

Can we define climate by means of an ensemble? A tale of time scales of convergence

Gábor Drótos^{1,2,3} and Tamás Bódai^{4,5,6}

¹MTA-ELTE Theoretical Physics Research Group, Pázmány Péter sétány 1/A, H-1117 Budapest, Hungary

²Instituto de Física Interdisciplinar y Sistemas Complejos (CSIC-UIB), Campus UIB, Carretera de Valldemossa, km 7,5, E-07122 Palma de Mallorca, Spain

³HUN-REN Institute for Nuclear Research, Bem tér 18/C, H-4026 Debrecen, Hungary

⁴Pusan National University, Busandaehak-ro 63beon-gil 2 (Jangeon-dong), Geumjeong-gu, 46241 Busan, Republic of Korea

⁵Center for Climate Physics, Institute for Basic Science, Busandaehak-ro 63beon-gil 2 (Jangeon-dong), Geumjeong-gu, 46241 Busan, Republic of Korea

⁶Department of Applied Statistics, Institute for Mathematics and Basic Sciences, Hungarian University of Agriculture and Life Sciences, Budapest, Hungary

Correspondence: Gábor Drótos (drotos@general.elte.hu)

Abstract. It is not really questioned today any more that climate can be described in theory by an ensemble of trajectories differing in their initial conditions, which is then translated to numerical ensembles in climate models. It is also widely accepted that any evolution observed within a few decades after initialization is not yet relevant to climate. Evolution at a later stage, instead, is then used to characterize climate and its change, under the implicit assumption that slower processes do not considerably contribute to differences between ensemble members, letting internal variability of climate be identified with these differences. However, a careful justification for this practice is as yet lacking. In particular, a definition of climate in support of this practice is outstanding, including the identification of the kind of time scales at play through providing an argumentation for their relevance. Our study aims at filling this gap. After pointing out that the most important criterion for a definition of climate is the uniqueness of the probability measure on which the definition relies, we first recall the naive proposal to represent such a probability measure by the distribution of ensemble members that has, loosely speaking, converged to the natural probability measure of the so-called snapshot or pullback attractor of the dynamics. We then consider the time scales of convergence and refine the proposal by taking a probability measure that is conditional on the (possibly time-evolving) state of modes characterized by convergence time scales longer than the horizon of a particular study. We design an ensemble simulation initialization scheme for studying convergence time scales and uniqueness of ensembles in Earth system models.

1 Motivation

In roughly the past two decades, $\mathcal{O}(10)$ – $\mathcal{O}(100)$ -member initial-condition ensemble simulations have become increasingly popular in single state-of-the-art Earth system models (e.g., Kay et al., 2015; Maher et al., 2019; Rodgers et al., 2021), and this trend is expected to continue (e.g., Mankin et al., 2020). The aim of these projects is the identification of forced responses and the exploration of the internal variability of climate under explicit time dependence (which the word ‘forcing’ will refer to

20 throughout this article). The motivation for this kind of investigation is based on the naive recognition (e.g., Deser et al., 2020, and references therein) that the Earth system (and models thereof) permits a plethora of equally plausible states (e.g., weather configurations) at any time instant, let this instant occur even during an active forcing scenario¹. From this point of view, the ensembles in question are intended to represent that set of states of the system; for this purpose, their different members have been generated with the same forcing scenario but stem from different initial conditions. Climatological mean values are
25 then identified with the ensemble mean, and it is more and more widespread to describe the internal variability of climate by further statistical quantifiers evaluated with respect to the ensemble. In turn, the time evolution of all these statistical quantifiers (including but not restricted to the mean) is usually identified with their forced response.

One assumption behind these interpretations is that climate, as a statistical description of weather (IPCC, 2021), is represented by the distribution of the ensemble members; if one likes, this may be regarded as a definition of climate (at least in the
30 given models; and this definition is then certainly more useful than those relying on temporal averaging). However, no careful justification for this assumption has been provided to date. In fact, not even a framework or construction has been identified that would allow for such a justification from “first principles”. The aim of the present article is to identify one; i.e., not a justification, but a way of analysing the system, an already existing mathematical description of probability measures associated with the system, that makes justification qualitatively well-posed.

35 We emphasize that we are looking for a kind of *ab initio* description. While no model can resolve all of the processes relevant to climate, the reasonably good performance of global climate or Earth system models in describing the real world supports the view that the climate system or the Earth system is basically a nonautonomous dissipative deterministic dynamical system; genuine stochasticity may enter on the smallest spatial and temporal scales only and is thus not directly relevant to describing the climate of the system. On the other hand, as these models are the most realistic ones that exist today, it
40 is a reasonable approach to try to define climate within these models, for which themselves the properties just mentioned hold. We are envisioning a mathematical analysis that takes into account all degrees of freedom, without any approximation, simplification or effective description, perhaps with the exception of stochastic processes on the smallest scales.

Building blocks of this analysis existed in the past. On the one hand, the authors of the present article, in collaboration with numerous coauthors, have already drawn attention in this context to the convergence process to the dynamics’ snapshot or pull-
45 back attractor under explicit time dependence (i.e., a forcing scenario) (Drótos et al., 2015; Herein et al., 2016; Drótos et al., 2017; Tél et al., 2020). On the other hand, one may prefer to distinguish between time scales corresponding to internal variability of climate and those on which relevant changes (which are possibly still due to internal dynamical processes) are interpreted as changes of climate itself; this distinction has been present in the literature for roughly half a century even in the context of ensemble interpretations (Lorenz, 1975; Leith, 1975a; Hasselmann, 1976). While this distinction is obviously an essential part
50 of any useful definition of climate in the context of human societies, it is usually *a priori* imposed on a system in existing studies rather than *a posteriori* identified.² To our knowledge, it has nowhere in the literature been specified in terms of the system’s

¹We shall use the term ‘scenario’ to refer to the form of time dependence of the forcing, irrespective of whether past or future forcing is considered.

²See e.g. Del Sarto and Flandoli (2024) as a particularly sophisticated example where the possibility of making this distinction, under the term ‘Hasselmann’s proposal’, is actually argued for in its Sections IIA-D, then a discussion very similar to ours follows; however, that study begins with *assuming* the

(close-to) *fundamental* dynamics what kind of time scales one refers to, what they characterize, and why that characterisation is relevant. We will see in the present article that our snapshot/pullback approach is suitable for addressing this issue.

Similarly to earlier work of ours, we will be guided by the uniqueness of the distribution underlying a possible definition of climate. In case this distribution is not unique but depends on personal choices (e.g., on the precise time and way of initialization), every research group creating an ensemble simulation may happen to define its own climate even using the same model and forcing scenario within the same predictable context (which we will define and explain later): such a climate is not characteristic to the system, the forcing scenario and perhaps other objective factors that define the predictable context, alone. Due to the dependence on the aforementioned personal choices, it is a “personal climate”, subjective in its nature, which should be avoided (cf. Werndl, 2019). We will assess in this article whether an “objective climate”, i.e., uniqueness can in principle be encountered, under which circumstances it appears in theory, and how we should be able to decide if it happens so in practice.

A historical overview of conceptual studies about an ensemble approach for representing climate is provided in Appendix A. By the end of this appendix, we arrive at our mentioned collaborations (Drótos et al., 2015; Herein et al., 2016; Drótos et al., 2017; Tél et al., 2020), which illustrate that uniqueness naturally emerges as ensembles converge (apparently, even forward in time) to the dynamics’ snapshot or pullback attractor. A tutorial discussion of the role of chaotic internal variability and the emergence of uniqueness in its statistics as a result of this convergence (which corresponds to forgetting the distribution of initial conditions) is given in Appendix C, which concludes in a naive proposal for defining climate.

In particular and in harmony with earlier work, this proposal would be to define climate at any time instant as the natural probability measure of the snapshot/pullback attractor of the underlying dynamics (including forcing) corresponding to that time instant, and temporal aspects of climate according to the underlying trajectories. We reiterate here that the mentioned probability measure is just traced out in a model by an initial-condition ensemble of numerical trajectories after some convergence time.

And we emphasize at this point that the above-mentioned issue of time scales poses a caveat about this naive definition. Namely, unpredictable long-time-scale internal variability of the full system is normally desirable to be excluded [cf. the Edinburgh effect in Stainforth (2023) and Box 1 in Lucarini and Chekroun (2023)]. To overcome this caveat, we propose an improved definition in this article, which relies on a hypothetical probability measure that is conditioned on the state of modes identified to be slow in terms of convergence. Whether this improved definition is directly applicable to the real Earth system and its fully coupled models will depend on the degree of separation between these time scales (i.e., the characteristic time scales of forgetting the distribution of initial conditions); an affirmative answer may perhaps be expected from existing analyses (e.g., Li and Jarvis, 2009; Olivié et al., 2012) which suggest a possible gap between convergence taking place in a few decades and a few centuries. In case the separation proves to be too small, our improved definition may still serve as a guidance for how to theoretically treat probabilistic aspects and how to design practically useful constructions. A closer look at the concept of forced response will also be taken, highlighting that a climate conditioned on the state of slow modes can change without the introduction of a forcing.

existence of variables *evolving* “appreciably” on different time scales, and neither it empirically verifies this assumption nor it argues for it based on the *fundamental* equations of motion governing the system.

85 Note that the process relevant to defining climate, as we will see, is described by terms of a sum with each term (apparently
and expectedly) converging in (forward) time in an exponential-like manner, so that any general conclusions of this study,
which do not consider particular models and particular variables, must remain qualitative and leave room for more concrete
findings resulting from follow-up work. We also point out here that we do not regard the task of the present article to formulate
mathematically precise statements up to all detail; we rather intend to sketch a roadmap leading to a mathematically sound
90 possible definition of climate along with providing instructions for a practical assessment in existing models.

2 A conditional definition

We will now lay down the basics of constructing a satisfactory definition of climate under realistic circumstances but still based
on the idea of convergence which makes memory about how initial conditions are distributed fade out. Realistic circumstances
mean a realistic range of time scales as numerical modeling indicates.

95 In numerical modeling, the natural probability density is approximately represented by the distribution of a finite-size ensemble.
While considerations for constructing the naive definition were based on the numerical example of our PlaSim configuration,
any global climate model or Earth system model shares the most relevant properties: they are nonautonomous dissipative
dynamical systems. Due to discretization, they have a finite number of variables. However, on the one hand, it is plausible to
think that a system described by partial differential equations should be obtained as a limit of infinitely many ordinary differ-
100 ential equations and thus exhibits qualitatively similar behavior; and, on the other hand, spatial autocorrelation may lead to an
effective number of degrees of freedom of a finite value. As a consequence, our theoretical considerations can presumably be
extended both to other models and even to the real Earth system. Cf. Lucarini and Chekroun (2023).

From the point of view of practical numerical modeling, however, we encounter a warning. In Appendix C, one may notice
the nonzero difference between the light blue ensemble and the natural probability density, represented by the dark gray en-
105 semble, up to some time after initialization in Fig. C2. As a universal implication, before convergence to the natural probability
density takes place (up to an accuracy prescribed with regard to some practical perspective), a numerical ensemble of trajec-
tories may *not* be generally expected to represent the statistical properties relevant for climate according to the criteria discussed
above (Drótos et al., 2017). Climate-related studies must, instead, utilize ensembles for which an appropriate convergence has
taken place. [With a term frequently used by the modeling community, the ‘ensemble spread’ must already correctly describe
110 the relevant probability density, with an accuracy imposed by some practical aspect; cf. Deser et al. (2025).]

This is the point where we also need to revisit the issue of what is relevant. If the convergence to the natural probability
density is longer than the horizon of a particular study, the natural probability density is obviously not relevant from a practical
point of view. In other words, taking into account the full (unpredictable) variability introduced by slow processes is not
desirable (remember the Edinburgh effect in Stainforth, 2023). Now, climate research *is* affected by this issue: as a prominent
115 example in the real Earth system and in fully coupled models thereof, time scales up to the order of 1000 years may appear
in association with the deep ocean (e.g., Li and Jarvis, 2009), while studies with immediate practical relevance concentrate on
the last and forthcoming few centuries.

A resolution to the issue *may* be provided depending on properties of this system; the conceptual novelty of our present work lies in an explicit analysis and conclusion about the sound nature of such a possibility. For the resolution, one needs to consider
120 properties of the process of the convergence itself.

Without the need to address all of the mathematical details, we discuss in Appendix D, with reference to the spectral theory of transfer operators (Györfyi and Szépfalusy, 1988; Lasota and Mackey, 1994; Keller and Liverani, 1999; Dellnitz et al., 2000; Blank et al., 2002; Froyland et al., 2010; Chekroun et al., 2014; Slegers, 2019; Chekroun et al., 2020; Navarra et al., 2021), that the convergence itself has its own characteristic time scales describing exponentially decaying contributions in
125 autonomous dynamical systems, and a generalization exists for nonautonomous dynamical systems. In particular, these time scales are provided by the real parts of eigenvalues of the Ruelle–Perron–Frobenius operator of an autonomous dynamics. These exponentially decaying individual contributions represent loss of memory about the distribution of the initial conditions of an (infinitely large) ensemble of trajectories (which is equivalent to the development of *unpredictability*).³ Then, if one such contribution “completely” decays while the next contribution has not yet started to decay “considerably”, *uniqueness* in the
130 distribution of ensemble members is reached in this time window. This is what *may* open the possibility to give a satisfactory definition of climate. The conditions are the existence of sufficiently well-defined time scales also in the nonautonomous dynamics under question and a sufficient separation of some of these time scales. The former condition may be weakened by requiring decay to be *bounded* by one with a well-defined time scale.

Note that the mentioned eigenvalues of the Ruelle–Perron–Frobenius operator have imaginary parts as well. They and their
135 generalized counterparts in nonautonomous systems describe *predictable* (kind of oscillatory) time evolution. As a consequence, one may generally observe “strong” time evolution, i.e., rapidly changing variable values or a rapid exploration of distant regions of the underlying attractor, even in association with processes or modes (reflected in the corresponding eigenfunctions in an autonomous case) that are associated with slow convergence (according to the real part of the eigenvalues in the autonomous case).

Now, how to define climate according to the above? In an ideal case, there would be an infinitely large separation or gap
140 between two consecutive time scales of convergence, and the time frame targeted by a climate study would just lie in between. In this case, one should consider a single realization in the state (‘slow variables’) of the modes falling on the upper side of the gap (‘slower-converging modes’, in what follows) and ensure complete decay in the other modes (‘faster-converging modes’). This would lead to a *conditional* but within that still *unique* probability measure and a corresponding *conditional definition* of
145 climate, conditioned on the chosen realization of the slower-converging modes (cf. Hawkins et al., 2016). Then we can use the very same idea if the mentioned gap is not infinitely but “sufficiently” large; see a bit later.

Note that any possible predictable time evolution in the slower-converging modes is irrelevant from the point of view of the possibility to give a sound definition. In the simplest case when there is no such evolution, the relevant probability measure would be the natural probability measure of the subsystem obtained by fixing the values of the slow variables [already proposed

³Loss of memory is reflected, for instance, in temporal autocorrelation functions, with the warning that it has two time arguments in a nonautonomous case (Tél et al., 2020); see also Herein et al. (2023) for a discussion of the rate of growth of uncertainty, which is, at the same time, described in our context by the time scales of convergence as per Appendix D.

150 in Drótos et al. (2015)]. However, this may not be expected to be the case in general, so that allowing for the time evolution of the corresponding slow variables will be necessary; the particular realization of this time evolution is what we refer to as the ‘predictable context’.

In the case of time-evolving slow variables, the probability measure defining climate evolves itself in time, even in the absence of forcing (i.e., explicit time dependence in the equations of motion). Note that explicitly identifying slow variables is not at all necessary; in case initialization is performed by applying a small perturbation to a (model) state (phase space position) 155 involving any arbitrary variables, we will end up with the desired probability measure after decay in the faster-converging modes. (At wish, one can perhaps call this measure a ‘conditional natural measure’.)

The only point still requiring attention is which (model) state one should choose for initialization in order to find the “relevant climate”. In most studies, this may be based on instrumental observations [cf. Stainforth et al. (2007); Daron and Stainforth 160 (2013); DelSole and Tippett (2018); de Melo Viríssimo et al. (2024) and Appendix B] and, in terms of an ensemble, it may correspond to one single or several but well-localized values of the slow variables. Achieving an appropriate representation of instrumental observations is, however, technically non-trivial, cf. the phenomenon of initialization shock (e.g., Doblas-Reyes et al., 2011).

We now return to the practical meaning of ‘unique’ and ‘sufficient’. Inevitably, they should be defined through some practical 165 criterion for accuracy, as even a fast convergence never becomes fully complete but only in an asymptotic limit (assuming that such a limit is meaningful, cf. Footnote 11). Here, we need to distinguish between two aspects: what kind of quantifier is used, and what level of accuracy is prescribed. For the former, any quantifier of distance between probability densities may be considered; e.g., absolute difference in ensemble mean or standard deviation (or in fact any statistical moment), a Wasserstein distance (Panaretos and Zemel, 2019), the Kullback–Leibler divergence (in fact not a metric) or some related 170 quantifier (van Erven and Harremos, 2014), or even some hypothesis test statistic as in de Melo Viríssimo et al. (2024). The choice among these, as well as the level of accuracy, however, may perhaps need to be determined case-by-case in particular applications. Notwithstanding, the exponential-like nature of convergence is expected to make the corresponding definition of climate rather robust against these choices. Actually, the distance from a suitably defined probability density may generally be thought to become negligible in practice after a few times the longest relevant time scale (approximate or, preferably, bounding 175 e -folding time) passes from initialization.

We must acknowledge that the separation between time scales may perhaps happen not to be sufficiently large in practice. Then it might not be possible to define a conditional probability measure that would be unique, neither mathematically nor in practice. DelSole and Tippett (2018) recognize the problem of possibly unseparated time scales and suggest to resolve it by conditioning the definition of climate on *observations* of the past; see Appendix B why this may not be fully satisfactory. See 180 further discussion in Section 4.

A further problem is posed by regime behavior (e.g., Franzke et al., 2015) in the slow variables even if the separation of time scales of convergence is sufficiently large from a practical point of view for most initializations. In particular, if initialization takes place during or shortly before a regime transition (such that convergence does not become “complete” before the regime transition), the evolving probability density may become strongly dependent on its initial condition even on a short time scale

185 (in the sense of how it is distributed between the two regimes, i.e., how many ensemble members fall into one regime or the other in a numerical investigation); and, according to the slow convergence, uniqueness may only be reached on the time scale of the slow variables. Thereby, uniqueness and thus a sound definition of climate will be lost on a short time scale. See Appendix E for a numerical illustration in a two-variable toy model where internal variability is modeled by stochastic terms.⁴⁵

If slower-converging modes have an influence on the time evolution of the practically relevant probability density, i.e., one that has already “completely” converged, we must emphasize that this is, if uniqueness is preserved, a climate change (according to the conditional definition of climate) but is not (entirely) a forced response (as it does not originate from an explicit time dependence of the equations of motion, at least alone). That is, climate change and forced response delineate in such situations. A forced response can only be identified relative to the time evolution observed in the absence of any time dependence in the relevant terms of the equations of motion. This is analogous to the practice of trying to remove spurious model drift (Gupta et al., 2013) in order to accurately determine trends of forced change. The implicit notion of an unforced climate change induced by variations in slower system components already exists, see, e.g., the relatively early work by Hawkins et al. (2016), especially points 2 and 4 in their summary section. Note that Hawkins et al. (2016) also contains hints to possible answers to our questions about predictability.

3 ~~The conditional definition in the Earth system~~

200 ~~The main focus of our discussion here will concern the real Earth system and its realistic models from the point of view of the possible conditional definition of climate introduced in the previous section.~~

~~We begin with acknowledging that we have no definite conclusion at present about the possibility of applying this definition. This is so because a major separation of time scales of convergence is not obvious in the real Earth system and in realistic models thereof; which, among other reasons, is so because time scales specifically characterizing convergence have rarely been assessed in these systems.~~

205 ~~Having said that, we illustrate through the example of the ocean that our concept nevertheless gives guidance already to decide what can possibly be regarded as climate and what cannot in a given context. For an appropriate assessment of the issue,~~

⁴An even further issue can be the presence of multiple stationary chaotic attractors with more or less intertwined basins of attraction, possibly including riddled basins (Alexander et al., 1992), and rate-dependent tipping (Ashwin et al., 2012) between them. In such a case, different initial densities will not converge even if their difference is relatively small, i.e., uniqueness will again be lost. This setup may be responsible for the experience of Vannitsem et al. (2021) which does suggest sensitive dependence of the resulting distribution to the choice of the initial distribution. See later our assessment of the issue in practice.

⁵One should also note that a chaotic snapshot attractor can split in the case of an underlying rate-dependent tipping (Kaszás et al., 2019), and only one of the branches will be relevant if observations for initialization are available after the splitting. Although the individual branches do not have their own, separate basins of attraction in the infinitely remote past, one can select those ensemble members that end up on the relevant branch, as in Kaszás et al. (2019). Even if the splitting takes place in the future, it is more natural to regard the two branches as two separate climates. The probability of ending up on a given branch is determined by the natural measure (or its conditional variant) at the time of the splitting, provided that initialization is performed sufficiently far in the past before the splitting. Otherwise, the determination of the probabilities is not unique.

we need to reiterate from Section 2 that covering “complete” variability must always be defined in terms of some practical criterion, according to the exponential-like nature of convergence.

210 Studies about *response time scales* in the ocean (which is a different but related property) suggest, on the one hand, that time scales associated with the mixed layer and with layers below the thermocline are separated by a factor of about 10, and there might be no other characteristic time scale in between (e.g., Li and Jarvis, 2009; Oliv   et al., 2012).⁶ Importantly, dynamical modes such as the Atlantic Meridional Overturning Circulation (AMOC; Buckley and Marshall, 2016) or the Atlantic Multidecadal Oscillation/Atlantic Multidecadal Variability (Delworth et al. (e.g., 2007) do not seem to introduce complications, see later. Such a separation might or might not be just enough to treat the deep ocean separately from the rest of the system: by the time “complete” variability would unfold in the mixed layer (in terms of its own time scales), possibly unpredictable variations in the deep ocean might just become large enough to be relevant for an appropriate description.

On the other hand, one characteristic time scale of the unfolding of variability can be supposed to be a few decades according to the results from Li and Jarvis (2009); Oliv   et al. (2012), which has recently been explicitly confirmed by Deser et al. (2025) with regard to AMOC: the authors report a *convergence* time (until convergence becomes practically complete) up to about 40 years. Therefore, it is clear that the variability associated with this convergence time *at least* needs to be regarded as the internal (chaotic) variability of climate, at least in studies concentrating on time spans around a century as pointed out by Deser et al. (2025) as well. In terms of the definition of DelSole and Tippet   (2018), a lead time of a few decades is required at least, and any shorter choice will inhibit a satisfactory interpretation of corresponding results.

225 Whether the unfolding of (unpredictable) variability on longer time scales is required to be taken into account remains an open question, according to what has been previously discussed. While a separation by a factor of 10 may rather suggest an answer of ‘no’, it is unclear why the intermediate time scale of the Atlantic Multidecadal Oscillation/Atlantic Multidecadal Variability, if such a mode exists (Vineze and J  nosi, 2011; Mann et al., 2021), does not appear in the referenced analyses and what role it plays in the unfolding of variability. We emphasize it here that the numerically observable time scales may depend on the particular choice of a variable.

230 Actually, one may also think of situations with little correlation between fluctuations of different system components, for which an example might be the relationship between the deep ocean and the surface-related processes, in which case the unfolding of deep oceanic variability would be irrelevant for most observables of practical interest. According to recent research (Singh et al., 2022), there are some signs that suggest this to be so for most of the globe but not for the Southern Ocean.

235 To be precise, time scales of convergence may not necessarily be associated with given system components or variables considered explicitly in the equations of motion. In this respect as well, one should rely on the framework discussed in Appendix D. This concerns also the degree of correlation between the internal variability of different system components; although, as explained in the previous Section, this is presumably relevant only to how a (model) state should be chosen for initialization (attending to the predictable context).

⁶Note that an apparent continuous dependence of the time scale on depth (Yang and Zhu, 2011) may well be a spurious result of an interplay between the two mentioned time scales. Furthermore, the smaller separation identified by Hogan and Sriver (2019) may originate from a suboptimal partitioning of the water column from the presently discussed point of view.

240 One also has to consider the possibility that the real Earth system or some realistic model thereof does *not* meet the prerequisites of the description of convergence described in Appendix D. Long-term persistence (polynomially decaying autocorrelation) and, more generally, scaling of fluctuations in climate-related time series are reviewed by Franzke et al. (2020); see Lovejoy (2023) for more details about the phenomenology as applied to and observed in the present context. As pointed out in sections 2.3 and 5 of Franzke et al. (2020), long-term persistence and scaling may be illusory or they may result from external forcing, which would be in harmony with the Markovian nature of (most of the) equations of motion.⁶ In any case, if there is a break in the power spectrum such that there is a non-scaling regime of considerable length beyond the break (as, e.g., in Vinceze and János, 2011), the decomposition of the convergence to (a generalization of) eigenmodes (as detailed in Appendix D) may retain its pertinence. While we have not identified relevant numerical modeling studies affected by the scaling issue, and the real-world Earth system may also be expected to be “well-behaved” at least in the sense last mentioned, real-world observational data as well as some theoretical aspects warn about possible complications (Franzke et al., 2020; Lovejoy, 2023). Careful future work, well beyond the scope of the present article, should address the question to what extent scaling laws in fluctuations have a spurious origin and which temporal or spatial scales are affected, and why, by scaling laws genuinely arising from the Earth system’s dynamics, and why model studies remain unaffected if they do so indeed.

Before summarizing, we need to take care about the issue of regime behavior as raised in Section 2. Actually, the absence of different qualitative behaviors in global variables between members of currently existing large ensembles suggests that such an effect is restricted to particular system components at most. Such an effect might nevertheless appear in some slow system components possible examples are related to Labrador Sea ice cover (Danabasoglu et al., 2020) and Southern Ocean variability (Gnanadesikan et al., 2020), and uniqueness might or might not be lost in these cases: it depends on the way or the time of initialization. Cf. the illustration in Appendix E. Notably, such an effect should be easy to identify, e.g. following our proposal in Section 4.2.1.⁶

Taken together, we can say that distinguishing between different time scales of convergence appears to be reasonable to be expected to provide a sound definition of climate in the sense of Section 2: *from the point of view of century-long investigations, we might expect to be able to meaningfully define climate through the probability measure obtained after a convergence time of a few decades* perhaps up to four, cf. Deser et al. (2025). Further gaps in the spectrum of convergence time scales may possibly give rise to sound definitions for investigations on other time scales as well, but whether such gaps exist is a fully open question at present.

While the potential relevance of separation between time scales in formulating a practically satisfying definition of climate was already discussed by Lorenz (1975) and was maybe born together with the very first attempts to define climate, please note that time scales of *convergence* are concerned in the conditional definition outlined in [Section 2](#)[this section](#). As a consequence, this definition allows external forcing to induce climate changes on *arbitrarily short* time scales (e.g., after volcanic eruptions),

⁶Note that the so-called Hurst effect does not imply long-term persistence (Franzke et al., 2015).

⁶Intertwined basins of attraction and transitions between basins would have an effect very similar to that of regime transitions. Accordingly, what is suggested by currently existing large ensembles applies to their issue as well; i.e., this issue may possibly be relevant but presumably in particular system components only, and it should be easy to identify. The possibility of its relevance is emphasized by the findings of Vannitsem et al. (2021), but the limited domain of the subject of that study leaves these findings in the range of “particular system components” as mentioned.

irrespective of possible slower modes of internal variability. This is different from early descriptions appearing in the literature; cf. Appendix A.

3 The conditional definition in the Earth system

275 Since our main analysis has a theoretical nature, we decided to follow requirements on article length by relegating a preliminary assessment of the issue in the real Earth system and its realistic models to Appendix F. We conclude there that we might expect to be able to meaningfully define climate in investigations of a length around a century through the probability measure obtained after a convergence time of a few decades, perhaps up to four (cf. Deser et al., 2025).

280 Besides the actual time scales, a main open question is posed by an observed scaling behavior in time series related to climate (Franzke et al., 2020; Lovejoy, 2023). Regime behavior and intertwined basins of attraction may also be relevant and call for attention.

4 Conclusions; proposal for an initialization scheme

4.1 What we can conclude about

As a main conclusion, attending to uniqueness does let one construct a possible definition of climate even in the presence of ‘slower-converging modes’ in a dynamical system like our fundamental models of the Earth system. Slower convergence is understood in comparison with a targeted time span. In fact, the meaningful nature of such a definition is due to the (expected) existence of a discrete set of modes describing the convergence of probability densities, having a characteristic convergence time each. Ensuring convergence in modes converging in a shorter time span than the targeted one is essential, while climate generally becomes conditional on the (possibly predictably time-evolving) state of the slower-converging modes (which we call the ‘predictable context’). Note that a suitable gap in the spectrum of time scales of convergence is needed for an actual possibility to construct a definition.

Regarding the context of centennial studies of the Earth system and based on modeling results available to date, an operational definition of climate might rely on a decadal-scale convergence of an ensemble⁶ within a basin of attraction to a (practically) unique but inevitably time-dependent probability density: this density could be identified with climate. In case the predictable context has a considerable impact on targeted aspects of the given probability density, which thus becomes conditional on the predictable context, this definition assumes

- (a) a sufficiently large separation between the relevant time scales of convergence, and
- (b) avoidance of a regime transition in association with the given slower-converging modes at initialization (as well as remaining sufficiently far from boundaries of basins of attraction).

For such a case, we spell out two issues that require further attention.

⁶This ensemble is infinitely large in principle but is sampled by a finite number of members in numerical modeling.

- 300 (i) For climate projections, initialization should rely on the observed state of the slower-converging modes (which is most easily achieved by using observations to initialize the whole system). Note that this is not so in current practice, but initial conditions in terms of slower-converging modes rely on one or more arbitrary time instants of a long control run: this may be problematic for the purpose of preparing climate projections regardless of how climate is defined. On the other hand, such a sampling of a control run remains important from a general point of view.
- 305 (ii) The concept of climate change may delineate from that of a forced response, i.e., climate change may become partially unforced.

We emphasize that the above definition meets our criterion about uniqueness described in Section 1: it only depends on the system, the forcing, and the predictable context. The latter is provided (in fact, together with the basin of attraction) by the state of the slower-converging modes and is objective from the point of view of century-long studies since, by definition, 310 slower-converging modes remain sufficiently *predictable* within this time span so that one can learn their state in principle. Should condition (a) be violated, observations of slower-converging modes will not provide the possibility of initializing the system in an objective way due to its sensitive dependence on initial conditions; see Appendix B.

In such a case, the notion of climate must remain subjective. Climate predictions can then, at best, be treated similarly to probabilistic (ensemble) weather forecasts (Gneiting and Raftery, 2005), i.e., probabilities can be associated with different outcomes 315 based on the (continually evolving) present state of the dynamical variables and the uncertainty in this knowledge (leading to what is discussed by Stainforth et al., 2007; Daron and Stainforth, 2013; DelSole and Tippett, 2018; de Melo Virissimo et al., 2024). However, this approach makes the comparison of past and future climates ambiguous, and impedes, e.g., a consistent evaluation of possible future climates as time passes; cf. Appendix B still. An alternative is to stick to fixed conventions and corresponding protocols for initialization.

320 We must recall once more that “sufficient” predictability in terms of slower-converging modes must allow for a “complete” convergence in faster-converging modes and during the time span of interest, where “complete” refers to an accuracy required by some practical aspect. As explained in Section 2, a high degree of robustness is expected against which practical aspect is chosen. Even if this is so, the choice of the practical aspect may be regarded as a subjective element in the definition. However, the resulting climate depends on it only in the sense that it will have an associated *precision* up to which it is defined, in the 325 sense that climates originating from different initializations may scatter within the range of accuracy up to which convergence is required to take place. This is thus an attribute through which the comparability of climates defined in different studies can be controlled or at least assessed⁷; other subjective factors do not give rise to such an attribute.⁸

Our assessment described so far assumes a kind of dynamics that does not give rise to scaling behavior in time series of its variables. As discussed in Section 3, there are some warnings that the real Earth system may be more complicated from this

⁷This is so even if different quantifiers of distance are used for defining the climates to be compared, since any quantifier can be evaluated in a posterior way.

⁸From a philosophical point of view, one could argue that fixing the time span of interest is also subjective. However, we regard it as part of the research question, similarly as the subjective choice of considering the Earth rather than some other planet. Of course, the question of how to define climate can be posed for different time spans or different planets as well. If we like, these are also attributes of a climate.

330 point of view (Franzke et al., 2020; Lovejoy, 2023), which is not obvious to explain and clearly requires further investigation; such an investigation is, however, well beyond the scope of the present article.

4.2 How to proceed to applications

4.2.1 An initialization scheme

~~Illustration of the initialization scheme proposed in the main text. The horizontal axis stands for time. The vertical axis differentiates between ensemble members which are represented by horizontal lines; these ensemble members are grouped according to the labels $Ef1$, $Ef1'$, etc. The blue realization (i.e., the control run) is unforced, whereas the red ensemble members are subject to the forcing scenario labelled by f and schematically represented in orange beneath the corresponding ensemble members. Note that the forcing scenario is continued without any interruption (e.g., any kind of restart) when initializing the ensembles with a label with a prime ($Ef1'$, etc.). Vertical light green lines indicate minor perturbations. Unforced ensembles are not covered by this schematic.~~

340 In a given model subjected to a given forcing (i.e., a given form of time dependence), such that convergence is described as in Appendix D, it is actually expected to be possible to decide if a (practically) unique probability density typically appears in a given variable or (derived) observable within a time span of interest, and to decide whether the state of slower-converging modes is relevant in an affirmative case. We ~~propose the following ensemble initialization scheme, pictured in Fig. G1, for this purpose.~~

~~In the first step, a single ensemble, $Ef1$, is initialized by small perturbations of a model state taken from a long control run. The trajectories of $Ef1$ are then integrated under the forcing of interest typically the historical forcing followed by some future scenario such as SSP3-7.0 (Gidden et al., 2019). The next step is the initialization of a second ensemble, $Ef1'$, by perturbations of a model state taken from an arbitrary member of the first ensemble, $Ef1$, similarly as in Herein et al. (2023). The delay in its initialization should be around the shortest time scale for which one may first be interested in convergence and uniqueness. Then the same procedure is repeated for a few further time instants of the control run, resulting in additional ensemble pairs ($Ef2, Ef2'$), ($Ef3, Ef3'$), etc., such that Efn is separated from $Ef(n-1)$ more than the length of the time span of interest (thereby sampling different states of modes with comparable convergence times at least). However, the pairs with subsequent indices n are different in that the delay in the initialization of Efn' with respect to Efn increases with n until the full time span of interest is approached (at least).~~

360 In terms of a given variable, uniqueness can be assessed by evaluating whether Efn' converges to Efn , for various indices n , with regard to a practical aspect. In case Efn' converges to Efn for any arbitrary n with a shorter time scale (approximate or bounding e -folding time) than the delay between Efn' and Efn for the given n , a “full” convergence with that time scale ensures uniqueness if “full” convergence is also observed up to some index m with Efm' separated from Efm by that “full” convergence time at least. The largest such m will define the longest time span (perhaps the time span of interest or, if studied, even longer) for which climate can be defined with the given convergence time.⁹ Once convergence is noticed to take considerably longer

⁹Actually, this longest time span provides an upper bound on the longest possible time span of interest that is meaningful to choose in the given study.

(for indices beyond that just mentioned largest m), it indicates unpredictability arising from modes with comparable time scales of convergence and precludes uniqueness for correspondingly long times.⁹

By intuition, we suspect that investigating convergence to an ensemble that has already begun spreading out from localized initial conditions is technically easier, especially from the point of view of slower converging modes, than to one initialized at the same time just localized around a different point in the phase space; this is one advantage of initializing Efn' from a member of Efn during an active forcing scenario rather than using a different time instant from the control run to initialize Efn' at the beginning of the scenario as well. This is also beneficial because properties of the convergence during the bulk of a forcing scenario may be different and have more practical relevance than those just after switching the scenario on. In fact, by virtue of the increase of the delay with n , different epochs within the scenario, possibly with different convergence properties, are covered by our scheme underline that no direct estimation of the spectrum of convergence time scales is needed for this purpose.

Having said that, however, we must note that our approach is not fully perfect as it does not explore such properties systematically and, at the same time, uses convergence with those possibly different properties for the assessment of the longest time span of a sound definition. In fact, it would be desirable to initialize many ensembles for each n at increasingly later time instants within the scenario, thereby introducing a new index serving for an assessment similar to but (by using later-initialized ensembles as additional references for convergence) more general than the one described earlier. Our scheme has been designed under the assumption, supported by some numerical experience (e.g., Herein et al., 2016), that convergence properties are not considerably different in different epochs of a climate forcing scenario.

Scenarios have a finite length. Even if it may reach beyond some given “time span of interest”, an assessment of arbitrary slower converging modes whether their mere existence or relevance to a given variable, implying issue (i) in an affirmative case is only possible with the help of a control run. This is what can be performed by evaluating the convergence of ensembles Efn with different indices n to each other.

All of what is discussed above needs a word of caution: certain ensembles of the many may happen to be initialized from very similar states of some mode by chance, which will make the effect of this mode perhaps even invisible when considering the convergence of these ensembles. Instead, initializing ensembles of trajectories in a carefully chosen way and comparing their convergence to each other. Without actually knowing these modes, this risk is unavoidable. It can be mitigated by increasing the number of initialized ensembles and carefully checking any suspicious case.

On the other hand, if convergence of Efn' to Efn fails for a few but only few values of n without being clustered to high n , it presumably indicates the problem of regime transitions or that of vicinity of basin boundaries; a violation of condition (b) in either case.

It is easy to amend the above scheme to study issue (ii). For this purpose, the ensembles Efn should be compared with ensembles generated with identical initialization but with no forcing (Eun ; ‘f’ and ‘u’ stand for forced and unforced, respectively). Notes should be sufficient. Initialization schemes described in the literature, however, that model drift (Gupta et al., 2013) has

⁹Beyond the interval of sufficient predictability, the “diverging” realizations of the slower converging modes define different possible climates.

395 ~~the same effect are not sufficiently sophisticated~~ from this point of view ~~as a possible predictable evolution of slower-converging~~
~~modes; but, at least, forced changes can always be discerned.~~

~~Our proposed initialization scheme is different from that of Hawkins et al. (2016), Hogan and Sriver (2019), the CanESM2~~
~~large ensemble (Kirchmeier-Young et al., 2017; Singh et al., 2022) and the CESM2 large ensemble (Rodgers et al., 2021) in~~
~~that pairs of ensembles are initialized by perturbations in association with each model state taken from the control run,~~
400 ~~thereby enabling assessment of convergence to a (possibly conditionally) unique probability density. Timing of initialization~~
~~is also important; different model states for different pairs are taken from far away time instants of the control run so that~~
~~slower-converging modes, with respect to the time span of interest, can also be studied. Most of the cited studies target a~~
~~century or so but apply a separation of just 50 years. Note that model drift (as in Singh et al., 2022) is not at all ideal for~~
~~assessing the impact of the state of slower-converging modes (the predictable context) on the climate defined after convergence~~
405 ~~in faster-converging modes, due to its inherently artificial nature.~~ In Appendix G, we describe a proposal of ours.

4.2.2 Final remarks

We emphasize here once more that the actual conclusions drawn from an investigation like the one described right above will generally depend on the particular choice of variable to study. This is a further aspect, besides the accuracy of convergence, that is determined by the particular goal of a study. This does not preclude an objective definition either; it just means that a
410 well-defined “climate of different variables” generally applies under different conditions.

As discussed earlier, what we call here a predictable context (the state of slower-converging modes) is also an objective factor setting the scene for climate. Preparing climate projections would require its incorporation into initialization according to the present state; in fact, such an initialization would be useful regardless whether climate can be defined uniquely or not. However, if appropriate observations are not available (as is presumably the case for the deep ocean), generating ensembles
415 from various such states is important for the purpose of mapping out different possibilities permitted by the system, each defining a different climate.⁹

One undesirable property of any ensemble-based definition of climate is its inaccessibility in single realizations (including the observed evolution of the real Earth system), but, strictly speaking, such an accessibility is not conceptually required in a probabilistic framework [unlike Werndl (2016) suggests].

420 By such a definition, an important deviation from standard terminology becomes necessary: as long as the underlying probability density remains (practically) unique, internal variability characterized by time scales shorter than the time scales of slow convergence will not be a source of uncertainty in the description of climate; not even when its future projections are considered. Instead, the very definition of climate should just be a full description of the (possibly conditional) statistics of the

⁹The picture could presumably be made much simpler in studies concerning time scales much longer than a century (e.g., paleoclimatic ones), in which convergence in terms of all of the relevant modes may be possible to ensure, so that a single climate can naturally be defined. At the same time, it may also be desirable to conform with definitions created with the purpose of characterizing climate and its change within a century. In fact, the time evolution of slower-converging modes can be included as a forcing if corresponding proxies are available.

system in terms of the mentioned probability density, including the statistics of this relatively “fast” internal variability just
425 mentioned.

In practice, this means that all statistical quantifiers should be evaluated with respect to the ensemble generated to sample this internal variability. This also implies that there is no need to “invent” newer and newer ensemble-based statistical quantifiers [such as correlation coefficients (Herein et al., 2017), empirical orthogonal functions (Haszpra et al., 2020), etc.] individually, one by one: instead, one should just follow this “recipe” for the evaluation of *any* quantifier. This does not exclude
430 evaluating ensemble-wise statistics of statistical quantifiers evaluated over time intervals in single realizations (“interval-wise taken” ensemble statistics in Drótos et al., 2015) — however, evaluating a statistical quantifier over a time interval and taking its ensemble mean is not the correct way to learn about the given statistical quantifier from a probabilistic point of view. This is related to the violation of Birkhoff’s ergodic theorem in a system with explicit dependence on time (Drótos et al., 2016), the manifestation of which was already recognized by Daron and Stainforth (2013); as a practical example, sources of
435 nonergodicity for teleconnections as cross-correlations are analyzed in Bódai et al. (2022).

Code and data availability. The data presented in Figs. C1 and C2 are from Drotos (2022). The data presented in Fig. E1 were created by Bodai (2022).

Appendix A: Historical overview

Traditionally, specific definitions of climate, beyond identifying it with a statistical description of the “climate system” or the
440 Earth system, were usually formulated in terms of *temporal* statistics of weather (e.g., IPCC, 2021), a practice that still determines the thinking of many researchers from various relevant fields such as mathematics (Flandoli et al., 2022) and geoscience (Nicklas et al., 2022). If no forcing acted on the Earth system, i.e., if its equations of motion (including boundary conditions) did not depend explicitly on time, the interval for evaluating temporal statistics could be extended to the infinite future, and these statistics would coincide with those evaluated with respect to an infinite-size ensemble distributed according to the natural
445 probability measure of the chaotic attractor of the system (Ott, 1993). [That is, Birkhoff’s ergodic theorem would hold; note that the Earth system is dissipative, and can generally be assumed to be chaotic (Herein et al., 2023).] The same would be true for a temporally periodic forcing handled in terms of a stroboscopic map (Tél and Gruiz, 2006). However, if the forcing is not periodic, such a construction is not available. In that case, statistical quantifiers evaluated with respect to time need not correspond to any probability measure that would be relevant to any particular time instant (Drótos et al., 2015). [This is most
450 easily seen if the system is forced by a monotonic parameter drift, cf. Jánosi and Tél (2024).]

While Lorenz (1963) recognized the phenomenon (namely chaos) that renders the evolution of the state of the system unpredictable, and a consecutive work also formulated that the permitted states of the system may be represented by an ensemble of realizations that differ in their initial conditions (Lorenz, 1975), the notion of climate in Lorenz (1975) still assumed that the statistical descriptors are approximately constant for shorter or longer times. In parallel, Hasselmann (1976) thought about a

455 statistical mechanical analogy that implied an ensemble to be relevant but without justification. Leith (1975a, b, 1978) worked out the same statistical mechanical analogy in more detail and also including the change of the ensemble statistics as a response to some external forcing (but assuming slow variation of the latter). An explicit motivation was the unpredictability of the individual realizations, but it was not discussed how the distribution of the ensemble members is determined and if it is unique.

460 Branstator and Teng (2010) revived the latter picture, allowed for variations in the forcing on any time scale, and also assumed (but not justified) the uniqueness of a distribution, already called *climatological* distribution, emerging from arbitrary initial conditions in the infinite past. This climatological distribution gave the reference basis for ensembles of trajectories (or abstract probability densities) initialized later, by which predictability was studied. Finally, DelSole and Tippet (2018) adopted the approach of Branstator and Teng (2010), discussed uniqueness in dynamical systems, and demonstrated it in a
465 stochastic model, providing thereby a well-established notion of the “standard” climatological distribution. Also interested in predictability, however, the work of DelSole and Tippet (2018) was not satisfied with this notion, and introduced a precise probabilistic framework that based a climatological distribution on observations of the more recent past. By such a definition, the unique nature of the concept of climate is again lost; see the separate Appendix B for further details.¹⁰

The same issue concerns the approach of Stainforth et al. (2007); Daron and Stainforth (2013); de Melo Viríssimo et al.
470 (2024). This work identified a changing climate early on with the image of an attractor traced out by an ensemble (Stainforth et al., 2007) and also emphasized the uniqueness of a distribution related to the attractor under stationary conditions (Stainforth et al., 2007; Daron and Stainforth, 2013). On the other hand, it aptly recognized that narrowing down initial uncertainty (the breadth of an initial distribution) in variables characterized by (some kind of) time scales that are long compared with the time span targeted by a particular study will result in a more relevant distribution than the full distribution permitted by the system
475 (Stainforth et al., 2007; Daron and Stainforth, 2013; Hawkins et al., 2016). However, it finally took “future climate as a distribution conditioned on our uncertain knowledge of the system’s current state (Stainforth et al 2007)” (Daron and Stainforth, 2013) [a view implicitly adopted in Hawkins et al. (2016) as well], similarly as DelSole and Tippet (2018), losing thus uniqueness. The most recent of this work (de Melo Viríssimo et al., 2024) still aligns with this view but in the explicit context of the snapshot/pullback framework (to be discussed below). A major critical remark, precisely that of long time scales, was included
480 here about this framework, which, however, had already been pointed out by a reviewer of Drótos et al. (2015) and which we had already addressed in an earlier preprint version of the present article (Drotos and Bodai, 2022).

Actually, Ghil et al. (2008); Chekroun et al. (2011) already drew attention to a climatological relevance of the concept of pullback attractors and the natural probability measures supported by them, corresponding to a unique probability density to which other densities converge at a given time t if their initialization time t_0 tends to $-\infty$ in a dissipative chaotic system even
485 with temporally varying parameters. This unique density was shown to be traced out by an ensemble of trajectories initialized in the remote past. However, Ghil et al. (2008); Chekroun et al. (2011) and related work (e.g., Ghil, 2014; Pierini et al., 2016) applied this concept to subsystems of the Earth, which were subject to randomly generated forcing with a fixed distribution

¹⁰Note that loss of predictability and convergence to a unique distribution are two facets of the same phenomenon: climate should, basically, be identified when predictability is lost.

instead of a drifting signal. Even the review by Ghil and Lucarini (2020) hardly mentions the latter possibility. Although Werndl (2016) applied the concept of pullback attractors to general forms of forcing, including drifting ones, that work discarded the corresponding definition of climate. Flandoli et al. (2022) did not restrict the form of forcing either, but based the analysis on a slow time dependence compared with the convergence of time averages to ensemble averages, which is generically not the case and is certainly not so during the ongoing global warming.

In fact, it was in Bódai and Tél (2012) that the applicability of the approach of Ghil et al. (2008) to the description of global climate changes was first recognized by writing that “climate change can be seen as the evolution of snapshot attractors”. (Note also that the correspondence between the rigorously defined pullback attractor and the so-called snapshot attractor, which was introduced to the physics literature by Romeiras et al. (1990), was pointed out there.) We discussed in detail in Drótos et al. (2015) that the uniqueness of the natural probability measure of a pullback or snapshot attractor (as represented by an initial-condition ensemble, although in a toy model) makes it the appropriate concept for describing the statistics, i.e., the climate, of a system forced by a drifting signal (both in terms of the “mean state” and in that of the internal variability, both of which respond to a forcing).

One of the most important numerical observations of Drótos et al. (2015), already formulated there in the text but in a preliminary form, is about a convergence towards an actual (rigorously defined) snapshot/pullback attractor during *forward time evolution*, i.e., in a push-forward sense [as opposed to the pullback sense which underlies both a pullback (Ghil et al., 2008) and a snapshot (Namenson et al., 1996) attractor and means tending to the remote past, eventually to $-\infty$, with the time of *initialization* t_0 while keeping the time instant of interest, t , fixed]. The particular observation is that forward convergence progresses in an approximately exponential way after some kind of transients, so that the actual snapshot/pullback attractor is arbitrarily approached within some “short” time. Therefore, constructing an ensemble numerically at the time instant of interest, t , such that it represents the snapshot/pullback attractor and its natural probability measure at t with a “sufficient” accuracy does *not* require initializing the ensemble in a very remote past. Instead, initializing a few approximate e -folding times earlier is usually satisfactory, and convergence is ensured by (and can be monitored during) forward time evolution.¹¹

In Herein et al. (2016) the applicability of this framework to an intermediate-complexity general circulation model was illustrated, which was also used in Herein et al. (2017) to extend the framework to variables describing spatial patterns. Drótos et al. (2017) also pointed out the practical relevance of the approach in the same model by investigating the convergence to the density of the natural probability measure from initial conditions obtained by a slight perturbation of a state of the system already located on the attractor. The conceptual power of the snapshot/pullback framework was underlined by Vincze et al. (2017) by applying it to a laboratory experiment. An overview of basics and applications of this framework to describe climate, along with discussing novel aspects, is provided in Tél et al. (2020), where a separate section is dedicated to generalizing (auto-) correlation functions to nonautonomous systems.

¹¹As a further implication, a forward-time limit of ∞ is not relevant; in fact, the forcing or the system need not even be defined in this limit.

Appendix B: Deficiencies of a definition relying on conditioning on observations

520 Conditioning the definition of climate on observations of the past may not generally be satisfactory, since it involves two kinds of ambiguity. First, such a climate depends (practically continuously) on how far in the past the observations (more precisely, initial conditions compatible with the observations) are prescribed: if the lead time may be chosen on a case-by-case basis, we will end up with “personal climates” as discussed in Section 1, and there will be no qualitative difference between the notion of climate and that of probabilistic weather forecast. Second, such a climate will depend on the precision of observations: 525 improving precision may narrow down the magnitude of its internal variability. Uniqueness, from the point of view of objective factors, will be lost due to these dependences.

These undesired dependences will disappear in practice if there is a considerable separation of time scales of convergence between the processes (modes) intended to be included in the unpredictable internal variability of climate and those intended to be excluded from it, *and* if initialization is chosen sufficiently far (relative to the fast-converging former processes) but not 530 too far (relative to the slow-converging latter processes) in the past. However, we will then recover the conditional definition introduced in Section 2, and initialization by observations (assuming it is possible) will be relevant only for the slow-converging processes (providing the predictable context). The conditional definition motivated by the concept of snapshot attractors and that of Stainforth et al. (2007); Daron and Stainforth (2013); DelSole and Tippett (2018); de Melo Viríssimo et al. (2024) based on observations become special cases of each other in this situation. Note, however, the ambiguity that may arise about how to 535 choose the predictable context if no reliable instrumental records are available. See also Section 4.2.2 in this respect.

Appendix C: A naive proposal for the definition of climate

In this section, we illustrate the basic idea behind reaching uniqueness through convergence. This will be done with the help of an intermediate-complexity climate model without fluid dynamics in the ocean, so that drawing final conclusions about the real Earth system or realistic models thereof will be impossible. The particular intermediate-complexity model is the Planet 540 Simulator (PlaSim) (Fraedrich et al., 2005) with a mixed-layer ocean. We use the same model output as in Drótos et al. (2017); please refer to this publication for more details about the configuration.

For this section, let us imagine that the dynamics of the Earth system is described perfectly by PlaSim in the mentioned configuration (including even the discretized nature of the model). This defines our dynamical system, which has a phase space of $\approx 10^5$ variables. Let our model Earth system be subject to the following forcing scenario:

$$545 \quad [\text{CO}_2](t) = \begin{cases} 360 & \text{for } t < 600, \\ 360 + 3.6(t - 600) & \text{for } t \geq 600, \end{cases} \quad (\text{C1})$$

where concentration is in ppm and time is in yr, and let us study the time evolution of the annual mean near-surface temperature at a particular grid point in the southern Pacific Ocean [similarly as in Drótos et al. (2017)]. Finally, let us suppose that our Earth system has followed the trajectory corresponding to the dark gray line of Fig. C1 up to $t_0 = 610\text{yr}$, which we identify with the present.

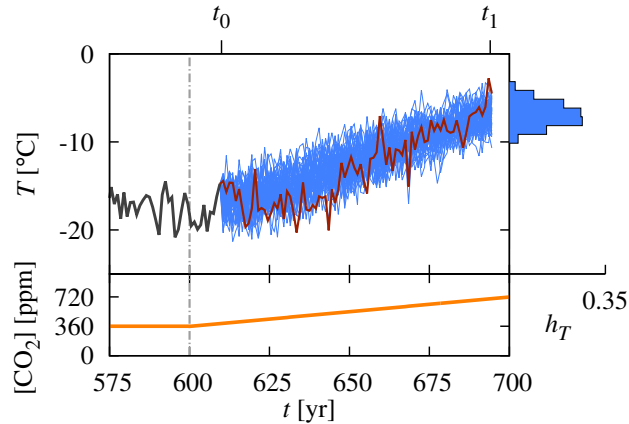


Figure C1. The annual mean near-surface temperature T of a single grid point in the southern Pacific Ocean (at 180°E and about 64°S) as a function of time. $t_0 = 610\text{yr}$ is identified with the present, and $t_1 = 694\text{yr}$ is an (almost) arbitrarily chosen year in the future. The dark gray line is a single PlaSim simulation and is regarded as the instrumental record in the model system. The red line is the continuation of the same simulation and is regarded as a prediction. The 48 light blue lines are alternative predictions, obtained from simulations initialized by slightly perturbing the surface pressure field of the simulation of the dark gray line at $t_0 = 610\text{yr}$ [see the main text and Drótos et al. (2017) for details]. The normalized histogram h_T constructed from the 48 values of the light blue lines and further 144 alternative predictions (a total of 192 values) at t_1 is shown on the right-hand-side of the main plot. The CO_2 concentration, through which the forcing scenario is defined, is also displayed (in orange). The vertical dot-dashed line in gray marks the beginning of the linear ramp in the CO_2 concentration. Data are from Drotos (2022).

550 A possible question about the *weather* of the future is, to take an (almost) arbitrarily chosen example, say, what the near-surface temperature at the given grid point will be at $t_1 = 694\text{yr}$. To obtain this temperature, we can use the $t = t_0$ phase space position of the dark gray trajectory of Fig. C1 as initial condition, and integrate PlaSim from t_0 to t_1 : the result for the near-surface temperature of the given grid point is the red trajectory shown in Fig. C1, which has, of course, a unique value at t_1 . However, this value is only one possible answer to the question.

555 Let us slightly perturb the $t = t_0$ phase space position of the dark gray trajectory 192 times to obtain 192 different initial conditions: the utilized random perturbation modifies the surface pressure field on the order of 10^{-3}hPa [see again Drótos et al. (2017) for the details], so that each perturbed initial condition can be regarded as realistic as the original one. In particular, these different initial conditions could not be distinguished by standard instrumental measurements on the real Earth (cf. de Melo Viríssimo and Stainforth, 2025). The integration of PlaSim from these initial conditions results in an ensemble of
560 trajectories, plotted in light blue in Fig. C1 (only 48 of the 192 for better visibility), each giving a separate answer for the *weather* at t_1 . All are possible answers, and the differences between them emerge as a result of the chaotic nature (unpredictable internal variability; Herein et al., 2023) of the system.

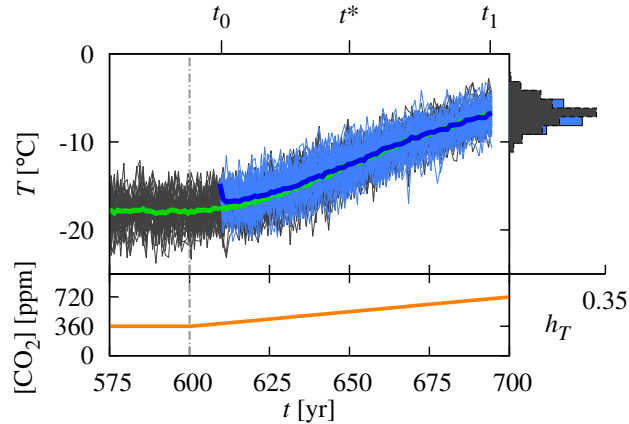


Figure C2. The light blue lines and histogram of Fig. C1 compared with an ensemble of the same size (marked by dark gray) initialized in the remote past (at $t = 0$, asymptotically far in practice from $t = 575$). The corresponding ensemble averages are also shown in dark blue and green, respectively. By t^* , the two ensemble averages practically coincide. Data are from Drotos (2022).

In fact, infinitely many of these answers would trace out a density, as indicated by the normalized histogram h_T in Fig. C1. We can ask now if this density provides a relevant characterization of the plethora of all possible near-surface temperature values at t_1 .

In particular, the question is about the uniqueness of this density. If we perturb the original initial condition at t_0 in a different way, will we end up with a different density at t_1 ? For example, if we take a larger or a smaller magnitude for the perturbation, will the final density be broader or more narrow? Or if we restrict the sign in the perturbation, will it shift the final density in some direction? Since the initial conditions can never be constrained to arbitrary precision, such perturbations can be regarded just as relevant as the original one.

In fact, for an infinitely large ensemble in a dissipative nonautonomous dynamical system exhibiting chaotic behavior (such as PlaSim), the kind of density in question corresponds to the natural probability measure of a snapshot (Romeiras et al., 1990) or pullback (Ghil et al., 2008; Chekroun et al., 2011) attractor if t_0 tends to $-\infty$, at least as long as the initial conditions remain in the basin of attraction of the same attractor¹². The density of the natural probability measure is unique, i.e., it will be the same for any generic set of infinitely numerous initial conditions within the basin of attraction of the given attractor. As we discussed in the introduction, an atmospherically motivated toy model was used in Drótos et al. (2015) to illustrate that a process of convergence to the natural probability measure (*forward in time*) is exponential-like on the long term (it is presumably faster than any power law, cf. Appendix D). In such a case, the convergence of a finite-size ensemble is practically accomplished (with “exponential precision”, according to some practical point of view; see later) within a finite amount of time. In Herein et al. (2016), this was found to be the case for PlaSim, too, and this finding was conjectured to be relevant to any global climate model or Earth system model. In our configuration of PlaSim, the convergence time proved to be a few decades.

¹²It is assumed here that the relevant basin of attraction existed in the infinitely remote past; cf. Section 2 and Footnote 5 in particular.

As a consequence, the *probability* density of the near-surface temperature at the investigated grid point in our configuration (and that of any other variable or set of variables) that is traced out by even a finite-size ensemble at t_1 will be practically the same regardless of how we choose initial conditions at t_0 — provided that two conditions are met. The first one is that t_0 has to
585 be sufficiently far, although not infinitely far, in the past from t_1 . The second is that the initial conditions must be chosen within such limits that ensure convergence to the desired attractor, avoiding a different one, e.g., a snowball Earth (cf. Kaszás et al., 2019; Ragon et al., 2022). Since t_0 and t_1 are separated by several decades in Fig. C1, this conclusion translates for our example as follows: within the mentioned limits, we cannot modify the initialization scheme at t_0 to obtain a substantially different probability density at t_1 ; that is, within a certain range of initialization schemes, we cannot “ruin” the quantitative
590 result. Among other initialization schemes (which need not rely on some particular trajectory of the system), an extremely small perturbation of the actually realized initial conditions (the $t = t_0$ phase space position corresponding to the dark gray line of Fig. C1 in our model Earth system, the equivalent of hypothetical perfect instrumental observations at t_0 in the real Earth system) will lead to this unique density, which corresponds to the natural probability measure, see Fig. C2.

These considerations suggest this unique density (the ‘natural probability density’ in what follows) to be the practically
595 relevant *a priori* (Paillard, 2008) probability density (i.e., existing independently of almost *any* observation about the system) that can be associated with the plethora of all possibilities at a future time instant t_1 that are permitted by the dynamics under the given forcing scenario (within the relevant basin of attraction, being the only observational constraint). That is, beyond being of practical interest, the natural probability density basically characterizes the *system* in statistical or probabilistic terms, instead of characterizing a particular situation (a “microstate” in a statistical mechanical analogy, e.g., a weather configuration
600 in the atmosphere¹³) permitted by the system. Note that this abstract density exists at *any* time instant: not only at t_1 , but also after and even before it (for instance, also at the time instant t_0 identified with the present, or even earlier; see the dark gray ensemble of Fig. C2). To practically obtain this density at a given time instant, one just needs to prescribe initial conditions for an ensemble in the sufficiently far past (within the relevant basin of attraction), and follow the time evolution of the ensemble until the desired time instant.

605 Of course, the natural probability density depends on time in the presence of a forcing (which is the generic situation and is so in the dynamics of the real Earth system; see again the dark gray ensemble of Fig. C2 for our model configuration). Otherwise, it is constant in time, and coincides with the natural probability density of the usual chaotic (stationary) attractor of the dynamical system (Ott, 1993).¹⁴¹⁵

¹³The prefix ‘micro’ is not to be confused with that used in the term ‘microinitialization’ where it refers to a certain type of (in particular, fast) variable which is perturbed (Stainforth et al., 2007).

¹⁴Note that the natural probability density should actually be defined in the full phase space of the system. The probability density of a given variable (such as the near-surface temperature at the selected grid point in the example of Figs. C1-C2) is the marginal density of this multivariate probability density. The full multivariate probability density carries information about the statistical relationships between different parts of the system (different variables, different geographical regions, etc.).

¹⁵As a subtlety, the natural probability measure is defined instantaneously in mathematical terms, whereas our numerical example considered an annual mean. This apparent discrepancy is easy to resolve. On the one hand, the individual time evolution of the trajectories composing the ensemble results in their phase space positions to be distributed according to the natural probability density at any time instant; on the other hand, since a particular solution of a dynamical system is unique, the time evolution of the trajectories also uniquely defines a probability density for temporal averages evaluated along these

We now recall that any definition of climate intends to capture statistical properties (including those related to temporal aspects). We also see that the relevant statistical properties of the system are, at root, described by the natural probability measure of the relevant snapshot or pullback attractor at any time instant. Therefore, *we would hereby naively suggest defining climate as the statistical properties determined by the (infinitely large) ensemble of trajectories evolving according to the natural probability measure under a given forcing scenario*. Then the expected value of some given variable will be the climatic mean value of that variable, and all higher-order moments will describe internal variability. With this definition, any particular realization of the dynamics performs a sampling of the probability density that defines climate [i.e., the ‘climatological distribution’ in the terminology of Branstator and Teng (2010)].

Appendix D: Identifying time scales of convergence

We discuss here how the spectral theory of transfer operators (Lasota and Mackey, 1994; Dellnitz et al., 2000; Slegers, 2019) should enable one to identify time scales of convergence.

Let us first consider an autonomous system, i.e., one without any explicit dependence on time, hypothetically describing a stationary climate:

$$\dot{\mathbf{x}} = \mathbf{F}(\mathbf{x}), \tag{D1}$$

where $\mathbf{x} \in X$ represents the vector composed of all dynamical variables of the system, with X being the d -dimensional phase space spanned by these vectors, and $\mathbf{F}$ defines the dynamics, which is assumed to be dissipative. Let $\mathcal{P}_{t_0}^t$ denote the Ruelle–Perron–Frobenius or transfer operator associated with the dynamics between time t_0 and $t_0 + t$ with $t \geq 0$ and defined with respect to the Lebesgue measure of X : the time evolution of probability densities f defined on X is described by the action of operators $\mathcal{P}_{t_0}^t$ on f with different values of $t \geq 0$. If the operators $\mathcal{P}_{t_0}^t$ are quasi-compact [i.e., if they have a finite number of isolated eigenvalues outside the essential radius; the universal belief that this is the typical case can be traced back to Keller and Liverani (1999); Blank et al. (2002)], the time evolution of an arbitrarily initialized probability density f from the class of square-integrable functions on X can be decomposed as

$$\mathcal{P}_{t_0}^t(f) = \sum_{i=1}^N c_i \lambda_i^{t/T} \varphi_i + r(t), \tag{D2}$$

where c_i , $i \in \{1, \dots, N\}$ are constant coefficients depending on f (such that $c_1 = 1$), T is the unit of time, N is the number of isolated eigenvalues of the operator $\mathcal{P}_{t_0}^T$, λ_i is the i th of these eigenvalues with $\lambda_1 = 1$ and $|\lambda_i| < 1$ for all $1 < i \leq N$ (note that these eigenvalues are generally complex and come in complex conjugate pairs), φ_i is the i th eigenfunction of the same operator (note that φ_1 is the density of the natural probability measure), and $r(t)$ is a residual decaying faster than $\lambda_N^{t/T}$

individual trajectories. (In fact, this is true not only for averages, but for any quantity derived from an interval of time or simply from more than one discrete time instant.) The latter construction was termed “interval-wise taken” in Drótos et al. (2015). For the annual mean near-surface temperature of our example, this density is what we called the natural probability density. We thus see that generalization to some finite time interval of interest (days, months, years, etc.) is straightforward.

(Gyögyi and Szépfalusy, 1988; Chekroun et al., 2014; Slegers, 2019; Navarra et al., 2021). If there are multiple attractors, which have separate basins of attraction in the phase space, the decomposition (D2) applies separately to densities f with an initial support falling in a single basin (Tantet et al., 2018a).

Eq. (D2) means that f converges to the density φ_1 of the corresponding natural probability measure, and the convergence
640 proceeds according to exponential terms with different time scales of decay, as per the real parts of the eigenvalues and accompanied by oscillations as determined by the imaginary parts, and a fast-decaying residual. The exponential contributions with different decay time scales may then form the basis of the conditional definition of climate: if the separation between decay time scales is sufficiently large, “complete” convergence is possible on one time scale without an influence from processes with longer convergence time scales. From a technical point of view, what remains after such a “complete” convergence is a
645 projection of the initial density f obtained by omitting the fast-converging terms in the decomposition (D2); eventually, this would be the mathematical object defining climate in an autonomous system.

A time scale of convergence may (possibly) be associated with a particular system component according to properties of the corresponding eigenfunction. Furthermore, the multivariate density shows how much fluctuations in different system components are correlated.

650 An important remark to be made here is that time evolution associated with a given eigenfunction may very well happen to be fast according to the oscillatory part, even if the corresponding decay is slow. In such a case, fast time evolution is *predictable*, as the autocorrelation function can be decomposed in terms of the same modes (one may refer, e.g., to Corollary 1 in Chekroun et al., 2020, which addresses a similar setup); cf. Section 2.

While the above considerations concern autonomous systems, generalization to periodically forced systems is easy by identifying the unit T of time with the period of the forcing. In this case, Eq. (D2) will hold for $t = nT$ with $n \in \mathbb{N}$. Similar results
655 also exist in systems with nonperiodic dependence on time (Froyland et al., 2010), but the eigenvalues are not constant in this case. Although tippings (Ashwin et al., 2012), especially crises of corresponding stationary chaotic attractors (Tantet et al., 2018a, b), may lead to complications at least for some eigenvalues, if most of the eigenvalues vary moderately enough, as might be expected under forcing scenarios relevant to century-long studies, these eigenvalues remain informative about the
660 time scales in the system (cf. Tantet et al., 2020). An attempt for computing an approximation of the spectrum of the relevant transfer operator is nevertheless beyond the scope of this article.

We mention that more sophisticated and, at the same time, more easily implementable techniques also exist for identifying some kinds of time scale separation; see, e.g., Froyland et al. (2014, 2016), and cf. Lucarini and Chekroun (2023). In contrast, our application appears to necessitate a very specific notion, for which these techniques are not sufficient: an explicit gap in the
665 real part of the eigenspectrum.

Appendix E: Regime transitions and uniqueness in slow-fast systems

For simplicity, we use a slow-fast system meant according to traditional terminology (Kuehn, 2015). In particular, we use the stochastic (in the Itô sense) equations of motion

$$dx = (-x^3 + 2x)dt + \sigma_{xx}dW_{xx,t} + \frac{1}{\epsilon}\sigma_{xy}ydt, \quad (\text{E1})$$

$$670 \quad dy = \frac{1}{\epsilon^2}(-y + cx)dt + \frac{1}{\epsilon}\sigma_{yy}dW_{yy,t}, \quad (\text{E2})$$

to demonstrate that an ensemble initialized by small perturbations during a regime transition cannot represent climate in the absence of its uniqueness on the short term. A time scale separation between the slow x and fast y variables is achieved as $\epsilon \rightarrow 0$ (Wouters and Gottwald, 2019). To represent a practical, realistic situation, we set $\epsilon = 0.2$. The slow subsystem is characterized by a symmetric quartic polynomial like in Bódai (2020). It is perturbed both by some noise ($\sigma_{xx} > 0$) and
675 the fast subsystem ($\sigma_{xy} > 0$), which latter features (internal) variability generated by white noise ($\sigma_{yy} > 0$) (Hasselmann, 1976; Wouters and Gottwald, 2019). Effectively, the time scale separation between y and some even faster z whose governing equation is eliminated is readily represented in the stochastic model. The noise $dW_{xx,t}$ is the slow subsystem’s own internal variability, unaffected by the fast subsystem. The fast variable is affected by the slow variable ($c \neq 0$), which is why when climate cannot be defined with respect to the slow variable, it carries over to the fast one. We use the parameter values of
680 $\sigma_{xx} = 0.6$, $\sigma_{xy} = 0.4$, $\sigma_{yy} = 1$, $c = 0.2$. The equations are integrated by the Euler–Maruyama integration scheme, using a time step of $\Delta t = 0.02$.

The quartic polynomial represents a double-well potential function and so gives rise to a saddle-type unstable fixed point, i.e., a saddle, at $x = y = 0$ in the unperturbed ($\sigma_{xx} = \sigma_{yy} = 0$) 2D system (E1)-(E2), also called a Melancholia state in Lucarini and Bódai (2020), whose stable manifold makes a finite angle with the $x = 0$ line owing to the coupling $\sigma_{xy} > 0$. With a weak perturbation (in the sense of Bódai, 2020), infrequent transitions between the potential wells, across the Melancholia state, take place
685 in terms of a long single-realization “control” run.

In order to examine the uniqueness of the converged ensemble, we initialize a pair of ensembles at some state of the control run. Namely, we “perturb” the fast variable as $y + \delta y$, $\delta y = +1$ for one ensemble and $\delta y = -1$ for the other ensemble.¹⁶ We do so in order to have two markedly different¹⁷ ensembles initially — the difference of the emanating densities has to disappear
690 on the fast time scale to allow for a unique definition of climate. When the dynamics is deterministic, a minute mismatch of the initial conditions within each of the ensembles — at least with respect to z — is needed for the spread of that ensemble. In our stochastic modeling, however, it is clearly not required. The various different realizations of the ensemble members are generated by various different realizations of the Wiener process $W_{yy,t}$.

Given the said inclination of the stable manifold, which is the basin boundary of the unperturbed system, the opposite
695 perturbations $y \pm \delta y$, even if taken in the y -direction only, could already achieve the placement of the two initial conditions

¹⁶Such a “perturbation” of a state for the purpose of *generating* initial conditions for trajectories is not to be confused with the “dynamic perturbation” of trajectories under the evolution equations as a result of $\sigma_{xy} > 0$, etc., in (E1)-(E2).

¹⁷Note that the standard deviation of y is comparable to $|\delta y| = 1$, and so the applied perturbations can correspond to extreme opposite, or, rather different, states of the fast process.

in the different basins of attraction, provided that the current (x, y) was not far off from the basin boundary. Then, ensuing perturbed ($\sigma_{yy} > 0$) realizations will much more likely end up in the near future in the potential well/regime where they started out from. As only the fast variable is perturbed to initialize an ensemble, the same single realization of $W_{xx,t}$ of the control run is used for all of the ensemble members. In effect, this gives rise to nonautonomous dynamics, i.e., explicit time dependence in the system. Therefore, it is the stable manifold of the corresponding snapshot saddle that will actually control the transitions (Bódai et al., 2013).

When we initialize the two ensembles during a transition (see the middle row of Fig. E1), the ensemble means of neither the fast nor the slow variable, $\langle y \rangle$ and $\langle x \rangle$, converge but remain well separated, indicating the lack of uniqueness, at least on the time scale of the fast variable y . I.e., the realized ensemble will depend on the particular initialization of the fast subsystem. This indeed prevents any of these to objectively represent climate. On the contrary, when we initialize the two ensembles sufficiently far away from a transition, whether before or right after it (see the top and bottom rows of Fig. E1, respectively), the ensemble means converge on the time scale of the fast variable, indicating uniqueness. Due to the “forcing” of the slow subsystem, $W_{xx,t}$, clusters of trajectories (a subset of all of the trajectories because of $\sigma_{xy} > 0$) suffer a regime transition in a coordinated manner, which is imprinted on the evolution of the ensemble means. That is, had the toy model reflected realistic characteristics of the climate system, regime transitions of the slow subsystem could give rise to the delineation of the concepts of climate change and forced response.

The case of $\sigma_{xx} = 0$, when the slow subsystem does not have an internal variability in isolation from the y system component, is qualitatively similar: there is no uniqueness on the fast time scale in association with initialization during a regime transition of x .

715 **Appendix F: The conditional definition in the Earth system: discussion**

The main focus of this discussion will concern the real Earth system and its realistic models from the point of view of the possible conditional definition of climate introduced in the previous section.

We begin with acknowledging that we have no definite conclusion at present about the possibility of applying this definition. This is so because a major separation of time scales of convergence is not obvious in the real Earth system and in realistic models thereof; which, among other reasons, is so because time scales specifically characterizing convergence have rarely been assessed in these systems.

Having said that, we illustrate through the example of the ocean that our concept nevertheless gives guidance already to decide what can possibly be regarded as climate and what cannot in a given context. For an appropriate assessment of the issue, we need to reiterate from Section 2 that covering “complete” variability must always be defined in terms of some practical criterion, according to the exponential-like nature of convergence.

Studies about *response time scales* in the ocean (which is a different but related property) suggest, on the one hand, that time scales associated with the mixed layer and with layers below the thermocline are separated by a factor of about 10, and there

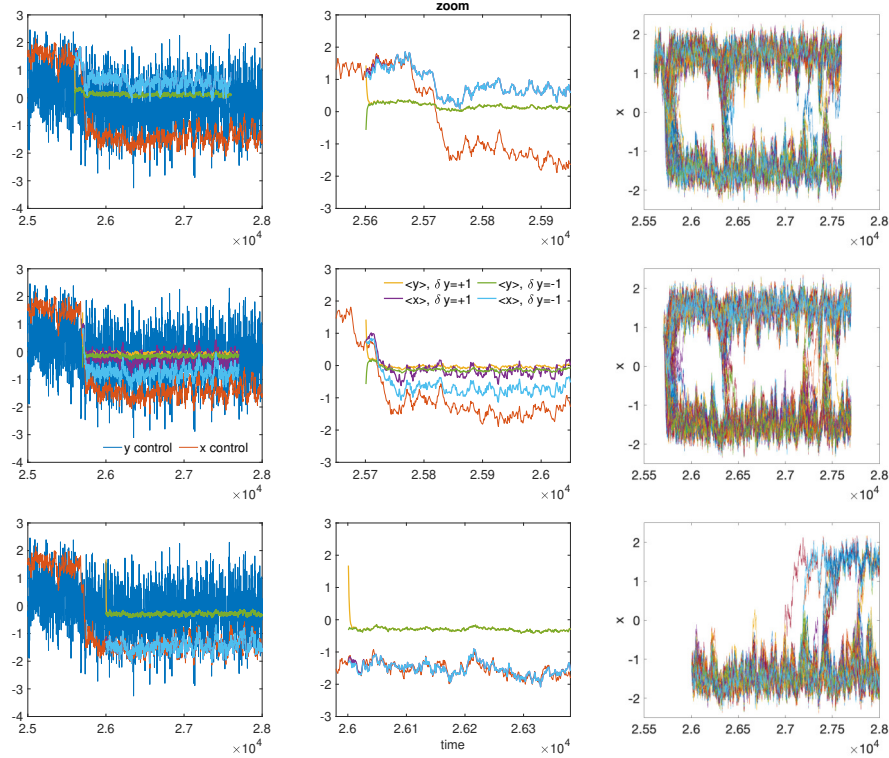


Figure E1. Initialization during and away from a regime transition precludes and supports the definition of climate, respectively. Two ensembles are initialized by perturbing the fast variable as $y + \delta y$, $\delta y = +1$ or -1 (see main text), upon which the ensemble means $\langle y \rangle$ or $\langle x \rangle$ are plotted in diagrams on the left to examine convergence, using 1000 ensemble members in each ensemble. Diagrams in the middle column provide zoomed pictures of those on the left with respect to time, excluding the y control time series, for the better visibility of uniqueness. The legend annotations in two panels apply to all in the left and middle columns. On the right, spaghetti diagrams of all of the 1000 slow time series are given for one of the ensembles ($\delta y = +1$). Data were produced by code by Bodai (2022).

might be no other characteristic time scale in between (e.g., Li and Jarvis, 2009; Olivié et al., 2012).¹⁸ Importantly, dynamical modes such as the Atlantic Meridional Overturning Circulation (AMOC; Buckley and Marshall, 2016) or the Atlantic Multidecadal Oscillation/Atlantic Multidecadal Variability Delworth et al. (e.g., 2007) do not seem to introduce complications, see later. Such a separation might or might not be just enough to treat the deep ocean separately from the rest of the system: by the time “complete” variability would unfold in the mixed layer (in terms of its own time scales), possibly unpredictable variations in the deep ocean might just become large enough to be relevant for an appropriate description.

On the other hand, one characteristic time scale of the unfolding of variability can be supposed to be a few decades according to the results from Li and Jarvis (2009); Olivié et al. (2012), which has recently been explicitly confirmed by Deser et al. (2025)

¹⁸Note that an apparent continuous dependence of the time scale on depth (Yang and Zhu, 2011) may well be a spurious result of an interplay between the two mentioned time scales. Furthermore, the smaller separation identified by Hogan and Srivier (2019) may originate from a suboptimal partitioning of the water column from the presently discussed point of view.

with regard to AMOC: the authors report a *convergence* time (until convergence becomes practically complete) up to about 40 years. Therefore, it is clear that the variability associated with this convergence time *at least* needs to be regarded as the internal (chaotic) variability of climate, at least in studies concentrating on time spans around a century [as pointed out by Deser et al. (2025) as well]. In terms of the definition of DelSole and Tippet (2018), a lead time of a few decades is required at least, and any shorter choice will inhibit a satisfactory interpretation of corresponding results.

Whether the unfolding of (unpredictable) variability on longer time scales is required to be taken into account remains an open question, according to what has been previously discussed. While a separation by a factor of 10 may rather suggest an answer of ‘no’, it is unclear why the intermediate time scale of the Atlantic Multidecadal Oscillation/Atlantic Multidecadal Variability, if such a mode exists (Vincze and János, 2011; Mann et al., 2021), does not appear in the referenced analyses and what role it plays in the unfolding of variability. We emphasize it here that the numerically observable time scales may depend on the particular choice of a variable.

Actually, one may also think of situations with little correlation between fluctuations of different system components, for which an example might be the relationship between the deep ocean and the surface-related processes, in which case the unfolding of deep oceanic variability would be irrelevant for most observables of practical interest. According to recent research (Singh et al., 2022), there are some signs that suggest this to be so for most of the globe but not for the Southern Ocean.

To be precise, time scales of convergence may not necessarily be associated with given system components or variables considered explicitly in the equations of motion. In this respect as well, one should rely on the framework discussed in Appendix D. This concerns also the degree of correlation between the internal variability of different system components; although, as explained in the previous Section, this is presumably relevant only to how a (model) state should be chosen for initialization (attending to the predictable context).

One also has to consider the possibility that the real Earth system or some realistic model thereof does *not* meet the prerequisites of the description of convergence described in Appendix D. Long-term persistence (polynomially decaying autocorrelation) and, more generally, scaling of fluctuations in climate-related time series are reviewed by Franzke et al. (2020); see Lovejoy (2023) for more details about the phenomenology as applied to and observed in the present context. As pointed out in sections 2.3 and 5 of Franzke et al. (2020), long-term persistence and scaling may be illusory or they may result from external forcing, which would be in harmony with the Markovian nature of (most of the) equations of motion.¹⁹ In any case, if there is a break in the power spectrum such that there is a non-scaling regime of considerable length beyond the break (as, e.g., in Vincze and János, 2011), the decomposition of the convergence to (a generalization of) eigenmodes (as detailed in Appendix D) may retain its pertinence. While we have not identified relevant numerical modeling studies affected by the scaling issue, and the real-world Earth system may also be expected to be “well-behaved” at least in the sense last mentioned, real-world observational data as well as some theoretical aspects warn about possible complications (Franzke et al., 2020; Lovejoy, 2023). Careful future work, well beyond the scope of the present article, should address the question to what extent scaling laws in fluctuations have a spurious origin and which temporal or spatial scales are affected, and why, by scaling laws genuinely arising from the Earth system’s dynamics, and why model studies remain unaffected if they do so indeed.

¹⁹Note that the so-called Hurst effect does not imply long-term persistence (Franzke et al., 2015).

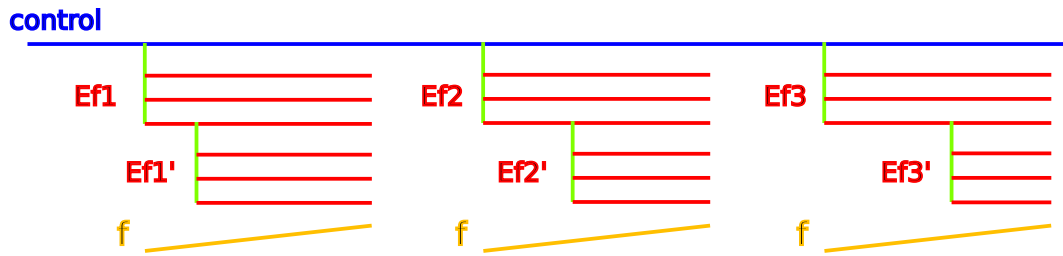


Figure G1. Illustration of the initialization scheme proposed in the main text. The horizontal axis stands for time. The vertical axis differentiates between ensemble members which are represented by horizontal lines; these ensemble members are grouped according to the labels Ef1, Ef1', etc. The blue realization (i.e., the control run) is unforced, whereas the red ensemble members are subject to the forcing scenario labelled by f and schematically represented in orange beneath the corresponding ensemble members. Note that the forcing scenario is continued without any interruption (e.g., any kind of restart) when initializing the ensembles with a label with a prime (Ef1', etc.). Vertical light green lines indicate minor perturbations. Unforced ensembles are not covered by this schematic.

770 Before summarizing, we need to take care about the issue of regime behavior as raised in Section 2. Actually, the absence of different qualitative behaviors in global variables between members of currently existing large ensembles suggests that such an effect is restricted to particular system components at most. Such an effect might nevertheless appear in some slow system components [possible examples are related to Labrador Sea ice cover (Danabasoglu et al., 2020) and Southern Ocean variability (Gnanadesikan et al., 2020)], and uniqueness might or might not be lost in these cases: it depends on the way or the

775 time of initialization. Cf. the illustration in Appendix E. Notably, such an effect should be easy to identify, e.g. following our proposal in Section 4.2.1.²⁰

Taken together, we can say that distinguishing between different time scales of convergence appears to be reasonable to be expected to provide a sound definition of climate in the sense of Section 2: *from the point of view of century-long investigations, we might expect to be able to meaningfully define climate through the probability measure obtained after a convergence time*

780 *of a few decades* [perhaps up to four, cf. Deser et al. (2025)]. Further gaps in the spectrum of convergence time scales may possibly give rise to sound definitions for investigations on other time scales as well, but whether such gaps exist is a fully open question at present.

Appendix G: An initialization scheme: description

²⁰Intertwined basins of attraction and transitions between basins would have an effect very similar to that of regime transitions. Accordingly, what is suggested by currently existing large ensembles applies to their issue as well; i.e., this issue may possibly be relevant but presumably in particular system components only, and it should be easy to identify. The possibility of its relevance is emphasized by the findings of Vannitsem et al. (2021), but the limited domain of the subject of that study leaves these findings in the range of “particular system components” as mentioned.

We propose the following ensemble initialization scheme, pictured in Fig. G1, for the purpose of deciding in a given model subjected to a given forcing whether a (practically) unique probability density typically appears in a given variable or (derived) observable within a time span of interest, and whether the state of slower-converging modes is relevant in an affirmative case

In the first step, a single ensemble, Ef_1 , is initialized by small perturbations of a model state taken from a long control run. The trajectories of Ef_1 are then integrated under the forcing of interest [typically the historical forcing followed by some future scenario such as SSP3-7.0 (Gidden et al., 2019)]. The next step is the initialization of a second ensemble, Ef_1' , by perturbations of a model state taken from an arbitrary member of the first ensemble, Ef_1 , similarly as in Herein et al. (2023). The delay in its initialization should be around the shortest time scale for which one may first be interested in convergence and uniqueness. Then the same procedure is repeated for a few further time instants of the control run, resulting in additional ensemble pairs (Ef_2, Ef_2') , (Ef_3, Ef_3') , etc., such that Ef_n is separated from $Ef_{(n-1)}$ more than the length of the time span of interest (thereby sampling different states of modes with comparable convergence times at least). However, the pairs with subsequent indices n are different in that the delay in the initialization of Ef_n' with respect to Ef_n increases with n until the full time span of interest is approached (at least).

In terms of a given variable, uniqueness can be assessed by evaluating whether Ef_n' converges to Ef_n , for various indices n , with regard to a practical aspect. In case Ef_n' converges to Ef_n for any arbitrary n with a shorter time scale (approximate or bounding e -folding time) than the delay between Ef_n' and Ef_n for the given n , a “full” convergence with that time scale ensures uniqueness if “full” convergence is also observed up to some index m with Ef_m' separated from Ef_m by that “full” convergence time at least. The largest such m will define the longest time span (perhaps the time span of interest or, if studied, even longer) for which climate can be defined with the given convergence time.²¹ Once convergence is noticed to take considerably longer (for indices beyond that just-mentioned largest m), it indicates unpredictability arising from modes with comparable time scales of convergence and precludes uniqueness for correspondingly long times.²²

By intuition, we suspect that investigating convergence to an ensemble that has already begun spreading out from localized initial conditions is technically easier, especially from the point of view of slower-converging modes, than to one initialized at the same time just localized around a different point in the phase space; this is one advantage of initializing Ef_n' from a member of Ef_n during an active forcing scenario rather than using a different time instant from the control run to initialize Ef_n' at the beginning of the scenario as well. This is also beneficial because properties of the convergence during the bulk of a forcing scenario may be different and have more practical relevance than those just after switching the scenario on. In fact, by virtue of the increase of the delay with n , different epochs within the scenario, possibly with different convergence properties, are covered by our scheme.

Having said that, however, we must note that our approach is not fully perfect as it does not explore such properties systematically and, at the same time, uses convergence with those possibly different properties for the assessment of the longest time span of a sound definition. In fact, it would be desirable to initialize many ensembles for each n at increasingly later time instants within the scenario, thereby introducing a new index serving for an assessment similar to but (by using

²¹ Actually, this longest time span provides an upper bound on the longest possible time span of interest that is meaningful to choose in the given study.

²² Beyond the interval of sufficient predictability, the “diverging” realizations of the slower-converging modes define different possible climates.

later-initialized ensembles as additional references for convergence) more general than the one described earlier. Our scheme has been designed under the assumption, supported by some numerical experience (e.g., Herein et al., 2016), that convergence properties are not considerably different in different epochs of a climate forcing scenario.

820 Scenarios have a finite length. Even if it may reach beyond some given “time span of interest”, an assessment of arbitrary slower-converging modes [whether their mere existence or relevance to a given variable, implying issue (i) in an affirmative case] is only possible with the help of a control run. This is what can be performed by evaluating the convergence of ensembles Efn with different indices n to each other.

825 All of what is discussed above needs a word of caution: certain ensembles of the many may happen to be initialized from very similar states of some mode by chance, which will make the effect of this mode perhaps even invisible when considering the convergence of these ensembles to each other. Without actually knowing these modes, this risk is unavoidable. It can be mitigated by increasing the number of initialized ensembles and carefully checking any suspicious case.

830 On the other hand, if convergence of Efn' to Efn fails for a few but only few values of n without being clustered to high n , it presumably indicates the problem of regime transitions [or that of vicinity of basin boundaries; a violation of condition (b) in either case].

It is easy to amend the above scheme to study issue (ii). For this purpose, the ensembles Efn should be compared with ensembles generated with identical initialization but with no forcing (Eun ; ‘f’ and ‘u’ stand for forced and unforced, respectively). Note, however, that model drift (Gupta et al., 2013) has the same effect from this point of view as a possible predictable evolution of slower-converging modes; but, at least, forced changes can always be discerned.

835 Our proposed initialization scheme is different from that of Hawkins et al. (2016), Hogan and Sriver (2019), the CanESM2 large ensemble (Kirchmeier-Young et al., 2017; Singh et al., 2022) and the CESM2 large ensemble (Rodgers et al., 2021) in that *pairs* of ensembles are initialized by perturbations in association with each model state taken from the control run, thereby enabling assessment of convergence to a (possibly conditionally) unique probability density. Timing of initialization is also important; different model states for different pairs are taken from far away time instants of the control run so that 840 slower-converging modes, with respect to the time span of interest, can also be studied. Most of the cited studies target a century or so but apply a separation of just 50 years. Note that model drift (as in Singh et al., 2022) is not at all ideal for assessing the impact of the state of slower-converging modes (the predictable context) on the climate defined after convergence in faster-converging modes, due to its inherently artificial nature.

Author contributions. Conceptualisation: GD; every other task: shared in various proportions.

845 *Competing interests.* The authors declare that they have no conflict of interest.

Acknowledgements. T. Tél deserves special thanks for his role in earlier work forming the basis of the present article. The contribution of M. Herein to preparing and running the numerical simulations is gratefully acknowledged. Kind feedback from the peer reviewers and the handling editor was essential to improve presentation and to appropriately describe open questions. GD is indebted to G. Froyland for his invaluable help by providing detailed explanations with regard to the spectral theory of transfer operators. Important insights into this theory
850 are also due to the kind help of M. Chekroun and G. Gottwald. Zs. Mihálka interpreted some of the above information to GD, which is highly appreciated. Useful discussions with C. Franzke, E. Hernández-García, A. Navarra, K. Rehfeld, B. Sándor, D. Stainforth, A. Tantet and T. Tél are acknowledged as well, along with helpful comments on the manuscript by B. Kaszás, T. Tél and M. Vincze. GD acknowledges financial support from the National Research, Development and Innovation Office (NKFIH, Hungary) under grant K125171, as well as from the European Union (European Social Fund and European Social Fund Plus) and the Government of the Balearic Islands through the
855 Margalida Comas and Vicenç Mut postdoctoral fellowships (grant numbers PD/020/2018 and PD-035-2023, respectively). TB was supported by the Institute for Basic Science (IBS) under grant IBS-R028-Y1 and by the Research Excellence Programme of the Hungarian University of Agriculture and Life Sciences.

References

- Alexander, J., Yorke, J., You, Z., and Kan, I.: Riddled basins, *International Journal of Bifurcation and Chaos*, 02, 795–813, <https://doi.org/10.1142/S0218127492000446>, 1992.
- Ashwin, P., Wieczorek, S., Vitolo, R., and Cox, P.: Tipping points in open systems: bifurcation, noise-induced and rate-dependent examples in the climate system, *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences*, 370, 1166–1184, <https://doi.org/10.1098/rsta.2011.0306>, 2012.
- Blank, M., Keller, G., and Liverani, C.: Ruelle–Perron–Frobenius spectrum for Anosov maps, *Nonlinearity*, 15, 1905, <https://doi.org/10.1088/0951-7715/15/6/309>, 2002.
- Bódai, T.: An Efficient Algorithm to Estimate the Potential Barrier Height from Noise-Induced Escape Time Data, *Journal of Statistical Physics*, 179, 1625–1636, <https://doi.org/10.1007/s10955-020-02574-4>, 2020.
- Bodai, T.: Regime transitions and uniqueness of ensembles in slow-fast systems. <https://doi.org/10.5281/zenodo.7277017>, <https://doi.org/10.5281/zenodo.7272665>, 2022.
- Bódai, T. and Tél, T.: Annual variability in a conceptual climate model: Snapshot attractors, hysteresis in extreme events, and climate sensitivity, *Chaos*, 22, 023 110, 2012.
- Bódai, T., Altmann, E. G., and Endler, A.: Stochastic perturbations in open chaotic systems: Random versus noisy maps, *Phys. Rev. E*, 87, 042 902, <https://doi.org/10.1103/PhysRevE.87.042902>, 2013.
- Branstator, G. and Teng, H.: Two Limits of Initial-Value Decadal Predictability in a CGCM, *Journal of Climate*, 23, 6292 – 6311, <https://doi.org/10.1175/2010JCLI3678.1>, 2010.
- Buckley, M. W. and Marshall, J.: Observations, inferences, and mechanisms of the Atlantic Meridional Overturning Circulation: A review, *Reviews of Geophysics*, 54, 5–63, <https://doi.org/10.1002/2015RG000493>, 2016.
- Bódai, T., Lee, J.-Y., and Aneesh, S.: Sources of Nonergodicity for Teleconnections as Cross-Correlations, *Geophysical Research Letters*, 49, e2021GL096 587, <https://doi.org/10.1029/2021GL096587>, 2022.
- Chekroun, M. D., Simonnet, E., and Ghil, M.: Stochastic climate dynamics: Random attractors and time-dependent invariant measures, *Physica D*, 240, 1685–1700, 2011.
- Chekroun, M. D., Neelin, J. D., Kondrashov, D., McWilliams, J. C., and Ghil, M.: Rough parameter dependence in climate models and the role of Ruelle–Pollicott resonances, *Proceedings of the National Academy of Sciences*, 111, 1684–1690, <https://doi.org/10.1073/pnas.1321816111>, 2014.
- Chekroun, M. D., Tantet, A., Dijkstra, H. A., and Neelin, J. D.: Ruelle–Pollicott Resonances of Stochastic Systems in Reduced State Space. Part I: Theory, *J. Stat. Phys.*, 179, 1366–1402, <https://doi.org/10.1007/s10955-020-02535-x>, 2020.
- Danabasoglu, G., Lamarque, J.-F., Bacmeister, J., Bailey, D. A., DuVivier, A. K., Edwards, J., Emmons, L. K., Fasullo, J., Garcia, R., Gettelman, A., Hannay, C., Holland, M. M., Large, W. G., Lauritzen, P. H., Lawrence, D. M., Lenaerts, J. T. M., Lindsay, K., Lipscomb, W. H., Mills, M. J., Neale, R., Oleson, K. W., Otto-Bliesner, B., Phillips, A. S., Sacks, W., Tilmes, S., van Kampenhout, L., Vertenstein, M., Bertini, A., Dennis, J., Deser, C., Fischer, C., Fox-Kemper, B., Kay, J. E., Kinnison, D., Kushner, P. J., Larson, V. E., Long, M. C., Mickelson, S., Moore, J. K., Nienhouse, E., Polvani, L., Rasch, P. J., and Strand, W. G.: The Community Earth System Model Version 2 (CESM2), *Journal of Advances in Modeling Earth Systems*, 12, e2019MS001 916, <https://doi.org/10.1029/2019MS001916>, 2020.

- Daron, J. D. and Stainforth, D. A.: On predicting climate under climate change, *Environmental Research Letters*, 8, 034021, <https://doi.org/10.1088/1748-9326/8/3/034021>, 2013.
- de Melo Viríssimo, F. and Stainforth, D. A.: Micro- and Macroparametric Uncertainty in Climate Change Prediction: A Large Ensemble Perspective, *Bulletin of the American Meteorological Society*, 106, E1319 – E1341, <https://doi.org/10.1175/BAMS-D-24-0064.1>, 2025.
- de Melo Viríssimo, F., Stainforth, D. A., and Bröcker, J.: The evolution of a non-autonomous chaotic system under non-periodic forcing: A climate change example, *Chaos: An Interdisciplinary Journal of Nonlinear Science*, 34, 013 136, <https://doi.org/10.1063/5.0180870>, 2024.
- Del Sarto, G. and Flandoli, F.: A non-autonomous framework for climate change and extreme weather events increase in a stochastic energy balance model, *Chaos: An Interdisciplinary Journal of Nonlinear Science*, 34, 093 122, <https://doi.org/10.1063/5.0223309>, 2024.
- Dellnitz, M., Froyland, G., and Sertl, S.: On the isolated spectrum of the Perron-Frobenius operator, *Nonlinearity*, 13, 1171, <https://doi.org/10.1088/0951-7715/13/4/310>, 2000.
- DelSole, T. and Tippet, M. K.: Predictability in a changing climate, *Climate Dynamics*, 51, 531–545, <https://doi.org/10.1007/s00382-017-3939-8>, 2018.
- Delworth, T. L., Zhang, R., and Mann, M. E.: Decadal to Centennial Variability of the Atlantic from Observations and Models, pp. 131–148, American Geophysical Union (AGU), <https://doi.org/https://doi.org/10.1029/173GM10>, 2007.
- Deser, C., Lehner, F., Rodgers, K. B., Ault, T., Delworth, T. L., DiNezio, P. N., Fiore, A., Frankignoul, C., Fyfe, J. C., Horton, D. E., Kay, J. E., Knutti, R., Lovenduski, N. S., Marotzke, J., McKinnon, K. A., Minobe, S., Randerson, J., Screen, J. A., Simpson, I. R., and Ting, M.: Insights from Earth system model initial-condition large ensembles and future prospects, *Nature Climate Change*, 10, 277–286, <https://doi.org/10.1038/s41558-020-0731-2>, 2020.
- Deser, C., Kim, W. M., Wills, R. C., Simpson, I. R., Yeager, S., Danabasoglu, G., Rodgers, K., and Rosenbloom, N.: Effects of macro vs. micro initialization and ocean initial-condition memory on the evolution of ensemble spread in the CESM2 large ensemble, *Climate Dynamics*, 63, 62, <https://doi.org/10.1007/s00382-024-07553-z>, 2025.
- Doblas-Reyes, F. J., Balmaseda, M. A., Weisheimer, A., and Palmer, T. N.: Decadal climate prediction with the European Centre for Medium-Range Weather Forecasts coupled forecast system: Impact of ocean observations, *Journal of Geophysical Research: Atmospheres*, 116, D19 111, <https://doi.org/https://doi.org/10.1029/2010JD015394>, 2011.
- Drotos, G.: The convergence of an ensemble to an attractor during climate change. <https://doi.org/10.5281/zenodo.7277017>, <https://doi.org/10.5281/zenodo.7277017>, 2022.
- Drotos, G. and Bodai, T.: On defining climate by means of an ensemble, *ESS Open Archive*, <https://doi.org/10.1002/essoar.10510833.3>, 2022.
- Drótos, G., Bódai, T., and Tél, T.: Probabilistic Concepts in a Changing Climate: A Snapshot Attractor Picture, *Journal of Climate*, 28, 3275–3288, 2015.
- Drótos, G., Bódai, T., and Tél, T.: Quantifying nonergodicity in nonautonomous dissipative dynamical systems: An application to climate change, *Phys. Rev. E*, 94, 022 214, <https://doi.org/10.1103/PhysRevE.94.022214>, 2016.
- Drótos, G., Bódai, T., and Tél, T.: On the importance of the convergence to climate attractors, *The European Physical Journal Special Topics*, 226, 2031–2038, <https://doi.org/10.1140/epjst/e2017-70045-7>, 2017.
- Flandoli, F., Pappalettera, U., and Tonello, E.: Nonautonomous attractors and Young measures, *Stochastics and Dynamics*, 22, 2240 003, <https://doi.org/10.1142/S0219493722400032>, 2022.
- Fraedrich, K., Jansen, H., Kirk, E., Luksch, U., and Lunkeit, F.: The Planet Simulator: Towards a user friendly model, *Meteorol. Z.*, 14, 299–304, 2005.

- Franzke, C. L. E., Osprey, S. M., Davini, P., and Watkins, N. W.: A Dynamical Systems Explanation of the Hurst Effect and Atmospheric Low-Frequency Variability, *Scientific Reports*, 5, 9068, <https://doi.org/10.1038/srep09068>, 2015.
- 935 Franzke, C. L. E., Barbosa, S., Blender, R., Fredriksen, H.-B., Laepple, T., Lambert, F., Nilsen, T., Rypdal, K., Rypdal, M., Scotto, M. G., Vannitsem, S., Watkins, N. W., Yang, L., and Yuan, N.: The Structure of Climate Variability Across Scales, *Reviews of Geophysics*, 58, e2019RG000 657, <https://doi.org/https://doi.org/10.1029/2019RG000657>, 2020.
- Froyland, G., Lloyd, S., and Santitissadeekorn, N.: Coherent sets for nonautonomous dynamical systems, *Physica D: Nonlinear Phenomena*, 239, 1527–1541, <https://doi.org/https://doi.org/10.1016/j.physd.2010.03.009>, 2010.
- 940 Froyland, G., Gottwald, G. A., and Hammerlindl, A.: A Computational Method to Extract Macroscopic Variables and Their Dynamics in Multiscale Systems, *SIAM Journal on Applied Dynamical Systems*, 13, 1816–1846, <https://doi.org/10.1137/130943637>, 2014.
- Froyland, G., Gottwald, G. A., and Hammerlindl, A.: A trajectory-free framework for analysing multiscale systems, *Physica D: Nonlinear Phenomena*, 328-329, 34–43, <https://doi.org/https://doi.org/10.1016/j.physd.2016.04.010>, 2016.
- Ghil, M.: A mathematical theory of climate sensitivity or, How to deal with both anthropogenic forcing and natural variability?, in: *Climate Change: Multidecadal and Beyond*, edited by Chang, C. P., Ghil, M., Latif, M., and Wallace, J. M., World Scientific Publ. Co./Imperial College Press, London, UK, 2014.
- 945 Ghil, M. and Lucarini, V.: The physics of climate variability and climate change, *Rev. Mod. Phys.*, 92, 035 002, <https://doi.org/10.1103/RevModPhys.92.035002>, 2020.
- Ghil, M., Chekroun, M. D., and Simonnet, E.: Climate dynamics and fluid mechanics: Natural variability and related uncertainties, *Physica D*, 237, 2111–2126, 2008.
- 950 Gidden, M. J., Riahi, K., Smith, S. J., Fujimori, S., Luderer, G., Kriegler, E., van Vuuren, D. P., van den Berg, M., Feng, L., Klein, D., Calvin, K., Doelman, J. C., Frank, S., Fricko, O., Harmsen, M., Hasegawa, T., Havlik, P., Hilaire, J., Hoesly, R., Horing, J., Popp, A., Stehfest, E., and Takahashi, K.: Global emissions pathways under different socioeconomic scenarios for use in CMIP6: a dataset of harmonized emissions trajectories through the end of the century, *Geoscientific Model Development*, 12, 1443–1475, <https://doi.org/10.5194/gmd-12-1443-2019>, 2019.
- 955 Gnanadesikan, A., Speller, C. M., Ringlein, G., Soucie, J. S., Thomas, J., and Pradal, M.-A.: Feedbacks Driving Interdecadal Variability in Southern Ocean Convection in Climate Models: A Coupled Oscillator Mechanism, *Journal of Physical Oceanography*, 50, 2227 – 2249, <https://doi.org/10.1175/JPO-D-20-0037.1>, 2020.
- Gneiting, T. and Raftery, A. E.: Weather Forecasting with Ensemble Methods, *Science*, 310, 248–249, <https://doi.org/10.1126/science.1115255>, 2005.
- 960 Gupta, A. S., Jourdain, N. C., Brown, J. N., and Monselesan, D.: Climate Drift in the CMIP5 Models, *Journal of Climate*, 26, 8597–8615, <https://doi.org/10.1175/JCLI-D-12-00521.1>, 2013.
- Györgyi, G. and Szépfalussy, P.: Relaxation processes in chaotic states of one dimensional maps, *Acta Physica Hungarica*, 64, 33–48, <https://doi.org/10.1007/BF03158518>, 1988.
- Hasselmann, K.: Stochastic climate models, Part 1, Theory, *Tellus*, 28, 473–485, 1976.
- 965 Haszpra, T., Topál, D., and Herein, M.: On the Time Evolution of the Arctic Oscillation and Related Wintertime Phenomena under Different Forcing Scenarios in an Ensemble Approach, *J. Climate*, 33, 3107–3124, <https://doi.org/10.1175/JCLI-D-19-0004.1>, 2020.
- Hawkins, E., Smith, R. S., Gregory, J. M., and Stainforth, D. A.: Irreducible uncertainty in near-term climate projections, *Climate Dynamics*, 46, 3807–3819, <https://doi.org/10.1007/s00382-015-2806-8>, 2016.

- Herein, M., Márffy, J., Drótos, G., and Tél, T.: Probabilistic concepts in intermediate-complexity climate models: A snapshot attractor picture, *J. Climate*, 29, 259–272, <https://doi.org/10.1175/JCLI-D-15-0353.1>, 2016.
- Herein, M., Drótos, G., Haszpra, T., Márffy, J., and Tél, T.: The theory of parallel climate realizations as a new framework for teleconnection analysis, *Scientific Reports*, 7, 44 529, <https://doi.org/10.1038/srep44529>, 2017.
- Herein, M., Tél, T., and Haszpra, T.: Where are the coexisting parallel climates? Large ensemble climate projections from the point of view of chaos theory, *Chaos: An Interdisciplinary Journal of Nonlinear Science*, 33, <https://doi.org/10.1063/5.0136719>, 031104, 2023.
- 975 Hogan, E. and Sriver, R. L.: The Effect of Internal Variability on Ocean Temperature Adjustment in a Low-Resolution CESM Initial Condition Ensemble, *Journal of Geophysical Research: Oceans*, 124, 1063–1073, <https://doi.org/10.1029/2018JC014535>, 2019.
- IPCC: Annex VII: Glossary, in: *Climate Change 2021: The Physical Science Basis. Contribution of Working Group I to the Sixth Assessment Report of the Intergovernmental Panel on Climate Change*, edited by Matthews, J., Möller, V., van Diemen, R., Fuglestedt, J., Masson-Delmotte, V., Méndez, C., Semenov, S., and Reisinger, A., p. 2215–2256, Cambridge University Press, Cambridge, United Kingdom and
- 980 New York, NY, USA, <https://doi.org/10.1017/9781009157896.022>, 2021.
- Jánosi, D. and Tél, T.: Overview of the advances in understanding chaos in low-dimensional dynamical systems subjected to parameter drift: Parallel dynamical evolutions and “climate change” in simple systems, *Physics Reports*, 1092, 1–64, <https://doi.org/https://doi.org/10.1016/j.physrep.2024.09.003>, 2024.
- Kaszás, B., Haszpra, T., and Herein, M.: The snowball Earth transition in a climate model with drifting parameters: Splitting of the snapshot
- 985 attractor, *Chaos: An Interdisciplinary Journal of Nonlinear Science*, 29, 113 102, <https://doi.org/10.1063/1.5108837>, 2019.
- Kay, J. E., Deser, C., Phillips, A., Mai, A., Hannay, C., Strand, G., Arblaster, J. M., Bates, S. C., Danabasoglu, G., Edwards, J., Holland, M., Kushner, P., Lamarque, J.-F., Lawrence, D., Lindsay, K., Middleton, A., Munoz, E., Neale, R., Oleson, K., Polvani, L., and Vertenstein, M.: The Community Earth System Model (CESM) Large Ensemble Project: A Community Resource for Studying Climate Change in the Presence of Internal Climate Variability, *Bulletin of the American Meteorological Society*, 96, 1333 – 1349, <https://doi.org/10.1175/BAMS-D-13-00255.1>, 2015.
- 990 Keller, G. and Liverani, C.: Stability of the spectrum for transfer operators, *Ann. Scuola Norm. Sup. Pisa Cl. Sci.*, 28, 141–152, 1999.
- Kirchmeier-Young, M. C., Zwiers, F. W., and Gillett, N. P.: Attribution of Extreme Events in Arctic Sea Ice Extent, *Journal of Climate*, 30, 553 – 571, <https://doi.org/10.1175/JCLI-D-16-0412.1>, 2017.
- Kuehn, C.: *Multiple Time Scale Dynamics*, Springer, Cham, Germany, <https://doi.org/https://doi.org/10.1007/978-3-319-12316-5>, 2015.
- 995 Lasota, A. and Mackey, M. C.: *Chaos, Fractals and Noise*, Springer, Berlin, Germany, 1994.
- Leith, C. E.: Appendix 2.2: The design of a statistical-dynamical climate model and statistical constraints on the predictability of climate, in: *The physical basis of climate and climate modeling: report of the International Study Conference in Stockholm*, 29 July - 10 August 1974, edited by Unknown, vol. 16 of *GARP Publications Series*, pp. 137–141, WMO, Geneva, Switzerland, https://library.wmo.int/index.php?lvl=notice_display&id=6943, 1975a.
- 1000 Leith, C. E.: Climate Response and Fluctuation Dissipation, *Journal of Atmospheric Sciences*, 32, 2022–2026, [https://doi.org/10.1175/1520-0469\(1975\)032<2022:CRAFD>2.0.CO;2](https://doi.org/10.1175/1520-0469(1975)032<2022:CRAFD>2.0.CO;2), 1975b.
- Leith, C. E.: Predictability of climate, *Nature*, 276, 352–355, <https://doi.org/10.1038/276352a0>, 1978.
- Li, S. and Jarvis, A.: Long run surface temperature dynamics of an A-OGCM: the HadCM3 4×CO₂ forcing experiment revisited, *Climate Dynamics*, 33, 817–825, <https://doi.org/10.1007/s00382-009-0581-0>, 2009.
- 1005 Lorenz, E. N.: Deterministic Nonperiodic Flow, *Journal of Atmospheric Sciences*, 20, 130 – 141, [https://doi.org/10.1175/1520-0469\(1963\)020<0130:DNF>2.0.CO;2](https://doi.org/10.1175/1520-0469(1963)020<0130:DNF>2.0.CO;2), 1963.

- Lorenz, E. N.: Appendix 2.1: Climatic predictability, in: The physical basis of climate and climate modeling: report of the International Study Conference in Stockholm, 29 July - 10 August 1974, edited by Unknown, vol. 16 of *GARP Publications Series*, pp. 132–136, WMO, Geneva, Switzerland, https://library.wmo.int/index.php?lvl=notice_display&id=6943, 1975.
- 1010 Lovejoy, S.: Review article: Scaling, dynamical regimes, and stratification. How long does weather last? How big is a cloud?, *Nonlinear Processes in Geophysics*, 30, 311–374, <https://doi.org/10.5194/npg-30-311-2023>, 2023.
- Lucarini, V. and Bódai, T.: Global stability properties of the climate: Melancholia states, invariant measures, and phase transitions, *Nonlinearity*, 33, R59–R92, <https://doi.org/10.1088/1361-6544/ab86cc>, 2020.
- Lucarini, V. and Chekroun, M. D.: Theoretical tools for understanding the climate crisis from Hasselmann’s programme and beyond, *Nat. Rev. Phys.*, 5, 744–765, <https://doi.org/10.1038/s42254-023-00650-8>, 2023.
- 1015 Maher, N., Milinski, S., Suarez-Gutierrez, L., Botzet, M., Dobrynin, M., Kornblueh, L., Kröger, J., Takano, Y., Ghosh, R., Hedemann, C., Li, C., Li, H., Manzini, E., Notz, D., Putrasahan, D., Boysen, L., Claussen, M., Ilyina, T., Olonscheck, D., Raddatz, T., Stevens, B., and Marotzke, J.: The Max Planck Institute Grand Ensemble: Enabling the Exploration of Climate System Variability, *Journal of Advances in Modeling Earth Systems*, 11, 2050–2069, <https://doi.org/https://doi.org/10.1029/2019MS001639>, 2019.
- 1020 Mankin, J. S., Lehner, F., Coats, S., and McKinnon, K. A.: The Value of Initial Condition Large Ensembles to Robust Adaptation Decision-Making, *Earth’s Future*, 8, e2012EF001 610, <https://doi.org/https://doi.org/10.1029/2020EF001610>, 2020.
- Mann, M. E., Steinman, B. A., Brouillette, D. J., and Miller, S. K.: Multidecadal climate oscillations during the past millennium driven by volcanic forcing, *Science*, 371, 1014–1019, <https://doi.org/10.1126/science.abc5810>, 2021.
- Namenson, A., Ott, E., and Antonsen, T. M.: Fractal dimension fluctuations for snapshot attractors of random maps, *Phys. Rev. E*, 53, 2287–2291, <https://doi.org/10.1103/PhysRevE.53.2287>, 1996.
- 1025 Navarra, A., Tribbia, J., and Klus, S.: Estimation of Koopman Transfer Operators for the Equatorial Pacific SST, *Journal of the Atmospheric Sciences*, 78, 1227 – 1244, <https://doi.org/10.1175/JAS-D-20-0136.1>, 2021.
- Nicklas, J. M., Fox-Kemper, B., and Lawrence, C. E.: A Novel Definition of Climate State Using Kalman Filtering and Application to Thresholds, *Earth ArXiv*, <https://doi.org/10.31223/X5FH2C>, 2022.
- 1030 Olivé, D. J. L., Peters, G. P., and Saint-Martin, D.: Atmosphere Response Time Scales Estimated from AOGCM Experiments, *Journal of Climate*, 25, 7956 – 7972, <https://doi.org/10.1175/JCLI-D-11-00475.1>, 2012.
- Ott, E.: *Chaos in Dynamical Systems*, Cambridge University Press, Cambridge, UK, 1993.
- Paillard, D.: From atmosphere, to climate, to Earth system science, *Interdisciplinary Science Reviews*, 33, 25–35, 2008.
- Panaretos, V. M. and Zemel, Y.: Statistical Aspects of Wasserstein Distances, *Annual Review of Statistics and Its Application*, 6, 405–431, <https://doi.org/https://doi.org/10.1146/annurev-statistics-030718-104938>, 2019.
- 1035 Pierini, S., Ghil, M., and Chekroun, M. D.: Exploring the Pullback Attractors of a Low-Order Quasigeostrophic Ocean Model: The Deterministic Case, *Journal of Climate*, 29, 4185 – 4202, <https://doi.org/10.1175/JCLI-D-15-0848.1>, 2016.
- Ragon, C., Lembo, V., Lucarini, V., Vêrad, C., Kasparian, J., and Brunetti, M.: Robustness of Competing Climatic States, *Journal of Climate*, 35, 2769 – 2784, <https://doi.org/10.1175/JCLI-D-21-0148.1>, 2022.
- 1040 Rodgers, K. B., Lee, S.-S., Rosenbloom, N., Timmermann, A., Danabasoglu, G., Deser, C., Edwards, J., Kim, J.-E., Simpson, I. R., Stein, K., Stuecker, M. F., Yamaguchi, R., Bódai, T., Chung, E.-S., Huang, L., Kim, W. M., Lamarque, J.-F., Lombardozzi, D. L., Wieder, W. R., and Yeager, S. G.: Ubiquity of human-induced changes in climate variability, *Earth System Dynamics*, 12, 1393–1411, <https://doi.org/10.5194/esd-12-1393-2021>, 2021.
- Romeiras, F. J., Grebogi, C., and Ott, E.: Multifractal properties of snapshot attractors of random maps, *Phys. Rev. A*, 41, 784, 1990.

- 1045 Singh, H. K., Goldenson, N., Fyfe, J. C., and Polvani, L. M.: Uncertainty in Pre-industrial Global Ocean Initialization Can Yield Irreducible Uncertainty in Southern Ocean Surface Climate, *Journal of Climate*, pp. 1 – 55, <https://doi.org/10.1175/JCLI-D-21-0176.1>, 2022.
- Slegers, W.: Spectral Theory for Perron-Frobenius operators, Master’s thesis, Uppsala University, Applied Mathematics and Statistics, 2019.
- Stainforth, D., Allen, M., Tredger, E., and Smith, L.: Confidence, uncertainty and decision-support relevance in climate predictions, *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences*, 365, 2145–2161, <https://doi.org/10.1098/rsta.2007.2074>, 2007.
- 1050 Stainforth, D. A.: Predicting Our Climate Future, Oxford University Press, Oxford, UK, 2023.
- Tantet, A., Lucarini, V., and Dijkstra, H. A.: Resonances in a Chaotic Attractor Crisis of the Lorenz Flow, *Journal of Statistical Physics*, 170, 584–616, <https://doi.org/https://doi.org/10.1007/s10955-017-1938-0>, 2018a.
- Tantet, A., Lucarini, V., Lunkeit, F., and Dijkstra, H. A.: Crisis of the chaotic attractor of a climate model: a transfer operator approach, *Nonlinearity*, 31, 2221–2251, <https://doi.org/10.1088/1361-6544/aaaf42>, 2018b.
- 1055 Tantet, A., Chekroun, M. D., Dijkstra, H. A., and Neelin, J. D.: Ruelle–Pollicott Resonances of Stochastic Systems in Reduced State Space. Part II: Stochastic Hopf Bifurcation, *J. Stat. Phys.*, 179, 1403–1448, <https://doi.org/10.1007/s10955-020-02526-y>, 2020.
- Tél, T. and Gruiz, M.: Chaotic Dynamics, Cambridge University Press, Cambridge, UK, 2006.
- Tél, T., Bódai, T., Drótos, G., Haszpra, T., Herein, M., Kaszás, B., and Vincze, M.: The Theory of Parallel Climate Realizations, *Journal of Statistical Physics*, 179, 1496–1530, <https://doi.org/10.1007/s10955-019-02445-7>, 2020.
- 1060 van Erven, T. and Harremos, P.: Rényi Divergence and Kullback-Leibler Divergence, *IEEE Transactions on Information Theory*, 60, 3797–3820, <https://doi.org/10.1109/TIT.2014.2320500>, 2014.
- Vannitsem, S., Demaeyer, J., and Ghil, M.: Extratropical Low-Frequency Variability With ENSO Forcing: A Reduced-Order Coupled Model Study, *Journal of Advances in Modeling Earth Systems*, 13, e2021MS002 530, <https://doi.org/https://doi.org/10.1029/2021MS002530>, 2021.
- 1065 Vincze, M. and Jánosi, I. M.: Is the Atlantic Multidecadal Oscillation (AMO) a statistical phantom?, *Nonlinear Processes in Geophysics*, 18, 469–475, <https://doi.org/10.5194/npg-18-469-2011>, 2011.
- Vincze, M., Borcia, I. D., and Harlander, U.: Temperature fluctuations in a changing climate: an ensemble-based experimental approach, *Scientific Reports*, 7, 254, <https://doi.org/10.1038/s41598-017-00319-0>, 2017.
- 1070 Werndl, C.: On Defining Climate and Climate Change, *The British Journal for the Philosophy of Science*, 67, 337–364, <https://doi.org/10.1093/bjps/axu048>, 2016.
- Werndl, C.: Initial-Condition Dependence and Initial-Condition Uncertainty in Climate Science, *The British Journal for the Philosophy of Science*, 70, 953–976, <https://doi.org/10.1093/bjps/axy021>, 2019.
- Wouters, J. and Gottwald, G. A.: Stochastic Model Reduction for Slow-Fast Systems with Moderate Time Scale Separation, *Multiscale Modeling & Simulation*, 17, 1172–1188, <https://doi.org/10.1137/18M1219965>, 2019.
- 1075 Yang, H. and Zhu, J.: Equilibrium thermal response timescale of global oceans, *Geophysical Research Letters*, 38, L14711, <https://doi.org/https://doi.org/10.1029/2011GL048076>, 2011.