The Fractional Energy Balance Equation

Shaun Lovejoy¹, Roman Procyk¹, Raphael Hébert¹-², Lenin del Rio Amador¹

¹ Physics, McGill University
3600 University st.
Montreal, Que. H3A 2T8
Canada

² Alfred-Wegener-Institut,
Helmholtz Centre for Polar and Marine Research
Telegrafenberg A45-2N 104 D-14473, Potsdam
Germany

Abstract

Energy conservation is a useful symmetry for understanding the climate: Energy Balance Models (EBMs) of Budyko-Sellers or box type. Extending these to time varying anomalies involves energy storage; possible mechanisms can be constrained by a second symmetry: scale invariance. Temporal storage scaling scale follows from the spatial scale invariance of the dynamics. Physically, it corresponds to a hierarchy of storage processes.

The classical Energy Balance Equations (EBEs) are differential equations of integer order ($H = 1$), here we generalize this to fractional orders: the Fractional EBE (FEBE, $0< H \leq 1$). In the FEBE, when the Earth is perturbed by a forcing, it relaxes to thermodynamic equilibrium via a slow power law process: $H = 1$ is the exceptional exponential case. Our FEBE derivation is phenomenological, it complements derivations based on the classical continuum mechanics heat equation (that imply $H = 1/2$ for the surface temperature) and of the more general Fractional Heat Equation which allows for $0 < H < 1$.

We solve the FEBE using Green’s functions, whose high and low frequency limits are power laws with a transition at the relaxation scale (several years). When stochastically forced the high frequency part of the internal variability are fractional Gaussian noises; when deterministically forced, the low frequency responses have been used for climate projections: the FEBE is linked to the Fluctuation Dissipation Theorem.

The FEBE introduces complex climate sensitivities that are convenient for handling periodic (especially annual) forcing. The FEBE obeys Newton’s law of Cooling, but the heat flux crossing a surface nonetheless depends on the fractional time derivative of temperature. The FEBE transient to equilibrium climate sensitivity ratio is compatible with GCM estimates. A simple ramp forcing model of the industrial epoch warming combining deterministic (external forcing) with stochastic (internal forcing), is statistically validated against centennial scale temperature series.
**Key Words:** Energy balance, fractional equations, scaling, climate response functions, internal variability, externally forced variability.

1. Introduction

The earth is a heterogeneous system with a dynamically evolving atmosphere, ocean and land surface, it has a huge number of degrees of freedom. While GCMs are routinely used to model its evolution, these are highly complex and each GCM has its own climate, and these are different from the real world climate. Simplified models have therefore been developed that exploit what at first sight appears to be a fairly simple yet fundamental constraint: the conservation of energy.

The corresponding Energy Balance Models (EBMs) are based on the (near) thermodynamic equilibrium of the earth with the sun and outer space. For example, anthropogenic forcings are currently around 2.4 W/m$^2$ above the long term mean solar forcing of $\approx$238 W/m$^2$ (both figures are for global averages), so that the perturbations away from equilibrium are of the order of 1%. This makes it plausible that the responses to external forcings are roughly linear. This was confirmed by examining the evolution of the spatial patterns of 32 CMIP5 GCM responses when forced by the various Representative Carbon Pathway (RCP) scenarios, [Hébert and Lovejoy, 2018]. Even when these included forcing are as high as 8.5W/m$^2$, i.e. 3.5% of the reference mean, the response (its regional distribution) of each model – although different from the other models - was (nearly) a linear projection of its own past. The extent of the linear regime is unclear, and it is certain that due to albedo-temperature feedbacks and other nonlinearities, it will eventually break down. Consequently, strongly nonlinear (chaotic) models have been proposed, see [Dijkstra, 2013] for a review. These are based on [Hasselmann, 1976]'s rather general and now classical mathematical decomposition of the dynamics into slow/fast components where the fast (weather) scales stochastically drive the slow (climate) component (see section 4 below).

Two different linear EBM modelling approaches have evolved, the earliest going back to [Budyko, 1969] and [Sellers, 1969] who applied the continuum mechanics heat equation. This yielded 1D (longitudinally averaged) models of the equilibrium surface temperature' latitudinal distribution, later extended to time [Dwyers and Petersen, 1975]. An attractive feature of the Budyko-Sellers models is that the system can become nonlinear if the albedo (and hence forcing) is coupled with the temperature (e.g. via the position of the snow line). This has enabled it to be used for modelling past and (possible) future climates.
A second EBM strand was inspired by “box models” that related CO₂ emissions to atmospheric concentrations. In energy balance box models, one or more boxes are used to model the regional or - more usually - the globally averaged temperature. The simplest such “zero-dimensional” box models, were first proposed by [Hasselmann et al., 1993] and involve a single uniform slab of material - the “box” – representing the Earth and that interacted radiatively with outer space according to Newton’s law of cooling (NLC). Although the Budyko-Sellers and box approaches are distinct, when the time-dependent Budyko-Sellers models are reduced to zero-dimensions and linearized, they are mathematically equivalent to a (single) box model (for a review see [North et al., 1981]).

A key difficulty for all EBMs is that some of the energy received from the sun today is stored and emitted to outer space only at a later time. If we consider models of the Earth averaged over (at least) macroweather time scales - i.e. longer than the typical lifetime of atmospheric (weather) structures [Lovejoy, 2013] - then some of the long wave / short wave radiative flux imbalances are stored in the subsurface (mostly in the oceans) and some are transported horizontally. Conversely, energy that was stored long ago – for example in deep ocean currents – may appear and contribute to today’s long wave emissions. The key to applying the energy balance principle is therefore to have a realistic model of the storage processes.

The box and Budyko-Sellers models handle the storage problem in quite different ways. The original Budyko-Sellers models had no storage at all, the flux imbalance was simply redirected meridionally away from the equator. In the box models, the temperature (T) of each box is spatially uniform and the energy stored (S) is the product of temperature with the box heat capacity (C). This means that energy fluxes into and out of the box instantaneously change the temperature of the entire box. In addition, energy fluxes across box surfaces are proportional to the temperature difference between the box and its surroundings (Newton’s Law of Cooling, NLC).

A recent series of papers showed how the realism of these Budyko-Sellers models could greatly be improved by introducing a vertical coordinate. Even though the aim is to improve the modelling of the horizontal temperature distribution, the inclusion of a vertical coordinate has two advantages. First, it allows us to apply the correct conductive-radiative surface boundary condition to determine the fraction of the radiative imbalance that is conducted into the subsurface and the fraction that is re-emitted at long wavelengths. Second, since the subsurface stores the heat, the third dimension simultaneously provides a precise storage mechanism. When this conductive radiative boundary condition was applied to the classical 3D heat equation (the Budyko-Sellers model with an extra vertical coordinate), a surprising consequence was that the surface temperature obeyed the long memory, Half-ordered Energy Balance Equation, (HEBE, [Lovejoy, 2020a; c], hereafter L1, L2). In the HEBE, the source of the surface temperature long memory is completely classical: it is simply a consequence of the slow diffusion of heat into and out of the subsurface.
Even if we use the standard integer-ordered heat equation, we will be forced to deal with fractional ordered surface temperature equations. But why restrict ourselves to the classical heat equation? Indeed, Budyko-Sellers type models are for macroweather Earth states i.e. they are averaged over a wide range of (weather) degrees of freedom, they are “effective material” type models, there is no compelling reason that they should be integer ordered. [Lovejoy, 2020b] (hereafter L3), therefore proposed that the Fractional Heat Equation (FHE) was a more appropriate macroweather energy balance model. Mathematically, the FHE is a fractional diffusion equation that has been studied in the statistical physics literature.

The classical heat equation is the special case of the FHE where the time derivative is integer ordered. For our present purposes, L3’s most important conclusion was that the equation for the surface temperature was the Fractional EBE or FEBE, a generalization of the HEBE from exponents (order of derivative) $H = 1/2$ to $0<H<1$. Like the HEBE, the FEBE is fundamentally a 2D macroweather temperature anomaly model, whose zero-dimensional version is a fractional relaxation equation, itself a fractional generalization of the classical zero-dimensional box models that are the $H = 1$ special cases. When the FEBE is driven by a Gaussian white noise, the result is fractional Relaxation noise (fRn) that generalizes the classical Ornstein-Uhlenbeck process and its high frequency limit is a fractional Gaussian noise process (fGn) that generalizes Brownian motion. The resulting FEBE can be used advantageously for both macroweather forecasts ([Lovejoy et al., 2015], [Del Rio Amador and Lovejoy, 2019]) and multidecadal projections ([Hebert, 2017], [Lovejoy et al., 2017], Procyk, work in progress).

While L3 derived the FEBE from a three dimensional continuum heat model, in this paper it is alternatively derived phenomenologically by arguing that the basic energy storage mechanisms respect a scaling principle and we treat the consequences of the FEBE for the Earth’s temperature. In section 2, we derive the zero-dimensional FEBE and in section 3, some of its mathematical properties, consequences and solutions. This includes the stochastic internal variability that was examined from a more mathematical point of view in [Lovejoy, 2019b], (hereafter L4). In section 4, we discuss model parameter estimation / validation (high frequency, low frequency, annual periodicity). We also propose a simple ramp and plateau model of the industrial epoch that includes both stochastic internal forcing and deterministic external forcing.

2. Energy Storage and the classical energy balance equation (EBE)

2.1 Multibox models and Climate Response Functions

A single box with its unique relaxation time is a very inflexible model. [Hasselmann et al., 1993; Hasselmann et al., 1997] already noted that it was desirable to go beyond this to use the more general linear response function framework. In this context, the response functions are called “climate response functions” (CRFs) and following [Hasselmann et al., 1993] we have a choice between the equivalent impulse and step CRFs, the former being the derivative of the latter (see below).
Unfortunately, without more assumptions or information, the linear framework with (nearly) arbitrary CRFs is unmanageably general. In order to make progress, [Hasselmann et al., 1997] proposed a response function consisting of a CRF equal to the sum of \( N \) exponentials corresponding to \( N \) boxes that mutually exchange heat, each with its own exponential relaxation time and relative weight. But \( N \) boxes requires \( 2N \) parameters - it is still too general - so that out of the practical necessity of fitting GCM outputs, [Hasselmann et al., 1997] ultimately chose \( N = 3 \). Following the more usual procedure of deriving the impulse responses from integer-ordered linear differential equations (where impulse CRFs are called “Green’s functions”), [Li and Jarvis, 2009] recalled that \( n^{th} \) order differential equations (with constant coefficients and with \( n \) an integer), can quite generally be reduced to sums of exponentials. However, in the application part of their paper, they nevertheless also used the value \( N = 3 \).

Even the \( N = 3 \) multibox model is difficult to manage and most authors have settled for \( N = 1 \) or 2, including the IPCC AR5 (2013, section 8.SM.11.2) that recommends the four parameter \( N = 2 \) model. Yet problems remain: to be taken seriously, the boxes and their time constants must have physical interpretations. For example, the role of each box is to store heat, but what do the boxes physically represent? Using the atmosphere, implies relaxation times of weeks which is too short for most applications. For the more popular two box model, there is a rough consensus that each represents part of the ocean: an upper “mixed layer” and a lower “deep” layer. By appropriately choosing the layer thicknesses, one can get relaxation times \( \tau \) spanning years to millenia. For example, single box models have \( \tau \) ranging from 4 to 40 years (e.g. [Schwartz, 2012], [Zeng and Geil, 2017], [Held et al., 2010]). Two box models typically have one box with \( \tau \) in the range 2 - 10 years and another with a \( \tau \) in the range 20 – 800 years ([Rypdal, 2012], [Geoffroy et al., 2013]) with the IPCC AR5 favouring \( \tau = 8.4 \) and \( \tau = 409.5 \) years.

### 2.2 Scaling storage

Rather than justifying a choice of CRF by invoking hypothetical homogeneous boxes, one can instead derive the CRF from physical considerations. The key is to exploit the wide range spatial scale invariance of geo-processes such as the ocean and atmospheric dynamics associated with energy storage (see e.g. [Lovejoy and Schertzer, 2013] for a review). The idea is to use the fact that from small to large spatial scales, there is a whole scaling hierarchy of storage processes (e.g. atmospheric or ocean eddies). Since the heat transfer times of each structure in the hierarchy depends on its spatial scale, it is reasonable to assume that the collective overall heat storage is also scaling.

The problem is that up until now, the scaling principle has not been applied to the storage, but instead directly to the overall temperature response ([Rypdal, 2012], [van Hateren, 2013], [Hebert, 2017] [Hebert et al 2020], [Nilsen et al., 2020]). Scaling impulse (Dirac \( \delta \) function) response CRFs are power laws:

\[
G_\delta(t) \propto t^{H-1}
\]  

(1)
where $H_F$ is a scaling exponent. However such pure power law CRFs suffer from the “runaway Green’s function effect”, ([Hébert and Lovejoy, 2015]); the fact that they imply nonphysical divergences either at low or at high frequencies (depending on whether $H_F > 0$ or $< 0$). This can readily be seen by considering the physically important integral of $G_\Theta$ the step response CRF: $G_\Theta(t) \propto t^{H_F}$. For example [Rypdal, 2012] proposed a CRF with $H_F > 0$ which implies that $G_\Theta(t)$ diverges as $t \to \infty$: such CRFs would yield infinite temperature responses to a doubling in CO$_2$: an infinite Equilibrium Climate Sensitivity (ECS). If forcings are carefully restricted so as to eventually return to zero, then the divergence can be avoided, but this stringent requirement is obviously unsatisfactory. [van Hateren, 2013] instead proposed a “fractal climate response” model that avoided divergences by using $N = 6$ boxes but used the scaling principle to link the amplitudes and time constants by power laws resulting in a four parameter model. As mentioned in [Hébert et al., 2020], his model featured unphysical logarithmic oscillations and an unnecessary low frequency cutoff.

Recently, [Myrvoll-Nilsen et al., 2020] have applied the original [Rypdal, 2012] infinite ECS model to 21st century warming scenarios claiming that their model with “not well-defined” ECS nevertheless “provides an accurate description of both forced and unforced surface temperature fluctuations”. Be that as it may, their further model justification as “a cost of the reduction of model complexity” cannot be sustained. For example, [Hebert, 2017; Hébert et al., 2020] proposed a scaling CRF but instead took $H_F < 0$ combined with an explicit high frequency cutoff to avoid divergences:

$$G_\Theta(t) = 1 - \left(1 + \frac{t}{\tau}\right)^{H_F}; \quad H_F < 0$$

At long times $t \gg \tau$, this step response CRF yields a power law approach to a constant (corresponding to thermodynamic equilibrium), and $G_\Theta(0) = 0$ so that there is no divergence at $t = 0$ either. The truncated scaling model (eq. 2) or the FEBE (below) both show that it is easy to avoid models with “not well-defined” ECS while continuing to maintain model simplicity.

The fundamental cutoff scale $\tau$ was estimated as $= 2$ years from the empirical coupling time of fluctuations in the ocean and atmosphere ([Hebert, 2017, Lovejoy et al., 2017]). Using historical forcings and temperature responses since 1880, they empirically estimated the exponent $H_F \approx -0.5 \pm 0.2$ as well as the climate sensitivity $\lambda = 2.4 \text{ K/CO}_2$ doubling (with 90% confidence interval [1.8-3.7], [Hébert et al., 2020]). Using the IPCC AR5 scenarios (the Representative Carbon Pathways, RCPs), the parameters were then used to make temperature projections to 2100. Since the source of the projection uncertainties were from the past data - not the large GCM “structural uncertainties” – they showed that the projections were compatible with, but much less uncertain than the CMIP5 alternatives.
In spite of its success in making climate projections, due to their truncations, the scaling CRF models are only approximations: they apply the scaling principle directly to the CRF whereas in fact it should only be applied to the storage term in the energy balance equation (EBE). It turns out that this model change can be made with a surprisingly small tweak to the usual EBE: it suffices to change the order of differential equation from integer to fractional order, the Fractional Energy Balance Equation (FEBE) discussed below. We will show that this seemingly trivial generalization can account for responses to both externally forced and internal variability, effectively explaining them both as different aspects of energy conservation.

2.3 The Earth’s Energy Balance

The earth’s energy balance can be expressed as:

\[ S(t) + E_{\text{tot},\uparrow}(t) = E_{\text{tot},\downarrow}(t) \]  
(3)

where the upward arrow indicates energy emitted to outer space and the downward arrow the energy received from the sun (and other sources external to the climate system); the difference is the energy stored, \( S(t) \); the subscript “tot” indicates the total over a long time, starting at a convenient reference baseline.

Before continuing, it is worth making a few comments:

a) This equation directly expresses conservation of energy; it is not yet the more usual rate equation obtained by taking derivatives (see below).

b) Equation 3 is “zero-dimensional” i.e. all the spatial degrees of freedom are averaged out, \( S \), \( E_{\text{tot},\uparrow} \), \( E_{\text{tot},\downarrow} \) are averages over the whole earth: energies per unit surface area (in standard units, J/m\(^2\)).

Now consider anomalies i.e. the deviations \( E_{\uparrow} \), \( E_{\downarrow} \) from the long term values \( E_{0,\uparrow}, E_{0,\downarrow} \):

\[ E_{\text{tot},\uparrow}(t) = E_{0,\uparrow}(t) + E_{\uparrow}(t) \]
\[ E_{\text{tot},\downarrow}(t) = E_{0,\downarrow}(t) + E_{\downarrow}(t) \]  
(4)

The long term radiances \( E_{0,\uparrow}, E_{0,\downarrow} \) could be taken as being proportional to long term average rates, outgoing and incoming radiances \( R_{0,\uparrow}, R_{0,\downarrow} \) so that \( E_{0,\uparrow}(t) = tR_{0,\uparrow} \)

\( E_{0,\downarrow}(t) = tR_{0,\downarrow} \).

Since the earth is nearly in thermodynamic equilibrium, \( R_{0,\uparrow} = R_{0,\downarrow} \) hence \( E_{0,\uparrow} = E_{0,\downarrow} \) so that the storage \( S(t) \) is the difference between the long and short wave radiative anomalies:

\[ S(t) + E_{\uparrow}(t) = E_{\downarrow}(t) \]  
(5)

(subtract the two equations in 4 and use eq. 3).
If the storage processes are scaling in space (i.e. hierarchical), and the characteristic storage time of each structure depends on its spatial extent in a power law manner, then it is reasonable to assume that the storage at a time \( t \) depends on the past in a power law manner:

\[
S_H(t) = \frac{C_H}{\Gamma(1-H)} \int_{t_0}^{t} T(s) \left( \frac{t-s}{\tau} \right)^{-H} \frac{ds}{\tau}; \quad 0 \leq H \leq 1
\]

(6)

where \( C_H \) is the (generalized) heat capacity per unit area of the earth, which in standard units is \( [C_H] = \frac{J}{Km^2} \), \( H \) is a fundamental scaling exponent and \( \tau \) is the (power law) relaxation time (see below). The power law kernel in eq. 6 relates the storage to the temperature via a Riemann-Liouville fractional integral of order \( 1-H \) with \( 0 \leq H \leq 1 \); \( \Gamma \) is the usual Gamma function. \( t_0 \) is the reference time from which the total storage is measured (here, taken to be in the distant past): \( S_H(t_0) = 0 \), later we will take \( t_0 = -\infty \). We have added the subscript \( H \) to emphasize that except in the classical \( H = 1 \) case, \( C_H \) is not a usual heat capacity.

The storage defined in eq. 6 is a generalization of the standard "box model" storage to which it reduces when \( H = 1 \). To see this, we integrate by parts and take \( T(t_0) = 0 \), this is equivalent to fixing the reference of our anomalies:

\[
S_H(t) = \frac{C_H \tau^{H-1}}{\Gamma(2-H)} \int_{t_0}^{t} T'(s)(t-s)^{1-H} ds; \quad 0 \leq H \leq 1
\]

(7)

Taking \( H = 1 \), we now obtain the usual \( H = 1 \) box storage result:

\[
S_1(t) = C_1 T(t); \quad H = 1
\]

(8)

A key difference between the \( H < 1 \) and \( H = 1 \) case is that only in the latter (box model) does the storage depend instantaneously on the temperature.

3. The fractional energy balance equation:

3.1 Fractional Storage

The usual energy balance equation is an equation for the rates, it is obtained by differentiation of eq. 5:

\[
\frac{dS_H(t)}{dt} + \lambda^{-1}T(t) = F(t); \quad F(t) = \frac{dE_\downarrow}{dt}; \quad \lambda^{-1}T(t) = \frac{dE_\uparrow}{dt}
\]

(9)

where \( F \) is the rate of anomalous energy input - the usual forcing (power per area) and \( \lambda \) is the climate sensitivity and \( T \) the temperature anomaly with respect to the long term equilibrium temperature.

The sensitivity is the temperature rise per forcing; in standard units, \( [\lambda] = \frac{Km^2}{W} \). When combined with the generalized heat capacity, it defines a time scale:
\[ \tau = \lambda C_H \]  

Since this dimensional combination is unique, \( \tau \) is the characteristic time for the system to relax to equilibrium when it is perturbed by a step function forcing. However, when \( H \neq 1 \), \( C_H \) does not correspond to our usual idea of a specific heat; it is really an average coefficient that expresses the collective heat transfer and storage characteristics of a hierarchy of storage mechanisms. Rather than treating \( C_H \) as fundamental and then using eq. 10 to determine the relaxation time, it is better to treat \( \tau \) as fundamental and \( C_H \) as a derived quantity: \( C_H = \tau / \lambda \), so that the storage can be written:

\[
S_H(t) = \frac{\tau^H}{\lambda \Gamma(1-H)} \int_{t_0}^{t} T(s)(t-s)^{-H} ds; \quad 0 \leq H \leq 1
\]  

In this model, beyond the dimensionless exponent \( H \), \( \tau \) is the fundamental physical quantity characterizing the storage processes.

We now recall the Riemann-Liouville derivative of order \( H \) that is defined as the ordinary derivative of the \( 1-H \) order fractional integral:

\[
t_0^t D_t^H T = \frac{d}{dt} \left( t_0^s D_s^{H-1} T \right) = \frac{d}{dt} \left[ \frac{1}{\Gamma(1-H)} \int_{t_0}^{t} (t-s)^{-H} T(s) ds \right] \quad 0 \leq H \leq 1
\]  

\[
= \frac{1}{\Gamma(1-H)} \int_{t_0}^{t} (t-s)^{-H} T'(s) ds;
\]

Where the bottom equality follows by integrating by parts and using \( T(t_0) = 0 \), the reference temperature. Therefore:

\[
\frac{dS_H}{dt} = \lambda^{-1} \tau^H \int_{t_0}^{\cdot} D_t^H T; \quad 0 \leq H \leq 1
\]  

Putting this into eq. 9 and taking \( t_0 \rightarrow -\infty \) we obtain:

\[
\tau^H \rightarrow \lambda D_t^H T + T = \lambda F
\]

This is the fractional energy balance equation, FEBE. The physical meaning of the fractional derivative is that when the forcing changes, the change in temperature is not instantaneous but rather depends on the entire past history of temperature changes.
Starting with the continuum mechanics heat equation, L1 derives this with the value $H = \frac{1}{2}$, while L3 derives the full FEBE (with $0 < H < 1$) from the fractional heat equation, itself a generalization of the continuum mechanics heat equation. The $t_0 = -\infty$ limit is convenient when the forcing is a stationary stochastic process (such as a white noise representing the internal variability) since the temperature response will then also be statistically stationary. Alternatively, it is also convenient if the forcing is periodic with mean zero. In the case of a deterministic forcing, that starts at a finite time (e.g. $t = 0$), we need only assume that $T(t) = 0, F(t) = 0$ for $t \leq 0$. Due to the infinite range of the fractional derivative, the FEBE eq. 16 is a Weyl fractional relaxation equation.

We derived the FEBE for temperatures and forcing that were taken as anomalies with respect to a loosely defined reference state. In fact, since the equation is linear, one can simply add or subtract reference forcings and temperatures for convenience. The same is true for deterministic and stochastic forcings whose solutions can be superposed when needed (section 4.5). If the temperature and forcing are nonlinearly coupled (e.g. via the albedo), then this freedom can be used to derive a nonlinear FEBE that could be used to model temperature/albedo feedbacks as well as the effects of orbital forcings, glacial-interglacial transitions and other low frequency behaviours. In the future, specific models and their stability may be investigated in this way.

**3.2 Newton’s law of Cooling and thermal impedance**

To understand the FEBE, we integrate both sides by order $H$ and rewrite it in the form:

$$T(t) = \int_{-\infty}^{t} \frac{1}{\Gamma(H)} \left( T_{eq}(s) - T(s) \right) \left( \frac{t - s}{\tau} \right)^{H-1} ds; \quad T_{eq}(t) = \lambda F(t)$$

$$0 < H < 1; \quad (15)$$

The temperature is thus the fractional integral of order $H$ of the difference between the temperature and the equilibrium temperature $T_{eq}$. If the forcing is constant at a value $F$, then at long enough times, the temperature will “relax” to the equilibrium value $T_{eq} = \lambda F$. When $H \neq 1$, this occurs in a power law way (due to the power law weight $\left( \frac{t - s}{\tau} \right)^{H-1}$ term). In the special $H = 1$ case – the conventional EBE – the relaxation is instead exponential.

If we interpret the forcing $\lambda F$ in terms of an effective external equilibrium temperature $T_{eq}$, then the FEBE satisfies Newton’s law of cooling (NLC) that states that a body’s rate of heat loss is directly proportional to the difference between its temperature and its environment. In these horizontally homogeneous models, it is the heat flux (energy rate/area $= Q_s$) across the surface into the body (see fig. 1) that is important so that the NLC can be written:

$$Q_s = \frac{1}{Z} (T_{eq} - T) \quad (16)$$
where $Z$ is a transfer coefficient sometimes called the “thermal impedance” (units: m$^2$K/W), its reciprocal $Y$ is the surface “thermal admittance”). Since $Q_s$ equals the rate of stored heat loss, we can rewrite eq. 9 as:

$$Q_s = \frac{dS_H}{dt} = \frac{\tau^H}{\lambda} \int_0^\infty D^H_t T = \frac{1}{\lambda} (T_{eq} - T)$$

(17)

By comparing this with eq. 16, we conclude a) that the FEBE satisfies the NLC, b) that the climate sensitivity is a thermal impedance.

**Newton’s law of cooling**

$$Q_s = \frac{1}{Z} (T_{eq} - T_s)$$

**Fig. 1:** A schematic showing Newton’s law of cooling (NLC) that relates the temperature difference across a surface to the heat flux crossing the surface, $Q_s$ (into the surface). $T_{eq}$ is the fixed outside temperature, heat will flow across the surface as long as the surface temperature $T_s \neq T_{eq}$. $Z$ is the thermal impedance, equal here to the climate sensitivity $\lambda$. To apply the NLC, we need to relate the heat flux to the surface temperature. The lower left shows the consequence of applying heat equation with conductive – radiative BC’s, the lower right shows the phenomenological assumption made by box models. The arrows represent heat fluxes, hence the factor $\lambda$ in the denominators. The system is assumed to be horizontally homogeneous.

### 3.2 Solving the FEBE using Green’s functions

Mathematically when $0 < H < 1$, the FEBE is a “fractional relaxation equation”, the extension to $1 < H \leq 2$, is a “fractional oscillation equation”. In the initial value problem, $F(t) = 0$ for $t \leq t_0$, the FEBE impulse Green’s function - the impulse CRF - is given in terms of Mittag-Leffler functions:
\[ G_{\zeta, H}(t) = G_{0, H}(t) = \tau^{-1} \left( \frac{t}{\tau} \right)^{H-1} E_{H, \zeta} \left( -\left( \frac{t}{\tau} \right)^H \right); \quad t \geq 0; \quad 0 \leq H \leq 2 \]

\[ G_{0, H}(t) = 0; \quad t < 0; \quad 0 \leq H \leq 2 \]  

(18)

Where:

\[ E_{\alpha, \beta}(z) = \sum_{n=0}^{\infty} \frac{z^n}{\Gamma(\alpha n + \beta)} \]  

(19)

is the \( \alpha, \beta \) order Mittag-Leffler function (these and most of the following results are in the notation of [Podlubny, 1999]). The condition \( G = 0 \) for \( t<0 \) is needed to respect causality. In the following this condition will be understood for all the \( G \) functions. As can be seen directly from the series expansion (eq. 19), the Mittag-Leffler functions are often called “generalized exponentials”. In particular, the classical \( H = 1 \) box model is recovered with help of the (exceptional) ordinary exponential:

\[ E_{1,1}(z) = e^z. \]

As usual, the Green’s function gives us the particular solution to the inhomogeneous equation for a Dirac \( \delta \) forcing. The full solution, including the solution to the homogeneous equation is:

\[ T(t) = \lambda \int_{t_0}^{t} G_{0, H}(t-t') F(t') dt' + T_0 E_{H,1} \left( -\left( \frac{t-t_0}{\tau} \right)^H \right); \quad T_0 = T(t_0); \quad t \geq t_0 \]

(20)

[Cheng and Chu, 2011]. Since \( E_{H,1}(0) = 1 \) we can confirm that this satisfies the initial condition \( T(t_0) = T_0 \).

Just as integer ordered linear differential equations with constant coefficients can generally be solved with exponentials, the analogous fractional ordered differential equations can generally be solved with generalized exponentials. However, the latter are based on power laws so that the usual exponentials correspond to the special case where all the orders of derivatives are restricted to integer values, power law Green’s functions are thus more general.

A convenient feature of Mittag-Leffler functions is that they can be easily integrated by any positive order \( \zeta \):

\[ G_{\zeta, H}(t) = \int_0^{t} D_t^{-\zeta} \left( G_{0, H}(t) \right) = \tau^{-1} \left( \frac{t}{\tau} \right)^{H-1+\zeta} E_{H, \zeta} \left( -\left( \frac{t}{\tau} \right)^H \right); \quad \zeta \geq 0; \quad 0 \leq H \leq 2 \]  

([Podlubny, 1999]). Since \( G_{0, H} \) is the response to a Dirac function forcing, with \( \zeta = 1 \) we obtain the response to the first order integral of the Dirac: the Heaviside (step function) forcing. Similarly, using \( \zeta = 2 \) we obtain the response to its integral, the “ramp” (=0 for \( t<0 \), = \( t \) for \( t\geq0 \), section 4.5), using \( \zeta = 3 \) we have the response to a parabolic forcing etc.:
where \( n \) is an integer. Fig. 2 shows the impulse, step and ramp responses for various \( H \)'s. The log-log plot is particularly illuminating since it shows directly that for \( H<1 \), there are two power law limits and that for the \( H = 1 \) case, that we obtain the usual exponential CRF. Note for the step response (middle graphs, and fig. 3), the approaches to the asymptotic value 1 (left) corresponding to thermodynamic equilibrium can be extremely slow.

The physically important step function response – the step CRF – is thus obtained by taking \( \zeta = 1 \):

\[
G_{0,1} (t) = G_{1,1} (t) = t \int_0^t G_{1,1} (t') dt' = \left( \frac{t}{\tau} \right)^H \left( - \left( \frac{t}{\tau} \right)^H \right) ; \quad t \geq 0; \quad 0 \leq H \leq 2
\]  

(23)

The \( H = 1 \) (box model) yields:

\[
G_{0,1} (t) = \tau^{-1} e^{-t/\tau}; \quad G_{1,1} (t) = 1 - e^{-t/\tau}
\]  

(24)

where we have used \( E_{1,2} (z) = \frac{e^z - 1}{z} \).

By integrating eq. 20 by parts, and using \( G_i (0) = 0 \), and the condition \( F(t_0) = 0 \), we can express the temperature response in terms of the step CRF rather than the impulse CRF:

\[
T(t) = \lambda \int_{t_0}^t G_{1,1} (t-s) F'(s) ds; \quad F'(t) = \frac{dF}{dt}
\]  

(25)

(we have taken \( T(t_0) = 0 \) for simplicity). Mathematically, the equivalence between eq. 20 and 25 arises because \( \delta(t) \) is the derivative of \( \Theta(t) \) and \( G_{0,1} (t) \) is the derivative of \( G_{1,1} (t) \). Expressing the solution in terms of the step CRF rather than the impulse CRF is advantageous. For example, \( G_{1,1} (t) \) is dimensionless so that \( \lambda \) has the usual dimensions of sensitivity, in addition, as emphasized by [Marshall et al., 2014] and by [Marshall et al., 2017] it directly expresses temperature in terms of the results of a \( \text{CO}_2 \) doubling experiment.
Fig. 2: The various response functions linear - linear (left column) and log - log right for $H = 1/10$ to 1 (exponential) in steps of 1/10. The lines bounding the envelope are $H=1/10$ (dashed), $H=1$ (thick). The middle value ($H=1/2$) is also thick. The top row shows the Dirac ($\delta(t)$) response functions ($G_0$), the middle row the step ($\Theta(t)$) responses ($G_1$), and the bottom row, the ramp ($t\Theta(t)$) responses ($G_2$).
Fig. 3: The FEBE response to a step function forcing (shown in thick black) for various values of $H$, using $\lambda = 1$ with the response for $H = 0.1, 0.3, 0.5, 0.7, 0.9$ (bottom to top). The red is for $H = 0.5$, the value close to the empirical exponent; the blue is for the box model value $H = 1$. The storage is the difference between the response and the forcing; both are shown with arrows for the case $H = 0.5$.

### 3.3 Complex climate sensitivities and the annual cycle

The essential FEBE behaviour can be understood by taking its Fourier transform ("F.T."). Using $\tilde{D}_t^H \leftrightarrow \left(i\omega\right)^H$ (e.g. [Podlubny, 1999]) we obtain:

$$\left(\left(i\omega\right)^H + 1\right)\tilde{T}(\omega) = \lambda \tilde{F}(\omega)$$

(26)

where the tilde indicates Fourier transform with conjugate variable $\omega$, the frequency. This shows that if the forcing is purely sinusoidal $F = F_s e^{i\omega t}$ that the temperature response is also sinusoidal with the same frequency but different phase: $T = T_s e^{i\omega t}$. Using the notion of thermal impedance identified with the climate sensitivity, we can follow engineering practice (useful for estimating diurnal temperature - heating lags in buildings and structures) and use complex thermal impedances i.e. complex climate sensitivities $\lambda(\omega)$:

$$T_s = \lambda(\omega) F_s; \quad \lambda(\omega) = Z(\omega) = \frac{\lambda_0}{1 + (i\omega\tau)^H}$$

(27)
Where we have used the notation $\lambda_0 = \lambda(0)$ for the static climate sensitivity (i.e. $\lambda_0 = \lambda$ in the previous notation). In the context of the Earth’s energy balance, it is more useful to think in terms of sensitivities than impedances so that when convenient we use $\lambda(\omega)$.

Complex climate sensitivities are useful in understanding the responses to periodic forcings, in particular when the forcing is due to the annual cycle. Figs. 4, 5 compare the phases and amplitudes of $\lambda(\omega)$ as functions of the relaxation time $\tau$ (with $\omega = 2\pi \text{ rad/yr}$) for various values of $H$ including the HEBE value ($H = 1/2$) and the EBE value $H = 1$. From fig. 4, the thin horizontal lines show that taking the empirical value of $\tau$ in the range 1 - 5 years, and the approximate empirical value $H \approx 0.4$ (dashed), that the lag (the phase of the sensitivity) is in the range 25 - 30 days, which is in the observed range for the lag between the summer forcing maximum and maximum temperatures over most land areas.

If the forcing was only due to conductive heating (rather than conductive-radiative forcing), then, as pointed out in L1, one obtains the classical ($\pi/4$) lag (corresponding to $H = 1/2$, $\tau = \infty$) obtained by [Brunt, 1932] for the diurnal lag. This pure heating lag corresponds to 46 days and is already too long for most of the globe. Indeed, from the detailed maps in [Donohoe et al., 2020] we estimate that in the extratropical regions, over land, the summer temperature maximum is typically 30 - 40 days after the solstice, but only 20 - 30 days after the maximum forcing (insolation) and for ocean, 60 - 70 days after the solstice but only 30 - 40 days after the maximum insolation. Similarly, the $H = 1$ EBE lag is in the range 82 to 91 days (for $\tau > 1$ year), also much too long. While it is true that over the ocean, the lag is typically longer than over land, this is probably because of the strong albedo periodicity associated with seasonal ocean cloud cover [Stubenrauch et al., 2006]. [Donohoe et al., 2020]. This delays the summer solstice forcing maximum over the ocean, potentially explaining the extra lag.

L1 tested out the implications of the HEBE ($H = 1/2$) model using a latitudinally averaged model and parameters taken from [North et al., 1981]. The model uses a 2$^{nd}$ order Legendre polynomial to take into account the latitudinal variations and a sinusoidal annual cycle with empirically fit parameters. Since it only models latitudinal variations, it effectively averages over land and ocean. In work in progress, we use modern satellite data to test the model at 2$^{o}$ spatial resolution. Future work will use these new results coupled with modern higher resolution data of outgoing longwave radiation to investigate this further and obtain more accurate local climate sensitivities.
Fig. 4: The (negative) phase lag of the temperature with respect to the forcing (the negative of the argument of $\lambda$) lag in days for $H = 1/10, 3/10, \frac{1}{2}, 7/10, 9/10$ (black), 4/10 (dashed middle), 1 (dashed top) as a function of $\tau$ measured in years. The horizontal lines are for 25 and 30 days, we see that the dashed $H = 4/10$ falls in this range for between one and 5 years. This range of lags is close to the typical extra-tropical lags found in [Donohoe et al., 2020].
4. Estimating H, τ

4.1 The low frequency response and climate projections

The FEBE has distinct high and low frequency behaviours that can be understood from the complex climate sensitivity:

\[
\tilde{T}(\omega) = \lambda(\omega) \tilde{F}(\omega) = \begin{cases} 
\lambda_0(i\omega\tau)^{-H} \tilde{F}(\omega); & \omega \gg \tau^{-1} \\
\lambda_0(1-(i\omega\tau)^{-H}) \tilde{F}(\omega); & \omega \ll \tau^{-1}
\end{cases}
\]  

\( \tilde{T} \) is the Fourier transform of the climate model response. The term \( \lambda_0 \) is a constant, \( \omega \) is the angular frequency, and \( \tau \) is the relaxation time. The complex exponent \( (i\omega\tau)^{-H} \) captures the different modes of the climate sensitivity, with \( H \) being a dimensionless parameter that characterizes the low frequency response. For large \( \omega \), the response is dominated by the high-frequency modes, while for small \( \omega \), the low-frequency modes become significant.

Fig. 5: The modulus of the climate sensitivity as a function of the relaxation time \( \tau \) for \( H \) ranging from 1/10, 3/10, 5/10, 7/10, 9/10 (solid, top to bottom) and \( H = 4/10 \) (middle dashed) and \( H = 1 \) (bottom dashed).
Where the approximations were obtained as usual by expanding the denominator of eq. 27. In the high frequency limit, \((i\omega t)^{-H} \tilde{F}(\omega)\) is the Fourier Transform of the \(H\) order fractional integral of \(F\), therefore at high frequencies the temperature integrates (smooths) the forcing whereas at low frequencies, the lowest order term (unity) corresponds to a Dirac function (associated with the equilibrium temperature) while the next order \((i\omega t)^{-H} \tilde{F}(\omega)\) term corresponds to a fractional differentiation of order \(H\) that determines the rate at which equilibrium is approached.

Alternatively and equivalently, we can use small and large \(t\) expansions of the Green’s functions (eq. 21). Let’s first consider the large \(t\) asymptotic expansions:

\[
G_{\xi,H}(t) = \tau^{\xi-1} \sum_{n=0}^{\infty} \frac{(-1)^n}{\Gamma(\xi-nH)} \left(\frac{t}{\tau}\right)^{\xi-nH} ; \ \zeta \geq 0
\]

[Podlubny, 1999]. The most important cases are for \(\zeta = 0, 1, 2\) corresponding to impulse, step and ramp responses:

\[
G_{0,H}(t) = -\frac{\tau^{-1}}{\Gamma(-H)} \left(\frac{t}{\tau}\right)^{-1-H} + \frac{\tau^{-1}}{\Gamma(-2H)} \left(\frac{t}{\tau}\right)^{-1-2H} - O\left(\frac{t}{\tau}\right)^{-1-3H}
\]

\[
G_{1,H}(t) = \left(1 - \frac{1}{\Gamma(1-H)} \left(\frac{t}{\tau}\right)^{-H} + \frac{1}{\Gamma(1-2H)} \left(\frac{t}{\tau}\right)^{-2H}\right) + O(t^{3H}) \quad \text{for} \quad t >> \tau ;
\]

\[
G_{2,H}(t) = t - \frac{\tau}{\Gamma(2-H)} \left(\frac{t}{\tau}\right)^{-1-H} + \frac{\tau}{\Gamma(2-2H)} \left(\frac{t}{\tau}\right)^{-1-2H} + O(t^{3H})
\]

We have used \(\Gamma(0) = \infty\) for \(\zeta = 0\), note that \(\Gamma(-H) < 0\) for \(0<H<1\). From these, we see that as required for the step response, \(G_{1,H}(0) = 0\) and \(G_{1,H}(\infty) = 1\) so that at large \(t\), the FEBE implies a power law relaxation to the new equilibrium temperature.

From either the low frequency Fourier approximation (eq. 28) or the asymptotic expansion (eq. 30), we see that the low frequency response \(T_{low}\) can be approximated by the first term: \(G_{low,H}(t) \approx G_{0,H}(t)\):

\[
T_{low,H}(t) = \lambda G_{low,H} * F; \quad G_{low,H}(t) = -\frac{\tau^{-1}}{\Gamma(-H)} \left(\frac{t}{\tau}\right)^{-1-H}
\]

\(T_{low}\) is expected to be dominated by the forced response to external forcings, appropriate for making multi-decadal temperature projections from anthropogenic forcings (in eq. 31, we assumed that the initial temperature is zero, otherwise the extra term in eq. 20 can be used). Comparing this with the [Hébert et al., 2020] truncated scaling CRF (eq. 1) we see that the FEBE has identical low frequency behaviour with:

\[
H = -H_F.
\]
Using the empirical value $H_F = -0.5 \pm 0.2$ that [Hébert et al., 2020] obtained by (Bayesian) fitting the global temperature series since 1880 to the IPCC $CO_{2eq}$ forcing, we find $H = 0.5 \pm 0.2$. Similarly, [Procyk et al., 2020] fit the FEBE (using the Mittag-Leffler Green’s function) to estimate $H = 0.38 \pm 0.06$. In the next section we consider the high frequency behaviour and use the internal variability (assumed to be forced by a Gaussian white noise) that gives the more precise estimate $H \approx 0.42 \pm 0.02$ [Del Rio Amador and Lovejoy, 2019].

We have seen (eq. 15) that $t$ quantifies the rate that the temperature relaxes to equilibrium following a perturbation. At the same time, according to the above analysis, $\tau$ represents the transition time between two power law regimes $(t/\tau)^{H-1}$ to $(t/\tau)^{H-1}$; using the empirical value $H \approx 0.4$ this is a transition from $(t/\tau)^{0.6}$ $(t<<\tau)$ to $(t/\tau)^{1.4}$ $(t>>\tau)$ behaviour with a relatively small change in the scaling. The exact FEBE solution (eq. 20, fig. 2, right hand column) confirms that the transition is very slow - i.e. it occurs over a wide range of scales. This very slow transition makes $\tau$ difficult to accurately estimate from observations of the temperature responses to forcings (see appendix A).

Based on the truncated power law CRF (eqs. 1, 2), [Hébert et al., 2020] used physical arguments to determine $\tau$. At short enough time scales, different parts of the globe have temperature fluctuations that vary more or less independently of each other: they are primarily responses to “internal” forcings, small scale storage mechanisms. At long enough scales, the variability is expected to be primarily due to the responses from external forcings. [Hébert et al., 2020] argued that $\tau$ could be estimated from the scale at which the whole global surface temperature starts to fluctuate as a unit. In particular, the atmosphere and ocean should have temperatures that fluctuate synchronously: $\tau$ should be the atmosphere-ocean coupling time scale.

$\tau$ was therefore estimated by considering the correlations between fluctuations in the atmospheric temperature averaged over land, and fluctuations in the SST averaged over the oceans. At scales of months, there is practically no correlation: the land and ocean fluctuate nearly independently of each other. However at long time scales, they are on the contrary highly correlated: they fluctuate together. The transition between independence and dependence occurs over a short interval; between about 6 months and 2 years. [Hébert et al., 2020] used this to infer that $\tau \approx 2$ years, a value that gave reasonable climate projections through to 2100. This time scale is essentially the same as the lifetime of planetary scale ocean gyres; the ocean weather – ocean macroweather transition scale ([Lovejoy et al., 2017]).

More recently, [Procyk et al., 2020] used a Bayesian approach and the FEBE CRF to find the optimum parameters $\lambda$, $\tau$, $H$ constrained by the estimated historic $CO_{2eq}$ forcings and the responses as estimated by five observational series of global average temperatures since 1880, they obtain a very similar value $\tau \approx 4.7$yrs (the 90% confidence interval was 2.4 - 7.0 years). Since the FEBE is in the same family as the EBE, as a Bayesian prior distribution for $\tau$, they used [Geoffroy et al., 2013] estimate of $4.1 \pm 1.1$ years for the box ($H = 1$) model based on a dozen different GCM model outputs.
4.2 The high frequency response: internal variability and macroweather forecasts

In section 4.1 we saw that if we take $H = -H_F$, that to leading order, the low frequency FEBE and truncated model impulse CRFs are both proportional to $t^{H-1}$, these give reasonable approximations to the low frequencies. But with the FEBE, we are no longer thinking in terms of a small number of homogeneous boxes; rather the FEBE is a phenomenological model of a hierarchy of storage processes from large to small. It is therefore logical to assume that the FEBE applies to both external and internal forcings. Mathematically this implies that both variabilities have the same Green’s function, opening up the possibility of estimating climate parameters directly from the internal variability.

The idea of using the internal variability to infer the response to external forcing goes back to at least [Leith, 1975], who proposed that the Fluctuation Dissipation Theorem (FDT) be used for this purpose. The FDT says that the impulse response CRF of a nonlinear dynamical system to (small) external forcings is equal to the normalized covariance function. The FDT is attractive since - if it is valid - it is independent of specific climate models. The trouble is that it is only valid for infinitesimal fluctuations of systems at thermodynamic equilibrium, and it is not at all obvious that the climate system is close enough to equilibrium for the FDT to apply. Nevertheless, starting with Leith, numerous such attempts have been made (especially [North et al., 1993], [Cionni and Visconti, 2004], and recently [Cox et al., 2018]). As we show in a future publication, it turns out that the FEBE does obey the FDT, but only in the low frequency regime where $t \gg \tau$ where the scaling storage term is small enough that the system is indeed close to thermodynamic equilibrium. In contrast, the EBE ($H = 1$) special case is exceptional: the FDT always holds, and the temperature is an Ornstein-Uhlenbeck process, [North et al., 1993], [Cox et al., 2018]).

Although at high frequencies, the scaling prevents the FDT from being valid, if there is information about the internal forcing, then the FEBE can still be used to infer the CRF and hence the system response to external forcing. For example, if we make the usual assumption that the internal forcing is a Gaussian white noise, then we can use the temperature statistics to deduce the exponent $H$, (see table 1, discussed here and in the next section). If the noise amplitude is also known, we can deduce $\lambda$ (see table 1, bottom row). Before discussing the full stochastically forced case, in this section we discuss the high frequency stochastic response dominated by scaling storage processes.

In order to investigate the high frequency behaviour, use the first term of the Mittag-Leffler expansion (eq. 20) to obtain:

$$T_{\text{high}} \approx \lambda G_{\text{high}} \ast F(t); \quad G_{\text{high}} = \frac{1}{\Gamma(H)} \left( \frac{t}{\tau} \right)^{H-1}; \quad t \ll \tau \quad (33)$$

This is the singular high frequency FEBE limit corresponding to a fractional integration of order $H$ (eq. 28).
Since we associate the high frequency behaviour primarily with internal variability, let’s consider the case where its source $f(t)$ is a white noise. Ignoring external forcings so that $T_{high}$ is its stochastic response:

$$T_{high}(t) = \frac{\lambda}{\tau^H \Gamma(H)} \int_{-\infty}^{t} (t-t')^{H-1} f(t') dt'$$

(34)

to ensure convergence, we assume that $f(t)$ is a statistically stationary process representing the “innovations” with mean $<f(t)> = 0$ and we take $T(-\infty) = 0$.

To understand this high frequency behaviour, it is useful to compare it with the definition of a fractional Gaussian noise process (fGn):

$$g_{H_i}(t) = \frac{K}{\Gamma(H_i+1/2)} \int_{-\infty}^{t} (t-t')^{H_i-1/2} \gamma(t') dt'; \quad -1 < H_i < 0; \quad K = \left( \frac{\pi}{2 \cos(\pi H_i) \Gamma(2-2H_i)} \right)^{1/2}$$

(35)

where $\gamma(t)$ is a “δ-correlated” unit white noise process, the innovations satisfying $<\gamma(t)\gamma(t')> = \delta(t-t')$ and $<\gamma> = 0$ (see e.g. [Biagini et al., 2008]). $\gamma(t)$ is a generalized function – a “noise” – that is properly defined over finite integrals. Often, it is defined in terms of a Wiener process $W$ with $dW = \gamma(t) dt$. $K$ is a standard normalization constant chosen for convenience.

Comparing eq. 34 and 35, we see that if $f(t)$ is a white noise (i.e. proportional to $\gamma(t)$), and if we take: $H_i = H - \frac{1}{2}$ then the high frequency temperature response is an fGn process with exponent $H_i$. This result is easy to understand since at high frequencies (eq. 28), the derivative term in the FEBE is dominant (i.e. $D^{\gamma}T \gg T$) and $g_{H_i}$ is indeed the solution of the resulting fractional differential equation:

$$\lim_{t \to -\infty} D^{\gamma}_{t+1/2} g_{H_i}(t) = K\gamma(t)$$

(36)

The standard assumption about internal variability is that it is forced by a Gaussian white noise. With this, we can use this high frequency behaviour to estimate $H$ from the empirical values of $H_i$. As mentioned above, according to the global scale Haar fluctuation analyses (see e.g. [Lovejoy et al., 2015] ; [Del Rio Amador and Lovejoy, 2019]), $H_i \approx -0.08 \pm 0.02$ so that $H = H_i + 1/2 \approx 0.4$ which is essentially the same as the [Hébert et al., 2020], [Procyk et al., 2020] value estimated from the low frequency response to external forcings (eq. 31).

We could note that [Cox et al., 2018] recently used an analogous method of exploiting the internal variability to estimate the ECS. However, they made the unrealistic Ornstein-Uhlenbeck ($H = 1$) assumptions that the autocorrelation was exponential and that the ratio of temperature variance to forcing variance follows the (strongly biased) $H = 1$ result (table 1 bottom row).

fGn processes are power law smoothed white noises (eq. 35), but when $-1/2 < H_i \leq 0$ ($0 < H \leq 1/2$), the smoothing is not enough to eliminate the small scale divergences. For example, if we define the “$\tau_r$ resolution fGn” by averaging over a time interval $\tau_r$:

$$g_{H_i,\tau_r}(t) = \frac{1}{\tau_r} \int_{t-\tau_r}^{t} g_{H_i}(t') dt'$$

(37)
Then, with the above choice of constant $K$ we have:

$$\left\langle g_{H_i,t}^2(t) \right\rangle = \tau_r^{2H_i}, \quad -1 < H_f < 0$$

(38)

If $H < 1/2$, then $H_f = H - 1/2 < 0$, implying a small scale divergence as the resolution $\tau_r$ is reduced: fGn processes have strong resolution dependencies. An important but regularly overlooked practical consequence of such divergences is that in macroweather, this leads to “space-time reduction factors” that multiplicatively bias macroweather anomalies and are important for example in accurately estimating the temperature of the earth [Lovejoy, 2017].

### 4.3 The stochastic FEBE, fractional Relaxation motion (fRm), fractional Relaxation noise (fRn)

In the previous section we considered the high frequency stochastic response: fGn which is the solution of the FEBE at high frequencies where the (fractional) storage term dominates. We noted that since the FEBE is a linear equation, we could separately model the deterministic and stochastic (internal) variability. In this section we give more information about this pure stochastic case driven by white noise innovations, we briefly summarize and expand upon some of the results in [Lovejoy, 2019b] (hereafter L4). We note that the stationary process generated by the stochastic FEBE – fractional Relaxation noise (fRn) generalizes Ornstein-Uhlenbeck processes from $H = 1$ to $0 < H < 2$.

Whereas the deterministic fractional relaxation equation has been well studied (see e.g. [Miller and Ross, 1993], [Podlubny, 1999]), the stochastic version has enjoyed much less attention although [West et al., 2003] reviews some applications of the Riemann-Liouville case to nonstationary random walks. In contrast, the stochastic infinite range (Weyl version, when $t_0 \to -\infty$) leading to stationary solutions has only recently been considered (L4) including the optimum predictor problem (L4). Below we discuss the main points and give the dimensional form of the main statistical properties.

The physical problem that we wish to solve is the noise driven FEBE with noise amplitude $\sigma$ and sensitivity $\lambda$:

$$\tau^H \frac{d^H}{dt^H} T + T = \lambda f(t); \quad f(t) = \sigma \gamma(t)$$

(39)

with initial conditions $T(-\infty) = 0$, $0 \leq H \leq 1$ and $\gamma(t)$ a unit Gaussian white noise so that

$$\left\langle f^2 \right\rangle^{1/2} = \sigma; \quad \left\langle f \right\rangle = 0, \quad \sigma$$ is the amplitude of the noise, eq. 39 is a fractional Langevin equation (see L4). To understand the statistical properties of the internal variability response, it suffices to study the nondimensional equation:

$$-\frac{D^H_t}{\tau^H} U^H + U^H = N^H \gamma(t)$$

(40)
Where for $0 < H \leq 1/2$, $N_H = K$ is the normalization constant in eq. 35 (for $\sqrt{2} < H < 2$, see eq. 42) and the nondimensional $U_H(t)$ function is called a fractional Relaxation noise (fRN) since it generalizes fGn (see L4). It also generalizes the $H = 1$ Ornstein-Uhlenbeck process.

Using $U_H$ and noting that $\gamma(t^d) = \frac{\tau}{\tau} \gamma(t)$, we can obtain the solution to the dimensional eq. 39 using:

$$ T(t) = \frac{\lambda \sigma}{N_H \tau^{(\gamma)}} U_H \left[ \frac{t}{\tau} \right] $$

(41)

"$=" means equality in a probability sense.

Since the fRN process is the solution of the fractional relaxation equation with a stationary, Gaussian, zero mean, white noise forcing, it is also stationary, Gaussian with zero mean. Its statistics are therefore fully characterized by its autocorrelation function. A complication in the calculation is that when $0 < H \leq 1/2$, in the small $r$ limit, the fractional term of eq. 39 dominates so that we obtain the (fGn) limit. The solution of eq. 40 is therefore - like $\gamma(t)$ - a generalized function; to obtain solutions with finite variances, we must take averages over finite resolutions $\tau_r$.

The resulting $\tau_r$ resolution autocorrelation function at lag $\Delta t$ is:

$$ R_{H,\tau_r}(\Delta t) = \left\langle U_{H,\tau_r}(t) U_{H,\tau_r}(t - \Delta t) \right\rangle = N_H^2 \int_0^\infty G_{0,H}(\Delta t + s) G_{0,H}(s) ds; $$

$$ R_{H,\tau_r}(0) = \left\langle U_{H,\tau_r}(t)^2 \right\rangle = \tau_r^{2H-1}; \quad 0 < H < 1/2 $$

$$ R_{H,\tau_r}(0) = 1; \quad 1/2 < H < 2 $$

$$ N_H = K; \quad 0 < H < 1/2 $$

$$ N_H = c; \quad c^{-2} = \int_0^\infty G_{0,H}(s)^2 ds; \quad 1/2 < H < 2 $$

(42)

Note that when $0 < H < 1/2$, the equation for $R_{H,\tau_r}(\Delta t)$ is only valid for $\Delta t \geq \tau_r$, also recall the classical Ornstein-Uhlenbeck process ($H = 1$) where $G_{0,1}(t) = e^t$ and the autocorrelation is: $R_1(\Delta t) = e^{-\Delta t} / 2$. The above formulae were given for $\Delta t > 0$ but since $G_{0,H}$ is causal, $G_{0,H}(t) = 0$ for $t < 0$, which implies $R_H(\Delta t) = R_H(-\Delta t)$. Table 1 gives the explicit dimensional formulae taken from L4 for the temperature autocorrelation. Also given in the table are the corresponding formulae for the temperature-forcing ($f = \sigma_f$) and temperature storage cross-correlations ($s = dS_H / dt = \tau^H \tau_H D^H T = \lambda f - T$).

Physically, the model is justified as long as the resolution $\tau_r$ is greater than the weather-macroweather transition scale $\tau_w$ ($\approx 10$ days). $\tau_w$ is the inner (smallest) scale over which the FEBE may expected to be valid. From this, we can obtain the high frequency fGn approximation (valid for $\Delta t \ll 1$ corresponding to $\Delta t \ll \tau$ in the dimensional equation):
\[ R_{H,T} (\Delta t) \approx H (2H + 1) \Delta t^{2H-1}; \quad \tau_r \ll \Delta t < 1; \quad 0 < H < \frac{1}{2} \]  \hspace{1cm} (43)

(see table 1 for \( \frac{1}{2} < H < 2 \)). At low frequencies, for \( \Delta t >> 1 \), we obtain:

\[ R_{H} (\Delta t) = -\frac{N_H^2}{\Gamma(-H)} \Delta t^{-1-H} + O(\Delta t^{-1-2H}): \quad 0 < H < 1; \quad \Delta t >> 1 \] \hspace{1cm} (44)

valid for all (non-negative) \( \Delta t \), see L4 and table 1 for the dimensional expressions and note that \( \Gamma(-H) < 0 \) for \( 0 < H < 1 \). This large \( \Delta t \) result is independent of the resolution \( \tau_r \) and that it holds over a wider range of \( H \) values \( (H = 1 \) is the exponential exception, see table 1). A technical point is that although when \( 0 < H < 1/2 \), the high frequency limit is the \( fGn \) with a huge memory, due to the slow \( \Delta t^{2H-1} \) fall-off in \( R_H (\Delta t) \) (eq. 43), the corresponding \( fRn \) is effectively truncated with a memory of the order of \( \tau \) (see L4).

Table 1 summarizes the (dimensionalized) versions of these relations along with the corresponding cross-correlation functions. Of particular note are parameter ranges where the resolution is important \( (0 < H \leq 1/2 \) for \( R_{TT,TR} \) and \( 0 < H < 1 \) for the temperature – forcing cross correlation \( R_{T_s,TR} \)) as well as the exceptionally particularly rapid fall-off at large \( \Delta t \) of for the temperature – storage cross correlation \( R_{T_s,TR} \). Other properties of solutions to the stochastic eq. 40 are given in L4.

A final interesting property is the predictability of \( fRn \), also discussed in L4. since for times \( < \tau \), \( fRn \) is close to the long memory \( fGn \) process, the two can both be well predicted (for \( fGn \), the skill –with infinite past data- becomes perfect in the limit \( H \to 1/2 \) \( (H_T > 0) \)), however, beyond the relaxation time \( \tau \), \( fRn \) cannot be well predicted, it approaches a white noise.
<table>
<thead>
<tr>
<th>General</th>
<th>$\Delta t \ll \tau$</th>
<th>$\Delta t \gg \tau$</th>
<th>Range</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Temperature Autocorrelation function:</strong> $\langle T_r(t) T_r(t - \Delta t) \rangle = R_{TT, \tau} (\Delta t)$</td>
<td>$\frac{\lambda^2 \sigma^2}{K^2 \tau^2} \left[ H \left( 1 + 2 H \frac{\Delta t}{\tau} \right)^{2H-1} + \ldots \right] ; \quad \tau &lt; \Delta t$</td>
<td>$-\frac{\lambda^2 \sigma^2}{\tau} \frac{(\Delta t / \tau)^{-H}}{\Gamma(-H)}$</td>
<td>$0 &lt; H &lt; 1/2$</td>
</tr>
<tr>
<td></td>
<td>$\frac{\lambda^2 \sigma^2}{K^2 \tau} \left( \frac{\tau}{\tau} \right)^{2H-1} + \ldots ; \quad \Delta t = 0$</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>$\sqrt{\Delta \tau} = \frac{\lambda \sigma}{c^2} \frac{1}{\Gamma(2 + H)} \left( \frac{\Delta t / \tau}{\tau} \right)^{1-2H} + \ldots$</td>
<td></td>
<td>$1/2 &lt; H &lt; 1$</td>
</tr>
<tr>
<td></td>
<td>$\sqrt{\Delta \tau} = \frac{\lambda \sigma}{c^2} \frac{1}{\Gamma(2 + H)} \left( \frac{\Delta t / \tau}{\tau} \right)^{H-1} + \ldots$</td>
<td></td>
<td>$1 &lt; H &lt; 3/2$</td>
</tr>
<tr>
<td></td>
<td>$\frac{\lambda^2 \sigma^2}{2 \tau} e^{-\Delta \tau \tau}$</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Temperature-Forcing cross-correlation function:</strong> $\langle T_r(t) f_r(t - \Delta t) \rangle = R_{TF, \tau} (\Delta t)$</td>
<td>$R_{TF, \tau} (\Delta t) \approx \frac{\lambda \sigma^2}{\tau} G_{0,\nu} \left( \frac{\Delta t}{\tau} \right) ; \quad \tau &lt; \Delta t ; \quad \tau &lt; \Delta t$</td>
<td>$0 &lt; H &lt; 2$</td>
<td></td>
</tr>
<tr>
<td></td>
<td>$R_{TF, \tau} (0) \approx \frac{\lambda \sigma^2}{\tau} \frac{1}{\Gamma(2 + H)} \left( \frac{\tau}{\tau} \right)^{1-2H} ; \quad \tau &lt; \Delta t$</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Temperature-storage cross-correlation function:</strong> $\langle T_r(t) s_r(t - \Delta t) \rangle = R_{TS, \tau} (\Delta t)$</td>
<td>$R_{TS, \tau} (\Delta t) \approx \frac{\lambda \sigma^2}{\tau} \frac{1}{\Gamma(2 + H)} \left( \frac{\tau}{\tau} \right)^{1+H} ; \quad \tau &lt; \Delta t$</td>
<td>$0 &lt; H &lt; 2$</td>
<td></td>
</tr>
<tr>
<td></td>
<td>$R_{TS, \tau} (0) = \frac{\lambda \sigma^2}{\tau} \frac{1}{\Gamma(2 + H)} \left( \frac{\tau}{\tau} \right)^{1+H} ; \quad \tau &lt; \Delta t$</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Variance Ratio of temperature versus forcing:</strong> $\langle T_r^2(t) \rangle / \langle f_r^2(t) \rangle$</td>
<td>$\frac{\lambda^2 K^{-2} \left( \frac{\tau}{\tau} \right)^{2H}}{\tau}$</td>
<td></td>
<td>$0 &lt; H &lt; 1/2$</td>
</tr>
<tr>
<td></td>
<td>$\frac{\lambda^2 c^{-2} \left( \frac{\tau}{\tau} \right)^{2H}}{\tau}$</td>
<td></td>
<td>$1/2 &lt; H &lt; 2$</td>
</tr>
</tbody>
</table>

**Table 1:** Above the thick horizontal line: Dimensional formulae summarizing auto and cross correlation functions of the white noise forced FEBE, the autocorrelations are dimensionalized version of those in L4. $\tau$ is the relaxation time, $\tau_r$ is the resolution. $T_r(t)$ is a resolution $\tau$, fRn process, the special $H = 1$ (EBE) is the Ornstein-Uhlenbeck process; in the $H = 1/2$ (HEBE) special case, there are logarithmic corrections to $R_{TT, \tau}$, see appendix B of L4.
To ensure small scale convergence, all quantities are taken at resolution $\tau_n$ so that
\[
\left\langle f_{\tau_n}^2 \right\rangle = \sigma^2 / \tau_r
\]
(for $R_{TTR}$, this is only important for $H \leq 0.5$; for $R_{TFC}$, it is important for $H<1$). Also shown (bottom row) is the temperature to the forcing variance ratio. The numerical constant $K$ is from eq. 35 and $c$ is from eq. 42. Note the exact value $c^2 = 2$ for $H = 1$ and that $\Gamma(-H) < 0$ for $0 < H < 1$. When $H = 1/2$ there are small scale logarithmic divergences.

### 4.4 ECS and TCR

In order to make accurate climate projections throughout the 21\textsuperscript{st} century, the single most important parameter is the climate sensitivity $\lambda$ since for any given forcing, this determines the ultimate (long time) temperature change. It is conventional to express $\lambda$ in terms of the change in temperature following a CO\textsubscript{2} doubling, this $\lambda$ is the “Equilibrium Climate Sensitivity” (ECS). This is effectively a change in units and can be effected using the canonical value 3.71 W/m\textsuperscript{2} for a CO\textsubscript{2} doubling. However, in order to make accurate projections through the 21\textsuperscript{st} century, it is also important to have accurate estimates of the memory, the “warming in the pipe” [Hansen et al., 2011]. The standard way of quantifying this is by comparing the ECS to the Transient Climate Response (TCR) that is defined as the temperature change following a CO\textsubscript{2} doubling with forcing linearly increasing over 70 years: since the CO\textsubscript{2} forcing is logarithmic in the CO\textsubscript{2} concentration, this is nearly exactly a 1\% increase per year. The smaller the TCR/ECS ratio, the larger the memory. Note that in the following, we only consider static climate sensitivities that we indicate by $\lambda$ (not $\lambda_0$).

The dimensionless ratio TCR/ECS is independent of $\lambda$, for the FEBE model, it is only a function of $\tau$, $H$ and theoretically – since the TCR is defined in terms of a “ramp”, the ratio is:

\[
\frac{TCR}{ECS} = G_{s,H}(\frac{\Delta t}{\tau}) \bigg/ \Delta t = 1 - \frac{1}{\Gamma(2-H)} \left( \frac{\Delta t}{\tau} \right)^{-H} + \frac{1}{\Gamma(2-2H)} \left( \frac{\Delta t}{\tau} \right)^{-2H} + \ldots \quad ; \quad \Delta t >> \tau
\]

where $\Delta t$ is the period over which the doubling occurs (conventionally 70 years).

Fig. 6 shows a contour plot of the ratio as a function of exponent $H$ and relaxation times $\tau$. The FEBE parameters (black circle) estimated from the historical data and IPCC AR5 forcing by [Procyk et al., 2020] with 90\% confidence intervals are shown (dashed rectangle). One can see that the ratio is in the range $= 0.65 - 0.75$ with mean $\approx 0.7$. On can also see that the single box ($H = 1$) model has ratios near 1 unless the relaxation time is 20 years or longer ($\Delta t/\tau \approx 3.5$).

As $H$ approaches zero, we see that the transition from the ramp response to the equilibrium response takes a longer and longer time. In the limit $H \rightarrow 0^+$; the ratio TCR/ECS=1/2, an exact result as can be easily seen by using $\lim_{H \rightarrow 0^+} D^H = 1$ in the FEBE (eq. 14).
The $H=1$ (exponential) case has too short a memory to be realistic so that most GCM outputs have been fit to double exponential (2 box) models, for example, the IPCC AR5 suggests box relaxation times of 8.5 and 409.5 years with TCR/ECS=0.58. Similarly, the IPCC AR5 (ch.9, see also [Yoshimori et al., 2016]), obtained TCR/ECS = 0.56±0.19 (90% confidence limits) from 23 CMIP5 GCMs. Fig. 7 shows these model based estimates compared to the FEBE estimates. We can see that the two are compatible although the FEBE data-based estimate has a much smaller uncertainty than the model based one.

Fig. 6: A contour plot of TCR/ECS for various $H$ values as well as the dimensionless ratio of the ramp time $\Delta t$ to the relaxation time $\tau$; the canonical value of $\Delta t$ is 70 years. The dashed lines show the 90% confidence intervals from the analysis in Procyk et al 2020 for both $\tau$ and $H$; (using $\Delta t = 70$ years) the black circle represents the median values.
Fig. 7: The TCR/ECS ratio for the FEBE model as a function of $H$ for various relaxation times using the canonical ramp time constant $\Delta t = 70$ years with relaxation time $\tau$. The black circle and rectangle shows the mean and 90% confidence interval values found by [Procyk et al., 2020]. The horizontal dashed lines show the 90% confidence limits (IPCC AR5 ch. 9) found using CMIP5 GCMs.

4.5 Modelling the internal and externally forced variability: a simple FEBE model from 1825-2100

We have seen that with a single exponent $H$, the FEBE correctly predicts both high and low frequency scaling regimes, allows it to convincingly model the internally and externally forced variability, the two are linked by the earth’s multiscale (and scaling) storage mechanisms. Let us therefore treat the total forcing $\mathcal{I}(t)$ as the sum of deterministic part $F(t)$ and a stochastic part $f(t)$ with $\langle f(t) \rangle = 0$:

$$\mathcal{I}(t) = F(t) + f(t)$$

(46)

and similarly for the temperature response:

$$T(t) = T_e(t) + T_i(t)$$

(47)

where the subscripts “e” and “i” are for externally forced and internal respectively.

We propose that $\mathcal{S}(t)$ satisfies the stochastic FEBE:

$$\tau'' \to D_i'' T + T = \lambda \mathcal{S}$$

(48)
If we identify the deterministic part with the externally forced response and the stochastic part with the internal variability, then the linearity of the FEBE implies that each satisfies the FEBE separately. First, take ensemble averages and use \( \langle f(t) \rangle = \langle T_i(t) \rangle = 0 \):

\[
\tau \left( \int_{-\infty}^{t} D_i^H T_i \right) + T_e = \lambda F(t); \quad F = \langle \mathcal{S} \rangle; \quad T_e = \langle T \rangle
\]  
(49)

(since \( T_i(-\infty) = 0 \)). Now, subtract this from eq. 48 to yield:

\[
\tau \left( \int_{-\infty}^{t} D_i^H T_i \right) + T_i = \lambda f(t)
\]  
(50)

Therefore in the stochastic FEBE, the externally forced and internal variability both simply represent sources or sinks of heat energy that are external to the radiating surface layer. Whether heat comes from internal storage or from outside the system, all that matters is that it appears at the surface where it can participate in the radiative part of the energy balance.

To illustrate how the FEBE models both the deterministic external and stochastic internal variability, let’s make a simple model that includes them both (see figs. 8, 9 and details in appendix A). The model has a resolution of the weather-macroweather transition scale (taken as 10 days) and it solves the FEBE over 100,000 days starting in the year 1825, up to the year 2100. The deterministic part of the forcing is taken to be zero up until 1880, this is followed by a linearly increasing forcing - a “ramp” - up until the year 2020 (\( \approx 51100 \) days) after which the forcing remains constant at that value until the year 2100. The stochastic part of the forcing is due to a Gaussian white noise whose amplitude is such that the RMS monthly anomaly is \( \pm 0.14K \) (as observed, e.g. [Lovejoy, 2017]). The forcing was adjusted so that the forced response was 1K in 2020 (close to the observed anthropogenic warming), according to this model, there is another 0.24K of warming “in the pipe”. More details including a scale by scale statistical analysis and comparison with statistical analyses of global temperature series are given in appendix A. In place of this “toy model”, [Procyk et al., 2020] uses IPCC AR5 forcings and makes quite realistic hindprojections (1880- present) and projections (through to 2100) that are within the uncertainty limits of the IPCC AR5 (CMIP5) GCMs.
Fig. 8: This shows the nondimensional temperature $T/T_\infty$ response (thin lines) for nondimensional forcing $F/F_\infty$ (thick line) with unit sensitivity ($\lambda = 1$). Each graph shows the results for $H = 0$ (bottom) to $H = 1$ (top, exponential), at intervals of 0.1; the red curve highlights the most realistic value $H = 0.4$. The plots are shown for the relaxation time $\tau = 10^2$, $10^3$, $10^4$ days (upper left to bottom as indicated). The resolution of the calculations were taken to be the weather – macroweather transition scale ($\tau_w = 10$ days), but this is only important for the stochastic internal variability; smooth deterministic forcings such as those here are insensitive to $\tau_w$ as long as it is much smaller than $\tau$. 
Fig. 9: Thin black line: A simulation of the monthly resolution temperatures with resolution \( \tau_w = 10 \) days, the sum of the externally forced model with \( H = 0.4, \tau = 1000 \) days (fig. 8, upper right) with the Gaussian white noise driven stochastic internal variability model averaged over a factor of three to simulate a monthly resolution series. The amplitude of the internal variability was chosen to match observations of the monthly resolution standard deviation of global temperatures \( \left( T_{\tau_w}^2 \right)^{1/2} = 0.14 \)C. Thick black line: at annual resolution (obtained by averaging the previous over an additional factor of 12).

5. Conclusions

Beyond the deterministic forecasting limit of about ten days, GCMs are effectively stochastic. In addition - at least with the external forcing scenarios used in the IPCC AR5 - the Representative Carbon Pathways - each GCM response to external forcings is quite linear [Hébert and Lovejoy, 2018]: to a good approximation, climate projections over the next century are a linear stochastic problem [Hebert, 2017], [Hébert et al., 2020], [Procyk et al., 2020]. In addition, stochastic, linear monthly, seasonal and interannual scale (macroweather) temperature forecasts have skill comparable – or better – than those of GCMs [Lovejoy et al., 2015], [Del Rio Amador and Lovejoy, 2019].
These facts – and the urgency of reducing forecast and (especially) projection uncertainties – militates for the rapid development of stochastic macroweather models constructed directly at macroweather scales. Up until now, such models have been based on the long memory of the Earth’s atmosphere and climate that was exploited by using the scaling principle: the fact that over wide ranges of spatial scales, atmospheric dynamics are scaling (see the extensive review [Lovejoy and Schertzer, 2013]). Predicting such long memory stochastic process requires a long series of historical data, therefore mathematically, these methods are effectively “past value” rather than conventional “initial value” type problems ([Lovejoy et al., 2015]).

At first, scaling was invoked to directly mediate the forcing and the response; a scaling Climate Response Function (CRF, [Rypdal, 2012], [van Hateren, 2013], [Hebert, 2017], [Hébert et al., 2020], [Myrvoll-Nilsen et al., 2020]). Recently, [Lovejoy, 2019b] and [Lovejoy, 2019a] made this phenomenological approach more physical (and realistic) by suggesting that rather than applying it to the CRF directly, that the scaling principle should instead be applied to the energy storage mechanisms. They pointed out that if this was done, that both high and low frequency memories and exponents would automatically be explained by a single more fundamental exponent $H$, the (fractional) order of the resulting Fractional Energy Balance Equation (FEBE).

Since then, another approach starting with energy balance models has reached the same conclusion and has shown how the zero-dimensional (global) models can be generalized to full 2D models for temperature anomalies. Perhaps the most surprising result [Lovejoy, 2020a; c] was that an apparently minor change to the classical Budyko-Sellers Energy Balance Models – the extension from 2D to the full (classical, integer-ordered) 3D continuum heat equation – generically implies that the surface temperature satisfies the half-order EBE (HEBE). Finally, a generalization to the (3D) Fractional Heat Equation [Lovejoy, 2020b] was shown to imply that the (2D) surface temperatures satisfy the FEBE.

In this paper, we pursued a more phenomenological approach by deriving the FEBE from the scaling principle applied to the energy storage processes. Mathematically, the FEBE is a fractional relaxation equation whose deterministic version has been studied for some while [Miller and Ross, 1993], [Podlubny, 1999] and whose stochastic version was studied in [Lovejoy, 2019b]. Basing ourselves on these results, we showed how the FEBE can be solved using Green’s functions and we linked the high and low frequency limits of the Green’s function to fractional Gaussian noise (high frequencies dominated by internal variability) and low frequencies to power law CRFs for climate projections. The fact that the stochastic FEBE is a good model for the internal variability means that the internal variability can be used to determine the response of the system to external forcing, a goal that has been pursued ever since [Leith, 1975] proposed using the Fluctuation Dissipation Theorem for this purpose.
We showed how the response to periodic forcing can be conveniently handled by considering complex climate sensitivities notably to take into account the phase lag between the annual maximum forcing and the temperature response. This opens the possibility of using the annual cycle to estimate the model parameters, exponents, relaxation times and usual (static) climate sensitivities. We showed that although the FEBE obeys Newton’s law of Cooling, that the heat flux crossing a surface nonetheless depends on the fractional temperature time derivative rather than the usual integer ordered one (this is an extension of the half-order result that follows quite classically from the continuum mechanics heat equation). We also derived the theoretical FEBE ratio of transient to equilibrium climate sensitivity (TCR/ECS) and showed that it was ≈0.70, i.e. within the 90% confidence interval of GCM estimates (IPCC AR5).

Finally, we put both the deterministic (external forcing) and stochastic (internal forcing) together into a simple 140 year ramp forcing model for the industrial epoch warming. The model assumed that between 1880 and 2020 there was 1K of warming and predicted that if the current forcing was held constant from 2020 onwards, that there would be another 0.24K “in the pipe”. We then statistically evaluated the model using Haar fluctuations and showed that the RMS fluctuations as functions of time scale were already fairly realistic, although precise parameter estimates are difficult.

Although in this paper we only considered the zero-dimensional FEBE, various extensions to 2D have already been proposed [Lovejoy, 2020b]. When forced by stochastic internal and deterministic external forcing, the FEBE thus constitutes a new class of low frequency atmospheric model that uses historical data and the long memory to make macroweather forecasts and climate projections. There are indications that these are more skillful and are less uncertain than conventional approaches [Del Rio Amador and Lovejoy, 2019], [Procyk et al., 2020]. The future challenge will be to improve this approach and extend it for predicting and projecting other atmospheric parameters - especially for precipitation.

7. Acknowledgements

This project received no specific funding, but S. Lovejoy benefitted from a small grant from the National Science and Engineering research Council (Canada). We acknowledge discussions with D. Clarke and C. Penland.

Appendix A:

To illustrate how the FEBE models both the deterministic external and stochastic internal variability, in section 4.5 we presented a simple model that includes them both (figs. 8, 9). In this appendix we give a few technical details and a statistical validation of the model using global temperature series.
First, consider the externally forced variability using a linearly increasing (ramp) model of the industrial epoch anthropogenic warming. Fig. 8 shows the forcing model (thick black) nondimensionalized with the long-time forcing $F_\infty$ with the response nondimensionalized by the long time (equilibrium) temperature $T_\infty$; since $\lambda = T_\infty / F_\infty$, the curves are independent of $\lambda$. The model has a resolution of the weather-macroweather transition scale taken as 10 days and it solves the FEBE up to 100,000 days starting in the year 1825, up to the year 2100. The forcing is taken to be zero up until 1880, this is followed by a linearly increasing forcing - a "ramp" - up until the year 2020 ($\approx 51100$ days) after which the forcing remains constant at that value until the year 2100.

In the historical part, the forcing is indeed roughly of this form, the amplitude estimated for 2020 at close to 2.4 W/m$^2$ (CO$_2$eq forcing, since the beginning of the industrial epoch). The resolution of the simulation was the typical weather-macroweather transition time $\tau_w \approx 10$ days, but as long as $\tau >> \tau_w$ the results for this smooth, deterministic simulation will be insensitive to this. However, as explained in section 4.2 (eq. 42, table 1), the resolution is fundamentally important for stochastic, noise driven simulations of internal variability. The future part of the model corresponds to a complete halt to all emissions (and other forcing) in 2020.

Before continuing, it’s worth making a quick comment about the numerical techniques that were used for the simulations. Numerical convolution algorithms are very efficient, so that we used the FEBE solution in convolution form. Since the FEBE Green’s functions have power law behaviour at both high and low frequencies, care is required at small and large scales. To avoid high frequency problems, (especially sensitive since below, the FEBE is stochastically forced with a Gaussian white noise), we therefore used the step response $G_{1,1}$ Green’s function that has a smoother small $t$, behaviour ($t^{1-H}$) rather than the impulse response $t^H$ behaviour. After the convolution, the resulting series was then numerically differentiated to obtain the solution. The low frequencies also have some issues because of the long memory: the storage. For simulations of pure fGn this is a nontrivial issue. Here however, for scales $t \gg \tau$, the memory - although still power law (eq. 31, $t^{1-H}$) - is much shorter than for pure fGn ($t^H$) so that it suffices to make simulations of total length a few times $\tau$. More details are given in [Lovejoy, 2019b].

Using these techniques, three plots are shown in fig. 8, one for each of the relaxation scales $\tau = 100, 1000, 10,000$, simulated days (upper left, upper right, bottom). The upper right simulation with $\tau = 1000$ days ($\approx 3$ years) has a value close to the various relaxation scale estimates and the curve corresponding to the realistic parameter $H = 0.4$ (red) is fairly close to the actual reconstructed warming over the period. These curves illustrate what might happen if emissions suddenly stopped, the parts beyond 2020 show the thermal “inertia” in the system. We see that with only some exceptions at the very long $\tau = 10,000$ ($\approx 30$ years, bottom) that the response is very linear over the ramp region, so that the response at the end of the 140 year ramp is a little larger than the TCR (defined after a 70 year ramp, see fig. 6, 7; the TCR/ECS contours in the figure are fairly horizontal, not very sensitive to $\Delta t/\tau$).
The $H = 0.4$ curve on the graph with the most realistic $\tau$ (=1000 days upper right) shows that $T(2020)/T(\infty) = 0.81$ (the TCR/ECS is $\approx 0.7$ see figs. 6, 7). If we take the current increase in temperature since pre-industrial epoch to be $T(2020) = 1$K, then the equilibrium temperature $T(\infty) = \text{ECS} = 1/0.81 = 1.24$K. If we take the corresponding forcing since the pre-industrial epoch to be 2.4 W/m$^2$, then $\lambda = 1.24/2.4 = 0.52$ K/W/m$^2$; using the value 3.71 W/m$^2$/ CO$_2$ doubling, we obtain $\lambda = 1.92$ C/CO$_2$ doubling. Comparing this with the IPCC AR5 range of 1.5 - 4.5 K/CO$_2$ doubling (90% confidence), we see that it is about 4% below the mean and easily within the range, it is close to the [Procyk et al., 2020] estimate $2.0 \pm 0.4$. This is surprisingly good confirmation for a model based only on the scaling storage assumption.

We can now combine the externally forced part of the model (fig. 8) with a simulation of the monthly resolution internally forced temperatures: the sum of using the $H = 0.4$, $\tau = 1000$ day model (middle upper left) averaged over a factor of three to simulate a monthly resolution series. Fig. 9 shows the result at monthly resolution (thin) and when this is averaged further to annual resolution (thick); the amplitude of the noise contribution was chosen to equal $\pm 0.14$K at monthly resolution, close to the real global statistics. Due to the (singular) small scale fGn behaviour, at annual resolution, the amplitude of the noise is thus reduced by the factor $12^H I = 12^{-0.1} = 0.78$ (eq. 38); i.e. the annual anomaly fluctuation amplitude is $0.11 = 12^H / 0.14$K.
Fig. 10a: The average Haar fluctuation structure function $S(\Delta t) = \left\langle \Delta T(\Delta t) \right\rangle^{1/2}$ for six globally averaged monthly resolution temperature time series (brown) compared with ramp model simulations with $H = 4/10$ ($H_1 = H - \frac{1}{2} = -1/10$) analyzed over the simulated period 1880-2012 (black; these are the series analysed in [Lovejoy, 2017]). The simulations have been calibrated so that their RMS variability at one year is the same as the data. Simulation results are shown for relaxation times varying from $10^{-2}$ days (top left) to $10^4$ days (dashed, bottom, left) each labelled at left by $\log_{10} t$. The reference (red) line has the theoretical slope ($H_1$) for fGn i.e. infinite relaxation time.
In spite of its simplicity, fig. 9 shows that the model is apparently quite realistic, yet proper statistical analysis is needed to confirm this. A convenient way to compare the model statistics with those of globally averaged temperature series is to estimate their root mean square Haar fluctuations. For a series \( T(t) \), the Haar fluctuation \( \Delta T(\Delta t) \) over an interval \( \Delta t \) is the difference between the average of the first and second halves of the interval. Fig. 10a,b shows the root mean square (RMS) Haar fluctuations \( S(\Delta t) = \left( \langle \Delta T(\Delta t)^2 \rangle \right)^{1/2} \), "< >" indicates averaging over all the available series and disjoint intervals) for six globally averaged series from 1880-2012 (thick brown) compared to simulated series with various relaxation times \( \tau \) ranging from \( 10^{-2} \) (top) to \( 10^4 \) days (bottom) for \( H = 4/10 \) (fig. 10a) and \( H = 1/2 \) (fig. 10b, the empirical curve is repeated on each). As mentioned, the simulations were "calibrated" by adjusting the amplitude of the Gaussian noise part of the forcing so that the amplitude of the model and empirical fluctuations were equal at annual scales. The monthly series are identical to those analysed in depth in [Lovejoy, 2017]; full details and references are given there.
The Haar fluctuations combine averaging and differencing so that their interpretation is straightforward. When the fluctuations increase with scale (i.e. regimes where the slopes in the plot are positive), and the raw fluctuations are multiplied by a factor 2 (as they are here, this is the canonical factor), the differencing dominates and the values are close to the RMS differences that quantify the typical change in temperature over a given time interval. For example, the figure shows that at centennial scales (far right), the typical fluctuations are of the order of 1K. In scale ranges over which the RMS fluctuations decrease, the averaging dominates and the values are close to the anomaly fluctuations i.e. the temporal mean of the series over $\Delta t$ after its overall mean has been removed.

Over the range of 2 months (the smallest $\Delta t$) to about 10 years, their approximate slope (indicated in the by the straight reference lines in fig. 10a, b) is about $H_t = -0.1$. In this decreasing macroweather regime, successive fluctuations tend to cancel each other out, the series appears to be stable. For $\Delta t \approx > 10$ years, the fluctuations start increasing: the climate regime. 10 years is the typical scale at which the response to anthropogenic forcing exceeds typical responses to internal forcing, at these scales, fluctuations no longer tend to cancel each other out, the series appears to be unstable.

According to the high frequency theory, we developed in section 4.2, for time scales much smaller than the relaxation scale $\tau$, the stochastic part of the response is an fGn with RMS exponent $H_t = H - 1/2$, so that when $\Delta t \ll \tau$ is large enough $S(\Delta t) \approx \Delta t^{H_t}$. Fig. 10a shows the simulation results for the empirically estimated $H = 4/10$ and the theoretical (large $\tau$) $H_t = -1/10$ reference line is shown in red. Even when $\tau = 10^4$ days ($\approx$ 30 years, dashed) the small $\Delta t$ slope of the simulation $S(\Delta t)$ is still a little high; although overall the fit is not bad. When we compare the simulation structure functions for decreasing $\tau$, we note a surprising feature: the $\log_{10} S(\Delta t)$ versus $\log_{10} \Delta t$ plots are remarkably linear over several orders of magnitude but with (absolute) slopes that systematically increase as $\tau$ decreases. Without the deterministic ramp that dominates the lower frequencies, this “pseudo-scaling” was investigated further in L4. Rather than being a true scaling regime, it is in reality a very slow transitional regime from high frequency $\Delta t^{H_t}$ to $\Delta t^{-1/2}$ (independent of $H$, $H_t$). Fig. 10b shows the same simulations but with $H = 1/2$, ($H_t = 0$). In this case, the effect is even more pronounced with the data being more compatible with $\tau \approx 10^3$ days (3 years). From these simulations, we see that when only a factor of 100 or 1000 in scale is available for analysis, it is may be difficult to accurately measure $H$ and $\tau$.

References


