# 7

# Reconciling Climate Model/Data Discrepancies: The Case of the 'Trees That Didn't Bark'

### Michael E. Mann

One way scientists attempt to validate theoretical models of Earth's climate is to measure their predictions against real-world observations. There is always the danger in this process, however, that the models may be artificially tuned, directly or indirectly, to get key climate attributes right. For example, there may be a tendency for scientists to choose values of uncertain parameters governing both the sensitivity of the climate to increasing greenhouse gas concentrations and the offsetting cooling impacts of industrial aerosol emissions in such a way that models correctly reproduce the observed warming trend of the past century. There is some evidence that such "compensation" may have led to artificially small spreads in the estimated uncertainty ranges in key climate parameters (Andreae et al. 2005).

It is therefore useful to employ a variety of observations from both the present and past, as independent constraints on climate model behavior. This is particularly true of efforts to estimate the equilibrium climate

---

M.E. Mann (✉)
Department of Meteorology, The Pennsylvania State University,
University Park, PA, USA

E.A. Lloyd, E. Winsberg (eds.), *Climate Modelling*,
https://doi.org/10.1007/978-3-319-65058-6_7

sensitivity ("ECS")—a key measure of our impact on the climate that is defined by the eventual warming we expect in response to a doubling of $CO_2$ concentrations relative to pre-industrial levels—levels we will see in a matter of decades under business-as-usual fossil fuel emissions. Various independent lines of evidence that can be brought to bear on the problem of estimating ECS include (Fig. 7.1) the ability of models to reproduce modern-day climatology, the cooling response of the climate to modern volcanic eruptions, the temperature changes during the last glacial

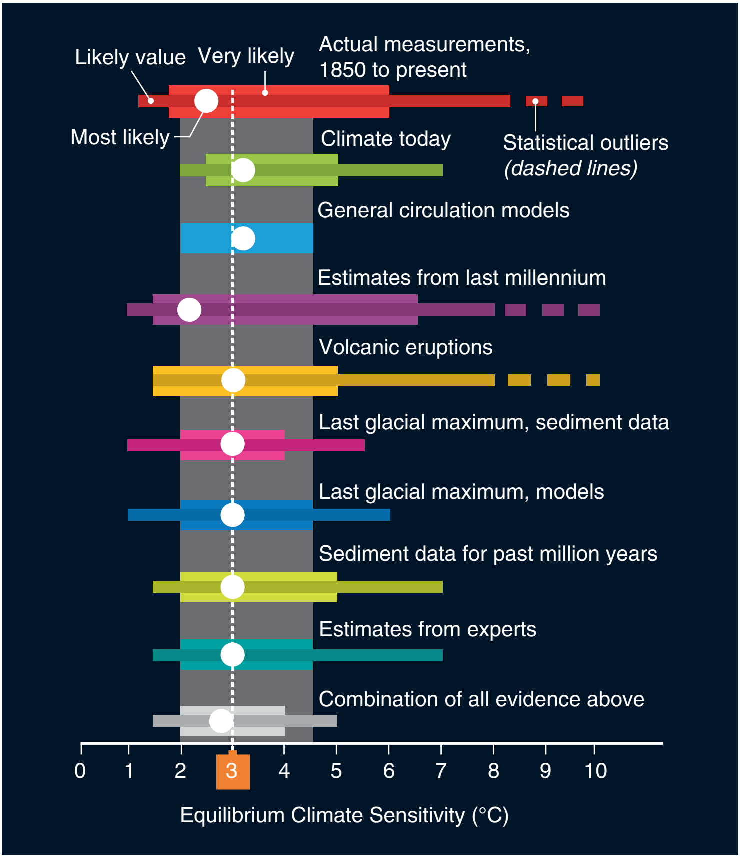

**Fig. 7.1** Estimates of the equilibrium climate sensitivity ("ECS") based on various independent lines of evidence summarized by Knutti and Hegerl (2008) (Modified from Mann 2014 Scientific American)

maximum period, and the changes in temperature associated with geological variations in greenhouse gas concentrations, among others. These different constraints point to a range for ECS of somewhere between 1.5 °C and 5 °C warming, with a mid-range/most likely value close to 3 °C. While most lines of evidence are broadly consistent with each other, there is at least one notable discrepancy: comparisons of simulations of temperature changes over the past millennium with paleoreconstructions of past temperature (the reconstructions are typically based primarily on tree rings, but they are often supplemented by information from corals, ice cores, lake sediments, and other climate "proxy" data). These comparisons (e.g., Hegerl et al. 2006) tend to suggest an ECS value toward the lower end of the range, closer to 2 °C than the mid-range of 3 °C.

This discrepancy is conspicuous enough to demand some level of additional scrutiny. In particular, it is important to consider what is driving the ECS estimate in these comparisons. In the centuries leading up to the industrial area of anthropogenic influence, the primary forcing of climate was from natural changes in radiative forcing associated with factors such as the gradual changes in the distribution of solar insolation associated with millennial-scale earth orbital variations, modest (small fraction of a percent) estimated changes in solar output on multidecadal and centennial timescales, and small but non-negligible natural fluctuations in greenhouse gas concentrations. The cooling effect of stratospheric aerosols (particles such as sulfates which reflect incoming sunlight) associated with intermittent but sizeable explosive volcanic eruptions, however, yields the greatest pre-anthropogenic radiative forcing of climate over the past millennium. The eruption of Tambora in 1815, for example, is estimated to have been twice as large, in terms of radiative forcing (~−4 W/m$^2$), as the largest eruptions recorded in the historical period (e.g., Krakatoa in 1883 and Pinatubo in 1991, both ~−2 W/m$^2$). The tropical eruption of AD 1258 is estimated as somewhere between three and four times as large (between −8 and −12 W/m$^2$). Volcanic forcing turns out to be by far the largest climate forcing in the pre-industrial era of the past millennium (see, e.g., Jansen et al. 2007). Hence, climate models driven by estimated natural radiative forcing changes over the past millennium yield temperature changes that are largely representative of the response to volcanic forcing. If either the model simulations or the paleoreconstructions misestimate the amplitude of this signal, estimates of ECS from those comparisons will accordingly be biased. Indeed, any errors in (a) the volcanic radiative forcing used to drive the climate models, (b) the model-estimated responses to that forcing, (c) the volcanic cooling as estimated by the paleoreconstructions, or (d) any combination thereof, will lead to biased estimates of ECS as inferred from model/data comparisons over the past millennium.

In this article, I summarize evidence that such biases do indeed exist. Specifically, I show that the paleoreconstructions may selectively underestimate the cooling signal associated with large explosive volcanic eruptions of the past millennium. I discuss my previously posed hypothesis (see Mann et al. 2012a) that the underestimation of volcanic cooling arises from a problem specific to the reliance of paleoreconstructions on tree-ring data from treeline-proximal environments, which leads to potential loss of sensitivity to large summer cooling events associated with major explosive volcanic eruptions. This loss of sensitivity potentially results in chronological errors in some subset of tree-ring records used to reconstruct past temperatures.

Requiring that model simulations match the resulting artificially muted volcanic cooling signal may lead to low-biased estimates of ECS. I review the challenges to our hypothesis that have been published, the additional work that we have done in response to those challenges that substantiates the viability of the hypothesis, and a recently proposed test that both proponents and critics of the hypothesis appear to agree would objectively determine whether chronological errors do compromise the integrity of tree-ring-based estimates of past volcanic cooling. Finally, I show that, regardless of the precise reason for the discrepancy, the mismatch between the paleoreconstructed and model-simulated volcanic cooling for a small number of large pre-industrial volcanic eruptions drives the anomalously low apparent values of ECS derived from comparisons of the past millennium. We demonstrate that there are ways to alleviate the impact of these events on the process of estimating ECS from model/data comparisons of the past millennium, and that doing so yields inferences more consistent with other independent lines of evidence.

## 7.1 Hypothesis Posed

Back in 2012, my co-authors and I published an article (Mann et al. 2012a—henceforth “MFR12”) providing a new hypothesis for the enigmatic discrepancy between the tree-ring reconstructed and climate model-predicted magnitude of volcanic cooling in the Northern Hemisphere (NH) mean temperatures during the pre-industrial era of the past millennium. Most notable among the discrepancies is the virtual absence of cooling in tree-ring reconstructions of NH mean temperatures during what ice core and other evidence suggest is the largest explosive volcanic eruption of the past millennium—the AD 1258 eruption (see Emile-Geay et al. 2008 for a review of evidence for a wide-spread global climate impact of this eruption). We suspected that the discrepancy (the trees that didn’t bark) might have something to do with the particular types of tree-ring information that were used to reconstruct past temperatures.

Tree rings are used as proxies for climate because trees create unique rings each year that often reflect the weather conditions that influenced the growing season that year. When seeking to reconstruct past temperature changes, tree-ring researchers (dendroclimatologists) typically seek trees growing at the boreal or alpine tree line, since temperature is most likely to be the limiting climate variable in that environment. This choice may prove problematic under certain conditions however. Trees in such environments are close enough to the summer temperature minimum threshold for growth that a lowering of temperatures by just a couple of degrees during the growing season may yield little or no growth and a consequent loss of sensitivity of tree growth to further cooling. In extreme cases, there may be no growth ring at all. If no ring is formed in a given year, that creates a further complication, introducing an error in the chronology established by counting rings back in time.

We investigated the potential impact of this problem by comparing a tree-growth model driven with climate model simulations of the past millennium with the model-simulated temperatures and tree-ring reconstructions of temperatures. The tree-growth model simulates the dependence of the thickness of growth rings on growing season temperature, based on an empirical growth response curve that accounts for the temperature thresholds governing tree growth (see Mann et al. 2012a for further details). Climate models were driven with estimated natural (volcanic+solar) and anthropogenic forcings over the past millennium. We employed two different climate model simulations: (1) the simulation of the NCAR CSM 1.4 coupled atmosphere-ocean General Circulation Model (GCM) analyzed by Ammann and Wahl (2007) and (2) simulations of a simple Energy Balance Model (EBM). While the GCM provides a more comprehensive and arguably realistic description of the climate system, the computational simplicity of the EBM lends itself to extensive sensitivity tests. As the target for our comparison, we used a state-of-the-art tree-ring-based NH mean temperature reconstruction of D'Arrigo et al. (2006—henceforth "D06"). The reconstruction was based on a composite of tree-ring annual ring width series from boreal and alpine tree-line sites across the NH, and made use of a very conservative ("RCS") tree-ring standardization procedure designed to preserve as much low-frequency climatic information as possible.

Interestingly, the long-term variations indicated by the model simulations compared remarkably well with those documented by the tree-ring reconstruction (Fig. 7.2), showing no obvious sign of the potential biases in the estimated low-frequency temperature variations that have been the focus of some previous work (see e.g., Jones and Mann 2004 for a discussion). Instead, the one glaring inconsistency was in the high-frequency variations, specifically, the cooling response to the largest few tropical eruptions, AD 1258/1259, 1452/1453 and the 1809 + 1815 double pulse of eruptions, which is sharply reduced in the reconstruction relative to the model predictions. Indeed, this was found to be true for any of several different published volcanic forcing series for the past millennium, regardless of the precise geometric scaling used to estimate radiative forcing from volcanic optical depth, and regardless of the precise climate sensitivity assumed.

Following the AD 1258 eruption, the climate model simulations predict a drop of 2 °C, but the tree-ring-based reconstruction shows only about a 0.5 °C cooling. Equally vexing, the cooling in the reconstruction occurs several years late relative to what is predicted by the model. The other large eruptions showed similar discrepancies. An analysis using

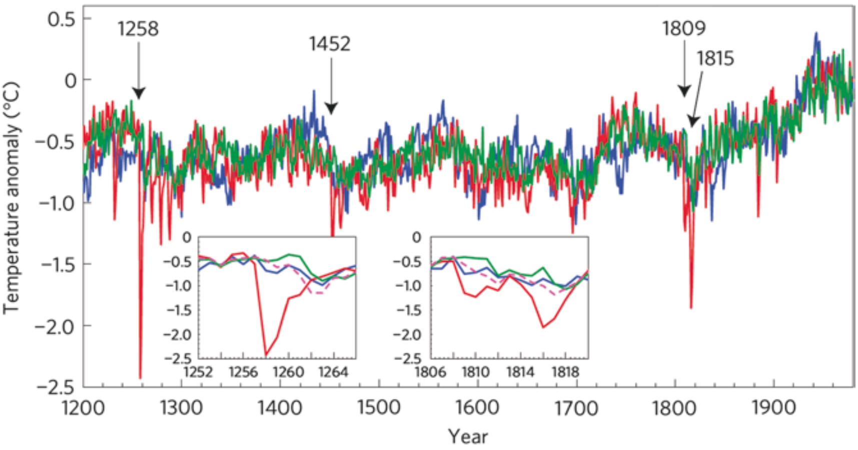

**Fig. 7.2** Shown in the above is the D'Arrigo et al. tree-ring-based NH reconstruction (*blue*) along with the climate model (NCAR CSM 1.4) simulated NH mean temperatures (*red*) and the "simulated tree-ring" NH temperature series based on driving the biological growth model with the climate model-simulated temperatures (*green*). The two insets focus on the response to the AD 1258 and AD 1809+1815 volcanic eruption sequences. Also shown in the insets are the results (dashed magenta) when the volcanic diffuse-light impact is ignored (From Mann et al. (2012a))

synthetic proxy data with spatial sampling density and proxy signal-to-noise ratios equivalent to those of the D06 tree-ring network (see MFR12 for further discussion) suggest that these discrepancies cannot be explained in terms of either the spatial sampling/extent or the intrinsic "noisiness" of the network of proxy records. However, using a tree-growth model that accounts for the temperature growth thresholding effects discussed above, combined with the complicating effects of chronological errors due to potential missing growth rings, explains the observed features remarkably well (see green curve in Fig. 7.2).

The attenuation of the response is produced primarily by the loss of sensitivity to further cooling for eruptions that place growing season temperatures close to the lower threshold for growth. The smearing and delay of the cooling, however, arises from another effect: when growing season lengths approach zero, we assume that no growth ring will be detectable for that year. That means that an age model error of one year will be introduced into the chronology counting back in time. As multiple large eruptions are encountered further back in time, these age model errors accumulate. This factor would lead to a precise chronological error, rather than smearing of the chronology, if all tree-line sites experienced the same cooling. However, stochastic weather variations will lead to differing amounts of cooling for synoptically distinct regions. That means that in any given year, some regions might fall below the "no ring" threshold, while other regions do not. That means that different chronological errors accumulate in synoptically distinct regions of the NH. In forming a hemispheric composite, these errors thus lead to a smearing out of the signal back in time as slightly different age model errors accumulate in the different regions contributing to the composite.

Accounting for this effect, our model accounts not only for the level of attenuation of the signal, but the delayed and smeared out cooling as well. This is particularly striking in comparing the behavior following both the AD 1258 and AD 1809 eruptions (compare the green and blue curves in the insets of the figure). Our model, for example, predicts the magnitude of the reduction of cooling following the eruptions and the delay in the apparent cooling evidence in the tree-ring record (i.e., in AD 1262 rather than AD 1258). We have also included a minor additional effect in these simulations. While volcanic aerosols cause surface cooling due to decreased shortwave radiation at the surface, they also lead to *increased* indirect, scattered light at the surface. Plant growth benefits from indirect sunlight, and past studies show that, e.g., a Pinatubo-sized eruption (roughly −2 W/m^2 radiative forcing) can result in a 30% increase in carbon assimilation by plants. This effect turns out to be relatively small because it is proportional in nature, and thus results in a very small absolute increase when growth is suppressed in the first place by limited growing seasons. However, *not* including this effect results in a slightly worse reproduction (purple dashed curves in the two insets of the figure) of the observed behavior.

As shown in MFR12, the central conclusions discussed above are insensitive to the precise details of the forcing estimates used, the volcanic scaling assumptions made, and the precise assumed climate sensitivity. They are also insensitive to the details of the biological tree-growth model over a reasonable range of model assumptions. Our conclusions would nonetheless soon be challenged by other scientists.

## 7.2 Hypothesis Challenged

The conclusion that tree-ring temperature reconstructions might suffer from age model errors due to missing rings is controversial, and it is important to recognize that it is only a working hypothesis for explaining some enigmatic features of tree-ring temperature reconstructions, more specifically, the *attenuation*, and the increasing (back in time) *delay* and *temporal smearing* in association with the response to past volcanic forcing. Were an equally successful and more parsimonious hypothesis to be provided for these features, we would be the first to concede to this alternative explanation. It was my hope that our hypothesis as presented in MFR12 would encourage a healthy discussion within the paleoclimate community, whether or not it ultimately stands up to additional scrutiny. In particular, it was my hope that dendroclimatologists might, in response to our work, go back and reassess their raw tree-ring chronologies more carefully, and critically assess the extent to which the artifacts we predicted might indeed be present in the underlying tree-ring data.

Initially, however, we instead encountered what might be considered a blanket dismissal of our hypothesis. A group comprised of the majority of leading tree-ring researchers in the United States and Europe published a comment (Anchukaitas et al. 2012—henceforth “A12”) that criticized various aspects of our analysis, but did not provide a plausible alternative explanation for the vexing problem we had identified. Our response (Mann et al. 2012b) appeared along with the comment. A12 suggested that our study represented a fundamental challenge to the validity of large-scale tree-ring-based reconstructions in general, but that is certainly not the case. As we noted in our response, in MFR12 we showed that tree-ring reconstructions effectively capture long-term temperature trends. We were simply questioning the ability of tree-ring width proxies to detect the short-term cooling associated with the largest few volcanic eruptions of the past millennium.

A12 criticized our study for not using more elaborate tree-growth models that include other influences (e.g., precipitation), but this rather misses the point. The fundamental assumption underlying tree-ring-based temperature reconstructions such as those we analyzed is that annual growth at temperature-limited tree-line locations yields an unbiased estimate of temperature changes exclusively. A12 further criticized our tree-growth parameter choices, and suggested that these parameter values yield an unrealistic prediction of missing twentieth-century tree rings. However, as we noted in our response, our analysis predicted no missing tree rings for the twentieth century. Our value of 10 °C as a threshold temperature for growth is at the upper end of the accepted 3–10 °C range, but this choice yields the closest fit to the observed tree-ring response, and we see qualitatively similar results for a lower temperature threshold value.

Addressing A12's criticism over the specifics of our tree-growth model, we demonstrated that similar results are obtained using the simplest possible (growing degree day) model, which involves a linear growth response above a threshold temperature. Using that model, we showed that the underestimation of volcanic cooling by tree rings is substantial for threshold values spanning the entire upper half of the 3–10 °C range, even using a conservative assumption of what constitutes a missing ring (a growing season of less than one week). Including the effect of increased diffuse light caused by volcanic aerosols—an important factor neglected by A12—leads to better agreement between our growth model and existing tree-ring reconstructions. For growth-model assumptions substantially different from those we adopted, however, the effect produces offsetting and spurious warming responses in the first few years following an eruption (see Mann et al. 2012a).

A12 sought to reconcile the lack of the expected cooling response to the AD 1258/1259 in the D06 tree-ring reconstruction by arguing that the radiative forcing might have been smaller than generally assumed. However, as we showed in MFR12, our findings are robust with respect to which of the various published volcanic forcing reconstructions or volcanic scaling assumptions are used. Moreover, changing the estimated radiative forcing associated with the AD 1258/1259 eruption would not explain other problematic features in the tree-ring reconstructed response. Our analysis, by contrast, provides a plausible explanation for why cooling is observed four years later than expected, and is greatly diminished in magnitude. Our hypothesis also explains a similar discrepancy between the tree-ring reconstruction and the cooling associated with the 1815

Tambora eruption. Importantly, this latter eruption is constrained by observational surface temperature data (Rohde et al. 2013). These data (a) confirm the model-estimated cooling and (b) contradict the muted/absent cooling in the tree-ring estimates.

Perhaps most importantly, we did not argue, as A12 seemed to suggest, that tree rings are uniformly recording the wrong year of the eruption in a way that can be diagnosed just by looking at composite series. Instead, we suggest that sufficiently many individual tree-ring records within the composites are likely to have dating errors due to potential missing/undetected rings following the largest volcanic eruptions that the cooling signal is muted and smeared in the large-scale averages.

One argument against the specific conclusion of missing growth rings is that trees are carefully cross-dated when forming regional chronologies, and this precludes the possibility of chronological errors. That, however, assumes that there are at least some trees within a particular region that will not suffer a missing ring during the years where our model predicts it. Yet our prediction is that *all trees* within a region of synoptic or lesser scale where growing season temperatures lie below the growth threshold will experience a missing ring. Thus, cross-dating within that region, regardless of how careful, cannot resolve the lost chronological information.

As we noted in our response, it should be possible to further investigate this hypothesis through a careful analysis of the detailed patterns of response to the largest eruptions among individual tree-ring chronologies distributed over the globe.

## 7.3 Additional Evidence

As we have seen, subsequent to the publication of MFR12 there was a vigorous debate about the viability of our hypothesis for the muted, delayed volcanic cooling signal in tree-ring composite-based reconstructions of hemispheric temperature change. Chief among the criticisms is that our hypothesis was based entirely on theoretical modeling, and that we had provided no empirical evidence for the claim of missing tree rings—an important component of our mechanism for the underestimation, smearing and delay of the volcanic cooling signal in tree-ring-based temperature reconstructions. In subsequent work (Mann et al. 2012b), we attempted to provide precisely that evidence.

It is necessarily more challenging to prove that something is missing than to prove it is present. Though local cross-dating of trees can be used to identify missing rings in individual cores contributing to local chronologies developed from nearby trees, it cannot reliably identify a coherent large-scale pattern of missing rings across an entire climatic region experiencing sub-growth limit summer temperatures, as MFR12 predicts to be the case following the largest few tropical volcanic eruptions. A more nuanced approach is required to detect the influence of missing rings.

We instead attempt to account for the effects of missing rings in some subset of the underlying tree-ring chronologies. We employed the original tree-ring data used by D06, which consists of a maximum of 66 distinct site chronologies representing 19 different regions back to 1686, decreasing to eight regions back to AD 1190 (we used the conventionally standardized tree ring series of D06, but broadly similar results were obtained using the alternative "RCS" standardization; see Mann et al. 2013). We performed Monte Carlo simulations using the MFR12 estimates of the timing and probabilities for a missing ring in a given year, yielding alternative versions of the D06 tree-ring series consistent with estimated chronological (age model) errors. Using these surrogate tree-ring series, we generated an ensemble of alternative *regional* composites consistent with estimated tree-age model uncertainties (e.g., the chance of a given region missing a ring in any particular realization is 90% in AD 1258, and 55% in AD 1816 as prescribed by MFR12—note that our net estimated age model errors amount to <1%, i.e., no more than 6 years out of 700+). This procedure was used to generate a large ensemble of surrogate hemispheric temperature reconstructions based on averaging the surrogate regional series emulating the procedures of D06 (see Mann et al. 2013 for further details). In principle, some subset of these surrogates should correct for the age model errors (i.e., missing rings).

As shown in Fig. 7.3, some of the surrogate reconstructions indeed suggest significantly greater cooling in association with the major volcanic eruptions. For the AD 1258 eruption, a large number of Monte Carlo

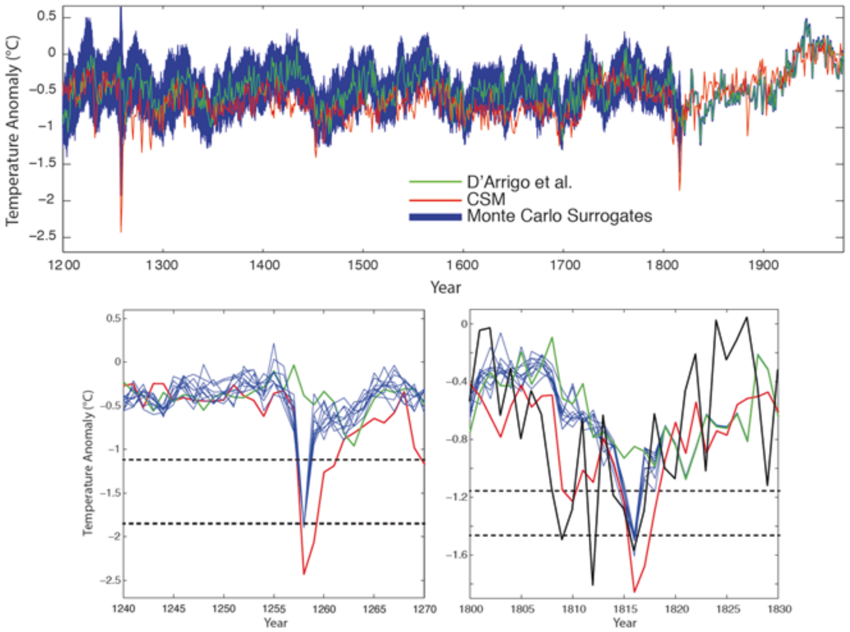

**Fig. 7.3** Ensemble of hemispheric tree-ring temperature reconstructions derived from available regional tree-ring composites resampled to account for predicted age model errors. Shown are the raw composite based on the D'Arrigo et al. (2006) tree-ring data (*green*), Monte Carlo surrogate reconstructions (8000 in total—blue curves), and GCM simulation (*red*). Insets: Expanded views of the response to the AD 1258/1259 and AD 1815 eruptions responses showing the 10 coldest surrogates (*blue*) for each eruptions and the 2 and 4 sigma significance thresholds for cooling (*dashed black*). Shown also for AD 1815 eruption is the recently back-extended instrumental NH land temperature record of Rohde et al. (2013) (*black*). Centering of all series is based on a 1961–1990 modern base period (From Mann et al. (2013))

surrogates point toward a distinct ~2 °C cooling in AD 1258 (lacking the enigmatic delayed and reduced 1260–62 cooling signal seen in the raw reconstruction). The increased AD 1258 cooling and disappearance of (likely spurious) AD 1260–62 cooling is seen to arise from a realignment of much larger cooling signals that are present in individual tree-ring series but interfere destructively before they are brought into alignment (see Mann et al. 2013). The year AD 1816 is far more consistent with its moniker as the "Year Without a Summer," with surrogates showing cooling of up to ~1.6 °C. The amplified cooling is not only far more consistent with the model-predicted cooling, but agrees far better with the available instrumental temperature record. These enhanced cooling responses that arise from permuting the tree-ring data within estimated age model errors are highly significant relative to the null hypothesis of chance occurrence due to random sampling variations from the Monte Carlo procedure (see Mann et al. 2013).

We thus argue that the missing rings in regional tree-ring temperature composites as hypothesized in MFR12 are not only plausible from a theoretical perspective, but appear to be detectable in the actual underlying regional tree-ring series and resulting hemispheric composites. Attempts to correct for the estimated chronological errors yield far greater post-volcanic cooling responses that agree with model predictions.

## 7.4 Wider Implications

I return now to the issue of why a seemingly technical and mundane matter involving tree rings and volcanic eruptions actually matters. As noted earlier, the apparent weak response of surface temperatures to the few largest eruptions of the past millennium as inferred from proxy temperature reconstructions is what drives estimates of relatively low ECS as derived from proxy reconstructions based either entirely or substantially upon tree-ring data (Hegerl et al. 2006). Hegerl et al. (2006) for example used comparisons during the pre-industrial period of EBM simulations and proxy temperature reconstructions based entirely or partially on tree-ring data to estimate ECS. Hegerl et al. (2006) ended up arguing for a substantially lower 5–95% range of ECS (1.5–6.2 °C) than is evident from other lines of evidence (see Fig. 7.1). As the primary radiative forcing during the pre-industrial period is from volcanic forcing, their conclusions were leveraged by the muted apparent response to very large past volcanic eruptions. If that muted response is an artifact, as our work suggests it to be, the resulting estimates of ECS are almost certainly downwardly biased. Moreover, this one potentially biased constraint on ECS (central value about 2.1 °C—see Fig. 7.1) is enough of an outlier (nearly all other lines of evidence point to an ECS value at or slightly above 3.0° C) that it ends up downwardly biasing the “combined” estimate of ECS (Fig. 7.1), taking it from 3.2 °C to roughly 2.8 °C, a nontrivial lowering of nearly 0.5 °C. Our findings therefore suggest that prevailing estimates of ECS from combinations of various lines of evidence (e.g., Knutti and Hegerl 2008) have likely underestimated the true climate sensitivity.

In Mann et al. (2013), we assessed the impact that the underestimation of volcanic cooling from tree-ring reconstructions as estimated by MFR12 would have on inferred values of ECS. Our analysis employed EBM simulations where the actual value of ECS is precisely known (it was set to the canonical mid-range value of 3 °C) and is then estimated using the simulated tree-ring response. We found that the truncation of volcanic cooling alone led to a decrease in apparent ECS from 3.0 °C to 1.7 °C in simulations of the pre-industrial interval AD 1200–1849. That calculation did not take into account the additional degradation by estimated chronological errors. When chronological errors are accounted for, the estimated ECS value drops to less than 1.0 °C—similar to the ECS value estimated using the D’Arrigo et al. (2006) tree-ring reconstruction. Using a later period AD 1300–1849, which eliminates the influence of the AD 1258 eruption, leads to a lesser but still large impact on ECS values (ECS ~2.0 °C without considering chronological errors, and ECS ~1.0 °C with chronological errors accounted for). These estimates pertain only to tree-ring-based temperature reconstructions. Most proxy-based reconstructions of past temperature instead use a mix of proxy data, including corals, ice cores, sediments, and other types of proxy information. For such reconstructions, we might expect a smaller underestimation of volcanic cooling than estimated for tree-ring only temperature reconstructions, and potentially a smaller bias in ECS estimates derived from the reconstructions. However, even if the estimated impact is reduced by a factor of two or three, it is large enough to explain the discrepancy between “last millennium” estimate of ECS and ECS estimates derived from the remaining lines of evidence (Fig. 7.1).

It is reasonable to ask whether our principal conclusions hold up even if the specifics of our hypothesis about the underestimation of volcanic cooling by tree-ring temperature reconstructions do not. We addressed that matter in additional work (Schurer et al. 2013) using an alternative approach. We employed a method wherein a large ensemble of state-of-the-art climate model simulations of the past millennium—the Coupled Model Intercomparison Project 5 (CMIP5) "past millennium" simulations—were used to estimate the "fingerprints" of the various natural radiative forcings of climate which include solar irradiance, Earth orbital changes, natural variations in greenhouse gas concentrations, and explosive volcanic eruptions. The amplitudes of those fingerprints were then estimated (via total least squares regression) from nine different proxy-based reconstructions of NH mean temperature spanning all or most of the past millennium. The amplitudes estimated from the paleoclimate reconstructions were then compared against the model-predicted amplitudes. The ratio of the two ("β") measures whether the reconstruction indicates a greater (β > 1), comparable (β ~ 1), or lesser (β < 1) amplitude than predicted by the models.

The procedure was performed using a variety of sub-intervals of the period 851–1950 as well as the full interval and the full pre-industrial interval AD 851–1850. With only one exception (a controversial reconstruction that exhibits far greater variability than all others), the reconstructions yielded estimates of β that are systematically less than unity (i.e., the entire uncertainty range for β lies below unity). However, if the few largest eruptions (which include the AD 1258, the AD 1453 Kuwae, and 1815 Tambora eruptions) are simply masked from the analysis (so that the analysis is based on the response to all other radiative forcing, i.e., moderate eruptions, solar irradiance changes, greenhouse gas concentrations, and Earth-Orbital changes), and the procedure is repeated, then remarkably, most of the β values are consistent with a value of unity within the associated error bars. In other words, if the largest eruptions of the past millennium are included in the analysis, the reconstructions indicate a response to forcing that is systematically smaller than predicted by the models. Yet if just that handful of eruptions is masked out, the reconstructions indicate a response that is consistent with the model simulations.

It is important to recognize that there are a number of sources of potential uncertainty and bias that contribute to these model/data comparisons in addition to potential biases in the proxy reconstructions. These include uncertainties or biases in the estimates of radiative forcings, and uncertainties or biases in the models' response to radiative forcings. This latter uncertainty/bias is tied in part to the uncertainty in the associated ECS, though there are also potential uncertainties and/or biases in climate responses that are specific to the way particular forcings are represented in the models. For example, in the case of volcanic radiative forcing there is some uncertainty in how volcanic aerosol size distributions are represented (see, e.g., MFR12; Mann et al. 2012b, 2013). Any combination of these uncertainties or biases can contribute to the model/data misfit.

That notwithstanding, the simplest interpretation of the above findings is that the climate models, including the ECS values that characterize their response to radiative forcing, are consistent with the paleoreconstructions if the response to the few largest volcanic eruptions are masked out in the analysis. That implies that the reduced apparent response to forcing in the reconstructions overall arises entirely from the discrepancy between the apparent and predicted response to volcanic radiative forcing. That finding, in turn, is consistent with the proposition that it is the specific discrepancy between the model-predicted and proxy reconstruction-estimated response to the few largest volcanic eruptions of the past millennium that leads to anomalously low values of apparent ECS in studies using paleoreconstructions of the past millennium such as Hegerl et al. (2006). That conclusion does not establish that the source of this discrepancy is the tree-growth saturation mechanism proposed by MFR12, but it provides independent support for the existence of some source of bias that is limited to the apparent response of the climate to the few largest volcanic eruptions of the past millennium.

## 7.5 The Gauntlet Is Laid Down

In a recent comment, Büntgen et al. (2014) provide a potential way forward to resolve definitively whether or not the specific tree-ring age model errors predicted by MFR12 (and further supported by Mann et al. 2013) can be established in the actual data. The authors demonstrate the existence of a distinct radiocarbon event during AD 774–775, which has consistently been recorded by trees in disparate locations including Japan, Germany, and the Alps, thus establishing that the dating of these trees is consistent and accurate.

Our hypothesis, as presented in MFR12, is that some trees growing near their thermal limits, as is the case with many trees selected for paleotemperature reconstructions which lie at the boreal or alpine tree line, can fail to produce an annual ring during unusually cold growing seasons following particularly large volcanic eruptions. The missing ring causes the year preceding the eruption to masquerade as the eruption year. Thus, the resulting chronology would not record the effects of the eruption because the ring from that year is missing, and all previous years in the chronology are shifted forward in time by the number of missing rings. This means that, even if the tree produced a growth ring following an older eruption, that ring would appear in the wrong year. The radiocarbon event of AD 774–775 provides a globally synchronous signature that ought to provide a unique, independent time marker that can be used to test our hypothesis.

As described by Rutherford and Mann (2014), we can make very specific predictions based on our hypothesis that can be tested using the radiocarbon event and existing tree-ring chronologies. With regard to the Alps series, the results from Mann et al. (2013) predict that there will be no missing rings in this region. The D'Arrigo et al. (2006) Alps regional series begins in AD 1350, and was included in our analysis of the climate response to the 1815/16 Tambora eruption sequence. Our "best match" surrogate ensembles for this eruption (Fig. 7.2 of Mann et al. 2013) use the Alps series on its original time scale. Our results are therefore consistent with the Büntgen et al. (2014) finding that there is no age model error with this series.

Of the 19 regional series used in D'Arrigo et al. (2006) and Mann et al. (2013), only three (Coastal Alaska, Tornestraesk, and Taymir) begin before AD 774 and can thus be directly tested using the AD 774/775 radiocarbon event. The results from Mann et al. (2013) predict the following minimum offsets for the event in these three series: the Coastal Alaska series should be four years too young, the Tornestraesk series should be one to five years too young, and the Taymir series should be one year too young (Fig. 7.4). In addition, the Mann et al. (2013) results predict that the "Icefields" series dates correctly, but as it begins in AD

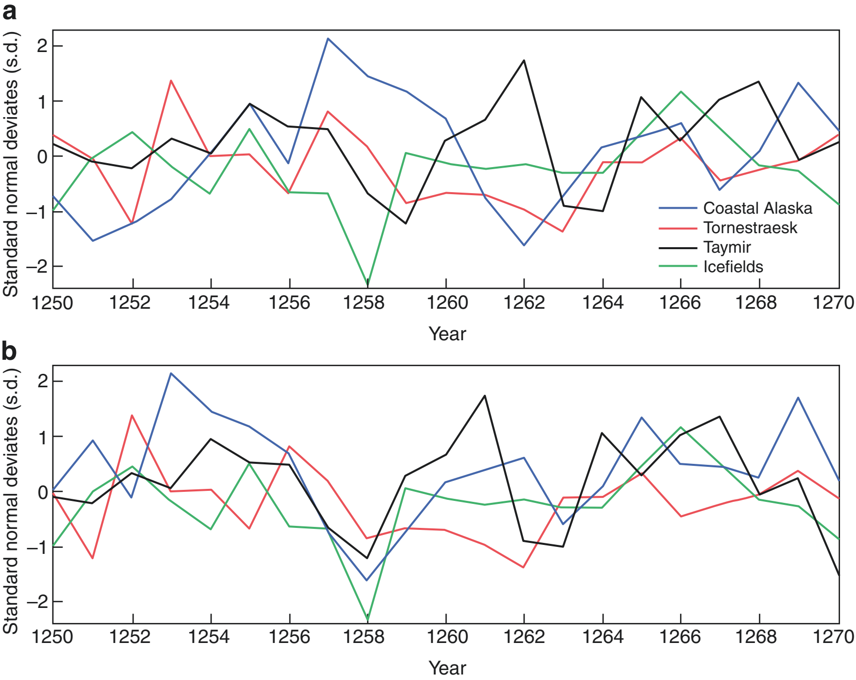

**Fig. 7.4** Tree-ring records across the AD1258 eruption. The three D'Arrigo et al. regional series that begin before AD774 (Coastal Alaska, Tornestraesk, and Taymir), along with the Icefields series for reference, are shown on their original time scale (**a**) and age-adjusted (**b**) in a way consistent with our hypothesis. The Icefields series is unaltered, the Coastal Alaska series is shifted four-years older (~0.6%), and the Tornestraesk and Taymir series are both shifted one year older (~0.1%) (From Rutherford and Mann (2014))

918, its age model cannot be validated with the AD 774/775 radiocarbon event.

Thus, the MFR12 hypothesis that missing growth rings due to unusually cold summers at tree line following the few largest volcanic eruptions of the past millennium is now testable. It will be up to dendroclimatologists and/or dendrochronologists to go back and examine the specific chronologies mentioned above which we predict to contain missing rings and check, using the AD 774/775 radiocarbon date to assess whether there are any age model errors in these chronologies.

## 7.6 Closing Thoughts

As alluded to by the title of this piece, what led to the hypothesis explored in this article isn't what was evident in paleoclimatic reconstructions of the past millennium, but instead, what *wasn't* evident. Much as with Sherlock Holmes and the "curious incident of the dog in the night-time [that didn't bark]," it is sometimes those things that we inexplicably can't see in the data that points to gaps in knowledge or understanding.

The scientific investigations summarized in this article grew out of an enigmatic observation that had bothered me for some time: paleoclimate reconstructions based partly or entirely on tree-ring data fail to show any evidence of large-scale cooling following what various lines of evidence indicate was the largest (from a radiative forcing standpoint) eruption of the past millennium, the AD 1258 tropical eruption. More generally, we found that the paleoclimate reconstructions indicate systematically less cooling following the largest volcanic eruptions than is predicted by climate models.

We are able to reproduce these observations based on simulations using a model of tree growth forced with climate model simulations of temperature over the past millennium. For values of the relevant parameters (i.e., the minimum temperature threshold for tree growth) within the cited range, we are able to reproduce the muted, delayed, and smeared cooling response to very large volcanic eruptions seen in tree-ring-based temperature reconstructions. These features are seen, in the simulations, to be an artifact of a maximum threshold on the cooling that can be recorded by tree-line-proximal trees, combined with the introduction of chronological age model errors in some subset of chronologies associated with a lack of growth during the growing season. The chronological errors accumulate differentially in different regions, leading to a smearing out of temperature signals in hemispheric composites that increases back in time.

While other researchers have raised various objections with our hypothesis and findings, we have been able to provide independent, indirect evidence that missing rings/chronological errors are indeed present in some subset of tree-ring chronologies based on Monte Carlo simulations that show that much larger volcanic cooling signals can be found in hemispheric composites when the estimated age model errors are taken into account. Ours is just one potential hypothesis for the model/data discrepancies in question, and as discussed in this article, at least one aspect of our hypothesis—the existence of chronological errors in some subset of tree-ring chronologies—can now potentially be tested based on the radiocarbon event of AD 774/775. We await with great interest the results of these tests.

Whether or not our specific hypothesis is correct, however, we have shown that some of our key conclusions appear to be robust. In particular, there is very compelling evidence that the discrepancies between model simulations and paleoclimate reconstructions over the past millennium appear to be associated almost exclusively with the response to the few largest volcanic eruptions of the past millennium. It is clear that if one simply masks these eruptions from any model/data comparisons, then the model simulations and reconstructions are consistent. A corollary of this conclusion is that previous studies arguing for relatively low (~2 °C) ECS based on model/data comparisons over the past millennium likely suffer from a bias related to the underestimation of volcanic cooling in the reconstructions. That would explain why this one line of evidence for ECS gives a substantially lower estimate of ECS than essentially every other line of evidence. Finally, these findings provide additional support for the contention that the most likely value of ECS is in the range of 3.0 °C, and that previous assessments that consider, even partly, evidence from the last millennium, may have underestimated ECS. This conclusion is hardly a trivial one, as it provides support for the contention that the climate system is substantially sensitive to carbon emissions, and that business-as-usual fossil fuel burning may have a profound impact on Earth's climate.

## References

Ammann, Caspar M., and Eugene R. Wahl. 2007. The Importance of the Geophysical Context in Statistical Evaluations of Climate Reconstruction Procedures. *Climatic Change* 85 (1–2): 71–88. https://doi.org/10.1007/s10584-007-9276-x.

Anchukaitis, Kevin J., Petra Breitenmoser, Keith R. Briffa, Agata Buchwal, Ulf Büntgen, Edward R. Cook, Rosanne D. D'Arrigo, et al. 2012. Tree Rings and Volcanic Cooling. *Nature Geoscience* 5: 836–837. https://doi.org/10.1038/ngeo1645.

Andreae, Meinrat O., Chris D. Jones, and Peter M. Cox. 2005. Strong Present-Day Aerosol Cooling Implies a Hot Future. *Nature* 435 (7046): 1187–1191.

Büntgen, Ulf, Lukas Wacker, Kurt Nicolussi, Michael Sigl, Dominik Güttler, Willy Tegel, Paul J. Krusic, and Jan Esper. 2014. Extraterrestrial Confirmation of Tree-Ring Dating. *Nature Climate Change* 4: 404–405. https://doi.org/10.1038/nclimate2240.

D'Arrigo, Rosanne, Rob Wilson, and Gordon Jacoby. 2006. On the Long-Term Context for Late Twentieth Century Warming. *Journal of Geophysical Research (Atmospheres)* 111: D03103. https://doi.org/10.1029/2005JD006352.

Emile-Geay, J., R. Seager, M.A. Cane, E.C. Cook, and G.J. Jaug. 2008. Volcanoes and ENSO Over the Past Millennium. *Journal of Climate* 21: 3134–3148.

Hegerl, Gabriele C., Thomas J. Crowley, William T. Hyde, and David J. Frame. 2006. Climate Sensitivity Constrained by Temperature Reconstructions over the Past Seven Centuries. *Nature* 440 (7087): 1029–1032. https://doi.org/10.1038/nature04679.

Jansen, Eystein, Jonathan Overpeck, and Keith R. Briffa. 2007. Paleoclimate. In *Climate Change 2007: The Physical Science Basis.* Working Contribution of Working Group 1 to the Fourth Assesment Report of the Intergovernmental Panel on Climate Change. Cambridge/New York: Cambridge University Press.

Jones, P.D., and M.E. Mann. 2004. Climate over Past Millennia. *Reviews of Geophysics* 42 (2): RG2002. https://doi.org/10.1029/2003RG000143.

Knutti, Reto, and Gabriele C. Hegerl. 2008. The Equilibrium Sensitivity of the Earth's Temperature to Radiation Changes. *Nature Geoscience* 1: 735–743. https://doi.org/10.1038/ngeo337.

Mann, Michael E. 2014. False Hope: The Rate of Global Temperature Rise May Have Hit a Plateau, but a Climate Crisis Still Looms in the Near Future. *Scientific American* 310: 78–81. https://doi.org/10.1038/scientificamerican0414-78.

Mann, Michael E., Jose D. Fuentes, and Scott Rutherford. 2012a. Underestimation of Volcanic Cooling in Tree-Ring-Based Reconstructions of Hemispheric Temperatures. *Nature Geoscience* 5: 202–205. https://doi.org/10.1038/ngeo1394.

———. 2012b. Reply to 'Tree-Rings and Volcanic Cooling'. *Nature Geoscience* 5 (12): 837–838.

Mann, Michael E., Scott Rutherford, Andrew Schurer, Simon F.B. Tett, and Jose D. Fuentes. 2013. Discrepancies Between the Modeled and Proxy-Reconstructed Response to Volcanic Forcing over the Past Millennium: Implications and Possible Mechanisms. *Journal of Geophysical Research: Atmospheres* 118 (14): 7617–7627. https://doi.org/10.1002/jgrd.50609.

Rohde, R., R.A. Muller, R. Jacobsen, E. Muller, S. Perlmutter, et al. 2013. A New Estimate of the Average Earth Surface Land Temperature Spanning 1753–2011. *Geoinfor Geostat: An Overview* 1: 1. http://static.berkeleyearth.org/papers/Results:Paper-Berkeley-Earth.pdf

Rutherford, Scott, and Michael E. Mann. 2014. Missing Tree Rings and the AD 774-775 Radiocarbon Event. *Nature Climate Change* 26: 648–649.

Schurer, Andrew P., Gabriele C. Hegerl, Michael E. Mann, Simon F.B. Tett, and Steven J. Phipps. 2013. Separating Forced from Chaotic Climate Variability over the Past Millennium. *Journal of Climate* 26 (18): 6954–6973. https://doi.org/10.1175/JCLI-D-12-00826.1.