Manuscript prepared for Earth Syst. Dynam.
with version 2015/04/24 7.83 Copernicus papers of the LaTeX class copernicus.cls.
Date: 10 January 2016

Global warming projections derived from an observation-based minimal model

K. Rypdal[1]

[1]Department of Mathematics and Statistics, UiT The Arctic University of Norway, Norway

Correspondence to: Kristoffer Rypdal
(kristoffer.rypdal@uit.no)

Appendix E: Divergences, causality and initial conditions

If $G(t)$ is a power law the integral over prehistory $t \in (-\infty, 0)$ may lead to paradoxes, such as divergences of the integral. The solution to the paradox is to interpret the power-law as an approximation, for instance to a superposition of exponential response kernels. For a white-noise forcing this corresponds to an aggregation of Ornstein-Uhlenbeck (OU) processes, which are known to have the potential to produce a process that is a very good approximation to a fractional Gaussian noise (fGn) up to the time scale corresponding to the OU process with the greatest correlation time (*Granger*, 1980).

The scaling properties on scales of decades and longer arise from the heat transport within the oceans. This transport exhibits a maximum response time, which will provide an upper (exponential) cut-off of the power-law response function, but the characteristic time of this cut-off may be centuries or millennia. *Fraedrich and Blender* (2003) state in their abstract: "Scaling up to decades is demonstrated in observations and coupled atmosphere-ocean models with complex and mixed-layer oceans. Only with the complex ocean model the simulated power laws extend up to centuries."

If we don't treat the power-law as an approximation we have to deal with the divergences of the integral

$$\Delta T(t) = \int_{-\infty}^{t} G(t-t')\,F(t')\,dt', \tag{1}$$

where $G(s) = s^{\beta_T/2-1}$. If we consider the unit step-function forcing $F(t) = H(t)$, and $\beta_T \neq 0$, the integral is

$$\Delta T(t) = \lim_{\epsilon \to 0^+} \int_{\epsilon}^{t} (t-t')^{\beta_T/2-1}\,dt' = \lim_{\epsilon \to 0^+} \int_{\epsilon}^{t} s^{\beta_T/2-1}\,ds = \lim_{\epsilon \to 0^+} \frac{2}{\beta_T}(t^{\beta_T/2} - \epsilon^{\beta_T/2}). \tag{2}$$

Clearly $\Delta T(t)$ diverges as $t \to \infty$ if $\beta_T > 0$, but it also diverges if $\beta_T < 0$ (as $\epsilon \to 0^+$). For $\beta_T = 0$ there is a logarithmic divergence in both limits.

For physically meaningful results the $\beta_T > 0$ case requires some sort of cut-off (e.g., an exponential tail) for sufficiently large t, and the $\beta_T < 0$ case requires an elimination of the strong singularity of $G(s)$ at $s = 0$. A shown in Appendix **??**, AOGCMs in the CMIP5 ensemble with step function forcing indicate a power-law response for large s at least up to 150 yr (and the GISS-E2-R model up to 2000 yr) with $\beta_T \approx 0.35$, so $\beta_T > 0$ is the case of interest for the global temperature response. The AOGCMs are also well approximated by an exponential response in the limit $s \to 0$ (for s up to a few years), so an exponential truncation in this high-frequency limit is also appropriate.

The truncation of the power-law kernels is a physical, and not a technical mathematical issue. It is an approximation to a hierarchy of exponential responses. With this interpretation the divergences evaporate. Below is a more detailed outline of this philosophy in an energy-balance context. Let us take as a starting point the simple zero-dimensional EBM before linearisation of the Stefan-Boltzmann law;

$$C\frac{dT}{dt} = -\epsilon\sigma_S T^4 + I(t), \tag{3}$$

where T is surface temperature in Kelvin, C is an effective heat capacity per area of the Earth's surface, σ_S is the Stefan-Boltzmann constant, ϵ is an effective emissivity of the atmosphere, and $I(t)$ is the incoming radiative flux density at the top of the atmosphere. Let $I_0 = I(0)$ be the initial incoming flux, $F(t) = I(t) - I_0$ is the radiative forcing, $T_{\text{eq}} = (I_0/\epsilon\sigma_S)^{1/4}$ is the equilibrium temperature at $t = 0$, $\Delta T(t) = T(t) - T_{\text{eq}}$ is the temperature anomaly measured relative to the initial equilibrium temperature, and $\Delta T_0 = \Delta T(0)$ is this anomaly at $t = 0$. Note that F here is the perturbation of the radiative flux with respect to the initial flux I_0 and not with respect to the flux $\epsilon\sigma_S T_0^4$ that would be in equilibrium with the initial temperature T_0. The linearised EBM for the temperature change relative to the temperature T_{eq} (the one-box model) is

$$\frac{d\Delta T}{dt} = -\nu\Delta T + \mathcal{F}(t), \quad \Delta T(0) = \Delta T_0, \tag{4}$$

where $\nu = 4\epsilon\sigma_S T_{\text{eq}}^3/C$, $\mathcal{F}(t) = F(t)/C$. By definition $\mathcal{F}(0) = [I(0) - I_0]/C = 0$. This is Eq. (**??**) and Eq. (**??**) with slightly different notation. The solution to the initial value problem (i.v.p.) Eq. (4), with the initial condition $\Delta T(0) = \Delta T_0$, takes the form

$$\Delta T_{\text{i.v.p.}} = \int_0^t G(t-t')\mathcal{F}(t')\,dt' + \Delta T_0 e^{-\nu t}, \tag{5}$$

where $G(s) = \exp(-\nu s)$. The generalisation to a linear, causal response model, where $G(s)$ is not necessarily exponential, involves extending the integration domain in Eq. (5) to the interval $(-\infty, t)$;

$$\Delta T_{\text{r.m.}}(t) = \int_{-\infty}^t G(t-t')\mathcal{F}(t')\,dt'. \tag{6}$$

From the initial condition $\Delta T(0)_{\text{r.m.}} = \Delta T_0$ Eq. (6) yields,

$$\Delta T_0 = \int_{-\infty}^{0} G(-t')\mathcal{F}(t')\,dt'. \tag{7}$$

For exponential response $G(s) = \exp(-\nu s)$ it is easy to verify that $\Delta T_{\text{i.v.p.}}(t) = \Delta T_{\text{r.m.}}(t)$, and Eq. (7) yields the following relation between the initial temperature anomaly and the forcing $\mathcal{F}(t)$ for $t \in (t, 0)$;

$$\Delta T_0 = \int_{-\infty}^{0} e^{\nu t'}\mathcal{F}(t')\,dt'. \tag{8}$$

For the exponential response there is no "divergence issue" in Eq. (6). Neither is there such an issue for the two-exponential solution to the two-box model (*Geoffroy et al.*, 2013). An "N-box model" exhibits a response function for the temperature in each box which is a superposition of exponentials; $G(s) = \sum_{i=1}^{N} a_i \exp(-\nu_i s)$. For the surface (mixed layer) box the temperature anomaly takes the form

$$\Delta T_{\text{r.m.}}(t) = \sum_{i=1}^{N} a_i e^{-\nu_i t} \int_{-\infty}^{t} e^{\nu_i t'}\mathcal{F}(t')\,dt'. \tag{9}$$

On the other hand, the N-box initial value problem has solution of the form

$$\Delta T_{\text{i.v.p.}}(t) = \sum_{i=1}^{N} a_i e^{-\nu_i t} \int_{0}^{t} e^{\nu_i t'}\mathcal{F}(t')\,dt' + \sum_{i=1}^{N} b_i e^{-\nu_i t}, \tag{10}$$

where the coefficients b_i are linearly related to the initial temperatures of each box; $b_i = \sum_{j=1}^{N} M_{ij} T_{0j}$. The condition $\tilde{T}_{\text{i.v.p.}}(t) = \tilde{T}_{\text{r.m.}}(t)$ now yields the relations between the initial temperatures and the prehistory of the forcing;

$$\sum_{j=1}^{N} M_{ij}\Delta T_{0j} = a_i \int_{-\infty}^{0} e^{\nu_i t'}\mathcal{F}(t')\,dt' \;\text{ for }\; i = 1,\ldots,N. \tag{11}$$

With a white-noise forcing $\mathcal{F}(t)$ the Eq. (4) is the Itô stochastic differential equation (in physics often called the Langevin equation). The solution is the Ornstein-Uhlenbeck (OU) stochastic process, which in discrete time corresponds to the first-order autoregressive (AR(1)) process. The power spectral density of this process is essentially a Lorentzian, which means that the high-frequency ($f \gg \nu$) part of the spectrum has the form $\sim f^{-2}$, and the low-frequency part $\sim f^0$. This means that if the climate response were well described by a one-box EBM we could use a power-law response model with $\beta_T \approx 2$ on time scales much shorter than the correlation time $\tau_c = \nu^{-1}$. On these time scales the stochastic process exhibits the characteristics of a Brownian motion (Wiener process), which is a self-similar process with spectral index $\beta = 2$. This process is non-stationary, and hence suffers from the divergences that we are worried about. But even though the Brownian motion diverges, the OU-process does not, because of the flattening of the spectrum for $f \ll \nu$.

Both observation data and AOGCMs indicate that the one-box EBM is inadequate, but the considerations above are equally valid for an N-box model, for which the white-noise forcing gives rise to an aggregation of OU-processes with different ν_i. Such an aggregation is known to be able to produce a process with approximate power-law spectrum with $0 < \beta < 2$ on time scales $\tau < \nu_{\text{min}}^{-1}$ (*Granger*, 1980).

Lovejoy et al. (2013) specifically argue that volcanic forcing may have a scaling exponent $\beta_F \approx 0.4$, and hence the convergence criterion $\beta = \beta_T + \beta_f < 1$ then requires $\beta_T < 0.6$. One remark to this is that the above discussion shows that the $\beta < 1$ criterion is not necessary on time scales shorter than $\tau < \nu_{\text{min}}^{-1}$. However, observation indicates that $\beta < 1$, so this does not invalidate their argument. More important is that in recent papers the response to volcanic forcing has been subtracted from both instrumental and multiproxy reconstruction data *Rypdal and Rypdal* (2014) and from millennium-long AOGCM simulations (*Østvand et al.*, 2014), and the residuals have been analysed for β without finding a detectable influence of the volcanic forcing on β. The same is seen by comparing control runs of the AOGCMs with those driven by volcanic forcing (*Østvand et al.*, 2014).

The importance of including the prehistory of the energy-flux imbalance when deriving projections for future change can be illustrated by considering a prehistory consisting of volcanic forcing $\mathcal{F}_V(t)$ only. The particular feature of volcanic forcing is that it consists of a succession of negative spikes in the radiation flux. If we assume that the time $t = 0$ is in a period with no volcanic forcing we can for illustration think of the forcing as a succession of negative forcing events of short duration, randomly distributed in time with typically longer waiting times between events than durations. Let us further assume that the climate response is so slow that $G(t)$ varies by a small amount over the mean waiting time. Hence, there exists time intervals of duration Δt which are short enough for $G(t)$ to be nearly constant over the interval, but long enough to have a sufficient number of large volcanic eruptions to estimate a mean volcanic forcing $\overline{\mathcal{F}}_V$. This assumption is not very good in practice, but let us use it for illustration. Under this assumption we can approximate the integral

$$\int_{t_1-\Delta t/2}^{t_1+\Delta t/2} G(t-t')\mathcal{F}_V(t')\,dt' \approx G(t-t_1)\int_{t_1-\Delta t/2}^{t_1+\Delta t/2} \mathcal{F}_V(t')\,dt = G(t-t_1)\overline{\mathcal{F}}_V\,\Delta t, \tag{12}$$

and hence from Eq. (6) the temperature anomaly due to the volcanic forcing is

$$\Delta T_V(t) = \overline{\mathcal{F}}_V \int_{-\infty}^{t} G(t-t_1)\,dt_1 = \overline{\mathcal{F}}_V \int_0^{\infty} G(s)\,ds \stackrel{\text{def}}{=} -\Delta T_{\text{volc}}. \tag{13}$$

This result is meaningful only if the integral $\int_0^\infty G(s)\,ds$ is finite, i.e., if power-law response kernels are properly truncated. The obvious, but still interesting, observation is that volcanic forcing keeps the temperature, when averaged over the time scale Δt, on a constant level $T_{\text{eq}} - \Delta T_{\text{volc}}$, i.e., the time-averged temperature is ΔT_{volc} lower than the temperature at which the climate system is in equilibrium during times with no volcanic forcing.

Assume some additional (e.g., anthropogenic) forcing $\mathcal{F}_A(t)$, for which $\mathcal{F}_A = 0$ for $t \leq 0$. Then the total temperature anomaly for $t > 0$ would be

$$\Delta T(t) = \Delta T_V(t) + \Delta T_A(t) = -\Delta T_{\text{volc}} + \int_0^t G(t-t')\mathcal{F}_A(t')\,dt', \tag{14}$$

implying that the temperature starts changing in response to this forcing from a non-equilibrium initial state. However, the statistics of volcanic forcing is more challenging than assumed above, and one has to consider the possibility of long periods with zero forcing, longer than the largest temperature relaxation time reflected in the response function $G(t)$. If such a quiet period starts at time t_q, then the temperature for $t > t_q$ is

$$\Delta T(t) = \overline{\mathcal{F}}_V \int_{t-t_q}^{\infty} G(s)\,ds + \int_0^t G(t-t')\mathcal{F}_A(t')\,dt', \tag{15}$$

and since the integral over the tail of $G(s)$ is assumed to be finite (there exists a maximum relaxation time constant τ_{max}) the first term on the right of Eq. (15) will vanish if $t > t_q + \tau_{\text{max}}$. In other words, if the time of observation has been preceded by a very long period of weak volcanic forcing the additionally forced temperature change may be unaffected by the non-equilibrium imposed by volcanic forcing. If we consider, as another example, that "normal" volcanic forcing is resumed at $t = 0$ after a pause of length $|t_q| > \tau_{\text{max}}$, then ΔT according to Eq. (15) grows from zero towards the expression in Eq. (14) as t grows beyond t_{max}. Hence, during the transient period $t \in (0, \tau_{\text{max}})$ there may be a volcanic cooling that counteracts anthropogenic warming, provided there has been long pause in volcanic forcing preceding the era of anthropogenic forcing.

The discussion made here serves to illustrate that the non-equilibrium of the radiative flux balance at $t = 0$ may influence the subsequent temperature evolution, and that volcanic forcing may be the source of such an imbalance. Knowledge about the of the history of volcanic forcing in the time interval $(-\tau_{\max}, t)$ can be helpful in assessing the influence of volcanic forcing on the long-term temperature evolution in the anthropocene. In the present paper the implicit assumption has been made that Eq. (14) is valid, i.e., that there is no long pause in volcanic forcing in the period extending from $1880 - \tau_{\text{max}}$ to 2200 CE. Hence this forcing only represents a constant downshift of the temperature. This assumption may be worth closer scrutiny.