UNIVERSITÄT
BERN

Faculty of Business, Economics
and Social Sciences

Department of Economics

Forecasting Global Temperatures by Exploiting Cointegration with Radiative Forcing

Luca Benati

23-08

May, 2023

DISCUSSION PAPERS

Schanzeneckstrasse 1
CH-3012 Bern, Switzerland
http://www.vwi.unibe.ch

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 most favorable median forecast predicts the land temperature anomaly to reach 5.6 Celsius degrees in 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 reached in the early years of the XXI century. From a methodological point of view, my evidence suggests that previous cointegration-based studies of climate change suffer from model mis-specification.
Keywords: Climate change; Bayesian VARs; cointegration; forecasting; conditional forecasts.
JEL Classification: E2, E3.

*I wish to thank Giuseppe Cavaliere, Lutz Kilian and Helmut Luetkepohl for useful suggestions. 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 summer of 2022, associated with heatwaves and droughts of an unprecedented spread and intensity especially across Europe and North America, highligthed in the starkest possible way the severity of the threat posed by climate change. 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?

In this paper I use Bayesian VARs in order to forecast global temperature anomalies for both the land and the ocean, and different latitudes, until the end of the XXI century, by exploiting their cointegration with the Joint Radiative Forcing (JRF) of the drivers of climate change. I obtain three main sets of results:

first, 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 fixed-coefficients I(2) cointegrated VECM models, 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.

Second, under a 'no change' scenario, the most favorable median forecast predicts the land temperature anomaly to reach 5.6 Celsius degrees in 2100, with the 90%-coverage credible set stretching from 3.6 to 8.7 degrees. In order to put these numbers into context it is worth recalling that the average increase in temperatures associated with the Paleocene-Eocene Thermal Maximum (PETM), about 55.5 million years ago, is estimated to have been between 5 and 8 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 5.6 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 for the year 2100 ranging from 2.8 Celsius degrees for the Equator, to 4 and 5.8 degrees for the 30 and 60 degrees North latitudes.

Third, 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 levels reached in the early years of the XXI century.

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 global temperatures forecasts based on comparatively small cointegrated VECMs, featuring at most four series. The contrast between the two approaches bears some similarities with the corresponding contrast, within Economics, between structural VAR (SVAR) methods and 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 time-series representation of the data, and impose upon this structure a minimal set of restrictions that allows to make 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. In particular, 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 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 (i.e. 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 misspecification. 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 data based on continuous, direct observations (which are available for the most recent past) with data that have been spline-interpolated based on discontinuous observations (which are the only data that are available for the more distant past); and the construction of the index of Joint Radiative Forcing. Section 3 presents statistical evidence on the stochastic properties of the series under investigation: the Joint Radiative Forcing 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 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, and outlines possible directions for future research.

2 The Data

2.1 Data sources

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 Climat Change (IPCC) is 1850-1900, I adjust the NOAA series by rescaling them accordingly.

By the same token, annual data since 1850 for the global 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 latitude for a grid of 38 possible longitudes, from -180 to +180 degrees.[1] 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.

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 levels of CO2, CH4, and N2O 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).

[1] 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.

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).[2] 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).[3]

2.2 Linking data based on continuous observations with interpolated data based on discontinuous observations

For three climate change drivers—CO2, NH4, and N2O—I link spline-interpolated data based on discontinuous observations with data based on continuous direct observations. 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.

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 '

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

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

to an interpolated series.) Finally, I construct the linked series for the overall period 1880-2022 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.

2.3 Construction of the index of Joint Radiative Forcing

Once each driver of climate change has been converted into radiative forcing, I construct the aggregate index of Joint Radiative Forcing (JRF) by simply summing up the individual components, with the single exception of El Niño and La Niña (ENSO), which I ignore for the reasons I discuss in Appendix A.1.[4] 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 2.4, the series for CO2, NH4, and N2O are in part stochastically simulated over the first portions of the respective sample periods, by adding to 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 sample, being based on direct 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 are clearly apparent from the figure. First, since 1850 CO2, CH4 and SOx have been by far the dominant drivers of climate change. Second,

[4] 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 in the JRF index the radiative forcing of ENSO 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.

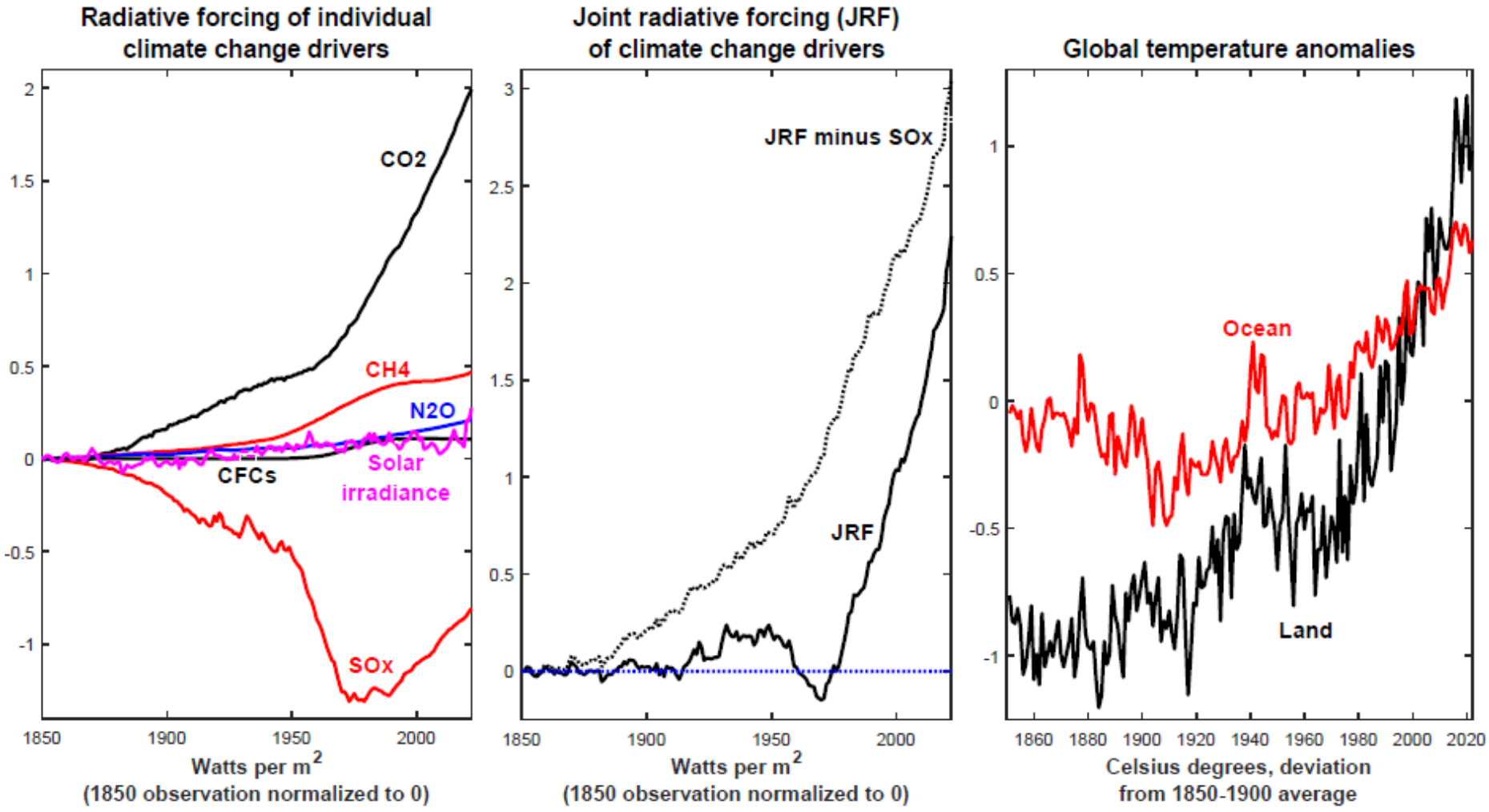

Figure 1 The raw data

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 JRF 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[5] 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 continuous 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.

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

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 $T^{\text{Land}} = \alpha + \beta JRF$, where the notation is obvious.

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.

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 of the land and ocean temperature anomalies, whereas Table A.1 in the Appendix shows the corresponding evidence for the temperature anomalies for the three mentioned latitudes. For the temperature anomalies series, for which we have data based on continuous observations for the entire sample since 1850, I perform the unit root 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.

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 six 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. The only partial exception is the temperature anomaly for the 0 degrees latitude, for which lack of rejection is very strong when a time trend is not included, whereas it is weak when it is included. The obvious explanation is that, as it is well known, the impact of climate change in terms of increases in temperatures is greater the higher the latitude, and it is therefore maximum at the Poles and minimum at the Equator.

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.

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

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 a reasonable characterization of the data, which has in fact been adopted by the vast majority of the 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.

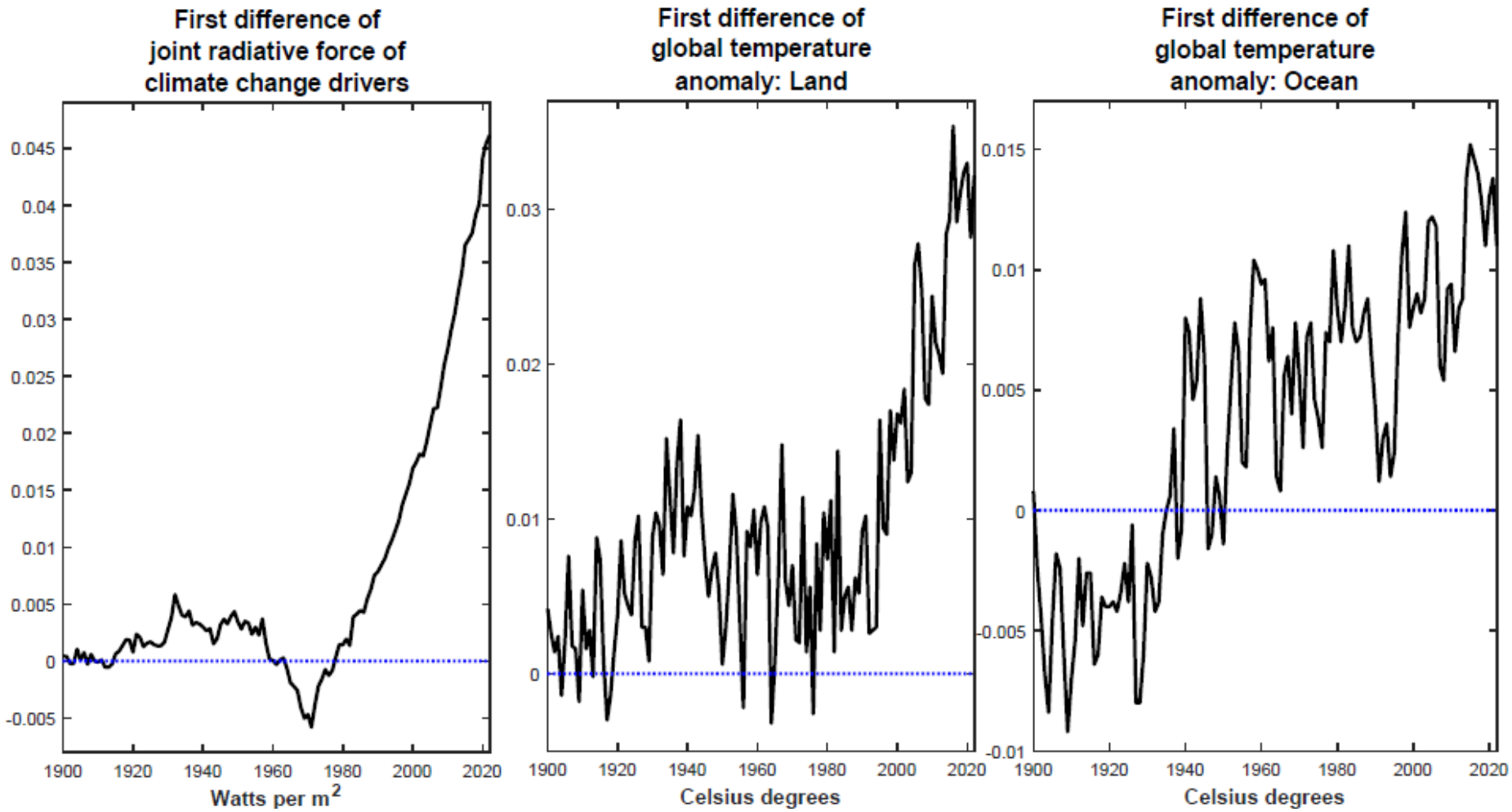

Figure 2 Fifty-years rolling averages of the first differences of the series

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 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. 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, as we would expect from the fact that the ocean anomaly behaves essentially as a low-frequency component of the land anomaly. Evidence for the temperature anomaly series pertaining to the three different latitudes is qualitatively the same, and it is available upkn 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). 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 in fact it is 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, Tables 2 and A.2 in the Appendix report evidence from Stock and Watson's (1996, 1998) tests of the null hypothesis of no time-variation

in the mean for the first differences of any of the series, against the alternative of random-walk time variation. 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, which are based on continuous measurements, 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.[6] 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

[6] A partial exception is the temperature anomaly series for the 0 latitude. As shown in Table A.2, strong evidence of random-walk time-variation is detected based on 25 and 33% trimming, but not based on 15% trimming. Once again, the obvious explanation is that the impact of climate change is greater the higher the latitude, and it is therefore minimum at the Equator.

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.2 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.

Table 3 reports results fom the following exercise. I generate, as previously described, 1,000 partially simulated series for the JRF index. Based on any of them and the land and ocean temperature anomalies I then estimate I(2) cointegrated VECM models 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.[7] 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 the 1 million models for samples of length equal to the length of the actual sample I am working with (1850-2022), and based on each of them I perform the same Stock and Watson's (1996, 1998) tests I have performed based on the actual data. In estimation I either impose the restriction that the land and ocean temperature anomalies share a common cointegration vector with the JRF index, or I allow the two cointegration vectors to be different. The results in Table 3 are based on allowing for distinct cointegration vectors, but the corresponding evidence based on imposing a common cointegration vector are qualitatively the same, and they are available upon request.

Table 3 reports, for any of the three series, the mean and the median of the Monte Carlo distribution of the simulated p-values, together with the fraction of simulations for which the p-values have been below 10%, out of the 1 million Monte Carlo simulations. The evidence in the table is very clear: if the true DGP had been the I(2) cointegrated VECM model, obtaining the results reported in Table 2 would have been extremely unlikely. For the sake of the argument, let us focus 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 an I(2) cointegrated VECM model.

[7] In particular, the main information I impose pertains to (i) the cointegration coefficient between the JRF index and either of the two temperature anomalies, with the prior being calibrated based on the estimates reported in Table II of Kaufmann, Kauppi, and Stock (2006); and (ii) the fact that, as it is well known, the reaction of the ocean anomaly to increases in radiative forcing is significantly slower than the corresponding reaction of the land anomaly.

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. Finally, evidence for the system featuring the JRF index and the three temperature anomalies pertaining to the 0, 30 degrees North, and 60 degrees North latitudes is qualitatively the same, and it is available upon request.

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 either 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, or the previously mentioned I(2) cointegrated VECM model. 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), (ii) they are cointegrated, and (iii) the cointegration residual in levels is I(0). 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) 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.

Based on the model featuring a multivariate random-walk specification for the time-varying equilibrium levels of the first differences of the three series and allowing for different cointegration vectors between the JRF index and either the land or the ocean temperature anomaly, the 90%-coverage confidence interval for the cointegration coefficient for the land anomaly is [-0.9521 -0.6000], whereas for the ocean anomaly it is [-2.2411 -1.4546], reflecting the much slower rate of warming of the oceans over the sample period. The corresponding confidence intervals based on the model imposing a common cointegration vector are [-0.9529 -0.5990] and [-2.2827 -1.4090], respectively. Notice that in either case the confidence intervals for the land and ocean anomaly are not overlapping. Taken at face value this would imply that the land and ocean anomaly, 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. Accordingly, in what follows I will consider both possibilities.

Evidence for the system featuring the JRF index and the temperature anomalies at different latitudes is qualitatively the same. In particular, 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 well-known fact that the higher the latitude, the greater the increase in temperatures has been.

Evidence based on the I(2) cointegrated VECM model is qualitatively the same, and it is available upon request. 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. 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 shown in the previous section, 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 (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 two 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.

Finally, as previously mentioned, I consider two alternative model specifications, in which I either allow b_{Land} to be different from b_{Ocean}, or I impose the restriction that they be the same.

4.3 Estimation

I estimate all of the models the model 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 the restricitions on the IRFs, 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).[8] 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 modfel's coefficients, as it uniquely involves the restricted likelihood of the data.[9] I calibrate the covariance matrix's scale factor based on the methodology proposed by Benati (2008,

[8] 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.

[9] 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

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

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. 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 non-negative impulse-response functions (IRFs) for all series 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,

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

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

The rationale for (I) is the well-known lag of the ocean temperature anomaly on its corresponding land counterpart. As for (II), its 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.

---

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.

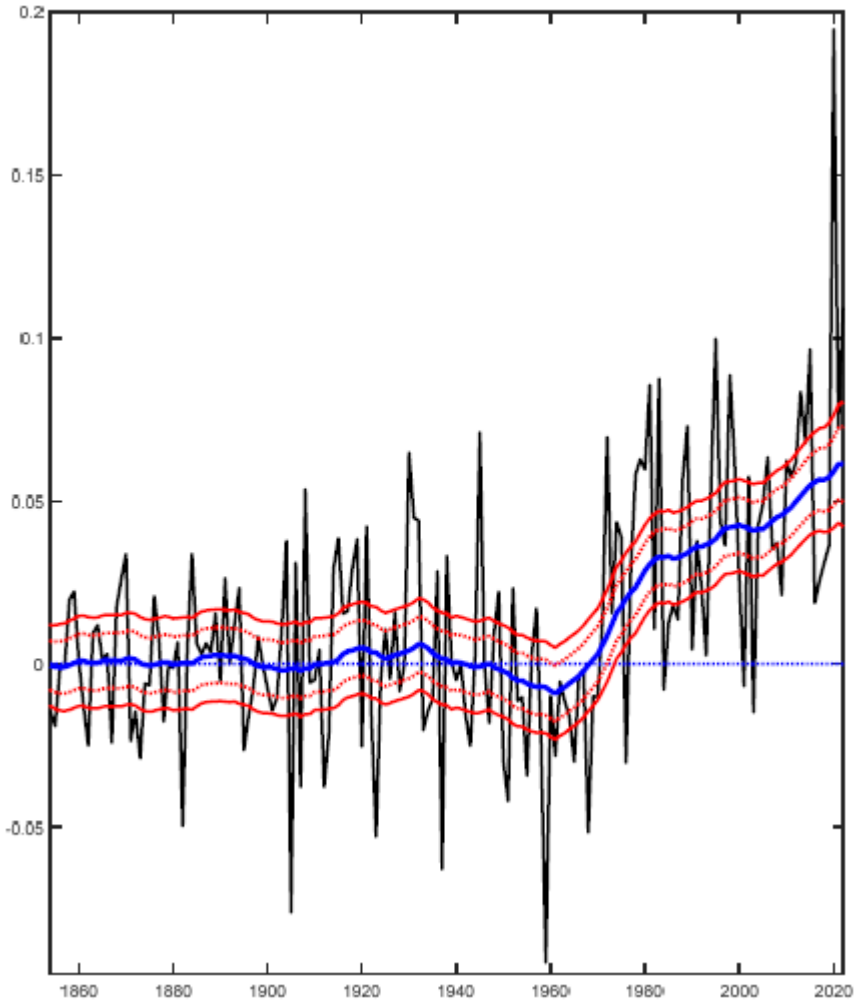

Figure 3 First difference of joint radiative force, and two-sided estimate of Mu(t) (median, and 16-84 and 5-95 credible set)

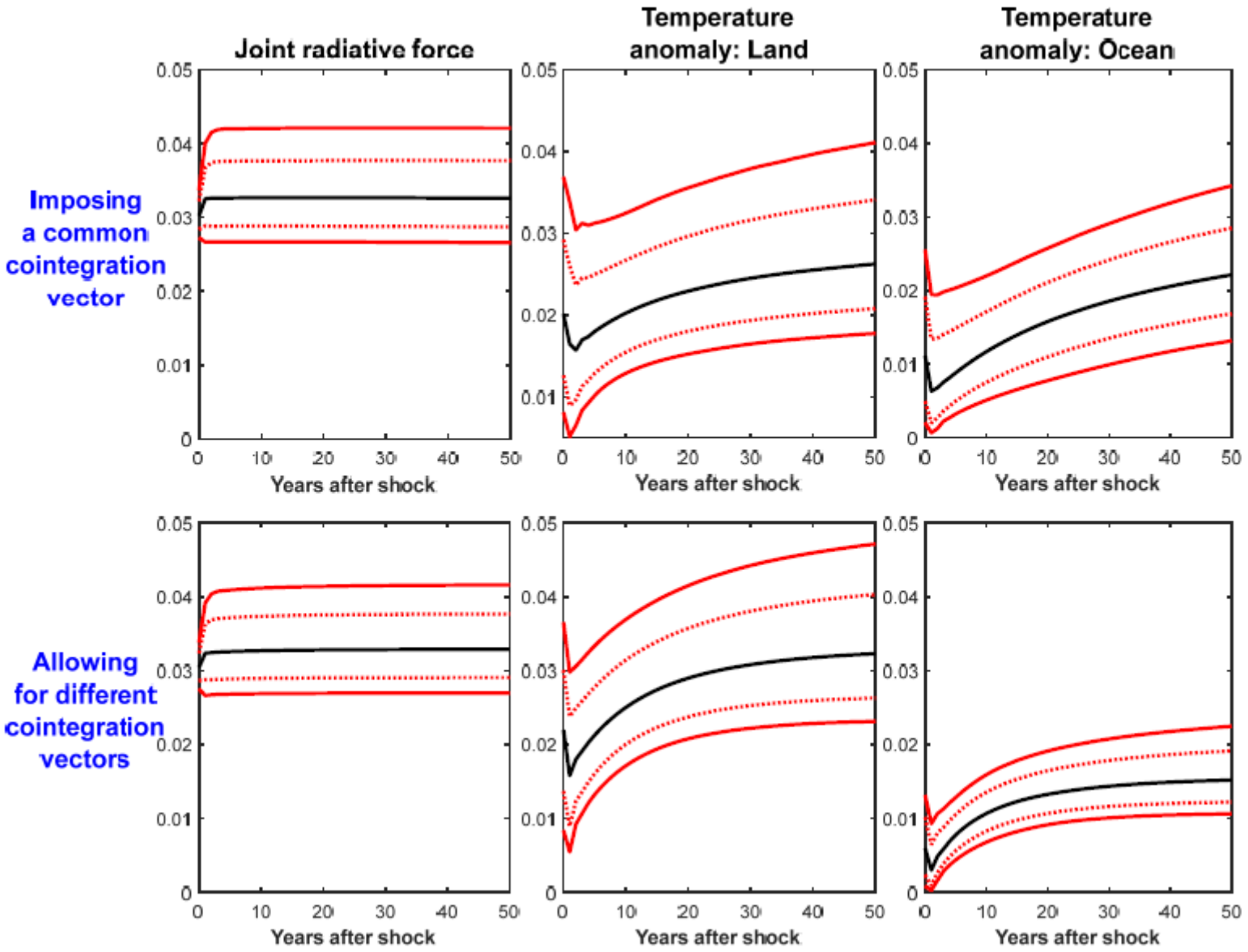

Figure 4 Impulse-response functions to a permanent shock to the level of joint radiative force (median, and 16-84 and 5-95 credible set)

As for the other systems, the only restriction I impose is that at all horizons the normalized IRFs of all temperature anomalies lie below the normalized IRF of the JRF index. On the other hand, I do not impose any restriction among the temperature anomalies' IRFs themselves.

I impose all of the previously discussed restrictions by rejecting all 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 credible sets of the posterior distribution.[10] 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 shock to the level of the JRF index. The response of the JRF index itself is essentially flat at all horizons, thus clearly suggesting that the JRF index is a pure unit root process. The responses of the two temperature anomalies series are as expected, and they reflect the restrictions I imposed upon them. In particular, the response of the ocean anomaly is slower than that of the land anomaly, which in turn converges to its new long-run equilibrium only about 40-50 years after the shock.

5.2 Unconditional forecasts based on data up to 2022

Figure 5 shows results from the following exercise. I 'freeze' the state of the system (in particular, the estimate of μ_t) to 2022, and I then stochastically simulate the model forward in time until the end of the XXI century conditional on data up to 2022. The evidence from the exercise is sobering. Starting from the model imposing a common cointegration vector for the land and ocean temperature anomalies (top row), the

[10] The estimate in Figure 3 is based on the model allowing for different cointegration vectors for the land and ocean anomalies. Estimates based on the alternative model imposing a common cointegration vector is qualitatively the same.

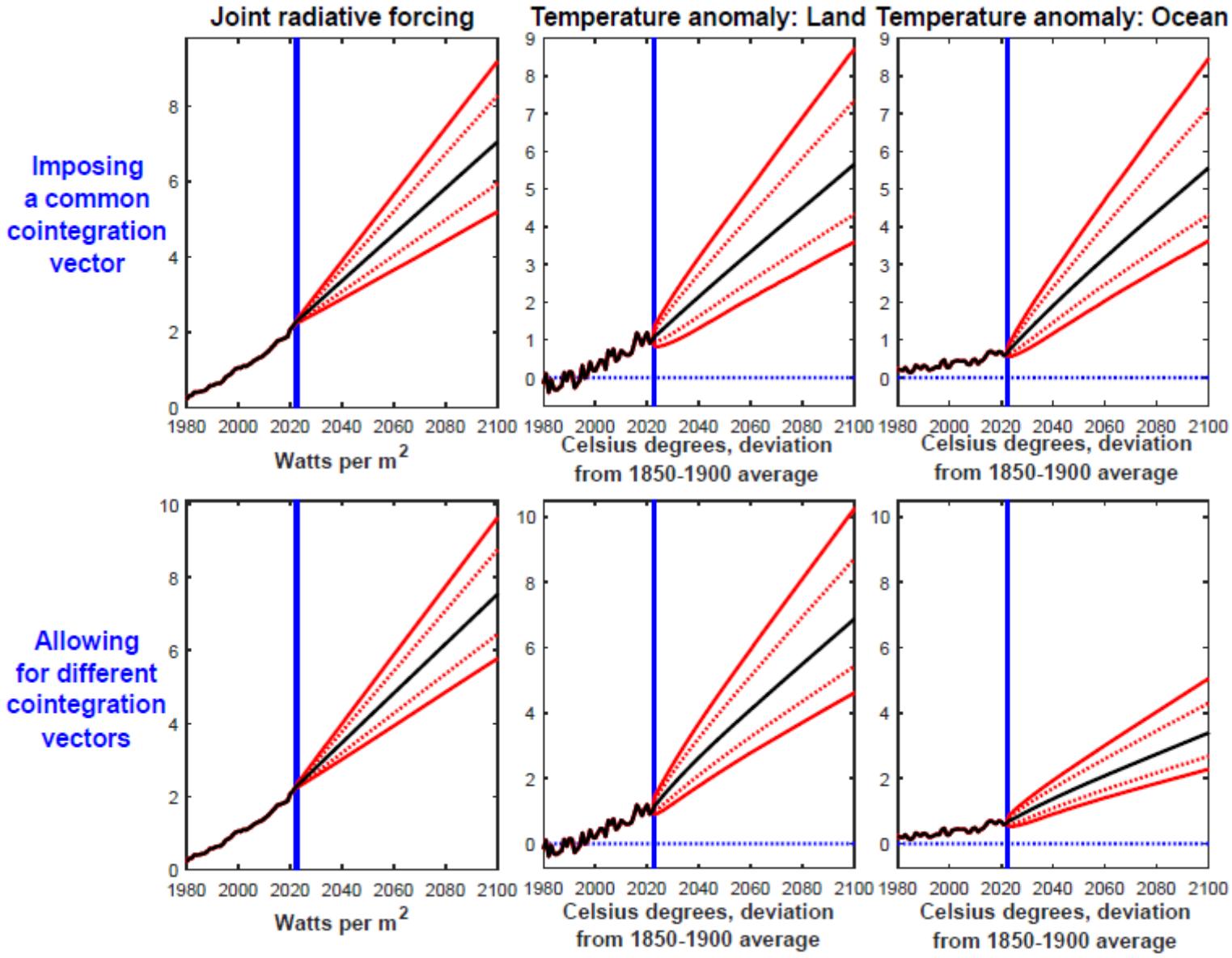

Figure 5 Forecasts under a 'no change' scenario conditional on data up to 2022 (median, and 16-84 and 5-95 credible set)

forecasts for the two series are, as expected, very close to each other, with median estimates for the year 2100 equal to 5.6 and 5.5, respectively, and 90%-coverage credible sets equal to [3.6, 8.7] and [3.6, 8.5], respectively. In order to put these numbers into context it is worth recalling that the average increase in temperatures associated with the Paleocene-Eocene Thermal Maximum (PETM), about 55.5 million years ago, is estimated to have been between 5 and 8 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 5.5-5.6 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.

The rationale for imposing a common cointegration vector between the two temperature anomalies and the JRF index is that basic physics suggests that, in the very long run, a given increase in the JRF should produce an identical increase in the land and the ocean temperature anomalies. For the purpose of producing comparatively short-horizon projections such as those in Figure 5, however, there is the risk that this may distort the forecasts. Intuitively, if the one-for-one long-run equilibrium between the land and the ocean anomalies reasserts itself over periods of centuries, the imposition of a common cointegration vector upon a model estimated based on just 172 years might distort the projections. At the end of the day, as the raw data for the two anomalies plotted in the third panel of Figure 1 clearly show, since 1850 the ocean anomaly has increased much less than the land anomaly. Even if we can be certain that (say) 10,000 years after a shock to the JRF index the two anomalies will have increased by exactly the same amount, the fact that they have materially deviated from one another for 172 years raises questions about the reliability of projections obtained by imposing a common cointegration vector.

Because of this it makes sense to also consider forecasts produced by the model allowing for different cointegration vectors for the two anomalies. These projections are shown in the bottom row of Figure 5. As it should be logically expected, they extrapolate into the future the marked divergence between the behaviour of the two series that has been going on since 1850. Whereas the median projection for the ocean anomaly in 2100 is equal to 'just' 3.4 Celsius degrees, the corresponding forecast for the land anomaly, at 6.9 degrees, is roughly around the middle of the range of the estimates of the increase in temperatures that had taken place during the PETM, about 55.5 million years ago.

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

Figures 6 and 7 show, based on the models imposing a common cointegration vector for the land and ocean anomalies and, respectively, allowing instead for different

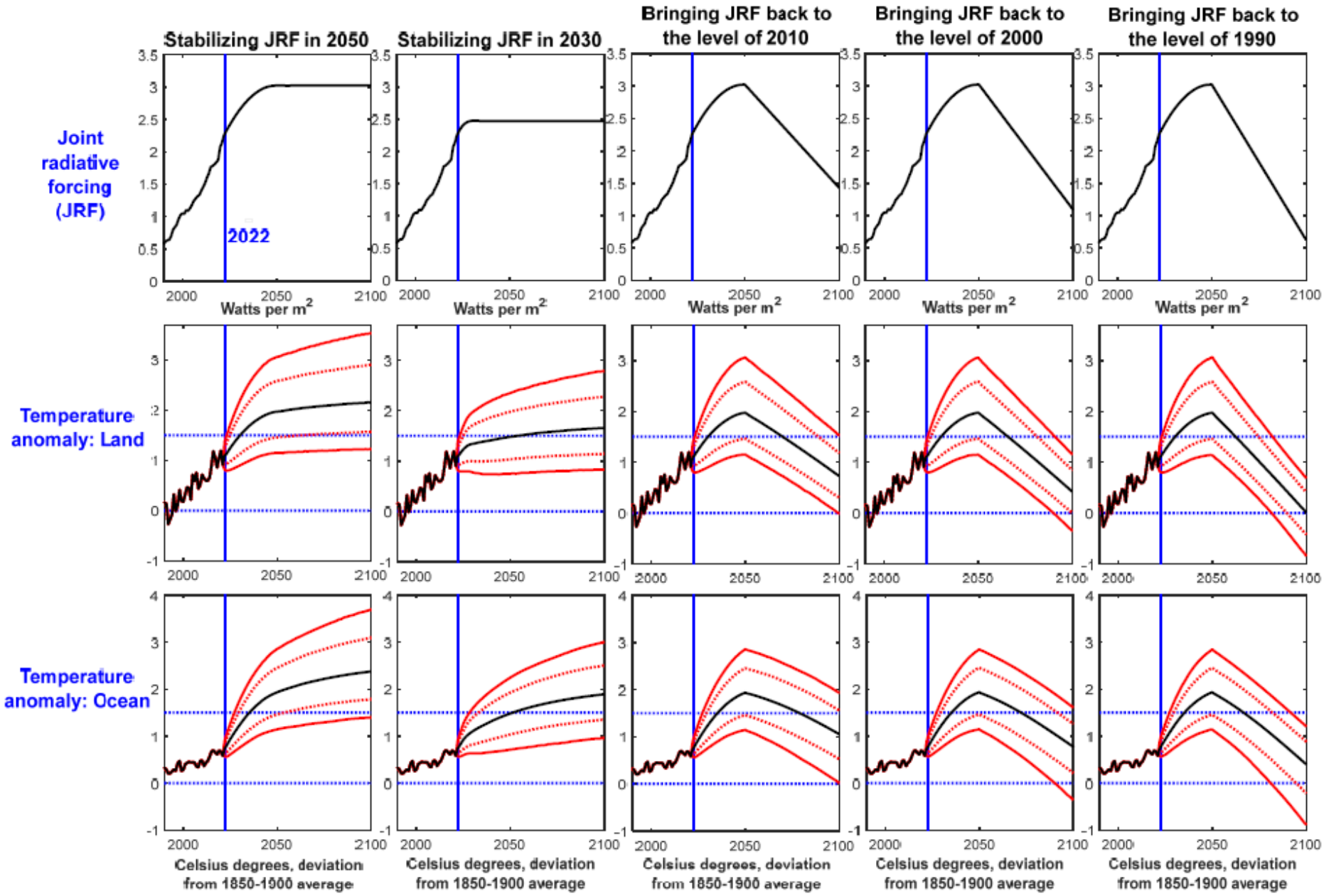

Figure 6 Cointegrated VECM with random-walk drift and a common cointegration vector for the land and ocean temperature anomalies: Forecasts conditional on data up to 2022 and alternative scenarios for the evolution of joint radiative force (median, and 16-84 and 5-95 credible set)

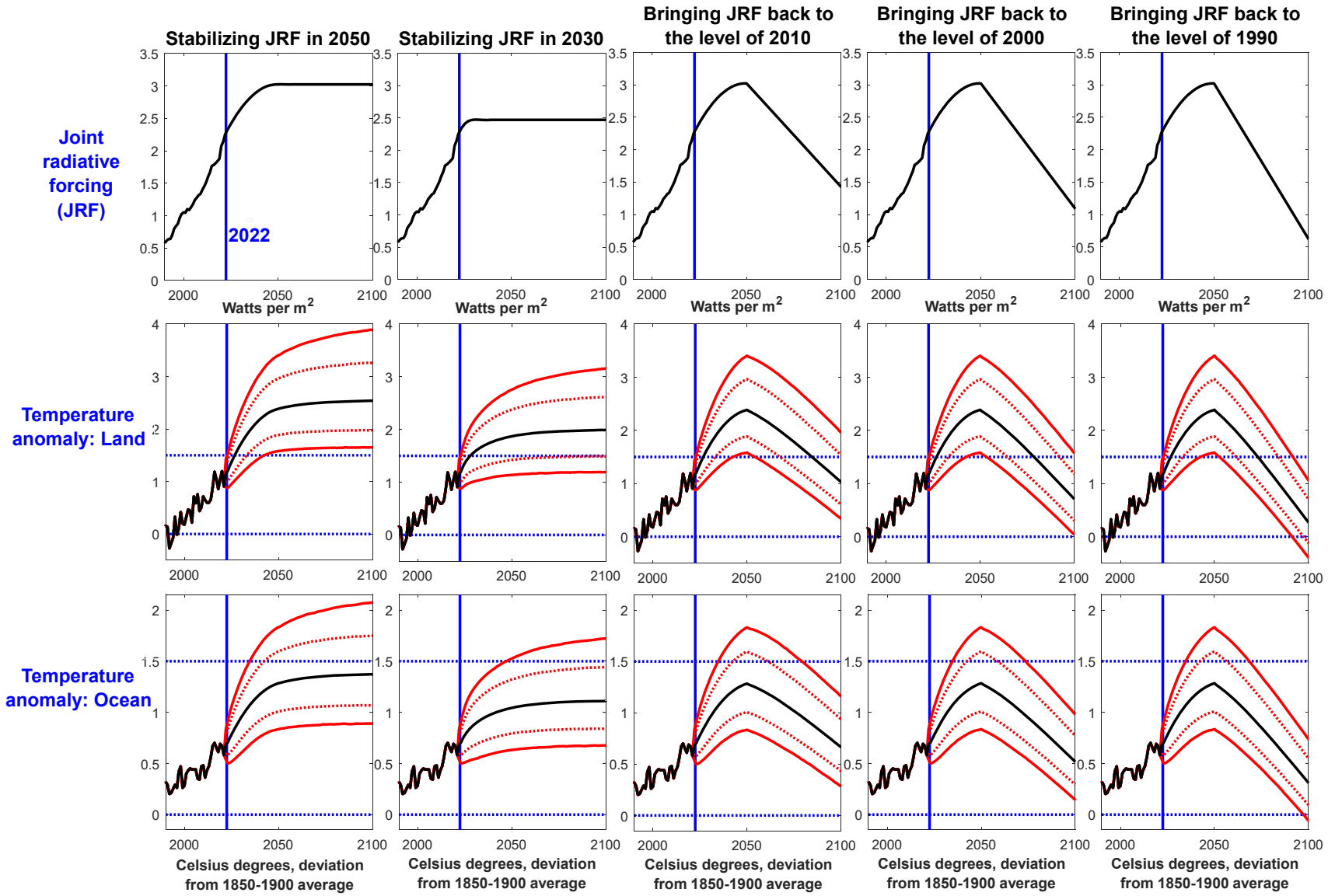

Figure 7 Cointegrated VECM with random-walk drift and different cointegration vectors for the land and ocean temperature anomalies: Forecasts conditional on data up to 2022 and alternative scenarios for the evolution of joint radiative force (median, and 16-84 and 5-95 credible set)

cointegration vectors, evidence from the following exercise. I 'freeze' the state of the system to 2022 as in the previous sub-section, and I then stochastically simulate the model forward in time until the end of the XXI century conditional on (1) data up to 2022, and (2) alternative possible paths for the evolution of the JRF until the end of the century. I compute the conditional forecasts as in Waggoner and Zha (1999). I consider five possible scenarios for the evolution of the JRF:

- the JRF being stabilized in 2050, with its first difference progressively decreasing starting in 2023, and reaching zero in 2050;
- a qualitatively similar scenario in which the stabilization takes place in 2030; 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.

Starting from the projections produced by the model imposing a common cointegration vector, stabilization of the JRF in either 2050 or even 2030 leaves open the possibility that warming will reach levels that in the climate science community are widely regarded as dangerous. For example, the 90%-coverage credible sets for the land and ocean temperature anomalies for the year 2100 associated with stabilizing the JRF in 2030 are equal to [0.8, 2.8] and [1.0, 3.0] respectively. Further, for both anomalies most of the 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. Evidence from the corresponding exercise in which the JRF is stabilized in 2050 are, as expected, even more ominous, with the 90%-coverage credible sets for the two anomalies for the year 2100 being equal to [1.2, 3.6] and [1.4, 3.7] respectively. Evidence from the corresponding exercise based on the model allowing for different cointegration vectors is, as expected, better for the ocean anomaly, but materially worse for the land anomaly. In particular, the 90%-coverage credible set for the land anomaly for 2100, equal to [1.6, 3.9], now lies *entirely* beyond 1.5 degrees.

Turning to the set of projections obtained by allowing the JRF index to peak in 2050, and then bringing it back to the levels it had reached in 1990, 2000, and 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, based on the model imposing a common cointegration vector, the 90%-coverage credible sets for the land and ocean temperature anomalies for the year 2100 are equal to [-0.3, 1.2] and [-0.4, 1.6] respectively if the JRF were to be brought back to the level of 2000, whereas the corresponding sets obtained by bringing it back to the level of 2010 are equal to [0, 1.5] and [0, 2] respectively. Based on the model allowing for different cointegration vectors the corresponding credible sets for the land anomaly obtained by bringing

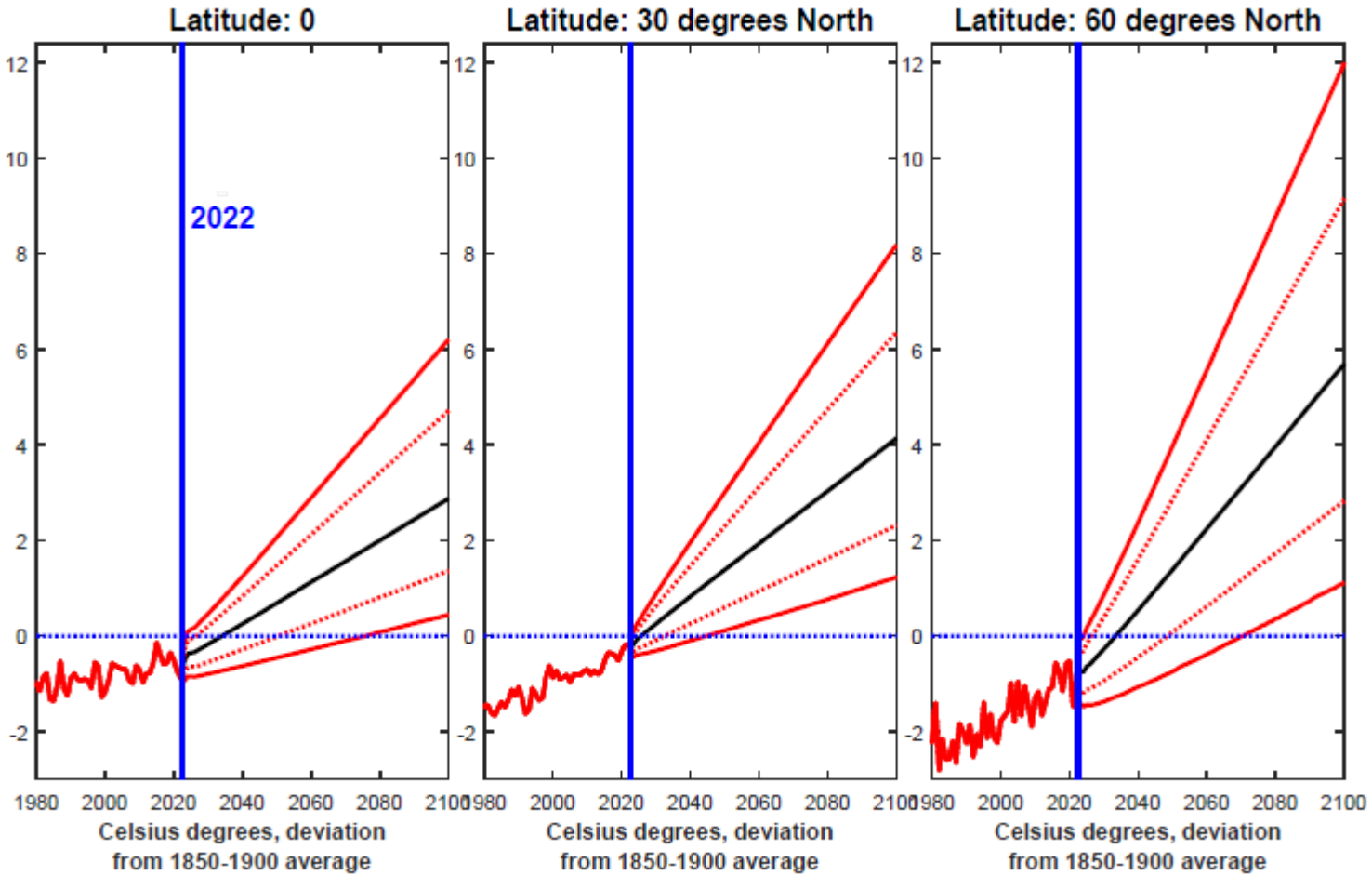

Figure 8 Cointegrated VECM with random-walk drift for the temperature anomalies at different latitudes: Forecasts under a 'no change' scenario conditional on data up to 2022 (median, and 16-84 and 5-95 credible set)

the JRF back to the levels of 2000 and 2010 are [0.1, 1.6] and [0.4, 2.0] respectively, whereas those for the ocean anomaly are [0.1, 1.0] and [0.3, 1.2]. 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 be the case.

6 Forecasts for different latitudes

Figure 8 shows results from the same 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 (latitude 0), and either 30 or 60 degrees North. The main finding in the figure pertains to the dramatic extent of variation in the predicted increases in temperatures at different latitudes under the 'no change' scenario. Whereas the median forecast for the Equator for the year 2100 is equal to 2.9 Celsius degrees, with a 90%-coverage credible set equal to [0.4, 6.2], the corresponding objects for the 30 and 60 degrees North latitudes are 4.2 and [1.2, 8.2], and 5.7 and [1.1, 12] respectively. This obviously reflects the fact that, as it is well known, the impact of climate change in terms of increases in temperatures is greater the higher the latitude, and it is therefore maximum at the Poles and minimum at the Equator.

7 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 temperature anomaly fully converges to its new long-run value in about 50 years, whereas the response of the ocean anomaly is even more drawn out. Under a 'no change' scenario, the most favorable median forecast predicts the land temperature anomaly to reach 5.6 Celsius degrees in 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 reached in the early years of the XXI century.

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.

8 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.

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.

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.

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.[11] 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

[11] 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.[12]

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.

[12] 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

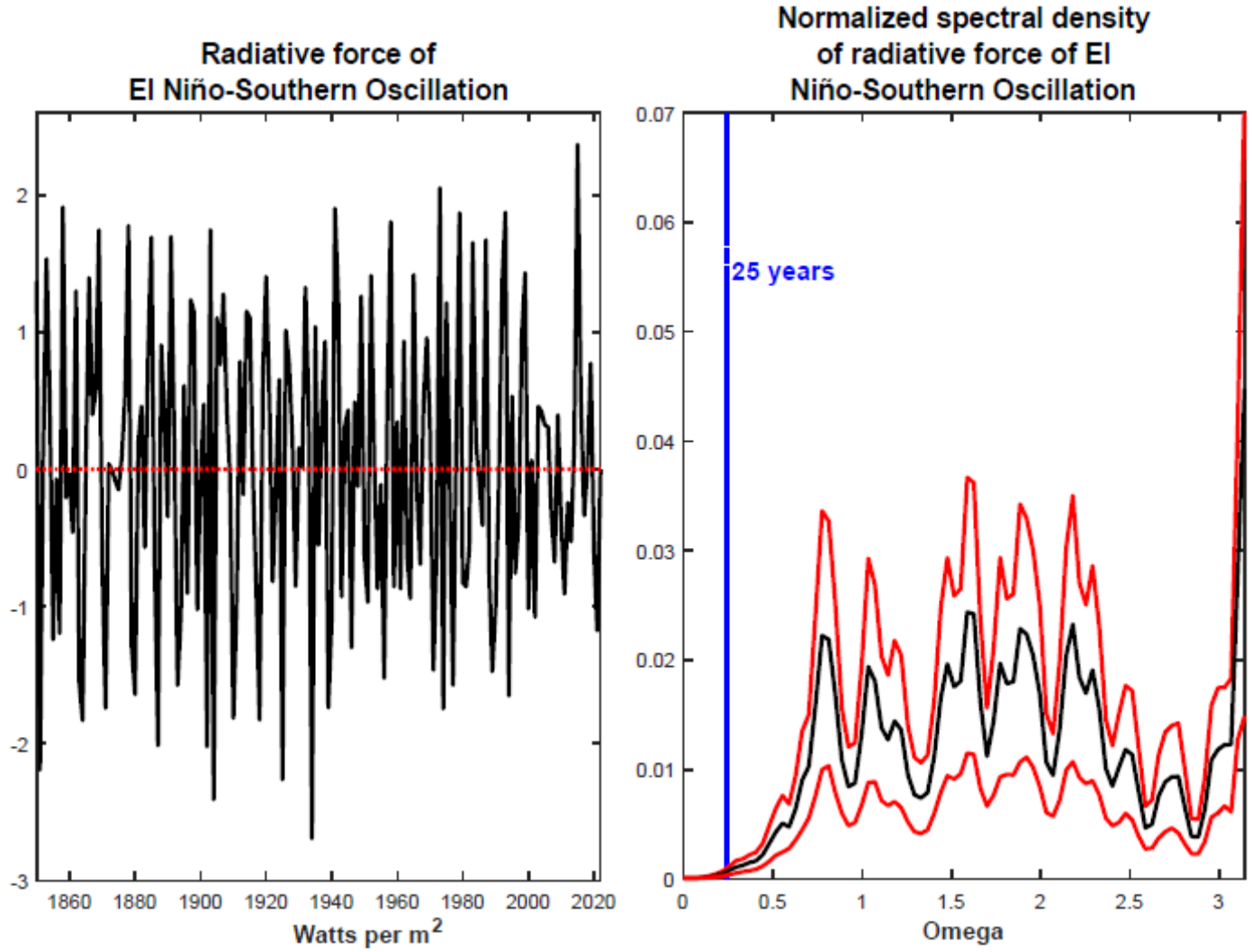

Figure A.1 Radiative force of El Niño-Southern Oscillation: raw series and normalized spectral density (with 90%-coverage bootstrapped confidence bands)

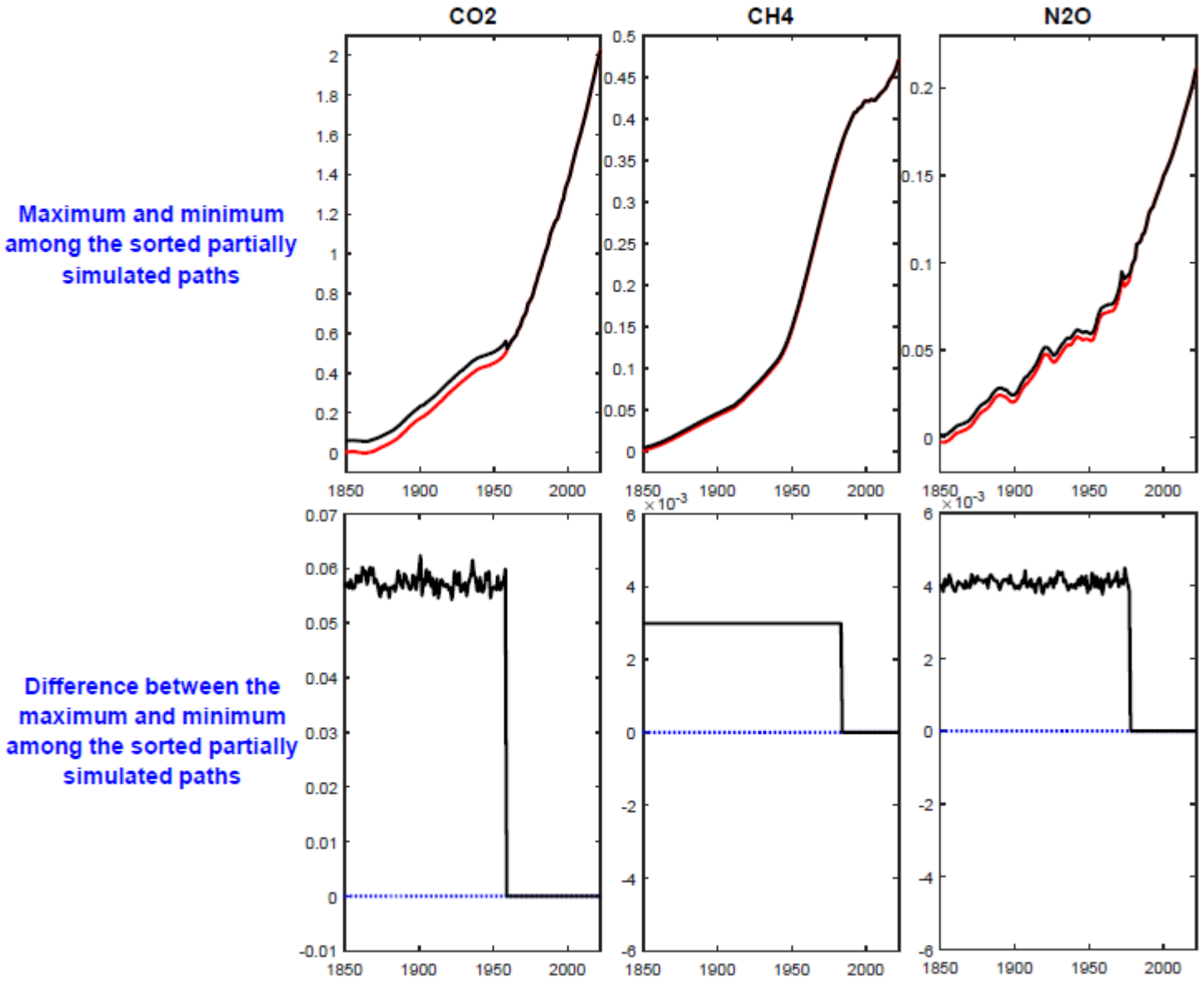

Figure A.2 Evidence on the close similarity between alternative partially simulated series for CO2, NH4, and N2O: maximum and minimum among the sorted partially simulated paths out of 100,000 simulations