



Evidence that global evapotranspiration makes a substantial contribution to the global atmospheric temperature slowdown

L. Mark W. Leggett¹ · David A. Ball¹

Received: 30 August 2016 / Accepted: 18 January 2018 / Published online: 14 February 2018
© Springer-Verlag GmbH Austria, part of Springer Nature 2018

Abstract

The difference between the time series trend for temperature expected from the increasing level of atmospheric CO₂ and that for the (more slowly rising) observed temperature has been termed the global surface temperature slowdown. In this paper, we characterise the single time series made from the subtraction of these two time series as the ‘global surface temperature gap’. We also develop an analogous atmospheric CO₂ gap series from the difference between the level of CO₂ and first-difference CO₂ (that is, the change in CO₂ from one period to the next). This paper provides three further pieces of evidence concerning the global surface temperature slowdown. First, we find that the present size of both the global surface temperature gap and the CO₂ gap is unprecedented over a period starting at least as far back as the 1860s. Second, ARDL and Granger causality analyses involving the global surface temperature gap against the major candidate physical drivers of the ocean heat sink and biosphere evapotranspiration are conducted. In each case where ocean heat data was available, it was significant in the models: however, evapotranspiration, or its argued surrogate precipitation, also remained significant in the models alongside ocean heat. In terms of relative scale, the standardised regression coefficient for evapotranspiration was repeatedly of the same order of magnitude as—typically as much as half that for—ocean heat. The foregoing is evidence that, alongside the ocean heat sink, evapotranspiration is also likely to be making a substantial contribution to the global atmospheric temperature outcome. Third, there is evidence that both the ocean heat sink and the evapotranspiration process might be able to continue into the future to keep the temperature lower than the level-of-CO₂ models would suggest. It is shown that this means there can be benefit in using the first-difference CO₂ to temperature relationship shown in Leggett and Ball (*Atmos Chem Phys* 15(20):11571–11592, 2015) to forecast future global surface temperature.

1 Introduction

A central issue in climate science has been that over the past two decades the observed global surface temperature trend has been lower than that expected from the majority of climate simulations.

This difference has been given various names including the pause (for example, Trenberth and Fasullo (2013), the hiatus (for example, Meehl et al. 2011, the Intergovernmental Panel on Climate Change Fifth Assessment Report (hereafter AR5) (IPCC 2013), the inconsistency between observed and simulated global warming (Fyfe et al. 2013) and the climate model/

temperature mismatch (Leggett and Ball 2015 (hereafter L&B 2015)).

Along with Yan et al. (2016), who consider the term hiatus a misnomer, we do not use the terms ‘pause’ or ‘hiatus’ because of their implication that the future is known. In this paper, we will formally use the term ‘global surface temperature simulation/observation inconsistency’. For brevity in the paper, we will use the term ‘the temperature gap’.

1.1 Expressing the temperature gap

L&B (2015) provided evidence that temperature does not follow the level of atmospheric carbon dioxide (CO₂) but first-difference CO₂ (that is, the change in CO₂ over time). This can be restated as “temperature expected from simulation minus observed temperature gives a gap (a “temperature gap”); and the time series of the level of CO₂ minus the time series for first-difference CO₂ gives a second series analogous to the temperature gap series (a “CO₂ gap”)”.

✉ L. Mark W. Leggett
mleggett.globalriskprogress@gmail.com

¹ Global Risk Policy Group Pty Ltd, Townsville, QLD, Australia

A key point is that to compare either of these gaps with other series as an outcome—dependent—variable it will greatly enable interpretation if we have a single series for the gap. Single series for each of the temperature and CO₂ gaps are therefore used in this study (For full details of derivation, see Section 4.2.1).

Each of these two gap series is a compound series derived from two underlying series. Examples of the use of climate time series derived in this way include a series of the “mean difference between ensemble members and a control simulation” (Notaro and Liu 2008); “the ensemble mean differences, INT minus CLI, to represent the signal of the influences of vegetation variability...” (Zhu and Zeng 2016); and evaluating the effect of rising atmospheric CO₂ based on “the differences between simulations” (Zeng et al. 2017).

1.2 The scope of this paper

L&B (2015) analysed this gap situation from the start of CO₂ recordings at Mauna Loa in 1959.

Since L&B (2015), other studies beyond those reviewed in that paper have provided explanations for the temperature gap. These studies used variables other than first-difference CO₂. Such variables have involved, for example, the Atlantic Multi-decadal Oscillation (AMO) (e.g. Pasini et al. (2016)). That noted, we have been able to find no further studies investigating the correlation between first-difference CO₂ and temperature.

This paper therefore deals with three questions relating to the correlation between first-difference CO₂ and temperature. First, what do the temperature and CO₂ gaps look like over a longer previous term; and over this longer period do temperature gaps appear frequently or is the current gap unusual? Second, are there potentially causal correlates between other, physical, climate factors and the temperature gap? These questions are explored both for the long term with annual data and, seeking corroboration, for series available only more recently, with monthly data. Third, if causal physical correlates between other climate factors and the temperature gap are found, do these have a capacity to further enlarge the temperature gap—that is, in the decades ahead to continue to keep the observed temperature trend below the trajectory of the future simulation trend based on level of CO₂?

The scope of this paper is therefore to further characterise and explore the specific topic of the relationship of first-difference CO₂ with temperature rather than to describe and compare the full range of published proposed causes of the temperature gap.

1.3 Candidate physical causal factors for the temperature gap in detail

As the gap is a temperature and therefore heat gap, the physical climate factors which might cause the gap are those which affect the global heat budget. Concerning the climate system components considered as actually or potentially affecting the global heat budget AR5 (p. 1451) states: “the climate system consists of

five major components: the atmosphere, the hydrosphere, the cryosphere, the lithosphere and the biosphere.” Of these, AR5 (p. 264) states that where energy budgets are concerned, the hydrosphere dominates: “Ocean warming dominates the total energy change inventory, accounting for roughly 93% on average from 1971 to 2010 (high confidence). Melting ice (including Arctic sea ice, ice sheets and glaciers) accounts for 3% of the total, and warming of the continents 3%. Warming of the atmosphere makes up the remaining 1%.”

It is noteworthy that the biosphere is not mentioned in this heat budget.

This view is still prominent today. In a review, Yan et al. (2016) argue that the slowing of the rate of global mean surface warming from 1998 to 2013 represents a redistribution of energy within the Earth system. They then go on to write “Improved understanding of *ocean* distribution and redistribution of heat will help us better monitor Earth’s energy budget.” [Present authors’ emphasis]

Similarly few instances are found in AR5 on the question of the role of vegetation as substantially affecting the heat budget. For example, AR5 states on page 707: “Overall, vegetation changes may have caused modest cooling at high latitudes and warming at low latitudes, but the uncertainties are large and confidence is very low.”

More widely, however, literature exists to suggest that global vegetation can have large effects on the global heat budget, particularly through evapotranspiration. This literature resulted initially from global climate modelling (GCM) and includes the following.

Kleidon et al. (2000) quantified the maximum possible influence of vegetation on the global climate by conducting two extreme climate model simulations: in a first simulation a ‘desert world’, and at the other extreme, a second simulation of a ‘green planet’. Land surface evapotranspiration more than triples in the presence of the ‘green planet’, and (resulting from the increase in latent heat flux) mean near-surface temperatures are lower by as much as 8 K.

Bounoua et al. (2000) reported that important effects of increased vegetation on climate are a cooling of about 1.8 K in the northern latitudes during the growing season and a slight warming during the winter, which is primarily due to the masking of high albedo of snow by a denser canopy, and a year-round cooling of 0.8 K in the Tropics.

Bounoua et al. (2010), conducting climate simulations with twice the current level of CO₂, found that increased evapotranspiration reduced land surface warming by 0.6 °C.

From Table 1 of Zhu and Zeng (2015), an almost doubling of evapotranspiration was modelled, and this was associated with a 2.74 K decrease in temperature.

Using an empirical approach, Zeng et al. (2017) have shown that of the overall sum of biophysical feedbacks related to the greening of the Earth, the net effect—some 79%—was

cooling. Of this 79%, evapotranspiration contributed the largest amount, some 41 percentage points.

Zeng et al. (2017) show that there is a net cooling from global vegetation. This net cooling effect is the sum of cooling from increased evapotranspiration (70%), changed atmospheric circulation (44%), decreased shortwave transmissivity (21%), and warming from increased longwave air emissivity (−29%) and decreased albedo (−6%).

Shen et al. (2015), using a statistical model, showed that “evaporative cooling over the Tibetan Plateau induced by vegetation growth may attenuate daytime warming by enhancing evapotranspiration (ET), a cooling process.”

How much of the evapotranspiration is from transpiration—that is, from plants? The percentage of total terrestrial evaporation which is evapotranspiration found from isotope observational methods is as follows: from run-off-based stable isotope techniques, 80 to 90% (Jasechko et al. 2013); and hybrid stable isotope techniques, $64 \pm 13\%$ (Good et al. 2015). It can be seen that from both estimates, the transpiration proportion is large.

In looking into the mechanism, Ban-Weiss et al. (2011) observe that evapotranspiration removes sensible heat by turning it into latent heat. Increased latent heat flux to the atmosphere has a local cooling influence known as ‘evaporative cooling’. However, as Ban-Weiss et al. point out, this energy will be released back to the atmosphere wherever the water condenses. That said, the above studies provide evidence that a longer lasting global cooling effect does occur by some mechanism. Using idealised climate model simulations, Ban-Weiss et al. suggest that a decrease in global mean surface air temperature of about 0.54 K can occur via this effect, largely as a consequence of planetary albedo increases associated with an increase in low elevation cloudiness caused by the increased evaporation.

This result leads us to the situation where a mechanism involving large but transitory heat sink behaviour may lead to a large long-term effect on the heat budget not as a heat sink but via reflection of fresh incoming energy away from the Earth’s surface.

In any event, the foregoing leads us, in what follows, to seek terrestrial vegetation evapotranspiration relationships, in addition to ocean heat relationships, with the temperature gap.

We now seek to determine the full list of candidate indicators for evapotranspiration and the ocean heat sink. Published evapotranspiration datasets from satellite sensing are available. However, they are short in duration, extending only from the 1980s (McCabe et al. 2016; Zhang et al. 2015). An alternate approach which provides markedly longer time series is as follows. Walter et al. (2004) state: “Evaporation or evapotranspiration (ET) trends can be identified by using published data and assuming that the balance among precipitation, streamflow, and ET dominates the terrestrial hydrological budget on annual and longer time scales.”

Hence, one indicator of global evapotranspiration is global precipitation minus global runoff and this data series is used in this paper. The normalised difference vegetation index (NDVI) can act as an indicator and is also used in this paper.

For the ocean heat sink, ocean heat content series are available from the National Oceanic and Atmospheric Administration (NOAA) and the Japan Meteorological Organisation (JMO) and are used.

A further major influence on climate is volcanic aerosols. Yan et al. (2016) state that intermittent events, such as the eruptions of major volcanoes like El Chichón (which began erupting in 1982) and Pinatubo (which began erupting in 1991), can temporarily counteract long-term warming. The effect of volcanic aerosols on the temperature gap is therefore also explored in this paper.

We note that all of the above research has looked at the effect on climate of only one or other of evapotranspiration and the ocean heat sink. In this paper, we look for the effects on the temperature gap of both factors together.

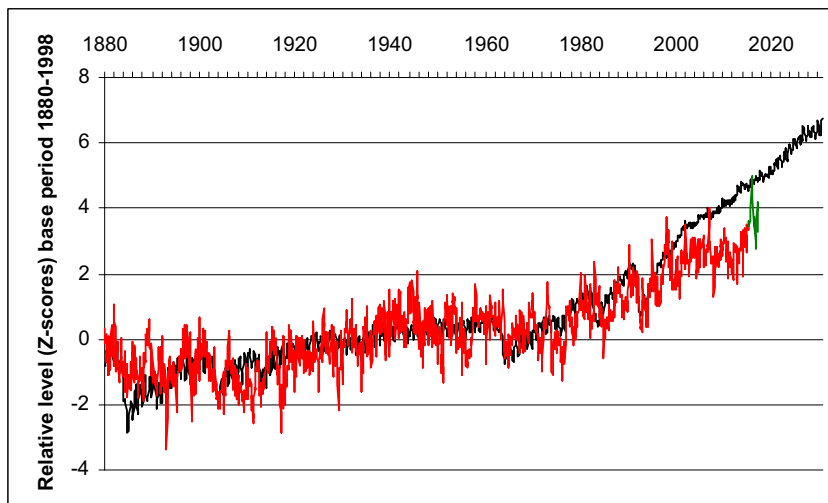
1.4 Structure of this paper

The structure of the rest of the paper is as follows. Section 2 provides a scan of the quantitative research methods available and chooses time series regression as the most suitable. In some instances, this is in the form of ordinary least squares (OLS) analysis but usually the autoregressive distributed lag (ARDL) and vector autoregression (VAR) methods from econometrics are used.

Section 3 explores the methods and data used to undertake our analysis. Using these methods, in Section 4 (Results), we first consider how far back in time the temperature gap and its associated CO₂ gap extend. Having determined this, we explore the potential drivers of the outcome variable, the temperature gap. We first explore the CO₂ gap as a driver of the temperature gap. This pair of series goes back to the 1860s, providing a considerable period of data for analysis.

The CO₂ gap being a non-physical construct, we then turn to assessing physical candidate drivers of the temperature gap. These drivers are evapotranspiration, the oceanic heat sink and volcanic aerosol emissions. We start by seeking long driver series commensurate with the CO₂ gap. Volcanic aerosols have a lengthy data series. But evapotranspiration and ocean heat series are shorter, commencing only in 1949 and 1950, respectively. However, we show that evapotranspiration correlates well with precipitation for which we have a series from 1901. So first we explore correlations of precipitation and volcanic aerosols with the temperature gap over the longer data series. We next explore correlation from 1950 for precipitation, ocean heat and volcanic aerosols. We then explore correlations for evapotranspiration ocean heat and volcanic aerosols.

Fig. 1 Monthly data, Z-scored to aid visual comparison (see Section 1). Shown are the output of an IPCC business-as-usual scenario model (CMIP5, RCP8.5 scenario) run for the IPCC fifth assessment report (IPCC 2014) (black curve). Global surface temperature (HadCRUT4.5) (red curve: over the period covered in L&B (2015) green curve: period since L&B (2015)



In Section 5, we explore how long the described conditions might last into the future. Section 6 provides a discussion and conclusion.

2 Methodological issues and objectives of the study

In this section, we seek to determine the best data analysis process with which to study the interactions of the preceding climate time series as potential causes and effects of each other.

2.1 Choice of overarching data analysis process for the study

We start by assessing the full continuum of data analysis models from which to choose.

Karplus (1977, 1992) has provided a framework for the characterisation and classification of models of systems. Enting (1987, 2010) has used this framework to assess model types used in climate studies.

In connection with his framework, Karplus (1992) observes that valid models of systems are the key to the successful prediction of the response (outputs) of systems to specified excitations (inputs). Karplus goes on to observe that there are numerous techniques, but all can be regarded as employing combinations of deduction and induction in varying proportions.

Karplus (1977) termed this deduction-induction range as a modelling spectrum. The position that each model took in the spectrum represented the degree of deduction as opposed to induction that was involved in the modelling process. The spectrum was described at the induction end as involving black-box models (characterised as being highly empirical, and only representing relations between inputs and outputs).

Fig. 2 Unsmoothed (raw) annual ice core level-of-CO₂ data in parts per million. Blanks in the series show the extent of missing data

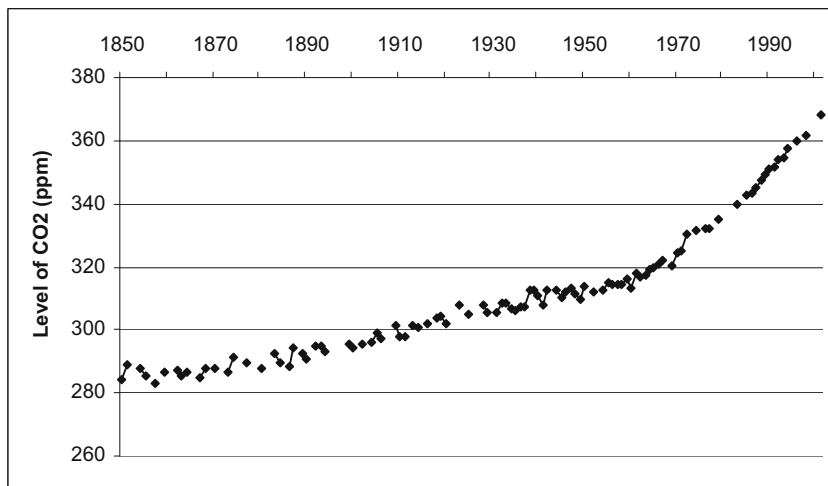
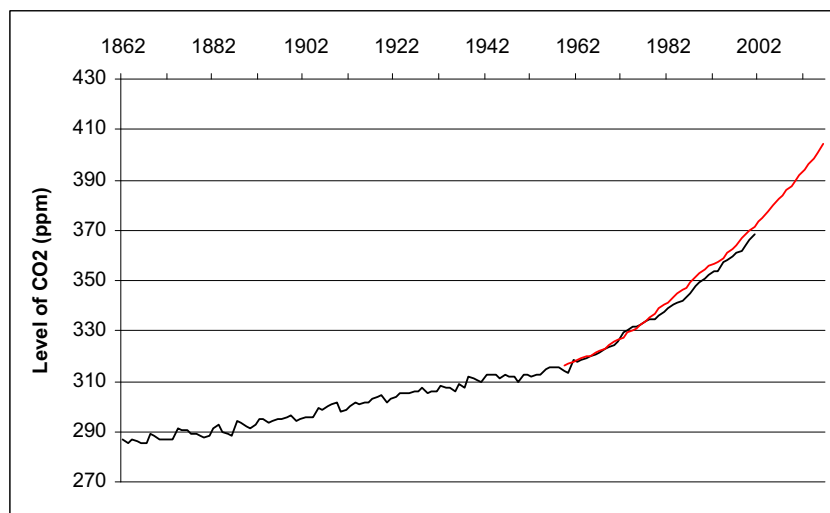


Fig. 3 Level of CO₂ data used. Black curve: unsmoothed (raw) annual ice core level-of-CO₂ data 1850–2000 in parts per million. The curve shows the result of linear interpolation from adjacent data points over the blanks shown in Fig. 2. Red curve: atmospheric CO₂ (ppm) 1959–2016 from Mauna Loa Observatory



The ‘curve-fitting’ model of a ‘black-box’ system is determined inductively from observations of the behaviour of the system. Curve-fitting, black-box models are typically used in fields such as economics.

At the deduction end of the spectrum are white-box models (characterised as having relations between inputs and outputs defined through processes involving internal states of the system expressed in mechanistic terms). White-box models (Enting (2010) considers that ‘glass-box models’ would have been a better term) are generally deterministic. White-box models are frequently used in the physical sciences. The behaviour of a ‘white-box’ system can be deduced directly from knowledge of the system structure and the basic physical laws that apply to it.

Climate studies have been assessed from the perspective of the Karplus spectrum by Enting (2010), who finds that both black-box and white-box models are used. Enting describes these as follows: black box—statistical fits which include regression-type analyses of CO₂ trends and cycles, correlation

studies relating CO₂ and ENSO and empirical fits of transfer relations connecting concentrations to emissions. These apply at the globally aggregated level; white box—Earth system models, built around global climate models (GCMs) for atmosphere and ocean.

In line with Karplus (1977, 1992), Enting (1987) argues that the spectrum concept gives a useful framework for comparing different types of modelling and forces an explicit recognition that possible uses of a model will depend on the type of model, i.e. its position within the spectrum.

Within individual climate papers in the literature, a range of terminology is used for model types. Rahman and Lateh (2015) divide the range of methods into two broad groups: simulation techniques and statistical models. In their survey of methods, Adams et al. (2013) state that the two endpoints on a theoretical continuum of mechanisms are process-based and empirical model types. This terminology is also used by Li (2017). Moore et al. (2013) defined two model categories: physically plausible models of reduced complexity that

Fig. 4 Black curve: first-difference annual ice core level-of-CO₂ data 1850–2000 in parts per million. Red curve: first-difference atmospheric CO₂ (ppm) 1959–2016 from Mauna Loa Observatory

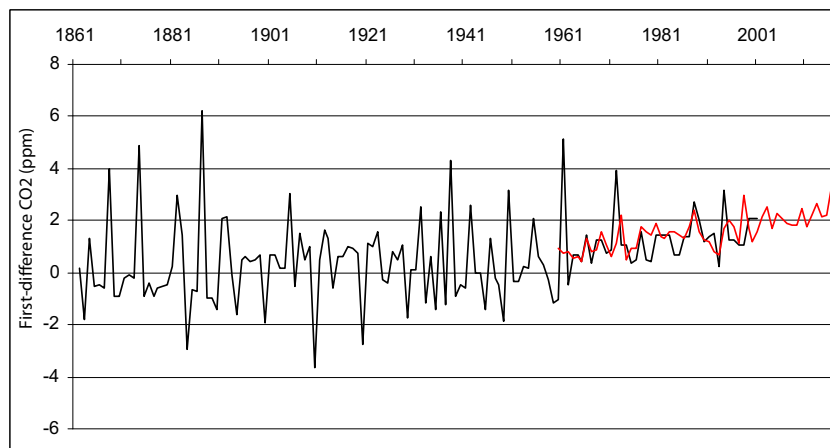
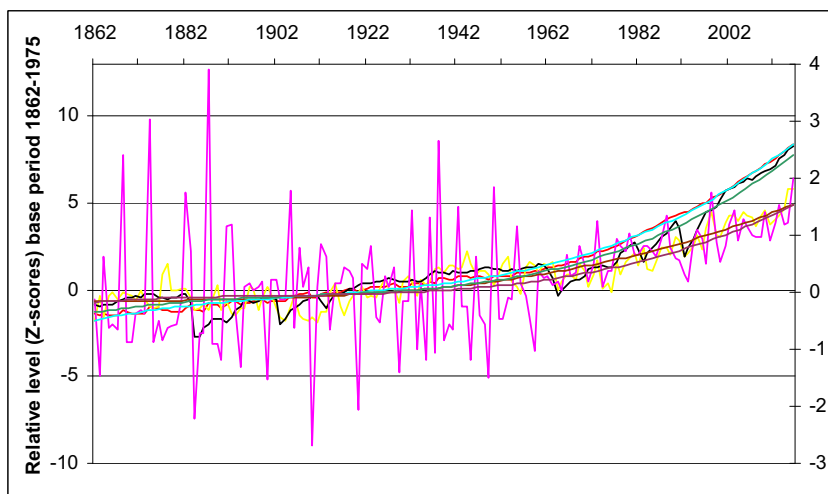


Fig. 5 Annual data, Z-score base 1862–1975. Left axis: modelled global surface temperature RCP8.5 model (black curve); level of CO₂ (red curve); observed global surface temperature (HadCRUT 4.5) (yellow curve). Right axis: first-difference ice core CO₂ (purple curve). To show the assortment of the core trends of four curves despite their disparities into two groupings, the four series are fitted with third-order polynomials



exploit statistical relationships between climate and climate forcing, and more complex physics-based models of the separate elements of the climate budget.

Pasini et al. (2016) use the term ‘GCM simulation’ for white box and ‘data-driven method’ for black box.

With regard to correct model specification, Grassi et al. (2013) note that process-based climate models (and too-simple empirical models) are often lacking due to the fact that climate series display complex statistical properties, and that modelling of these must be correctly specified to provide valid statistical inference.

Further on the strengths of black-box models, Pasini et al. (2016), referring to Pasini et al. (2012) and Triacca et al. (2014), note that black-box models can point to aspects not always clearly addressed by white-box modelling.

Pasini and Mazzocchi (2015) and Mazzocchi and Pasini (2017) argue that many attribution results coming from non-GCM studies align with those from GCM studies, implying the robustness of each. These include

identifying anthropogenic forcings as the main drivers of temperature change.

Such non-GCM methods include neural network investigations which find substantiations of the importance of greenhouse gases in driving the recent global warming (see, for instance, Pasini et al. 2006, 2017; Verdes 2007; Schonwiese et al. 2010). Granger causality analyses have also established the major role of anthropogenic forcings, even in comparison with natural forcings and drivers of natural variability (see Attanasio et al. (2012) and Pasini et al. (2012); Stern and Kaufmann (2014) for specific results; and Attanasio et al. (2013) for a brief review of attribution studies via Granger causality analyses).

That said, a key issue with GCMs is that, as mentioned at the outset of this paper, the observed global surface temperature trend has been lower than that expected from the majority of climate simulations.

The key approach to check for model adequacy is termed ‘validation’ (Montgomery et al. 2008). A common method of

Fig. 6 Main series break in trend in ratio of Z-first-difference CO₂ to Z-level of CO₂ (black curve). The break is indicated by both a third-order polynomial (red curve) and approximately by two straight lines (black lines)

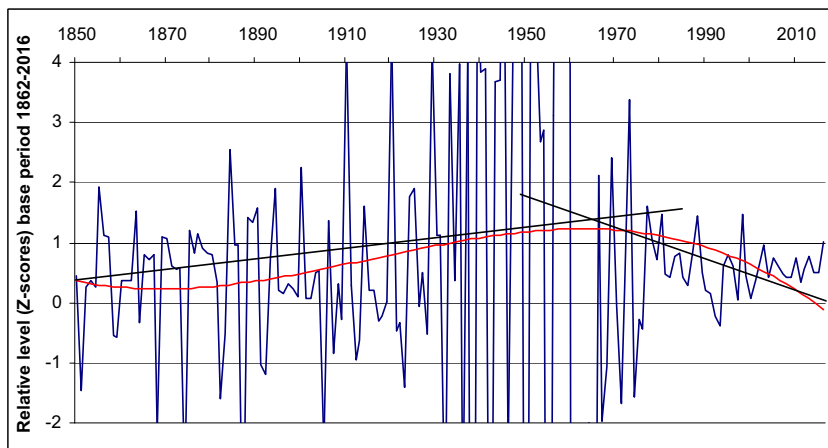
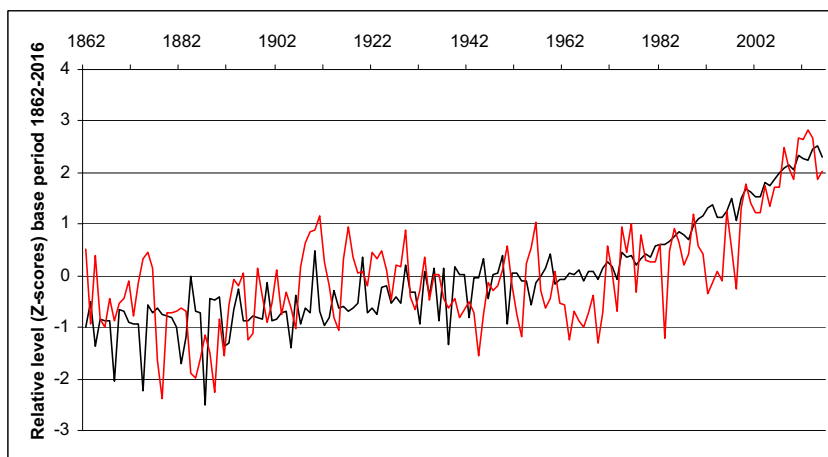


Fig. 7 Z-scored CO₂ gap (black curve) and temperature gap (red curve)



validation is to check the ability of the simulation to correctly predict outputs caused by inputs other than those used in constructing the model. This is done by ‘saving’ some system observations in order to use them later for validation. In statistics, this is termed the ‘split-sample’, ‘cross-validation’ or ‘train-test’ approach (Montgomery et al. 2008).

Using the split-sample approach, Newman (2013) finds that an empirical time series model shows global surface temperature prediction skill better than that of the major process-based climate models (phase 5 of the Coupled Model Intercomparison Project (CMIP5)). To Newman, these results suggested that current coupled model decadal forecasts may not yet have much prediction skill beyond that captured by multivariate, predictable linear dynamics.

In the AR5 (IPCC 2013) chapter titled Evaluation of Climate Models, Flato et al. (2013) state (p. 826) “...many (GCM) studies have failed to find strong relationships between observables and projections”; and (p. 772): “Almost all CMIP5 historical simulations do not reproduce the observed recent warming hiatus.”

One of the reasons for these problems when process models are used to model climate may be that, as Enting (2010) points out, process models are vulnerable to neglect of processes—the ‘Kelvin error’. Enting writes: “The term ‘Kelvin error’ refers to the risk of missing a process from the modelling, taking its name from Lord Kelvin’s underestimates of the ages of the earth and sun due to neglect of nuclear processes.”

By contrast, a black-box model, while its component processes are not specified at all, by definition contains *all the component processes* of the reality under study.

As we consider this point is the deepest to come out of this review, we will use the terms *white box* and *black box* for the two extremes of the modelling continuum and take the position that black-box modelling is the superior choice of model for the assessment of the global-temperature simulation/observation inconsistency.

Our previous study (Leggett and Ball, 2015) utilised such a black-box approach, explaining the temperature simulation/observation inconsistency in terms of a highly statistically significant relationship between first-difference atmospheric CO₂ and global surface temperature. Pasini et al. (2016) also

Fig. 8 As for Fig. 7 with addition of Z-scored reverse volcanic aerosols series (green curve)

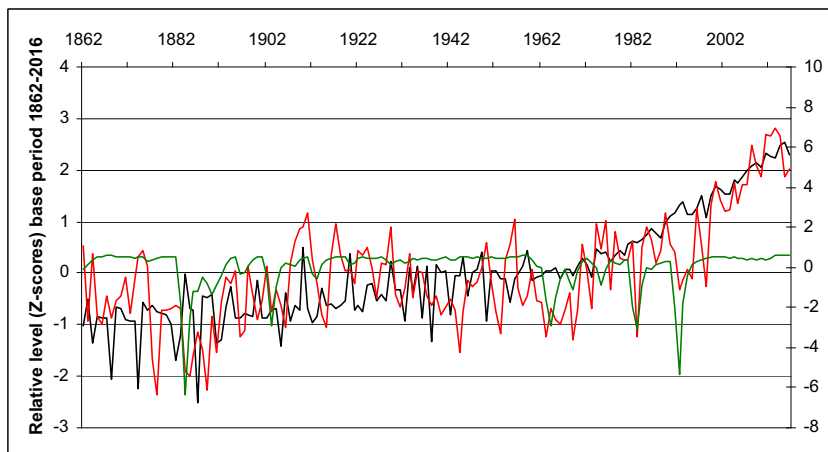


Table 1 Annual data from 1860 to 2014: results of ADF tests for trend stationarity allowing for both drift and trend for CO₂ gap, temperature gap and reverse volcanic aerosols

	ADF test statistic	P value	ADF test interpretation
a. level of series			
CO ₂ gap	-1.0845	0.9302	Non-stationary
Temperature gap	2.58709	0.2863	Non-stationary
Reverse volcanic aerosols	-5.08623	0.0001	Trend stationary
b. first difference of series			
CO ₂ gap	-8.94112	1.01E-15	Stationary
Temperature gap	-8.65338	1.11E-14	Stationary

used a black-box model—involving the level of CO₂ and an index related to natural variability patterns, the Atlantic Multidecadal Oscillation (AMO)—which was also able to correlate with the inconsistency.

This paper therefore will use a black-box method to assess the roles of level and first-difference CO₂ in relation to temperature.

2.2 Choice of the type of black-box model

Finding the best method within the realm of black-box modelling involves several considerations—the full range of black-box models from which to choose; and given that different models may have different strengths and weaknesses, consideration of the trade-offs required so that an optimised modelling procedure can be chosen.

Within the realm of black-box modelling, there are two main types of methods—univariate and multivariate (Greene 2012).

Multivariate modelling involves the utilisation of correlation between the outcome variable and at least one causal variable. Given the evidence for causality between atmospheric CO₂ and global surface temperature demonstrated using multivariate modelling (in its bivariate form) (Leggett and Ball 2015), the multivariate modelling approach is used in this study.

Within the realm of multivariate modelling, the base model is prepared by regression analysis—often termed ‘ordinary least squares analysis’ (Greene 2012).

There are two broad sub-categories: ordinary least squares (OLS) regression and regression corrected for a range of statistical issues that arise when the variables involved are time series. Time series models (Greene 2012) differ from ordinary regression models in that the results are in a sequence. Hence, the dependent variable cannot only be influenced by the independent variables, but also by *prior values of the dependent variable itself*.

Table 2 Eviews ARDL estimation output for temperature gap as a function of the CO₂ gap

Model-dependent variable	Number of models evaluated	Selected model	Long-run relationship		Short-run dynamic relationship			
			Bounds test F-statistic	Significance for both I(0) and I(1) orders of integration	Autocorrelation (out to 24 lags)	Adjusted R-squared	F-statistic	P value
a. bounds test and short-run model	100	ARDL (1, 0, 0)	16.35800	< 1%	Nil	0.657802	99	0***
b. long-run model	Coefficient	Std. error	t-statistic	Prob.				
CO ₂ gap	0.671595	0.091636	7.328958	0***				
Reverse volcanic aerosols	0.428447	0.095967	4.46453	0***				
C	0.001007	0.091352	0.011026	0.9912				

*** < 1% significance level

Fig. 9 Z-scores, 1862 to 2016. Observed temperature gap (red curve) and temperature gap ARDL-modelled from independent variables of CO₂ gap and reverse volcanic aerosols (black curve)

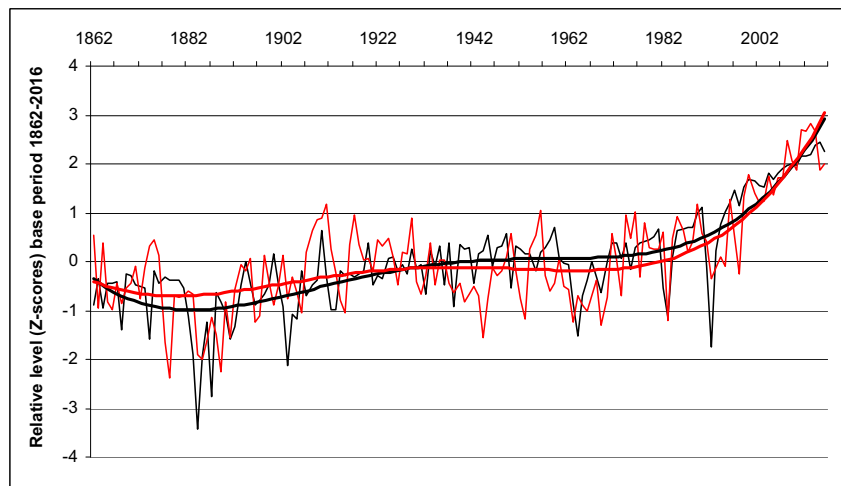


Table 3 Pairwise Granger causality analysis involving variables of temperature gap and CO₂ gap. Annual data from 1862 to 2016

Null hypothesis:	Obs	Chi-sq	df	Prob.
CO ₂ gap does not Granger cause temperature gap	350	9.883641	5	0.0786*
Temperature gap does not Granger cause CO ₂ gap		4.132023	5	0.5306

*10% significance level

This is termed autocorrelation between measured values. This serial nature of the measurements must be addressed by careful examination of the lag structure of the model. This type of OLS regression is termed ‘time series analysis’ (Greene 2012).

Given that our data are time series, the choice between OLS and time series analysis is straightforward—time series analysis is selected.

2.3 Choice of the type of time series model

The major issue in the realm of the analysis of the particular time series we are studying concerns what is termed the ‘order of integration’ of each of the series used.

Greene (2012) states: “The series y_t is said to be integrated of order one, denoted $I(1)$, because taking a first difference produces a stationary process. A non-stationary series is integrated of order d , denoted $I(d)$, if it becomes stationary after being first differenced d times. An $I(1)$ series in its raw

Fig. 10 Global precipitation (black curve) and global runoff (red curve) in Sverdrups (Sv), 1949 to 2012

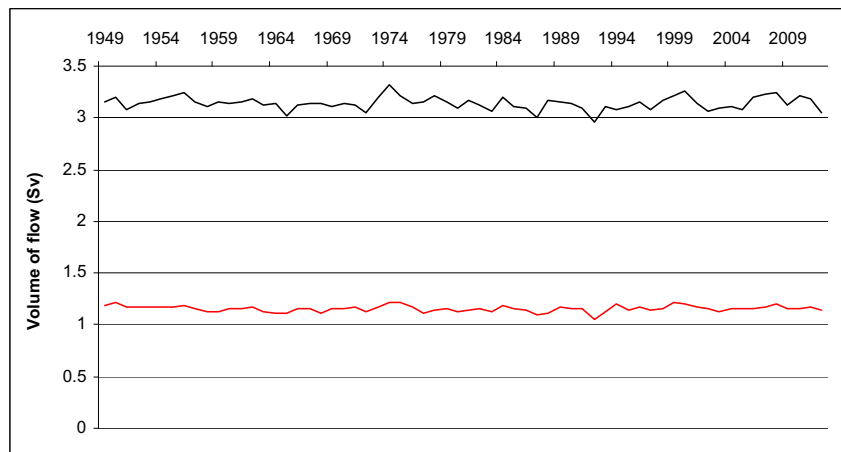
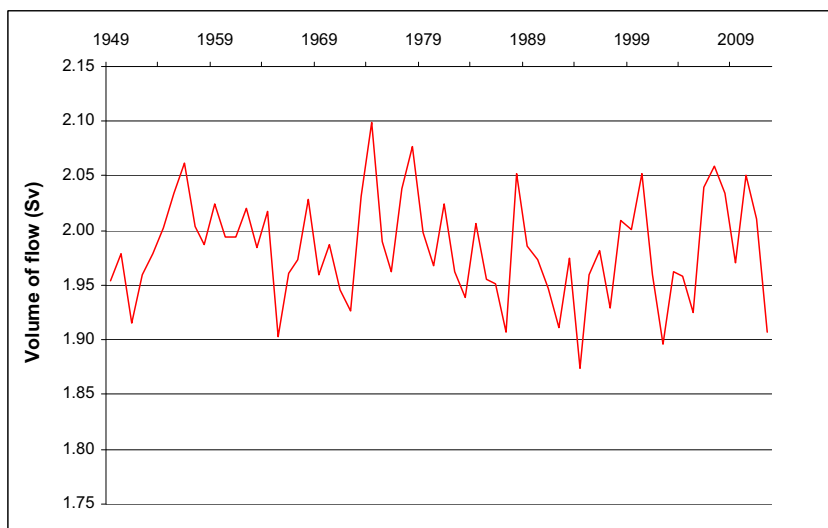


Fig. 11 Global evapotranspiration in Sverdrups (Sv), 1949 to 2012 (Dai 2016)



(undifferenced) form will typically be constantly growing, or wandering about with no tendency to revert to a fixed mean.”

It is not straightforward to deal with mixtures of series with different orders of integration. This was dealt within L&B (2015) within the vector autoregressive framework by the Toda and Yamamoto (1995) method. In returning to this question more comprehensively in this paper, we note that Greene (2012) (p. 999) and also, for example, Janjua et al. (2014) and Ahmad and Du (2017) show that the ARDL method (Pesaran et al. 2001) is the most comprehensive current way to address this question.

This is because this method can be used whether variables are purely of order of integration $I(0)$, purely $I(1)$ or a mixture of both $I(0)$ and $I(1)$.

For ARDL, the stationarity or otherwise of each series must still be assessed. This is to ensure that there are no $I(2)$ or higher series present. ARDL does not work for such series (Pesaran et al. 2001). To assess stationarity, a range of tests exists. In this study, the augmented Dickey-Fuller (ADF) test is used.

The ARDL method gives reliable results whether variables are purely of order of integration $I(0)$, purely $I(1)$ or a mixture of both $I(0)$ and $I(1)$ because it uses special significance tests (called ‘bounds tests’) (Pesaran et al. 2001). These test the significance of results against both an $I(0)$ realm and an $I(1)$ realm. If the result passes the test for each realm, a defensible model is obtained.

If the outcome of the bounds testing is positive, one next estimates a long-run ‘levels model’. These results are then used to measure short-run dynamic effects.

With this achieved, a model must be established in which any autocorrelation in the relationship, if present, is fully accounted for by use of an optimal lag structure. In this study, this is done within the ARDL process by use of an information criterion. Of the several types of information criteria available, following Greene (2012), we use the Schwarz criterion, which with its heavier penalty for degrees of freedom lost, “...leans toward a simpler model. All else given, simplicity does have some appeal.”

Fig. 12 Z-scores: global precipitation (CRUTS) 1901–2016 (black curve) compared with evapotranspiration 1949 to 2012 (Dai 2016) (red curve)

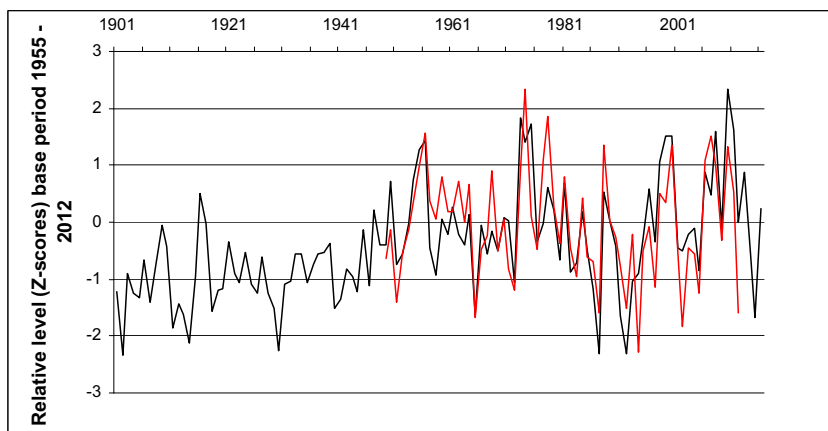


Table 4 Annual data from 1949 to 2012: level of series

	ADF test statistic	P value	ADF test interpretation
Precipitation	-5.4359	2.11E-05	Stationary
Evapotranspiration	-6.18452	1.42E-06	Stationary

2.4 Granger causality

Concerning the issue of causality in L&B (2015) we wrote:

One method using correlational data, however, approaches more closely the quality of information derived from random placement into experimental and control categories. The concept is that of Granger causality (Granger 1969). According to Stern and Kaufmann (2014), a time-series variable “x” (e.g. atmospheric CO₂) is said to “Granger-cause” variable “y” (e.g. surface temperature) if past values of x help to predict the current level of y, better than just the past values of y do, given all other relevant information.

Amblard and Michel (2013) point out that there is great closeness between Granger causality and information theory. We note the centrality to information theory of the case where information is being transmitted against a background containing a noise—a stochastic—element. With the noise idea uppermost, therefore, if some of the signature (‘signal’) of series a is seen at a later point in the makeup of the signal of series b (and the opposite is not true), this becomes the focus. Other differences between the series need not primarily concern us.

With this in mind, we seek observational climate series which precisely measure the variable we are interested in, and are extensive. It will be seen that this frequently requires using series which are ‘noisy’. We nonetheless use these and apply the ARDL and Granger causality analyses to determine whether useful information is obtained from their interactions.

As in L&B (2015), Granger causality analysis is implemented in the study by using a standard VAR model. As discussed, the Schwarz information criterion (SIC) is used to select an optimal maximum lag length (*k*) for the variables in the VAR. This lag length is then lengthened, if necessary, to ensure that firstly the estimated model was dynamically stable (i.e. all of the inverted roots of the characteristic equation lie inside the unit circle), and secondly, the errors of the equations are serially independent.

Granger causality results in this study are only reported if the VAR models meet the criteria in the preceding paragraph.

This study requires testing for Granger causality between the levels of some of the data series. In this case, the Granger

Table 5 OLS regression between evapotranspiration and precipitation, for annual data for the period 1949–2012

Dependent variable	Independent variables	Independent variable regression coefficient	Independent variable P value	Whole model adjusted R-squared	Whole model P value	Whole model LM test for autocorrelation up to order 4 P value	Whole model CUSUM test for parameter stability P value
Evapotranspiration	Const	-0.3093776	0.0014119***	0.5350152	4.04E-12	0.0736871	0.5715436
Interpretation	Precip_CRUTS_Z0116	0.726378	4.04E-12***			No significant autocorrelation	No instability

*** < 1% significance level

Table 6 Eviews ARDL estimation output for temperature gap as a function of precipitation and volcanic aerosols

Model-dependent variable	Number of models evaluated	Selected model	Long-run relationship		Short-run dynamic relationship			
			Bounds test <i>F</i> -statistic	Significance for both <i>I</i> (0) and <i>I</i> (1) orders of integration	Autocorrelation (out to 36 lags)	Adjusted <i>R</i> -squared	<i>F</i> -statistic	<i>P</i> value
a. bounds test and short-run model								
Temperature gap	12,500	ARDL (1, 0, 0, 1, 0)	19.79110	< 1%	Nil	0.783079	59.79101	0***
b. long-run model								
Long-run model independent variables	Coefficient	Std. error	<i>t</i> -statistic	Prob. *				
Dummy to 2003	-1.07554	0.219512	-4.899683	0***				
Dummy to 1937	0.848413	0.157018	5.403301	0***				
Dummy to 1980	-1.082175	0.156957	-6.894715	0***				
Precipitation	0.583684	0.08982	6.498402	0***				
Reverse volcanic aerosols	0.16765	0.061559	2.723415	0.0075***				
C	1.442896	0.177263	8.139878	0				

***< 1% significance level

causality testing procedure must be modified to allow for the differences in the orders of integration of the data series. Here, for each VAR model, the maximum lag length (k) is determined, but then one additional lagged value of each of the two variables is included in each equation of the VAR. However, the Wald test for Granger non-causality is applied only to the coefficients of the original k lags of CO₂. Toda and Yamamoto (1995) show that this modified Wald test statistic will still have an asymptotic distribution that is chi-square, even though the level of CO₂ is non-stationary, and the Granger causality test will be reliable.

We note that for reasons of avoiding duplication of results, from ARDL we focus on the long-run model results; short-term dynamics, which are at the heart of VAR Granger causality analysis, are dealt with there.

3 Methods and data

3.1 Specifics of implementation of ARDL modelling in this study

This method, in its output, selects the best lags for all variables and determines whether a significant model is possible.

Pesaran et al. (2001) point out that ARDL modelling is "... also based on the assumption that the disturbances ... are serially uncorrelated. It is therefore important that the lag order p of the underlying VAR is selected appropriately. There is a delicate balance between choosing p sufficiently large to mitigate the residual serial correlation problem and, at the same time, sufficiently small so that the conditional ECM is not unduly over-parameterized, particularly (when) limited time series data ... are available."

We therefore provide the ARDL method with an adequately large pool of runs from which to seek the best specified model—that is, one with the lowest Schwarz information criterion. Only models which pass all the tests for model specification listed in Section 3 above are used. Hence, each ARDL table is set up with a common structure as follows.

Each table has an 'a' and a 'b' section. The specific way in which each model passes the model specification tests is shown in the 'a' section of the table. The 'a' section is provided to illustrate that an adequately well-specified model has been achieved.

As outlined above, the ARDL logic is that significance is tested through the long-run relationship. If long-run significance is shown, the short-run relationship will also be significant. Hence, for the reason of showing the significance result, the long-run section precedes the short-run section in the 'a' section of each table.

The specific information sought from each well-specified ARDL model is then given in the 'b' section of each table. This relates to each of the potential driving variables, its

Fig. 13 Same as Fig. 12 but showing linear regression line for global precipitation

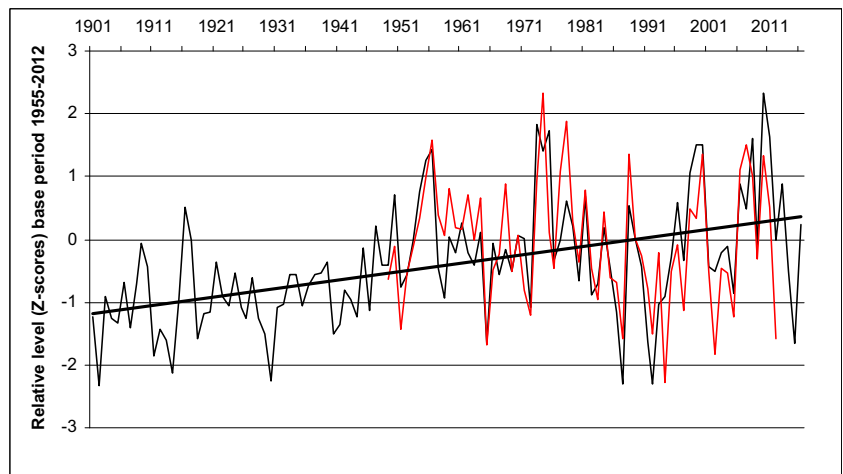


Fig. 14 Same as Fig. 13 but with addition of ocean heat to 700 m (purple curve), evapotranspiration (Zhang et al. 2015) (green curve) and NDVI (green curve)

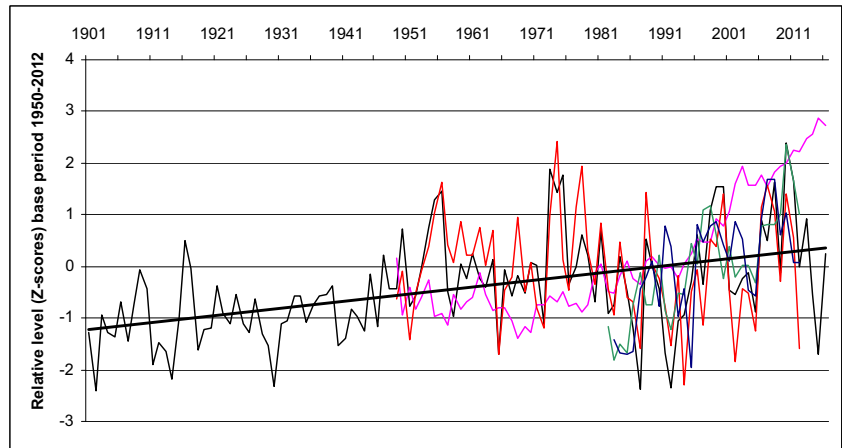


Fig. 15 Same as Fig. 13 but with addition of reverse volcanic aerosols series (bright green curve)

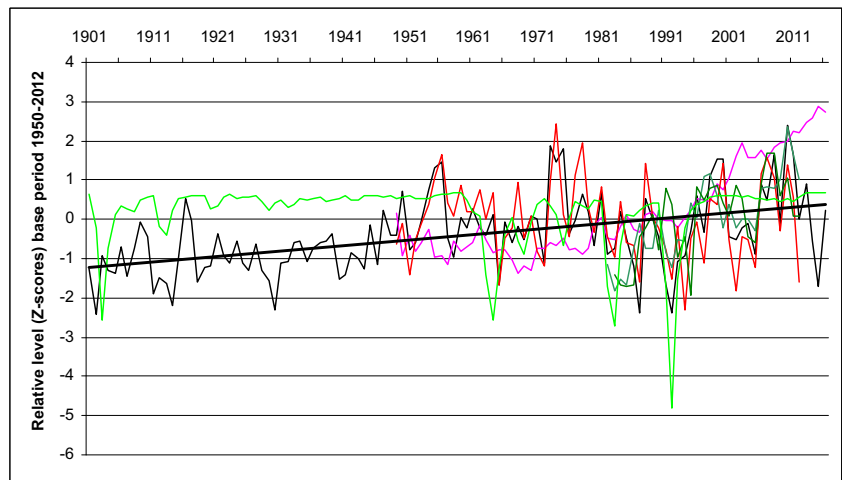
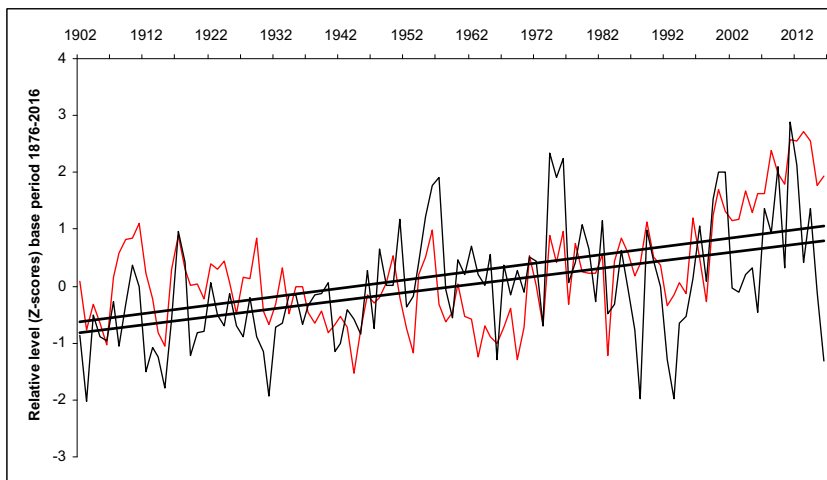


Fig. 16 Z-scores: the temperature gap (red curve) compared with precipitation (black curve). Linear regression trend lines for each series are also shown



degree of statistical significance and the relative percentage of the total driving task that it achieves.

3.2 Data, data sources and data terminology used

For global surface temperature, we used the Hadley Centre–Climate Research Unit combined Landsat and SST surface temperature series (HadCRUT) version 4.5.0.0 (Morice et al. 2012). In the tables, figures and text in the paper, this series is termed ‘global surface temperature’.

A data series projected from a business as usual global climate model (GCM) for global surface temperature, the CMIP5, RCP8.5 scenario model (Taylor et al. 2012) was used.

For atmospheric CO₂ data from 1958 to the present, the US Department of Commerce National Oceanic and Atmospheric Administration Earth System Research Laboratory Global Monitoring Division Mauna Loa, Hawaii, annual CO₂ series (Keeling et al., 2009) is used. For CO₂ data prior to 1958, ice core data is used (Rubino et al. 2013) (unsmoothed version of

ice core data series provided by Etheridge, personal communication, 2017). In the paper, this series is termed ‘atmospheric CO₂’.

Data for the evapotranspiration and ocean heat sink indicators are as follows. Global evapotranspiration is indicated by three data series (i) derived from global precipitation minus global runoff, (ii) from remote sensing (satellite) and (iii) the NDVI.

Annual global precipitation and runoff data from 1949 to 2012 is from Figure 2.8a of Dai (2016), data series depicted provided by Dai, personal communication (2017). A longer annual global precipitation data series from 1901 to 2016 which is used is from the CRU TS4.01 dataset (Harris et al., 2014). In the paper, these series are termed ‘precipitation’ and the specific precipitation series used in a particular instance is identified by its start and end dates.

The satellite data series for evapotranspiration is digitised from Figure 1a of Zhang et al. (2015). This series is termed ‘evapotranspiration’ (Zhang et al. 2015) in the paper. NDVI monthly data from 1980 to 2006 are from the Global

Fig. 17 The close match between the observed temperature gap (red curve) and the long-run ARDL model of the gap using precipitation, reverse volcanic aerosols and three dummy variables (black curve)

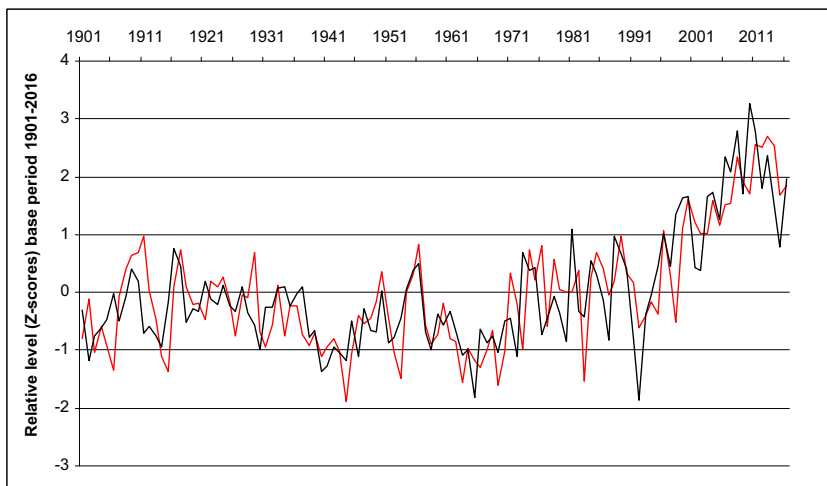


Table 7 Pairwise Granger causality analysis involving variables of temperature gap and precipitation. Annual data from 1901 to 2016

Null hypothesis:	Obs	Chi-sq	df	Prob.
Precipitation does not Granger cause temperature gap	109	40.95265	6	0.0000***
Temperature gap does not Granger cause CO ₂ gap		4.132023	5	0.5306

***< 0.001% significance level

Inventory Modeling and Mapping Studies (GIMMS) dataset (Tucker et al. 2005); NDVI data from 2006 to 2013 were provided by the Institute of Surveying, Remote Sensing and Land Information, University of Natural Resources and Life Sciences, Vienna. In the paper, the two series are merged for use and the merged series is termed ‘NDVI’.

According to AR4 (IPCC 2007) (p. 387), two-thirds of ocean-stored energy is absorbed between the surface and a depth of 700 m. Data for ocean heat content from the surface to 700 m is available (i) from the NOAA National Oceanographic Data Center (NODC) (Levitus et al. 2012) for the period 1955 to the present and (ii) from the Japan Marine Agency (JMA) (Ishii and Kimoto 2009) for the period 1950 to the present. For annual data, the longer JMA series is used. NOAA also provides quarterly data for world oceans from 1955 to the present. An approximately monthly series was prepared for use in this paper by expanding the quarterly data series. This was done by using the value for each quarter to stand for each month in the quarter. The ocean heat to 700 m series as used in the paper is termed ‘ocean heat’.

Volcanic aerosol data is from the National Aeronautic and Space Administration Goddard Institute for Space Studies Stratospheric Aerosol Optical Thickness series (Sato et al. 1993). As mentioned in Section 1.3, increasing volcanic aerosols correlate with reduced temperature. In radiative forcing terms, the forcing from volcanic aerosols

is – 27 times the optical thickness (Stern and Kaufmann 2014; Pasini et al. 2017). Hence, the volcanic aerosol series used in the paper is reversed and is termed ‘reverse volcanic aerosols’.

To make it easier to assess the relationship between the key climate variables visually, the data were normalised using statistical Z-scores or standardised deviation scores (expressed as “relative level” in the figures). In a Z-scored data series, each data point is part of an overall data series that sums to a zero mean and variance of 1, enabling comparison of data having different native units. Hence, when several Z-scored time series are depicted in a graph, all the time series will closely superimpose, enabling visual inspection to clearly discern the degree of similarity or dissimilarity between them. Individual figure legends contain details on the series lengths.

A regression using Z-scored variables provides standardised regression coefficients. These coefficients report how much change a one standard deviation change in the independent variable produces in the dependent variable. Although comparisons between these coefficients must be interpreted with care, a standardised coefficient of 4 for variable a, for example, indicates that independent variable a is twice as influential upon the dependent variable as another independent variable that has a standardised coefficient of 2 (Allen 1997).

In the time series analyses, the temperature gap and global atmospheric surface temperature are the dependent variables. Variability is explored using either interannual (yearly) or monthly data. The period covered in the figures is sometimes shorter than that used in the data preparation because of the loss of some data points due to calculations of differences and of moving averages.

Assessments were carried out using the time series statistical software packages Gnu Regression, Econometrics and Time-series Library (GRET, <http://gret.sourceforge.net>) and IHIS EViews (2017).

Fig. 18 Z-scores. Ocean heat series to 700 m. Sources: Japan Marine Agency (JMA) (black curve) and US National Oceanic Atmospheric Administration (NOAA) (red curve)

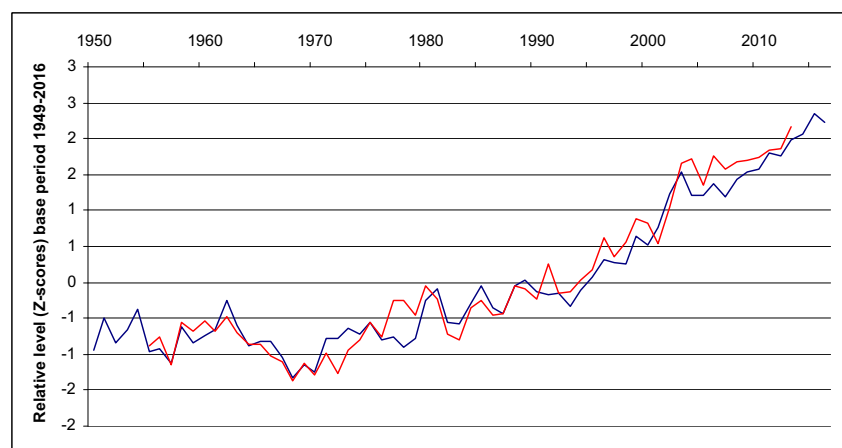


Table 8 Eviews ARDL estimation output for temperature gap as a function of precipitation, ocean heat and volcanic aerosols: bounds test and short-run model

Model-dependent variable	Number of models evaluated	Selected model	Long-run relationship		Short-run dynamic relationship				
			Bounds test <i>F</i> -statistic	Significance for both <i>I</i> (0) and <i>I</i> (1) orders of integration	Autocorrelation (out to 5 lags)	Adjusted <i>R</i> -squared	<i>F</i> -statistic	<i>P</i> value	
a. bounds test and short-run model									
Temperature gap	500	ARDL (1, 0, 0, 1, 0)	6.896771	< 1%	Nil	0.811587	48.3823	0***	
b. long-run model									
Long-run model independent variables	Coefficient	Std. error	<i>t</i> -statistic	Prob.*					
Ocean	0.783182	0.081911	9.561386	0***					
Precipitation	0.345264	0.117953	2.927126	0.0048***					
Reverse volcanic aerosols	0.076923	0.087531	0.878814	0.383					
C	0.017046	0.076143	0.223863	0.8236					

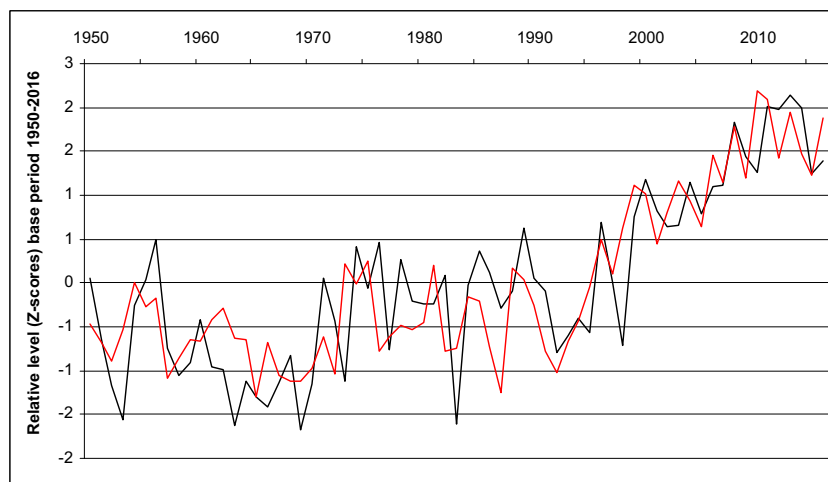
***<1% significance level

Table 9 Eviews ARDL estimation output for temperature gap as a function of precipitation and ocean heat

Model-dependent variable	Number of models evaluated	Selected model	Long-run relationship		Short-run dynamic relationship				
			Bounds test <i>F</i> -statistic	Significance for both <i>I</i> (0) and <i>I</i> (1) orders of integration	Autocorrelation (out to 5 lags)	Adjusted <i>R</i> -squared	<i>F</i> -statistic	<i>P</i> value	
a. Temperature gap	100	ARDL (1, 0, 2) (SIC)	8.464329	< 1%	Nil	0.812329	58.13568	0***	
b. Long-run model independent variables	Coefficient	Std. error	<i>t</i> -statistic	Prob.*					
Ocean	0.801983	0.080971	9.904548	0***					
Precipitation	0.382988	0.112499	3.404356	0.0012***					
C	0.018063	0.077526	0.232987	0.8166					

***< 1% significance level

Fig. 19 Observed temperature gap (red curve) and that from an ARDL long-run model involving precipitation and ocean heat (black curve)



3.3 Presentation

We note that to assist readability in text involving repeated references, atmospheric CO₂ is sometimes referred to simply as ‘CO₂’ and global surface temperature as ‘temperature’.

The time period covered and the frequency—annual or monthly—of the time series used in each table or figure are given in the title of the table or figure.

In the tables of results, statistical significance at the 10, 5 and 1% levels is indicated by the symbols *, ** and ***, respectively.

4 Results

4.1 Current status of temperature simulation/observation inconsistency

In L&B (2015), we illustrated the temperature simulation/observation inconsistency using data up to May 2015. Here, we assess the situation in the succeeding period to determine the extent to which it is still an issue. Figure 1 shows the situation including data up to May 2017. The figures show (using monthly data) the IPCC CMIP5 business-as-usual scenario RCP8.5, and temperature, labelled as H45.

The figure shows that the inconsistency continues past the period covered in our 2015 paper. Hence, the matters raised in L&B (2015) concerning the relative roles of level and rate of change of atmospheric CO₂ in correlations with global temperature are still a salient issue.

4.2 The CO₂ and temperature gaps in perspective from 1860

We now place the findings for the period 1959 to 2014 of L&B (2015) in a longer time perspective to address the

question of whether something relatively unprecedented is happening with the present CO₂ and temperature gaps. Annual data from 1860 to 2016 is used.

4.2.1 Calculation of CO₂ gap and temperature gap

The ingredients for the calculation of a single series for the atmospheric CO₂ gap are level of and first-difference CO₂.

To take the CO₂ data back before the start of the Mauna Loa record in 1958, ice core CO₂ data is used. Unsmoothed (raw) annual ice core level-of-CO₂ data is provided by Etheridge (personal communication, 2017) and is depicted in Fig. 2. There are a number of missing years: of the 152-year span of the annual series, data was present for 100 instances, or 65.8% of the total.

The time series modelling we are using requires continuous evenly spaced series (Mudelsee 2010). The missing data points are therefore interpolated. Linear interpolation from the adjacent data points is used (Mudelsee 2010). The resulting data series is shown in Fig. 3. Annual atmospheric CO₂ data is also shown at the same scale of parts per million (ppm) (Keeling et al. 2009).

Figure 3 shows the relationship in parts per million between the raw ice core data and atmospheric data (Mauna Loa). It can be seen that there is considerable similarity but some divergence in trend over the shared period. The ice core data is also considerably less smooth than the atmospheric data.

A single CO₂ series from 1861 to 2016 is prepared by using ice core CO₂ data in parts per million from 1861 to 1957 and atmospheric CO₂ data in parts per million thereafter from 1958 to 2016.

The issue of the differences between the ice core and the atmospheric CO₂ series becomes more marked when first-difference series are considered (Fig. 4).

Table 10 Eviews ARDL estimation output for temperature gap as a function of evapotranspiration, ocean heat and volcanic aerosols

Model-dependent variable	Number of models evaluated	Selected model	Long-run relationship		Short-run dynamic relationship			
			Bounds test <i>F</i> -statistic	Significance for both <i>I</i> (0) and <i>I</i> (1) orders of integration	Autocorrelation (out to 26 lags)	Adjusted <i>R</i> -squared	<i>F</i> -statistic	<i>P</i> value
a. bounds test and short-run model								
Temperature gap	500	ARDL (1, 1, 0, 0) (SIC)	14.42752	< 1%	Nil	0.751061	37.80802	0***
b. long-run model								
Long-run model independent variables	Coefficient	Std. error	<i>t</i> -statistic	Prob.*				
Evapotranspiration	0.38593	0.09958	3.875575	0.0003***				
Ocean heat	0.853227	0.085589	9.968877	0***				
Reverse volcanic aerosols	0.134912	0.073985	1.823513	0.0736*				
C	0.001314	0.070254	0.018707	0.9851				

*** < 1% significance level

* 10% significance level

Here, before the period of overlap, the amplitude of the variation in the ice core series is greater than for the atmospheric series. Further, the amplitude of the ice core series is lower later in its series, although still higher than the coincident atmospheric series.

We now assess whether or not the use of ‘smoothed’ series would be a better approach. According to Cook and Peters (1981) concerning time series from tree rings, “...smoothing ... is not a panacea for removing non-climatic variance in forest ring-width series. A certain amount of climatic information will always be lost due to the shape of the frequency response curves and where the signal and noise spectra overlap in the lower frequencies. These are problems common to any filtering operation.”

Given this point, we therefore wish to maximise the signal in the relationships we are investigating even at the expense of increased noise. This is consistent with the point made in Section 2.4 about the way in which Granger causality analysis can detect signals in noise. We therefore use the raw unsmoothed ice core series before the start of the atmospheric series in 1958 in this study.

4.2.2 The series used to derive the CO₂ and temperature gaps

From the CO₂ series above, merged series are produced, the merge occurring at the start of the atmospheric series in 1959. These CO₂ series are shown in Fig. 5. Also shown are the analogous temperature series: the data series projected from a simulated business-as-usual global climate model (GCM) for global surface temperature, the CMIP5, RCP8.5 scenario (Taylor et al. 2012) and the observed global surface temperature (HadCRUT 4.50).

The figure shows the marked difference in character of the first-difference CO₂ series from the other series. Despite this, however, it is seen that polynomial trendlines (indicated as ‘Poly.’ in the Key) show similarities in core trends between the level of CO₂ and RCP8.5 and between first-difference CO₂ and observed temperature.

4.3 The two gaps constructed and compared

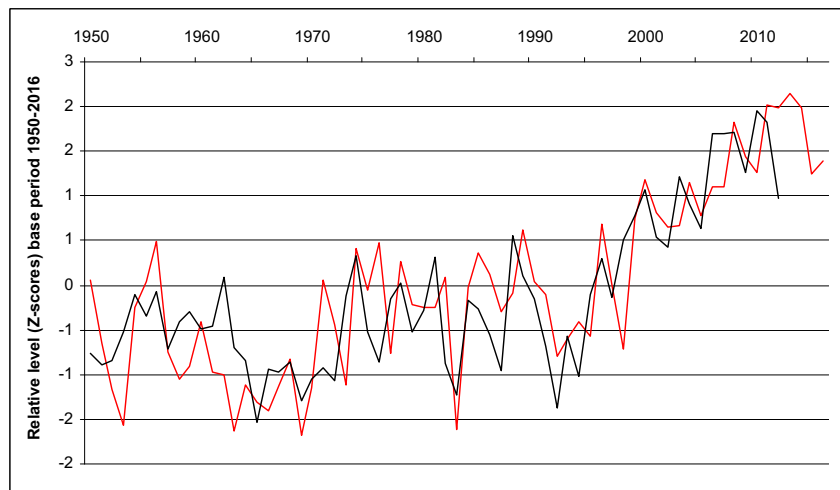
We take two approaches to determining the best standardisation to specify each gap. First, seeking the largest climate shifts from 1900, we find (Swanson and Tsonis 2009) shifts at 1912, 1942, 1976/1977 and 2001/2002. Of these, the most recent major shift is 1976/1977.

A second way is to seek the major point of departure for the ratio of first-difference CO₂ to CO₂. This is shown in Fig. 6.

Despite being very noisy, the figure shows that the major shift over the entire period occurs in the 1970s, a result in line with the above-mentioned 1976/1977 climate shift.

Hence, the standardisation chosen to specify the two gaps is to Z-score all four series shown in Fig. 5 using a base period

Fig. 20 Observed temperature gap (red curve) and that from an ARDL long-run model involving evapotranspiration, ocean heat and reverse volcanic aerosols (black curve)



from the start of the data to 1975—that is, the period before the 1976/1977 climate shift. The so Z-scored first-difference CO₂ series is then subtracted from the Z-scored level of CO₂ series to produce the CO₂ gap, and the Z-scored observed temperature series is subtracted from the Z-scored RCP8.5 simulation series to produce the temperature gap. These series are then themselves re-Z-scored, this time across the entire period (1862 to 2016). The two gaps so produced are shown in Fig. 7.

Figure 7 shows the following. First, due to the nature of the new gap series, the early large ice core CO₂ variance is reduced. Second, it is notable that there is great similarity between the two gaps.

Figure 8 shows that, further, some of the differences which do exist between the two gaps are matched by the volcanic aerosol series.

4.3.1 ARDL analysis of the relationship between the temperature gap and the CO₂ gap

Let us now use ARDL analysis to assess the relationship between the CO₂ and temperature gaps.

First, as required for the ARDL method, we check that none of the series are *I*(2). ADF tests assessing this are presented in Table 1a and b.

Table 11 Pairwise Granger causality analysis involving variables of temperature gap and evapotranspiration. Annual data from 1949 to 2016

Null hypothesis:	Obs	Chi-sq	df	Prob.
Evapotranspiration does not Granger cause temperature gap	60	10.73041	3	0.0133**
Temperature gap does not Granger cause evapotranspiration		5.26254	3	0.1536

**0.05% significance level

Table 1a shows that Volc is stationary in levels; Table 1b shows that the other two series are stationary when differenced—so no series is *I*(2), and the ARDL method can be used.

Table 2a and b show the Eviews ARDL estimation output for temperature gap as the dependent variable and CO₂ gap as the independent variable. Table 2a shows that both bounds test and residual diagnostics indicate that a well-specified model has been achieved. Table 2a also shows that the overall model is highly statistically significant. Table 2b shows that the CO₂ gap has a highly statistically significant relationship with the temperature gap. Further, Volc is significant in the model, although with a smaller regression coefficient than the CO₂ gap.

Figure 9 depicts the observed temperature gap and that fitted from the long-run model.

Notably, Fig. 9 provides clear evidence that the rise in both gaps seen since about 1960 is markedly greater in amplitude than any of the cyclic processes seen over the entire prior period back to the 1860s. As such, the present rise in both gaps is exceptional.

4.3.2 Granger causality analysis of the relationship between the two gaps

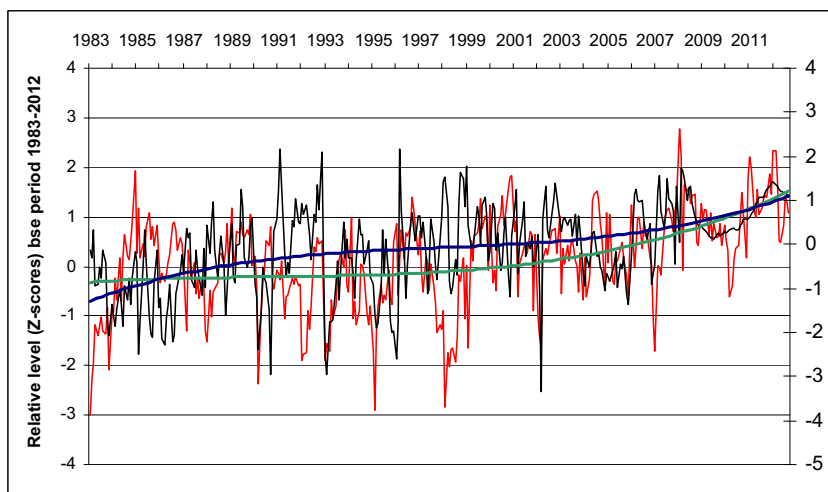
Next, we turn to Granger causality. As at least some series are not *I*(0), the Toda-Yamamoto variant to the VAR is employed

Table 12 Pairwise Granger causality analysis involving variables of temperature gap and ocean heat. Annual data from 1950 to 2016

Null hypothesis:	Obs	Chi-sq	df	Prob.
Ocean heat does not Granger cause temperature gap	66	6.391205	2	0.0409**
Temperature gap does not Granger cause ocean heat		1.390955	2	0.4988

**10% significance level

Fig. 21 The monthly temperature gap (red curve) in comparison with NDVI (black curve). Also given are third-order polynomials for temperature gap (green curve) and NDVI (blue curve)



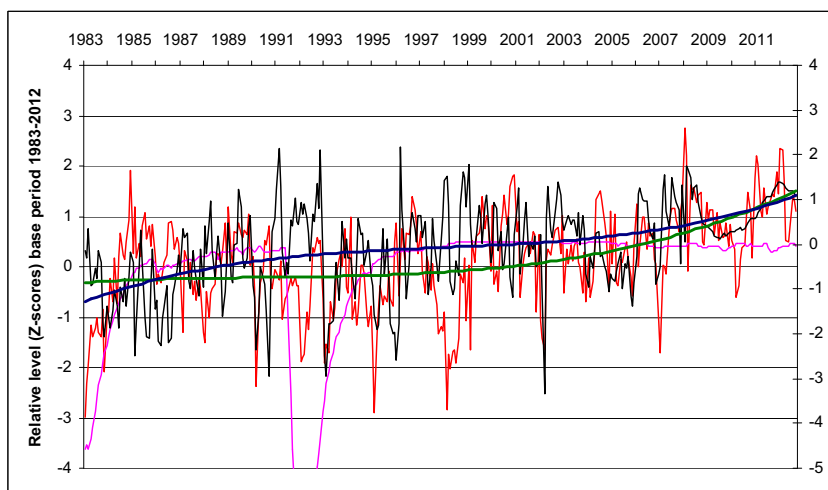
(see Section 3 above). Table 3 shows that causality from CO_2 to temperature is almost significant in the model.

This significance is lower than for the equivalent analysis shown in L&B (2015). This result is not surprising when it is recalled that L&B (2015) used monthly data and showed the driver first-difference CO_2 leading temperature by only a few months. Such a small lead cannot be expected to be strongly reflected in annual data. There is no evidence for causality in the other direction, from temperature gap to CO_2 gap.

4.4 Relationship between temperature gap and evapotranspiration and ocean heat sink

We now explore the involvement of candidate physical causes. The CO_2 gap being a non-physical construct, and the temperature gap being the final outcome, in what follows we investigate the relationship of the candidate physical causes simply with the temperature gap.

Fig. 22 As for Fig. 21, with addition of reverse volcanic aerosols series (purple curve)



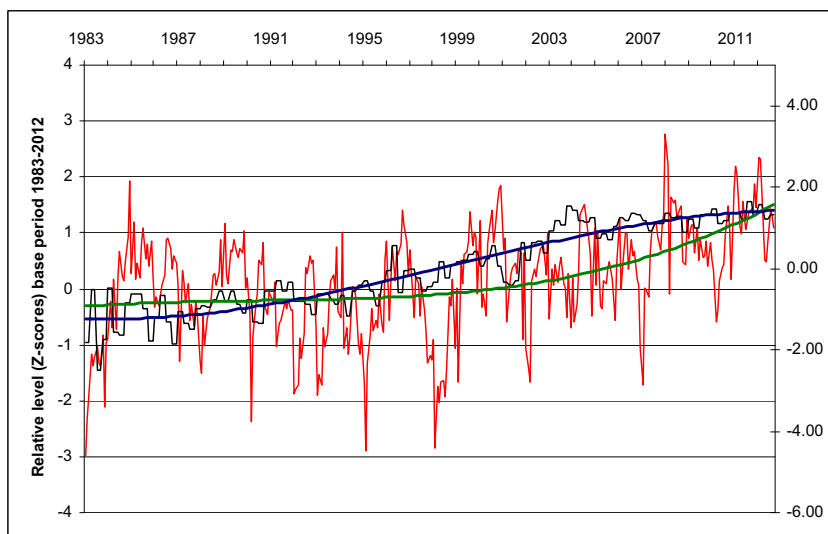
As outlined above, the potential causes of a temperature gap must affect the global heat budget. There are two accepted such factors—evapotranspiration via the terrestrial vegetation and the ocean.

4.4.1 Evapotranspiration

In Section 1, we cited evidence that evapotranspiration has the capacity to affect the global heat budget. Is there a source for an adequately long global evapotranspiration time series? Published evapotranspiration datasets are available. However, they are short in duration, not extending back before the 1980s (McCabe et al. 2016; Zhang et al. 2015).

An alternate approach is as follows. Walter et al. (2004) state: “Evaporation or evapotranspiration (ET) trends can be identified ... by using published data and assuming that the balance among precipitation, streamflow, and ET dominates the terrestrial hydrological budget on annual and longer time scales.”

Fig. 23 The monthly temperature gap (red curve) in comparison with ocean heat to 700 m (NOAA series) (black curve) gap. Third-order polynomials are given for temperature gap (green curve) and ocean heat (blue curve)



Hence, evapotranspiration can be derived by subtracting streamflow runoff from precipitation. Figure 10 shows global precipitation and global runoff from the longest global time series available—1949 to 2012 (Dai 2016).

It is noteworthy that the figure shows that runoff is only about a third of the size of precipitation, showing the marked size of evapotranspiration. Figure 11 shows the evapotranspiration series derived from this data.

Given that we have determined above that the CO₂ and temperature gaps are both rising, the fact that this putative gap driver series—the evapotranspiration series—is decreasing, is on the face of it inconsistent. This is explored further in the next section.

4.4.2 Evapotranspiration compared with precipitation

Recalling from Figs. 10 and 11 that, compared with runoff, precipitation will be the dominant series in the resulting ET series, Fig. 12 compares ET and precipitation, both series having been Z-scored.

The figure shows that there is a very close apparent similarity between the two series.

The degree of similarity is quantified as follows. First, to determine which of OLS or ARDL analysis should be used, we determine the order of integration of the series (Table 4).

Table 13 Monthly data from 1983 to 2012: results of ADF tests for trend stationarity allowing for both drift and trend for temperature gap, NDVI, oceanic heat and reverse volcanic aerosols

	ADF test statistic	P value	ADF test interpretation
Temperature gap	-6.53604	4.91E-08	Trend-stationary
LED14M NDVI	-6.87923	5.52E-09	Trend-stationary
Ocean heat	-3.88048	1.28E-02	Trend-stationary
Reverse volcanic aerosols	-3.72526	0.02066	Trend-stationary

The table shows both series are stationary over the period assessed, so OLS analysis can be used. This is provided in Table 5.

In particular, note the CUSUM test result which shows no significant series breaks skewing the relationship.

The results from Table 5 provide evidence that precipitation can be used as a surrogate for evapotranspiration. This now gives us a data series back to 1901.

Figure 13 is the same as Fig. 12, but with a linear regression line through precipitation. It can be seen that, despite much variation, overall, over the period 1901 to 2015, precipitation is rising.

It can now be seen that, if it were available, a longer-run evapotranspiration series overall might well also be rising.

4.4.3 All candidate drivers of temperature gap compared

Figure 14 adds more candidate drivers of the temperature gap. The drivers are ocean heat, a satellite measure of evapotranspiration and NDVI.

They all rise themselves over the periods shown and show various degrees of similarity in signature.

When volcanic aerosols (series reversed) are added (Fig. 15), correlations with part of the depression in the signatures from 1960 to 1990 are seen.

4.4.4 The longest-series candidate driver—the evapotranspiration surrogate, precipitation—compared with the temperature gap

We now begin our formal assessments. We commence by comparing the precipitation driver candidate with the temperature gap.

We first note that the gap and precipitation each display not only an overall net rising trend but also that the slope of the

Table 14 Eviews ARDL estimation output for temperature gap as a function of NDVI, ocean heat and volcanic aerosols

Model-dependent variable	Number of models evaluated		Selected model	Long-run relationship		Short-run dynamic relationship				
				Bounds test <i>F</i> -statistic	Signif. for both <i>I</i> (0) and <i>I</i> (1) orders of integration	Autocorrelation (out to 36 lags)	Adjusted <i>R</i> -squared	<i>F</i> -statistic	<i>P</i> value	
a. bounds test and short-run model										
Temperature gap	4		ARDL (2, 0, 0, 0)	10.28399	< 1%	Nil	0.598103	114	0***	
b. long-run model										
Long-run model independent variables	Coefficient		Std. error	<i>t</i> -statistic	Prob.*					
Led14M NDVI	0.275921		0.121836	2.26468	0.0241**					
Reverse volcanic aerosols	0.158744		0.096165	1.650756	0.0996*					
Ocean heat	0.23989		0.126301	1.899362	0.0583*					
C	0.231184		0.105615	2.188928	0.0292**					

*** < 1% significance level

* 10% significance level; ** 5% significance level

linear trend of each is very similar. Second, the gap series shows descending sections like the precipitation series (and it will be recalled ET (Fig. 5)) over the 1950–1990 period.

4.4.5 Data from 1901: temperature gap as a function of precipitation and volcanic aerosols

First, we use the OLS relationship between precipitation and the temperature gap to determine by rolling Chow testing the series breaks that are present and, hence, any dummy variables that are called for. Series breaks are found and the dummy variables arising are seen in Table 6b. It is notable that the dummy variables recommended by Chow testing are very close to the three of the four largest climate shifts from 1900 shown by Swanson and Tsonis (2009) as being at 1912, 1942, 1976/1977 and 2001/2002 (Fig. 16).

We now run an ARDL model with these independent variables and these dummy variables. Results are given in Table 6a and b.

Table 6a shows that a well-specified model is obtained. Precipitation is significant and substantive in the model (Table 6b). Figure 17 shows the close match between the above long-run ARDL model and the temperature gap.

We now turn to assessing the presence of Granger causality between precipitation and the temperature gap. As at least some series are not *I*(0), the Toda-Yamamoto extra step to the VAR method is employed. Results are given in Table 7.

Notably, the table shows that, over the period from 1901 to 2016, precipitation is strongly Granger causal of the temperature gap. There is no evidence for causality in the other direction.

We now turn to the period from 1950 to 2016, over which we can include ocean heat data alongside that for precipitation and volcanic aerosols.

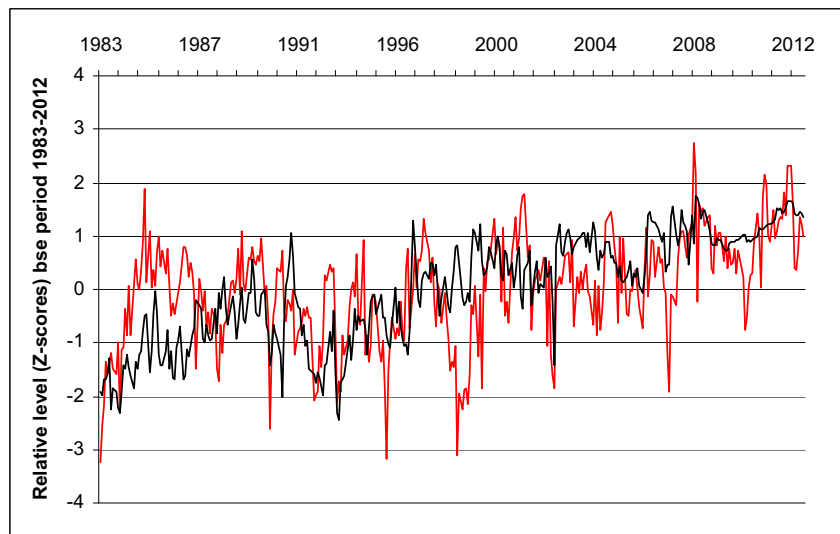
For annual data, there are two main published ocean heat series available, one from the National Oceanic and Atmospheric Administration (NOAA) and the other from the Japan Marine Agency (JMA) (see Section 3.2 above). Figure 18 shows that these series agree closely over their common years. For annual data, we will use the longer JMA series.

4.4.6 Data from 1950: temperature gap as a function of precipitation, ocean heat and volcanic aerosols

In this section, the relationship between the temperature gap and precipitation, ocean heat and volcanic aerosols is assessed by ARDL and Granger causality analysis. Table 8a and b provide the results of the ARDL analysis.

Table 8a and b show that with ocean heat included, volcanic aerosol is now no longer significant in the model. Hence, a further ARDL model is run without volcanic aerosol. These results are given in Table 9a and b.

Fig. 24 Observed temperature gap (red curve) and that from an ARDL model involving NDVI, ocean heat and reverse volcanic aerosols (black curve)



It is seen that both precipitation and ocean heat are significant in the model, with the strength of the precipitation variable being about half that of the ocean heat variable.

Figure 19 depicts the above model in comparison with the temperature gap.

4.4.7 Data from 1950: temperature gap in relation to evapotranspiration, ocean heat and volcanic aerosols

We now substitute the true evapotranspiration series for the precipitation series and conduct further ARDL and Granger causality analysis. Table 10a and b provide the results of the ARDL analysis.

Table 10b shows that evapotranspiration, like precipitation, is significant in the ARDL model.

Referring to Table 9b for precipitation and Table 10b for evapotranspiration, it is observed that both precipitation and evapotranspiration have coefficients of very similar value, and that this occurs against a coefficient for ocean heat which is of very similar value in both ARDL assessments.

Figure 20 depicts the above model in comparison with the temperature gap.

We now turn to Granger causality analysis of the relationship between the temperature gap and, respectively,

Table 15 Pairwise Granger causality analysis involving variables of temperature gap and NDVI. Monthly data from January 1983 to September 2012

Null hypothesis:	Obs	Chi-sq	df	Prob.
NDVI does not Granger cause temperature gap	350	8.400914	2	0.0150**
Temperature gap does not Granger cause NDVI		3.357008	2	0.1867

**5% significance level

evapotranspiration and ocean heat. As the temperature gap series is $I(1)$, the Toda-Yamamoto extra step to the VAR method is employed. The result of the analysis for evapotranspiration is shown in Table 11.

The table shows that, despite its lack of an upward trend, evapotranspiration is Granger causal of the temperature gap, and that there is no evidence for causality in the opposite direction.

The presence of Granger causality between ocean heat and the temperature gap is next assessed. As it can be shown that both series are $I(1)$, the Toda-Yamamoto extra step to the VAR method is employed. The result of the analysis is shown in Table 12.

Table 12 shows that, as for evapotranspiration, ocean heat is Granger causal of the temperature gap, and that there is no evidence for causality in the opposite direction.

4.5 Monthly data from 1983: temperature gap in relation to vegetation index, ocean heat and volcanic aerosols

We now test to see whether the relationships seen above using annual data can also be seen using monthly data, albeit over the more recent period for which such data is available and using NDVI as a substitute for evapotranspiration.

First, let us conduct a visual inspection of the several data series. Figure 21 shows the monthly temperature gap in comparison with NDVI.

In Fig. 21, polynomial curves show that, over the shorter time frame, an increasing temperature gap as shown in the annual data results can be seen. As expected from the increasing annual evapotranspiration series over the period from 1983 (see Fig. 14, monthly NDVI is also seen to increase. The best curve fit for NDVI is observed when it leads the temperature gap by 14 months.

Table 16 Pairwise Granger causality analysis involving variables of temperature gap and ocean heat

Null hypothesis:	Obs	Chi-sq	df	Prob.
Ocean heat does not Granger cause temperature gap	350	12.90040	7	0.0746*
Temperature gap does not Granger cause ocean heat		10.47154	7	0.1634

*10% significance level

Figure 22 adds in the volcanic aerosol series and shows how some of the difference between NDVI and the temperature gap is matched by the volcanic aerosol series.

Figure 23 shows the temperature gap and the ocean heat series. As expected, the figure shows ocean heat also rising.

4.5.1 ARDL and Granger causality analysis

In this section, the relationship between the temperature gap and NDVI, ocean heat and volcanic aerosols is assessed by ARDL and Granger causality analysis.

Table 13 displays the results of the ADF analysis. The table provides the stationarity characteristics of the series.

Table 13 also shows that all series are trend-stationary. Hence, OLS could be used for the analysis. That said, for consistency with the other analyses presented in the paper, analysis is carried out using ARDL. Table 14a and b provide the results of the ARDL analysis.

Table 14a shows that the model passes bounds tests and shows no autocorrelation in near months and, hence, that a well-specified ARDL model is obtained.

Table 14b shows that NDVI is statistically significant in the model. Ocean heat and volcanic aerosols are significant at the 0.1 level. Both NDVI and ocean heat show regression coefficients of similar value.

Figure 24 shows the observed temperature gap and the model derived from the ARDL analysis.

Tables 15 and 16 provide the results of pairwise Granger causality analyses for the temperature gap with (i) NDVI and (ii) ocean heat variables.

Tables 15 and 16 also show that, for monthly data over the period studied, each of NDVI and ocean heat show Granger causality of the observed temperature gap (ocean heat at the lower 10% statistical significance level). Causality in the other direction is not shown in either case.

The results in this section show that the processes observed at the annual level are also seen at the monthly level.

5 Future prospects

What then of the future? We have provided evidence that evapotranspiration and the ocean heat sink are responsible for the observed temperature gap so far. To what extent do these two factors have the capacity to maintain the gap into the future, especially under the scenario of a further increase in atmospheric CO₂?

For evapotranspiration, Dai (2016) in his Figure 16 summarises research which shows that with rising atmospheric CO₂, substantially increased evapotranspiration is expected over the remainder of the twenty-first century. For the ocean heat sink, the IPCC AR4 (IPCC 2007) p. 389 states that the ocean's heat capacity is about 1000 times larger than that of the atmosphere.

These two prospects suggest that evapotranspiration and the ocean heat sink might have the capacity to continue to maintain the temperature gap into the future. However, there is evidence that cyclic factors affect heat storage by the ocean. In a review of evidence for potential mechanisms for the functioning of the oceanic heat sink as it relates to the global surface temperature slowdown, Yan et al. (2016) list such cyclic factors as the concurrent effects of changing amplitudes of the Pacific Decadal Oscillation (PDO) or the AMO. Another potential mechanism (suggested as a main cause for the global surface temperature slowdown) involves the movement of heat to deeper layers of the Atlantic and Southern Oceans, one possible explanation for this increased heat storage being a salinity-driven mechanism.

Table 17 Summary of results of ARDL and VAR Granger causality analyses of relationships between proposed driver variables and temperature gap

Independent variable	1901–2016		1950–2016		1983–2012	
	ARDL	Granger causality	ARDL	Granger causality	ARDL	Granger causality
Precipitation	Yes	Yes	Yes	Yes	(b)	(b)
Reverse volcanic aerosols	Yes	(b)	Yes	(b)	Yes	Yes
Ocean heat	(a)	(a)	Yes	Yes	Yes	Yes
Evapotranspiration	(a)	(a)	Yes	Yes	Yes	(a)
NDVI	(a)	(a)	(a)	(a)	Yes	Yes

(a) Not available; (b) not in scope of assessment

The study proposing this salinity-driven mechanism (Chen and Tung 2014) also noted the presence of oscillatory cycles in this effect.

Under one scenario, these cycles could create a situation in which to a greater or lesser degree some of the heat sequestered in the deep ocean would be returned in the future to the surface and to the atmosphere (Trenberth and Fasullo 2010), being a driver of further warming. In this case, evapotranspiration becomes crucial as the remaining driver which would tend to keep the observed temperature lower than the level-of-CO₂ models would suggest.

6 Discussion

This paper has provided three further pieces of evidence concerning the global surface temperature slowdown.

Expressing the slowdown as a temperature gap, and developing an analogous CO₂ gap, the first piece of evidence is that that the present size of the global surface temperature and CO₂ gaps seen since about the 1960s is unprecedented over a period starting at least as far back as the 1860s.

Second, Table 17 summarises the findings for the ARDL and Granger causality analyses involving the global surface temperature gap against all physical candidate drivers over all periods assessed and for both annual and monthly data series.

Table 17 shows that in all cases where data was available, evapotranspiration or its argued surrogate, precipitation, showed ARDL correlation with and Granger causality of the temperature gap. This was true initially in isolation when ocean heat data was not available. In each later case where ocean heat data was available, evapotranspiration or precipitation was also significant ('required') in models alongside ocean heat. In terms of relative scale, while the standardised regression coefficient for evapotranspiration was repeatedly about half that for ocean heat, the standardised regression coefficients for both these variables were always larger than that for the volcanic aerosol variable. In all the circumstances, we consider that the foregoing shows that, alongside the ocean heat sink, evapotranspiration is likely to be making a substantial contribution to the global atmospheric temperature outcome.

Third, there is evidence that evapotranspiration and the ocean heat sink might be able to continue to maintain the temperature gap into the future—in other words to continue keeping the temperature lower than the level-of-CO₂ models would suggest.

This last point means there can be benefit in using the straightforward first-difference CO₂ to temperature relationship shown in L&B (2015) to forecast future global surface temperature. This will be the subject of a future paper.

Acknowledgements The authors would like to acknowledge with appreciation the support and advice of J. Gordon and C. Dawson and the helpful and constructive comments of an anonymous referee of the paper.

References

- Adams HD, Williams AP, Xu C, Rauscher SA, Jiang X, McDowell NG (2013) Empirical and process-based approaches to climate-induced forest mortality models. *Front Plant Sci* 4:438
- Ahmad N, Du L (2017) Effects of energy production and CO₂ emissions on economic growth in Iran: ARDL approach. *Energy* 123:521–537. <https://doi.org/10.1016/j.energy.2017.01.144>
- Allen MP (1997) Understanding regression analysis. Plenum, New York
- Amblard P-O, Michel OJ (2013) The relation between Granger causality and directed information theory: a review. *Entropy* 15:113–143
- Attanasio A, Pasini A, Triacca U (2012) A contribution to attribution of recent global warming by out-of-sample Granger causality analysis. *Atmos Sci Lett* 13(1):67–72. <https://doi.org/10.1002/asl.365>
- Attanasio A, Pasini A, Triacca U (2013) Granger causality analyses for climatic attribution. *Atmos Clim Sci* 3:515–522
- Ban-Weiss GA, Bala G, Cao L, Pongratz J, Caldeira K (2011) Climate forcing and response to idealized changes in surface latent and sensible heat. *Environ Res Lett* 6:1–8
- Bounoua L, Collatz GJ, Los SO, Sellers PJ, Dazlich DA, Tucker CJ, Randall DA (2000) Sensitivity of climate to changes in NDVI. *J Clim* 13(13):2277–2292. [https://doi.org/10.1175/1520-0442\(2000\)013<2277:SOCTCI>2.0.CO;2](https://doi.org/10.1175/1520-0442(2000)013<2277:SOCTCI>2.0.CO;2)
- Bounoua L, Hall FG, Sellers PJ, Kumar A, Collatz GJ, Tucker CJ, Imhoff ML (2010) Quantifying the negative feedback of vegetation to greenhouse warming: A modeling approach. *Geophys Res Lett* 37(23):n/a-n/a
- Chen X, Tung K-K (2014) Varying planetary heat sink led to global-warming slowdown and acceleration. *Science* 345(6199):897–903. <https://doi.org/10.1126/science.1254937>
- Cook ER, Peters K (1981) The smoothing spline: a new approach to standardizing forest interior tree-ring width series for dendroclimatic studies. *Tree-Ring Bull* 41:45–53
- Dai A (2016) Historical and future changes in streamflow and continental runoff: a review. *Terrestrial water cycle and climate change: natural and human-induced impacts*. *Geophys Monogr* 221:17–37
- Enting IG (1987) A modelling spectrum for carbon cycle studies. *Math Comput Simul* 29(2):75–85. [https://doi.org/10.1016/0378-4754\(87\)90099-1](https://doi.org/10.1016/0378-4754(87)90099-1)
- Enting IG (2010) Inverse problems and complexity in earth system science. In: Dewar RL, Detering F (eds) *Complex physical, biophysical and econophysical systems*. Singapore, World Scientific
- Flato G, Marotzke J, Abiodun B, Braconnot P, Chou SC, Collins W, Cox P, Driouech F, Emori S, Eyring V, Forest C, Gleckler P, Guilyardi E, Jakob C, Kattsov V, Reason C, Rummukainen M (2013) Evaluation of climate models. In: Stocker TF, Qin D, Plattner G-K, Tignor M, Allen SK, Boschung J, Nauels A, Xia Y, Bex V, Midgley PM (eds) *Climate change 2013: The physical science basis. Contribution of working group I to the fifth assessment report of the intergovernmental panel on climate change*. Cambridge University Press, Cambridge
- Fyfe JC, Gillett NP, Zwiers FW (2013) Overestimated global warming over the past 20 years. *Nat Clim Chang* 3(9):767–769. <https://doi.org/10.1038/nclimate1972>
- Good SP, Noone D, Bowen G (2015) Hydrologic connectivity constrains partitioning of global terrestrial water fluxes. *Science* 349(6244):175–177
- Granger CWJ (1969) Investigating Causal Relations by Econometric Models and Cross-spectral Methods. *Econometrica* 37(3):424

- Granger CWJ (1980) Testing for causality: a personal viewpoint. *J Econ Dyn Control* 2:329–352. [https://doi.org/10.1016/0165-1889\(80\)90069-X](https://doi.org/10.1016/0165-1889(80)90069-X)
- Grassi S, Hillebrand E, Ventosa-Santaulària D (2013) The statistical relation of sea-level and temperature revisited. *Dyn Atmos Oceans* 64: 1–9. <https://doi.org/10.1016/j.dynatmoce.2013.07.001>
- Greene WH (2012) *Econometric analysis*, 7th edn. Prentice Hall, Boston
- Harris I, Jones PD, Osborn TJ, Lister DH (2014) Updated high-resolution grids of monthly climatic observations—the CRU TS3.10 dataset. *Int J Climatol* 34:623–642
- IHS EViews: EViews 9.5, IHS Global Inc., Irvine, California, 2017. available at: <http://www.eviews.com/download/download.shtml> (last accessed 20 July 2017)
- IPCC (2007) Solomon S, Qin D, Manning M, Chen Z, Marquis M, Averyt KB, Tignor M, Miller HL (eds) *Climate change 2007: the physical science basis. Contribution of working group I to the fourth assessment report of the intergovernmental panel on climate change*. Cambridge University Press, Cambridge
- IPCC (2013) Stocker TF, Qin D, Plattner G-K, Tignor M, Allen SK, Boschung J, Nauels A, Xia Y, Bex V, Midgley PM (eds) *Climate change 2013: the physical science basis. Contribution of working group I to the fifth assessment report of the intergovernmental panel on climate change*. Cambridge University Press, Cambridge
- Ishii M, Kimoto M (2009) Reevaluation of historical ocean heat content variations with time-varying XBT and MBT depth bias corrections. *J Oceanogr* 65:287–299 Ocean heat data used available at http://www.data.jma.go.jp/gmd/kaiyou/data/english/ohc/ohc_global.txt (last accessed 12 October 2017)
- Janjua PZ, Samad G, Khan N (2014) Climate change and wheat production in Pakistan: an autoregressive distributed lag approach. *NJAS-Wageningen J Life Sci* 68:13–19. <https://doi.org/10.1016/j.njas.2013.11.002>
- Jasechko S, Sharp ZD, Gibson JJ, Birks SJ, Yi Y, Fawcett PJ (2013) Terrestrial water fluxes dominated by transpiration. *Nature* 496(7445):347–350
- Karplus WJ (1977) The spectrum of mathematical modelling and systems simulation. *Math Comput Simul* 19(1):3–10. [https://doi.org/10.1016/0378-4754\(77\)90034-9](https://doi.org/10.1016/0378-4754(77)90034-9)
- Karplus WJ (1992) *The heavens are falling: the scientific prediction of catastrophes in our time*. Plenum, New York. <https://doi.org/10.1007/978-1-4899-6024-5>
- Keeling RF, Piper SC, Bollenbacher AF, Walker SJ (2009) Carbon Dioxide Research Group, Scripps Institution of Oceanography (SIO), University of California, La Jolla, California USA 92093-0444, available at: <http://cdiac.ornl.gov/ftp/trends/CO2/maunaloa.CO2> (last accessed 14 September 2017)
- Kleidon A, Fraedrich K, Heimann M (2000) A green planet versus a desert world: estimating the maximum effect of vegetation on the land surface climate. *Clim Chang* 44(4):471–493. <https://doi.org/10.1023/A:1005559518889>
- Leggett LMW, Ball DA (2015) Granger causality from changes in level of atmospheric CO₂ to global surface temperature and the El Niño–Southern Oscillation, and a candidate mechanism in global photosynthesis. *Atmos Chem Phys* 15(20):11571–11592. <https://doi.org/10.5194/acp-15-11571-2015>
- Levitus S et al (2012) World ocean heat content and thermosteric sea level change (0–2000 m), 1955–2010. *Geophys Res Lett* 39:L10603 Ocean heat data used available at <https://www.nodc.noaa.gov/OC5/3M%5fHEAT%5fCONTENT/index.html>, last accessed 12 October 2017
- Mazzocchi F, Pasini A (2017) Climate model pluralism beyond dynamical ensembles. *WIREs Clim Change* 8(6):e477. <https://doi.org/10.1002/wcc.477>
- McCabe MF, Ershadi A, Jiménez C, Miralles DG, Michel D, Wood EF (2016) The GEWEX LandFlux project: evaluation of model evaporation using tower-based and globally gridded forcing data. *Geosci Model Dev* 9(1):283–305. <https://doi.org/10.5194/gmd-9-283-2016>
- Meehl GA, Arblaster JM, Fasullo JT, Hu A, Trenberth KE (2011) Model-based evidence of deep-ocean heat uptake during surface-temperature hiatus periods. *Nat Clim Chang* 1(7):360–364. <https://doi.org/10.1038/nclimate1229>
- Montgomery DC, Jennings CL, Kulahci M (2008) *Introduction to time series analysis and forecasting*. Wiley, New York
- Moore JC, Grinsted A, Zwinger T, Jevrejeva S (2013) Semi-empirical and process-based global sea level projections. *Rev Geophys* 51(3):484–522. <https://doi.org/10.1002/rog.20015>
- Morice CP, Kennedy JJ, Rayner NA, Jones PD (2012) Quantifying uncertainties in global and regional temperature change using an ensemble of observational estimates: The HadCRUT4 data set. *J Geophys Res: Atmos* 117(D8):n/a–n/a
- Mudelsee M (2010) *Climate time series analysis*. Springer, Dordrecht. <https://doi.org/10.1007/978-90-481-9482-7>
- Newman M (2013) An empirical benchmark for decadal forecasts of global surface temperature anomalies. *J Clim* 26(14):5260–5269. <https://doi.org/10.1175/JCLI-D-12-00590.1>
- Notaro M, Liu Z (2008) Statistical and dynamical assessment of vegetation feedbacks on climate over the boreal forest. *Clim Dyn* 31(6): 691–712. <https://doi.org/10.1007/s00382-008-0368-8>
- Pasini A, Mazzocchi F (2015) A multi-approach strategy in climate attribution studies: is it possible to apply a robustness framework? *Environ Sci Pol* 50:191–199. <https://doi.org/10.1016/j.envsci.2015.02.018>
- Pasini A, Lore M, Ameli F (2006) Neural network modelling for the analysis of forcings/temperature relationships at different scales in the climate system. *Ecol Model* 191(1):58–67. <https://doi.org/10.1016/j.ecolmodel.2005.08.012>
- Pasini A, Triacca U, Attanasio A (2012) Evidence of recent causal decoupling between solar radiation and global temperature. *Environ Res Lett* 7:034020 (6pp)
- Pasini A, Triacca U, Attanasio A (2016) Evidence for the role of the Atlantic multidecadal oscillation and the ocean heat uptake in hiatus prediction. *Theor Appl Clim*. (published online). <https://doi.org/10.1007/s00704-016-1818-6>
- Pasini A, Racca P, Amendola S, Cartocci G, Cassardo C (2017) Attribution of recent temperature behaviour reassessed by a neural-network method. *Sci Rep* 7(1):17681. <https://doi.org/10.1038/s41598-017-18011-8>
- Pesaran MH, Shin Y, Smith RJ (2001) Bounds testing approaches to the analysis of level relationships. *J Appl Econ* 16(3):289–326. <https://doi.org/10.1002/jae.616>
- Rahman MR, Lateh H (2015) Climate change in Bangladesh: a spatio-temporal analysis and simulation of recent temperature and rainfall data using GIS and time series analysis model. *Theor Appl Climatol* 128(1-2):27–41. <https://doi.org/10.1007/s00704-015-1688-3>
- Rubino M, Etheridge DM, Trudinger CM, Allison CE, Battle MO, Langenfelds RL, Steele LP, Curran M, Bender M, White JWC, Jenk TM, Blunier T, Francey RJ (2013) A revised 1000 year atmospheric d13C-CO₂ record from Law Dome and South Pole, Antarctica. *J Geophys Res Atmos* 118(15):8482–8499. <https://doi.org/10.1002/jgrd.50668>
- Sato M, Hansen JE, McCormick MP, Pollack JB (1993) Stratospheric aerosol optical depths, 1850–1990. *J Geophys Res* 98:22987–22994 Data used available at: http://data.giss.nasa.gov/modelforce/strataer/tau.line_2012.12.txt, last accessed 10 August 2014
- Schonwiese C-D, Walter A et al (2010) Statistical assessments of anthropogenic and natural global climate forcing. An update. *Meteorol Z* 19(1):3–10. <https://doi.org/10.1127/0941-2948/2010/0421>
- Shen M, Piao S, Jeong S-J, Zhou L, Zeng Z, Ciais P, Chen D, Huang M, Jin C-S, Li LZ, Li Y, Myneni RB, Yang K, Zhang G, Zhang Y, Yao

- T (2015) Evaporative cooling over the Tibetan Plateau induced by vegetation growth. *Proc Nat Acad Sci* 112(30):9299–9304
- Stern DI, Kaufmann RK (2014) Anthropogenic and natural causes of climate change. *Clim Chang* 122(1–2):257–269. <https://doi.org/10.1007/s10584-013-1007-x>
- Swanson KL, Tsonis AA (2009) Has the climate recently shifted? *Geophys Res Lett* 36:L06711–L06714
- Taylor KE, Stouffer RJ, Meehl GA (2012) An overview of CMIP5 and the experiment design. *Bull Am Meteorol Soc* 93:485–498 CMIP5 RCP 8.5 data used available at: http://climexp.knmi.nl/data/icmip5_tas_Amon_modmean_rcp45_0-360E_-90-90N_n_+++a.txt, last accessed 3 October 2017
- Toda HY, Yamamoto T (1995) Statistical inference in vector autoregressions with possibly integrated processes. *J Econ* 66(1–2):225–250. [https://doi.org/10.1016/0304-4076\(94\)01616-8](https://doi.org/10.1016/0304-4076(94)01616-8)
- Trenberth KE, Fasullo JT (2010) Tracking Earth's Energy. *Science* 328(5976):316–317
- Trenberth KE, Fasullo JT (2013) An apparent hiatus in global warming? *Earth's Futur* 1(1):19–13. <https://doi.org/10.1002/2013EF000165>
- Triacca U, Pasini A, Attanasio A, Giovannelli A, Lippi M (2014) Clarifying the roles of greenhouse gases and ENSO in recent global warming through their prediction performance. *J Clim* 27(20):7903–7910. <https://doi.org/10.1175/JCLI-D-13-00784.1>
- Verdes PF (2007) Global warming is driven by anthropogenic emissions: a time series analysis approach. *Phys Rev Lett* 90:048501
- Walter MT, Wilks DS, Parlange J-Y, Schneider RL (2004) Increasing evapotranspiration from the conterminous United States. *J Hydrometeorol* 5(3):405–408. [https://doi.org/10.1175/1525-7541\(2004\)005<0405:IEFTCU>2.0.CO;2](https://doi.org/10.1175/1525-7541(2004)005<0405:IEFTCU>2.0.CO;2)
- Yan X-H, Boyer T, Trenberth K, Karl TR, Xie S-P, Nieves V, Tung K-K, Roemmich D (2016) The global warming hiatus: slowdown or redistribution? *Earth's Futur* 4(11):472–482. <https://doi.org/10.1002/2016EF000417>
- Zeng Z, Piao S, Li LZ, Zhou L, Ciais P, Wang T, Li Y, Xu L, Wood EF, Friedlingstein P, Mao J, Estes LD, Myneni RB, Peng S, Shi X, Seneviratne SI, Wang Y (2017) Climate mitigation from vegetation biophysical feedbacks during the past three decades. *Nat Clim Chang* 7(6):432–436. <https://doi.org/10.1038/nclimate3299>
- Zhang K, Kimball JS, Nemani RR, Running SW, Hong Y, Gourley JJ et al (2015) Vegetation greening and climate change promote multidecadal rises of global land evapotranspiration. *Sci Rep* 5: 15956 (9pp)
- Zhu J, Zeng X (2015) Comprehensive study on the influence of evapotranspiration and albedo on surface temperature related to changes in the leaf area index. *Adv Atmos Sci* 32(7):935–942. <https://doi.org/10.1007/s00376-014-4045-z>
- Zhu J, Zeng X (2016) Influences of the interannual variability of vegetation LAI on surface temperature. *Atmos Oceanic Sci Lett* 9(4):292–297. <https://doi.org/10.1080/16742834.2016.1189800>