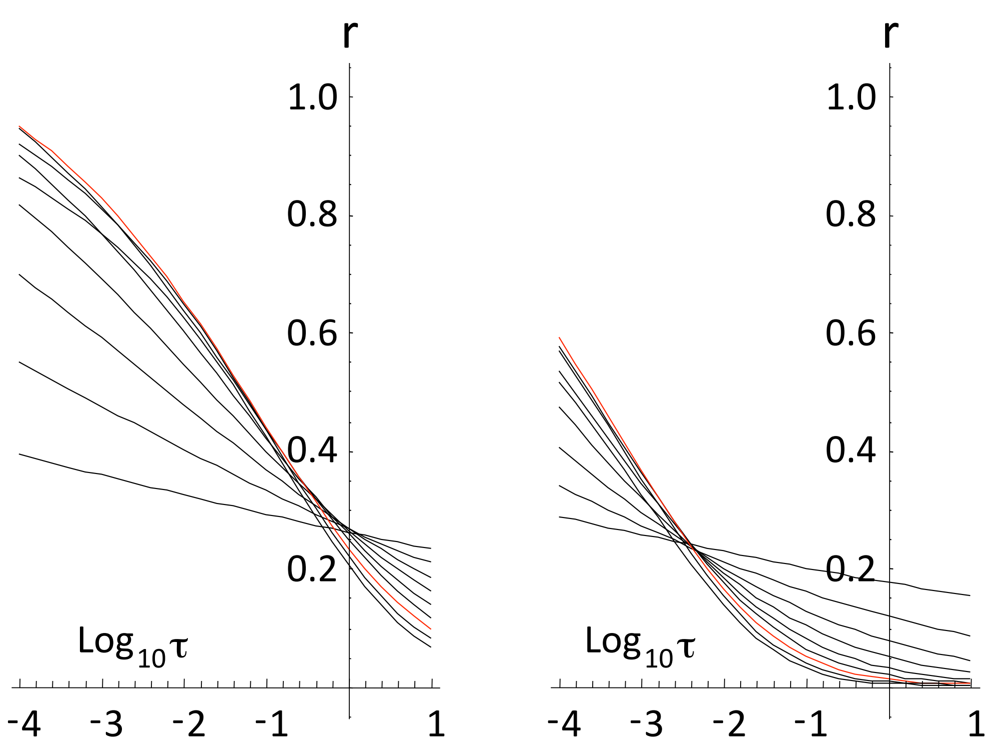

Fig. 9: The ratio of fRn skill to fGn skill (left: one step horizon, right: ten step forecast horizon) as a function of resolution τ for H increasing from (at left) bottom to top (H = 1/20, 2/20, 3/20…9/20); the H = 9/20 curves (close to the empirical value) is shown in red.

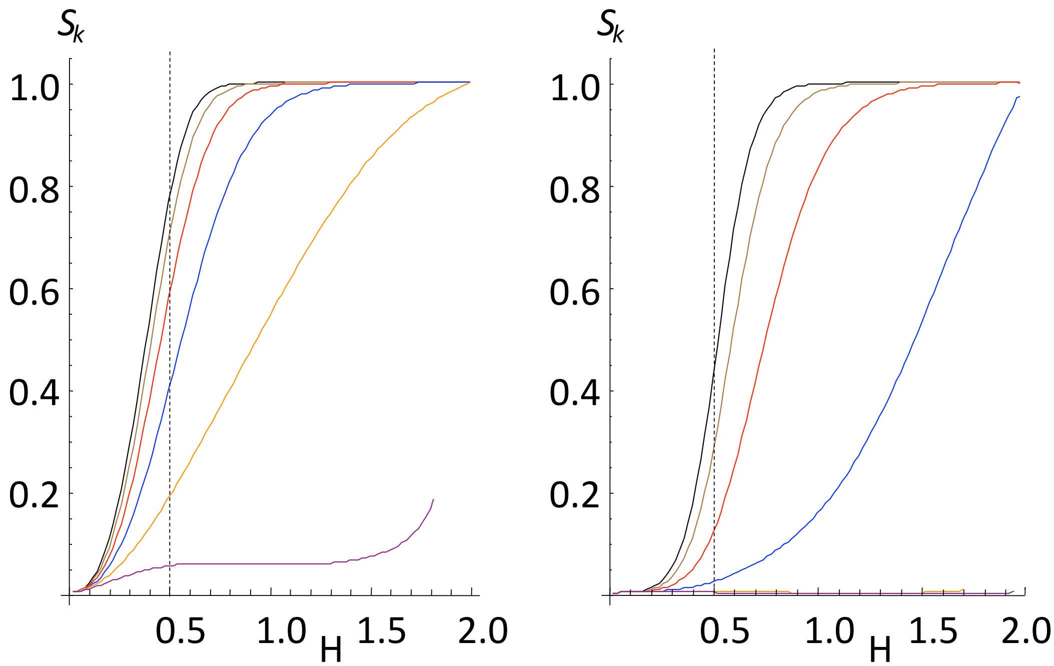

Fig. 10: The one step (left) and ten step (right) fRn forecast skill as a function of H for various resolutions (τ) ranging from $\tau = 10^{-4}$ (black, left of each set) through $\tau = 10^{-3}$ (brown) 10^{-2} (red), 0.1 (blue), 1 (orange), 10 (purple). In the right set $\tau = 1$ (orange), 10 (purple) lines are nearly on top of the $S_k = 0$ line. Again red ($\tau = 10^{-2}$) is the more empirical relevant value for monthly data. Recall that the regime $H < 1/2$ (to the left of the vertical dashed lines) corresponds to the overlap with fGn.

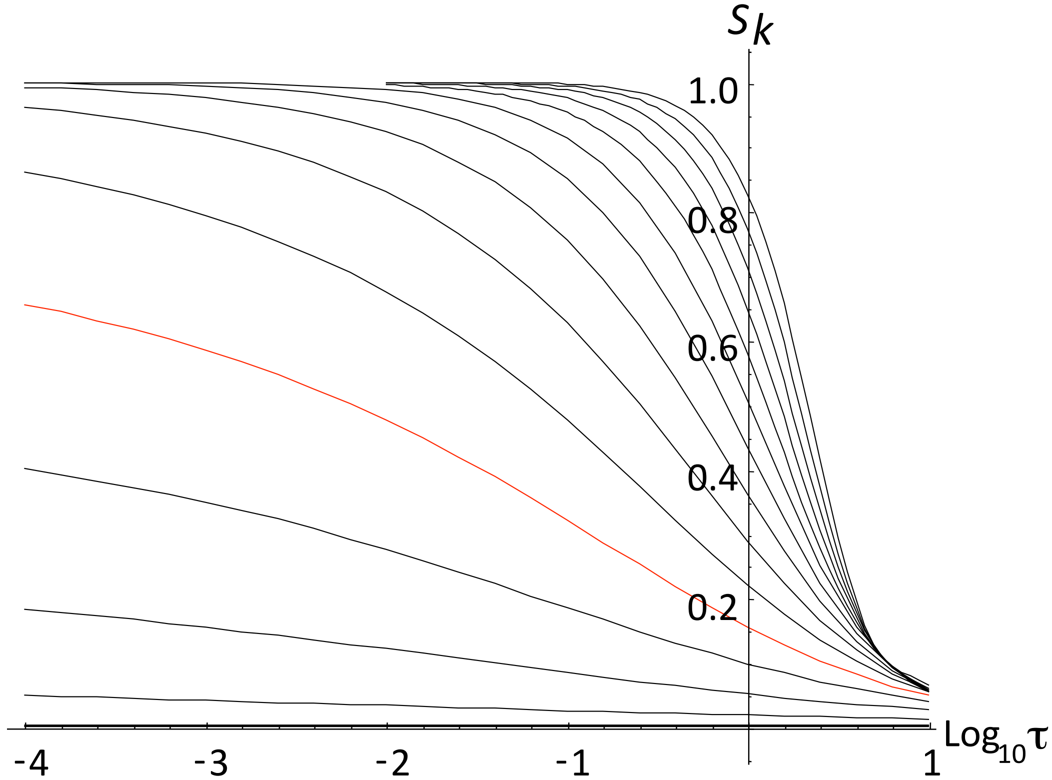

Fig. 11a: One step fRn prediction skills as a function of resolution for *H*'s increasing from 1/20 (bottom) to 29/20 (top), every 1/10. Note the rapid transition to low skill, (white noise) for $\tau > 1$. The curve for $H = 9/20$ is shown in red.

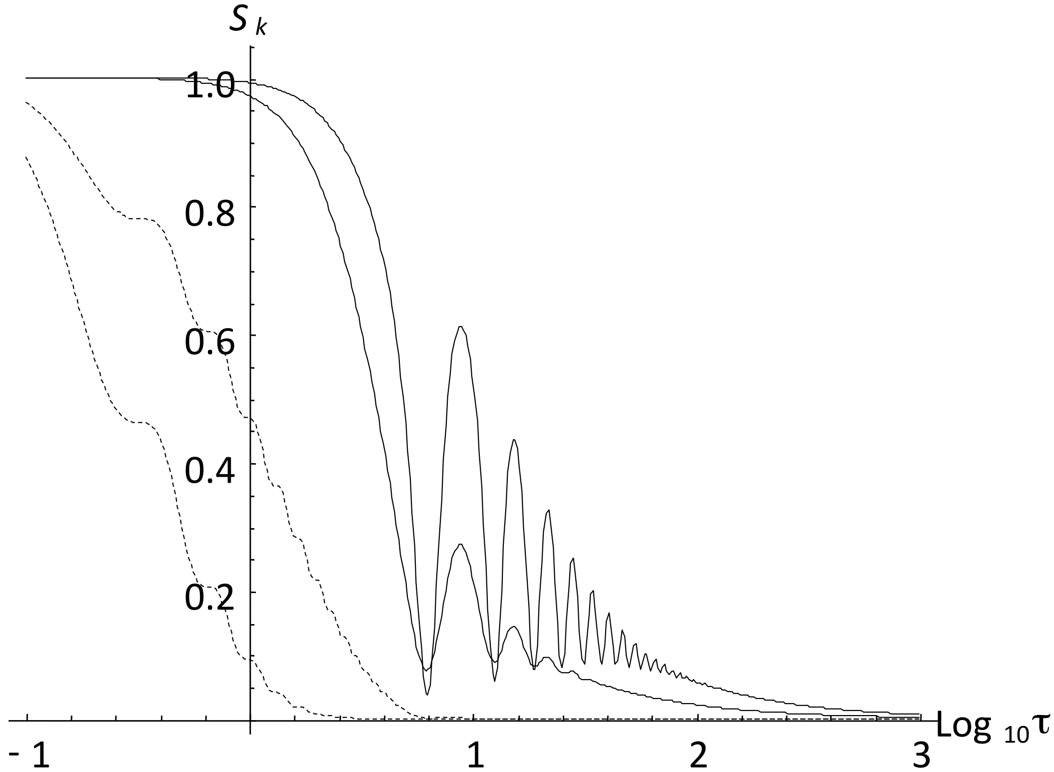

Fig. 11b: Same as fig. 11a except for $H = 37/20$, $39/20$ showing the one step skill (black), and the ten step skill (dashed). The right hand dashed and right hand solid lines, are for $H = 39/20$, they clearly show that the skill oscillates in this fractional oscillation equation regime. The corresponding left lines are for $H = 37/20$.

4. Conclusions:

Ever since [*Budyko*, 1969] and [*Sellers*, 1969], the energy balance between the earth and outer space has been modelled by the Energy Balance Equation (EBE) which is an ordinary first order differential equation for the temperature (Newton's law of cooling). In the EBE, the integer ordered derivative term accounts for energy storage. Physically, it corresponds to storage in a uniform slab of material. To increase realism, one may introduce a few interacting slabs (representing for example the atmosphere and ocean mixed layer; the Intergovernmental Panel on Climate Change recommends two such components [*IPCC*, 2013]). However due to spatial scaling, a more realistic model involves a continuous hierarchy of storage mechanisms and this can easily be modelled by using fractional rather than integer ordered derivatives: the Fractional Energy Balance Equation (FEBE, [*Lovejoy et al.*, 2020]).

The FEBE is a fractional relaxation equation that generalizes the EBE. When forced by a Gaussian white noise, it is also a generalization of fractional Gaussian noise (fGn) and its integral generalizes fractional Brownian motion (fBm). Over the parameter range $0 < H < 1/2$ (H is the order of the fractional derivative), the high frequency FEBE limit (fGn) has been used as the basis of monthly and seasonal temperature forecasts [*Lovejoy et al.*, 2015], [*Del Rio Amador and Lovejoy*, 2019], [*Del Rio Amador and Lovejoy*, 2020]. For multidecadal time scales the low frequency limit has been used as the basis of climate projections through to the year 2100 [*Hebert*, 2017], [*Lovejoy et al.*, 2017], [*Hébert et al.*, 2020], [*Procyk et al.*, 2020]. The success of these two applications with different exponents but with values predicted by the FEBE with the same empirical underlying $H \approx 0.4$, is what originally motivated the FEBE, and the work reported here. The statistical characterizations – correlations, structure functions Haar fluctuations and spectra as well as the predictability properties are important for these and other FEBE applications.

While the deterministic fractional relaxation equation is classical, various technical difficulties arise when it is generalized to the stochastic case: in the physics literature, it is a Fractional Langevin Equation (FLE) that has almost exclusively been considered as a model of diffusion of particles starting at an origin. This requires t = 0 (Riemann-Liouville) initial conditions that imply that the solutions are strongly nonstationary. In comparison, the Earth's temperature fluctuations that are associated with its internal variability are statistically stationary. This can easily be modelled by Weyl fractional derivatives, i.e. initial conditions at $t = -\infty$.

Beyond the proposal that the FEBE is a good model for the Earth's temperature, the key novelty of this paper is therefore to consider the FEBE as a Weyl fractional Langevin equation and proceed to give the fundamental statistical properties including series expansions about the origin and infinity (asymptotic), as well as the theoretical predictability skill. When driven by Gaussian white noises, the solutions are a new stationary process – fractional Relaxation noise (fRn). Over the range $0<H<1/2$, we show that the small scale limit is a fractional Gaussian noise (fGn) – and its integral - fractional Relaxation motion (fRm) - has stationary increments and which generalizes fractional Brownian motion (fBm). Although at long enough times, the fRn tends to a Gaussian white noise, and fRm to a standard Brownian motion, this long time convergence is slow (it is a power law).

The deterministic FEBE has two qualitatively different cases: $0< H <1$ and $1< H <2$ corresponding to fraction relaxation and fractional oscillation processes respectively. In comparison, the stochastic FEBE has three regimes: $0< H < 1/2$, $1/2 < H < 3/2$, $3/2 < H <2$, with the lower ranges ($0 < H < 3/2$) having anomalous high frequency scaling. For example, it was found that fluctuations over scales smaller than the relaxation time can either decay or grow with scale - with exponent H - 1/2 (section 3.5) - the parameter range $0 < H < 3/2$ has the same scaling as the (stationary) fGn ($H < 1/2$) and the (nonstationary) fBm ($1/2 < H < 3/2$), so that processes that have been empirically identified with either fGn or fBm on the basis of their scaling, may in fact turn out to be (stationary) fRn processes; the distinction is only clear at time scales beyond the relaxation time.

Although the basic approach could be more applied to a range of FLEs, we focused on the fractional relaxation-oscillation equation. Much of the effort was to deduce the asymptotic small and large scale behaviours of the autocorrelation functions that determine the statistics and in verifying these with extensive numeric simulations. An interesting exception was the $H = 1/2$ special case which for fGn corresponds to an exactly 1/f noise. Here, we were able to find exact mathematical expressions for the full correlation functions, showing that they had logarithmic dependencies at both small and large scales. The resulting Half order EBE (HEBE) has an exceptionally slow transition from small to large scales (a factor of a million or more is needed) and empirically, it is quite close to the global temperature series over scales of months, decades and possibly longer.

Beyond improved monthly, seasonal temperature forecasts and multidecadal projections, the stochastic FEBE opens up several paths for future research. One of the more promising of these is to follow up on the special value $H = 1/2$ that is very close to that found empirically and that can be analytically deduced from the classical Budyko-Sellers energy transport equation by improving the mathematical treatment of the radiative boundary conditions [*Lovejoy*, 2020a; b]. In the latter case, one obtains a partial fractional differential equation for the horizontal space-time variability of temperature anomalies over the Earth's surface, allowing regional forecasts and projections. Generalizations include the nonlinear albedo-temperature feedbacks needed for modelling of transitions between different past climates.

Acknowledgements:

I thank L. Del Rio Amador, R. Procyk, R. Hébert, C. Penland, N. Watkins for discussions. I also acknowledge an exchange with K. Rypdal. This work was unfunded, there were no conflicts of interest.

Appendix A: The small and large scale fRn, fRm statistics:

A.1 $R_H(t)$ as a Laplace transform

In section 2.3, we derived general statistical formulae for the auto-correlation functions of motions and noises defined in terms of Green's functions of fractional operators. Since the processes are Gaussian, autocorrelations fully determine the statistics. While the autocorrelations of fBm and fGn are well known (and discussed in section 3.1), those for fRm and fRn are new and are not so easy to deal with since they involve quadratic integrals of Mittag-Leffler functions.

In this appendix, we derive the basic power law expansions valid as well as large t (asymptotic) expansions, and we numerically investigate their accuracy. For simplicity, we consider the unnormalized autocorrelation and V functions.

It seems simplest to start with the Fourier expression for the autocorrelation function for the unit white noise forcing (section 3.5), eq. 65, 66. First convert the inverse Fourier transform (eq. 66) into a Laplace transform. For this, consider the integral over the contour C in the complex plane:

$$I(z)=\frac{1}{2\pi}\int_C \frac{e^{zt}}{\left(1+z^H\right)\left(1+(-z)^H\right)}dz \qquad \text{(A1)}$$

We take C to be the closed contour obtained by integrating along the imaginary axis (this part gives R_H(t), eq. 66), and closing the contour along an (infinite) semicircle over the second and third quadrants. When 0<H<1, there are no poles in these quadrants, but we must integrate around a branch cut on the -ve real axis. When 1<H<2, we must take into account two new branch cuts and two new poles in the -ve real plane. In a polar representation $z=re^{i\theta}$, the additional branch cuts are along the rays $z=re^{\pm i\pi/H}$; r>1, circling around the poles at $z=e^{\pm i\pi/H}$. The branch cuts give no net contribution, but the residues of the poles do make a contribution ($P_H \neq 0$ below). We can express both cases with the formula:

$$R_H(t)=-\frac{1}{\pi}\mathrm{Im}\int_0^\infty \frac{e^{-xt}dx}{\left(1+x^H\right)\left(1+x^H e^{i\pi H}\right)}+P_{H,+}(t); \quad t>0 \qquad \text{(A2)}$$

"Im" indicates the imaginary part and:

$$P_{H,\pm}(t)=0; \qquad 0<H<1$$

$$P_{H,\pm}(t)=-e^{t\cos\left(\frac{\pi}{H}\right)}\frac{\sin\left(\pm\frac{\pi}{H}+\frac{H\pi}{2}+t\sin\left(\frac{\pi}{H}\right)\right)}{H\sin\left(\frac{\pi H}{2}\right)}; \quad 1<H<2 \qquad \text{(A3)}$$

While the integral term is monotonic, the P_H term oscillates with frequency $\omega=2\pi/\sin(\pi/H)$. P_H accounts for the oscillations visible in figs. 2, 3, 5b although since when 1<H<2, cos(π/H)<1, they decay exponentially. When H>1, this pole contribution dominates $R_H(t)$ for a wide range of t values around t = 1, although as we see below, eventually at large t, power law terms come to the fore. When $H = 1$, we obtain the classical Ornstein-Uhlenbeck autocorrelation: $R_1(t) = \frac{1}{2}e^{-|t|}$.

A.2 Asymptotic expansions:

An advantage of writing $R_H(t)$ as a Laplace transform is that we can use Watson's lemma to obtain an asymptotic expansion (e.g. [*Bender and Orszag*, 1978]). The idea is that an expansion of eq. A.2 around $x = 0$ can be Laplace transformed term by term to yield an asymptotic expansion for large t. Defining the convenient coefficient:

$$D_n = (-1)^{n+1}\frac{\cos\left(\left(n-\frac{1}{2}\right)\pi H\right)-\cos\left(\frac{\pi H}{2}\right)}{2\pi\sin\left(\frac{\pi H}{2}\right)} = (-1)^n\frac{\sin\left(nH\frac{\pi}{2}\right)\sin\left((n-1)H\frac{\pi}{2}\right)}{\pi\sin\left(H\frac{\pi}{2}\right)} \qquad \text{(A4)}$$

The x expansion of the integrand can be expressed in terms of D_{-n} as:

$$-\frac{1}{\pi}\mathrm{Im}\frac{1}{\left(1+x^H\right)\left(1+x^H e^{Hi\pi}\right)} = -2\sum_{n=1}^{\infty}D_{-n}x^{nH} \qquad \text{(A5)}$$

Therefore, taking the term by term Laplace transform and using Watson's lemma:

$$R_H(t) = -2\sum_{n=1}^{\infty}D_{-n}\Gamma(1+nH)t^{-(1+nH)} + P_{H,+}(t);\quad t>>1 \qquad \text{(A6)}$$

Where we have used $\Gamma(1+Hn)\sin(nH\pi) = -\pi/\Gamma(-nH)$, and have included the exponentially decaying residue $P_{H,+}$ that contributes when $1<H<2$. The first two terms (without P_{H+}) are explicitly:

$$R_H(t) = -\frac{1}{\Gamma(-H)}t^{-(1+H)} + \frac{2+\sec(H\pi)}{2\Gamma(-2H)}t^{-(1+2H)} + \ldots;\quad t>>1 \qquad \text{(A7)}$$

Note that for $0<H<1$ $\Gamma(-H)<0$.

For the motions (fRm), we need the expansion of V_H(t), it can be obtained by using $R_H(t) = \frac{1}{2}\frac{d^2V_H(t)}{dt^2}$ (eq. 35). Integrating R_H twice, we have:

$$V_H(t) = t + a_H - 4\sum_{n=1}^{\infty}D_{-n}\Gamma(-1+nH)t^{1-nH} + 2P_{H,-}(t);\quad t>>1 \qquad \text{(A8)}$$

Where the $t + a_H$ terms come from the constants of integration and P_{H-} from the poles when $1<H<2$. The unit coefficient of the leading t term is a consequence of $\lim_{t\to\infty}\frac{\partial V_H}{\partial t} = 1$. This can be shown by considering the derivative of V_H from eq. 24:

$$\frac{\partial V_H}{\partial t} = J(t) + G_{1,H}(t)^2; \quad J(t) = \int_t^\infty G_{0,H}(u)\left(G_{1,H}(u) - G_{1,H}(u-t)\right)du \tag{A9}$$

Since for $0<H<2$, $G_{0,H}(t)>0$ and $0< G_{1,H}(t)<2$ we obtain:

$$\left|J(t)\right| < \int_t^\infty G_{0,H}(u)\left|G_{1,H}(u) - G_{1,H}(u-t)\right|du < 2\int_t^\infty G_{0,H}(u)\,du \tag{A10}$$

For large t, $G_{0,H}(t) \approx t^{-1-H}$, $\lim_{t\to\infty} J(t) = 0$, in addition $\lim_{t\to\infty} G_{1,H}(t)^2 = 1$ so that $\lim_{t\to\infty}\frac{\partial V_H}{\partial t} = 1$ and to leading order $V_H(t) \approx t$ for large t. If needed, the constant term a_H can be obtained numerically.

A.3 Power series expansions about the origin:

For many applications one is interested in the behavior of R_H(t) for scales of months which is typically less than the relaxation time, i.e. $t<1$. It is therefore important to understand the small t behaviour. We again consider the Laplace integral for the $0<H<1$ case. In this case, we can divide the range of integration in two parts:

$$R_H(t) = -\frac{\mathrm{Im}}{\pi}\int_0^1 \frac{e^{-xt}dx}{\left(1+x^H\right)\left(1+e^{i\pi H}x^H\right)} - \frac{\mathrm{Im}}{\pi}\int_1^\infty \frac{e^{-xt}dx}{\left(1+x^H\right)\left(1+e^{i\pi H}x^H\right)} \tag{A11}$$

and then use the binomial expansions:

$$\frac{1}{\left(1+x^H\right)\left(1+x^H e^{i\pi H}\right)} = \frac{1}{e^{i\pi H}-1}\sum_{n=0}^\infty (-1)^n\left(e^{in\pi H}e^{i\pi H}-1\right)x^{nH}; \quad x<1$$

$$\frac{1}{\left(1+x^H\right)\left(1+x^H e^{i\pi H}\right)} = -\frac{1}{e^{i\pi H}-1}\sum_{n=1}^\infty (-1)^n\left(e^{-in\pi H}e^{i\pi H}-1\right)x^{-nH}; \quad x>1 \tag{A12}$$

We can now integrate each term seperately using:

$$\int_0^1 e^{-xt}x^{nH}\,dx = \sum_{j=1}^\infty \frac{(-1)^{j-1}}{(Hn+j)\Gamma(j)}t^{j-1}$$

$$\int_1^\infty e^{-xt}x^{-nH}\,dx = E_{nH}(t) = \pi\frac{t^{-1+Hn}}{\sin(\pi nH)\Gamma(Hn)} + \sum_{j=1}^\infty \frac{(-1)^{j-1}}{(Hn-j)\Gamma(j)}t^{j-1} \tag{A13}$$

Where E_{nH} is the exponential integral function. Adding the two integrals and summing over n, we obtain:

$$R_H(t) = \sum_{n=2}^\infty D_n\Gamma(1-Hn)t^{-1+Hn} + \sum_{j=1,odd}^\infty F_j\frac{t^{j-1}}{\Gamma(j)} \tag{A14}$$

(note the appearance of D_n with $n>0$) and:

$$F_j = -\frac{1}{\pi}\cot\left(\frac{\pi H}{2}\right)\sum_{n=-\infty}^{\infty}\frac{(-1)^n}{nH+j} = -\frac{1}{\pi H}\cot\left(\frac{\pi H}{2}\right)\left(\Phi\left(-1,1,1-\frac{j}{H}\right)+\Phi\left(-1,1,\frac{j}{H}\right)\right) \tag{A15}$$

where Φ is the Hurwitz-Lerch phi function $\Phi(z,s,a) = \sum_{n=0}^{\infty} z^n (n+a)^{-s}$.

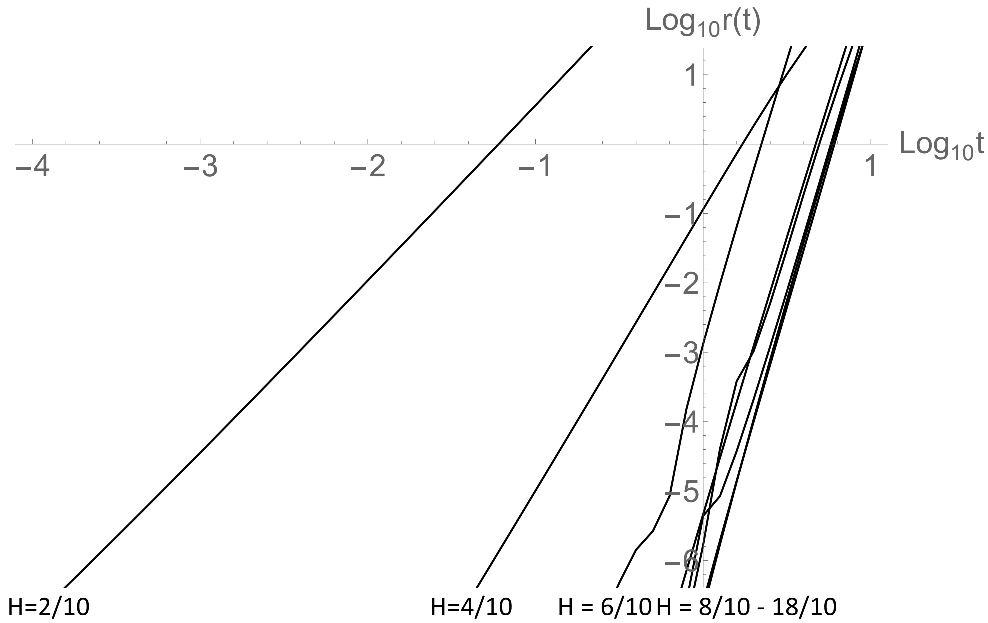

Fig. A1: This shows the logarithm of the relative error in the $R_H^{(10,10)}(t)$ approximation (i.e. with 10 fractional terms and 10 integer order terms) with respect to the deviation from the fGn $R_H(t)$ $r = \log_{10}\left|1-\left(R_H^{fGn}(t)-R_H^{(10,10)}(t)\right)/\left(R_H^{fGn}(t)-R_H(t)\right)\right|$. The lines are for H = 2/10, 4/10,… ,16/10, 18/10 (excluding the exponential case $H = 1$), from left to right (note convergence is only for irrational H, therefore an extra 10^{-4} was added to each H). For the low H values the convergence is particularly slow, but is believed for H.

Comments:

1) These and the following formulae are for $t>0$; R_H is symmetric for $t \rightarrow -t$.

2) Each integer term of the expansion F_j is itself obtained as an infinite sum, so that the overall result for $R_H(t)$ is effectively a doubly infinite sum. This procedure implicitly swaps the order of the summation and apparently explains the fact that while the expansions

were derived for the case $0<H<1$, the final expansion is valid for the full range $0<H<2$: numerically, it accurately reproduces the oscillations when $H>1$.

3) The fGn correlation function is given by the single $n = 2$ term:

$$R_H^{(fGn)}(t) = D_2\Gamma(1-2H)t^{-1+2H} = \frac{\sin(H\pi)}{\pi}\Gamma(1-2H)t^{-1+2H} \qquad \text{(A16)}$$

When $0 < H < 1/2$, it is divergent at the origin; since it corresponds to fGn, the normalization constant is $N_H^{-2} = K_H^{-2} = 2D_2\Gamma(-1-2H)$. When $½< H <2$, it is still the leading term fractional term, but the constant F_1 dominates at small t.

4) The Hurwitz-Lerch phi function $\Phi\left(-1,1,1-\frac{j}{H}\right)$ needed for F_j, diverges for $H = j/n$ where, n is an integer. The overall sum over all j thus diverges for all rational H. For irrational H, the convergence properties are not easy to establish, although due to the Γ functions, these series apparently converge for all $t\geq 0$, but the convergence is rather slow. Fig. A1 shows some numerical results showing the convergence of the 10^{th} order fractional 10^{th} order integer power approximation ($n_{max} = j_{max} =10$). Since the fGn term diverges for small t when $H \leq 1/2$ it is more useful to consider the convergence of the difference with respect to the fGn term (i.e. $R_{fGn}(t) - R_{H,a}(t)$ is the sum in eq. A.15 from $n = 3$ to 10 and odd j ≤ 9). Fig. A1 shows the logarithm of the ratio of the approximation with respect to the true value: $r = \log_{10}\left|1-\left(R_{fGn}(t)-R_{H,a}(t)\right)/\left(R_{fGn}(t)-R_H(t)\right)\right|$ (to avoid exact rationals, 10^{-4} was added to the H values). From the figure we see that the approximation is satisfactory except for small H, we return to this below.

5) For $H>1/2$, when $t = 0$, the only nonzero term is from the constant F_1: $R_H(0) = F_1$, this gives the normalization constant (section 3.2). Comparing with eq. 67, we therefore have:

$$\begin{aligned} R_H(0) &= \int_0^\infty G_{0,H}(s)^2\, ds = F_1 \\ &= -\frac{1}{\pi H}\cot\left(\frac{\pi H}{2}\right)\left(\Phi\left(-1,1,1-\frac{1}{H}\right)+\Phi\left(-1,1,\frac{1}{H}\right)\right) \end{aligned} \qquad 1/2 < H < 2 \qquad \text{(A17)}$$

Similarly, when $H >3/2$, we can apply Parseval's theorem to the derivative $G'_{0,H}$, where it gives the coefficient of the t^2 term so that:

$$\int_0^\infty G'_{0,H}(s)^2\, ds = -F_3 = \frac{1}{\pi H}\cot\left(\frac{\pi H}{2}\right)\left(\Phi\left(-1,1,1-\frac{3}{H}\right)+\Phi\left(-1,1,\frac{3}{H}\right)\right) \qquad \text{(A18)}$$

(when $H<3/2$, the left hand side diverges while the right hand side remains finite).

6) The expression for $V_H(t)$ can be obtained by integrating twice noting that $V_H(0) = 0$, $V'_H(0)=0$:

$$V_H(t) = 2\sum_{n=2}^{\infty} D_n \Gamma(-1-Hn) t^{1+Hn} + 2\sum_{j=1,odd}^{\infty} F_j \frac{t^{j+1}}{\Gamma(j+2)}; \quad 0 < H < 2 \tag{A19}$$

3.4 A Convenient approximation

The expansion for R_H is the sum of a fractional and an integer ordered series. Partial sums appear to converge (fig. A1), albeit slowly. Examination of partial sums shows that the integer ordered and fractional ordered terms tend to cancel, the difficulty due to the term $\Phi\left(-1,1,1-\frac{j}{H}\right)$ that comes from the exponential integral. This suggests an alternative way of expressing the series:

$$R_H(t) = \sum_{n=2}^{\infty} D_n E_{nH}(t) + \sum_{j=1}^{\infty} C_j \frac{(-1)^{j-1}}{\Gamma(j)} t^{j-1}; \quad C_j = \sum_{n=2}^{\infty} \frac{D_n}{(Hn+j)} \tag{A20}$$

Where D_n is given by eq. A.4 and the n sums start at $n = 2$ since $D_1 = 0$. C_j can be expressed as:

$$C_j = -\frac{ie^{-iH\pi}}{2\pi H\left(e^{iH\pi}-1\right)}\left(-\left(e^{iH\pi}+e^{2iH\pi}\right)\Phi\left(-1,1,1+\frac{j}{H}\right)+\Phi\left(e^{iH\pi},1,1+\frac{j}{H}\right)+e^{3iH\pi}\Phi\left(e^{-iH\pi},1,1+\frac{j}{H}\right)\right) \tag{A21}$$

We can also expand the exponential integral:

$$E_{nH}(t) = \pi \frac{t^{-1+Hn}}{\sin(\pi nH)\Gamma(Hn)} + \sum_{j=1}^{\infty} \frac{(-1)^{j-1}}{(Hn-j)\Gamma(j)} t^{j-1} \tag{A22}$$

For the j_{max} and n_{max} partial sums, we have:

$$R_{H,n_{\max},j_{\max}}(t) = \sum_{n=2}^{n_{\max}} D_n \Gamma(1-nH) t^{-1+Hn} + \sum_{j=1}^{j_{\max}} F_{j,n_{\max}} \frac{(-1)^{j-1}}{\Gamma(j)} t^{j-1}; \quad F_{j,n_{\max}} = C_j + \sum_{n=2}^{n_{\max}} \frac{D_n}{Hn-j} \tag{A23}$$

Now define the (j_{max}, n_{max}) approximation by:

$$R_{H,n_{\max},j_{\max}}(t) = \frac{R_H^{(n_{\max}+1,j_{\max})}(t) + R_H^{(n_{\max},j_{\max})}(t)}{2} \tag{A24}$$

This has the effect of adding in half the next higher n term and is more accurate; overall, j_{max} and n_{max} may now be taken to be much smaller than in the previous approximation. For example putting n_{max} =2, j_{max} = 1, we get with the partial sum:

$$R_{H,2,1}(t) = R_H^{(fGn)}(t) + \frac{D_3}{2}\Gamma(1-3H)t^{-1+3H} + F_1 \tag{A25}$$

Where:

$$F_1 = C_1 + \frac{D_2}{2H-1} + \frac{D_3}{2(3H-1)}$$

$$D_2 = \frac{\sin(\pi H)}{\pi}; \quad D_3 = -\frac{\sin(\pi H)(1+2\cos(\pi H))}{\pi} \qquad \text{(A26)}$$

To understand the behaviour, fig. A2 shows the behaviour of coefficient of the t^{-1+3H} term $\frac{D_3}{2}\Gamma(1-3H)$, the constant term F_1 and the coefficient of the next integer (linear in t) term $F_2 = C_2 + \frac{D_2}{2H-2} + \frac{D_3}{2(3H-2)}$. Up until the end of the fGn region (H = 1/2), the t^{-1+3H} and F_1 terms have opposite signs and tend to cancel. In addition, we see that for $t \approx<1$ and $H<1$, they dominate over the (omitted) linear term. Fig. B3 shows that the $R_{H,2,1}$ approximation is surprisingly good for $H<1$ and is still not so bad for $1< H <2$. This approximation is thus useful for monthly resolution macroweather temperature fields that have relaxation times of years or longer and where H is mostly over the range $0< H <1/2$, but over some tropical ocean regions can increase to as much as $H \approx 1.2$ ([*Del Rio Amador and Lovejoy*, 2020]). Fig. A2 shows that the (2,1) approximation is reasonably accurate for $t \approx<1$, especially for $H<1$.

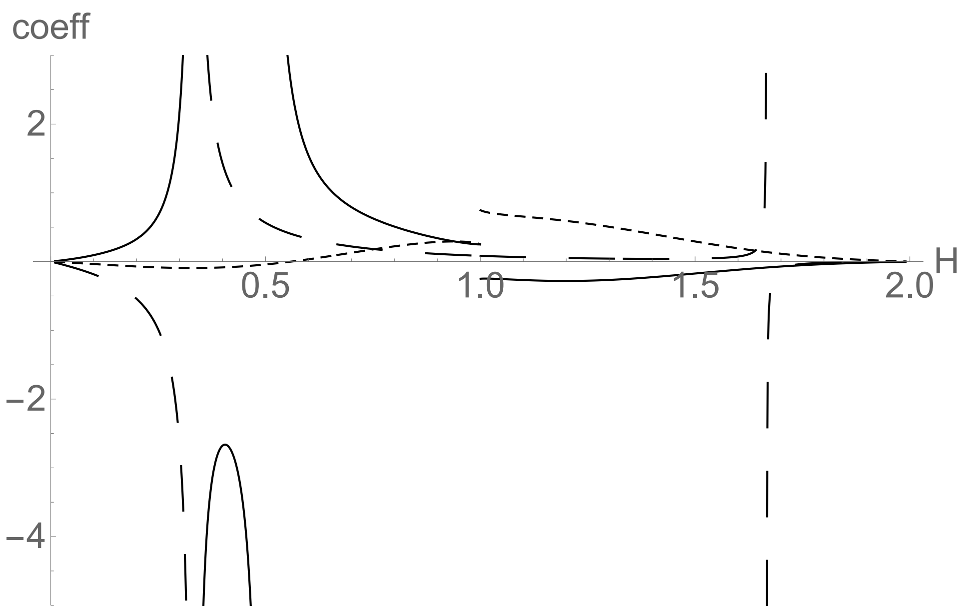

Fig. A2: The solid line is the constant term F_1, the long dashes are the coefficients $\frac{D_3}{2}\Gamma(1-3H)$ of the fractional power, the short dashes are the coefficients of the linear term: $F_2 = C_2 + \frac{D_2}{2H-2} + \frac{D_3}{2(3H-2)}$.

We can see that the contribution of the linear term (used in the $R_{H,2,2}(t)$ approximation) for H<1 and t<1 is fairly small; whereas for 1<H<2, it is larger and the $R_{H,2,2}(t)$ approximation is significantly better than the $R_{H,2,1}(t)$ approximation (see fig. B3).

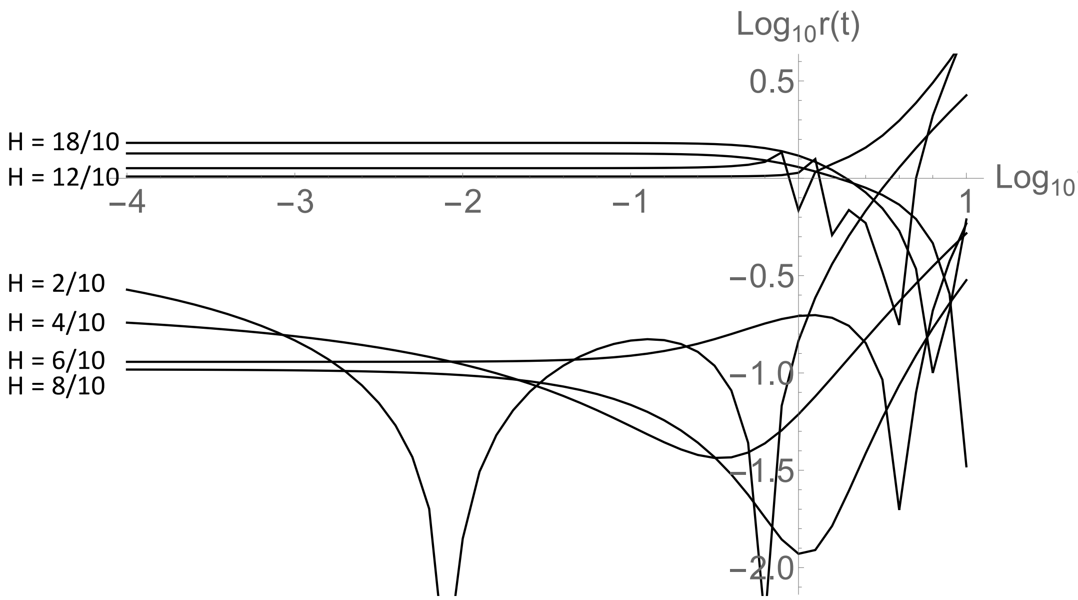

Fig. A3: This shows the logarithm of the relative error in the (2,1) approximation with respect to the deviation from the fGn R_H(t) ($r=\log_{10}\left|1-\left(R_H^{fGn}(t)-R_{H,2,1}(t)\right)/\left(R_H^{fGn}(t)-R_H(t)\right)\right|$). For H<1, t<0 it is of the order ≈ 30% whereas for H>1, it of the order 100%. The H = 1 (exponential) curve is not shown although when t<0 the error is of order 60%.

Appendix B: The H=1/2 special case:

When $H = 1/2$, the high frequency fGn limit is an exact "1/f noise", (spectrum ω^{-1}) it has both high and low frequency divergences. The high frequency divergence can be tamed by averaging, but the not the low frequency divergence, so that fGn is only defined for $H<1/2$. However, for the fRn, the low frequencies are convergent over the whole range $0 < H < 2$, and for $H = 1/2$ we find that the correlation function has a logarithmic dependence at both small and large scales. This is associated with particularly slow transitions from high to low frequency behaviours. The critical value $H = 1/2$ corresponds to the HEBE that was recently proposed [*Lovejoy*, 2020a; b]where it was shown that the value $H = 1/2$ could be derived analytically from the classical Budyko-Sellers energy balance equation.

For fRn, it is possible to obtain exact analytic expressions for R_H, V_H and the Haar fluctuations; we develop these in this appendix, for some early results, see [*Mainardi and Pironi*, 1996]. For simplicity, we assume the normalization $N_H = 1$.

The starting point is the expression:

$$E_{1/2,1/2}(-z) = \frac{1}{\sqrt{\pi}} - z e^{z^2} erfc(z) \qquad erfc(z) = \frac{2}{\sqrt{\pi}} \int_z^\infty e^{-s^2}\, ds$$
$$E_{1/2,3/2}(-z) = \frac{1 - e^{z^2} erfc(z)}{z} \quad , \tag{B1}$$

(e.g. [*Podlubny*, 1999]). From this, we obtain the impulse and step Green's functions:

$$G_{0,1/2}(t) = \frac{1}{\sqrt{\pi t}} - e^t erfc\left(t^{1/2}\right)$$
$$G_{1,1/2}(t) = 1 - e^t erfc\left(t^{1/2}\right) \quad , \tag{B2}$$

(see eq. 16). The impulse response $G_{0,H}(t)$ can be written as a Laplace transform:

$$G_{0,1/2}(t) = \frac{1}{\pi} \int_0^\infty \frac{\sqrt{p}}{1+p} e^{-tp}\, dp \quad . \tag{B3}$$

Therefore, the correlation function is:

$$R_{1/2}(t) = \int_0^\infty G_{0,1/2}(t+s)\, G_{0,1/2}(s)\, ds = \frac{1}{\pi^2} \int_0^\infty ds e^{-s(p+q)} \int_0^\infty \int_0^\infty \frac{\sqrt{qp}}{(1+p)(1+q)} e^{-qt}\, dp\, dq \tag{B4}$$

Performing the s and p integrals we have:

$$R_{1/2}(t) = \frac{1}{2\pi} \int_0^\infty \left[\frac{1}{(1+q)} + \frac{\sqrt{q}}{(1+q)} - \frac{1}{\left(1+\sqrt{q}\right)} \right] e^{-qt}\, dq \quad . \tag{B5}$$

Finally, this Laplace transform yields:

$$R_{1/2}(t) = \frac{1}{2}\left(e^{-t} erfi\sqrt{t} - e^t erfc\sqrt{t}\right) - \frac{1}{2\pi}\left(e^t Ei(-t) + e^{-t} Ei(t)\right) \quad , \tag{B6}$$

where:

$$Ei(z) = -\int_{-z}^{\infty} e^{-u} \frac{du}{u}, \tag{B7}$$

and:

$$erfi(z) = -i\left(erf(iz)\right); \quad erf(z) = \frac{2}{\sqrt{\pi}} \int_0^z e^{-s^2} ds. \tag{B8}$$

To obtain the corresponding V_H use:

$$V_{1/2}(t) = 2\int_0^t \left(\int_0^s R_{1/2}(p)\, dp \right) ds. \tag{B9}$$

The exact $V_{1/2}(t)$ is:

$$V_{1/2}(t) = G_{3,4}^{2,2}\left[t \middle| \begin{matrix} 2, & 2, & 5/2 & \\ 2, & 2, & 0, & 5/2 \end{matrix} \right] + \frac{e^t}{\pi}\left(Shi(t) - Chi(t)\right) + \left(e^{-t} erfi\left(\sqrt{t}\right) - e^t erf\left(\sqrt{t}\right)\right)$$
$$+ t\left(1 + \frac{\gamma_E - 1}{\pi}\right) - 4\sqrt{\frac{t}{\pi}} + \frac{(1+t)\log t}{\pi} + 1 + \frac{\gamma_E}{\pi} \tag{B10}$$

where $G_{3,4}^{2,2}$ is the MeijrG function, Chi is the CoshIntegral function and Shi is the SinhIntegral function.

We can use these results to obtain small and large t expansions:

$$R_{1/2}(t) = -\left(\frac{2\gamma_E + \pi + 2\log t}{2\pi}\right) + \frac{2\sqrt{t}}{\sqrt{\pi}} - \frac{t}{2} - \left(\frac{3 + 2\gamma_E + \pi + 2\log t}{4\pi}\right)t^2 + O\left(t^{3/2}\right); \quad t << 1 \tag{B11}$$

$$R_{1/2}(t) = \frac{1}{2\sqrt{\pi}} t^{-3/2} - \frac{1}{\pi} t^{-2} + \frac{15}{8\sqrt{\pi}} t^{-7/2} + O\left(t^{-4}\right); \quad t >> 1,$$

where γ_E is Euler's constant = 0.57… and:

$$V_{1/2}(t) = -\frac{t^2 \log t}{\pi} + \frac{191 - 156\gamma_E - 78\pi}{144\pi} + \frac{16}{15\sqrt{\pi}} t^{5/2} - \frac{t^3}{6} - \frac{t^4 \log t}{12\pi} + O\left(t^{3/2}\right); \quad t << 1 \tag{B12}$$

$$V_{1/2}(t) = t + \frac{\pi + 2\gamma_E}{\pi} + \frac{2\log t}{\pi} - \frac{4}{\sqrt{\pi}} t^{1/2} + \frac{1}{\sqrt{\pi}} t^{-1/2} - \frac{2}{\pi} t^{-2} + \frac{15}{4\sqrt{\pi}} t^{-3/2} + O\left(t^{-4}\right); \quad t >> 1.$$

We can also work out the variance of the Haar fluctuations:

$$\left\langle \Delta U_{1/2}^2(\Delta t) \right\rangle = \frac{\Delta t^2 \log \Delta t}{4\pi} + \frac{6\pi + 12\gamma_E - \log 16 + 960 \log 2}{240\pi} + \frac{512\left(\sqrt{2} - 2\right)}{240\sqrt{\pi}} \Delta t^{1/2} + \frac{\Delta t}{3} + O\left(\Delta t^{3/2}\right); \quad \Delta t << 1 \tag{B13}$$

$$\left\langle \Delta U_{1/2}^2(\Delta t) \right\rangle = 4\Delta t^{-1} - \frac{32\sqrt{2}}{\sqrt{\pi}} \Delta t^{-3/2} + \frac{3t^{-2} \log \Delta t}{\pi} + O\left(\Delta t^{-2}\right); \quad \Delta t >> 1.$$

Figure B1 shows numerical results for the fRn with $H = ½$, the transition between small and large t behaviour is extremely slow; the 9 orders of magnitude depicted in the figure are barely enough. The extreme low $(R_{1/2})^{1/2}$ (dashed) asymptotes at the left to a slope zero (a square root logarithmic limit, eq. B11), and to a -3/4 slope at the right. The RMS Haar fluctuation (black) changes slope from 0 to -1/2 (left to right). This is shown more clearly in fig. B2 that shows the logarithmic derivative of the RMS Haar (black) compared to a regression estimate over two orders of magnitude in scale (blue; a factor 10 smaller and 10 larger than the indicated scale was used). This figure underlines the gradualness of the transition from $H = 0$ to $H = -1/2$. If empirical data were available only over a factor of 100 in scale, depending on where this scale was with respect to the relaxation time scale (unity in the plot), the RMS Haar fluctuations could have any slope in the range 0 to -1/2 with only small deviations.

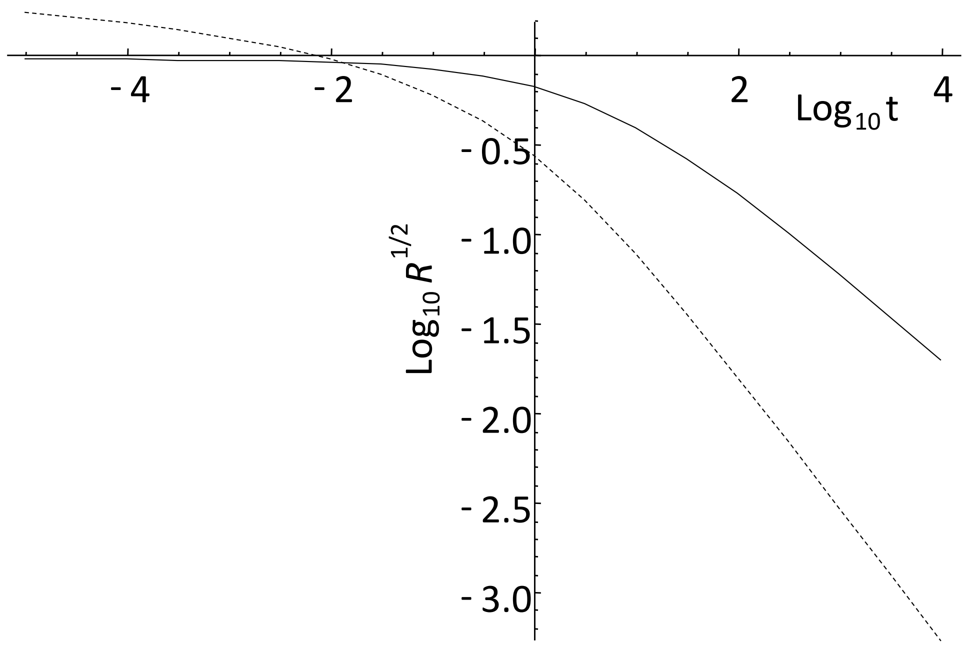

Fig. B1: fRn statistics for $H = 1/2$: the solid line is the RMS Haar fluctuation, the dashed line is the root correlation function $(R_{1/2})^{1/2}$ (the normalization constant = 1, it has a logarithmic divergence at small t).

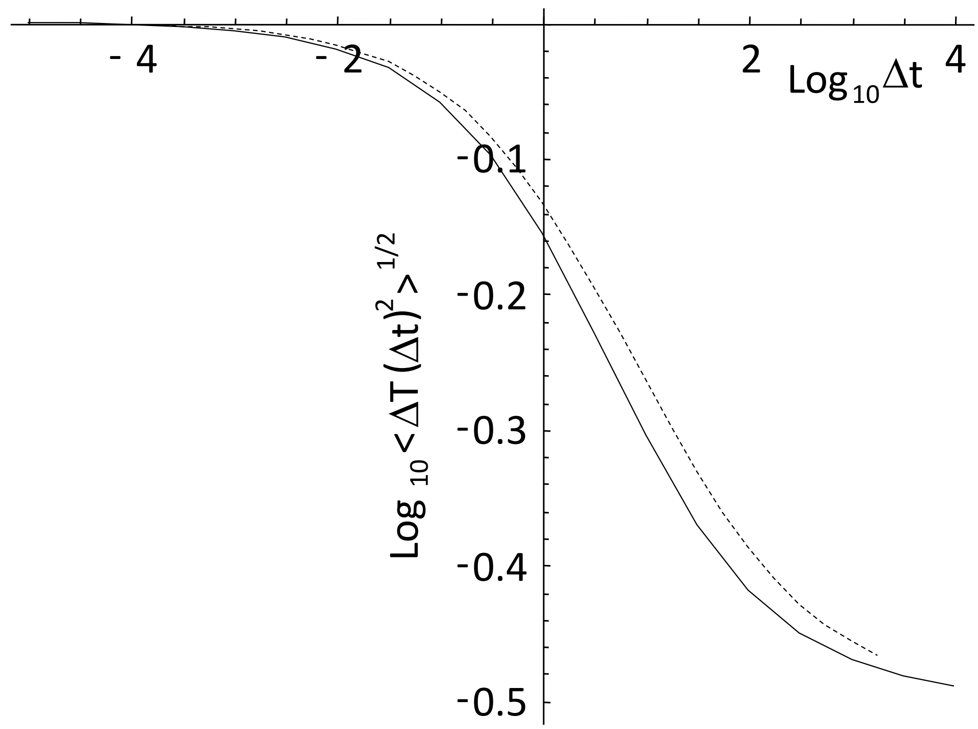

Fig. B2: The logarithmic derivative of the RMS Haar fluctuations (solid) in fig. B1 compared to a regression estimate over two orders of magnitude in scale (dashed; a factor 10 smaller and 10 larger than the indicated scale was used). This plot underlines the gradualness of the transition from $H = 0$ to $H = -1/2$: over range of 100 or so in scale there is approximate scaling but with exponents that depend on the range of scales covered by the data. If data were available only over a factor of 100 in scale, the RMS Haar fluctuations could have any slope in the fGn range 0 to -1/2 with only small deviations.

References:

Atanackovic, M., Pilipovic, S., Stankovic, B., and Zorica, D., *Fractional Calculus with applications in mechanics: variations and diffusion processes*, 313 pp., Wiley, 2014.

Bender, C. M., and Orszag, S. A., *Advanced mathematical methods for scientists and engineers*, Mc Graw Hill, 1978.

Biagini, F., Hu, Y., Øksendal, B., and Zhang, T., *Stochastic Calculus for Fractional Brownian Motion and Applications*, Springer-Verlag, 2008.

Budyko, M. I., The effect of solar radiation variations on the climate of the earth, *Tellus*, *21*, 611-619, 1969.

Buizza, R., Miller, M., and Palmer, T. N., Stochastic representation of model uncertainties in the ECMWF Ensemble Prediction System, *Q. J. Roy. Meteor. Soc.*, *125*, 2887–2908, 1999.

Chekroun, M. D., Simonnet, E., and Ghil, M., Stochastic Climate Dynamics: Random Attractors and Time-dependent Invariant Measures *Physica D*, *240*, 1685-1700 2010.

Coffey, W. T., Kalmykov, Y. P., and Titov, S. V., Characteristic times of anomalous diffusion in a potential, in *Fractional Dynamics: Recent Advances*, edited by J. Klafter, S. Lim and R. Metzler, pp. 51-76, World Scientific, 2012.

Del Rio Amador, L., and Lovejoy, S., Predicting the global temperature with the Stochastic Seasonal to Interannual Prediction System (StocSIPS) *Clim. Dyn.* doi: org/10.1007/s00382-019-04791-4., 2019.

Del Rio Amador, L., and Lovejoy, S., Using scaling for seasonal global surface temperature forecasts: StocSIPS *Clim. Dyn.*, *under review*, 2020.

Dijkstra, H., *Nonlinear Climate Dynamics*, 357 pp., Cambridge University Press, 2013.

Franzke, C., and O'Kane, T. (Eds.), *Nonlinear and Stochastic Climate Dynamics*, Cambridge University Press, Cambridge, 2017.

Gripenberg, G., and Norros, I., On the Prediction of Fractional Brownian Motion *J. Appl. Prob.*, *33*, 400-410, 1996.

Hasselmann, K., Stochastic Climate models, part I: Theory, *Tellus*, *28*, 473-485, 1976.

Hebert, R. (2017), A Scaling Model for the Forced Climate Variability in the Anthropocene, MSc thesis, McGill University, Montreal.

Hébert, R., Lovejoy, S., and Tremblay, B., An Observation-based Scaling Model for Climate Sensitivity Estimates and Global Projections to 2100, *Climate Dynamics*, *(in press)*, 2020.

Herrmann, R., *Fractional Calculus: an Introduction for Physicists*, World Scientific, 2011.

Hilfer, R. (Ed.), *Applications of Fractional Calculus in Physics* World Scientific, 2000.

Hipel, K. W., and McLeod, A. I., *Time series modelling of water resources and environmental systems*, 1st edn. ed., Elsevier, 1994.

Hurst, H. E., Long-term storage capacity of reservoirs, *Transactions of the American Society of Civil Engineers*, *116*, 770-808, 1951.

IPCC, *Climate Change 2013: The Physical Science Basis. Contribution of Working Group I to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change.*, Cambridge University Press: Cambridge, 2013.

Jumarie, G., Stochastic differential equations with fractional Brownian motion inputs, *Int. J. Systems. Sci.*, *24*, 1113, 1993.

Karczewska, A., and Lizama, C., Solutions to stochastic fractional relaxation equations, *Phys. Scr.*, *T136* 7pp doi: 10.1088/0031-8949/2009/T136/014030 2009.

Kobelev, V., and Romanov, E., Fractional Langevin Equation to Describe Anomalous Diffusion *Prog. of Theor. Physics Supp.*, *139*, 470-476, 2000.

Kou, S. C., and Sunney Xie, X., Generalized Langevin Equation with Fractional Gaussian Noise: Subdiffusion within a Single Protein Molecule, *Phys. Rev. Lett.*, *93*, 4 doi: 10.1103/PhysRevLett.93.180603, 2004.

Lovejoy, S., Using scaling for macroweather forecasting including the pause, *Geophys. Res. Lett.*, *42*, 7148–7155 doi: DOI: 10.1002/2015GL065665, 2015.

Lovejoy, S., The spectra, intermittency and extremes of weather, macroweather and climate, *Nature Scientific Reports*, *8*, 1-13 doi: 10.1038/s41598-018-30829-4, 2018.

Lovejoy, S., *Weather, Macroweather and Climate: our random yet predictable atmosphere*, 334 pp., Oxford U. Press, 2019.

Lovejoy, S., The Half-order Energy Balance Equation, Part 2: The inhomogeneous HEBE and 2D energy balance models, *Earth Sys. Dyn. Disc.* doi: https://doi.org/10.5194/esd-2020-13, 2020a.

Lovejoy, S., The Half-order Energy Balance Equation, Part 1: The homogeneous HEBE and long memories, *Earth Syst .Dyn. Disc.* doi: https://doi.org/10.5194/esd-2020-12, 2020b.

Lovejoy, S., and Schertzer, D., *The Weather and Climate: Emergent Laws and Multifractal Cascades*, 496 pp., Cambridge University Press, 2013.

Lovejoy, S., Del Rio Amador, L., and Hébert, R., Harnessing butterflies: theory and practice of the Stochastic Seasonal to Interannual Prediction System (StocSIPS), , in *Nonlinear Advances in Geosciences,* , edited by A. A. Tsonis, pp. 305-355, Springer Nature, 2017.

Lovejoy, S., del Rio Amador, L., and Hébert, R., The ScaLIng Macroweather Model (SLIMM): using scaling to forecast global-scale macroweather from months to Decades, *Earth Syst. Dynam.*, *6*, 1–22 doi: www.earth-syst-dynam.net/6/1/2015/, doi:10.5194/esd-6-1-2015, 2015.

Lovejoy, S., Procyk, R., Hébert, R., and del Rio Amador, L., The Fractional Energy Balance Equation, *Quart. J. Roy. Met. Soc.* , *(under revision)*, 2020.

Lutz, E., Fractional Langevin equation, *Physical Review E*, *64*, 4 doi: 10.1103/PhysRevE.64.051106, 2001.

Mainardi, F., and Pironi, P., The Fractional Langevin Equation: Brownian Motion Revisited, *Extracta Mathematicae 10* 140—154, 1996.

Mandelbrot, B. B., *The Fractal Geometry of Nature*, Freeman, 1982.

Mandelbrot, B. B., and Van Ness, J. W., Fractional Brownian motions, fractional noises and applications, *SIAM Review*, *10*, 422-450, 1968.

Mandelbrot, B. B., and Wallis, J. R., Computer Experiments with fractional gaussian noises: part 3, mathematical appendix, *Water Resour Res*, *5*, 260–267 doi: org/10.1029/WR005 i001p 00260, 1969.

Metzler, R., and Klafter, J., The Random Walks Guide To Anomalous Diffusion: A Fractional Dynamics Approach, *Physics Reports*, *339*, 1-77, 2000.

Newman, M., An Empirical Benchmark for Decadal Forecasts of Global Surface Temperature Anomalies, *J. of Clim.* , *26*, 5260-5269 doi: DOI: 10.1175/JCLI-D-12-00590.1, 2013.

Nonnenmacher, T. F., and Metzler, R., Applications of fractional calculus techniques to problems in biophysics, in *Fractional Calculus in physics*, edited by R. Hilfer, pp. 377-427, World Scientific, 2000.

Palma, W., *Long-memory time series*, Wiley, 2007.

Palmer, T. N., and Williams, P. (Eds.), *Stochastic physics and Climate models*, 480 pp., Cambridge University PRess, Cambridge, 2010.

Papoulis, A., *Probability, Random Variables and Stochastic Processes*, Mc Graw Hill, 1965.

Penland, C., A stochastic model of IndoPacific sea surface temperature anomalies *Physica D*, *98*, 534-558, 1996.

Penland, C., and Magorian, T., Prediction of Nino 3 sea surface temperatures using linear inverse modeling, *J. Climate*, *6*, 1067–1076, 1993.

Podlubny, I., *Fractional Differential Equations*, 340 pp., Academic Press, 1999.

Procyk, R., Lovejoy, S., and Hébert, R., The Fractional Energy Balance Equation for Climate projections through 2100, *Earth Sys. Dyn. Disc.*, *under review* doi: org/10.5194/esd-2020-48 2020.

Rypdal, K., Global temperature response to radiative forcing: Solar cycle versus volcanic eruptions, *J. Geophys. Res.*, *117*, D06115 doi: 10.1029/2011JD017283, 2012.

Sardeshmukh, P., Compo, G. P., and Penland, C., Changes in probability assoicated with El Nino, *J. Climate*, *13*, 4268-4286, 2000.

Sardeshmukh, P. D., and Sura, P., Reconciling non-gaussian climate statistics with linear dynamics, *J. of Climate*, *22*, 1193-1207, 2009.

Schertzer, D., and Lovejoy, S., Physical modeling and Analysis of Rain and Clouds by Anisotropic Scaling of Multiplicative Processes, *Journal of Geophysical Research*, *92*, 9693-9714, 1987.

Schertzer, D., LarchevÍque, M., Duan, J., Yanovsky, V. V., and Lovejoy, S., Fractional Fokker-Planck equation for nonlinear stochastic differential equation driven by non-Gaussian Levy stable noises, *J. of Math. Physics*, *42*, 200-212, 2001.

Schiessel, H., Friedrich, C., and Blumen, A., Applications to problems in polymer physics and rheology, in *Fractional Calculus in physics*, edited by R. Hilfer, pp. 331-376, World Scientific, 2000.

Sellers, W. D., A global climate model based on the energy balance of the earth-atmosphere system, *J. Appl. Meteorol.*, *8*, 392-400, 1969.

van Hateren, J. H., A fractal climate response function can simulate global average temperature trends of the modern era and the past millennium, *Clim. Dyn.* , *40*, 2651 doi: https://doi.org/10.1007/s00382-012-1375-3, 2013.

Vojta, T., Skinner, S., and Metzler, R., Probability density of the fractional Langevin equation with reflecting walls, *Phys. Rev. E 100,* , 042142 doi: 10.1103/PhysRevE.100.042142, 2019.

Watkins, N., Fractional Stochastic Models for Heavy Tailed, and Long-Range Dependent, Fluctuations in Physical Systems, in *Nonlinear and Stochastic Climate Dynamics*, edited by C. Franzke and O. K. T., pp. 340-368, Cambridge University Press, 2017.

Watkins, N., chapman, s., Klages, R., Chechkin, A., Ford, I., and stainforth, d., Generalised Langevin Equations and the Climate Response Problem, *Earth and Space Science Open Archive* doi: doi:10.1002/essoar.10501367.1, 2019.

Watkins, N. W., Chapman, S. C., Chechkin, A., Ford, I., Klages, R., and Stainforth, D. A., On Generalized Langevin Dynamics and the Modelling of Global Mean Temperature, *arXiv:2007.06464v1*, *[cond-mat.stat-mech]* 2020.

West, B. J., Bologna, M., and Grigolini, P., *Physics of Fractal Operators*, 354 pp., Springer, 2003.