

Foundations of attribution in climate-change science

Article

Published Version

Creative Commons: Attribution 4.0 (CC-BY)

Open Access

Lloyd, E. A. and Shepherd, T. G. ORCID:
<https://orcid.org/0000-0002-6631-9968> (2023) Foundations of attribution in climate-change science. Environmental Research Climate, 2 (3). 034014. ISSN 2752-5295 doi: 10.1088/2752-5295/aceea1 Available at
<https://centaur.reading.ac.uk/112907/>

It is advisable to refer to the publisher's version if you intend to cite from the work. See [Guidance on citing](#).

To link to this article DOI: <http://dx.doi.org/10.1088/2752-5295/aceea1>

Publisher: Institute of Physics

All outputs in CentAUR are protected by Intellectual Property Rights law, including copyright law. Copyright and IPR is retained by the creators or other copyright holders. Terms and conditions for use of this material are defined in the [End User Agreement](#).

www.reading.ac.uk/centaur

CentAUR

Central Archive at the University of Reading

Reading's research outputs online

PAPER • OPEN ACCESS

Foundations of attribution in climate-change science

To cite this article: Elisabeth A Lloyd and Theodore G Shepherd 2023 *Environ. Res.: Climate* **2** 035014

View the [article online](#) for updates and enhancements.

You may also like

- [X-RAY EMISSION FROM OPTICALLY SELECTED RADIO-INTERMEDIATE AND RADIO-LOUD QUASARS](#)

B. P. Miller, W. N. Brandt, D. P. Schneider et al.

- [A CHANDRA SURVEY OF THE X-RAY PROPERTIES OF BROAD ABSORPTION LINE RADIO-LOUD QUASARS](#)

B. P. Miller, W. N. Brandt, R. R. Gibson et al.

- [The composite X-ray spectra of radio-loud and radio-quiet SDSS quasars](#)

Min-Hua Zhou and Min-Feng Gu

ENVIRONMENTAL RESEARCH CLIMATE



PAPER

Foundations of attribution in climate-change science

OPEN ACCESS

RECEIVED

17 March 2023

REVISED

4 August 2023

ACCEPTED FOR PUBLICATION

9 August 2023

PUBLISHED

25 August 2023

Original content from this work may be used under the terms of the [Creative Commons Attribution 4.0 licence](#).

Any further distribution of this work must maintain attribution to the author(s) and the title of the work, journal citation and DOI.

Elisabeth A Lloyd¹ and Theodore G Shepherd^{2,3,*}

¹ Department of History and Philosophy of Science and Medicine, Indiana University, Bloomington, IN 47401, United States of America

² Department of Meteorology, University of Reading, Reading RG6 6BB, United Kingdom

³ Jülich Supercomputing Centre, Forschungszentrum Jülich, 52428 Jülich, Germany

* Author to whom any correspondence should be addressed.

E-mail: theodore.shepherd@reading.ac.uk

Keywords: attribution, causality, regional climate change, storylines

Abstract

Attribution—the explanation of an observed change in terms of multiple causal factors—is the cornerstone of climate-change science. For anthropogenic climate change (ACC), the central causal factor is evidently ACC itself, and one of the primary tools used to reveal ACC is aggregation, or grouping together, of data, e.g. global mean surface temperature. Whilst this approach has served climate-change science well, the landscape is changing rapidly. First, there is an increasing focus on regional or local aspects of climate change, and on singular or unprecedented events, which require varying degrees of disaggregation. Relatedly, climate change is increasingly apparent in observations at the local scale, which is challenging the primacy of climate model simulations. Finally, the explosion of climate data is leading to more phenomena-laden methodologies such as machine learning. All this demands a re-think of how attribution is performed and causal explanations are constructed. Here we use Lloyd's 'Logic of Research Questions' framework to show how the way in which the attribution question is framed can strongly constrain its possible and responsive answers. To address the Research Question 'What was the effect of ACC on X?' (RQ1), scientists generally consider the question 'What were the causal factors leading to X, and was ACC among them?'. If the causal factors include only external forcing and internal variability (RQ2), then answering RQ2 also answers RQ1. However, this unconditional attribution is not always possible. In such cases, allowing the causal factors to include elements of the climate system itself (RQ3)—the conditional, storyline approach—is shown to allow for a wider range of possible and responsive answers than RQ2, including that of singular causation. This flexibility is important when uncertainties are high. As a result, the conditional RQ3 mitigates against the sort of epistemic injustice that can arise from the unconditional RQ2.

1. Introduction

There is currently much talk of 'useable' or 'actionable' climate-change science, especially in the adaptation context (Coen 2021, Jebeile and Roussos 2023). To be effective, action to address the impacts of climate change needs to be anchored in attribution (in the sense of a causal explanation), otherwise the action could be misdirected. By helping to understand the historical record, attribution not only is relevant for liability (Burger *et al* 2020, Lloyd and Shepherd 2021, Wentz *et al* 2023) but also provides the basis for understanding current and future climate risk and undertaking action to mitigate that risk. As a result, the attribution statements made by the Intergovernmental Panel on Climate Change (IPCC) are both keenly awaited and the subject of intense scrutiny. In particular, the role of anthropogenic climate change (ACC) in observed changes in impactful hazards is a central focus of attribution science.

Physical climate science, especially as it is used to inform regional climate risk and adaptation planning, is fast becoming a data-driven field: climate model data output is becoming increasingly heterogeneous and voluminous; and climate change is increasingly apparent not just in the traditional global metrics (e.g. global

mean surface temperature, global mean sea level) but also in observed data at the regional scale (Hegglin *et al* 2022), including extreme events (see Chapter 11 of the IPCC Working Group (WG) I Sixth Assessment Report (AR6), Seneviratne *et al* (2021)). There are increasing calls for artificial intelligence and machine learning (AI/ML) methods to be applied to make sense of all this data (Arribas *et al* 2022), and climate service providers are already using such methods (e.g. Bouwer 2022).

Although physical climate science has long been a data-centric field, it has largely been data-centric in what Lloyd *et al* (2022) call a *phenomena-agnostic* way: data generally consist of time series of basic variables defined on a (possibly irregular) spatial grid, and users of the data are free to configure the data in whatever way they want. This contrasts with data-centric fields such as biology, which need to use *phenomena-laden* data-handling methodologies—that is, where data are packaged into recognized entities, such as categories, theoretical entities, and biological phenomena—in order to make the data interpretable by users. This has potential pitfalls, because choices made by those constructing the data repositories can have unanticipated and sometimes exclusionary downstream consequences (Leonelli 2016). For example, those choices could restrict the exchange of data to certain kinds of researchers, thereby marginalizing some groups and rewarding others (Lloyd *et al* 2022).

Consider, for example, the field of animal behavior, and the filming or videotaping of the behaviors of the animals in question as phenomena. Such video data is transmissible as a whole, but needs to be accompanied by a sophisticated narration of events to be fully understood. Not all educational institutions or locales will have access to such sophisticated narrators whom their students could question and learn from, learning how to categorize and catalogue behaviors, thus disadvantaging those far from the privileged centers of such research in the Global North. Similar considerations would apply to animations of selected climate phenomena. On the other hand, standardized climate data repositories with predefined variable formats might not allow for the incorporation of traditional knowledge, so phenomena-laden approaches also offer advantages in this respect.

As physical climate science becomes more data-driven from the explosion of climate data, the need to combine different lines of evidence (regional and global climate models of varying spatial resolutions, observations, and theory) will increasingly require the use of phenomena-laden data-handling methodologies, which are inherent in AI/ML methods. This changing landscape demands a rethink of how climate-change attribution is performed and how causal explanations are constructed, in order to avoid falling into some of the pitfalls that can arise with phenomena-laden methodologies. To avoid the situation of epistemic injustice—the unfair, irresponsible, and inequitable sourcing of knowledge—discussed by Leonelli (2016) and Lloyd *et al* (2022), people must be able to make sense of their own local situation, using the available combination of climate model output and observational data, in collaboration with domain experts on the phenomena of interest for their region (Rodrigues and Shepherd 2022).

The premise of this paper is that to achieve this goal it is necessary to return to the foundations of attribution in climate-change science. Particular scientific practices which may have been appropriate in one context can become normative within a discipline, after which researchers tend to forget about the original reasons for those practices and apply them inappropriately in another context (e.g. Arnet 2019). To avoid such problems, it is important to pay close attention to how research questions are posed, because that constrains the range of possible and responsive answers (Lloyd 2015). Sometimes separate scientific communities appear to be asking the same question, using the same terms and language, but because of their differing background and assumptions, they are actually asking very distinct questions. In such cases, Lloyd's (2015) 'Logic of Research Questions' (LRQ) analytic tool can be used to help unpack and highlight the differences between relevant theoretical approaches and frameworks. The LRQ tool is applicable to any scientific field that experiences controversy about methods and inference, and has been applied to controversies in neuroscience, climate science, and paleontology.

A simple example comes to us from evolutionary biology (Lloyd 2021). There was a 50 year-long debate concerning what appeared to be a flaw with the 'adaptationist program' taken up by many sociobiologists, behavioral ecologists and evolutionary psychologists, among others, but it turned out to be difficult to pin down. With LRQ the flaw became obvious. The adaptationist program's sole RQ, '*What is the function of this trait?*', has only a single type of possible and responsive answer: 'The function of this trait is A; the function of this trait is B...' The chief competitor of the adaptationist program, followed by evolutionists in other research communities such as evolutionary genetics, hierarchical genetics theory and developmental evolutionary theory, took an interest in adaptations but also in a variety of other causes of evolutionary outcomes, such as genetic linkage, phyletic history, developmental factors, drift, migration, and so on. Their RQ was often: 'What evolutionary factor(s) account for the form and distribution of this trait?' and also, more specifically, '*Does this trait have a function?*'. The possible and responsive answers to the more detailed version of this RQ include all of the evolutionary causes just mentioned, and more.

Now compare the two italicized RQs from the two communities of inquiry. They look almost identical! But they have vastly different sets of possible and responsive answers. So it is easy now to see why the evolutionary psychologists, sociobiologists, and animal behaviorists were criticized for their method: it constrained the possible evolutionary answer to only a single possible cause, the selection of an evolutionary adaptation, i.e. ‘a function’, while it remained unclear why or whether any other possible cause, such as genetic linkage, drift, migration, and developmental factors, would ever be considered. As a result, when one possible function failed, the adaptationists just moved to the next possible and responsive function, and did not look elsewhere (Lloyd 2021).

In this paper, we use the LRQ tool by considering the possible and responsive answers to the following attribution questions:

RQ1: What was the effect of ACC on X?

RQ2: What were the independent causal factors leading to X, and was ACC among them?

RQ3: What were the causal factors leading to X (including external drivers), and was ACC among them?

Here we follow the IPCC Guidance Paper on Detection and Attribution (Hegerl *et al* 2010) who use the term ‘external driver’ to refer to a causal factor external to X, but internal to the climate system, and in contrast to external forcing. RQ1 is a question asked by various publics, including the IPCC, and it is also how many climate science publications present their results. Yet in order to answer RQ1, it is necessary to first answer either RQ2 or RQ3 (Hegerl *et al* 2010). This follows from the fact that in order to convert data into evidence for or against a putative explanation for that data, one must consider all other possible explanations for the data as well (see e.g. Shepherd 2021). However, there is a fundamental difference between *unconditional* attribution (RQ2), where the causal factors are independent and include only external forcing and internal variability, and *conditional* attribution (RQ3), where the causal factors also include elements of the climate system itself, such as changes in sea-surface temperature (SST) patterns or in large-scale atmospheric temperature gradients, which are not independent of either ACC or internal variability.

We illustrate the difference by considering three examples of X: a global-scale thermodynamic phenomenon, a regional climate phenomenon, and an extreme event. Our point is not that one RQ is right and the other wrong, as different actors will have legitimate reasons, with their own goals, to pursue different research questions, including those contrasted here. Our point is simply that they are different, and those differences constrain the possible and responsive answers to the RQ. Failing to appreciate this fact can lead to confusion and misunderstanding. We focus here on attribution, but briefly discuss the implications for risk (understood broadly), which differ significantly between RQ2 and RQ3.

2. Aggregation and disaggregation

Climate is generally understood to comprise the statistics of weather, so to analyze changes in climate it is necessary to estimate those statistics from data samples, which are grouped in some way. This process is called ‘aggregation’ of data. In theoretical probability distributions, which arise from well-defined chance processes, the data are usually assumed to be independent and identically distributed (iid), meaning that the data values are exchangeable and do not depend on each other, and differ from each other only by chance.

Data samples in the real world are never iid. Nevertheless, real-world data can sometimes follow such theoretical distributions remarkably well. A famous example is the analysis by Bortkiewicz (1898) of the number of Prussian cavalry units suffering the death of a soldier by horsekick in a given year, collected over a 20 years period (figure 1). The data are seen to closely follow a Poisson distribution, which describes the number of rare events over a given period, given a fixed likelihood of occurrence. Does this mean that those deaths occurred by chance? From the perspective of the Prussian army, who would be interested in the total number of deaths, that is a reasonable perspective.

However, each of those deaths would have a tragic causal story behind it, involving a particular horse and an individual man. Doubtless, the commanding officer of such a unit, if confronted by the family of the dead soldier, could have thought of various things that could have been done differently that day, which might have avoided the death. The nature of the attribution (the possible and responsive answer to the RQ) is very different in the two cases.

Both perspectives are of course valid (Pandit 2016), and this sort of dialectic between aggregate and individual questions occurs in many domains, e.g. public health vs clinical treatment. It is well understood that a factor that increases risk across a population, while of obvious relevance for anybody interested in effects on entire populations, cannot be reliably applied to individuals within that population (Bueno de Mesquita and Fowler 2021). Thus, whilst the pivot from attribution to risk is straightforward for entire populations, it is not so for individuals within that population. The same principle clearly applies to climate



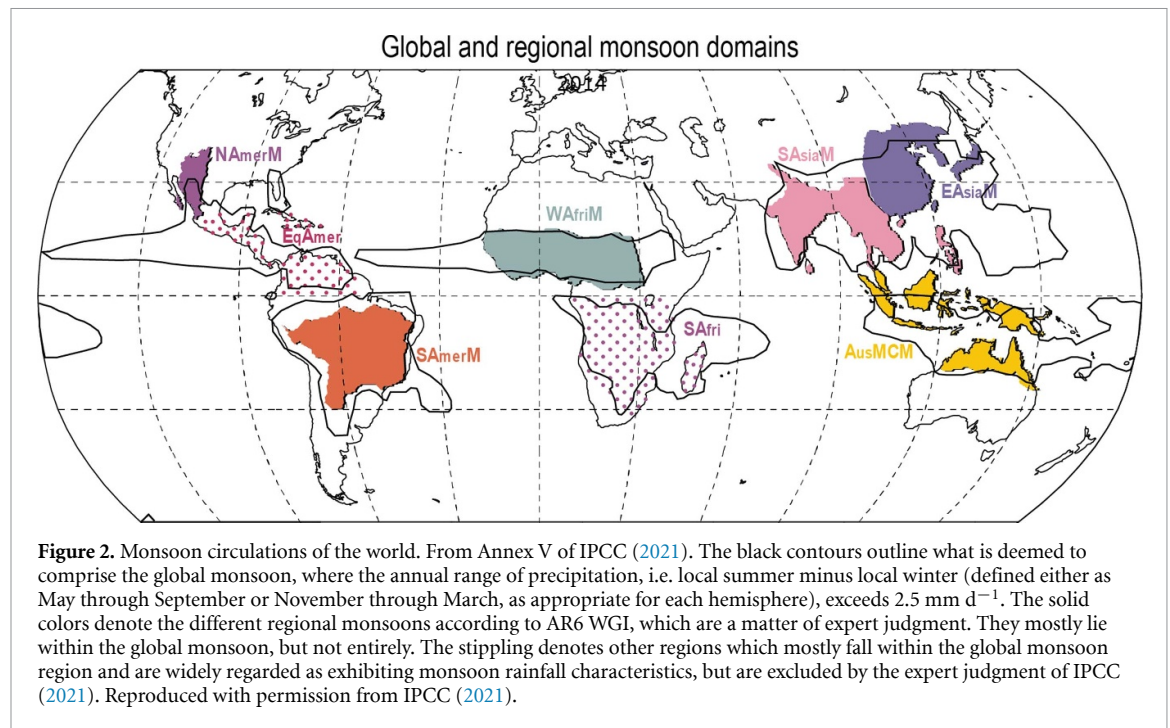
Figure 1. The number of cavalry units experiencing a particular number of deaths by horsekick in a given year, collected over a 20 years period. Data from von Bortkiewicz (1898), reproduced from Pandit (2016). Pandit J J 2016. John Wiley & Sons. © 2015 The Association of Anaesthetists of Great Britain and Ireland.

data as well (Brown 2022, Perkins-Kirkpatrick *et al* 2022). In particular, for a singular extreme event (e.g. Hurricane Sandy; see Sobel 2014), where explanations invariably focus on detailed causal processes, it is generally more informative to disaggregate rather than to aggregate, and construct ‘storyline’ explanations (Shepherd 2016). The key point is that the kind and form of data needed to address and answer different research questions depends completely on which research question(s) are under consideration.

Both aggregation and disaggregation, even for their own legitimate purposes, have their perils. Because the real world is not iid, any process of aggregation will inevitably include systematic as well as random variations in the data. There is thus a trade-off involved: aggregating over larger samples will beat down the noise from random variations, but also obscure the systematic variations which may be of scientific interest. This trade-off arises frequently in climate science, e.g. in the process of taking annual (Zappa *et al* 2015a) or zonal averages (Heggin *et al* 2008). Unfortunately, however, the systematic variations remain in the data; they are just hidden. Drawing inferences from aggregated data relies on the ability to remove the effect of (or ‘control for’, to use statistical terminology) those systematic variations, which represent potentially confounding factors.

This problem is exemplified in recent studies of storm-track shifts in the Southern Hemisphere (SH), a focus of much attention. Climate scientists have sought to constrain the uncertainty in model predictions through the use of ‘emergent constraints’ (Hall *et al* 2019), which are statistical relationships between an observable property of the climate system and a predicted property from the same model-type. Hall *et al* (2019) emphasize the importance of these statistical relationships being causal rather than merely correlational. Kidston and Gerber (2010) claimed to have found such an emergent constraint between the magnitude of the zonally-averaged SH storm-track shift and the timescale of its internal variability. Such a relationship has theoretical support in the fluctuation–dissipation theorem of statistical physics (Breul *et al* 2022), which relates the forced response of a fluctuating system to its internal variability timescales, so has some physical plausibility. However, Simpson and Polvani (2016) showed that the correlational relationship found by Kidston and Gerber (2010) was a statistical artifact, due to a failure to control for the confounding effect of the seasonal cycle. They proposed another emergent constraint for the winter season only, between the magnitude of the storm-track shift and its climatological location. Again, such a relationship has plausible theoretical support (Simpson and Polvani 2016). However, Breul *et al* (2023) showed that Simpson and Polvani’s (2016) correlational relationship arises from aggregating over two very different behaviors, one in the Pacific basin (where there are two jets, not just one) and another elsewhere, and does not reflect a causal physical relationship. This can be seen as a failure to control for the confounding effect of longitudinal asymmetry, which is quite pronounced in the SH winter season.

There is no end to confounding factors; when you control for one, there will be yet more lurking. Out-of-sample testing does not help, since the same systematic error will just be repeated. The hope is that one can make the remaining systematic variations small enough to be ignorable, but that can only ever be a



hope. For this reason, the gold standard for statistical inference is generally considered to be randomized controlled trials, which are widely applied in situations where controlled trials are possible, such as tests for the efficacy of a new drug. But randomized controlled trials are not an option in climate-change science. Yet in climate science, confounding factors are everywhere. Here be dragons.

On the other hand, causal analysis of singular events runs the risk of telling ‘just-so stories’ (Shepherd and Lloyd 2021). An example is explanations from atmospheric dynamics of various aspects of climate change, which are themselves singular events. Typically, if a theory predicts a certain effect (e.g. a poleward shift of the storm tracks in response to greenhouse gas forcing), then its proponent considers that to be an explanation. But there is a 50/50 chance of getting the sign of such an effect right, so that is not a very high bar. Even though the poleward shift of storm tracks is an accepted prediction of climate change, there are something like two dozen competing theories for it (Shaw 2019).

Yet sometimes there is no avoiding moving from generalizations to particular, singular cases. Consider the example of regional monsoon circulations, which are a crucial aspect of the climate system. There are many regional monsoon circulations around the planet (figure 2), which are explainable in general terms from climate dynamics (Geen *et al* 2020). However, their manifestation in each region is different, because of differences in the extent of the land region and the role of local mountain ranges. Moreover, the definition of regional monsoon used in figure 2 is a matter of expert judgment, and excludes some regions, shown by the stippling, which are otherwise widely recognized as monsoons. Thus, for regional aspects of climate change, each monsoon circulation has to be treated differently. This is a problem in singular causation, which requires distinctive methods.

3. Three examples

We now consider the Research Questions posed in the Introduction, applied to three examples.

3.1. Example 1: Increase of global-mean surface temperature (GMST)

Our first example is attribution of global warming itself, as evident in the time series of GMST. In this case, RQ1 (box 1) amounts to whether there is a detectable, i.e. statistically significant, increase in GMST (meaning an increase that cannot be explained by internal variability), which is attributable to ACC. Following the classic detection-attribution (D&A) framework used by IPCC WGI (Hegerl *et al* 2010), this question is converted into the unconditional RQ2 (box 1), which considers the effect of other external forcings as well as ACC and internal variability. The answer to RQ2 is provided by predicting, with one or more climate models, how GMST responds to anthropogenic forcings. This can be done by subtracting the simulations with only natural forcings from those with both natural and anthropogenic forcings (figure 3). Essential elements in the argumentation are that the climate models are understood to be capable of representing the relevant physics

Box 1. Research Questions and their Possible and Responsive Answers for Example 1.

RQ1: Was there an observed change in GMST, and if so, can it be attributed to ACC?

A: There was a statistically significant increase in GMST, which can be attributed to ACC

RQ2: What were the independent causal factors leading to the observed change in GMST, and was ACC among them? (Answers are mutually compatible)

A: Internal variability contributed x_1 (percent or absolute amount)

A: ACC contributed x_2

A: Other known forcings contributed x_3

RQ3: What were the causal factors (including external drivers) leading to the observed change in GMST, and was ACC among them? (Answers are mutually compatible)

A: Changes in ENSO and AMO contributed x_1 (percent or absolute amount)

A: ACC (conditional on ENSO and AMO) contributed x_2

A: Other known forcings (conditional on ENSO and AMO) contributed x_3

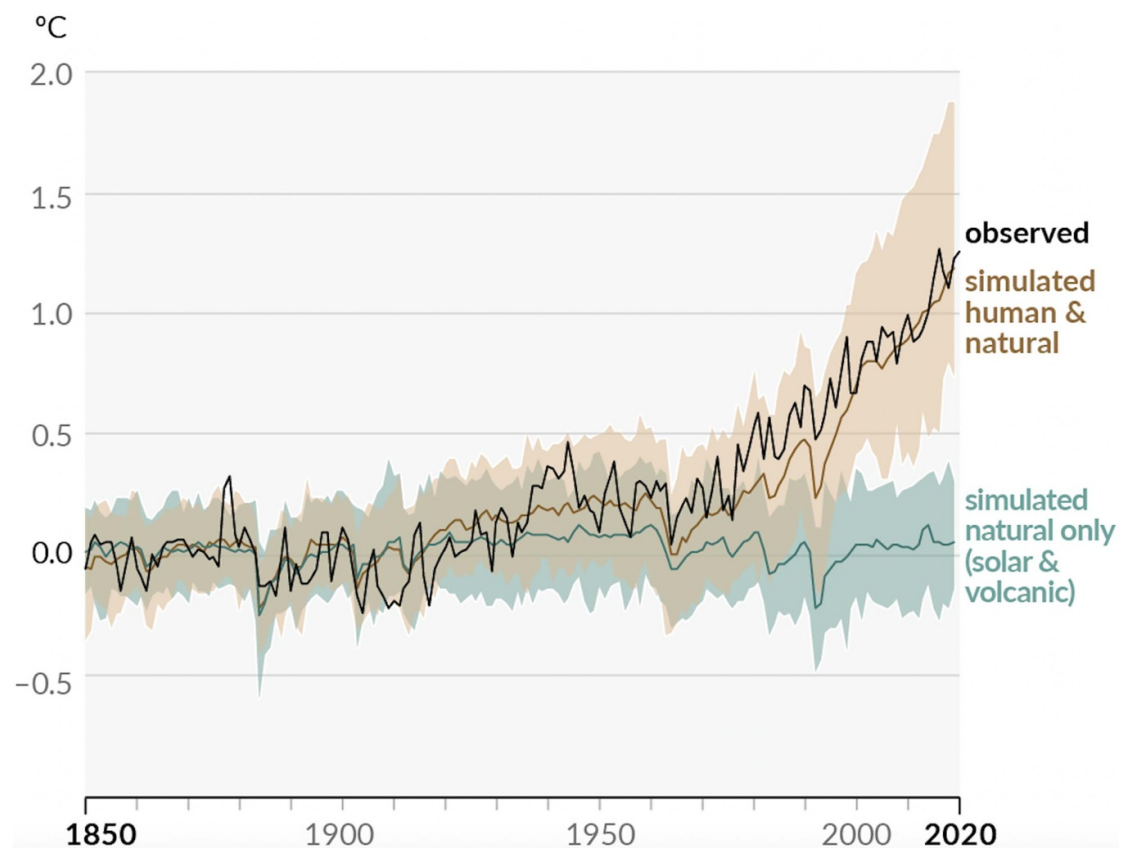
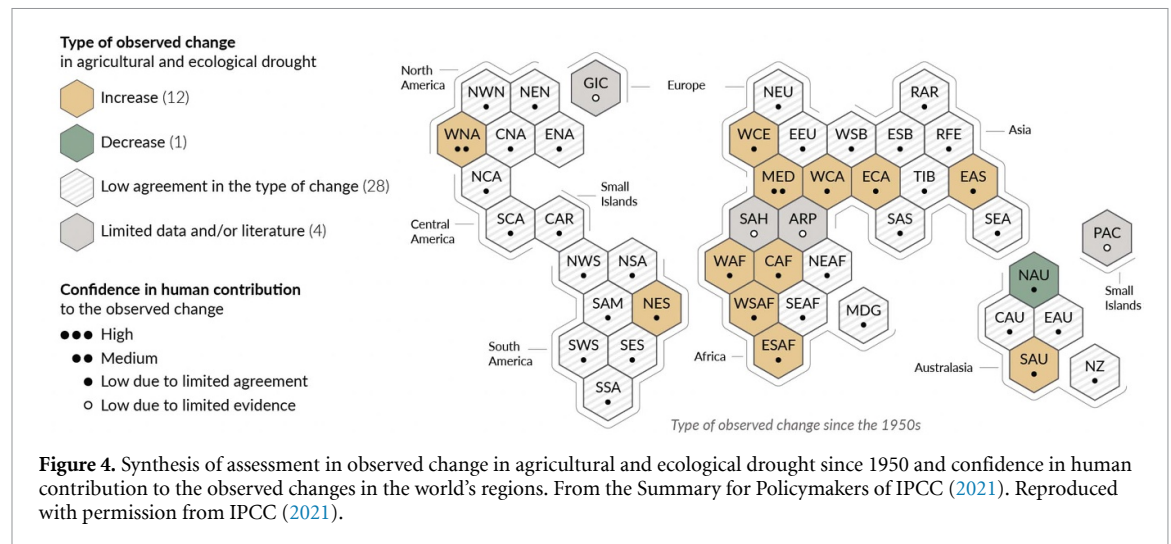


Figure 3. Observed global warming (black) relative to 1850–1900, and as simulated by climate models with only natural forcings (green) and with both natural and anthropogenic forcings (brown). From the Summary for Policymakers of IPCC (2021). Reproduced with permission from IPCC (2021).

of global-mean warming, and that the applied forcings are broadly realistic (Lloyd 2010). To answer RQ1, it is also essential to be able to argue that the estimated effect could not be the result of internal variability (or sampling uncertainty, to use statistical language). That is partly shown by the gap between the shaded regions in figure 3, and partly by the additional fact that the spatial fingerprint of the observed climate response, with warming throughout the oceans and the troposphere, and cooling only in the stratosphere, is, at least to first order, distinct from the surface warming that would occur from internal variability, which would be compensated elsewhere in the climate system, e.g. in the ocean (Hegerl and Zwiers 2011).



The IPCC AR6 WGI Summary for Policymakers (SPM) answers RQ2 as follows (IPCC 2021, paragraph A.1.3):

‘The likely range of total human-caused global surface temperature increase from 1850–1900 to 2010–2019 is 0.8 °C–1.3 °C, with a best estimate of 1.07 °C. It is likely that well-mixed GHGs contributed a warming of 1.0 °C–2.0 °C, other human drivers (principally aerosols) contributed a cooling of 0.0 °C–0.8 °C, natural drivers changed global surface temperature by –0.1 °C to +0.1 °C, and internal variability changed it by –0.2 °C to +0.2 °C.’

That attribution statement does not directly answer RQ1, but it is straightforward to construct a responsive answer to RQ1 from the information provided. The comparatively small magnitude of the contribution from internal variability shows that the observed increase in GMST is statistically significant, whilst the attribution of the warming to ACC comes from the comparatively small magnitude of the contribution from other forcings, and the match between the ACC contribution and the observed warming. Thus, for this example, the two RQs lead to essentially the same conclusion. Moreover, there are direct implications for future risk in both cases, since there is a high confidence that the effect of ACC on GMST will not only continue but also increase in the future.

Since an unconditional attribution of long-term GMST increase to ACC is sought by the IPCC, and is found, RQ2 is preferred to RQ3 for this purpose. However, modes of climate variability such as El Niño Southern Oscillation (ENSO) and the Atlantic Multidecadal Oscillation (AMO) have a discernible effect on GMST, and may be considered as external drivers of GMST. In RQ2 those phenomena are implicitly included in the internal variability, but several studies have asked RQ3 using multiple linear regression of the GMST timeseries itself (see Imbers *et al* 2013 and references therein). Since neither ENSO nor AMO exhibit any long-term trend this move does not affect the attribution of long-term GMST increase to ACC. However, by separately considering the predictable component of internal variability, RQ3 improves the detection of small forced signals such as those arising from solar variability, helps in the understanding of decadal variations in GMST, and provides some near-term predictability of GMST (Imbers *et al* 2013).

3.2. Example 2: Regional trends in drought

Our second example is observed changes in drought (figure 4). Drought is an inherently regional phenomenon so one cannot aggregate globally, as with GMST. Instead, in this figure the aggregation is done across regions defined by the IPCC, which are multi-purpose (i.e. they are not defined for drought alone) and which represent a compromise between physical and social factors (Iturbide *et al* 2020). In particular, there will invariably be a trade-off between the level of disaggregation desired by governments (who the IPCC reports are designed to serve) and the increased confidence that can usually be assigned to more aggregated statements. Moreover, there are also various physical definitions of drought that could be used, as drought is a complex, multifaceted phenomenon which affects different societal sectors (e.g. water resources, energy, agriculture, ecosystems) differently. In contrast to GMST, which is a dependent variable in even the simplest physics-based models of climate, drought is an emergent property of climate which has no clear physical definition (Zaitchik *et al* 2023). For the construction of figure 4 a definition of drought had to be chosen, but as with the regional monsoons in figure 2, it is a matter of judgment, especially when applied across a region.

As well as the ambiguity in the definition of drought and the choice of regions over which to aggregate, there are other differences with GMST which are relevant for attribution. The first is that the sign of the response to ACC is not obvious a priori, as evidenced by the fact that figure 4 allows for both increase and decrease. (Although there is a general expectation of increase in agricultural and ecological drought from increased evaporation due to warming—which is borne out by the fact that the cases of increase in figure 4 far exceed those of decrease—for any particular region these expectations can certainly fail (van Garderen and Mindlin 2022)). The second is that, relatedly, changes in atmospheric circulation are highly relevant for drought, yet the response of atmospheric circulation to ACC projects strongly on the modes of internal variability. This is expected theoretically from the aforementioned fluctuation–dissipation theorem, and is also seen in climate model simulations (e.g. Zappa *et al* 2015b). As a result, it is not as easy to distinguish signal from noise as it is for GMST, especially since the signal itself is so uncertain (Shepherd 2014).

The information shown in figure 4 provides the IPCC's answer to RQ1 for this example, which is arrived at using the unconditional RQ2 (box 2). The possible and responsive answers for each region are increase or decrease in the severity of a basket of metrics for agricultural and ecological drought (e.g. soil moisture), together with a confidence level in the attribution to ACC. The majority of the regions are seen to fall into one of the two 'unknown' categories, which can be seen as an inability to provide a responsive answer to RQ2, and hence to RQ1. This could be because the observational record is considered inadequate to characterize the changes, or because the changes have not been documented or analyzed in a peer-reviewed publication, or because the climate models do not provide a sufficiently clear prediction of the expected changes to allow for the hypothesis testing that is inherent in the classic IPCC D&A framework. Note that in each of those cases, disciplinary norms are involved.

Box 2. Research Questions and their Possible and Responsive Answers for Example 2.

RQ1: Was there an observed change in drought in region Y, and if so, can it be attributed to ACC?
(Answers are mutually exclusive)

A: There was a statistically significant increase in drought in region Y, which can be attributed to ACC

A: There was a statistically significant decrease in drought in region Y, which can be attributed to ACC

RQ2: What were the independent causal factors leading to the observed change in drought in region Y, and was ACC among them? (Answers are mutually compatible)

A: Internal variability contributed x_1 (percent or absolute amount)

A: ACC contributed x_2

A: Other known forcings contributed x_3

RQ3: What were the causal factors (including external drivers) leading to the observed change in drought in region Y, and was ACC among them? (Answers are mutually compatible)

A: Changes in large-scale temperature gradients contributed x_1 (percent or absolute amount)

A: ACC (conditional on those large-scale temperature gradients) contributed x_2

A: Other known forcings (conditional on those large-scale temperature gradients) contributed x_3

The inability to provide a responsive answer to RQ2 (and thus to RQ1) over most of the world's regions is an example of what Shepherd (2019) referred to as the 'confidence straightjacket' that arises in situations of high uncertainty. Yet it is not as if there is no information relevant to those regions. Section 10.4.2 of the IPCC AR6 WGI report (Doblas-Reyes *et al* 2021) assessed three regional examples of observed precipitation changes (a major driver of drought): Sahel drought and recovery (straddling SAH, WAF, CAF and NEAF in figure 4), south-eastern South America summer wetting (SES), and south-western North America drought (straddling WNA and NCA). Their strong confidence levels, taking into account multiple lines of evidence

Table 1. Summary of selected findings from Section 10.4.2 of Doblas-Reyes *et al* (2021). Bold font is used for emphasis.

There is very high confidence (robust evidence and high agreement) that patterns of 20th-century ocean and land surface temperature variability have caused the **Sahel drought and subsequent recovery** by adjusting meridional gradients. There is high confidence (robust evidence and medium agreement) that the changing temperature gradients that perturb the West African monsoon and Sahel rainfall are themselves driven by anthropogenic emissions: warming by GHG emissions was initially restricted to the tropics but suppressed in the North Atlantic due to nearby emissions of sulphate aerosols, leading to a reduction in rainfall. The North Atlantic subsequently warmed following the reduction of aerosol emissions, leading to rainfall recovery.

There is high confidence that **South-Eastern South America summer precipitation has increased** since the beginning of the 20th century. Since AR5, science has advanced in the identification of the drivers of the precipitation increase in South-Eastern South America since 1950, including GHG through various mechanisms, stratospheric ozone depletion and Pacific and Atlantic variability. There is high confidence that anthropogenic forcing has contributed to the South-Eastern South America summer precipitation increase since 1950, but very low confidence on the relative contribution of each driver to the precipitation increase.

There is high confidence (robust evidence and medium agreement) that most (>50%) of the anomalous atmospheric circulation that caused the **south-western North America negative precipitation trend** can be attributed to teleconnections arising from tropical Pacific SST variations related to PDV. There is high confidence (robust evidence and medium agreement) that anthropogenic forcing has made a substantial contribution (about 50%) to the south-western North America warming since 1980.

(table 1), stand in contrast to the largely agnostic perspective of figure 4 with regard to those regional phenomena. Crucially, the findings allow for conditional attribution to known physical drivers of drought in each region—most notably to changes in large-scale temperature gradients, either in the atmosphere or as manifested in particular SST patterns—which might or might not be attributable to ACC. Thus, those authors were answering RQ3 for this example (box 2). Because the changes in large-scale temperature gradients were not themselves fully attributed to external forcing or internal variability, the conditional attribution statements provided by RQ3 do not answer RQ1, i.e. they do not permit a straightforward pivot to future risk (see Chapter 12 of the IPCC AR6 WGI report (Ranasinghe *et al* 2021), where the stated uncertainties in these cases are high). That is the price one inevitably pays for the wider range of possible and responsive answers offered by the conditional RQ3 compared to the unconditional RQ2. However, the conditional nature of the attribution allows for additional or new information to be subsequently brought to bear by other researchers concerned with those phenomena, and to be combined with knowledge about regional vulnerabilities and exposure, which is an essential part of any risk analysis.

3.3. Example 3: An extreme event

Our third and final example is an extreme event, namely the failure of the South American monsoon rainfall in the austral summer of 2013/14 discussed by Rodrigues and Shepherd (2022). That was a compound extreme event induced by an anomalous blocking anticyclone, which cut off the seasonal flow of moisture to the region, led to land and marine heat waves and drought, and affected the food-water-energy nexus. If X is that specific event, then there is no possible and responsive answer to RQ2, because the sample size is one, and unconditional attribution requires aggregation. Moreover, there is no detection step involved, as in the first two examples, because the very posing of the question is conditional on the fact that the event occurred. In order to allow a possible and responsive answer to RQ2, X must be changed to *events of that type*, by the creation of an event class (Shepherd 2016) (reflected in RQ1 in box 3). This move allows a probabilistic attribution, which can follow the classic, unconditional IPCC D&A framework and which directly links attribution with risk. As noted earlier, however, the results cannot be directly applied to that particular event, hence the answer is not actually responsive to the original question. In contrast, singular causation is a possible and responsive answer to RQ3 (box 3). Moreover, as with Example 2 and figure 4, the aggregation involved in the definition of ‘events of that type’ involves a compromise between scientific and social factors (Philip *et al* 2020).

Box 3. Research Questions and their Possible and Responsive Answers for Example 3.

RQ1: What was the effect of ACC on the failure of the South American monsoon rainfall in the austral summer of 2013/14? (Answers are mutually exclusive)

A: ACC increased the likelihood/magnitude of events of that type

A: ACC decreased the likelihood/magnitude of events of that type

RQ2: What were the independent causal factors leading to the failure of the South American monsoon rainfall in the austral summer of 2013/14, and was ACC among them? (Answers are mutually compatible)

A: Internal variability contributed x_1 (percent or absolute amount) to events of that type

A: ACC contributed x_2 to events of that type

A: Other known forcings contributed x_3 to events of that type

RQ3: What were the causal factors (including external drivers) leading to the failure of the South American monsoon rainfall in the austral summer of 2013/14, and was ACC among them? (Answers are mutually compatible)

A: A blocking anticyclone was the physical driver of the event

A: ACC (conditional on the blocking anticyclone) contributed $x_1/y_1/z_1$ to the observed land heatwave/marine heatwave/drought

A: Other known forcings (conditional on the blocking anticyclone) contributed $x_2/y_2/z_2$ to the observed land heatwave/marine heatwave/drought

A probabilistic attribution performed by Martins *et al* (2018) of the multi-year period encompassing the drought, covering the southern parts of region NES in figure 4, concluded that ‘we could not find sufficient evidence that ACC increased drought risk’. What they meant was that they were not able to reject the null hypothesis of no anthropogenic effect, which is a RQ1 framing⁴. This despite the fact that both observations and CMIP5 climate simulations generally supported *increased drought risk* in that region, as shown in figure 4 and as also concluded in section 12.4.4.2 of the IPCC AR6 WGI report (Ranasinghe *et al* 2021). However, Martins *et al* (2018) included two versions of the Met Office model in their analysis which indicated decreased drought risk (see their figure 13.2(i)), and which prevented them from rejecting the null hypothesis. But absence of evidence is not evidence of absence (Shepherd 2021). Insisting that all lines of evidence agree sets an impossibly high bar for such a singular event, especially when climate models can struggle to represent the phenomena in question (Pereima *et al* 2022).

Most climate scientists would surely place more weight on the combined evidence from observations and the CMIP5 ensemble than on two versions of the same climate model, no matter what model that was. Thus, one option for Martins *et al* (2018) would have been to loosen the evidentiary standard to ‘more likely than not’ (Lloyd *et al* 2021), which would have allowed an attribution statement to be made in this case. However, that would still have been an aggregated statement, not applicable to that particular event. Rodrigues and Shepherd (2022) discuss the various causal factors affecting drought risk in the region based on the published literature—notably teleconnections from SST patterns in the equatorial Indian and Pacific oceans—which could be used to develop a possible and responsive answer to RQ3 in this case. In particular, the known thermodynamic factors such as increasing likelihood of heat waves (for a given circulation regime) could then be invoked to support the conditional attribution (Hegerl *et al* 2010).

4. Discussion

The three examples above illustrate that by asking the unconditional RQ2, which if answered would immediately answer RQ1, one seeks ACC as *the* explanation of a climate phenomenon and regards all other factors as potentially confounding. This approach forces one into a probabilistic perspective, and often leads to null hypothesis significance testing, where the null hypothesis is chance together with natural forcings. In contrast, by asking the conditional RQ3, one seeks all relevant and possibly interacting causal factors, of which ACC will only be one, and the other factors may be equally if not more important. The latter is the

⁴ Philip *et al* (2020) indicates four possible answers to RQ1 for probabilistic attribution: the two listed in box 3, plus ‘ACC did not change events of that type’, and ‘assessment not possible in current state of knowledge’. We regard the first of these as logically impossible, and the second as a failure to answer RQ1.

storyline approach, where the other factors (including local inhomogeneities across a region) may be more appropriately treated as mediators, rather than confounders (Lloyd and Shepherd 2021). The unconditional RQ2 has been the traditional approach of IPCC WGI. Indeed, the IPCC D&A framework (Hegerl *et al* 2010) offers the following guidance:

To avoid selection bias in studies, it is vital that the data are not preselected based on observed responses, but instead chosen to represent regions/phenomena/timelines in which responses are expected, based on process-understanding.

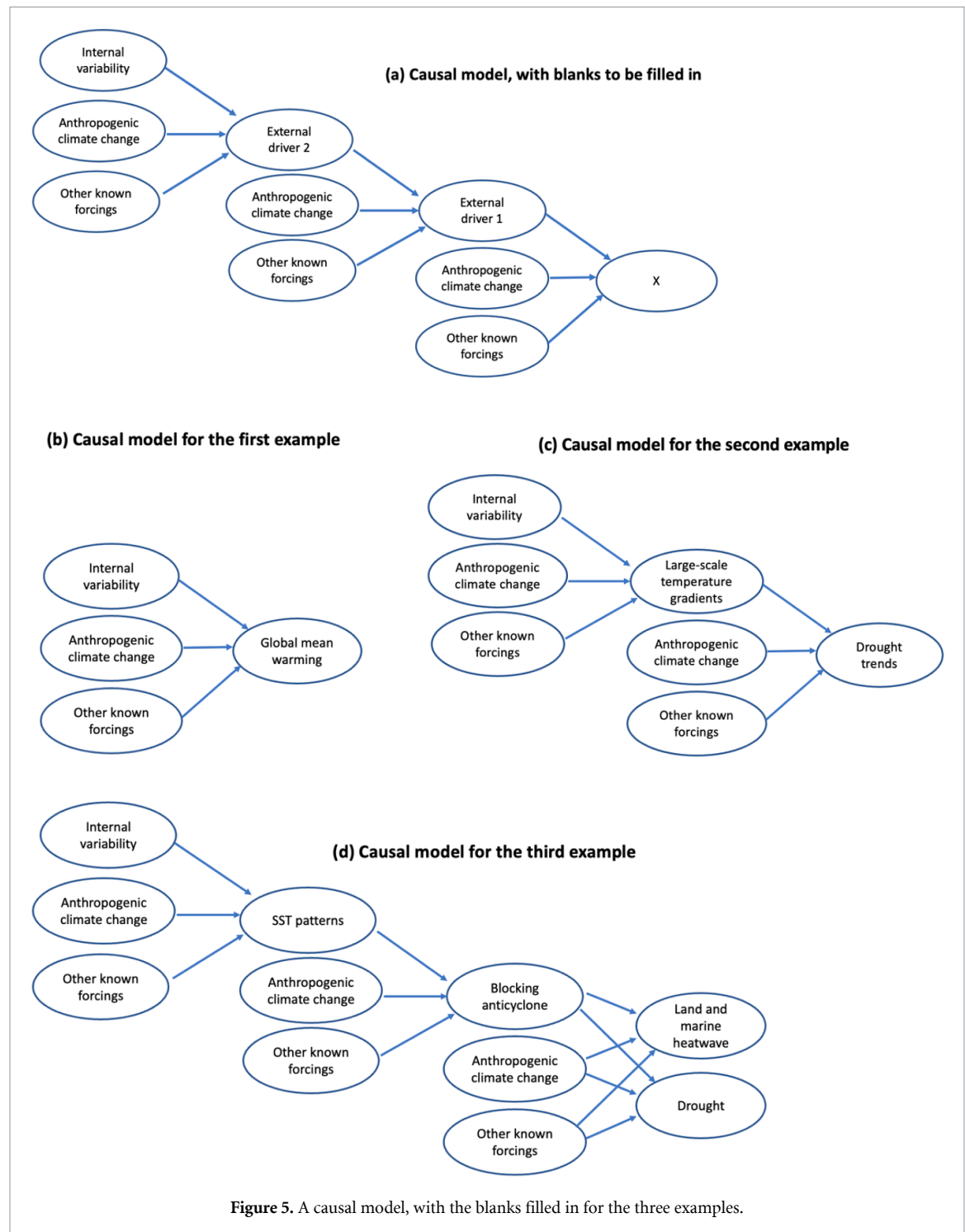
Confounding factors (or influences) should be explicitly identified and evaluated where possible.

Both statements are standard forms of guidance for statistical inference based on aggregated sample populations. The first guards against the multiple testing problem, and the second guards against confounding factors. This guidance is entirely appropriate when seeking an unconditional attribution. However, the guidance demands a model-first approach rather than an observations-first approach, and furthermore is challenged by complex local environments. Both can lead to epistemic injustice (Shepherd and Sobel 2020). The conditional attribution provided by RQ3 finesses these limitations by allowing the articulation of various explanations (which may be plural), and the assessment of the strength of evidence behind them. Conditional attribution aligns more closely with how climate change is treated from the lens of the Intergovernmental Science-Policy Platform on Biodiversity and Ecosystem Services (Lloyd and Shepherd 2020, Shepherd and Lloyd 2021) or the Sustainable Development Goals.

But what other causal factors should be considered in climate-change attribution? As noted earlier, as well as other external forcings (such as volcanic eruptions), the IPCC attribution framework (Hegerl *et al* 2010) allows for external *drivers*, where external does not mean external to the climate system but only external to the feature of interest. This device is especially useful for the climate impacts considered by IPCC WGII, as it allows WGII to focus on the linkage between changes in climate (from whatever origin) and their impacts, which the observed data can more convincingly answer. But the same device can be usefully deployed within the WGI domain as well. Thus, for example, observed changes in regional climatic extremes and impacts may be attributable in part to changes in remote climatic drivers (as in table 1), whose response to ACC may be highly uncertain. Indeed, that is how climate scientists would normally begin an attribution investigation of regional climate change. But since climatic drivers are phenomenological, this approach to attribution inevitably depends on how the phenomena are defined. In other words, it is *phenomena-laden*. That is not in itself a bad thing, because it reflects the fact that the relevant phenomena are emergent rather than being definable *ab initio* from the laws of physics, but it does raise the key issue of who controls how the phenomena are defined.

RQ3 works backward from the phenomenon of interest to seek conditional explanations. One can see this as constructing a model outline, with the blanks to be filled in (Lloyd 2021), which is illustrated for our case by figure 5. In panel (a), changes in the phenomenon of interest X are seen to be caused by external driver 1 which is known to affect X in the internal variability, as well as by ACC and any other known forcings. If this dependence is represented in a causal network, then the influence of the forcings is conditional on the state of external driver 1 (Shepherd 2019). However, external driver 1 may itself be affected by ACC and any other known forcings, as well as by external driver 2. One can replicate this structure as far back as one wishes, stopping the process with ‘internal variability’. (Of course, in a causal network, internal variability is also implicit within each of the causal linkages, which are understood to be noisy.)

We can now fill in the blanks for our three examples, based on the possible and responsive answers to RQ2 or 3 (boxes 1–3). In the first example (panel (b)), we can stop at the first step with internal variability, making the attribution unconditional (RQ2), which is why we obtain a possible and responsive answer to RQ1 as well. In the second example (panel (c)), it is clear that observed drought trends in many regions of the planet are driven in large measure by trends in large-scale temperature gradients, which represent external drivers, but there is generally considerable uncertainty in the extent to which such trends are attributable to ACC or other forcings (see Lee *et al* 2022 for the case of tropical SST patterns). In such a situation, dynamical storylines of the external drivers can provide possible and responsive answers to RQ3 (see Zappa and Shepherd 2017 for the case of atmospheric temperature gradients, and Ghosh and Shepherd 2023 for the case of SST patterns). In the third example (panel (d)), which is of a singular event, we may start with a very local external driver, the blocking anticyclone. Such a move explicitly distinguishes between thermodynamic and dynamical aspects of changes in extremes, which is a device that has been widely used in the literature and reflects the different levels of uncertainty in the two aspects (Shepherd 2016, Pfahl *et al* 2017). It can be done statistically for classes of events (Cattiaux *et al* 2010, Horton *et al* 2015), in which case dynamical storylines of future risk can be constructed (Bevacqua *et al* 2020). It can also be done for singular events (Reed *et al* 2020, van Garderen *et al* 2021), where the attribution question can be reframed as ‘What



would a similar extreme event have looked like in the past? However, in this case the pivot to future risk will generally just be thermodynamic. Introducing the additional layer in the causal chain allows the consideration of compound risk in a spatially coherent way, e.g. the conjunction of land and marine heat waves and drought across a region. This structure is elaborated in figure 3 of Rodrigues and Shepherd (2022).

5. Conclusion

Returning to the original motivation of epistemic justice, the conditionality of the causal explanations permitted by RQ3 is a simplifying assumption (Rodrigues and Shepherd 2022). It further allows the construction of regional climate information, which will inevitably be phenomena-laden (in terms of the external drivers described in figure 5 as well as the target phenomenon X), to be based on local knowledge and domain experts. Together, those features help meet the three criteria of context-sensitivity and

decision-relevance, clarity and comprehensibility, and trustworthiness, which were identified by Schwab and von Storch (2018) in their study of a particular use case as being the criteria that matter most to stakeholders. Because attribution is increasingly practiced by such local communities of users with specific research questions, their interests must be incorporated into both the judgments about the definitions of phenomena of interest and the sets of possible and responsive answers to the research questions that are posed (Jézéquel *et al* 2019, Stone *et al* 2021). This may be especially true in cases using AI/ML, where the phenomena introduced to control and categorize big data must be managed in such a way that local users' interests, values, and rights are respected and maintained in the information processing practices (Leonelli 2016, Lloyd *et al* 2022). Note that AI/ML methods need not be incompatible with these local approaches, so long as they are not 'black boxes' (Arribas *et al* 2022). Finally, RQ3 opens the door to further research in a much more effective way than does RQ2. That is also an important benefit, since the value of research is not only in generating findings, but also in opening up new fields of research.

Data availability statement

No new data were created or analysed in this study.

Acknowledgments

EAL was funded by NSF grant 'Big Data and Climate Science: A Philosophical and Methodological Exploration', SES-#1754740. TGS was funded by the European Commission Horizon 2020 project XAIDA (Extreme Events: Artificial Intelligence for Detection and Attribution), Grant Agreement No. 101003469. The authors would like to thank Gabi Hegerl and an anonymous reviewer for helping us to sharpen our arguments.

ORCID iDs

Elisabeth A Lloyd  <https://orcid.org/0000-0001-9349-9805>

Theodore G Shepherd  <https://orcid.org/0000-0002-6631-9968>

References

- Arnet E 2019 Conwy Lloyd Morgan, methodology, and the origins of comparative psychology *J. Hist. Biol.* **52** 433–61
- Arribas A *et al* 2022 Climate risk assessment needs urgent improvement *Nat. Commun.* **13** 4326
- Bevacqua E, Zappa G and Shepherd T G 2020 Shorter cyclone clusters modulate changes in European wintertime precipitation extremes *Environ. Res. Lett.* **15** 124005
- Bouwer L 2022 Machine learning helps to improve climate services *Open Access Government January 2023* pp 428–9 (available at: www.openaccessgovernment.org/article/machine-learning-helps-to-improve-climate-services/149065/) (Accessed 28 February 2023)
- Breul P, Ceppi P and Shepherd T G 2022 Relationship between Southern Hemispheric jet variability and forced response: the role of the stratosphere *Weather Clim. Dyn.* **3** 645–58
- Breul P, Ceppi P and Shepherd T G 2023 Revisiting the wintertime emergent constraint of the Southern Hemispheric midlatitude jet response to global warming *Weather Clim. Dyn.* **4** 39–47
- Brown P 2022 Overestimating the human influence on the economic costs of extreme weather events (available at: <https://patrickbrown.org/2022/02/14/overestimating-the-human-influence-on-the-economic-costs-of-extreme-weather-events/>) (Accessed 28 February 2023)
- Bueno de Mesquita E and Fowler A 2021 *Thinking Clearly with Data: A Guide to Quantitative Reasoning and Analysis* (Princeton University Press)
- Burger M, Wentz J and Horton R 2020 The law and science of climate change attribution *Columbia J. Environ. Law* **45** 57
- Cattiaux J, Vautard R, Cassou C, Yiou P, Masson-Delmotte V and Codron F 2010 Winter 2010 in Europe: a cold extreme in a warming climate *Geophys. Res. Lett.* **37** L20704
- Coen D R 2021 A brief history of usable climate science *Clim. Change* **167** 51
- Doblas-Reyes F J *et al* 2021 Linking global to regional climate change *Climate Change 2021: The Physical Science Basis. Contribution of Working Group I to the Sixth Assessment Report of the Intergovernmental Panel on Climate Change* ed V Masson-Delmotte *et al* (Cambridge University Press) pp 1363–512
- Geen R, Bordoni S, Battisti D S and Hui K 2020 Monsoons, ITCZs, and the concept of the global monsoon *Rev. Geophys.* **58** e2020RG000700
- Ghosh R and Shepherd T G 2023 Storylines of maritime continent dry period precipitation changes under global warming *Environ. Res. Lett.* **18** 034017
- Hall A, Cox P, Huntingford C and Klein S 2019 Progressing emergent constraints on future climate change *Nat. Clim. Change* **9** 269–78
- Hegerl G C, Hoegh-Guldberg O, Casassa G, Hoerling M, Kovats S, Parmesan C, Pierce D and Stott P 2010 Good practice guidance paper on detection and attribution related to anthropogenic climate change *Report of the IPCC Expert Meeting on Detection and Attribution Related to Anthropogenic Climate Change* (World Meteorological Organization) (available at: www.ipcc.ch/library/)
- Hegerl G C and Zwiers F W 2011 Use of models in detection and attribution of climate change *WIREs Clim. Change* **2** 570–91
- Hegglin M I *et al* 2022 Space-based Earth observation in support of the UNFCCC Paris Agreement *Front. Environ. Sci.* **10** 941490

- Hegglin M I, Boone C D, Manney G L, Shepherd T G, Walker K A, Bernath P F, Daffer W H, Hoor P and Schiller C 2008 Validation of ACE-FTS satellite data in the upper troposphere/lower stratosphere (UTLS) using non-coincident measurements *Atmos. Chem. Phys.* **8** 1483–99
- Horton D E, Johnson N C, Singh D, Swain D L, Rajaratnam B and Diffenbaugh N S 2015 Contribution of changes in atmospheric circulation patterns to extreme temperature trends *Nature* **522** 465–9
- Imbers J, Lopez A, Huntingford C and Allen M R 2013 Testing the robustness of the anthropogenic climate change detection statements using different empirical models *J. Geophys. Res.* **118** 3192–9
- IPCC 2021 *Climate Change 2021: The Physical Science Basis. Contribution of Working Group I to the Sixth Assessment Report of the Intergovernmental Panel on Climate Change* ed V Masson-Delmotte *et al* (Cambridge University Press)
- Iturbide M *et al* 2020 An update of IPCC climate reference regions for subcontinental analysis of climate model data: definition and aggregated datasets *Earth Syst. Sci. Data* **12** 2959–70
- Jebeile J and Roussos J 2023 Usability of climate information: toward a new scientific framework *WIREs Clim. Change* **14** e833
- Jézéquel A, Yiou P and Vanderlinden J-P 2019 Comparing scientists and delegates perspectives on the use of extreme event attribution for loss and damage *Wea. Clim. Extremes* **26** 100231
- Kidston J and Gerber E P 2010 Intermodel variability of the poleward shift of the austral jet stream in the CMIP3 integrations linked to biases in 20th century climatology *Geophys. Res. Lett.* **37** L09708
- Lee S, L'Heureux M, Wittenberg A T, Seager R, O'Gorman P A and Johnson N C 2022 On the future zonal contrasts of equatorial Pacific climate: perspectives from observations, simulations, and theories *npj Clim. Atmos. Sci.* **5** 1–15
- Leonelli S 2016 *Data-centric Biology: A Philosophical Study* (University of Chicago Press)
- Lloyd E A 2010 Confirmation and robustness of climate models *Phil. Sci.* **77** 971–84
- Lloyd E A 2015 Adaptationism and the logic of research questions: how to think clearly about evolutionary causes *Biol. Theor.* **10** 343–62
- Lloyd E A 2021 *Adaptation* (Cambridge University Press)
- Lloyd E A, Lusk G, Gluck S and McGinnis S 2022 Varieties of data-centric science: regional climate modeling and model organism research *Phil. Sci.* **89** 802–23
- Lloyd E A, Oreskes N, Seneviratne S I and Larson E J 2021 Climate scientists set the bar of proof too high *Clim. Change* **165** 55
- Lloyd E A and Shepherd T G 2020 Environmental catastrophes, climate change, and attribution *Ann. New York Acad. Sci.* **1469** 105–24
- Lloyd E A and Shepherd T G 2021 Climate change attribution and legal contexts: evidence and the role of storylines *Clim. Change* **167** 28
- Martins E S P R, Coelho C A S, Haarsma R, Otto F E L, King A D, van Oldenborgh G J, Kew S, Philip S, Vasconcelos F C Jr and Cullen H 2018 A multimethod attribution analysis of the prolonged northeast Brazil hydrometeorological drought (2012–16) [in “Explaining Extreme Events of 2016 from a Climate Perspective”] *Bull. Am. Meteorol. Soc.* **99** S65–9
- Pandit J J 2016 Deaths by horsekick in the Prussian army—and other ‘Never Events’ in large organisations *Anaesthesia* **71** 7–11
- Pereima M F R, Chaffe P L B, Amorim P B and Rodrigues R R 2022 A systematic analysis of climate model precipitation in southern Brazil *Int. J. Clim.* **42** 4240–57
- Perkins-Kirkpatrick S E, Stone D A, Mitchell D M, Rosier S, King A D, Lo Y T E, Pastor-Paz J, Frame D and Wehner M 2022 On the attribution of the impacts of extreme weather events to anthropogenic climate change *Environ. Res. Lett.* **17** 024009
- Pfahl S, O'Gorman P A and Fischer E M 2017 Understanding the regional pattern of projected future changes in extreme precipitation *Nat. Clim. Change* **7** 423–7
- Philip S *et al* 2020 A protocol for probabilistic extreme event attribution analyses *Adv. Stat. Clim. Meteorol. Oceanogr.* **6** 177–203
- Ranasinghe R *et al* 2021 Climate change information for regional impact and for risk assessment *Climate Change 2021: The Physical Science Basis. Contribution of Working Group I to the Sixth Assessment Report of the Intergovernmental Panel on Climate Change* ed V Masson-Delmotte *et al* (Cambridge University Press) pp 1767–926
- Reed K A, Stansfield A M, Wehner M F and Zarzycki C M 2020 Forecasted attribution of the human influence on hurricane florence *Sci. Adv.* **6** eaaw9253
- Rodrigues R R and Shepherd T G 2022 Small is Beautiful: climate-change science as if people mattered *PNAS Nexus* **1** pgac009
- Schwab M and von Storch H 2018 Developing criteria for a stakeholder-centred evaluation of climate services: the case of extreme event attribution for storm surges at the German Baltic Sea *Meteorol. Hydrol. Water Manag.* **6** 27–35
- Seneviratne S I *et al* 2021 Weather and climate extreme events in a changing climate *Climate Change 2021: The Physical Science Basis. Contribution of Working Group I to the Sixth Assessment Report of the Intergovernmental Panel on Climate Change* ed V Masson-Delmotte *et al* (Cambridge University Press) pp 1513–766
- Shaw T A 2019 Mechanisms of future predicted changes in the zonal mean mid-latitude circulation *Curr. Clim. Change Rep.* **5** 345–57
- Shepherd T G 2014 Atmospheric circulation as a source of uncertainty in climate change projections *Nat. Geosci.* **7** 703–8
- Shepherd T G 2016 A common framework for approaches to extreme event attribution *Curr. Clim. Change Rep.* **2** 28–38
- Shepherd T G 2019 Storyline approach to the construction of regional climate change information *Proc. R. Soc. A* **475** 20190013
- Shepherd T G 2021 Bringing physical reasoning into statistical practice in climate-change science *Clim. Change* **169** 2
- Shepherd T G and Lloyd E A 2021 Meaningful climate science *Clim. Change* **169** 17
- Shepherd T G and Sobel A H 2020 Localness in climate change *Comp. Stud. South Asia Afr. Middle East* **40** 7–16
- Simpson I R and Polvani L 2016 Revisiting the relationship between jet position, forced response, and annular mode variability in the southern midlatitudes *Geophys. Res. Lett.* **43** 2896–903
- Sobel A 2014 *Storm Surge: Hurricane Sandy, Our Changing Climate, and Extreme Weather of the past and Future* (HarperWave)
- Stone D A, Rosier S M and Frame D J 2021 The question of life, the universe and event attribution *Nat. Clim. Change* **11** 276–8
- van Garderen L, Feser F and Shepherd T G 2021 A methodology for attributing the role of climate change in extreme events: a global spectrally nudged storyline *Nat. Hazards Earth Syst. Sci.* **21** 171–86
- van Garderen L and Mindlin J 2022 A storyline attribution of the 2011/2012 drought in Southeastern South America *Weather* **77** 212–8
- von Bortkiewicz L 1898 *Das Gesetz der kleinen Zahlen [The law of small numbers]* (B G Teubner)
- Wentz J, Merner D, Franta B, Lehmen A and Frumhoff P C 2023 Research priorities for climate litigation *Earth's Future* **11** e2022EF002928
- Zaitchik B F, Rodell M, Biasutti M and Seneviratne S I 2023 Wetting and drying trends under climate change *Nat. Water* **1** 502–13
- Zappa G, Hoskins B J and Shepherd T G 2015a Improving climate change detection through optimal seasonal averaging: the case of the North Atlantic jet and European precipitation *J. Clim.* **28** 6381–97
- Zappa G, Hoskins B J and Shepherd T G 2015b The dependence of wintertime Mediterranean precipitation on the atmospheric circulation response to climate change *Environ. Res. Lett.* **10** 104012
- Zappa G and Shepherd T G 2017 Storylines of atmospheric circulation change for European regional climate impact assessment *J. Clim.* **30** 6561–77