Representation Uncertainty in the Earth Sciences

Claire Bulgin[1], Christopher M Thomas[2], Joanne Alexandra Waller[3], and Emma R Woolliams[4]

[1]University of Reading
[2]UK Met Office
[3]Met Office
[4]National Physical Laboratory

November 21, 2022

Abstract

The first Joint Workshop on Representation Uncertainty in the Earth Sciences was held in March 2021. This brought together the Earth observation, data assimilation and forecast verification and post-processing communities, alongside metrologists to discuss the definition and quantification of representation uncertainty within the Earth Sciences. The aim of the workshop was to facilitate cross-disciplinary discussion, establishing where existing methodologies could be shared and to foster future collaboration. A key outcome of the workshop was a working definition of representation uncertainty applicable across all the Earth Science communities, which is presented in this white paper. The cross-disciplinary discussions at the workshop highlighted the need for scientists to work with metrologists to establish a common vocabulary for uncertainties, accessible to Earth Science applications. Further recommendations included regular workshops to discuss progress in defining and quantifying representation uncertainty and awareness of cross-disciplinary funding opportunities to further address representation uncertainty issues.

C. E. Bulgin[1,2], C. M. Thomas[3], J. A. Waller[3] and E.R. Woolliams[4]

[1]University of Reading, Reading, UK,[2]National Centre for Earth Observation, Leicester, UK,[3]UK Met Office, Exeter, UK,[4]National Physical Laboratory, Teddington, UK.

Corresponding author: Claire Bulgin (c.e.bulgin@reading.ac.uk)

Key Points:

- Representation uncertainty can be defined as occurring at a comparison interface between two quantities.
- A common definition of representation uncertainty can be applied across the Earth Science community.
- Collaboration between scientists and metrologists is recommended to define an accessible and relevant vocabulary for uncertainty analysis.

Abstract

The first Joint Workshop on Representation Uncertainty in the Earth Sciences was held in March 2021. This brought together the Earth observation, data assimilation and forecast verification and post-processing communities, alongside metrologists to discuss the definition and quantification of representation uncertainty within the Earth Sciences. The aim of the workshop was to facilitate cross-disciplinary discussion, establishing where existing methodologies could be shared and to foster future collaboration. A key outcome of the workshop was a working definition of representation uncertainty applicable across all the Earth Science communities, which is presented in this white paper. The cross-disciplinary discussions at the workshop

highlighted the need for scientists to work with metrologists to establish a common vocabulary for uncertainties, accessible to Earth Science applications. Further recommendations included regular workshops to discuss progress in defining and quantifying representation uncertainty and awareness of cross-disciplinary funding opportunities to further address representation uncertainty issues.

Plain Language Summary

This paper describes a workshop which brought together experts from different Earth science disciplines to discuss and attempt to define the term "representation uncertainty". We make observations of the Earth using satellites, ground based instruments (such as weather stations) and instruments on ocean buoys or aircraft. These observations are used in their own right, and also by computer models used to generate weather forecasts. The observations themselves are imperfect and we quantify these imperfections using the term "uncertainty". In this paper we discuss the uncertainty that happens when we compare two different sets of observations, two different models, or observations and models. As well as the uncertainties inherent in models and observations, there is also an uncertainty due to the fact that the two things being compared are not representing exactly the same conditions. For example, a satellite observation may represent an average value over a few hundred metres, while an instrument on the surface measures only at a single point, and the model represents an area of several kilometres. Understanding those differences is essential to be able to properly combine different sets of observations, and observations with models.

1 Introduction

The term "representation uncertainty" is used widely in the Earth Sciences, often to describe an uncertainty that occurs at the comparison interface of two different representations of the same physical quantity, although each subcommunity has its own ways of characterizing and evaluating these quantities. The first Joint Workshop on Representation Uncertainty in the Earth Sciences was held online from 23$^{\text{rd}}$ to 25$^{\text{th}}$ March 2021, sponsored by the National Centre for Earth Observation. This discussion-based workshop was open to scientists working in the fields of Earth observation (EO), data assimilation (DA), forecast verification and post-processing (FVPP), and metrology. The workshop was designed to maximize discussion time, beginning with short presentations on where representation uncertainty might occur within the Earth sciences and introducing the field of metrology, which underpins the science of uncertainty. Discussion groups were held first within each discipline, and then across the different disciplines, posing the questions 'Where are we now?' and 'What could we do going forward?'. Science talks and poster contributions were made, showcasing ongoing research towards quantifying representation uncertainty. The workshop concluded with a panel discussion summarizing the key outcomes identified within each discipline. The workshop promoted consistent use of terminology across all disciplines to facilitate discussion.

2 Workshop Outcomes

2.1 Defining representation uncertainty

One of the main challenges of the meeting was to find a definition of representation uncertainty that was widely applicable across the disciplines represented at the workshop. The 'working definition' was regularly revisited throughout the meeting and during the concluding panel discussion, the following definition was proposed:

"The uncertainty associated with a comparison of two quantities (that are themselves uncertain). This contribution to the uncertainty only occurs at the comparison interface."

Representation uncertainty is recognised as a wider term that includes many different sources of uncertainty, each of which is often described by a more specific term. This broad definition, examples of which are shown in Figure 1, should be applicable across all disciplines, as it allows the detail of how representation uncertainty arises to be different for each comparison, which we acknowledge is often complex.

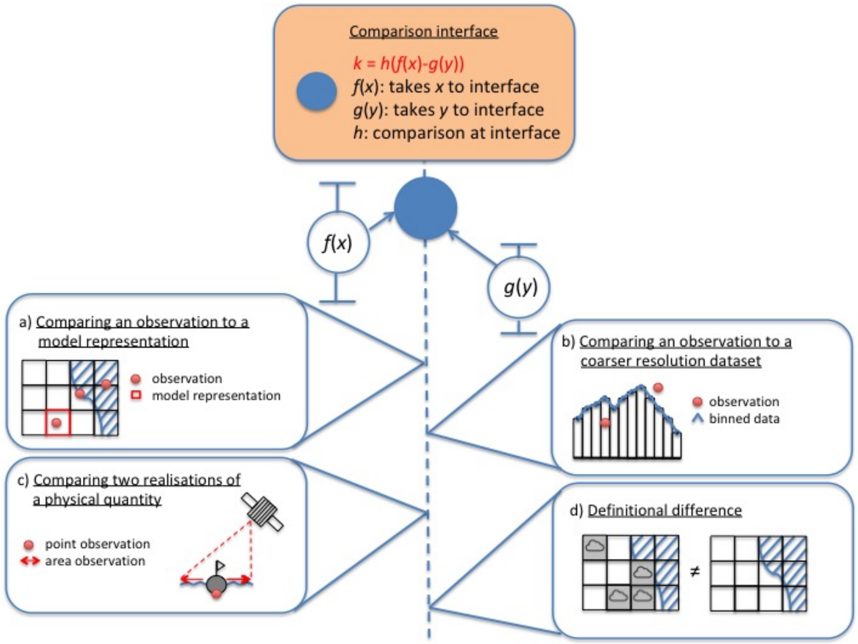

Figure 1: Schematic detailing examples of the comparison interface at which representation uncertainty occurs. At this interface, k is the result of a comparison of observed or modelled quantities, x or y. These quantities are transformed through the functions $f\,(x)$ and $g\,(y)$ to bring them to the comparison interface. The comparison is performed as some function of the difference between the transformed quantities, producing the result$k = h\,(f\,(x) - g\,(y))$. Both$f\,(x)$ and $g\,(y)$ have associated uncertainty and at the point of comparison, there is also the representation uncertainty, identified with the uncertainty on k at the comparison interface. The details of the representation uncertainty are specific to a given comparison, and may arise from a) comparison of an observation to a geophysical model equivalent, b) comparison of an observation to a coarser resolution data set, c) comparison of two realisations of the same physical quantity or d) a definitional difference (which could be a comparison of a measured or modelled data set to a conceptual, more complete, data set or the difference between what users would like and what data producers can provide). The schematics within each box provide just a single example of the possible comparisons.

2.2 Metrology and uncertainty in Earth Science applications

Metrology ensures international consistency and century-long stability of measurements for both science and trade by providing traceability to the International System of Units, the SI. More recently, metrologists have also started to apply the principles of metrological traceability to provide confidence in data derived from modelling. Metrological traceability is built on two core concepts: uncertainty analysis, defined by the Guide to the Expression of Uncertainty in Measurement (the GUM; JCGM 100:2008), and comparisons, formalized to validate uncertainty assessments. Metrologists have worked with the EO community for several decades but have had limited interactions with the DA and FVPP communities.

The Earth sciences often have data value chains, where one community's output becomes another community's input. Each stage of the chain, whether based on instruments or modelling, can be described by a measurement model (equation) with terms that include the previous stage's output, and new terms introduced. Additionally, there will be assumptions inherent in the form of the model which Mittaz et al. (2019) described as the "plus zero uncertainty", written by including a+0 into the measurement model. In laboratory metrology, this+0 includes concepts such as representing an integral with a trapezoidal-rule summation, or assuming that the instrument does not change its properties when moved into the field.

Thus there are two types of uncertainty associated with a value calculated by a measurement model. First, the propagated uncertainty comes from the input quantities, through the GUM methods, to an uncertainty associated with the calculated value. This describes the spread of probable values around the measured value where the true calculated value is expected to lie. Then there is the uncertainty associated with the extent to which the true calculated value represents the quantity of interest – the +0 term.

2.3 Current status within each research discipline

In this section we include some examples of where representation uncertainty occurs within the different disciplines and relate these back to the definition given in Section 2.1, using the shorthand notation ('a'), ('b'), ('c') and ('d') to refer to the four types of representation uncertainty as defined in Figure 1.

2.3.1. Earth Observation (remotely sensed and in situ)

Earth observation includes both in situ and remote sensing measurements., There are "plus zero uncertainties" identified in EO data production that represent a conceptual comparison between the evaluated quantity and the desired quantity ('d'), or a definitonal difference between what users want and what data producers can provide ('d'). Representation uncertainty can also occur at the comparison interface between two datasets, for example ground-based and satellite observations ('b' and/or 'c'), where representation uncertainty arises both from making point-to-area average comparisons, and because in situ data are typically sparse; hence error statistics from comparisons with satellite data are not necessarily globally representative. Error statistics may also mask data complexity ('d', Povey and Grainger, 2019). Representation uncertainty is also inherent in the extent to which the measured variable matches the target geophysical quantity, either due to imperfections in the measurement equation or user requirements ('d', Stier, 2016).

Considering in-situ measurements, representation uncertainty often occurs when a continuous data field in time or space is inferred from discrete measurements. This uncertainty may change temporally as sampling capacity for global measurements increases (Good, 2016). When gap-filling data to provide spatially complete data fields, representation uncertainty can be introduced both by the choice of interpolation method (Dodd, 2015) and by the extent to which the available data represent areas that are not observed. Work presented at the workshop highlighted the challenges in calculating an uncertainty budget where the appropriate representation uncertainty has a dependence on what the user wants to observe: point measurements, spatial or temporal averages, anomalies or trends.

In remote sensing data, representation uncertainty can occur as a result of instrument sampling ('d'), and gridding data ('d' or 'b'). For some data streams, a sub-sample of the full resolution satellite data is passed to the ground receiving station introducing sampling uncertainties in the Level 1 data (Belward et al., 1994). Representation uncertainty is also common at Level 3 when regularly gridding clear-sky only data; these products will include a sampling uncertainty when compared to all-sky observations (Bulgin et al, 2016). Further sources of representation uncertainty are associated with Level 4 products involving data composites from different instruments with different satellite overpass times ('c') (Good et al. , 2020).

2.3.2. Data Assimilation

Data assimilation is the process by which observations are combined with model data, weighted by their respective uncertainties and accounting for physical constraints, in order to provide an optimised estimate of a geophysical state. Often the observations represent different variables from those modeled, so the model space must be mapped into observation space using an observation operator. There are two main comparison interfaces where representation uncertainty may arise: comparison of an observation with its model counterpart ('a'), and, if a processed form of an observation is assimilated, the comparison of this processed observation to the conceptual perfectly processed observation ('d') (Janjic et al., 2018).

There are two main contributors to representation uncertainty at the observation-model interface. First is the error due to unresolved scales and processes, which arises when the observations represent different spatial and temporal scales than those of the assimilating model. Second is the observation operator error arising when the observation operator is approximated either to reduce computational complexity and cost or because of unknown parameters and processes. At the interface between the observation and its processed form, representation uncertainty will arise due to the errors introduced and propagated through the processing chain, sometimes considered as a measurement uncertainty attributed to the assimilated observation. Observations may also be subject to quality control procedures; inaccuracy or occasional failure of these can be an additional source of uncertainty.

Representation uncertainty is most commonly accounted for in DA through including the representation error covariance matrix, F , along with instrument error statistics, E , in the observation error covariance matrix, R = E + F . For this reason, methods to estimate the full R matrix, where the measurement error covariance matrix is known, have been commonly used to isolate representation uncertainty (Desroziers et al (2005), Hollingsworth and Lönnberg (1986)). Alternatively, individual sources of representation uncertainty can be estimated from the error statistics of a comparison between two values: e.g. either two representations of a variable or using two observation operators with differing levels of approximation (Schutgens et al., 2016; Saunders et al., 2018, Waller et al., 2021). Although it is most common for the representation error to be included in R , work presented at the workshop showed that other methods exist that account for uncertainties via updates in small-scale background uncertainties and model uncertainties (e.g. Janjic 2006; Bell et al., 2020). These approaches highlight the difficulty in separating representation uncertainties from other types of uncertainty inherent in the DA process.

2.3.3. Forecast verification and post-processing

Forecast post-processing attempts to correct, combine and exploit the information contained within existing forecasts to produce optimal products for dissemination to the public and other customers. The related field of forecast verification quantifies the success of a forecast by comparing its predictions to observations independent of those predictions. Both of these fields rely on the comparison of quantities at different spatial and temporal scales, and hence will have representation uncertainties associated with various steps in the processing chain. In addition, post-processing must often provide outputs at different scales depending on the user requirements (‘d’).

As discussed in the workshop, the representation uncertainty related to the difference between point observations and grid values (‘a’) is an important topic in FVPP; scale mismatch uncertainties can occur due to sub-grid variability that is not modelled. A new probabilistic post-processing system incorporating verification, the Integrated Model post-PROcessing and VERification (IMPROVER, Met Office, 2019), considers this representation uncertainty. Roberts et al. (pers. comm.) have demonstrated the importance of accounting for local topography when producing forecasts. Topography may not be well represented in a relatively coarse model, but accounting for this in the post-processing step can improve forecast skill. Adjusting the rate of change of temperature with height can also provide better agreement between point and grid values. Ben Bouallegue et al. 2020 used a statistical parameterisation to quantify the representation uncertainty related to the difference between point and grid values. The results can be used in ensemble verification and to represent sub-grid variability that is not present in the model.

A further example of representation uncertainty in FVPP relates to differences in the spatial position between observations of local weather phenomena and the forecast equivalent. This is in distinction to scale mismatch uncertainties; a meteorological feature could be modelled to extremely high precision but be located in the incorrect position (‘c’). This is mitigated in FVPP by using neighbourhood methods and forecast ensembles.

2.4 Existing collaborations between disciplines

The Earth sciences are multidisciplinary, with one community’s output often being another community’s input. While different communities perform their own uncertainty analysis (to differing levels of formality), uncertainties may not be fully transferred and representation uncertainty generated at a comparison interface, may not be fully considered. Projects such as GAIA-CLIM have attempted to address these gaps by bringing the EO, DA and metrology communities together. Some of the residual-based methods commonly used in DA have begun to be used by the EO community. For example, Merchant et al (2020) applied the diagnostics of Desroziers et al. (2005), along with additional bias correction, to estimate error covariance parameters for

SST retrievals.

Another example is the combination of several EO products (remote sensing and/or in situ) to quantify uncertainties in the Earth's energy and water cycles with inverse modelling methods (L'Ecuyer et al. 2015, Rodell et al. 2015, Thomas et al. 2020). The outputs of the inverse modelling procedure can potentially be used to evaluate the accuracy of Global Climate Model products, providing an opportunity to collaborate with the modelling community. For this evaluation to be effective, the uncertainties, including any representation uncertainties, must be accurately determined.

3 Future opportunities for collaboration and community requirements

One barrier to effective collaboration on uncertainty in the Earth sciences is communication, particularly where similar words take different meanings in different groups. Integrating metrologists into this multidisciplinary community can help to bridge this gap, clarifying vocabulary and the distinction between terms such as "error" and "uncertainty" (Mittaz et al, 2019). Regular communication on the definition, sources, quantification and mitigation of representation uncertainty will lead to more efficient transfer of information, better inventory of uncertainty (by considering other perspectives) and more traceability. Ideally the various communities will formulate a common language that can be used consistently in published papers. We can draw parallels with the adoption of the notation in Ide et al 1997 by the DA community.

It was identified in the workshop that collaboration between the DA and FVPP communities could be beneficial, using statistical methods such as the Deroziers et al (2005) diagnostics to account for temporal uncertainties in FVPP. These uncertainties are important given the serial correlation present in many atmospheric and ocanic phenomena. Furthermore, the acquisition of crowdsourced data to provide a 'ground truth' for forecast verification was discussed. Such data are also of interest to the DA community but in order to use these data effectively, extensive quality control and validation would be required. Another example is the use of dynamical-model-dependent propagation of uncertainties for daily satellite products. The ability to apply different methods to a single scenario would provide multiple estimates of representation uncertainty and could help provide confidence in those estimates.

The workshop has begun information dissemination between the Earth science disciplines and future workshops could benefit from inclusion of model developers/validators, instrument manufacturers, and space agencies. Model developers understand which scales and processes are poorly represented by models and may be able to aid quantification of representation uncertainties that are large when data are assimilated. Instrument manufacturers could also create common methods for documenting instrument calibration and data-processing procedures to optimise data use, where commercial sensitivities allow.

4 Recommendations

We encourage all the communities represented at the workshop to continue to work towards a mutually agreed understanding of representation uncertainty, starting with the definition presented in this paper, and acknowledging in many cases more specialised terms will be used to characterise the individual component source(s) of uncertainty. Many of the discussions in the workshop highlighted the need to continue this conversation in a multidisciplinary format, perhaps facilitated by regular workshops, expanded to include other relevant communities. In continuing this conversation, the development of a consistent vocabulary on uncertainties that is relevant to the Earth Sciences community is essential. Such a vocabulary would be supported by metrologists to aid communication and underpin future discussions and collaborations. A cross-community paper developing the ideas outlined in this white paper could be the first target for such future collaboration and could define both a formal vocabulary and make specific recommendations for collaborative research. As funding is often a barrier to inter-disciplinary research, the establishment of a community awareness of projects where representation uncertainty issues may be addressed would be beneficial and "sandpit events" (where scientists work together in a focused way to answer a specific challenge) may also be useful to foster in-depth consideration of inter-disciplinary knowledge transfer on representation uncertainty.

Acknowledgments, Samples, and Data

We acknowledge the many excellent contributions from all workshop participants. In particular we would like to thank the panel, the speakers, session chairs and rapporteurs and the poster authors. We would also like to thank Maurice Cox, John Eyre, Paul Green, Keith Haines, Christopher Merchant, Jonathan Mittaz, Simon Pinnock, Adam Povey, Nigel Roberts and Bruce Wright for their constructive feedback on the white paper draft.

None of the authors has expressed any conflict of interest.

Funding acknowledgements: The organizing committee would like to thank the National Centre for Earth Observation (NCEO), which sponsored this event, providing technical and organizational support throughout from grant NE/R016518/1 of the Natural Environment Research Council. Emma Woolliams acknowledges funding in part from the Instrument Data quality Evaluation and Assessment Service - Quality Assurance for Earth Observation (IDEAS-QA4EO) contract funded by ESA-ESRIN (n. 4000128960/19/I-NS) and in part from UK Government's Department for Business, Energy and Industrial Strategy (BEIS) through the UK's National Measurement System programmes.

Data Availability Statement

This article reports proceedings of a discussion-based online workshop. There are no associated datasets to be made available in conjunction with this commentary.

References

Bell, Z., Dance, S.L., & Waller, J.A. (2020) Accounting for observation uncertainty and bias due to unresolved scales with the Schmidt-Kalman filter, *Tellus A: Dynamic Meteorology and Oceanography* , 72:1, 1-21, doi: 10.1080/16000870.2020.1831830

Belward, A. S., Kennedy, P. J. and Grégoire, J. M. (1994), The limitations and potential of AVHRR GAC data for continental scale fire studies. *International Journal of Remote Sensing, 15* (11), 2215-2234. doi:10.1080/01431169408954239.

Ben Bouallegue, Z., Haiden, T., Weber, N. J., Hamill, T. M., & Richardson, D. S. (2020). Accounting for Representativeness in the Verification of Ensemble Precipitation Forecasts, *Monthly Weather Review* , *148* (5), 2049-2062. Retrieved May 5, 2021, from https://journals.ametsoc.org/view/journals/mwre/148/5/mwr-d-19-0323.1.xml

Bulgin, C. E., Embury, O. and Merchant, C. J. (2016), Sampling uncertainty in gridded sea surface temperature products and Advanced Very High Resolution Radiometer (AVHRR) Global Area Coverage (GAC) data.*Remote Sensing of Environment, 117* , 287-294. doi:10.1016/j.rse.2016.02.021.

Desroziers G, Berre L, Chapnik B, Poli P. 2005. Diagnosis of observation, background and analysis-error statistics in observation space. *Q. J. R. Meteorol* . Soc. 131 3385–3396.

Dodd, E. M. A, Merchant, C. J., Rayner, N. A., and Morice, C. P. (2015), An investigation into the impact of using various techniques to estimate Arctic surface air temperature anomalies. *Journal of Climate, 28* , 1743-1763.

Good, S. A. (2016), The impact of observational sampling on time series of a 0-700m ocean average temperature: a case study. *International Journal of Climatology, 37* (35), 2260-2268. doi:10.1002/joc.4654.

Good, S. A., Fiedler, E., Mao, C., Martin, M. J., Maycock, A., Ried, R., Roberts-Jones, J., Searle, T., Waters, J., While, J., Worsfold, M. (2020), The current configuration of the OSTIA system for operational production of foundation sea surface temperature and ice concentration analyses. *Remote Sensing, 12* , 720.

Hollingsworth A, Lönnberg P. 1986. The statistical structure of short-range forecast errors as determined from radiosonde data. Part I: The wind field *Tellus 38A* 111–136.

Janjić, T. and Cohn, S. E. (2006) Treatment of observation error due to unresolved scales in atmospheric data assimilation. *Monthly Weather Review* , 134, 2900–2915.

Janjić, T, Bormann, N, Bocquet, M, Carton, JA, Cohn, SE, Dance, SL, Losa, SN, Nichols, NK, Potthast, R, Waller, JA, Weston, P. (2018) On the representation error in data assimilation, *Q J R Meteorol Soc* .; 144: 1257– 1278. https://doi.org/10.1002/qj.3130

JCGM 100:2008 Guide to the expression of Uncertainty in Measurement, GUM (Sevres, Paris: BIPM) p 100 (www.bipm.org)

L'Ecuyer, T. S., et al. (2015) The observed state of the energy budget in the early twenty-first century. *J. Climate* , 28, 8319–8346, https://doi.org/10.1175/JCLI-D-14-00556.1

Merchant, C. J., Saux-Picart, S., & Waller, J. (2020). Bias correction and covariance parameters for optimal estimation by exploiting matched in-situ references. Remote Sensing of Environment, 237, 111590.

Met Office (2019). IMPROVER: Probabilistic meteorological post-processing , .

Mittaz J, Merchant CJ, Woolliams ER. 2019. Applying principles of metrology to historical Earth observations from satellites.*Metrologia* , 56, 032002. https://doi.org/10.1088/1681-7575/ab1705

Oke, P. R. and Sakov, P. (2008) Representation error of oceanic observations for data assimilation. *Journal of Atmospheric and oceanic technology* , 25, 1004–1017

Povey, A.C. and Grainger, R.G. (2019) Toward More Representative Gridded Satellite Products. *IEEE Geoscience and Remote Sensing Letters* , 16(5), 672-676, doi: 10.1109/LGRS.2018.2881762.

Rodell, M., et al. (2015) The observed state of the water cycle in the early twenty-first century. *J. Climate* , **28** , 8289–8318, https://doi.org/10.1175/JCLI-D-14-00555.1

Saunders, R., Hocking, J., Turner, E., Rayer, P., Rundle, D., Brunel, P., Vidot, J., Roquet, P., Matricardi, M., Geer, A., Bormann, N., and Lupu, C. (2018) An update on the RTTOV fast radiative transfer model (currently at version 12), *Geosci. Model Dev* ., 11, 2717–2737, https://doi.org/10.5194/gmd-11-2717-2018.

Schutgens, N. A. J., Gryspeerdt, E., Weigum, N., Tsyro, S., Goto, D., Schulz, M. and Stier, P. (2016) Will a perfect model agree with perfect observations? the impact of spatial sampling. *Atmospheric Chemistry and Physics* , 16, 6335–6353

Stier, P. (2016) Limitations of passive satellite remote sensing to constrain global cloud condensation nuclei. *Atmospheric Chemistry and Physics,* 16(10), 6595-6607, doi:10.5194/acp-16-6595-2016.

Thomas, C. M., Dong, B., & Haines, K. (2020). Inverse Modeling of Global and Regional Energy and Water Cycle Fluxes using Earth Observation Data, *Journal of Climate* , *33* (5), 1707-1723, https://doi.org/10.1175/JCLI-D-19-0343.1.

Van Weele, M. (editor). GAIA-CLIM report, deliverable 6.11, Gaps Assessment and Impacts

Document. Available at: http://www.gaia-clim.eu/page/gaid .

Waller, J.A., Dance, S.L. and Lean, H.W. (2021), Evaluating errors due to unresolved scales in convection permitting numerical weather prediction. *Q J R Meteorol Soc* , *147* , 2657-2669. https://doi.org/10.1002/qj.4043

C. E. Bulgin[1,2], C. M. Thomas[3], J. A. Waller[3] and E.R. Woolliams[4]

[1]University of Reading, Reading, UK, [2]National Centre for Earth Observation, Leicester, UK, [3]UK Met Office, Exeter, UK, [4]National Physical Laboratory, Teddington, UK.

Corresponding author: Claire Bulgin (c.e.bulgin@reading.ac.uk)

Key Points:

- Representation uncertainty can be defined as occurring at a comparison interface between two quantities.
- A common definition of representation uncertainty can be applied across the Earth Science community.
- Collaboration between scientists and metrologists is recommended to define an accessible and relevant vocabulary for uncertainty analysis.

Abstract

The first Joint Workshop on Representation Uncertainty in the Earth Sciences was held in March 2021. This brought together the Earth observation, data assimilation and forecast verification and post-processing communities, alongside metrologists to discuss the definition and quantification of representation uncertainty within the Earth Sciences. The aim of the workshop was to facilitate cross-disciplinary discussion, establishing where existing methodologies could be shared and to foster future collaboration. A key outcome of the workshop was a working definition of representation uncertainty applicable across all the Earth Science communities, which is presented in this white paper. The cross-disciplinary discussions at the workshop highlighted the need for scientists to work with metrologists to establish a common vocabulary for uncertainties, accessible to Earth Science applications. Further recommendations included regular workshops to discuss progress in defining and quantifying representation uncertainty and awareness of cross-disciplinary funding opportunities to further address representation uncertainty issues.

Plain Language Summary

This paper describes a workshop which brought together experts from different Earth science disciplines to discuss and attempt to define the term “representation uncertainty”. We make observations of the Earth using satellites, ground based instruments (such as weather stations) and instruments on ocean buoys or aircraft. These observations are used in their own right, and also by computer models used to generate weather forecasts. The observations themselves are imperfect and we quantify these imperfections using the term “uncertainty”. In this paper we discuss the uncertainty that happens when we compare two different sets of observations, two different models, or observations and models. As well as the uncertainties inherent in models and observations, there is also an uncertainty due to the fact that the two things being compared are not representing exactly the same conditions. For example, a satellite observation may represent an average value over a few hundred metres, while an instrument on the surface measures only at a single point, and the model represents an area of several kilometres. Understanding those differences is essential to be able to properly combine different sets of observations, and observations with models.

1 Introduction

The term "representation uncertainty" is used widely in the Earth Sciences, often to describe an uncertainty that occurs at the comparison interface of two different representations of the same physical quantity, although each subcommunity has its own ways of characterizing and evaluating these quantities. The first Joint Workshop on Representation Uncertainty in the Earth Sciences was held online from 23rd to 25th March 2021, sponsored by the National Centre for Earth Observation. This discussion-based workshop was open to scientists working in the fields of Earth observation (EO), data assimilation (DA), forecast verification and post-processing (FVPP), and metrology. The workshop was designed to maximize discussion time, beginning with short presentations on where representation uncertainty might occur within the Earth sciences and introducing the field of metrology, which underpins the science of uncertainty. Discussion groups were held first within each discipline, and then across the different disciplines, posing the questions 'Where are we now?' and 'What could we do going forward?'. Science talks and poster contributions were made, showcasing ongoing research towards quantifying representation uncertainty. The workshop concluded with a panel discussion summarizing the key outcomes identified within each discipline. The workshop promoted consistent use of terminology across all disciplines to facilitate discussion.

2 Workshop Outcomes

2.1 Defining representation uncertainty

One of the main challenges of the meeting was to find a definition of representation uncertainty that was widely applicable across the disciplines represented at the workshop. The 'working definition' was regularly revisited throughout the meeting and during the concluding panel discussion, the following definition was proposed:

> *"The uncertainty associated with a comparison of two quantities (that are themselves uncertain). This contribution to the uncertainty only occurs at the comparison interface."*

Representation uncertainty is recognised as a wider term that includes many different sources of uncertainty, each of which is often described by a more specific term. This broad definition, examples of which are shown in Figure 1, should be applicable across all disciplines, as it allows the detail of how representation uncertainty arises to be different for each comparison, which we acknowledge is often complex.

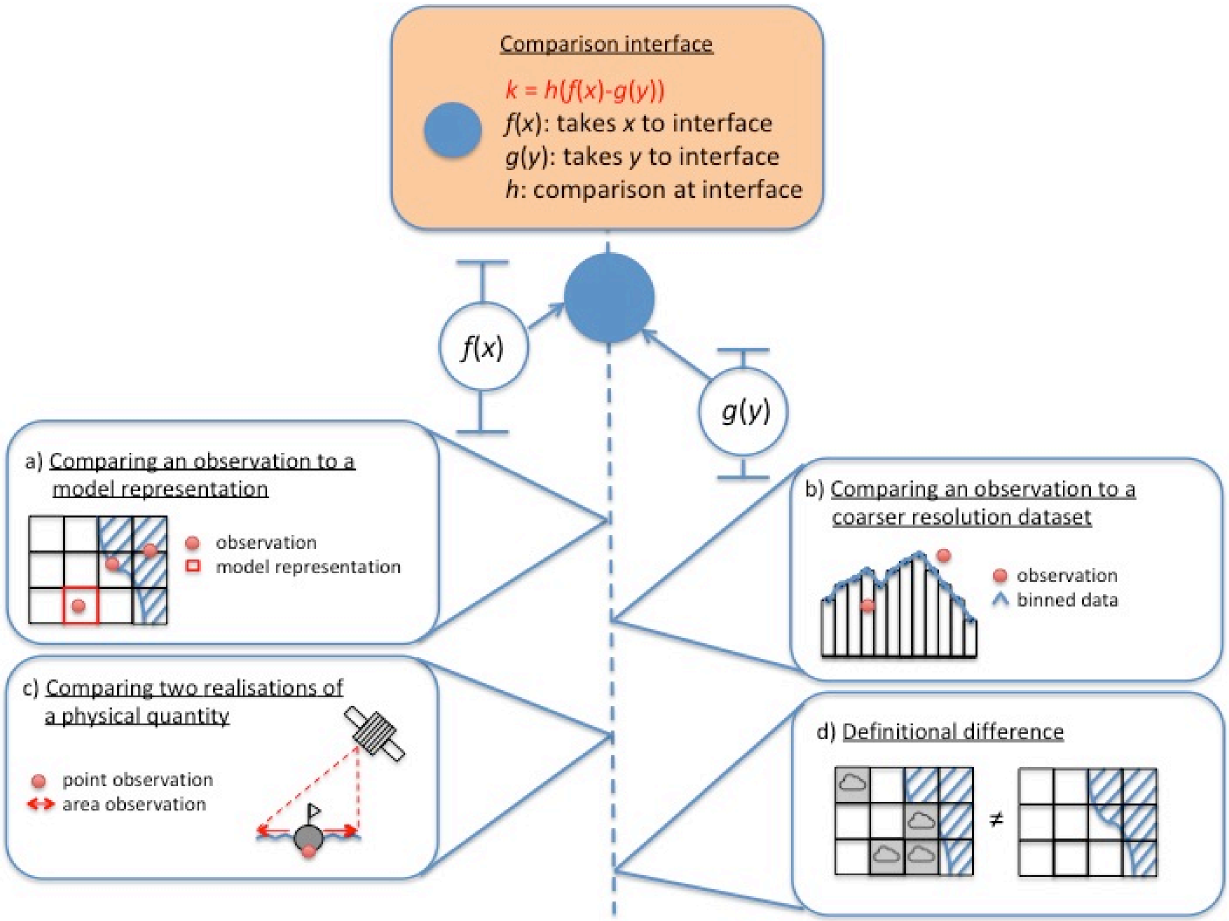

Figure 1: Schematic detailing examples of the comparison interface at which representation uncertainty occurs. At this interface, k is the result of a comparison of observed or modelled quantities, x or y. These quantities are transformed through the functions $f(x)$ and $g(y)$ to bring them to the comparison interface. The comparison is performed as some function of the difference between the transformed quantities, producing the result $k = h\left(f(x) - g(y)\right)$. Both $f(x)$ and $g(y)$ have associated uncertainty and at the point of comparison, there is also the representation uncertainty, identified with the uncertainty on k at the comparison interface. The details of the representation uncertainty are specific to a given comparison, and may arise from a) comparison of an observation to a geophysical model equivalent, b) comparison of an observation to a coarser resolution data set, c) comparison of two realisations of the same physical quantity or d) a definitional difference (which could be a comparison of a measured or modelled data set to a conceptual, more complete, data set or the difference between what users would like and what data producers can provide). The

schematics within each box provide just a single example of the possible comparisons.

2.2 Metrology and uncertainty in Earth Science applications

Metrology ensures international consistency and century-long stability of measurements for both science and trade by providing traceability to the International System of Units, the SI. More recently, metrologists have also started to apply the principles of metrological traceability to provide confidence in data derived from modelling. Metrological traceability is built on two core concepts: uncertainty analysis, defined by the Guide to the Expression of Uncertainty in Measurement (the GUM; JCGM 100:2008), and comparisons, formalized to validate uncertainty assessments. Metrologists have worked with the EO community for several decades but have had limited interactions with the DA and FVPP communities.

The Earth sciences often have data value chains, where one community's output becomes another community's input. Each stage of the chain, whether based on instruments or modelling, can be described by a measurement model (equation) with terms that include the previous stage's output, and new terms introduced. Additionally, there will be assumptions inherent in the form of the model which Mittaz et al. (2019) described as the "plus zero uncertainty", written by including a +0 into the measurement model. In laboratory metrology, this +0 includes concepts such as representing an integral with a trapezoidal-rule summation, or assuming that the instrument does not change its properties when moved into the field.

Thus there are two types of uncertainty associated with a value calculated by a measurement model. First, the propagated uncertainty comes from the input quantities, through the GUM methods, to an uncertainty associated with the calculated value. This describes the spread of probable values around the measured value where the true calculated value is expected to lie. Then there is the uncertainty associated with the extent to which the true calculated value represents the quantity of interest – the +0 term.

2.3 Current status within each research discipline

In this section we include some examples of where representation uncertainty occurs within the different disciplines and relate these back to the definition given in Section 2.1, using the shorthand notation ('a'), ('b'), ('c') and ('d') to refer to the four types of representation uncertainty as defined in Figure 1.

2.3.1. Earth Observation (remotely sensed and in situ)

Earth observation includes both in situ and remote sensing measurements., There are "plus zero uncertainties" identified in EO data production that represent a conceptual comparison between the evaluated quantity and the desired quantity ('d'), or a definitonal difference between what users want and what data producers can provide ('d'). Representation uncertainty can also occur at the comparison interface between two datasets, for example ground-based and satellite observations ('b' and/or 'c'), where representation uncertainty arises both from making point-to-area average comparisons, and because in situ data are typically sparse; hence error statistics from comparisons with satellite data are not necessarily globally representative. Error statistics may also mask data complexity ('d', Povey and Grainger, 2019). Representation uncertainty is also inherent in the extent to which the measured variable matches the target geophysical quantity, either due to imperfections in the measurement equation or user requirements ('d', Stier, 2016).

Considering in-situ measurements, representation uncertainty often occurs when a continuous data field in time or space is inferred from discrete measurements. This uncertainty may change temporally as sampling capacity for global measurements increases (Good, 2016). When gap-filling data to provide spatially complete data fields, representation uncertainty can be introduced both by the choice of interpolation method (Dodd, 2015) and by the extent to which the available data represent areas that are not observed. Work presented at the workshop highlighted the challenges in calculating an uncertainty budget where the appropriate representation uncertainty has a dependence on what the user wants to observe: point measurements, spatial or temporal averages, anomalies or trends.

In remote sensing data, representation uncertainty can occur as a result of instrument sampling ('d'), and gridding data ('d' or 'b'). For some data streams, a sub-sample of the full resolution satellite data is passed to the ground receiving station introducing sampling uncertainties in the Level 1 data (Belward et al., 1994). Representation uncertainty is also common at Level 3 when regularly gridding clear-sky only data; these products will include a sampling uncertainty when compared to all-sky observations (Bulgin et al, 2016). Further sources of representation uncertainty are associated with Level 4 products involving data composites from different instruments with different satellite overpass times ('c') (Good et al., 2020).

2.3.2. Data Assimilation

Data assimilation is the process by which observations are combined with model data, weighted by their respective uncertainties and accounting for physical constraints, in order to provide an optimised estimate of a geophysical state. Often the observations represent different variables from those modeled, so the model space must be mapped into observation space using an observation operator. There are two main comparison interfaces where representation uncertainty may arise: comparison of an observation with its model counterpart ('a'), and, if a processed form of an observation is assimilated, the comparison of this processed observation to the conceptual perfectly processed observation ('d') (Janjic et al., 2018).

There are two main contributors to representation uncertainty at the observation-model interface. First is the error due to unresolved scales and processes, which arises when the observations represent different spatial and temporal scales than those of the assimilating model. Second is the observation operator error arising when the observation operator is approximated either to reduce computational complexity and cost or because of unknown parameters and processes. At the interface between the observation and its processed form, representation uncertainty will arise due to the errors introduced and propagated through the processing chain, sometimes considered as a measurement uncertainty attributed to the assimilated observation. Observations may also be subject to quality control procedures; inaccuracy or occasional failure of these can be an additional source of uncertainty.

Representation uncertainty is most commonly accounted for in DA through including the representation error covariance matrix, $\boldsymbol{F}$, along with instrument error statistics, $\boldsymbol{E}$, in the observation error covariance matrix, $\boldsymbol{R} = \boldsymbol{E} + \boldsymbol{F}$. For this reason, methods to estimate the full $\boldsymbol{R}$ matrix, where the measurement error covariance matrix is known, have been commonly used to isolate representation uncertainty (Desroziers et al (2005), Hollingsworth and Lönnberg (1986)). Alternatively, individual sources of representation uncertainty can be estimated from the error statistics of a comparison between two values: e.g. either two representations of a variable or using two observation operators with differing levels of approximation (Schutgens et al., 2016; Saunders et al., 2018, Waller et al., 2021). Although it is most common for the representation error to be included in $\boldsymbol{R}$, work presented at the workshop showed that other methods exist that account for uncertainties via updates in small-scale background uncertainties and model uncertainties (e.g. Janjic 2006; Bell et al., 2020). These approaches highlight the difficulty in separating representation uncertainties from other types of uncertainty inherent in the DA process.

2.3.3. Forecast verification and post-processing

Forecast post-processing attempts to correct, combine and exploit the information contained within existing forecasts to produce optimal products for dissemination to the public and other customers. The related field of forecast verification quantifies the success of a forecast by comparing its predictions to observations independent of those predictions. Both of these fields rely on the comparison of quantities at different spatial and temporal scales, and hence will have representation uncertainties associated with various steps in the processing chain. In addition, post-processing must often provide outputs at different scales depending on the user requirements (‘d’).

As discussed in the workshop, the representation uncertainty related to the difference between point observations and grid values (‘a’) is an important topic in FVPP; scale mismatch uncertainties can occur due to sub-grid variability that is not modelled. A new probabilistic post-processing system incorporating verification, the Integrated Model post-PROcessing and VERification (IMPROVER, Met Office, 2019), considers this representation uncertainty. Roberts et al. (pers. comm.) have demonstrated the importance of accounting for local topography when producing forecasts. Topography may not be well represented in a relatively coarse model, but accounting for this in the post-processing step can improve forecast skill. Adjusting the rate of change of temperature with height can also provide better agreement between point and grid values. Ben Bouallegue et al. 2020 used a statistical parameterisation to quantify the representation uncertainty related to the difference between point and grid values. The results can be used in ensemble verification and to represent sub-grid variability that is not present in the model.

A further example of representation uncertainty in FVPP relates to differences in the spatial position between observations of local weather phenomena and the forecast equivalent. This is in distinction to scale mismatch uncertainties; a meteorological feature could be modelled to extremely high precision but be located in the incorrect position ('c'). This is mitigated in FVPP by using neighbourhood methods and forecast ensembles.

2.4 Existing collaborations between disciplines

The Earth sciences are multidisciplinary, with one community's output often being another community's input. While different communities perform their own uncertainty analysis (to differing levels of formality), uncertainties may not be fully transferred and representation uncertainty generated at a comparison interface, may not be fully considered. Projects such as GAIA-CLIM have attempted to address these gaps by bringing the EO, DA and metrology communities together. Some of the residual-based methods commonly used in DA have begun to be used by the EO community. For example, Merchant et al (2020) applied the diagnostics of Desroziers et al. (2005), along with additional bias correction, to estimate error covariance parameters for SST retrievals.

Another example is the combination of several EO products (remote sensing and/or in situ) to quantify uncertainties in the Earth's energy and water cycles with inverse modelling methods (L'Ecuyer et al. 2015, Rodell et al. 2015, Thomas et al. 2020). The outputs of the inverse modelling procedure can potentially be used to evaluate the accuracy of Global Climate Model products, providing an opportunity to collaborate with the modelling community. For this evaluation to be effective, the uncertainties, including any representation uncertainties, must be accurately determined.

3 Future opportunities for collaboration and community requirements

One barrier to effective collaboration on uncertainty in the Earth sciences is communication, particularly where similar words take different meanings in different groups. Integrating metrologists into this multidisciplinary community can help to bridge this gap, clarifying vocabulary and the distinction between terms such as "error" and "uncertainty" (Mittaz et al, 2019). Regular communication on the definition, sources, quantification and mitigation of representation uncertainty will lead to more efficient transfer of information, better inventory of uncertainty (by considering other perspectives) and more traceability. Ideally the various communities will formulate a common language that can be used consistently in published papers. We can draw parallels with the adoption of the notation in Ide et al 1997 by the DA community.

It was identified in the workshop that collaboration between the DA and FVPP communities could be beneficial, using statistical methods such as the Deroziers et al (2005) diagnostics to account for temporal uncertainties in FVPP. These uncertainties are important given the serial correlation present in many atmospheric and ocanic phenomena. Furthermore, the acquisition of crowdsourced data to provide a 'ground truth' for forecast verification was discussed. Such data are also of interest to the DA community but in order to use these data effectively, extensive quality control and validation would be required. Another example is the use of dynamical-model-dependent propagation of uncertainties for daily satellite products. The ability to apply different methods to a single scenario would provide multiple estimates of representation uncertainty and could help provide confidence in those estimates.

The workshop has begun information dissemination between the Earth science disciplines and future workshops could benefit from inclusion of model developers/validators, instrument manufacturers, and space agencies. Model developers understand which scales and processes are poorly represented by models and may be able to aid quantification of representation uncertainties that are large when data are assimilated. Instrument manufacturers could also create common methods for documenting instrument calibration and data-processing procedures to optimise data use, where commercial sensitivities allow.

4 Recommendations

We encourage all the communities represented at the workshop to continue to work towards a mutually agreed understanding of representation uncertainty, starting with the definition presented in this paper, and acknowledging in many cases more specialised terms will be used to characterise the individual component source(s) of uncertainty. Many of the discussions in the workshop highlighted the need to continue this conversation in a multidisciplinary format, perhaps facilitated by regular workshops, expanded to include other relevant communities. In continuing this conversation, the development of a consistent vocabulary on uncertainties that is relevant to the Earth Sciences community is essential. Such a vocabulary would be supported by metrologists to aid communication and underpin future discussions and collaborations. A cross-community paper developing the ideas outlined in this white paper could be the first target for such future collaboration and could define both a formal vocabulary and make specific recommendations for collaborative research. As funding is often a barrier to inter-disciplinary research, the establishment of a community awareness of projects where representation uncertainty issues may be addressed would be beneficial and "sandpit events" (where scientists work together in a focused way to answer a specific challenge) may also be useful to foster in-depth consideration of inter-disciplinary knowledge transfer on representation uncertainty.

Acknowledgments, Samples, and Data

We acknowledge the many excellent contributions from all workshop participants. In particular we would like to thank the panel, the speakers, session chairs and rapporteurs and the poster authors. We would also like to thank Maurice Cox, John Eyre, Paul Green, Keith Haines, Christopher Merchant, Jonathan Mittaz, Simon Pinnock, Adam Povey, Nigel Roberts and Bruce Wright for their constructive feedback on the white paper draft.

None of the authors has expressed any conflict of interest.

Funding acknowledgements: The organizing committee would like to thank the National Centre for Earth Observation (NCEO), which sponsored this event, providing technical and organizational support throughout from grant NE/R016518/1 of the Natural Environment Research Council. Emma Woolliams acknowledges funding in part from the Instrument Data quality Evaluation and Assessment Service - Quality Assurance for Earth Observation (IDEAS-QA4EO) contract funded by ESA-ESRIN (n. 4000128960/19/I-NS) and in part from UK Government's Department for Business, Energy and Industrial Strategy (BEIS) through the UK's National Measurement System programmes.

Data Availability Statement

This article reports proceedings of a discussion-based online workshop. There are no associated datasets to be made available in conjunction with this commentary.

References

Bell, Z., Dance, S.L., & Waller, J.A. (2020) Accounting for observation uncertainty and bias due to unresolved scales with the Schmidt-Kalman filter, *Tellus A: Dynamic Meteorology and Oceanography*, 72:1, 1-21, doi: 10.1080/16000870.2020.1831830

Belward, A. S., Kennedy, P. J. and Grégoire, J. M. (1994), The limitations and potential of AVHRR GAC data for continental scale fire studies. *International Journal of Remote Sensing, 15*(11), 2215-2234. doi:10.1080/01431169408954239.

Ben Bouallegue, Z., Haiden, T., Weber, N. J., Hamill, T. M., & Richardson, D. S. (2020). Accounting for Representativeness in the Verification of Ensemble Precipitation Forecasts, *Monthly Weather Review*, *148*(5), 2049-2062. Retrieved May 5, 2021, from https://journals.ametsoc.org/view/journals/mwre/148/5/mwr-d-19-0323.1.xml

Bulgin, C. E., Embury, O. and Merchant, C. J. (2016), Sampling uncertainty in gridded sea surface temperature products and Advanced Very High Resolution Radiometer (AVHRR) Global Area Coverage (GAC) data. *Remote Sensing of Environment, 117*, 287-294. doi:10.1016/j.rse.2016.02.021.

Desroziers G, Berre L, Chapnik B, Poli P. 2005. Diagnosis of observation, background and analysis-error statistics in observation space. *Q. J. R. Meteorol.*

Soc. 131 3385–3396.

Dodd, E. M. A, Merchant, C. J., Rayner, N. A., and Morice, C. P. (2015), An investigation into the impact of using various techniques to estimate Arctic surface air temperature anomalies. *Journal of Climate, 28*, 1743-1763.

Good, S. A. (2016), The impact of observational sampling on time series of a 0-700m ocean average temperature: a case study. *International Journal of Climatology, 37*(35), 2260-2268. doi:10.1002/joc.4654.

Good, S. A., Fiedler, E., Mao, C., Martin, M. J., Maycock, A., Ried, R., Roberts-Jones, J., Searle, T., Waters, J., While, J., Worsfold, M. (2020), The current configuration of the OSTIA system for operational production of foundation sea surface temperature and ice concentration analyses. *Remote Sensing, 12*, 720.

Hollingsworth A, Lönnberg P. 1986. The statistical structure of short-range forecast errors as determined from radiosonde data. Part I: The wind field *Tellus 38A* 111–136.

Janjić, T. and Cohn, S. E. (2006) Treatment of observation error due to unresolved scales in atmospheric data assimilation. *Monthly Weather Review*, 134, 2900–2915.

Janjić, T, Bormann, N, Bocquet, M, Carton, JA, Cohn, SE, Dance, SL, Losa, SN, Nichols, NK, Potthast, R, Waller, JA, Weston, P. (2018) On the representation error in data assimilation, *Q J R Meteorol Soc.*; 144: 1257– 1278. https://doi.org/10.1002/qj.3130

JCGM 100:2008 Guide to the expression of Uncertainty in Measurement, GUM (Sevres, Paris: BIPM) p 100 (www.bipm.org)

L'Ecuyer, T. S., et al. (2015) The observed state of the energy budget in the early twenty-first century. *J. Climate*, 28, 8319–8346, https://doi.org/10.1175/JCLI-D-14-00556.1

Merchant, C. J., Saux-Picart, S., & Waller, J. (2020). Bias correction and covariance parameters for optimal estimation by exploiting matched in-situ references. Remote Sensing of Environment, 237, 111590.

Met Office (2019). IMPROVER: Probabilistic meteorological post-processing , .

Mittaz J, Merchant CJ, Woolliams ER. 2019. Applying principles of metrology to historical Earth observations from satellites. *Metrologia*, 56, 032002. https://doi.org/10.1088/1681-7575/ab1705

Oke, P. R. and Sakov, P. (2008) Representation error of oceanic observations for data assimilation. *Journal of Atmospheric and oceanic technology*, 25, 1004–1017

Povey, A.C. and Grainger, R.G. (2019) Toward More Representative Gridded

Satellite Products. *IEEE Geoscience and Remote Sensing Letters*, 16(5), 672-676, doi: 10.1109/LGRS.2018.2881762.

Rodell, M., et al. (2015) The observed state of the water cycle in the early twenty-first century. *J. Climate*, **28**, 8289–8318, https://doi.org/10.1175/JCLI-D-14-00555.1

Saunders, R., Hocking, J., Turner, E., Rayer, P., Rundle, D., Brunel, P., Vidot, J., Roquet, P., Matricardi, M., Geer, A., Bormann, N., and Lupu, C. (2018) An update on the RTTOV fast radiative transfer model (currently at version 12), *Geosci. Model Dev.*, 11, 2717–2737, **https://doi.org/10.5194/gmd-11-2717-2018**.

Schutgens, N. A. J., Gryspeerdt, E., Weigum, N., Tsyro, S., Goto, D., Schulz, M. and Stier, P. (2016) Will a perfect model agree with perfect observations? the impact of spatial sampling. *Atmospheric Chemistry and Physics*, 16, 6335–6353

Stier, P. (2016) Limitations of passive satellite remote sensing to constrain global cloud condensation nuclei. *Atmospheric Chemistry and Physics,* 16(10), 6595-6607, doi:10.5194/acp-16-6595-2016.

Thomas, C. M., Dong, B., & Haines, K. (2020). Inverse Modeling of Global and Regional Energy and Water Cycle Fluxes using Earth Observation Data, *Journal of Climate*, *33*(5), 1707-1723, https://doi.org/10.1175/JCLI-D-19-0343.1.

Van Weele, M. (editor). GAIA-CLIM report, deliverable 6.11, Gaps Assessment and Impacts

Document. Available at: http://www.gaia-clim.eu/page/gaid .

Waller, J.A., Dance, S.L. and Lean, H.W. (2021), Evaluating errors due to unresolved scales in convection permitting numerical weather prediction. *Q J R Meteorol Soc*, *147*, 2657-2669. https://doi.org/10.1002/qj.4043