Forecasting Global Temperatures by Exploiting Cointegration with Radiative Forcing*

Luca Benati
University of Bern†

Abstract

I use Bayesian VARs to forecast global temperatures anomalies until the end of the XXI century by exploiting their cointegration with the Joint Radiative Forcing (JRF) of the drivers of climate change. Under a 'no change' scenario the land and ocean temperature anomalies are projected to reach nearly 6 and 3 Celsius degrees, respectively, by 2100. Forecasts conditional on alternative paths for the JRF show that, given the extent of uncertainty, bringing climate change under control will require to bring the JRF back to the level of the early XXI century. In recent years the JRF has exhibited a marked acceleration, which by 2023 has not been fully reflected in temperature anomalies yet, thus pointing towards their corresponding acceleration going forward.
Keywords: Climate change; Bayesian VARs; cointegration; forecasting; conditional forecasts.
JEL Classification: E2, E3.

*I wish to thank Jennifer Castle, Giuseppe Cavaliere, Marco Del Negro, Jesus Gonzalo, Lutz Kilian, Helmut Luetkepohl, Bill McGuire and participants to the IAAE 2023 and IWEEE 2024 meetings for useful discussions, suggestions and/or comments. Special thanks to Jonathan Wright for very useful suggestions on testing for the null of cointegration with I(2) data. Usual disclaimers apply.

†Department of Economics, University of Bern, Schanzeneckstrasse 1, CH-3001, Bern, Switzerland. Email: luca.benati@vwi.unibe.ch

1 Introduction

For more than a decade global temperatures have been consistently breaking records nearly every year. Against this background, the scorching summers of 2022 and especially 2023, characterized by heatwaves, droughts, wildfires and floods of an unprecedented spread and intensity, have highligthed in the starkest possible way the severity of the threat posed by climate change.

In this paper I use Bayesian VARs in order to forecast temperature anomalies for both the land and the ocean, different latitudes, individual continents, and individual cities until the end of the XXI century, by exploiting their cointegration with the Joint Radiative Forcing (JRF) of the drivers of climate change. My goal is to provide tentative answers to the following questions: What is the increase in global temperatures that is already implied by the level of greenhouse gases' emissions reached in 2023? How will global temperatures evolve going forward? And what are the reductions in greenhouse gases' emissions that will be required in order to bring climate change under control?

I obtain four main sets of results:

(I) Under a 'no change' scenario, median forecasts predict the land and ocean temperature anomalies to reach nearly 6 and 3 Celsius degrees by 2100, respectively, with the corresponding 90%-coverage credible sets equal to [4.0; 7.9] and [2.1; 3.9] degrees. In order to put these numbers into context it is worth recalling that the lower bound of the estimates for the increase in temperatures associated with the Paleocene-Eocene Thermal Maximum (PETM), about 55.5 million years ago, is about 5 Celsius degrees. During that period Antarctica was covered with tropical forests, and Arctic waters pullulated with alligators. Further, and crucially, the period of sustained carbon increase that led to the PETM is estimated to have lasted between 20 thousand and 50 thousand years. If the land temperature anomaly were to reach nearly 6 Celsius degrees (or possibly even higher values) within less than eight decades, the extent to which society could adapt—or whether it could adapt at all—is entirely open to question. Forecasts for alternative latitudes highlight a dramatic extent of variation, with median projected increases by the year 2100 ranging from 3.2 Celsius degrees for the Equator, to 6.0 and 7.4 degrees for the 60 degrees North latitudes and the Arctic, respectively. Forecasts for individual continents point towards a non-negligible extent of heterogeneity, with the median projections for 2100 being equal to about 6 degrees for Europe and Asia, about 5 for North America, and about 4 for Africa, Oceania and South America.

(II) Since the end of the 1970s the rate of increase of the JRF has exibited a consistent, continuous *acceleration*. Given the long lags required for temperature anomalies to reach their new long-run equilibria in response to a permanent increase in the JRF (which I estimate at several decades) this is consistent with the dramatic worsening of climate change over the most recent years. Further, consistent with the acceleration of the increase in the JRF in recent years, over the last decade the JRF has

progressively drifted in the upper tail of the distribution of the forecast conditional on data up until 2013. On the other hand, over the same period both the land and the ocean temperature anomalies have been mostly in line with the median projection conditional on data up to 2013, and in fact in the most recent years both of them have been *below* the median forecast. The logical implication is that in recent years temperature anomalies have not kept up with the rapid acceleration in the rate of increase in the JRF. In turn, this implies that, to the extent that going forward the intrinsic dynamics of the system-Earth will re-establish the long-run equilibrium relationship between the JRF and temperature anomalies, the latter will have to increase *faster* than they have over the decade 2013-2023, in order to catch up with the increases in the JRF.

(III) Forecasts conditional on alternative paths for the JRF show that, given the extent of statistical uncertainty, bringing climate change under control (which I take it to mean keeping temperatures' increases below 1.5 Celsius degrees) will require to bring the JRF back to the levels reached in the early years of the XXI century. Consistent with this, evidence suggests that, even if we were somehow able to prevent any increase in the JRF after 2023, still, the intrinsic dynamics of the system in response to past JRF increases would produce potentially dangerous levels of warming going forward. This evidence provides clear proof that, once taking into account of statistical uncertainty, the JRF *has already exceeded*, in recent years, the level climate scientists regard as safe.

(IV) From a methodological point of view, evidence suggests that previous cointegration-based studies of climate change suffer from model mis-specification. There are two issue involved. First, Stock and Watson's (1996, 1998) tests applied to the first differences of climate change series uniformly and strongly suggest that they contain a non-negligible random-walk component, so that their levels are in fact I(2). The vast majority of previous studies, however, have not considered this possibility, and they have rather assumed that the series are only integrated of order one. Second, Monte Carlo evidence shows that standard fixed-coefficients I(2) cointegrated VECM models,[1] which have been used in a small number of previous studies, are also at odds with the data, whose first differences exhibit random-walk time-variation in the mean. I model this feature of the data via a multivariate random-walk specification for the means of the first differences of the series, subject to the restrictions imposed by cointegration between their levels, a feature that is in fact compatible with the data.

In the climate science literature, long-horizon forecasts for global temperatures are routinely produced via (ensembles of) large-scale models that describe in great detail, and with a significant extent of granularity, a large array of features of the dynamics of Earth's climate. Within the present work, on the other hand, I produce temperatures forecasts based on comparatively small cointegrated VECMs, featuring at most four series. The contrast between the two approaches bears some similarities

[1] See e.g. Juselius (2006).

with the corresponding contrast, within Economics, between structural VAR (SVAR) methods and Dynamic Stochastic General Equilibrium (DSGE) models. Broadly conceptually in line with large-scale climate science models, DSGE models aim to provide a detailed description of all of the interactions taking place in the economy. In line with the approach adopted in the present work, on the other hand, SVAR methods start with a plausible reduced-form time-series representation of the data, and then impose upon this structure a minimal set of restrictions that allows to produce meaningful inference.

Although, to the very best of my knowledge, all previous cointegration-based studies of climate change have been based on Classical methods, there are several reasons behind my adoption of a Bayesian approach. First and foremost, this approach allows a researcher to reject models (i.e. draws from the posterior distribution) exhibiting implausible features. For example, it is well known that on average ocean temperatures have reacted to the increase in the JRF much more slowly than land temperatures. This feature can be easily imposed in estimation by (e.g.) rejecting all models for which the impulse-response function (IRF) of the ocean temperature anomaly to a permanent shock to the JRF index converges to the new steady-state faster than the corresponding IRF of the land anomaly. A Bayesian approach therefore allows to narrow down the set of plausible models, thus producing in principle comparatively more precise inference and forecasts. Further, Bayesian methods provide a natural way of incorporating information from previous studies. This could pertain (e.g.) to the long-run equilibrium relationship between the JRF index and global temperatures (which is related to what in the climate science literature is referred to as 'climate sensitivity'). In the present work I do not exploit this possibility since, as previously pointed out, my evidence suggests that previous cointegration-based studies of climate change suffer from model mis-specification, which raises doubts about the reliability of their estimates. In principle, however, this is an important advantage of a Bayesian approach.

The paper is organized as follows. The next section discusses the data sources; how I address the issue of linking series based on regular, direct observations (which are available for the most recent past) with series that have been spline-interpolated based on irregular observations (which are the only data available for the more distant past); and the construction of the index of JRF. Section 3 presents statistical evidence on the stochastic properties of the series under investigation, i.e. the JRF index and the temperature anomalies. I present evidence from unit root tests; Stock and Watson's (1996, 1998) tests of the null hypothesis of time-invariance against the alternative of random-walk time-variation applied to the first differences of the series; and Wright's (2000) tests of the null hypothesis of cointegration between the JRF index and the temperarture anomalies. Section 4 discusses my econometric approach, paying particular attention to the issue of modelling the common I(2) component shared by the series. Section 5 discusses the evidence: impulse-response functions to a permanent shock to the level of the JRF index; forecasts up to the end of the XXI

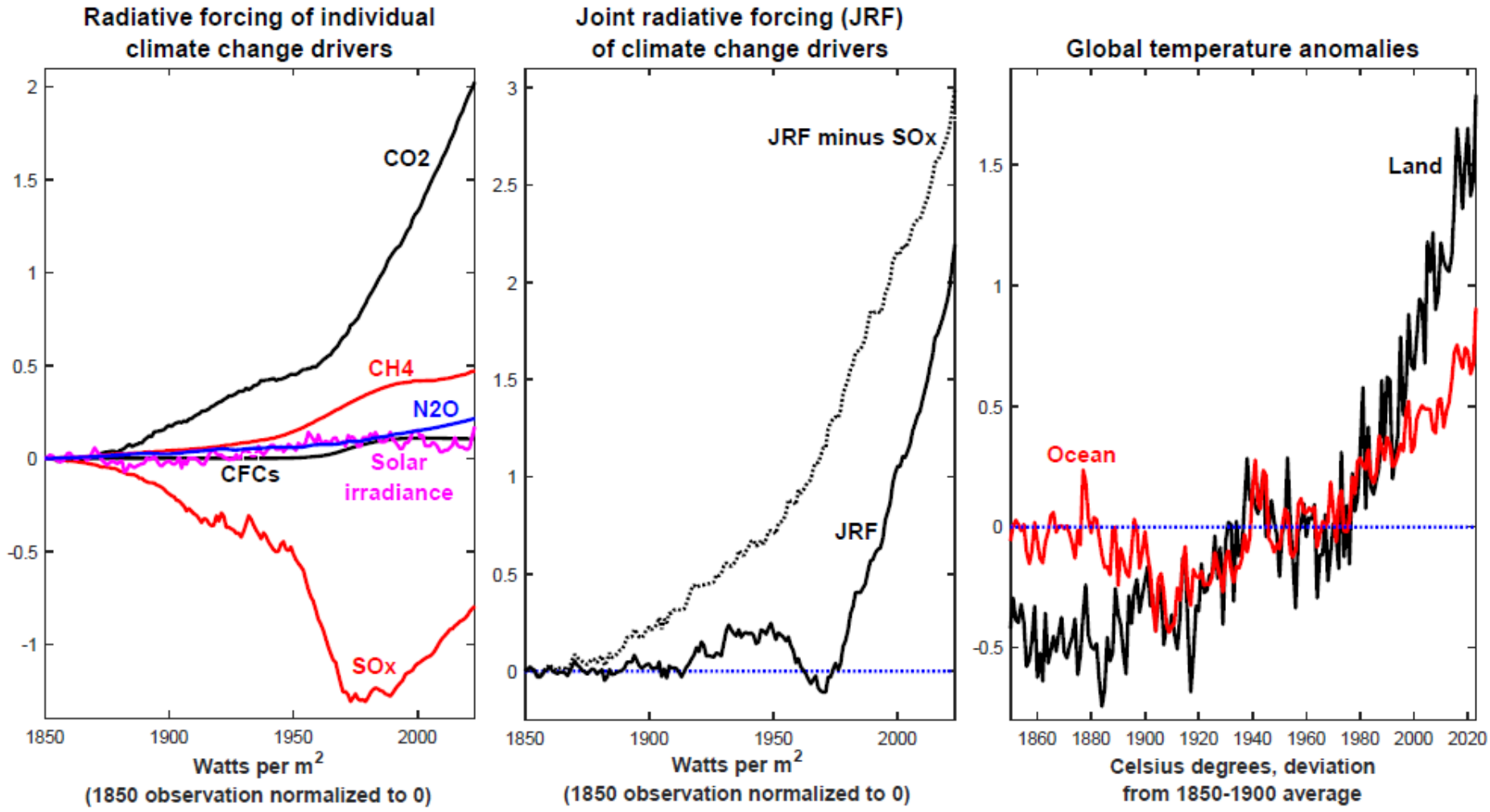

Figure 1 The raw data

century conditional on data up to 2022 but without imposing any other restriction; and forecasts over the same time span conditional on alternative possible paths for the future evolution of the JRF index. Section 6 concludes.

2 The Data

2.1 Data sources

2.1.1 Temperature anomalies

Annual data for the global land and ocean temperature anomalies since 1850 (January-December averages, in Celsius degrees) are from the website of the National Oceanic and Atmospheric Administration (NOAA) at: https://www.ncei.noaa.gov/. They are expressed as deviations from the 1901-2000 average. Since the standard reference period used first and foremost by the Intergovernmental Panel on Climate Change (IPCC) is 1850-1900, I adjust the NOAA series by rescaling them accordingly.

By the same token, annual data since 1850 for the temperature anomalies at three different latitudes (0, and either 30 or 60 degrees North) are from the website of NOAA. Any of the three series has been computed as the average of the global temperature anomalies at that specific latitude for a grid of 38 possible longitudes, from -180 to +180 degrees.[2] Since all of the series for the temperature anomalies pertaining to alternative combinations of latitude and longitude are expressed as deviations from the 1991-2020 average, I rescale them in such a way that, once again, they are expressed as deviations from the standard reference period 1850-1900.

Annual data since 1910 for the temperature anomalies for individual continents (Africa, Asia, Europe, North America, South America, and Oceania, therefore excluding Antarctica) are also from the website of NOAA.

Finally, annual data since 1850 for the temperature anomalies for individual cities are from the website of NOAA.[3] Again, since all of the series for the temperature anomalies pertaining to alternative combinations of latitude and longitude are expressed as deviations from the 1991-2020 average, I rescale them in such a way that, once again, they are expressed as deviations from the standard reference period 1850-1900.

[2] So, to be clear, e.g., the global temperature anomaly for the 30 degrees North latitude has been computed as the average of the temperature anomalies for 30 degrees North and -180 degrees; 30 degrees North and -170 degrees; ...; 30 degrees North and +170 degrees; and 30 degrees North and +180 degrees.

[3] Specifically, NOAA provides temperature anomalies' data for each combination of latitude and longitude on the planet since 1850. Based on individual cities' latitude and longitude their temperature anomalies' series can then be immediately recovered.

2.1.2 Drivers of climate change

Data sources for CO2, CH4, and N2O are as follows. As for CO2, data before 1958 have been spline-interpolated based on the data retrieved from the Scripps CO2 Program (at http://scrippsco2.ucsd.edu). Since 1958, they are based on direct measurements from the Mauna Loa observatory. As for CH4, until 1997 data are from Robertson at el. (2001). Since then they are from NOAA. As for N2O, until 2017 data are from https://www.n2olevels.org/. Since then they are from NOAA. The concentrations of CO2, CH4, and N2O in the atmosphere have been converted into radiative forcing (expressed in Watts per square meter) based on the formulas found in Table 1 of Butler and Montzka (2018).

Data on the radiative forcing of chlorofluorocarbons (CFC11 and CFC12) are from Stern and Kaufmann (2014), and they have been updated based on data from NOAA and the formulas for radiative forcing found in Stern and Kaufmann (2000, p. 435).

A series for the radiative forcing of anthropogenic sulfur emissions (SOx) is from Stern and Kaufmann (2014), and it has been updated based on data from the OECD and the formulas for radiative forcing found in Stern and Kaufmann (2000, p. 435).

The radiative forcing of El Niño and La Niña (El Niño-Southern Oscillation, henceforth ENSO) is from Dergiades, Kaufmann, and Panagiotidis (2016) until 2011, and it has been updated based on data from NOAA.

Data on solar irradiance are from Coddington et al. (2015) and Kopp et al. (2016) until 2014. Since then they are from the SORCE Total Irradiance Monitor (TIM).[4] I convert the resulting index of solar irradiance into radiative forcing based on the formula found on p. 435 of Stern and Kaufmann (2000), which in turn is based on the IPCC (see Shine et al. 1991). Since solar irradiance features an 11-years cycle which is irrelevant for the present purposes, I remove it via the band-pass filter proposed by Christiano and Fitzgerald (2003).[5]

2.2 Linking data based on regular measuring with interpolated data based on irregular observations

For three climate change drivers—CO2, NH4, and N2O—I link spline-interpolated series based on irregular observations with series based on regular direct measuring. One obvious concern with doing this is that the two types of data that are being linked are not exactly comparable, and performing econometrics based on the resulting linked series may therefore produce unreliable results. As it is routinely done in the climate science literature, for either CO2, NH4, or N2O I therefore address this issue as follows.

[4] Details of the TIM design and calibrations are given in Kopp and Lawrence (2005) and Kopp et al. (2005).

[5] Specifically, since I am working at the annual frequency, I remove the frequency band associated with fluctuations between 10 and 12 years.

To fix ideas, let us focus on CO2 (the logic for NH4 and N2O is the same). The spline-interpolated series based on data from the Scripps CO2 program is available until 2018, whereas the series based on direct measurements from the Mauna Loa observatory is available since 1958. Over the common sample period, 1958-2018, I estimate an AR(1) process for the difference between the two series. Then, based on standard resampling methods, I bootstrap (i.e., stochastically simulate) the estimated AR(1) process for a sample equal to the length of the sample for the spline-interpolated data (i.e. 1880-2018) and I add it to the spline-interpolated series. In this way I obtain for the 1880-2018 period a series that mimics the stochastic properties of the series based on direct measurement from Mauna Loa for the period since 1958. (This is what in the climate science literature is labelled as 'adding red noise ' to an interpolated series.) Finally, I construct the linked series for the overall period 1880-2023 by linking the thus constructed, partially simulated series (until 1957) and the Mauna Loa series (since 1958). For NH4 and N2O I proceed in the same way.

By construction, over the first part of the sample period the linked series for CO2, NH4, and N2O are random, as they depend on the specific realizations of the bootstrapped red noise processes. In Section 2.5 below I discuss how I address this issue via Monte Carlo integration, by integrating out the randomness originating from the bootstrapped red noise processes.

2.3 Construction of the JRF index

Once each driver of climate change has been converted into radiative forcing, I construct the aggregate JRF index as in Kaufmann, Kauppi, and Stock (2006) by summing up the individual components. The single exception is El Niño and La Niña (ENSO), which I ignore for the reasons I discuss in Appendix A.1.[6] As shown by Kaufmann, Kauppi, and Stock (2006, see Table II and the discussion on page 261), it is indeed not possible to reject the null hypothesis that 'the temperature effect of a unit of radiative forcing (e.g. W/m^2) is equal across forcings'.

2.4 A look at the raw data

Figure 1 shows the radiative forcing of individual climate change drivers; the JRF index, either including or excluding the radiative forcing of anthropogenic sulfur emissions (SOx); and the global land and ocean temperature anomalies.

As I discuss in Section 2.5, the series for CO2, NH4, and N2O are in part stochastically simulated over the first portions of the respective sample periods, by adding to

[6] In brief, (1) ENSO features virtually no spectral power at frequencies beyond 25 years, and (ii) it is extraordinarily noisy compared to the other drivers of climate change. The implication is that including the radiative forcing of ENSO in the JRF index would uniquely add a large amount of comparatively high-frequency noise, whereas it would bring essentially *no information* about the long-horizon developements that are the focus of the present work.

spline-interpolated data bootstrapped (i.e. stochastically simulated) red noise based on estimated AR(1)'s. With the exception of solar irradiance and the CFCs, all of the series shown in the first two panels of Figure 1 pertain therefore to a single stochastic simulation. It is to be stressed, however, that because of the comparatively small magnitude of the estimated red noise compared to the level of the series, the difference between individual stochastic simulations is very small (then, for either CO2, NH4, or N2O the second part of the respective samples, being based on direct regular measurements, is by definition the same). This means that, in practice, what is shown in Figure 1 is representative of the entire universe of simulations. Figure A.2 in the Appendix provides simple evidence on this. The figure shows, for CO2, NH4, and N2O, the maximum and the minimum among the sorted partially simulated paths out of 100,000 simulations, together with the difference between them. Evidence is very clear: over the first portions of the sample (until 1960, 1980, and 1980, respectively), the simulated paths had been very close. This had especially been the case for CH4, and just slighty less so for CO2 and N2O.

Two main findings clearly emerge from Figure 1. First, since 1850 CO2, CH4 and SOx have been by far the dominant drivers of the evolution of the JRF index. Second, until about the early 1990s SOx had been playing an important moderating role in the overall increase in the JRF index. Since then, however, its previous moderating contribution has gone into reverse, as efforts to remove anthropogenic sulfur emissions from the atmosphere have started to bear fruits. As a result, over the last three decades the evolution of the SOx radiative forcing has contributed to the overall increase in the JRF index.

The second panel of Figure 1 illustrates this point in an especially stark way. Normalizing the JRF index to zero in 1850, excluding the impact of SOx the index would have increased much faster than it has historically been the case. To the extent that efforts to remove anthropogenic sulfur emissions from the atmosphere will continue and will be successful, the radiative forcing of SOx shown in the first panel will converge to zero, and the overall JRF index will therefore be more and more dominated by the remaining drivers.

Finally, the third panel illustrates the well-known lag that the ocean temperature anomaly has consistently exhibited over the last five decades compared to its land counterpart. Whereas the two anomalies had been fluctuating pretty much in synch between the early XX century and the 1970s, since then a sizeable divergence has developed, with the ocean anomaly consistently lagging behind its land counterpart.

2.5 Integrating out simulated red noise via Monte Carlo integration

As discussed, over the first part of the sample period the linked series for CO2, NH4, and N2O are random, as they depend on the specific realizations of the bootstrapped red noise processes. I therefore address this issue as follows.

For $k = 1, 2, 3, ..., K$, with $K = 1{,}000$, I generate partially simulated[7] series for the concentration of CO2, NH4, and N2O in the atmosphere, I convert them into radiative forcing, and I sum them to the radiative forcings of the remaining drivers of climate change, thus obtaining a partially simulated series for the JRF index. Based on this and on the series for the temperature anomalies (which are based on regular direct observations over the entire sample period) I then estimate the cointegrated VECM models I describe in Section 4, and I compute the median, and the 16-84 and 5-95 percentiles for all of the objects if interest (impulse-response functions to a permanent shock to the level of JRF; unconditional and conditional forecasts; etc.). Finally, I integrate out the uncertainty deriving from the fact that for CO2, NH4, and N2O the first part of the sample has been partially simulated by computing the average (corresponding to the expected value) of the objects of interest across all of the K simulations. This Monte Carlo integration procedure allows to perform the empirical analysis by effectively controlling for the fact that three of the radiative forcing series have been partially stochastically simulated.

3 Preliminary Statistical Analysis

My econometric approach, which is based on Bayesian cointegrated VARs, is predicated on the following two assumptions:

(I) both the JRF index and all of the temperature anomalies series are integrated processes. In particular, they are all I(2).

(II) The JRF index is cointegrated with any of the temperature anomaly series. In particular, cointegration pertains to the *levels* of the series, so that (e.g.) for the land anomaly in a long-run equilibrium $T^{\text{Land}} = \alpha + \beta JRF + \theta_t$, where the notation is obvious and θ_t is an I(0) cointegration residual.

As I discuss in the next four sub-sections, these assumptions are either validated by, or at the very least clearly compatible with, evidence from unit root tests; Stock and Watson's (1996, 1998) tests for the null of time-invariance against the alternative of random-walk time-variation in the mean of the first differences of the series; and Wright's (2000) tests of the null of cointegration. For reasons of space in this section I only report evidence for the JRF index and the land and ocean temperature anomalies. The corresponding evidence for the temperature anomalies pertaining to different latitudes, individual continents, and individual cities is reported in Online Appendix A.1, and will only be briefly referenced in what follows. That evidence is near-uniformly in line with the evidence discussed herein. The main exception pertains to the cointegration tests between the JRF index and the temperature anomalies for individual continents, which detect cointegration only for Europe. A possible explanation for this result is that the sample period for individual continents' temperature

[7] Based on the previous discussion, 'partially simulated' refers to the first part of the sample, for which we only have spline-interpolated data.

anomalies (112 years) is materially shorter than for either the land and ocean anomalies, or the anomalies for different latitudes and individual cities (in either instance, 172 years). Accordingly, in what follows I will proceed under the assumption that temperature anomalies for individual continents are also cointegrated with the JRF index.

3.1 Unit root tests

Tables 1*a*-1*b* show bootstrapped *p*-values for Elliot, Rothenberg, and Stock's (1996) unit root tests for both the levels and the first differences of the JRF index and the land and ocean temperature anomalies. For the temperature anomalies series, for which we have raw data for the entire sample since 1850, I perform the tests in the standard way, bootstrapping them as in Diebold and Chen (1996) based on the first difference of the series that is being tested (i.e., either the level or the first difference of either of the two anomalies). I consider five possible lag orders, from 1 to 5 years.

Table 1*a* Bootstrapped *p*-values for Elliot, Rothenberg, and Stock's (1996) unit root tests for the land and ocean temperature anomalies

	p=1	p=2	p=3	p=4	p=5
	Land				
In levels, without time trend	0.6824	0.9052	0.9750	0.9904	0.9940
In levels, with time trend	0.1526	0.3048	0.7162	0.7180	0.7308
In first differences, without time trend	0.0000	0.0000	0.0000	0.0000	0.0000
	Ocean				
In levels, without time trend	0.4682	0.7256	0.9120	0.9296	0.9588
In levels, with time trend	0.1278	0.4878	0.5440	0.5558	0.6766
In first differences, without time trend	0.0000	0.0000	0.0000	0.0000	0.0000

For the JRF index, on the other hand, I generate 10,000 partially simulated series as previously described (i.e., by adding bootstrapped red noise to the spline-interpolated data for the first part of the sample), and based on each of them I perform the same unit root tests I perform for the temperature anomalies. Table 1*b* reports the means and the medians of the Monte Carlo distributions of the bootstrapped *p*-values across the 10,000 simulations, together with the fraction of Monte Carlo replications for which the *p*-values are smaller than 10%.

For any of the three series the null of a unit root cannot be rejected in levels, either including or not including a time trend, and based on any of the five lag orders. In differences evidence is mixed: for all of the temperature anomalies a unit root is strongly rejected, whereas for the JRF index it is rejected at the 10 per cent level only for lag orders smaller than or equal to 3.

Tables A.1.1, A.2.1 and A.3.1 in the Online Appendix report the corresponding evidence for the temperature anomalies for different latitudes, individual continents,

and individual cities respectively. This evidence is near-uniformly in line with that in Tables 1*a*-1*b*, with only three exceptions. For both the latitude 0, and South America and Oceania, the null of a unit root can be rejected based on models featuring a time trend. Since, arguably, the most meaningful tests are those not including a time trend,[8] in what follows I will ignore these results, and I will proceed under the assumption that these three series all feature a unit root.

Table 1*b* Bootstrapped *p*-values for Elliot, Rothenberg, and Stock's (1996) unit root tests for the JRF index[a]

	p=1	p=2	p=3	p=4	p=5
	Mean of Monte Carlo distribution of p-values				
In levels, without a time trend	1.0000	1.0000	1.0000	1.0000	1.0000
In levels, with a time trend	0.9995	0.9994	0.9992	0.9997	0.9999
In first differences, without time trend	0.0001	0.0032	0.0379	0.1313	0.3018
	Median of Monte Carlo distribution of p-values				
In levels, without a time trend	1.0000	1.0000	1.0000	1.0000	1.0000
In levels, with a time trend	0.9995	0.9995	0.9995	1.0000	1.0000
In first differences, without time trend	0.0000	0.0030	0.0375	0.1305	0.3010
	Fraction of Monte Carlo distribution of p-values smaller than 10 per cent				
In levels, without a time trend	0.0000	0.0000	0.0000	0.0000	0.0000
In levels, with a time trend	0.0000	0.0000	0.0000	0.0000	0.0000
In first differences, without time trend	1.0000	1.0000	1.0000	0.0425	0.0000

[a] Based on10,000 Monte Carlo simulations of joint radiative forcing.

Based on the evidence in Tables 1*a*-1*b* and A.1.1-A.3.1 a reasonable characterization of the data, which has in fact been adopted by the vast majority of cointegration-based studies on climate change, is that all of the series are I(1). As the evidence in the next sub-section shows, this conclusion would however most likely be incorrect, since Stock and Watson's (1996, 1998) tests applied to the first differences of the series clearly suggest that they all contain a random-walk component.

3.2 Searching for random-walk time-variation in the first differences of the series

Figure 2 shows rolling averages of the first differences of the JRF index and of the land and ocean temperature anomalies for 50-years samples. The evidence is unmistakeable: the average level of all of the three series has exhibited a broad upward trend

[8] The reason is that in terms of both physics and basic logic it is not clear why temperature anomalies should feature a deterministic time trend.

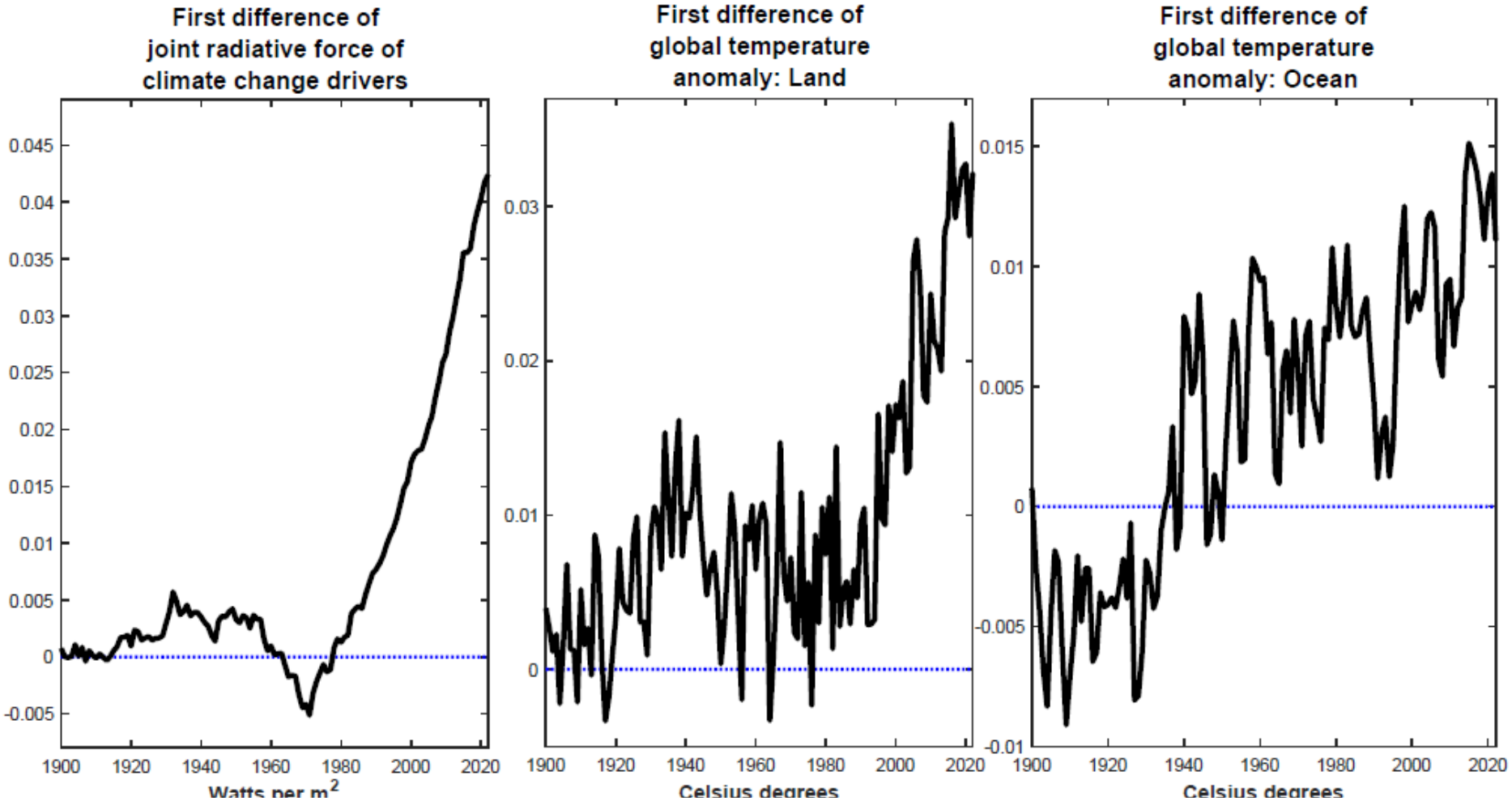

Figure 2 Fifty-years rolling averages of the first differences of the series

over the sample period. For both the JRF index and the land temperature anomaly this is especially apparent since the end of the 1970s, when the contribution of SOx to the JRF index turned from negative to positive (see Section 2.4). As for the ocean anomaly the range of variation is significantly smaller than for the land anomaly, but the overall pattern of increase is more regular. This is what we would expect from the fact that, once appropriately scaled, the ocean anomaly behaves essentially as a low-frequency component of the land anomaly. Evidence for the temperature anomalies pertaining to different latitudes, individual continents and individual cities is qualitatively the same, and it is available upon request.

This evidence questions the notion that the first differences of the series are in fact I(0), and therefore that the series themselves are I(1). Indeed, from I(0) series one would not expect such a consistent pattern of progressive increase at the very low frequencies over such a long period of time. Rather, the evidence in Figure 2 naturally suggests that the first differences of the series feature an I(1) component which is too small to be detected by standard unit root tests, but it is in fact sufficiently sizeable to induce a progressive increase in the average level of the series' first differences.

Table 2 Simulated p-values for Stock and Watson's tests for the null of time-invariance against the alternative of random-walk time-variation in the mean of the first differences of the series

	JRF index[a]			Temperature anomalies	
HAC correction:	Mean	Median	Fraction below 10%	Land	Ocean
	Trimming: 0.15				
Newey and West (1987)	0.0448	0.0148	0.8590	0.000	0.049
Andrews (1991)	0.0531	0.0209	0.8590	0.000	0.112
	Trimming: 0.25				
Newey and West (1987)	0.0409	0.0120	0.8480	0.001	0.041
Andrews (1991)	0.0480	0.0171	0.8480	0.000	0.088
	Trimming: 0.33				
Newey and West (1987)	0.0754	0.0488	0.7740	0.003	0.044
Andrews (1991)	0.0864	0.0591	0.7700	0.001	0.094

[a] Mean and median of the Monte Carlo distribution of p-values, and fraction of p-values smaller than 10%.

In order to explore this issue, Table 2 reports evidence from Stock and Watson's (1996, 1998) tests of the null hypothesis of no time-variation in the mean of the first difference of any of the series, against the alternative of random-walk time variation. Tables A.1.2, A.2.2 and A.3.2 in the Online Appendix report the corresponding evidence for the temperature anomalies for different latitudes, individual continents, and individual cities respectively. In implementing Stock and Watson's approach I closely follow Stock and Watson (1996, 1998). The methodology is described in detail

in Appendix B, and it is exactly the same I used in Benati (2007). I consider three alternative values of 'trimming', i.e. the standard 15% and, in order to give more power to the tests, either 25% or 33%. I control for the possible autocorrelation and/or heteroskedasticity of the residuals via either Newey and West's (1987) or Andrews' (1991) covariance matrix estimator. For any of the temperature anomalies, for which we have raw data for the entire sample since 1850, I simply report the simulated p-values produced by Stock and Watson's methodology. For the JRF index, which as discussed is partially stochastically simulated, I report the mean and the median of the Monte Carlo distribution of the simulated p-values, together with the fraction of the p-values smaller than 10% across all of the 10,000 Monte Carlo simulations.

Overall, I detect strong evidence of random-walk time-variation for all series.[9] This suggests that although the I(1) component is too small to be detected based on standard unit root tests, in fact it is sufficiently large to be detected based on the approach proposed by Stock and Watson (1996, 1998).

Table 3 Monte Carlo evidence on the plausibility that a fixed-coefficients I(2) cointegrated VECM model may have produced the results in Table 2: mean and median of the Monte Carlo distribution of the simulated p-values for Stock and Watson's tests, and fraction of replications for which the p-values are smaller than 10%

HAC correction:	Newey and West (1987)			Andrews (1991)		
	JRF index	Land anomaly	Ocean anomaly	JRF index	Land anomaly	Ocean anomaly
	Trimming: 0.15					
Mean	0.5884	0.0490	0.1705	0.8228	0.0388	0.2027
Median	0.5750	0.0140	0.1130	0.8950	0.0080	0.1280
Fraction below 10%	0.0050	0.8659	0.4615	0.0270	0.9049	0.4324
	Trimming: 0.25					
Mean	0.6845	0.0897	0.2463	0.6399	0.0713	0.3094
Median	0.6720	0.0250	0.1580	0.8040	0.0140	0.2300
Fraction below 10%	0.0010	0.7417	0.3604	0.2072	0.8078	0.3003
	Trimming: 0.33					
Mean	0.6765	0.1059	0.2628	0.6483	0.0862	0.3270
Median	0.6620	0.0290	0.1800	0.7960	0.0170	0.2530
Fraction below 10%	0.000	0.7177	0.3433	0.1872	0.7708	0.2853

[9] The only exception is the temperature anomaly for the latitude 0, for which the simulated p-values are consistently greater than 10 per cent. The most plausible explanation is that the impact of climate change is greater the higher the latitude, and it is therefore maximum at the Poles and minimum at the Equator. As a consequence, for the latitude 0 the extent of random-walk time-variation is likely too small to be detected by Stock and Watson's tests. In what follows I will therefore proceed under the assumption that the temperature anomaly for the Equator, too, features random-walk time-variation in the mean.

3.3 Interpreting the results from Stock and Watson's (1996, 1998) tests

The natural interpretation of the evidence in Figure 2 and especially Tables 2, and A.1.2-A.3.2 in the Online Appendix, is that the equilibrium levels of the first differences of the series follow a multivariate random-walk process driven by the progressive increase in the first difference of the JRF index. In principle, however, an alternative possible interpretation of this evidence might be that the data-generation process (DGP) is a time-invariant I(2) VAR. As I now show via Monte Carlo, however, this clearly seems not to be the case.

Tables 3, and A.1.3-A.3.3 in the Online Appendix, report results from the following exercise. I generate, as previously described, 1,000 partially simulated series for the JRF index. Based on any of them and the temperature anomalies' series I then estimate a standard fixed-coefficients I(2) cointegrated VECM model[10] via Bayesian methods as described in Appendix C, by combining the likelihood of the data with a minimal set of priors designed in order to impose a meaningful structure upon the VECM.[11] For each of the 1,000 samples this produces 1,000 draws from the posterior distribution, thus resulting in a total of 1 million models. I then stochastically simulate each of these models for samples of length equal to the length of the actual sample I am working with (1910-2022 for continents' anomalies, and 1850-2022 for all other anomalies), and based on each of them I perform the same Stock and Watson's (1996, 1998) tests I performed based on the actual data.

Table 3 reports, for the JRF index and the land and ocean anomalies, the means and the medians of the Monte Carlo distributions of the simulated p-values, together with the fractions of simulations for which the p-values have been below 10%, out of the 1 million Monte Carlo simulations. Tables A.1.3-A.3.3 in the Online Appendix report the corresponding evidence for the different latitudes, individual continents, and individual cities respectively. The evidence in the four tables is very clear: if the true DGP had been a standard fixed-coefficients I(2) cointegrated VECM model, obtaining the results reported in Tables 2 and A.1.2-A.3.2 would have been extremely unlikely. For the sake of the argument, let us focus in Table 3 on the results based on 25% trimming and Newey and West's (1987) HAC correction. For two series out of three (the JRF index and the ocean anomaly) both the mean and the median of the Monte Carlo distributions of the p-values are materially beyond 10%, and the fractions of Monte Carlo simulations for which the p-values are below 10% are equal to 0.0010 for the JRF index, and to 0.3604 for the ocean anomaly. Only for the land anomaly the evidence in Table 3 is compatible with the notion that the DGP is a standard I(2) cointegrated VECM model. The evidence in Tables A.1.2-A.3.2 is

[10] The model is discussed e.g. in Juselius (2006).

[11] As I discuss below, the main restriction I impose is that the reaction of the ocean anomaly to increases in the JRF is slower than the corresponding reaction of the land anomaly, which I express in terms of their normalized impulse-response functions to a permanent JRF shock.

qualitatively the same.

Evidence based on non-cointegrated I(2) VAR models is qualitatively the same, and it is available upon request. My preference for the evidence based on I(2) cointegrated VECM models originates from the fact that, as I show in the next sub-section, the data quite clearly point towards cointegration between the JRF index and temperature anomalies, so that non-cointegrated I(2) VAR models are significantly less empirically plausible.

I now turn to testing for cointegration between the JRF index and either temperature anomaly.

3.4 Evidence from Wright's (2000) cointegration tests

Based on the climate science literature, the relevant null hypothesis to be tested is that the level of the JRF index is cointegrated with the levels of any of the temperature anomalies, so that even if the series are I(2), the residual from the cointegrating regressions

$$T_t = a + bJRF_t + u_t, \tag{1}$$

where T_t is one of the temperature anomaly series, is I(0). I therefore proceed as follows.

To fix ideas, let us focus on the land temperature anomaly. I start by generating, as previously described, 1,000 partially simulated series for the JRF index, JRF_t^j, with $j = 1, 2, ..., 1{,}000$. Then,

(1) based on each pair $\{JRF_t^j, T_t^{\text{Land}}\}$, $j = 1, 2, ..., 1{,}000$, I perform a Wright (2000) test for the null hypothesis of cointegration between JRF_t^j and T_t^{Land}.

(2) Based on each triplet $\{JRF_t^j, T_t^{\text{Land}}, T_t^{\text{Ocean}}\}$, $j = 1, 2, ..., 1{,}000$, I estimate the cointegrated VECM model discussed in the next section, featuring a multivariate random-walk specification for the time-varying equilibrium levels of the first differences of the three series. For each sample $j = 1, 2, ..., 1{,}000$, this produces $d = 1, 2, ..., 1{,}000$ draws from the posterior distribution.

(3) For each $j = 1, 2, ..., 1{,}000$, I then stochastically simulate each of the 1,000 d models (i.e. draws from the posterior), thus obtaining a Monte Carlo distribution of Wright's (2000) test under the null hypothesis that (i) the three series are I(2) and (ii) they are cointegrated in levels. Based on this, I compute as in Wright (2000) the 90%-coverage confidence intervals for the cointegration coefficient.

For either the land or the ocean temperature anomaly Wright's (2000) tests cannot reject the null hypothesis of cointegration for any $j = 1, 2, ..., 1{,}000$. Finally, as discussed in Section 2.5 I integrate out the randomness associated with the simulated red noise I have added to the JRF index over the first part of the sample period by computing the average, across all j's, of the confidence intervals for the cointegration coefficient.

The average across the 1,000 simulations of the 90%-coverage confidence interval for the cointegration coefficient for the land anomaly is [-0.9521 -0.6000], whereas

the corresponding object for the ocean anomaly it is [-2.2411 -1.4546], reflecting the significantly slower rate of warming of the oceans over the sample period. It is to be noticed that the confidence intervals for the land and ocean anomalies are not overlapping. Taken at face value this would imply that the land and ocean anomalies, although both cointegrated with the JRF index, exhibit different long-run equilibrium relationships with it. Another possible interpretation is that, in response to a permanent increase in the JRF index, the two temperature anomalies ultimately increase by exactly the same amount, so that they share the same cointegration vector with the JRF, but that the sample period since 1850 is simply too short to capture this. As I will discuss in the next section, climate science suggests that the former interpretation is significantly more plausible, and that in a long-run equilibrium the land and the ocean exhibit different responses to a permanent increase in the JRF.

Evidence for the systems featuring the JRF index and temperature anomalies either for different latitudes, or pertaining to individual continents or cities, discussed in Online Appendices A.1.1-A.1.3, is qualitatively the same. In particular, the averages across the 1,000 simulations of the 90%-coverage confidence intervals for the cointegration coefficients for the latitudes 0, 30 degrees North, and 60 degrees North are equal to [-3.9105 -0.6002], [-3.0715 -0.5049], and [-1.7727 -0.3265] respectively, reflecting the fact that the higher the latitude, the greater the increase in temperatures.

I now turn to discussing my econometric approach.

4 The Econometric Approach

I start by discussing the standard cointegrated VECM model for I(1) series detailed (e.g.) in Hamilton (1994), and I then turn to the modification I propose in order to take into account of the fact that, as previously shown, the first differences of the series feature random-walk time-variation in their means.

4.1 The I(1) cointegrated VECM model

Let the standard cointegrated VECM representation for a $(N\times 1)$ vector of I(1) series Y_t be

$$\Delta Y_t = B_0 + B_1 \Delta Y_{t-1} + ... + B_p \Delta Y_{t-p} + \alpha\beta' Y_{t-1} + u_t, \tag{2}$$

where β is the matrix of the cointegration vectors, α is the matrix of the loading coefficients, $E[u_t' u_t] = \Sigma$, and the rest of the notation is standard. By defining as M the time-invariant unconditional mean of ΔY_t, with $B_0 = [I_N - B_1 - ... - B_p]M$, this expression can be rewritten as

$$\Delta Y_t - M = B_1(\Delta Y_{t-1} - M) + ... + B_p(\Delta Y_{t-p} - M) + \alpha\beta' Y_{t-1} + u_t. \tag{3}$$

As shown in Section 3.2, evidence from Stock and Watson's (1996, 1998) tests suggests that within the present context M features random-walk time-variation, which implies

that (2) does not provide a meaningful description of the DGP.[12] The next sub-section describes how I address this issue.

4.2 An I(2) cointegrated VECM model with random-walk time-variation in the means of the series' first differences

The natural way of modelling random-walk time-variation in M is to postulate that it evolves according to a multivariate random walk specification, subject to the restrictions imposed by cointegration between the levels of the series. To fix ideas, let us focus on the trivariate system featuring the JRF index and the land and ocean temperature anomalies, so that $Y_t = [JRF_t, T_t^{\text{Land}}, T_t^{\text{Ocean}}]'$. For systems featuring alternative temperature anomalies indices the logic is exactly the same. As previously shown, the levels of both temperature anomalies are cointegrated with the level of the JRF index. This implies that the system features a single I(2) stochastic trend, originating from the progressive increase of the JRF, and two cointegration vectors. Within the present context, the natural rotation of the cointegration space is obtained by defining the matrix of the cointegration vectors as

$$\beta = \begin{bmatrix} 1 & 1 \\ -b_{\text{Land}} & 0 \\ 0 & -b_{\text{Ocean}} \end{bmatrix}. \tag{4}$$

which implements the previously discussed restrictions.

As discussed e.g. in Kleibergen and van Dijk (1994) and Bauwens and Lubrano (1996), for the r cointegration vectors to be uniquely identified, each of the r columns of β ought to feature at least r restrictions. Within the present context this is indeed the case, as each of the two columns features two restrictions, and in fact either of them depends on a single cointegration coefficient.

Finally, the fact that the three series are cointegrated, with the matrix of cointegration vectors given by (4), imposes the following restrictions on M_t:

$$M_t = \begin{bmatrix} 1 \\ \frac{1}{b_{\text{Land}}} \\ \frac{1}{b_{\text{Ocean}}} \end{bmatrix} \mu_t \tag{5}$$

$$\mu_t = \mu_{t-1} + \epsilon_{\mu,t} \tag{6}$$

where μ_t is a scalar random-walk process capturing the common stochastic trend driving the frequency-zero dynamics of ΔY_t, with $\epsilon_{\mu,t} \sim N(0, \sigma_\mu^2)$. Since, in equation

[12] In fact, attempts to estimate the I(1) cointegrated VECM (2) were uniformly fruitless. On the one hand, Johansen's estimator consistently produced estimated models featuring explosive eigenvalues. On the other hand, direct MLE estimation based on numerical methods, along the lines of the discussion in Section 5.3 below, consistently proved to be highly problematic, with the problem boiling down once again to the difficulty of obtaining non-explosive estimates. In line with the previous discussion, the natural interpretation of all this is that the data are in fact I(2).

(5), μ_t has been normalized on the JRF index, in fact it captures the I(1) stochastic trend in the first difference of JRF_t, i.e. the sources of anthropogenic climate change. When jointly considered, (4), (5) and (6) imply that the JRF index and the temperature anomalies maintain a long-run equilibrium relationship in response to permanent shocks to both the level of the JRF index, and its first difference.

4.3 Estimation

I estimate all models via Bayesian methods, by combining the log-likelihood of the data with a minimal set of inequality restrictions (discussed in the next sub-section) on the impulse-response functions (IRFs) of the series to a permanent shock to the JRF index. In practice, this means that I perform MLE estimation subject to the restriction that I reject models (i.e. draws) that do not satisfy the restrictions on the IRFs. So, although I adopt a Bayesian approach, which allows me to reject draws that do not satisfy these restrictions, in fact I do not specify a *single* prior for *any* parameter.

I numerically maximize the restricted log-likelihood of the data (where by 'restricted' I mean that it is subject to the just-mentioned restrictions on the IRFs) via simulated annealing exactly as in Benati (2008). Following Goffe, Ferrier, and Rogers (1994) I implement simulated annealing via the algorithm proposed by Corana, Marchesi, Martini and Ridella (1987).[13] I then stochastically map the restricted log-likelihood based on Random Walk Metropolis (RWM). In implementing the RWM algorithm I exactly follow An and Schorfheide (2007, Section 4.1), with the only difference that the jump to the new position in the Markov chain is accepted or rejected based on a rule which does not involve any Bayesian priors on the model's coefficients, as it uniquely involves the restricted likelihood of the data.[14] I calibrate the covariance matrix's scale factor for RWM based on the methodology proposed by Benati

[13] I set the key parameters to T_0 =100,000, $r_T = 0.9$, $N_t = 5$, $N_s = 20$, $\epsilon = 10^{-6}$, and $N_\epsilon = 4$, where T_0 is the initial temperature, r_T is the temperature reduction factor, N_t is the number of times the algorithm goes through the N_s loops before the temperature starts being reduced, N_s is the number of times the algorithm goes through the function before adjusting the step size, is the convergence (tolerance) criterion, and N_ϵ is the number of times convergence is achieved before the algorithm stops.

[14] So, to be clear, the proposal draw for β, $\tilde{\beta}$, is accepted with probability min[1, $r(\beta_{s-1}, \tilde{\beta} \mid Y, X)$], and rejected otherwise, where β_{s-1} is the current position in the Markov chain, and

$$r(\beta_{s-1}, \tilde{\beta} \mid Y, X) = \frac{L(\tilde{\beta} \mid Y, X)}{L(\beta_{s-1} \mid Y, X)}$$

which uniquely involves the restricted likelihood. With Bayesian priors it would be

$$r(\beta_{s-1}, \tilde{\beta} \mid Y, X) = \frac{L(\tilde{\beta} \mid Y, X)P(\tilde{\beta})}{L(\beta_{s-1} \mid Y, X)P(\beta_{s-1})}$$

where $P(\cdot)$ would encodes the priors about β.

(2008, Appendix C).

I run a burn-in pre-sample of 1,000,000 draws which I then discard. I then generate 10,000,000 draws, which I 'thin' by sampling every 1,000 draws in order to reduce their autocorrelation. This leaves 10,000 draws from the ergodic distribution which I use for inference. For all models the fraction of accepted draws is very close to the ideal one, in high dimensions, of 0.23 (see Gelman, Carlin, Stern, and Rubin, 1995).

I check convergence of the Markov chain based on Geweke's (1992) inefficiency factors (IFs) of the draws from the ergodic distribution for each individual parameter. The IFs are defined as the inverse of the relative numerical efficiency measure of Geweke (1992),

$$RNE = (2\pi)^{-1}\frac{1}{S(0)}\int_{-\pi}^{\pi} S(\omega)d\omega, \tag{7}$$

where $S(\omega)$ is the spectral density of the sequence of draws from RWM for the parameter of interest at frequency ω. I estimate the spectral densities via the lag-window estimator as described in chapter 10 of Hamilton (1994). I also considered an estimator based on the fast-Fourier transform, and results were very similar. For all parameters the IFs are equal to at most 3-4, well below the values of 20-25 which are typically taken to indicate problems in the convergence of the Markov chain.

4.4 Restrictions imposed in estimation

In estimation I impose the restrictions that for each parameters' draw a permanent shock to the level of the JRF index generates a non-negative impulse-response function (IRF) for JRF at all horizons. Further, for the system featuring the JRF index and the land and ocean temperature anomalies I impose the restriction that, once the individual series' IRFs have been normalized by their respective long-run impacts, at all horizons

(1) the response of the ocean temperature anomaly is slower than the response of the land temperature anomaly (i.e. the normalized IRF of the latter lies below the normalized IRF of the former at all horizons), and

(2) the response of the land temperature anomaly is slower than the response of the JRF index (i.e. the normalized IRF of the latter lies below the normalized IRF of the former at all horizons).

The rationale for (1) is the well-known slower responsiveness of the ocean temperature anomaly to increases in the JRF compared to its land counterpart. As for (2) the rationale is that, as matter of logic, the JRF index ought to respond faster to its own shocks than either of the two temperature anomalies.

As for the other systems, beyond the restriction that a permanent shock to JRF produces a non-negative IRF for JRF itself at all horizons, I impose the restriction that, at all horizons, the normalized IRFs of all temperature anomalies lie below the normalized IRF of the JRF index.

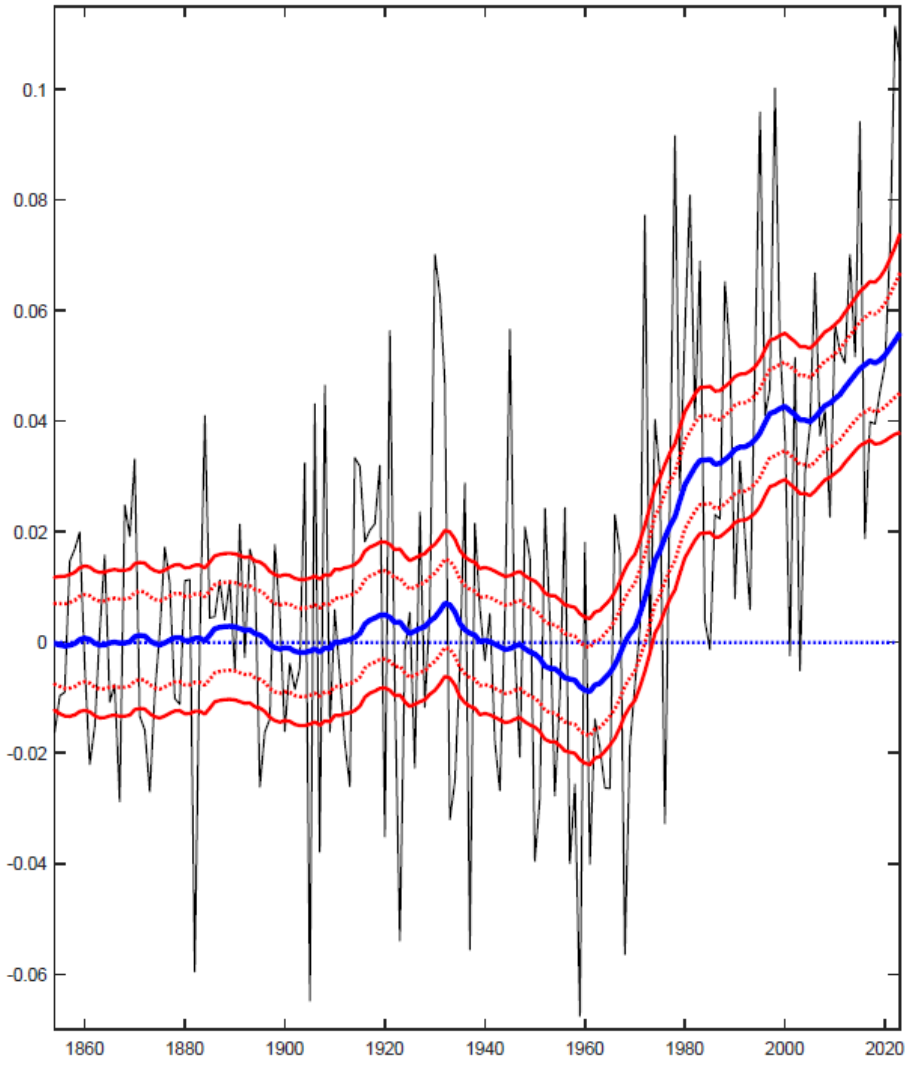

Figure 3 First difference of joint radiative force, and two-sided estimate of Mu(t) (median, and 16-84 and 5-95 credible set)

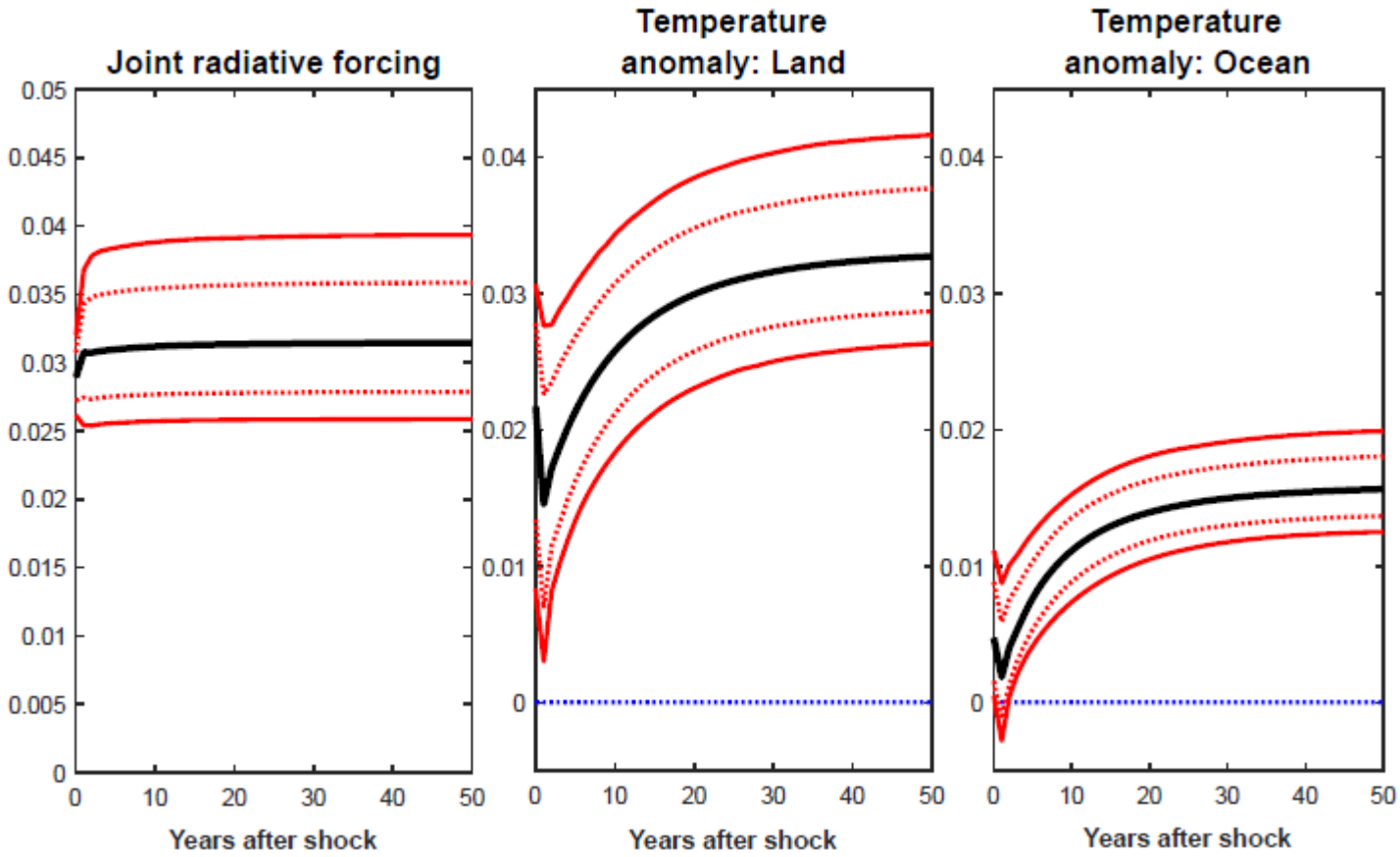

Figure 4 Impulse-response functions to a permanent shock to the level of joint radiative forcing (median, and 16-84 and 5-95 credible set)

I impose all of the previously discussed restrictions by rejecting the draws that do not satisfy them. I now turn to discussing the evidence.

5 Evidence

Figure 3 shows the first difference of the JRF index and the two-sided median estimate of μ_t, together with the 16-84 and 5-95 per cent credible sets of the posterior distribution, based on the model with $Y_t = [JRF_t, T_t^{\text{Land}}, T_t^{\text{Ocean}}]'$. The estimate of μ_t has been computed via the Monte Carlo integration procedure proposed by Hamilton (1986). Two facts are readily apparent from the figure. First, as it was to be expected based on the normalization in (5), μ_t behaves as a sort of time-varying equilibrium level of the first difference of the JRF index. Second, μ_t had been fluctuating around zero until about the 1970s, whereas since then it has been increasing rapidly. This is consistent with the dramatic acceleration of climate change over the most recent period.

5.1 Impulse-response functions to a permanent shock to the JRF index

Figure 4 shows the IRFs of the JRF index and of the land and ocean temperature anomalies to a permanent one-standard deviation shock to the level of the JRF. The response of the JRF is essentially flat at all horizons, thus clearly suggesting that in fact the index is a pure unit root process. The responses of the two temperature anomalies are as expected, and they partly reflect the restrictions imposed upon them. In particular, the response of the ocean anomaly is slower and more drawn out than that of the land anomaly, which in turn converges to its new long-run equilibrium only beyond the 50 years ahead horizon considered herein. It is to be noticed, however, that in line with a recent body of evidence from the climate science literature discussed e.g. in Mann (2023), the bulk of temperatures' response takes place within about 20 years.

5.2 Unconditional forecasts based on data up to 2023

5.2.1 Land and ocean temperature anomalies

Figure 5 shows results from the following exercise. I 'freeze' the state of the system (in particular, the estimate of μ_t) to 2023, and I then stochastically simulate the model featuring the land and ocean anomalies forward in time until the end of the XXI century conditional on data up to 2023. The evidence from the exercise is sobering: under such 'no change' scenario, median forecasts predict the land and ocean temperature anomalies to reach nearly 6 and 3 Celsius degrees by 2100, respectively, with the corresponding 90%-coverage credible sets equal to [4.0; 7.9] and [2.1; 3.9] degrees. In

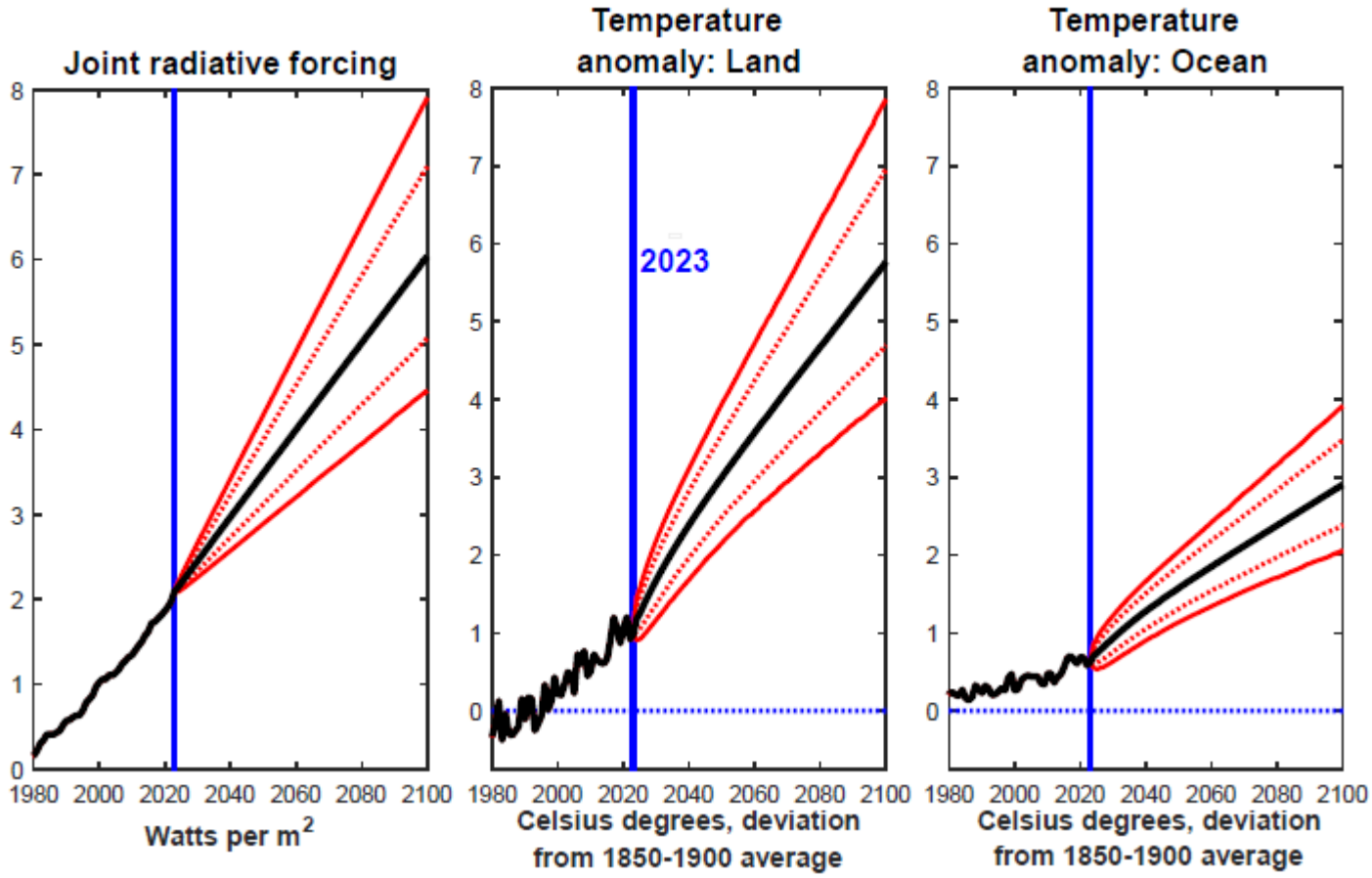

Figure 5 Forecasts under a 'no change' scenario conditional on data up to 2023 (median, and 16-84 and 5-95 credible set)

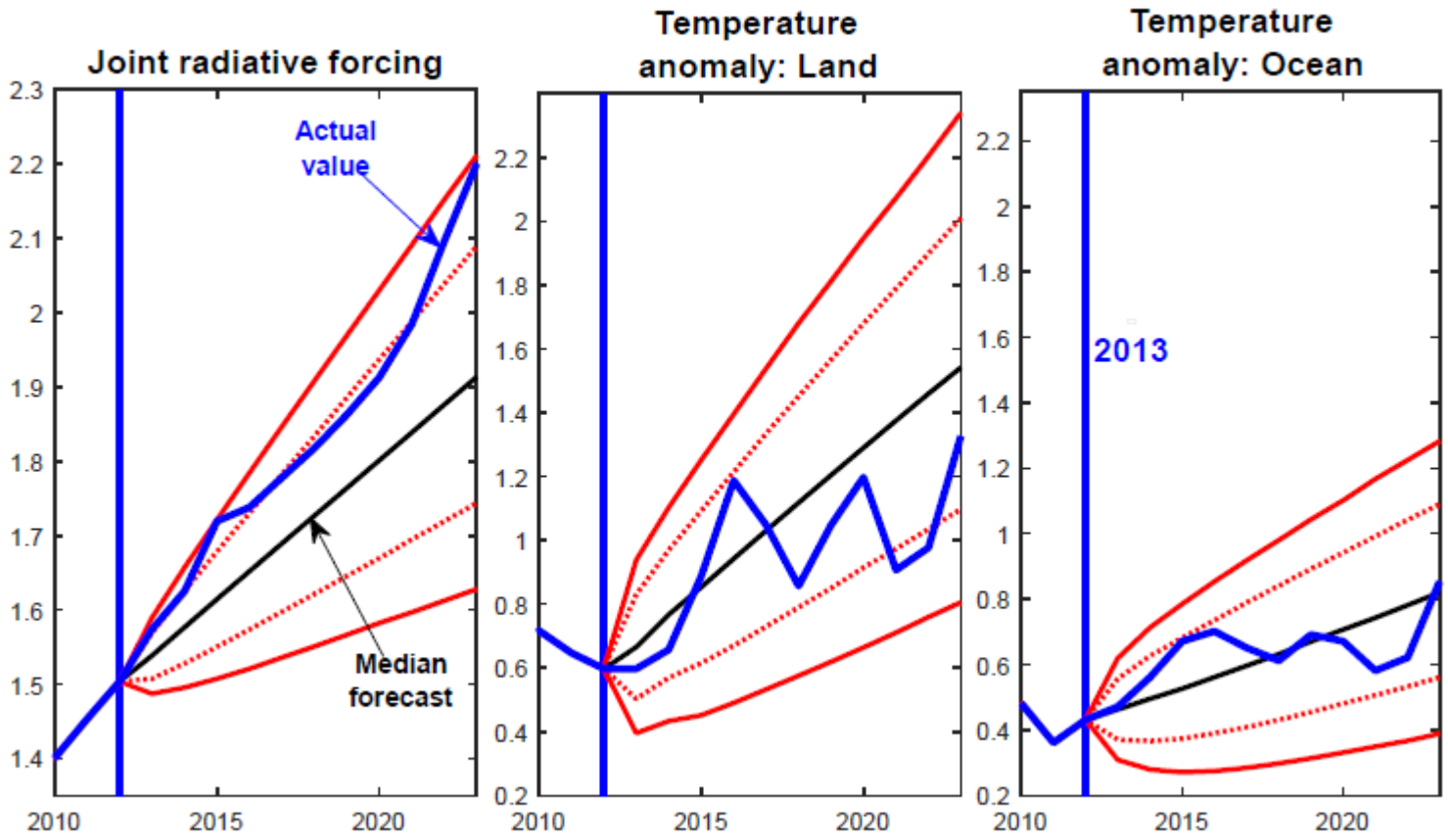

Figure 6 Forecasts under a 'no change' scenario conditional on data up to 2013 (median, and 16-84 and 5-95 credible set)

order to put these numbers into context it is worth recalling that the lower bound of the estimates for the increase in temperatures associated with the Paleocene-Eocene Thermal Maximum (PETM), about 55.5 million years ago, is about 5 Celsius degrees. During that period Antarctica was covered with tropical forests, and Arctic waters pullulated with alligators. Further, and crucially, the period of sustained carbon increase that led to the PETM is estimated to have lasted between 20 thousand and 50 thousand years. If the land temperature anomaly were to reach nearly 6 Celsius degrees (or possibly even higher values) within less than eight decades, the extent to which society could adapt—or whether it could adapt at all—is entirely open to question.

Figure 6 shows, based on the same model, results from an out-of-sample forecasting exercise. I estimate the model exactly as before based on data up to 2013, and I then stochastically simulate it into the future up until 2023 as described in the previous paragraph. The figure shows the same objects shown in Figure 5 together with the actual values of the three series. Two main findings emerge from the figure. First, consistent with the finding in Figure 3 of a continuous acceleration in the rate of change of the JRF index since the early 1980s, between 2013 and 2023 the actual value of the JRF has progressively drifted in the upper tail of the distribution of the forecast conditional on data up until 2013. On the other hand, over the same period both the land and the ocean temperature anomalies have been mostly in line with the median projection based on data up to 2013, and in fact in the most recent years both of them have been *below* the median forecast. The logical implication is that in recent years temperature anomalies have not kept up with the rapid acceleration in the rate of increase in the JRF. In turn, this implies that, to the extent that going forward the intrinsic dynamics of the system-Earth will re-establish the long-run equilibrium relationship between the JRF and temperature anomalies, the latter will have to increase *faster* than they have over the decade 2013-2023, in order to catch up with the increases in the JRF.

5.2.2 Different latitudes

Figure 7 shows results from the same unconditional exercise as in Figure 5, but this time for the global temperature anomalies (i.e., jointly for the land and the ocean) at three different latitudes, the Equator (i.e. latitude 0), 60 degrees North, and the Arctic,[15] together with the IRFs to a permanent shock to the JRF index. In line with the well-known fact that the impact of climate change gets systematically stronger at higher and higher latitudes, and it is therefore maximum at the Poles and minimum at the Equator, two main findings clearly emerge from the figure. First, the long-run response of temperatures to a one standard deviation permanent shock to the JRF index exhibits dramatic differences across latitudes: one hundred years after the

[15]Notice that NOAA does not define the Arctic as the latitude 90 degrees North, but rather as a geographical area.

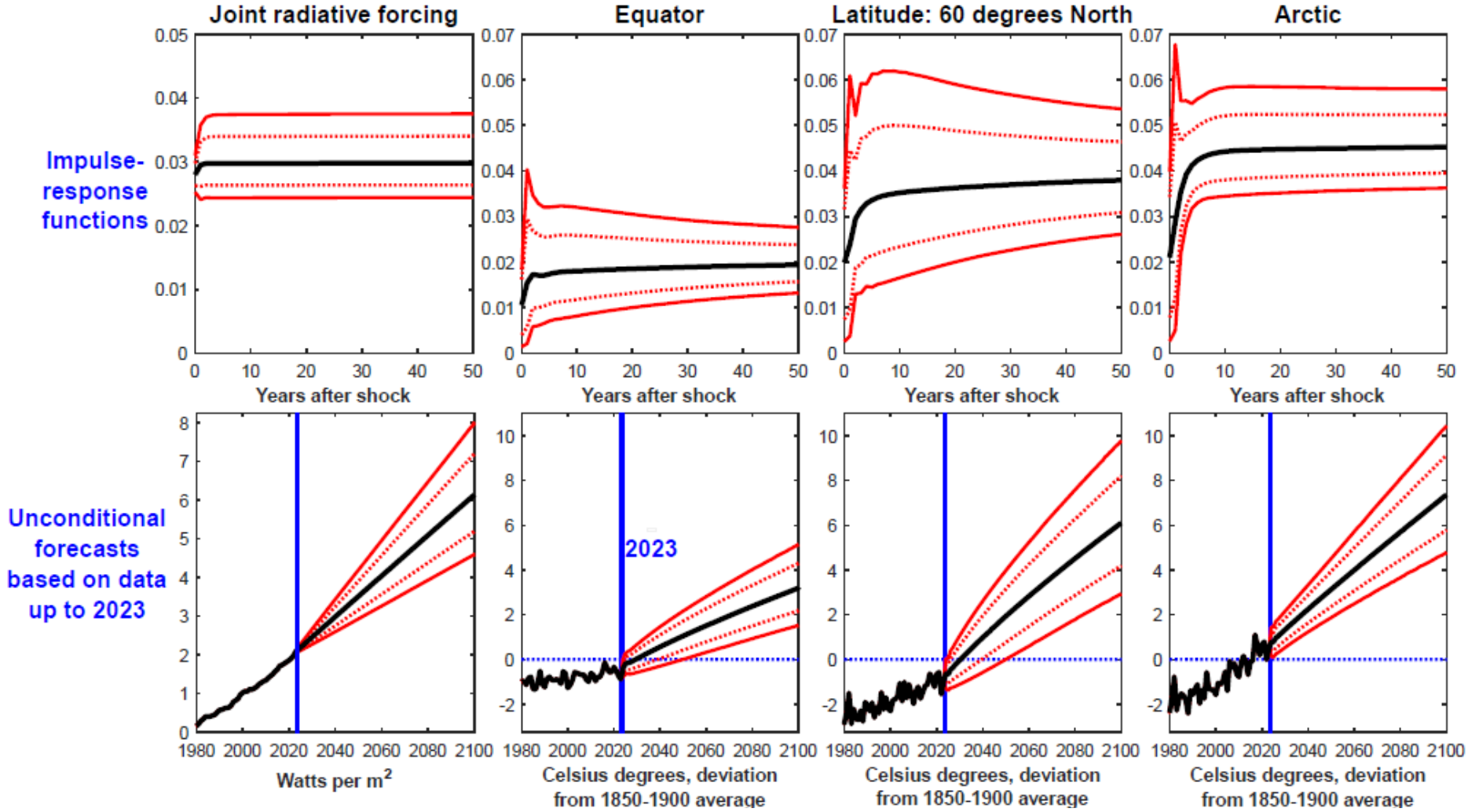

Figure 7 Northern emisphere: Impulse-response functions to a permanent shock to joint radiative forcing, and forecasts under a 'no change' scenario conditional on data up to 2023 for different latitudes (medians, and 16-84 and 5-95 credible set)

shock it is equal to about 0.02 Celsius degrees at the Equator, whereas at the 60 degrees North latitude and at the Arctic it is equal to about 0.04 and respectively 0.045 degrees. Second, as a consequence of this, unconditional forecasts under the 'no change' scenario based on data up to 2022 exhibit a corresponding significant extent of variation across latitudes. Whereas the median forecast for the Equator for the year 2100 is equal to 3.2 Celsius degrees, with a 90%-coverage credible set equal to [1.5, 5.1], the corresponding objects for the 60 degrees North latitude and the Arctic are 6.0 and [2.9, 9.8], and 7.4 and [4.7, 10.4] respectively.

5.2.3 Individual continents

Figure 8 shows the corresponding projections for individual continents. Since continents' temperature anomalies' data start in 1910, within the present context it is not possible to take as reference period the standard sample considered by the IPCC, 1850-1900. Because of this, in Figure 8 I take as reference the year 2023. The main finding in the figure is that the forecasts point towards a non-negligible extent of heterogeneity, with the median projections for 2100 being equal to about 6 degrees for Europe and Asia, about 5 for North America, and about 4 for Africa, Oceania and South America.

Table 4 Median and 16-84 and 5-95 per cent-coverage credible sets of climate sensitivity

	Land and ocean
Land	0.995 [0.844 1.227] [0.763 1.461]
Ocean	0.472 [0.394 0.588] [0.351 0.705]
	Alternative latitudes:
Equator	0.669 [0.640 0.705] [0.628 0.730]
60 degrees North	1.300 [1.249 1.340] [1.209 1.354]
Arctic	1.507 [1.457 1.576] [1.435 1.632]
	Individual continents:
Africa	0.717 [0.651 0.778] [0.602 0.820]
Asia	1.057 [0.959 1.153] [0.885 1.223]
Europe	1.022 [0.927 1.118] [0.860 1.186]
North America	0.862 [0.780 0.945] [0.723 1.002]
South America	0.705 [0.596 0.810] [0.514 0.885]
Oceania	0.624 [0.531 0.713] [0.463 0.778]

5.2.4 Climate sensitivity

Tables 4 and 5 report the median and the 16-84 and 5-95 per cent-coverage credible sets of climate sensitivity, which I define as the increase in temperature anomalies

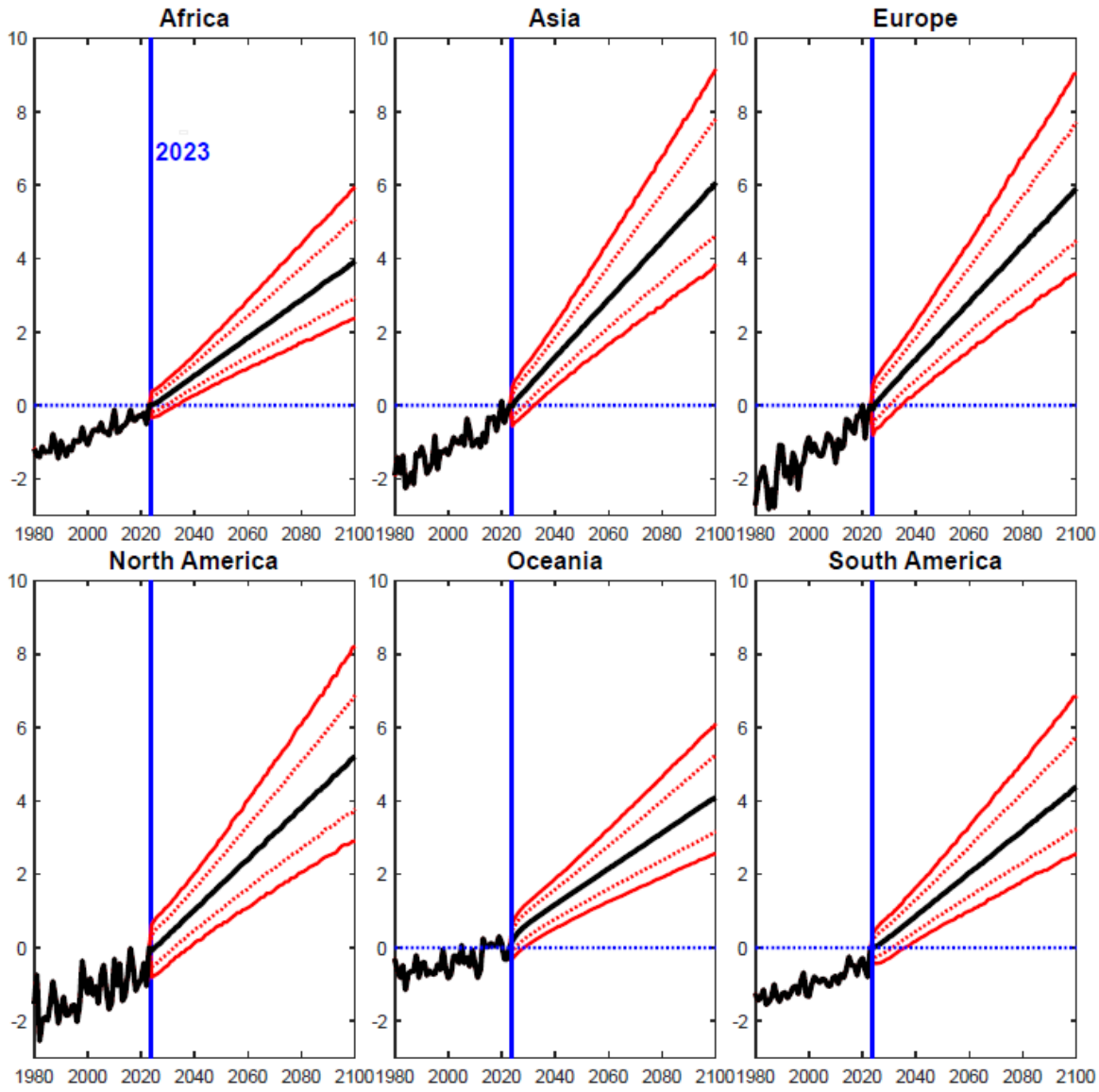

Figure 8 Forecasts under a 'no change' scenario conditional on data up to 2022 for individual continents (median, and 16-84 and 5-95 credible set: reference year for computation of temperature anomaly is 2023)

associated with a unitary increase in the JRF index.[16] Focusing on median estimates, the main findings in the table can be summarized as follows:

first, whereas the land anomaly increases virtually one-for-one with the JRF, the ocean anomaly increases only slightly less than half as much.

Second, the higher the latitude, the greater the climate sensitivity: whereas at the equator it is equal to about 1.5, at 30 degrees North it is equal to nearly 1.9, and at 60 degrees North it is equal to 2.8.

Table 5 Median and 16-84 and 5-95 percentiles of the posterior distribution of climate sensitivity	
Athens	0.5834 [0.3329 0.9410] [0.2114 1.2460]
Atlanta	0.4798 [0.2648 0.8169] [0.1599 1.1865]
Beijing	1.0767 [0.5991 1.6970] [0.3591 2.2241]
Berlin	1.1413 [0.6336 1.8776] [0.3987 2.5585]
Bern	1.3216 [0.7642 2.0567] [0.4994 2.6685]
Cape Town	0.4566 [0.1803 0.8215] [0.0448 1.0861]
Casablanca	0.9170 [0.5512 1.3526] [0.3437 1.7214]
Copenhagen	0.8938 [0.5004 1.4974] [0.3186 2.0841]
Dublin	0.5305 [0.2609 0.9169] [0.1337 1.2441]
Helsinki	1.1714 [0.6844 1.9411] [0.4527 2.7102]
Il Cairo	0.6008 [0.2868 1.0269] [0.1381 1.3666]
Lisbon	0.7761 [0.4610 1.2059] [0.3003 1.5533]
London	0.7956 [0.4088 1.3705] [0.2262 1.9088]
LosAngeles	0.7161 [0.2791 1.3311] [0.0690 1.8433]
Madrid	0.8800 [0.5238 1.3526] [0.3445 1.7218]
Mexico City	0.5824 [0.3324 0.9124] [0.1891 1.1514]
New Delhi	0.5908 [0.2497 1.1149] [0.0790 1.6083]
Oslo	0.8947 [0.5032 1.5013] [0.3198 2.0764]
Paris	1.2209 [0.7055 1.9344] [0.4517 2.5318]
Riad	1.3052 [0.8105 1.9300] [0.5508 2.4227]
Rio De Janeiro	0.8538 [0.4811 1.3210] [0.2683 1.6709]
Rome	0.7434 [0.4512 1.1403] [0.3060 1.4820]
Santiago	0.3204 [0.1006 0.7014] [-0.0012 1.0900]
Sydney	0.7108 [0.3517 1.1728] [0.1877 1.5771]
Stockholm	0.8803 [0.4924 1.4763] [0.3213 2.1006]
Vienna	1.2850 [0.7306 2.0857] [0.4759 2.7726]
Warsaw	1.1733 [0.6671 1.9581] [0.4365 2.7186]

Third, there is a significant extent of heterogeneity across continents. Asia and

[16] My definition is different from that found in the climate science literature, where climate sensitivity is defined as the increase in temperatures associated with a doubling of the concentration of CO2 in the atmosphere.

Europe are the continents characterized by the largest climate sensitivity, in both cases (based on median estimates) just slightly greater than one, so that for practical purposes in the long run the JRF and temperatures in the two continents increase virtually one-for-one. At the other extreme is Oceania, with a median estimate equal to just 0.624. It is to be noticed that the 90 per cent-coverage credible set for Oceania does not overlap with the corresponding sets for Asia end Europe, thus pointing to large and material differences between continents' climate sensitivity. In between are Africa, and North and South America, with median estimates ranging between 0.705 and 0.862, and the 90 per cent-coverage credible sets mostly overlapping with those of Asia, Europe, and Oceania.

Fourth, the range of variation for individual cities is quite extraordinary, from a minimum of 0.3 for Chile's Santiago to a maximum of 1.3 for the Swiss capital, Bern.

5.3 Conditional forecasts based on alternative assumptions about the future path of the JRF

Figure 9 shows evidence from the following exercise. I 'freeze' once again the state of the system at 2023, and I then stochastically simulate the model forward in time until the end of the XXI century conditional on (1) data up to 2023, and (2) alternative possible paths for the evolution of the JRF until 2100. I compute the conditional forecasts as in Waggoner and Zha (1999). I consider four possible scenarios for the evolution of the JRF after 2023, specifically

- the JRF being stabilized in 2050, with its first difference progressively decreasing starting in 2024 and reaching zero in 2050, and
- three scenarios in which the JRF peaks in 2050 and it is then brought back in 2100 to the level it had reached in 1990, 2000, and 2010 respectively.

I consider 2050 as the year in which the JRF either gets stabilized, or it peaks and then starts decreasing, for the following reason. Bringing the JRF back to the level it had reached in either 1990, 2000, or 2010, and possibly even stabilizing it in 2050, will require removing huge quantities of CO2 from the atmosphere. This, in turn, will require enormous amounts of (clean) energy. Under this respect, at the moment the only possible energy source that might be used for this purpose would seem to be nuclear fusion, which however is not predicted to become available (in particular at the scale required) before mid-century.

Stabilization of the JRF in 2050 leaves open the possibility that warming will reach levels that in the climate science community are widely regarded as dangerous. Specifically, the 90%-coverage credible sets for the land and ocean temperature anomalies for the year 2100 are equal to [1.9, 3.2] and [1.1, 1.7], with the corresponding median projections being equal to 2.6 and 1.4. Further, for the land anomaly (i.e.,

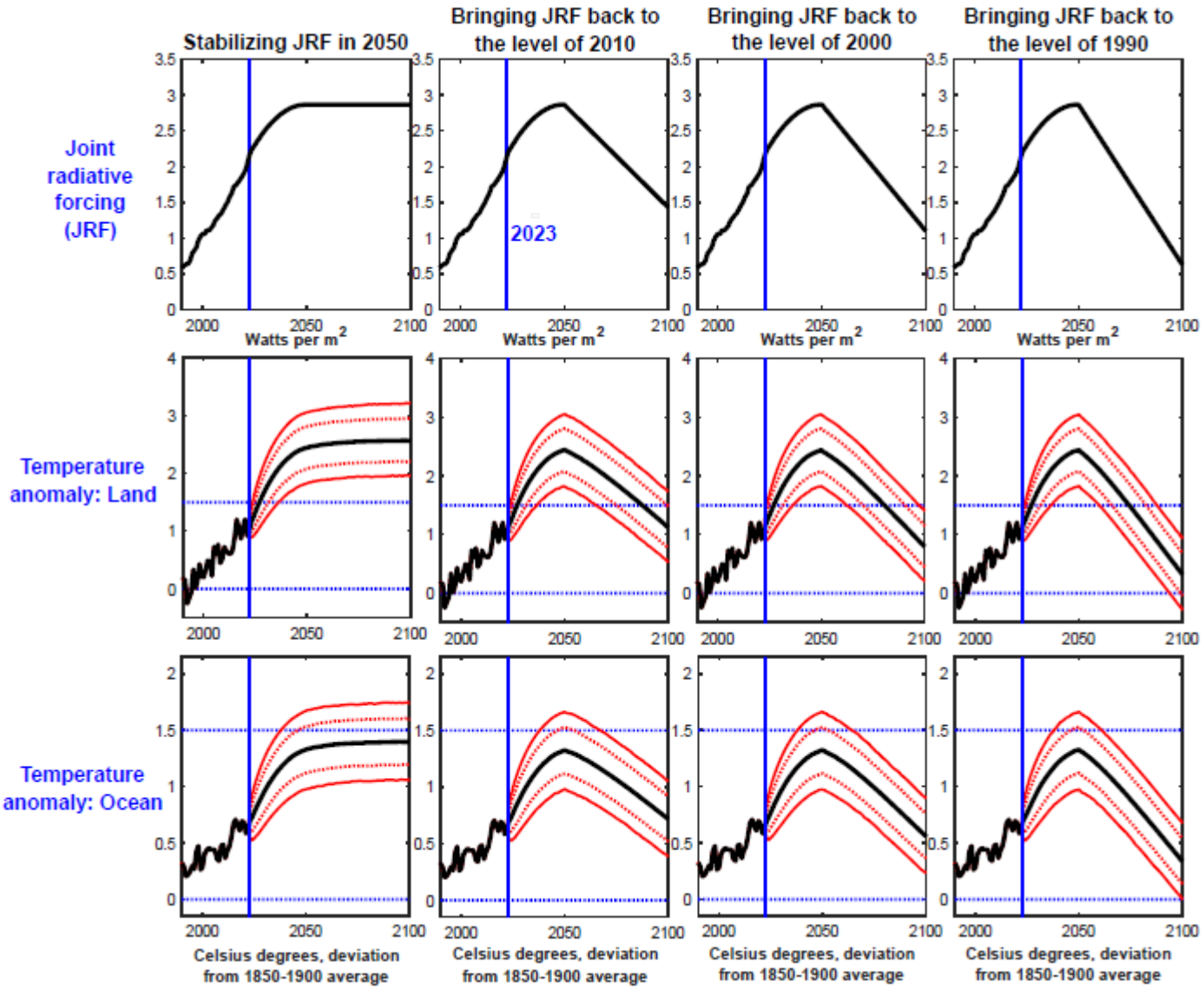

Figure 9 Forecasts conditional on data up to 2023 and alternative scenarios for the evolution of joint radiative force (median, and 16-84 and 5-95 credible set)

that which is most relevant for us humans), the entire mass of the posterior distribution lies beyond 1.5 degrees, which rightly or wrongly is regarded as the benchmark threshold beyond which warming will truly become dangerous.

Turning to the set of projections obtained by allowing the JRF index to peak in 2050, and then to decrease back to the level it had reached in either 1990, 2000, or 2010, evidence suggests that given the extent of uncertainty involved, bringing climate change under control will require to scale the JRF back to the level it had reached in the early years of the XXI century. Specifically, the 90%-coverage credible sets for the land anomaly for the year 2100 obtained by bringing the JRF back to the levels of 2000 and 2010 are [0.2, 1.4] and [0.6, 1.8] respectively, whereas the corresponding sets for the ocean anomaly are [0.2, 0.9] and [0.4, 1.1]. Clearly, to the extent that bringing climate change under control requires keeping the increases in global temperatures below 1.5 degrees, only a level of the JRF equal to that reached in the early years of the XXI century can provide sufficient reassurance that this will in fact be the case.

5.3.1 'Freezing' the JRF at the level reached in 2023

Figure 10 shows results from the corresponding exercise in which the JRF is 'frozen' at the level it reached in 2023. The evidence is sobering. Even if we were able to somehow prevent any increase in the JRF starting from 2024, still the intrinsic dynamics of the system in response to past JRF increases would produce potentially dangerous levels of warming going forward. Focusing on the land temperature anomaly, the posterior probability that it would exceed 1.5 Celsius degrees by the end of the century is equal to about two-thirds, with a median projection for the year 2100 equal to 1.8 Celsius degrees.

The evidence in Figure 10 provides stark proof that, once taking into account of statistical uncertainty, the JRF *has already exceeded*, in recent years, the level climate scientists regard as safe, and it reinforces the previous section's point that only scaling it back to the level reached in the early years of the XXI century would bring climate change under control.

6 Conclusions

In this paper I have used Bayesian VARs in order to forecast global temperature anomalies until the end of the XXI century, by exploiting their cointegration with the Joint Radiative Forcing of the drivers of climate change. My main results can be summarized as follows. The response of the JRF index to a permanent shock to its own level, which is essentially flat at all horizons, clearly suggests that the JRF is very close to a pure unit root process. By contrast, the responses of global land and ocean temperature anomalies are delayed and drawn out. In particular, in response to the shock the land and ocean temperature anomalies fully converge to their new long-run values in more than 50 years. Under a 'no change' scenario

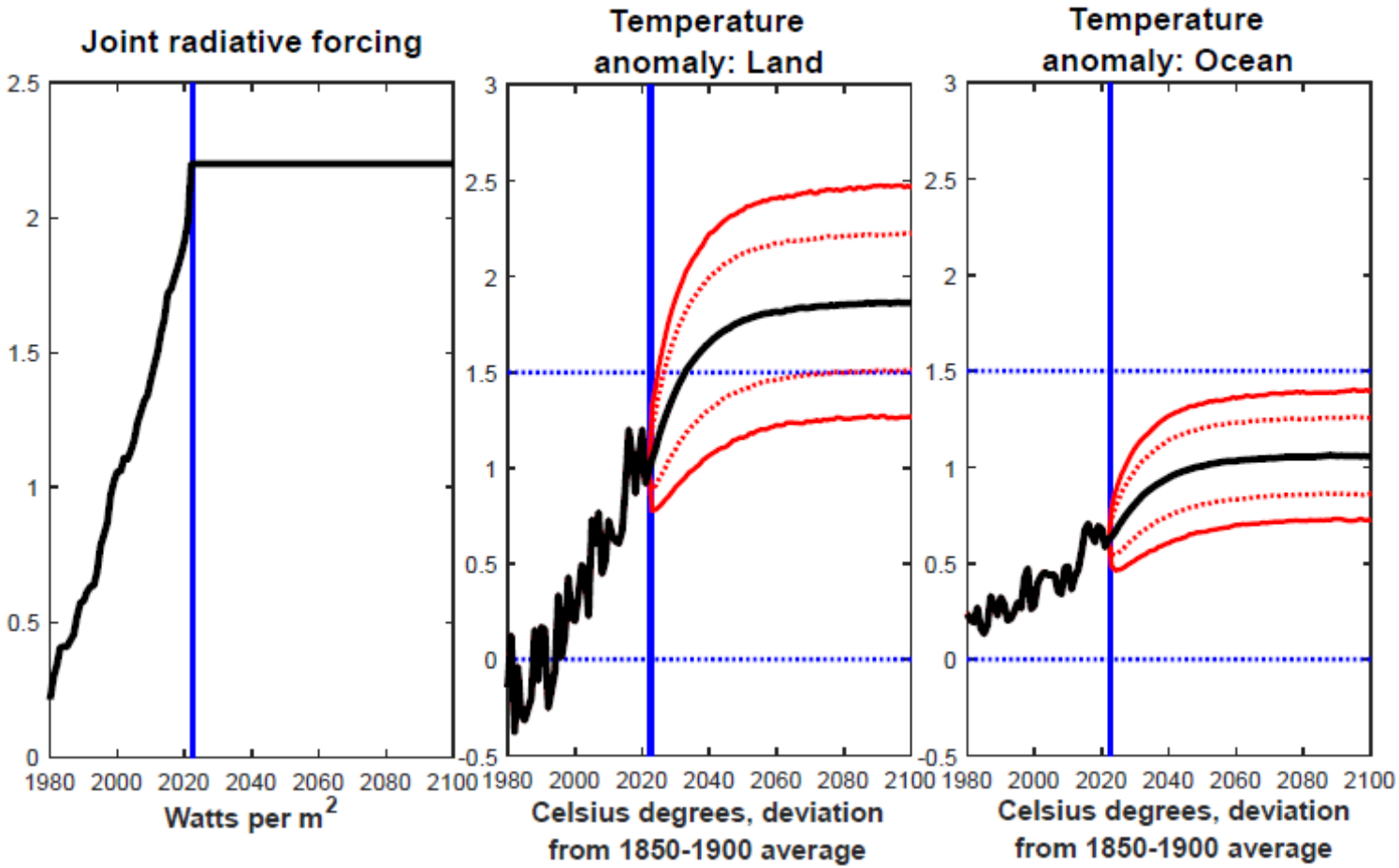

Figure 10 Forecasts conditional on data up to 2023, with joint radiative forcing frozen at the level reached in 2023 (median, and 16-84 and 5-95 credible set)

the land and ocean temperature anomalies are projected to reach nearly 6 and 3 Celsius degrees, respectively, by 2100. Forecasts conditional on alternative paths for the JRF show that, given the extent of uncertainty, bringing climate change under control will require to bring the JRF back to the level of the early XXI century. In recent years the JRF has exhibited a marked acceleration, which by 2023 has not been fully reflected in temperature anomalies yet, thus pointing towards their corresponding acceleration going forward. From a methodological point of view, my evidence suggests that previous cointegration-based studies of climate change suffer from model mis-specification. First, climate change series are clearly I(2), whereas the vast majority of studies have not tested for this possibility, and have rather assumed that they are only integrated of order one. Second, evidence suggests that fixed-coefficients I(2) cointegrated VECMs are at odds with the data, whose first differences exhibit random-walk time-variation in the mean. I model this feature via a multivariate random-walk specification for the means of the first differences, subject to the restrictions imposed by cointegration between the levels.

References

An, S., and Schorfheide, F. (2007): "Bayesian Analysis of DSGE Models", *Econometric Reviews*, 26, 113-172.

Anderson, T.W. (1951), "Estimating Linear Restrictions on Regression Coefficients for Multivariate Normal Distributions", *Annals of Mathematical Statistics*, 22, 327-351.

Bauwens, L., and M. Lubrano (1996): "Identification Restrictions and Posterior Densities in Cointegrated Gaussian VAR Systems", in Advances in Econometrics 11, Part B (JAI Press, Greenwich), 3-28

Beltrao, K. and P. Bloomfield (1987): "Determining the Bandwidth of a Kernel Spectrum Estimate", *Journal of Time Series Analysis*, 8(1), 21-38.

Benati, L. (2007): "Drift and Breaks in Labor Productivity", *Journal of Economic Dynamics and Control*, 31, 2847-2877.

Benati, L. (2008): "Investigating Inflation Persistence Across Monetary Regimes", *Quarterly Journal of Economics*, 123(3), 1005-1060.

Bruns, S.B., Csereklyei, Z. and Stern, D.I. (2020): "A Multicointegration Model of Global Climate Change", *Journal of Econometrics*, 214, 175-197.

Butler, J.H., and Montzka, S.A. (2018): "The NOAA Annual Greenhouse Gas Index (AGGI)", NOAA Earth System Research Laboratory, Boulder, CO

Cavaliere, G., A. Rahbek, and A. M. R. Taylor (2012): "Bootstrap Determination of the Cointegration Rank in Vector Autoregressive Models", *Econometrica*, 80(4), 1721-1740.

Cochrane, J.H. (1988): "How Big Is the Random Walk in GNP?", *Journal of Political Economy*, 96(5), 893-920

Coddington, O., Lean, J.L., Pilewskie, P., Snow, M., Lindholm, D. (2015): "A Solar Irradiance Climate Data Record", *Bulletin of the American Meteorological Society*, p. 1265-1282.

Corana, A., Marchesi, M., Martini,C., and Ridella, S. (1987): "Minimizing Multimodal Functions of Continuous Variables with the Simulated Annealing Algorithm," *ACM Transactions on Mathematical Software*, 13.

Dergiades, T., Kaufmann, R.K., Panagiotidis, T. (2016): "Long-Run Changes in Radiative Forcing and Surface Temperature: The Effect of Human Activity Over the Last Five Centuries", *Journal of Environmental Economics and Management*, 76, 67-85.

Diebold, F.X. and Chen, C. (1996): "Testing Structural Stability with Endogenous Breakpoint: A Size Comparison of Analytic and Bootstrap Procedures", *Journal of Econometrics*, 70(1), 221-241.

Elliot, G., T.J. Rothenberg and J.H. Stock (1996): "Efficient Tests for an Autoregressive Unit Root", *Econometrica*, 64(4), 813-836.

Engle, R. F., and C. W. Granger (1987): "Cointegration and Error Correction: Representation, Estimation, and Testing", *Econometrica*, 55(2), 251-276.

Franke, J. and W. Hardle (1992): "On Bootstrapping Kernel Spectral Estimates", *Annals of Statistics*, 20(1), 121-145.

Gelman, A., Carlin, J.B., Stern, H.S., and Rubin, D. (1995): *Bayesian Data Analysis*, New York, Chapman and Hall.

Geweke, J. (1992): "Evaluating the accuracy of sampling-based approaches to the calculation of posterior moments", in J. M. Bernardo, J. Berger, A. P. Dawid and A. F. M. Smith (eds.), *Bayesian Statistics*, Oxford University Press, Oxford, pages 169-193.

Goffe, W.L., Ferrier, G., and Rogers, J. (1994): "Global Optimization of Statistical Functions with Simulated Annealing", *Journal of Econometrics*, 60, 65-99.

Hamilton, J.D. (1986): A Standard Error for the Estimated State Vector of a State-Space Model, *Journal of Econometrics*, 33(3), 387-397.

Kaufmann, R.K., H. Kauppi, and J.H. Stock (2006): "Emissions, Concentrations, and Temperature: A Time Series Analysis", *Climatic Change*, 77: 249-278.

Kleibergen, F. and H.K. van Dijk (1994): "On the Shape of the Likelihood/Posterior in Cointegration Models", *Econometric Theory*, 10, 514-551.

Koop, G., Strachan, R., van Dijk, H., and Villani,M. (2006): "Bayesian Approaches to Cointegration", in K. Patterson and T. Mills, editors, *The Palgrave Handbook of Theoretical Econometrics*, Palgrave MacMillan

Koop, G., Léon-González, R., and Strachan, R.W. (2010): "Efficient Posterior Simulation for Cointegrated Models with Priors on the Cointegration Space", *Econometric Reviews*, 29(2), 224-242

Kopp, G. and G. Lawrence (2005): "The Total Irradiance Monitor (TIM): Instrument Design", *Solar Physics*, 230(1), 91-109.

Kopp, G, K. Heuerman, and G. Lawrence (2005): "The Total Irradiance Monitor

(TIM): Instrument Calibration", *Solar Physics*, 230(1), 111-127.

Kopp, G., Krivova, N., Lean, J., and C.J. Wu (2016): "The Impact of the Revised Sunspot Record on Solar Irradiance Reconstructions", *Solar Physics*, p. 1-18.

Johansen, S. (1988), "Statistical Analysis of Cointegration Vectors", *Journal of Economic Dynamics and Control*, 12, 231-254.

Johansen, S. (1991), "Estimation and Hypothesis Testing of Cointegration Vectors in Gaussian Vector Autoregressive Models", *Econometrica*, 69, 111-132.

Juselius, K. (2006), *The Cointegrated VAR Model: Methodology and Applications*, Oxford University Press.

Liu, H., and Rodriguez, G. (2005): "Human Activities and Global Warming: A Cointegration Analysis", *Environmental Modelling & Software*, 20, 761-773.

Luetkepohl, H. (1991): *Introduction to Multiple Time Series Analysis*, 2nd edition. Springer-Verlag.

Mann, M. (2023): *Our Fragile Moment: how lessons from the Earth's past can help us survive the climate crisis*

Nyblom, J. (1989), "Testing for the Constancy of Parameters Over Time", *Journal of the American Statistical Association*, 84(405), 223-230.

Robertson, A., Overpeck, J., Rind, D., Mosley-Thompson, E., Zielinski, G., Lean, J., Koch, D., Penner, J., Tegen, I., and Healy, R. (2001): "Hypothesized Climate Forcing Time Series for the Last 500 Years", *Journal of Geophysical Resesearch Atmosphere*, Vol. 106(D14), p. 14, 783.

Shine, K.P.R.G., Derwent, D.J., Wuebbles, D.J., and Mocrette, J.J. (1991): "Radiative Forcing of Climate", in Houghton, J.T., Jenkins, G.J., and Ephramus, J.J., editors, *Climate Change: The IPCC Scientific Assessment*, Cambridge University Press, Cambridge, pp. 47-68.

Stern, D.I. and Kaufmann, R.K. (2000): "Detecting a Global Warming Signal in Hemispheric Temperature Series: A Structural Time Series Analysis", *Climatic Change*, 47, 411-438.

Stern, D.I. and Kaufmann, R.K. (2014): "Anthropogenic and Natural Causes of Climate Change", *Climate Change*, 122, 257-269.

Stock, J. and Watson, M. (1996): "Evidence of Structural Instability in Macroeconomic Time Series Relations", *Journal of Business and Economic Statistics*, 14(1), 11-30.

Stock, J. and Watson, M. (1998): "Median-Unbiased Estimation of Coefficient Variance in a Time-Varying Parameter Model", *Journal of the Americal Statistical Association*, 93(441), 349-358.

Strachan, R. and Inder, B. (2004): "Bayesian Analysis of the Error Correction Model", *Journal of Econometrics*, 123, 307-325.

Waggoner, D.F. and Zha, T. (1999): "Conditional Forecasts in Dynamic Multivariate Models", *Review of Economics and Statistics*, 81(4), 639-651.

A Why Excluding El Niño and La Niña

Figure A.1 shows the radiative forcing of El Niño and La Niña (ENSO), together with its estimated normalized spectral density with 90%-coverage bootstrapped confidence bands.[17] Two main findings are clearly apparent from the figure:

(1) the radiative forcing of ENSO is extraordinarily noisy compared to the radiative forcing of the other drivers of climate change. For example in Figure 1 in the main text of the paper the radiative forcing of the dominant driver of climate change, CO2, goes from zero (by normalization) in 1880 to nearly 2 in 2022. By contrast, the ENSO radiative forcing in Figure 2 has a standard deviation of 1.0742, and since 1850 it has oscillated from a minimum of -2.6940 to a maximum of 2.3704.

(2) As the second panel of Figure A.1 clearly illustrates, ENSO's radiative forcing has essentially no spectral power at frequencies beyond 25 years.

The implication of (1) and (2) is that, for the present purposes, including in the JRF index the radiative forcing of ENSO shown in the first panel of Figure A.1 would uniquely add a large amount of comparatively high-frequency noise, whereas it would bring essentially *no information* about the long-horizon, low-frequency developements that are the focus of the present work. To put it differently, this would uniquely complicate the analysis, whereas it would not bring any benefit whatsoever. Because of this, in the construction of the JRF index I have decided to ignore the El Niño and La Niña phenomenon.

B Stock and Watson's (1996, 1998) Methodology for Searching for Random-Walk Time-Variation

Section 3.2 in the main text of the paper presents evidence from tests for the null hypothesis of time-invariance against the alternative of random-walk time-variation for the first difference of either the JRF index, the land temperature anomaly, or the ocean temperature anomaly, based on Stock and Watson's (1996, 1998) TVP-MUB methodology applied to the AR(p) model

$$y_t = \mu + \phi_1 y_{t-1} + \phi_2 y_{t-2} + ... + \phi_p y_{t-p} + u_t = \theta' z_t + u_t \tag{B.1}$$

where y_t is the first difference of any of the three series. I select the lag order, p, as the maximum among the lag orders selected by the Akaike and Schwartz information criteria, for a maximum possible number of lags P=20 years. In implementing the

[17] I estimate the spectral density by smoothing in the frequency domain the Fast-Fourier-Transform (FFT)-based estimator of the series' periodogram via a Bartlett spectral window. The bandwidth is selected automatically via the procedure proposed by Beltrao and Bloomfield (1987). Spectral bootstrapping is implemented via the procedure proposed by Franke and Hardle (1992). I implement 10,000 bootstrap replications.

TVP-MUB methodology I closely follow Stock and Watson (1996, 1998). Letting θ_t=[μ_t, $\phi_{1,t}$, ..., $\phi_{p,t}$]′, the time-varying parameters version of (B.1) is given by:

$$y_t = \theta_t' z_t + u_t \tag{B.2}$$

$$\theta_t = \theta_{t-1} + \eta_t \tag{B.3}$$

with η_t *iid* $N(0_{p+1}, \lambda^2\sigma^2 Q)$, with 0_{p+1} being a (p+1)-dimensional vector of zeros; σ^2 being the variance of u_t; Q being a covariance matrix; and $E[\eta_t u_t]$=0. Following Nyblom (1989) and Stock and Watson (1996, 1998), I set Q=$[E(z_t z_t')]^{-1}$. Under such a normalisation, the coefficients on the transformed regressors, $[E(z_t z_t')]^{-1/2} z_t$, evolve according to a (p+1)-dimensional standard random walk, with λ^2 being the ratio between the variance of each 'transformed innovation' and the variance of u_t.[18]

The point of departure is the OLS estimate of θ in (B.1), $\hat{\theta}_{OLS}$. Conditional on $\hat{\theta}_{OLS}$ I compute the residuals, $\hat{u}_t$, and the estimate of the innovation variance, $\hat{\sigma}^2$, and I perform an *exp*-Wald test for a single break in the mean of y_t at an unknown point the sample as in (e.g.) Bai and Perron (1998, 2003) by regressing y_t on a constant, using either Newey and West's (1987) or Andrews' (1991) covariance matrix estimator to control for possible autocorrelation and/or heteroskedasticity in the residuals. I estimate the matrix Q as in Stock and Watson (1996) as

$$\hat{Q} = \left[T^{-1} \sum_{t=1}^{T} z_t z_t' \right]^{-1}.$$

I consider a 50-point grid of values for λ over the interval [0, 0.15], which I call Λ. For each $\lambda_j \in \Lambda$ I compute the corresponding estimate of the covariance matrix of η_t as $\hat{Q}_j$=$\lambda_j^2 \hat{\sigma}^2 \hat{Q}$, and conditional on $\hat{Q}_j$ I simulate model (B.2)-(B.3) 10,000 times as in Stock and Watson (1996, section 2.4), drawing the pseudo innovations from pseudo random *iid* $N(0, \hat{\sigma}^2)$. For each simulation, I compute an *exp*-Wald test (obviously, without however applying the HAC correction) thus building up its empirical distribution conditional on λ_j. Based on the empirical distributions of the test statistic I then compute the median-unbiased estimate of λ as that particular value of λ_j which is closest to the statistic I previously computed based on the actual data. I compute the p-value based on the empirical distribution of the test conditional on λ_j=0. Finally, for reasons of robustness I consider three alternative values of trimming, 15, 25, and 33 per cent.

In line with the previous discussion, I partially simulate the JRF index 10,000 times, and I implement the previously described procedure based on each partially simulated series. For the JRF index the table therefore reports the median and the 5th and 95th percentiles of the Monte Carlo distribution of the p-values.

[18] To be precise, given that the Stock-Watson methodology is based on local-to-unity asymptotics, λ is actually equal to the ratio between τ, a small number which is fixed in each sample, and T, the sample length.

C The I(2) Cointegrated VECM

Let the standard I(1) cointegrated VECM representation for a $(N \times 1)$ vector of I(1) series Y_t be (abstracting form the intercept term)

$$\Delta Y_t = \Gamma_1 \Delta Y_{t-1} + ... + \Gamma_{p-1} \Delta Y_{t-p+1} + \Pi Y_{t-1} + u_t, \tag{C.1}$$

with $\Pi = \alpha\beta'$, where β is the matrix of the cointegration vectors, α is the matrix of the loading coefficients, $E[u_t' u_t] = \Sigma$, and the rest of the notation is standard. Expression (C.1) implies the following restricted VAR(p) representation in levels for Y_t,

$$Y_t = \Pi_1 Y_{t-1} + ... + \Pi_p \Delta Y_{t-p} + u_t, \tag{C.2}$$

The I(2) cointegrated VECM representation for Y_t is then given by

$$\Delta^2 Y_t = \Psi_1 \Delta^2 Y_{t-1} + ... + \Psi_{p-2} \Delta^2 Y_{t-p+2} + \Pi Y_{t-1} - \Gamma \Delta Y_{t-1} + u_t, \tag{C.3}$$

with

$$\Psi_i = -\sum_{j=i+1}^{p-1} \Gamma_j \tag{C.4}$$

for $i = 1, 2, ..., p\text{-}2$. Based on Γ and the Ψ_i's in (C.3), it is possible to recover the Γ_i's in (C.1) as follows. Since

$$\Gamma = I_N - \sum_{i=1}^{p-1} \Gamma_i \tag{C.4}$$

it can be shown that

$$\begin{aligned}
\Gamma_1 &= \Psi_1 - \Gamma + I_N \\
\Gamma_2 &= \Psi_2 - \Psi_1 \\
\Gamma_3 &= \Psi_3 - \Psi_2 \\
&... \\
\Gamma_{p-2} &= \Psi_{p-2} - \Psi_{p-1} \\
\Gamma_{p-1} &= -\Psi_{p-2}
\end{aligned}$$

I estimate the model via Bayesian methods, by combining the log-likelihood of the data with a minimal set of inequality restrictions (discussed in sub-section 4.4 of the main text) on the impulse-response functions (IRFs) of the three series to a permanent shock to the JRF index. In practice, this means that I perform MLE estimation subject to the restriction that I reject models (i.e. draws) that do not satisfy the restrictions on the IRFs. So, although I adopt a Bayesian approach, which allows me to reject draws that do not satisfy the restricitions on the IRFs, in fact I do not specify a single prior for any parameter. I numerically maximize the log-likelihood of the data via simulated annealing exactly as in Benati (2008). Following

Goffe, Ferrier, and Rogers (1994) I implement simulated annealing via the algorithm proposed by Corana, Marchesi, Martini and Ridella (1987). I then stochastically map the log-likelihood based on Random Walk Metropolis (RWM). In implementing the RWM algorithm I exactly follow An and Schorfheide (2007, Section 4.1), with the only difference that the jump to the new position in the Markov chain is accepted or rejected based on a rule which does not involve any Bayesian priors, as it uniquely involves the likelihood of the data. I calibrate the covariance matrix's scale factor based on the methodology proposed by Benati (2008, Appendix C).

I run a burn-in pre-sample of 1,000,000 draws which I then discard. I then generate 10,000,000 draws, which I 'thin' by sampling every 1,000 draws in order to reduce their autocorrelation. This leaves 10,000 draws from the ergodic distribution which I use for inference. The fraction of accepted draws is very close to the ideal one, in high dimensions, of 0.23 (see Gelman, Carlin, Stern, and Rubin, 1995).

I check convergence of the Markov chain based on Geweke's (1992) inefficiency factors (IFs) of the draws from the ergodic distribution for each individual parameter. For all parameters the IFs are equal to at most 3-4, well below the values of 20-25 which are typically taken to indicate problems in the convergence of the Markov chain.

Table A.1 Bootstrapped *p*-values for Elliot, Rothenberg, and Stock's (1996) unit root tests for temperature anomalies for different latitudes					
	p=1	p=2	p=3	p=4	p=5
	Latitude: 0				
In levels, without time trend	0.0002	0.0485	0.2324	0.4663	0.6551
In levels, with time trend	0.0000	0.0015	0.0084	0.0499	0.1236
In first differences, without time trend	0.0000	0.0000	0.0000	0.0000	0.0000
	Latitude: 30 North				
In levels, without time trend	0.6392	0.8624	0.9605	0.9796	0.9810
In levels, with time trend	0.3206	0.6153	0.7910	0.8376	0.8038
In first differences, without time trend	0.0000	0.0000	0.0000	0.0000	0.0000
	Latitude: 60 North				
In levels, without time trend	0.0627	0.3521	0.6822	0.8909	0.9656
In levels, with time trend	0.0027	0.0280	0.1427	0.3346	0.4500
In first differences, without time trend	0.0000	0.0000	0.0000	0.0000	0.0000

Table A.2 Simulated *p*-values for Stock and Watson's tests for the null of time-invariance against the alternative of random-walk time-variation in the mean of the first differences of the series			
	Latitude:		
HAC correction:	0	30 North	60 North
	Trimming: 0.15		
Newey and West (1987)	0.4040	0.0026	0.0008
Andrews (1991)	0.3680	0.0022	0.0000
	Trimming: 0.25		
Newey and West (1987)	0.3618	0.0064	0.0002
Andrews (1991)	0.3302	0.0054	0.0000
	Trimming: 0.33		
Newey and West (1987)	0.3514	0.0158	0.0030
Andrews (1991)	0.3232	0.0118	0.0000
[a] Mean and median of the Monte Carlo distribution of p-values, and fraction of p-values smaller than 10%.			

Figures for Appendix

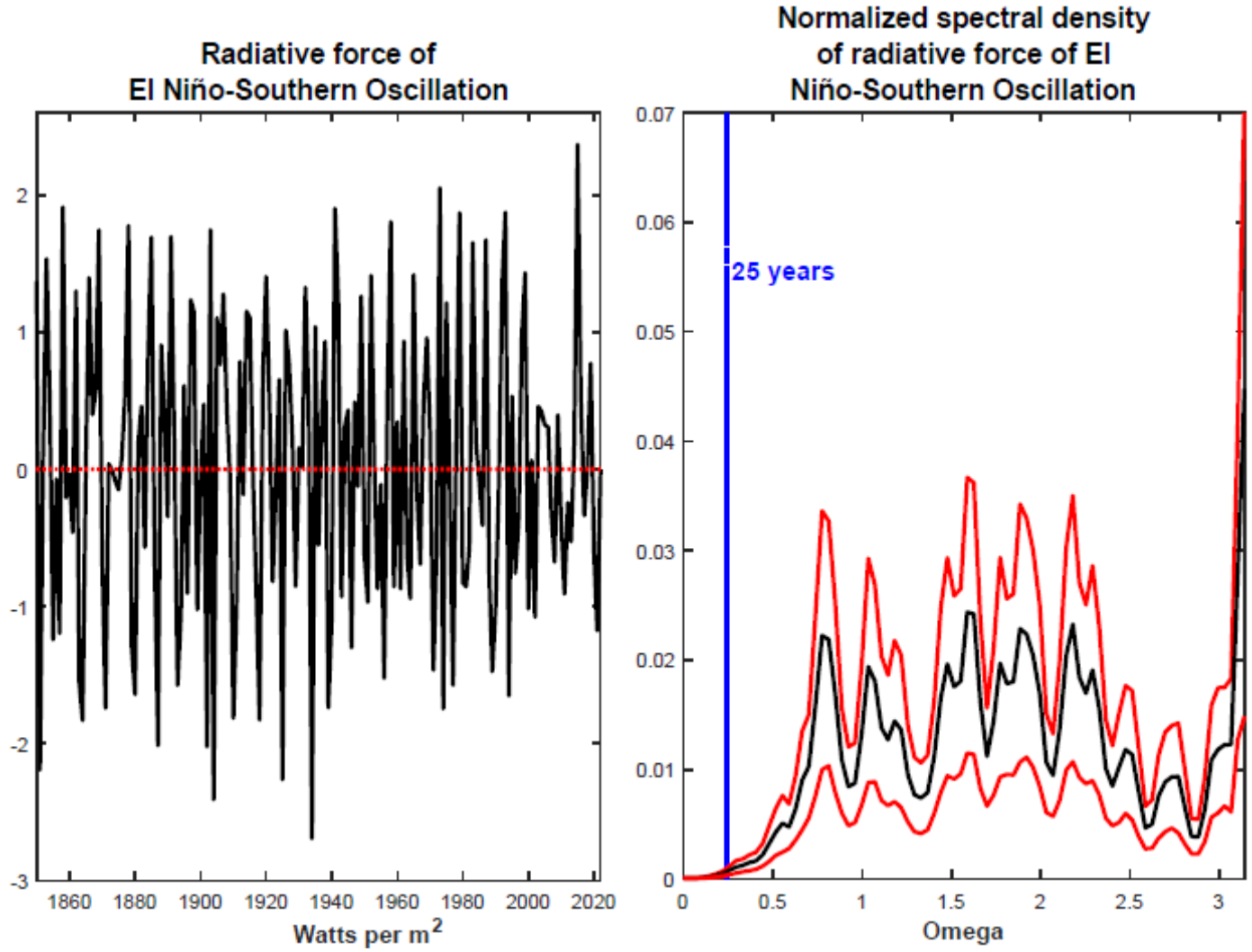

Figure A.1 Radiative force of El Niño-Southern Oscillation: raw series and normalized spectral density (with 90%-coverage bootstrapped confidence bands)

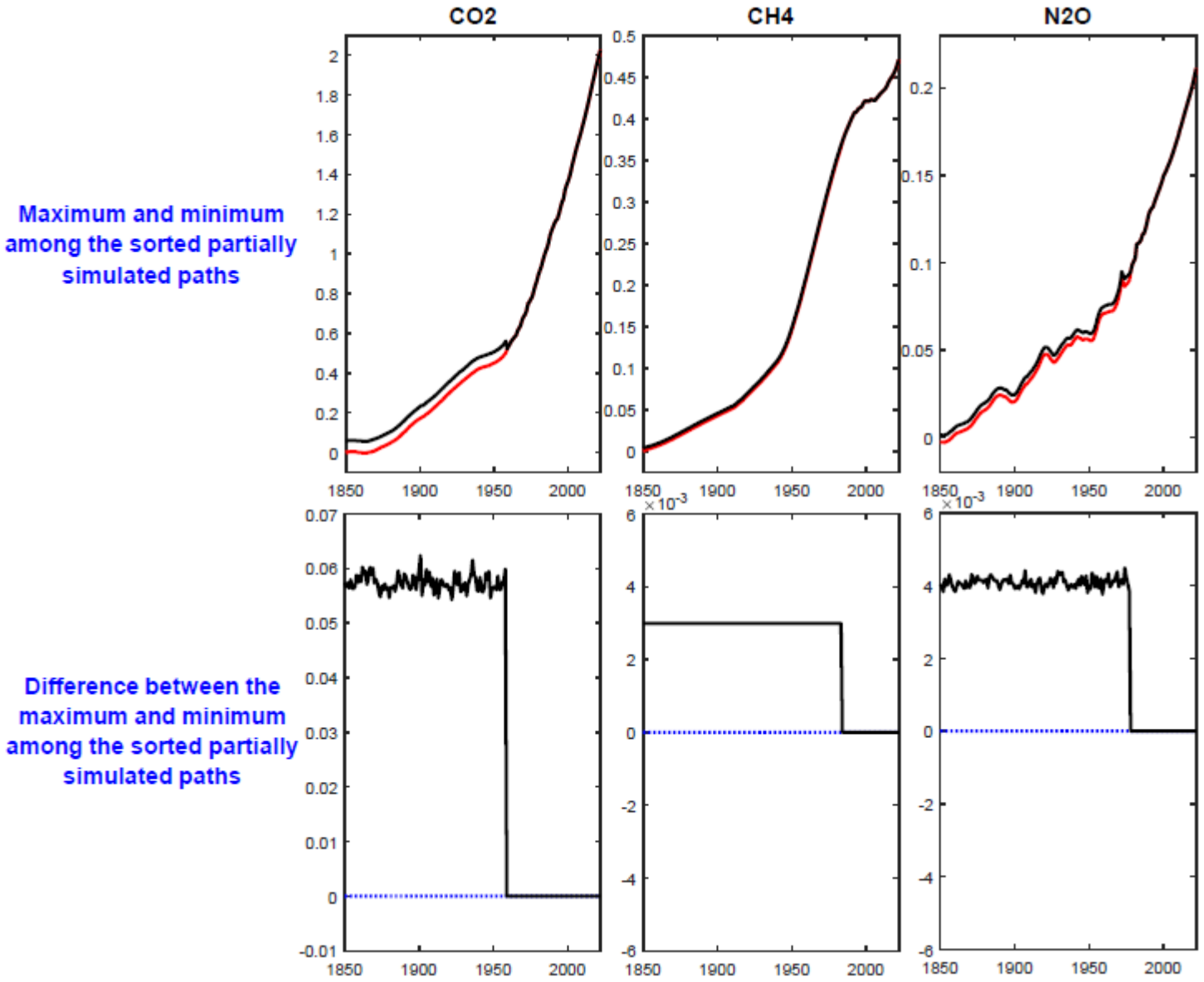

Figure A.2 Evidence on the close similarity between alternative partially simulated series for CO2, NH4, and N2O: maximum and minimum among the sorted partially simulated paths out of 100,000 simulations