



The Importance of Scale in the Definition of Uncertainties: How do we Best Communicate this to Data Users?

Claire E. Bulgin^{1,2} · Paul Green³ · Alexander Gruber⁴ · Claire Macintosh⁵ ·
Jonathan Mittaz¹ · Ashley Ramsay³ · Nick A. Rayner⁶

Received: 16 December 2024 / Accepted: 15 September 2025
© The Author(s) 2025

Abstract

Climate services often require observational climate data to inform decision-making on mitigation and adaptation activities. Understanding the uncertainties in the climate datasets that are used for this purpose, and how these uncertainties relate to the context of the climate service is critical to making well-informed decisions. Recent developments in the production of climate-relevant satellite datasets has focused on characterising uncertainties from a bottom-up perspective with a high degree of mathematical rigour. Using the example of three essential climate variables: sea surface temperature, soil moisture and carbon dioxide we discuss how to translate the highly-detailed uncertainty information provided with high-resolution datasets into something appropriate to the scale of a climate service, where the decision-making context might be local, regional or global. Close engagement between climate data producers and climate service providers is essential to ensure we have the best possible platform to make decisions as we adapt to climate change.

Keywords Uncertainties · Scale · Climate applications · Decision making

1 Introduction

In order to draw meaningful conclusions when using remote sensing data (or any form of numerical data), the user needs to know the degree of confidence the data provider has in their data product and the degree to which it can be trusted in their application. At the most fundamental level, whether or not a particular source of information is suitable for use in a particular decision context can be the first question asked. In many cases, the plausible range of the information provided is the key information needed. This level of confidence in the data is crucial to helping data users decide on the appropriate responses to the outcomes of their data analysis (Gruber et al. 2024; Zeng et al. 2019; Yang et al. 2022). This manuscript focuses on the use of Earth observation data, often in the form of long-term climate data records (CDRs) as applied to decision making in the context of climate services and climate modelling.

Climate services is a broad term that spans a continuum between the provision of data products alone, through to services which have a specific decision-making context

in mind (within which users would ideally have engaged in co-developing the service) (Nightingale et al. 2018). In the decision-making space, climate information and its uncertainty is typically just one part of the information needed, and it is rare for a service to be based directly on climate information alone (although such services do exist, e.g. drought payouts by the Africa Risk Capacity). CDRs will likely therefore be used in combination with climate model predictions (e.g. annual, seasonal or decadal forecasts) and projections (e.g. on 50–100-or-more-year timescales), together with non-climate information. Uncertainties in CDRs help to put uncertainties in the forecasts or future climate projections into context, or to constrain those uncertainties through careful comparison, e.g. in the case of climate sensitivity, (Forster et al. 2021) where improved uncertainty estimates have been instrumental in the constraint of climate sensitivity estimates derived from observations. Recent studies have looked at the applicability of CDRs to particular climate applications (Roebeling et al. 2025) but focus primarily on the signal autocorrelation in homogenized gap-filled datasets or selection of appropriate products from existing databases (Dee et al. 2024). In this article our focus is on the implications of scale for product uncertainties, the relative importance of different error sources and considerations for the rescaling of data to different spatiotemporal resolutions for use in a climate context.

Decision making can have high financial stakes: e.g. do we build more flood defences to protect against extreme weather? Should we routinely install air conditioning in office buildings or shopping centres due to rising temperatures? Would urban areas be cooler with more roof-top gardens and green spaces? Decision making also has an impact on human lives and livelihoods: e.g. if we don't build more flood defences, will our agricultural industry suffer significant losses in livestock and crops? If we don't install air conditioning in new office buildings will these become uninhabitable in hot weather, decreasing productivity? If we don't think more sustainably about urban form and green spaces, will we increase the morbidity of vulnerable people groups such as the elderly and very young children? High-stakes decision making requires the best possible understanding of the confidence in the data used and there is currently often a missing link in translating the uncertainty in remote sensing products as calculated by the data producer to the uncertainty in the data as applied to a specific application or decision-making context.

Historically, uncertainties in climate data records (CDRs) have been provided with reference to another data source, e.g. in situ data, often averaged over broad space and time scales and calculated under the assumption that the reference data were themselves without uncertainty (Le Borgne et al. 2011; Noyes et al. 2006). Within the European Space Agency (ESA) Climate Change Initiative (CCI) programme (Plummer et al. 2018), significant time and effort has been invested in identifying error sources in the measurement and retrieval processes for a number of Essential Climate Variables (ECVs), and quantifying the associated uncertainty in the resultant products, e.g. (Bulgin et al. 2016a; Sayer et al. 2020; Barnoud et al. 2023; Araza et al. 2022; Khvorostovsky et al. 2020; Gruber et al. 2019). The approach across different ECVs has been varied, reflecting the varying maturity of the measurement and retrieval processes.

In the case of ECVs where the uncertainty budget is well-developed, the analysis undertaken has been mathematically rigorous, developed in consultation with the metrology (science of measurement) community. Per-observation uncertainty estimates are provided (Ablain et al. 2019; Dorigo et al. 2023) with a breakdown of the uncertainty components (Bulgin et al. 2016a; Ghent et al. 2019), which has resulted in a very detailed understanding of the data confidence up to the point at which the data provider makes the data products publicly available. The trade-off in this case is that the uncertainty

information provided with these data products is mathematically complex and this can be a barrier to user uptake (Aldred et al. 2023; Good and Veal 2023; Good et al. 2021).

The correct propagation of this uncertainty information requires a high level of mathematical literacy on behalf of the data user and both the desire and time to understand the information provided and how to use it appropriately. Feedback from users has suggested a preference for quality level information or a binary screening process for ‘good/good enough for my purpose’ and ‘bad’ data (Aldred et al. 2023). From a user perspective, this is often a much easier entry point to using ‘good quality’ data as, in practice, many users will take these data and adapt them for their own purposes; regridding at coarser resolution, sub-sampling or combining with other datasets. Ensembles can also be used to provide a measure of uncertainty (Kennedy et al. 2019) although this is more typically used with climate models than Earth observation data (Haughton et al. 2014; Deser 2020).

Mathematically speaking, the reason for providing a breakdown of uncertainty components for a data product is that the correlation length scale of the different error sources is important when propagating the uncertainty into new, derived products. A second important aspect is that the process of deriving these new data products can also add new sources of uncertainty (Mittaz et al. 2019; Bulgin et al. 2016a; Pasik et al. 2023). This manuscript discusses the importance of scale (both spatial and temporal) in the quantification of uncertainties on a given data product, e.g. what happens to the uncertainty if we calculate large-scale averages of the geophysical variable over large space or time scales? Are there sources of uncertainty that arise through the process of making this average? Do the dominant types of uncertainty change? The manuscript illustrates the correct way to propagate the uncertainties defined by the data provider for three case studies: sea surface temperature, soil moisture and carbon dioxide. It then discusses the implications of the findings for how we translate sophisticated understanding of carefully-quantified uncertainty information into higher-level understanding of the suitability of a product for use in a climate services context.

The remainder of this manuscript is arranged as follows: in Sect. 2 we take three example ECVs (sea surface temperature, soil moisture and carbon dioxide) and examine how the magnitude of the different components of the uncertainty budget scale over time and space aggregates of the data. In Sect. 3 we provide an extended discussion, comparing the approaches, scales and magnitudes of uncertainty in the three ECVs considered. We consider the implications of the variability in uncertainty components with scale for modelling, climate services and decision making. In Sect. 4 we provide a forward-look for how we might bridge the gap between data providers with detailed uncertainty information on the one side and data users applying these data in decision-making contexts on the other.

2 Examples of How Uncertainties Change with Scale

2.1 Sea Surface Temperature

All sea surface temperature (SST) products produced as part of the ESA CCI programme include per-datum uncertainties (Bulgin et al. 2016b, a). The data include both the total uncertainty and a breakdown of the total into three different components, characterised by their correlation length scale: independent, structured and common. The independent part of the uncertainty budget is uncorrelated between one SST measurement and the next and includes error sources such as instrument noise (Bulgin et al. 2016a). The structured component relates

to uncertainties that are correlated over synoptic scales, arising from errors in the specification of atmospheric state in the SST retrieval (Bulgin et al. 2016a). The common part of the uncertainty budget is fully correlated over the entire satellite mission, and is related primarily to errors arising from instrument calibration (Bulgin et al. 2016a).

For this illustration we use daily SST products from the Sea and Land Surface Temperature Radiometer (SLSTR) at 0.05° resolution. The daily data are spatially incomplete for two reasons: (1) the satellite does not achieve full global coverage in a 24-hour period and (2) SST retrievals cannot be made where clouds obscure the Earth's surface. As many users require products on coarser spatiotemporal resolutions for their applications, we consider here how the partitioning of uncertainties between independent, structured and common components varies as we average the data over spatial scales of 0.05° , 0.1° , 0.25° , 0.5° , 1.0° and 2.0° at temporal resolutions of daily, 5-day, 10-day and 28-day averages. Note that 0.05° paired with daily is the native resolution of the input data. Given the 'gappy' nature of the input data due to incomplete coverage, scaling to coarser resolutions requires calculation of an additional sampling uncertainty, which forms an additional part of the independent uncertainty component. For a full derivation of this sampling uncertainty please refer to Appendix A as this material is not published elsewhere.

Figure 1 shows the uncertainty component partitioning with scale. Data are split between night (a–c), daytime data at tropical latitudes between 40°S – 40°N (d–f) and mid-to-high latitudes ($<40^\circ\text{S}$ and $>40^\circ\text{N}$, g–i). Data are split in this way as the structured uncertainty component is larger at tropical latitudes due to the nature of the synoptic-scale weather systems compared with mid-to-high latitudes (note that this split does not correspond to the latitudinal split in spatial correlation length scales discussed in Appendix A).

The dominant source of uncertainty at high spatial resolutions is the independent component. As the spatiotemporal scales increase, the dominant source of uncertainty shifts towards the structured component as the independent component averages down. The 'gappy' nature of the input data is important in determining how quickly the structured uncertainties begin to dominate. The majority of coarser resolution grid cells are only partially sampled, resulting in an addition of the sampling uncertainty to the propagated part of the independent component. The second peak in the daytime sampling uncertainty at the 10-day temporal resolution (discussed in Appendix A) is evident in both daytime plots. At night, this feature is absent due to reduced spatial variability in SST then (Fig. 1).

At the maximum spatiotemporal averaging considered here, the ratio of the independent to structured components is $\sim 10/90\%$ at tropical latitudes and $\sim 30/70\%$ at mid-to-high latitudes, reflecting the larger source of structured uncertainties at tropical latitudes related to atmospheric conditions. At night, the reduction in the independent uncertainty component is most obvious with coarsening spatial resolution, although at the coarser spatial scales, the reduction with decreasing temporal resolution is also evident. Sampling uncertainties are lower at night and decrease less rapidly with scale. Consequently, the transition between dominance of the independent to structured uncertainty components is slower.

2.2 Soil Moisture

Uncertainty assessments of remotely-sensed soil moisture products usually focus on the estimation of independent uncertainty on a pixel level. Common uncertainties (i.e., biases) are usually not quantified due to a lack of reference data for absolute soil moisture levels over large areas (Gruber et al. 2020). Estimates of 'independent' uncertainty are usually lumped estimates that do not distinguish independent from structured components. Error

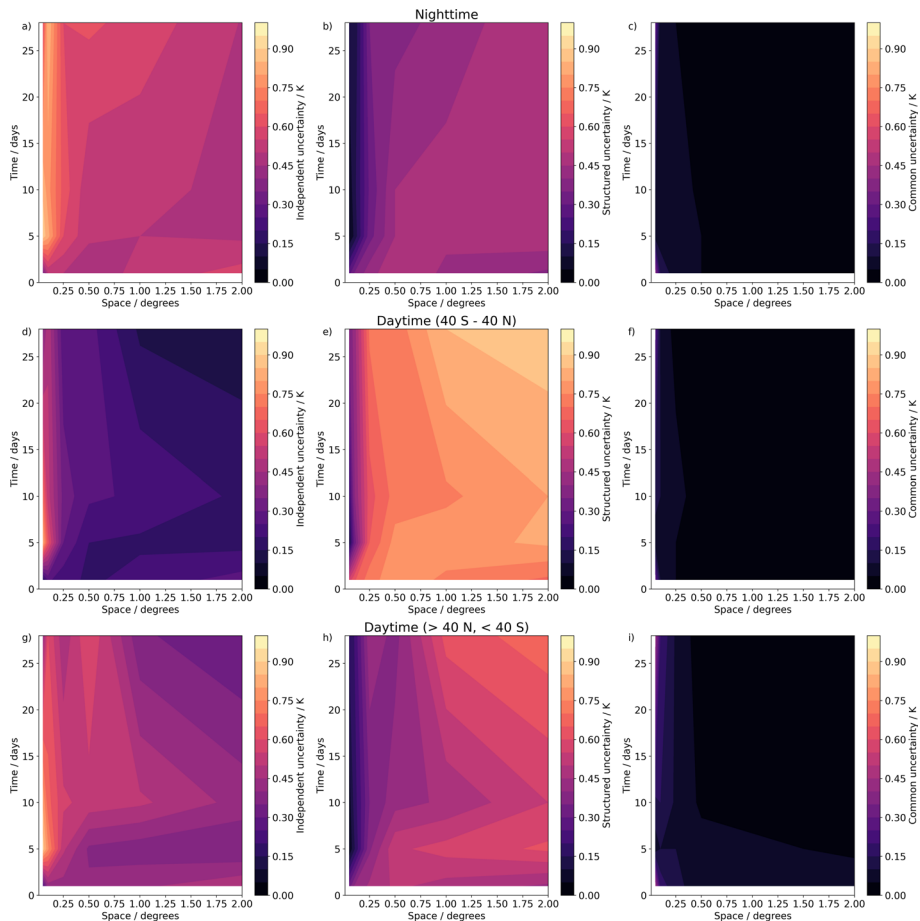


Fig. 1 Spatiotemporal partitioning of SST uncertainty components (independent, structured, common) for nighttime data (a–c), daytime data between 40 S–40N (d–f) and daytime data > 40 N and < 40 S (g–i)

correlations have been found to exist both in time (Zwieback et al. 2013) and space (Zwieback et al. 2018), and need to be considered when pixel-level soil moisture estimates are aggregated. In this example, we demonstrate how such error correlations can be estimated for different soil moisture products, and over which time and length scales these may be important.

Independent uncertainties of global soil moisture products (and other variables) are usually estimated using triple collocation analysis (TCA) (Gruber et al. 2016a). Moreover, several variations to TCA have been proposed to estimate not only the random error variances of individual products, but also error covariances between products (Gruber et al. 2016b) and error auto-covariances in time (Zwieback et al. 2013) and space (Gruber et al. 2015). Here we use these TCA variants to derive spatial and temporal error correlation length scales of two common satellite soil moisture products, ASCAT (H SAF H119 CDR; https://dx.doi.org/10.15770/EUM_SAF_H_0009) and SMAP (SPL2SMP v8; <https://dx.doi.org/10.5067/LPJ8F0TAK6E0>), together with those of a modeled data set, GLDAS-Noah (GLDAS_NOAH025_3H v2.1; <https://dx.doi.org/10.5067/E7TYRXPJKWOQ>). The

SMAP product maps contiguous observations from 40 km resolution radiometer footprints onto the 36 km EASEv2 grid (Brodzik et al. 2012). The ASCAT product maps contiguous observations from 25 km resolution scatterometer footprints onto a 12.5 km discrete global grid (Bartalis et al. 2006). GLDAS uses a contiguous 0.25° regular modelling grid (Rodell et al. 2004). The ASCAT and GLDAS data are matched to the SMAP grid using a nearest-neighbour approach.

TCA estimates independent uncertainty from the (co)variances between three products with mutually uncorrelated errors as:

$$u_i^2 = \sigma_{ii} - \frac{\sigma_{ij}\sigma_{ik}}{\sigma_{jk}} \quad (1)$$

where u_i^2 is the estimated independent uncertainty (error variance) of data set i ; and $\sigma_{..}$ denotes the temporal covariances between the data sets at a given location. Similarly, error auto-covariances of the individual products can be derived as:

$$u_{ii'} = \sigma_{ii'} - \frac{\sigma_{ij'}\sigma_{i'k}}{\sigma_{j'k}} \quad (2)$$

where i' , j' , and k' are the lagged data sets i , j , and k , respectively, shifted by a lag of arbitrary distance; and $u_{ii'}$ is the error auto-covariance for the chosen lag distance. This lag can be chosen in time, in which case $u_{ii'}$ yields an estimate of the temporal auto-covariance, or in space, in which case $u_{ii'}$ yields an estimate of the spatial error auto-covariance. Finally, estimates of error auto-covariances can be converted to estimates of error auto-correlation as

$$R_{ii'}^u = \frac{u_{ii'}}{u_i u_{i'}} \quad (3)$$

Figure 2 shows the estimates of the temporal and spatial error auto-correlation of ASCAT, SMAP, and GLDAS for different lags, averaged over the Contiguous United States (CONUS). Notable error correlations exist both in time and in space. As expected, temporal error correlations in the satellite soil moisture retrievals, which are independent, consecutive observations, drop faster than those in the model simulations, which are dynamically propagated states (e -folding times of roughly 5 days for ASCAT and SMAP as opposed to about 25 days for GLDAS). Somewhat surprisingly, spatial error correlations of ASCAT retrievals and GLDAS simulations are remarkably similar and persist over multiple (0.25°) grid cells whereas those in SMAP retrievals are notably smaller, even though SMAP footprint resolution is substantially coarser (40 km) than that of ASCAT (25 km). ASCAT error correlation length scales are, nevertheless, consistent with those found in previous studies (Gruber et al. 2015).

The existence of these error correlations has important implications, in particular when soil moisture estimates are averaged in time or space. This is done, for example, in drought studies which often look at 10-daily or monthly averages, or climate studies which often work on grids that are coarser than those of observational data (e.g., 0.5°–1°). Averaging soil moisture estimates also reduces their uncertainty, but the greater the correlation between errors, the less efficient this uncertainty reduction becomes. This is illustrated in Fig. 3, which shows the fractional uncertainty reduction when averaging soil moisture observations as a function of spatial and temporal averaging distance. This uncertainty reduction is shown both for an idealized scenario where error correlations are neglected

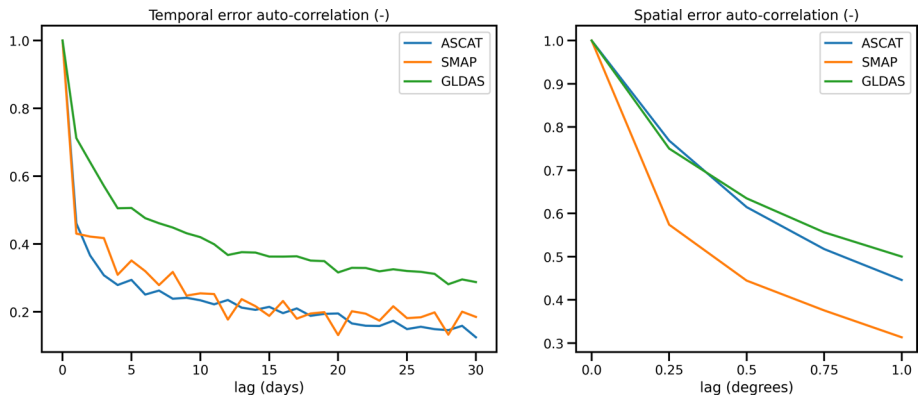


Fig. 2 Temporal (left) and spatial (right) error auto-correlation for different soil moisture products, estimated for different lags

(dashed lines), and for a real-life example when considering the actual error correlations that were estimated for ASCAT, SMAP, and GLDAS in Fig. 2. If errors were fully independent, averaging soil moisture estimates of 30 consecutive days or four neighbouring 0.25° grid cells would—in theory—reduce their uncertainty to near-zero. In reality, however, one can only expect a reduction by about 50 % at best. Note that Figs. 2 and 3 consider the average spatial and temporal error auto-correlation over the entire study domain (the CONUS). However, error correlations can vary significantly across regions depending on land surface and climatic conditions, and should thus be quantified more specifically for each domain of interest.

2.3 Carbon Dioxide

This analysis of column-averaged dry air CO_2 mole fraction (XCO_2) retrieval uncertainty on a range of spatial and temporal scales is based on a bottom-up assessment of the core uncertainty contributions available from the literature. The initial analysis builds on uncertainty quantification for the JAXA Greenhouse Gases Observing Satellite (GOSAT)

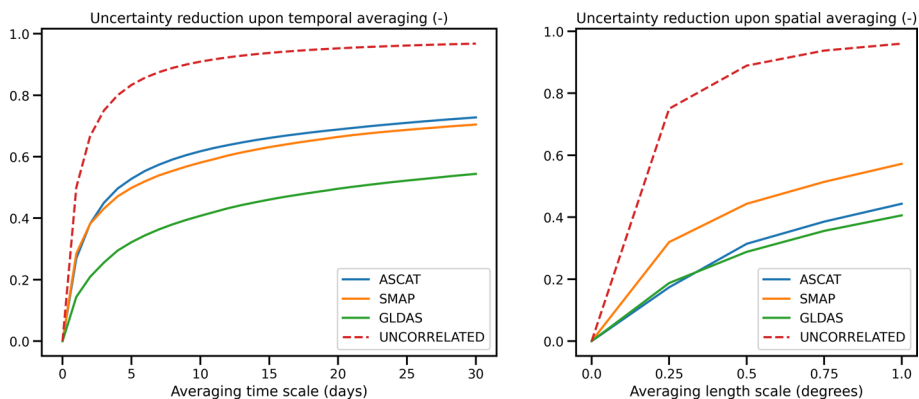


Fig. 3 Fractional uncertainty reduction upon temporal (left) and spatial (right) averaging with and without temporal/spatial error auto-correlation

mission XCO₂ product (Kuze et al. 2009) supplemented by uncertainty estimates from the NASA Orbiting Carbon Observatory (OCO-2) mission (Crisp et al. 2008) particularly to quantify forward model errors when characterising the atmosphere through which the retrieval is made. There are a number of XCO₂ products currently being generated utilising GOSAT, OCO-2 or TanSat (sometimes referred to as CarbonSat) (Ran and Li 2019) data with a variety of retrieval codes, e.g. (O'Dell et al. 2018), (Noël et al. 2021) and (Yang et al. 2020). Consequently, the analysis presented here is devised to detail considerations for a nominal XCO₂ product rather than be specific to one operational product, and it is not the authors' intention to replicate the detailed analysis performed by the respective product teams. The uncertainty estimates used in this analysis are taken from (Cogan et al. 2012; Connor et al. 2016, 2008). As a bottom-up approach, the analysis assumes continuous data availability over the spatial and temporal scales explored. In practice, the ground sampling differs between instruments; GOSAT has a Fourier transform spectrometer aboard, sampling non-contiguous 10 km footprints as projected on the Earth's surface. TanSat and OCO use spectrometers capable of higher-resolution and more continuous XCO₂ retrievals. Between two-thirds and three-quarters of the measurements are then excluded in the quality filtering process, e.g. (Parker et al. 2020) and (Noël et al. 2021) and thus the geographical distribution of the available data is varied. The effect of this sampling uncertainty is not presently defined either by the XCO₂ data providers or the user community and has therefore not been included here.

The correlation length scales in space and time have been estimated from an understanding of the uncertainty component and sampling regime, and typically categorised as a) independent b) structured, where there are short-term correlations (typically more prominent over local spatial extents than temporal extents) or c) common and persistent. Measurement noise is considered independent at a per-pixel level in both space and time dimensions, as is the majority share of the post-processing uncertainty. The common category of uncertainty contributions is dominated by elements of the forward model, including forward model spectroscopy uncertainties and the prior errors in quantities such as CO₂, temperature and aerosol optical depth. Figure 4 estimates the typical combined uncertainty behaviour in space and time from an individual/instantaneous sample to the global scale of 10,000 km, and the temporal scale of multiple years.

The uncertainty behaviour relatively quickly drops from the independent dominated per-pixel peak above 5 ppm falling to values around 3.5 ppm at scales of a few 100s of km and beyond synoptic time scales. At longer spatial and temporal scales all the initial structured effects are treated as independent and the behaviour asymptotes to be dominated by the common terms. A fuller consideration of the sampling uncertainty and geographical structure of quality-filtered measurements would likely inflate the overall uncertainty, particularly at medium scales (a week to months, and 100s km) especially in regions where seasonally-persistent cloudiness prevails.

The combined common term of approximately 2.5 ppm is a little above some of the literature-quoted accuracies of 1–2 ppm when validated against the Total Carbon Column Observing Network (TCCON) e.g. (Yoshida et al. 2013). However, the estimate here is bottom-up and additional correlations that have not been considered here could reduce the overall estimate. Previous studies to reconcile bottom-up approaches to observed uncertainties typically show the bottom-up approach does more often than not overestimate the bias, suggesting there are possible unconsidered correlations at play.

Validation of the XCO₂ product against TCCON sites is an interesting point to address in terms of overall uncertainty considerations. Some products use the ground network for validation, whereas others perform a final bias correction based on comparisons between

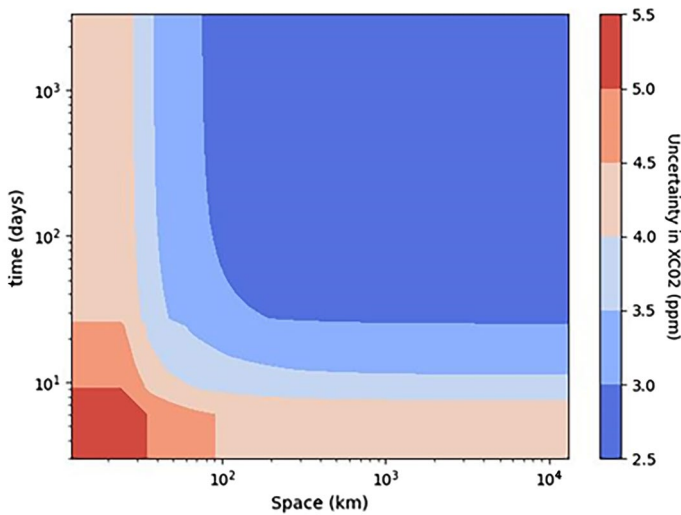


Fig. 4 Typical change in total combined uncertainty in a XCO₂ product (ppm) over lengthened spatial and temporal scales

the satellite and TCCON observations. Both approaches are valid and a bias correction via this method will clearly improve the dominant systematic contribution on the 10,000 km and annual scales over which such comparisons are typically performed. In this case, an additional common uncertainty term would then need to be added to account for the assumptions in the bias correction process and remaining residual.

At a global scale, mass balance considerations also act as a viable constraint on the overall atmospheric CO₂ budget. Depending on how these large spatial scale validation activities are used, it would be conceivable that the overall uncertainty would drop further in the far top right corner of Fig. 4. However, no common approach exists to the authors' knowledge so this is difficult to quantify in a generalised case. Additionally, there will be some variation in errors due to macroscale effects in specific regions of the globe, such as tropical convection that disrupts the smooth decline with scale seen in Fig. 4, but this is again difficult to quantify without a detailed analysis, beyond the scope of this study.

3 Discussion

3.1 Key Messages from the Sea Surface Temperature, Soil Moisture and Carbon Dioxide Examples

The three examples in Sect. 2 illustrate that although a common uncertainty framework can be applied to multiple ECVs, e.g. considering individual uncertainty components and the associated correlations, the question of propagating uncertainties to different spatiotemporal scales is nuanced. Indeed, the degree to which various uncertainty components are considered to be important varies considerably between ECVs. Of the examples presented here: the SST product includes the most rigorous attempt to quantify all uncertainties from all possible error sources (Bulgin et al. 2016a, b; Embury et al. 2024), the soil moisture community is overly optimistic in the treatment

of all uncertainties as independent and the CO₂ community do not directly address the issue of sampling uncertainty when coarsening the data resolution (although this is considered when applied to inversion schemes on a regional scale, e.g. (Chevallier et al. 2014)).

Key factors about the input product to be scaled will determine to some extent the proportion of different uncertainty components with scale. Independent uncertainties in spatially continuous products will rapidly reduce with scale, whilst independent uncertainties in ‘gappy’ data (as in the SST and XCO₂ case) are magnified with scale by the addition of sampling uncertainty, particularly on the shorter spatiotemporal scales. The way in which this would be quantified depends on the characteristics of the sensor ground sampling.

The nature of the starting product will also dictate the spatiotemporal length scales over which aggregation is possible or sensible given the sampling frequency. For the SST example, at spatiotemporal scales of 2 km and 28-days, sampling frequency was such that a coverage >50% was difficult to achieve. Expanding the spatiotemporal scales further would lead to further reductions in sampling frequency, which may be undesirable depending on the intended application. The scientific questions that can be answered also change with scale: with large spatiotemporal averages details such as small scale variability from localised heavy rainfall increasing soil moisture, the location of CO₂ sources or the diurnal variability in SST will be lost. If sampling is also in some way systematic, e.g. more frequent sampling of an ocean basin or land area due to orbit characteristics or cloud cover this may also affect the apparent answer to the question posed.

With fully gap-filled products, much larger scale averages can be calculated without addition of sampling uncertainty, although the process of gap-filling can introduce another source of uncertainty if the data are not continuous by nature. It should be noted that gap-filled L4 SST and soil moisture products do both exist and contain uncertainty information (such as the one used as a reference for calculating sampling uncertainty in the SST example, (Good and Embury 2024; Embury et al. 2024), and for soil moisture Preimesberger and Stadiotti (2024)), but in both cases uncertainty information is not propagated directly and fully from the input data through the analysis, so there is a break in the uncertainty traceability chain.

Assumptions made about correlation length scales will also impact the total uncertainty budget and a naive assumption of independence can lead to a false confidence in the resulting spatiotemporal averages as illustrated in the soil moisture example. While error correlations are generally considered to be a “problem”, because they render observations redundant for averaging or aggregation purposes, they do provide an opportunity for data assimilation techniques to propagate observational information across space or time: if, for example, an observation at location x suggests that a model forecast at location x is too low and if *errors* are correlated in space, then this suggests that a model forecast at location y is probably too low as well. This is true regardless of whether or not the modelled or observed *states* at locations x and y are correlated. Thus, a single observation at location x can help improve model forecasts at both locations x and y (Reichle and Koster 2003). This also means that error covariances enable data assimilation techniques to improve model simulations in entirely unobserved regions (Gruber et al. 2018). Nevertheless, such “two-dimensional filtering” requires reliable error covariance estimates for both the model simulations and the observations, and both over- and underestimated error covariances (i.e., a “conservative” guess) can deteriorate model skill instead of improving it (Gruber et al. 2015).

3.2 Bridging the Gap Between Data Producers and Data Users for Climate Services

Fundamentally, the suitability of data for use in a particular decision context is key to ensuring data can be used in services. This requires first, a nuanced understanding of the decision context or service by the data provider coupled with detailed understanding of a given observational product and what question it was designed to answer, and only then can the suitability of the data be clearly assessed. Each component of uncertainty needs to be carefully considered in the context of the question that is being asked: what is important at daily, local scales may not be for global, decadal scale applications. As every application of Earth Observation data will require information on the space and time resolution that is relevant to the question being asked of the data, flexibility in provision of uncertainty information is key to allowing its use. From the perspective of the data provider, the prerequisite then is a comprehensive mapping of uncertainty sources and their relationships to each other. The propagation of the mapping frameworks developed by the FIDUCEO project (Mittaz et al. 2019) is an excellent example of efforts by data providers to fulfil this need (Embury et al. 2024; Gorroño et al. 2024; Ablain et al. 2019).

Climate services will often utilise larger-scale spatiotemporal averages of ECV data e.g. local, district or regional averages or alternatively derived indices such as indicators of extreme events. In the case of decisions that are affected by extreme events, estimates of likelihood or return periods can be important, e.g. for extreme rainfall events in the generation of hydropower. The confidence with which those return periods can be estimated is affected by data quality. Here then, careful quantification of uncertainty in underlying data and the appropriate propagation of this through into the indicators provides greater confidence in estimates of derived quantities such as return periods, or emerging instabilities close to tipping points (Lenton et al. 2024).

The absence of relevant information limits our ability to make decisions. In the case of return times, e.g. drought onset in Sect. 2.2, consideration of uncertainties can have a direct impact on constraint of the relevant indicator. If our estimates are less well constrained, then we have less information on appropriate actions to take in the mitigation of possible future events. In climate modelling, a lack of well-characterised uncertainty information limits our ability to effectively evaluate model representation of the climate system, which increases uncertainty in process representation, and, potentially, in projections of future climate.

For some use cases, providing an estimate of the total uncertainty on each value is important—what is the interval in which the real value lies? It should be very clear what is and is not included—are there key uncertainties that are missing (e.g. sampling uncertainty in the XCO₂ case discussed above)? Knowledge of missing error components can be the most important information to convey because underestimated uncertainties can lead to incorrect use of the data. Note that limitations of the observing system may also be a critical component of uncertainty for some applications, particularly where there may be systematic omissions—as in the case of small fires, where fires that are not captured by lower resolution (500 m) observations represent almost half of the observable burned area when they can be observed by higher resolution (20 m) instruments (van der Velde et al. 2024; Mota et al. 2019; Wooster et al. 2015).

Communicating uncertainty clearly, in a way that is accessible to all data users, is paramount to facilitating its use in downstream applications (irrespective of scale).

Data producer-user partnerships typically fall into three broad categories: 1) users who have no contact with the data producers, 2) data users with high maths-literacy and special, mathematically nuanced applications who can more easily use highly-technical information supplied by data producers about their products and 3) intermediaries, where data are being used in a decision context with some contact between data user and producer. In all cases, clarity and consistency in the communication of uncertainty information is essential. Particularly for case 1, the documentation about the data product should be readily accessible, clearly presented, consistent in the use and definition of terminology (Loew et al. 2017; Merchant et al. 2017; Strobl et al. 2024) and communicated in such a way that it is easily accessible to the non-expert. Providing recipes, or pre-calculated uncertainties on indicators provides a quick look uncertainty estimate, captures expert understanding of the data and enables users to cross-check pre-calculated numbers against their own calculations. It also potentially enables use of the uncertainty information without the user having to perform the formal propagation (e.g. <https://surftemp.net>). In case 2, more detailed product specifications (e.g. CCI documentation, see <https://climate.esa.int/en/projects/>) may be required including fully mathematical derivations; but these would only be an accessible entry point for data users with high maths-literacy. Case 3 involves more communication between data users and producers, with the additional requirement of clear verbal communication and joint understanding of the decision context in which the data are to be used (Gruber et al. 2024). These cases may take the form of research projects, for example the Methane Emissions Detection using Satellites Assessment (MEDUSA) project (<https://climate.esa.int/en/projects/medusa>), which seeks to understand which aspects of uncertainty associated with a complex observing landscape are relevant to specific user requirements.

4 Outlook

This paper demonstrates via case studies that the scale of interest of the scientific question under consideration affects the relative importance of contributions to observational uncertainty, and that this has the potential to affect decisions made on the basis of derived information. Although the global picture still remains important in terms of long-term projections and large-scale energy budgets, there is simultaneously a move towards an increasingly mature landscape of information for action and pathways to adaptation and mitigation at the scale of people and communities, on the shorter term and at local, higher resolution. Processes such as gap filling, that are an established component of global analyses, may be critical in some contexts to ensure some information content but can become actively undesirable by adding artificial/smoothed data at local or sub-annual scales (e.g. for tipping points single sensor analyses can be beneficial (Lenton et al. 2024)).

Good quality decision making, assuring a sensible and measured response to the question posed, requires three components: (1) a clear articulation of the question that must be answered including mapping of its components, constraints, and input requirements, (2) for each component, mapping and propagation of relevant data and associated uncertainty and (3) integration of (2) into (1) such that a decision can be made.

For a single decision—say, flood defence planning, links between communities that often have very different scientific languages, and to communities experiencing impacts, take time, effort and—associated—sustainable funding to build. Well-informed decision

making at scale further requires the generalisation of local-scale decisions across the domain of interest. This has been identified in the Technical Dialogue of the first Global Stocktake (Secretariat 2023), [key finding 10] as a barrier in the adaptation landscape. However, there is some evidence that coordinated approaches are beginning to take hold in some sectors, for example the Methane Alert and Response System (UNEP-MARS) (Programme 2022) and the associated Methane Emissions Detection Using Satellites Assessment (MEDUSA) project, as well as coordinated adaptation platforms such as Copernicus Emergency Management Service (<https://emergency.copernicus.eu/>). Development of pipelines from data to decision making at scale must become a priority if we are to make progress in closing the implementation gap [(Secretariat 2023), key finding 4] to fulfilling the aims of the Paris Agreement.

Ultimately, if good quality data are not developed for decision making, then poorer information will be used—this may provide a weaker constraint and we will, on average, make less-well informed decisions.

Appendix A Calculating SST Uncertainties with Scale

SLSTR data at 0.05° resolution are used to calculate the relative contribution of the three uncertainty components (independent, structured, common) to the total uncertainty budget with scale. Daytime and nighttime data are considered independently in this analysis, with orbits split between the descending node (night) and ascending node (day), as in Fig. 5. The resultant partitioning will depend on two factors: (1) the correlation length scales of the existing uncertainty components, relative to the spatiotemporal scales over which the data are averaged. Independent uncertainties will average down whilst fully correlated (common) uncertainties will not. Propagation of the uncertainty components follows the laws of uncertainty propagation as outlined in the Guide to Uncertainty in Measurement (Metrology 2008, 2007). (2) As illustrated by Fig. 5, the input data are spatially incomplete, and therefore any re-gridding at coarser spatiotemporal resolutions requires the calculation of an additional sampling uncertainty. This sampling uncertainty will arise whenever the region over which the mean SST is calculated is not fully sampled, either in time or space.

A.1 Calculating Sampling Uncertainty

To calculate the sampling uncertainty we require a spatially and temporally complete reference dataset. For this purpose we use the global daily L4 SST product (Good and Embury 2024; Embury et al. 2024). This is an analysis product derived from satellite inputs from multiple sensors, with a background prior that enables gap filling using an optimal interpolation technique (Good et al. 2020). The resultant spatiotemporal variability in the L4 product is therefore significantly lower than in the corresponding L3 data.

To calculate sampling uncertainty (S), we first subtract the daily climatological SST (Embury et al. 2024) from our L3 and L4 SST data so that all calculations are done in anomaly space. We then divide the anomaly data into two sets; n and m . Set n contains all the cells at the target resolution where both the L3 and L4 data are fully sampled (all observations are clear-sky). Set m contains the cells where the L3 data contain only a subsample of the complete data available in the L4 equivalent (some observations are cloud-affected). We use set n to calculate the underestimation of

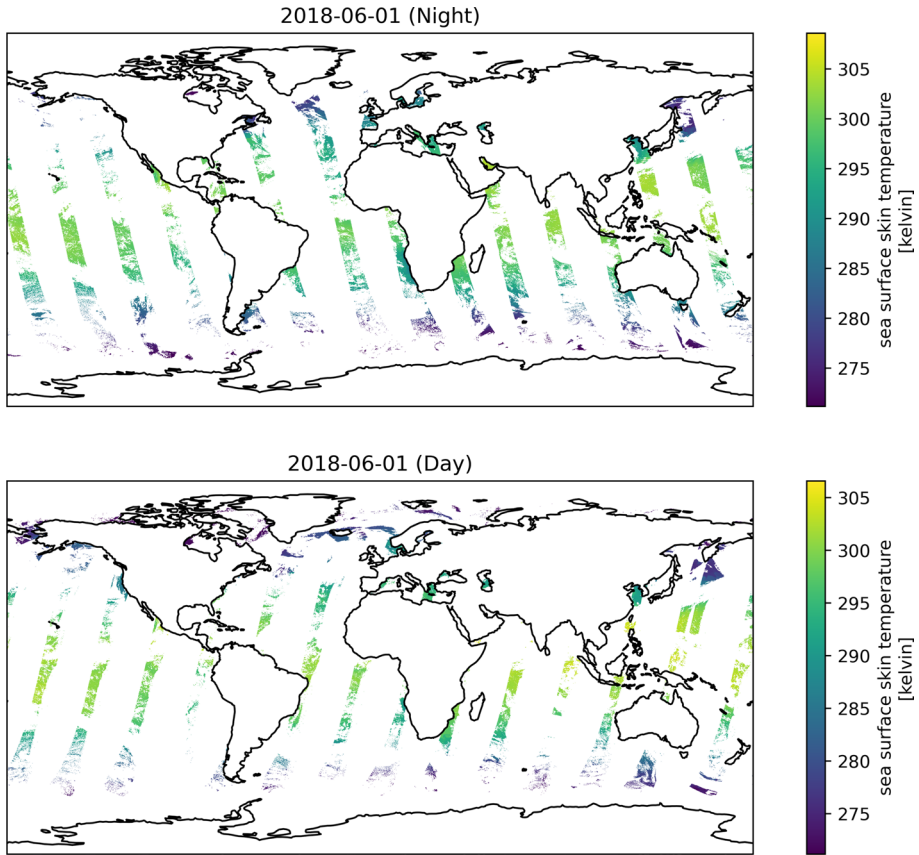


Fig. 5 SLSTR-A L3C SSTs (K) from the ESA CCI dataset for 01-06-2018. Data are split over a 24 UTC time period into night (top) and day (bottom)

variability in the L4 data relative to the L3 data (K). To do this we calculate the ratio of the variance in the L3 data (σ_{L3}^2) after first subtracting the noise ($\sigma_{L3_ind}^2$), to the variance in the corresponding L4 data (σ_{L4}^2). We do this for each cell (i) in set (n).

$$K_i = \sqrt{\frac{\sigma_{L3}^2 - \sigma_{L3_ind}^2}{\sigma_{L4}^2}} \quad (A1)$$

The inflation factor for the sampling uncertainty at the given resolution (K) is the mean value over the n cells.

$$K = \sqrt{\frac{\sum_0^n K_i^2}{n}} \quad (A2)$$

We then analyse the SST anomaly differences (d) for each subsampled cell (j) in set m . Thus, for a given cell j , with p total observations in the L4 data and q observations in the L3 subsample:

$$d_j = \frac{\sum_0^q (SST_{L3})}{q} - \frac{\sum_0^p (SST_{L4})}{p} \quad (A3)$$

Set m is then categorised according to the fraction of the cell sampled by the L3 data. For a given fraction (l), the sampling uncertainty is the standard deviation of the SST differences for ($j \in l$).

$$S_l = K\sigma(d_j \in l) \quad (A4)$$

A.2 Propagating Locally Correlated Uncertainties

We employ the following methodology to calculate the correlation length scales for the structured component of the uncertainty budget, which arises from errors in specification of the atmospheric state in our SST retrieval. We simulate SLSTR orbital data globally using hourly ERA5 atmospheric profiles (Hersbach et al. 2020) and the ESA CCI L4 SST products (Good and Embury 2024; Embury et al. 2024). Both the ERA5 and L4 products are interpolated to the higher 1 km spatial scale, matching the SLSTR spatial resolution and ERA5 data are also temporally interpolated to the SLSTR instantaneous pixel times. These data are then passed to the RTTOV v13.1 radiative transfer model (Hocking et al. 2022) to estimate the top-of-atmosphere brightness temperature in the 3.7, 11 and 12 micron channels. To omit cloudy pixels in this clear-sky simulation, we use the nearest clear-sky atmospheric profile in locations where the ERA5 total cloud cover exceeds 10 %. This threshold should ensure that significantly cloud-affected profiles are not used in these clear-sky simulations.

We retrieve the SST from the simulated SLSTR channels using retrieval coefficients derived using the same methodology as the ESA CCI project (Embury et al. 2024), including a normalisation step to apply the coefficients to the simulated data. We estimate retrieval error by directly comparing the SST retrieved using the coefficients, with the input SST used in the simulation. Once we have the SST errors at every point we can then estimate both the spatial and temporal correlation length scales. For the spatial term we define the correlation length scale as the point at which the error has dropped by a half; scanning out from the central point in annular sections. For the temporal case we use a linear model between observed points (we have currently simulated data once a week for a year) to estimate the time at which the error has dropped by a half. A full description of the methodology deployed here and a detailed description of the outcomes will be the subject of a further publication.

Based on the global distribution of the correlation length scales calculated using the simulated SLSTR data, we assume a fixed temporal correlation length scale of 3-days, over which the structured component of the uncertainty budget is considered to be fully correlated. For the spatial correlation length scale we use 50 km (approximated to be 0.5°) for latitudes between 60° S– 75° N. For the extreme polar regions, the correlation length scales are larger than the upper bound of the spatial scales considered in this manuscript (300 km, approximated as 3°). When gridding directly from 0.05 daily data to a spatiotemporal resolution exceeding either of the spatial or temporal correlation length scales, we calculate the approximate number of pieces of information we have n_b , by dividing the spatiotemporal region of interest (r) by the size of the correlated spatiotemporal ‘box’ (b). We then scale

the structured uncertainty component $1/\sqrt{n_b}$, to reflect some averaging down of the uncertainty component. This is an approximation of the metrological propagation of structured uncertainties, applied here as the full propagation at all space and time scales considered is extremely complex to implement, involving very large covariance matrices.

$$n_b = \frac{r}{b} \quad (\text{A5})$$

It should be noted, that where the spatiotemporal scales exceed the correlation length scales (beyond 3 days and above 0.5° for the majority of the globe), for the purposes of further propagation this structured component would then be added to the independent partition as it would no longer be correlated (not shown in this paper). As spatiotemporal scales increase beyond those represented here, the relative contribution of the common uncertainty component will increase.

A.3 SST Uncertainties with Scale

We sample the SLSTR-A data beginning with the first day in every month of 2018, and according to the temporal length scale of the required coarsening. For the one, five and ten-day timeframes we use data from all months, but for the twenty-eight day sampling, February and September are omitted due to missing data in the data record. Where ten or more completely sampled grid cells are available, we use these to calculate K as described in A.0.1. Tables 1 and 2 show the number of grid boxes from which the statistics are calculated, the number of complete cells available and the calculated value of K . Where insufficient complete grid cells are available to calculate K , its value is assumed with reference to neighbouring resolutions that include complete data. With increasing spatiotemporal resolution, K reduces in size, and is typically lower at night than during the day. This is intuitive as solar heating in the day is likely to enhance SST variability.

The resultant sampling uncertainty, calculated as a function of the sampled cell fraction at each given resolution is shown in Fig. 6. Each subplot shows a different temporal resolution (1, 5, 10 and 28 days) with the sampling uncertainty plotted for each coincident spatial resolution considered. Sampling uncertainty is always zero when the sampled fraction is one. The sampled fraction range is a function of the grid cell size at the given spatiotemporal resolution and data binning. For some of the higher sampled fractions at large spatiotemporal resolutions no data are available to calculate the sampling uncertainty due to the ‘gappy’ nature of the input data.

Generally, the largest sampling uncertainties occur for the smallest sampled fractions (with the exception of 0.05° at 10 days as discussed further below), reaching a maximum of 2.5, 4.0, 0.56 and 0.32 K for the daily, 5-day, 10-day and 28-day data, respectively, for the smallest sampled fractions. The sampling uncertainty for the 10-day data shows a second spike for sampling fractions between 0.5–0.7 (daytime only), consistent across the different spatial resolutions. A second peak is also seen in a similar location for the 28-day data, (where higher sampled fractions are observed in the data). These peaks relate to a sampling artefact of using these data. The equatorial repeat time for observations by a single SLSTR instrument is 1.8 days (Donlon et al. 2012), giving a maximum sampled fraction of 50 % close to the equator, meaning that the data with higher sampled fractions will be predominantly located at mid to high latitudes. SST variability is larger at these latitudes (Bulgin et al. 2020), inflating the sampling

Table 1 Number of observations, fully populated samples and value of **K** for each spatiotemporal scaling of daytime data

	0.05°	0.1°	0.25°	0.5°	1.0°	2.0°
	Obs/Full/K	Obs/Full/K	Obs/Full/K	Obs/Full/K	Obs/Full/K	Obs/Full/K
Daily	1.2×10 ⁷ /–	3.5×10 ⁶ /2.1×10 ⁶ /6.8	8.2×10 ⁵ /1.8×10 ⁵ /3.1	2.6×10 ⁵ /19,786/1.8	84,989/1329/1.3	27,485/47/1.1
5-day	5.2×10 ⁷ /33,616/4.6	1.6×10 ⁷ /6444/2.7	3.3 × 10 ⁶ /531/2.0	9.9×10 ⁵ /51/1.6	2.9 × 10 ⁵ /–/1.1	78,472/–/1.1
10-day	8.2 × 10 ⁷ /39/1.4	2.4×10 ⁷ /–/1.1	4.6×10 ⁶ /–/1.1	1.3×10 ⁶ /–/1.1	3.5×10 ⁵ /–/1.1	86,665/–/1.1
28-day	1.1×10 ⁸ /–/1.1	2.9×10 ⁷ /–/1.1	5.0×10 ⁶ /–/1.1	1.3×10 ⁶ /–/1.1	3.1×10 ⁵ /–/1.1	73,320/–/1.1

Where fewer than ten fully sampled cells are available, **K** is estimated based on the spatiotemporal resolution

Table 2 Number of observations, fully populated samples and value of **K** for each spatiotemporal scaling of nighttime data

	0.05°	0.1°	0.25°	0.5°	1.0°	2.0°
	Obs/Full/K	Obs/Full/K	Obs/Full/K	Obs/Full/K	Obs/Full/K	Obs/Full/K
Daily	1.3×10 ⁷ /–	3.8×10 ⁶ /2.1×10 ⁶ /5.7	9.0×10 ⁵ /1.8×10 ⁵ /2.5	2.9×10 ⁵ /19,970/1.5	91,024/1543/1.2	28,892/58/1.1
5-day	5.5×10 ⁷ /29,083/2.5	1.7×10 ⁷ /6549/2.1	3.5×10 ⁶ /760/1.7	1.0×10 ⁶ /122/1.4	2.9×10 ⁵ /14/1.2	79,337/–/1.1
10-day	8.5×10 ⁷ /–/1.1	2.5×10 ⁷ /–/1.1	4.8×10 ⁶ /–/1.1	1.3×10 ⁶ /–/1.1	3.5×10 ⁵ /–/1.1	87,099/–/1.1
28-day	1.1×10 ⁸ /–/1.1	2.9×10 ⁷ /–/1.1	5.0×10 ⁶ /–/1.1	1.3×10 ⁶ /–/1.1	3.1×10 ⁵ /–/1.1	73,299/–/1.1

Where fewer than ten fully sampled cells are available, **K** is estimated based on the spatiotemporal resolution

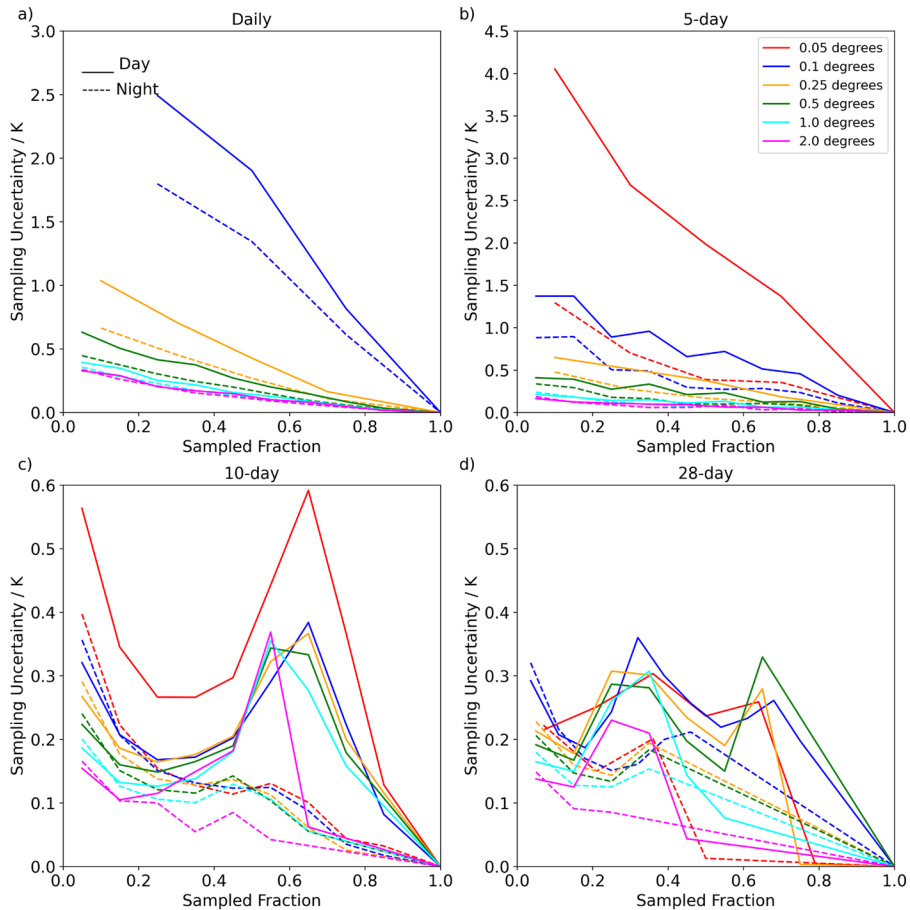


Fig. 6 Sampling uncertainty (K) as a function of sampled fraction for **a** daily, **b** 5-day, **c** 10-day and **d** 28-day temporal resolutions. Sampling uncertainty is plotted as a function of spatial resolution and day (solid) and night (dashed)

uncertainty relative to lower sampled fractions which also include tropical data. Once calculated, the sampling uncertainty is added to the independent uncertainty component, as it is uncorrelated with the sampling uncertainty in neighbouring cells.

Acknowledgements This paper is an outcome of the Workshop “Remote Sensing in Climatology: Essential Climate Variables and their Uncertainties” held at the International Space Science Institute (ISSI) in Bern, Switzerland (13–17 November 2023).

Funding This study has received funding through the National Environmental Research Council’s (NERC’s) support of the National Centre for Earth Observation (NCEO) with contract no. NE/R016518/1. The study has also received funding from the European Space Agency’s Fiducial Reference Measurements for Soil Moisture (FRM4SM) project (ESA Contract No: 4000135204/21/I-BG). NR was supported by the Met Office Hadley Centre Climate Programme funded by DSIT. NR is grateful to a number of colleagues for discussing their experience of using uncertainty information in different contexts. PG and AR were supported by the UK Government’s Department for Science, Innovation and Technology through the UK’s National Measurement System programmes.

Declarations

Conflict of interest The authors declare no conflict of interest.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

References

- Ablain M, Meysignac B, Zawadzki L, Jugier R, Ribes A, Spada G, Benveniste J, Cazenave A, Picot N (2019) Uncertainty in satellite estimates of global mean sea-level changes, trend and acceleration. *Earth Syst Sci Data* 11:1189–1202
- Aldred F, Good E, Bulgin CE, Rayner N (2023) User requirements document: Wp1.1. - del-d1.1. Technical report, European Space Agency Land Surface Temperature Climate Change Initiative
- Araza A, Bruin S, Herold M, Quegan S, Labriere N, Rodriguez-Veiga P, Avitabile V, Santoro M, Mitchard ETA, Ryan CM, Phillips OL, Willcock S, Verbeeck H, Carreiras J, Hein L, Schelhaas M-J, Pacheco-Pascagoza AM, Conceicao Bispo P, Laurin GV, Vieillard G, Slik F, Wijaya A, Lewis SL, Morel A, Liang J, Sukhdeo H, Schepaschenko D, Cavlovic J, Gilani H, Lucas R (2022) A comprehensive framework for assessing the accuracy and uncertainty of global above-ground biomass maps. *Remote Sens Environ* 272
- Barnoud A, Picard B, Meyssignac B, Marti F, Ablain M, Roca R (2023) Reducing the uncertainty in the satellite altimetry estimates of global mean sea level trends using highly stable water vapour climate data records. *J Geophys Res Oceans*. <https://doi.org/10.1029/2022JC019378>
- Bartalis Z, Kidd R, Scipal K (2006) Development and implementation of a discrete global grid system for soil moisture retrieval using the MetOp ASCAT scatterometer. In: 1st EPS/MetOp RAO Workshop, vol. ESA SP-618. ESRIN, Frascati, Italy
- Brodzik MJ, Billingsley B, Haran T, Raup B, Savoie MH (2012) EASE-grid 2.0: incremental but significant improvements for earth-gridded data sets. *ISPRS Int J Geo-Inf* 1(1):32–45. <https://doi.org/10.3390/ijgi1010032>
- Bulgin CE, Merchant CJ, Ferreira D (2020) Tendencies, variability and persistence of sea surface temperature anomalies. *Sci Rep*. <https://doi.org/10.1038/s41598-020-64785-9>
- Bulgin CE, Embury O, Corlett G, Merchant CJ (2016a) Independent uncertainty estimates for coefficient based sea surface temperature retrieval from the Along-Track Scanning Radiometer. *Remote Sens Environ* 178:213–222
- Bulgin CE, Embury O, Merchant CJ (2016b) Sampling uncertainty in gridded sea surface temperature products and Advanced Very High Resolution Radiometer (AVHRR) Global Area Coverage (GAC) data. *Remote Sens Environ* 177:287–294
- Chevallier F, Palmer PI, Feng L, Boesch H, O'Dell CW, Bousquet P (2014) Toward robust and consistent regional CO₂ flux estimates from in situ and spaceborne measurements of atmospheric CO₂. *Geophys Res Lett* 41(3):1065–1070. <https://doi.org/10.1002/2013GL058772>
- Cogan A, Boesch H, Parker R, Feng L, Palmer P, Blavier J-F, Deutscher NM, Macatangay R, Notholt J, Roehl C et al (2012) Atmospheric carbon dioxide retrieved from the Greenhouse gases Observing SATellite (GOSAT): Comparison with ground-based TCCON observations and geos-chem model calculations. *J Geophys Res Atmos* 117(D21)
- Connor BJ, Boesch H, Toon G, Sen B, Miller C, Crisp D (2008) Orbiting carbon observatory: inverse method and prospective error analysis. *J Geophys Res Atmos*. <https://doi.org/10.1029/2006JD008336>
- Connor B, Bösch H, McDuffie J, Taylor T, Fu D, Frankenberg C, O'Dell C, Payne VH, Gunson M, Pollock R et al (2016) Quantification of uncertainties in OCO-2 measurements of XCO₂: simulations and linear error analysis. *Atmos Meas Tech* 9(10):5227–5238
- Crisp D, Miller CE, DeCola PL (2008) NASA Orbiting Carbon Observatory: measuring the column averaged carbon dioxide mole fraction from space. *J Appl Remote Sens* 2(1):023508
- Dee D, Obregon A, Buontempo C (2024) Are our climate data fit for your purpose? *Bull Am Meteorol Soc*. <https://doi.org/10.1175/BAMS-D-23-0295.1>

- Deser C (2020) Certain uncertainty: the role of internal climate variability in projections of regional climate change and risk management. *Earth's Future*. <https://doi.org/10.1029/2020EF001854>
- Donlon C, Berruti B, Buongiorno A, Ferreira M-H, Femenias P, Frerik J, Goryl P, Klein U, Laur H, Mavrocordatos C, Nieke J, Rebhan H, Seitz B, Stroede J, Sciarra R (2012) The Global Monitoring for Environment and Security (GMES) sentinel-3 mission. *Remote Sens Environ* 120:37–57
- Dorigo W, Preimesberger W, Stradiotti P, Kidd R, Schalie R, Vilet M, Rodríguez-Fernandez N, Madelon R, Baghdadi N (2023) Algorithm Theoretical Baseline Document (ATBD) supporting product version 08.1. Technical report, ESA Climate Change Initiative Plus - Soil Moisture
- Embury O, Merchant CJ, Good SA, Rayner NA, Hoyer JL, Atkinson C, Block T, Alerskans E, Pearson KJ, Worsfold M, McCarroll N, Donlon C (2024) Satellite-based time-series of sea-surface temperature since 1980 for climate applications. *Sci Data* 11
- Forster P, Storelvmo T, Armour K, Collins W, Dufresne J-L, Frame D, Lunt DJ, Mauritsen T, Palmer MD, Watanabe M, Wild M, Zhang H (2021) The Earth's Energy Budget, Climate Feedbacks, and Climate Sensitivity. In: *Climate change 2021: the physical science basis. Contribution of Working Group I to the Sixth Assessment Report of the Intergovernmental Panel on Climate Change* [Masson-Delmotte, V., P. Zhai, A. Pirani, S.L. Connors, C. P'Aan, S. Berger, N. Caud, Y. Chen, L. Goldfarb, M.I. Gomis, M. Huang, K. Leitzell, E. Lonnoy, J.B.R. Matthews, T.K. Maycock, T. Waterfield, O. Yelekçi, R. Yu, and B. Zhou (eds.)], pp. 923–1054. Cambridge University Press, Cambridge, United Kingdom and New York, NY, USA., ??? (2021). <https://doi.org/10.1017/9781009157896.009>.
- Ghent D, Veal K, Trent T, Dodd E, Sembhi H, Remedios J (2019) A new approach to defining uncertainties for MODIS land surface temperature. *Remote Sens*. <https://doi.org/10.3390/rs11091021>
- Good sA, Embury O (2024) ESA sea surface temperature climate change initiative (SST_{cci}): level 4 analysis product, version 3.0. Technical report, NERC EDS Centre for Environmental Data Analysis
- Good E, Aldred F, Mottram R, Hoyer J, Karagali I, Sismanidis P, Bechtel B, Walther S, Cheval S, Dumitrescu A, Mallick K, Bossung C, Hemming D, King R, Niclos R (2021) Climate assessment report: WP5.1 - DEL-CAR. Technical report, European Space Agency Land Surface Temperature Climate Change Initiative
- Good E, Veal K (2023) 2022 user workshop report: WP5 - DEL-5.3. Technical report, European Space Agency Land Surface Temperature Climate Change Initiative
- Good S, Fiedler E, Mao C, Martin MJ, Maycock A, Reid R, Roberts-Jones J, Searle T, Waters J, White J, Worsfold M (2020) The current configuration of the OSTIA system for operational production of foundation sea surface temperature and ice concentration analyses. *Remote Sens* 12:720
- Gorroño J, Guanter L, Valentin Graf L, Gascon F (2024) A framework for the estimation of uncertainties and spectral error correlation in Sentinel-2 level-2a data products. *IEEE Trans Geosci Remote Sens* 62:1–13. <https://doi.org/10.1109/TGRS.2024.3435021>
- Gruber A, Bulgin CE, Dorigo W, Embury O, Formanek M, Merchant C, Mittaz J, Muñoz-Sabater J, Pöppel F, Povey A, Wagner W (2025) Making sense of uncertainties: ask the right question. *Surv Geophys*. <https://doi.org/10.1007/s10712-025-09889-5>
- Gruber A, Crow W, Dorigo W, Wagner W (2015) The potential of 2D Kalman filtering for soil moisture data assimilation. *Remote Sens Environ* 171:137–148. <https://doi.org/10.1016/j.rse.2015.10.019>
- Gruber A, De Lannoy G, Albergel C, Al-Yaari A, Brocca L, Calvet J-C, Colliander A, Cosh M, Crow W, Dorigo W, Draper C, Hirschi M, Kerr Y, Konings A, Lahoz W, McColl K, Montzka C, Muñoz-Sabater J, Peng J, Reichle R, Richaume P, Rüdiger C, Scanlon T, van der Schalie R, Wigneron J-P, Wagner W (2020) Validation practices for satellite soil moisture retrievals: what are (the) errors? *Remote Sens Environ* 244:111806. <https://doi.org/10.1016/j.rse.2020.111806>
- Gruber A, Su C-H, Zwieback S, Crow W, Dorigo W, Wagner W (2016) Recent advances in (soil moisture) triple collocation analysis. *Int J Appl Earth Obs Geoinf* 45:200–211. <https://doi.org/10.1016/j.jag.2015.09.002>
- Gruber A, Su C-H, Crow W, Zwieback S, Dorigo W, Wagner W (2016) Estimating error cross-correlations in soil moisture data sets using extended collocation analysis. *J Geophys Res Atmos* 121(3):1208–1219. <https://doi.org/10.1002/2015JD024027>
- Gruber A, Crow W, Dorigo W (2018) Assimilation of spatially sparse in situ soil moisture networks into a continuous model domain. *Water Resour Res* 54(2):1353–1367. <https://doi.org/10.1002/2017WR021277>
- Gruber A, Scanlon T, Schalie R, Wagner W, Dorigo W (2019) Evolution of the ESA cci soil moisture climate data records and their underlying merging methodology. *Earth Syst Sci Data* 11(2):717–739. <https://doi.org/10.5194/essd-11-717-2019>
- Haughton N, Abramowitz G, Pitman A, Phipps SJ (2014) On the generation of climate model ensembles. *Clim Dyn*. <https://doi.org/10.1007/s00382-014-2054-3>

- Hersbach H, Bell B, Berrisford P, Hirahara S, Horanyi A, Munoz-Sabater J, Nicolas J, Peubey C, Radu R, Schepers D, Simmons A, Soci C, Abdalla S, Abellan X, Balsamo G, Bechtold P, Biavati G, Bidlot J, Bonavita M, De Chiara G, Dahlgren P, Dee D, Diamantakis M, Dragani R, Flemming J, Forbes R, Fuentes M, Geer A, Haimberger L, Healy S, Hogan EJ, van der Holm Janiskova M, Keeley S, Laloyauz P, Lopez P, Lupu GC, Radnoti, Rosnay P, Rozum I, Vamborg F, Villaume S, Thépaut J-N (2020) The ERA5 global reanalysis. *Quart J R Meteorol Soc*
- Hocking J, Saunders R, Geer A, Vidot J (2022) RTTOV v13 Users Guide. Technical report, EUMETSAT Satellite Application Facility on Numerical Weather Prediction
- Kennedy JJ, Rayner NA, Atkinson CP, Killick RE (2019) An ensemble data set of sea-surface temperature change from 1850: the Met Office Hadley Centre HadSST.4.0.0.0 data set. *J Geophys Res Atmos*. <https://doi.org/10.1029/2018JD029867>
- Khvorostovsky K, Hendricks S, Rinne E (2020) Surface properties linked to retrieval uncertainty of satellite sea-ice thickness with upward-looking sonar measurements. *Remote Sens*. <https://doi.org/10.3390/rs12183094>
- Kuze A, Suto H, Nakajima M, Hamazaki T (2009) Thermal and near infrared sensor for carbon observation Fourier-transform spectrometer on the greenhouse gases observing satellite for greenhouse gases monitoring. *Appl Opt* 48(35):6716–6733
- Le Borgne P, Roquet H, Merchant CJ (2011) Estimation of sea surface temperature from the Spinning Enhanced Visible and Infrared Imager, improved using numerical weather prediction. *Remote Sens Environ*. <https://doi.org/10.1016/j.rse.2010.08.004>
- Lenton TM, Abrams JF, Bartsch A, Bathiany S, Boulton CA, Buxton JE, Conversi A, Cunliffe AM, Headden S, Laverne T, Poulter B, Shepherd A, Smith T, Swingedouw D, Winkelmann R, Boers N (2024) Remotely sensing potential climate change tipping points across scales. *Nat Commun*. <https://doi.org/10.1038/s41467-023-44609-w>
- Loew A, Bell W, Brocca L, Bulgin CE, Burdanowitz J, Calbet X, Donner RV, Ghent D, Gruber A, Kaminski T, Kinzel J, Klepp C, Lambert J-C, Schaepman-Strub G, Schröder M, Verholst T (2017) Validation practices for satellite-based Earth observation data across communities. *Rev Geophys* 55:779–817
- Merchant CJ, Paul F, Popp T, Ablain M, Bontemps S, Defourny P, Hollman R, Laverne T, Laeng A, Leeuw G, Mittaz J, Poulsen C, Povey AC, Reuter M, Sathyendranath S, Sandven S, Sofieva VF, Wagner W (2017) Uncertainty information in climate data records from Earth observation. *Earth Syst Sci Data* 9:511–527
- Metrology JC (2008) Evaluation of Measurement Data - Guide to the Expression of Uncertainty in Measurement. Organisation Internationale de Metrologie Legale
- Metrology JC (2007) International Vocabulary of Metrology - Basic and General Concepts and Associated Terms (VIM). Organisation Internationale de Metrologie Legale, ???
- Mittaz J, Merchant CJ, Woolliams ER (2019) Applying principles of metrology to historical Earth observations from satellites. *Metrologia*. <https://doi.org/10.1088/1681-7575/ab1705>
- Mota B, Gobron N, Cappucci F, Morgan O (2019) Burned area and surface albedo products: assessment of change consistency at global scale. *Remote Sens Environ* 225:249–266
- Nightingale J, Boersma KF, Muller JP, Compennolle S, Lambert JC, Blessing S, Giering R, Gobron N, De Smedt I, Coheur P, George M, Schulz J, Wood A (2018) Quality assurance framework development based on six new ECV data products to enhance user confidence for climate applications. *Remote Sens*. <https://doi.org/10.3390/rs10081254>
- Noël S, Reuter M, Buchwitz M, Borchardt J, Hilker M, Bovensmann H, Burrows JP, Noia A, Suto H, Yoshida Y, Buschmann M, Deutscher NM, Feist DG, Griffith DWT, Hase F, Kivi R, Morino I, Notholt J, Ohyama H, Petri C, Podolske JR, Pollard DF, Sha MK, Shiomi K, Sussmann R, Té Y, Velasco VA, Warneke T (2021) xCO₂ retrieval for GOSAT and GOSAT-2 based on the focal algorithm. *Atmos Meas Tech* 14(5):3837–3869. <https://doi.org/10.5194/amt-14-3837-2021>
- Noyes EJ, Minnett PJ, Remedios JJ, Corlett GK, Good SA, Llewellyn-Jones DT (2006) The accuracy of the AATSR sea surface temperatures in the Caribbean. *Remote Sens Environ*. <https://doi.org/10.1016/j.rse.2005.11.011>
- O'Dell C, Eldering A, Wennberg P, Crisp D, Gunson M, Fisher B, Frankenberg C, Kiel M, Lindqvist H, Mandrake L et al (2018) Improved retrievals of carbon dioxide from Orbiting Carbon Observatory-2 with the version 8 ACOS algorithm. *Atmos Meas Tech* 11:6539–6576
- Parker RJ, Webb A, Boesch H, Somkuti P, Barrio Guillo R, Di Noia A, Kalaitzi N, Anand JS, Bergamaschi P, Chevallier F, Palmer PI, Feng L, Deutscher NM, Feist DG, Griffith DWT, Hase F, Kivi R, Morino I, Notholt J, Oh Y-S, Ohyama H, Petri C, Pollard DF, Roehl C, Sha MK, Shiomi K, Strong K, Sussmann R, Té Y, Velasco VA, Warneke T, Wennberg PO, Wunch D (2020) A decade of GOSAT proxy satellite CH₄ observations. *Earth Syst Sci Data* 12(4):3383–3412. <https://doi.org/10.5194/essd-12-3383-2020>
- Pasik A, Gruber A, Preimesberger W, Santis D, Dorigo W (2023) Uncertainty estimation for a new exponential-filter-based long-term root-zone soil moisture dataset from Copernicus Climate Change

- Service (C3S) surface observations. *Geosci Model Dev* 16(17):4957–4976. <https://doi.org/10.5194/gmd-16-4957-2023-corrigendum>
- Plummer S, Lecomte P, Doherty M (2018) The ESA Climate Change Initiative (CCI): a European contribution to the generation of the global climate observing system. *Remote Sens Environ* 203
- Preimesberger W, Stadiotti P (2024) ESA CCI SM GAPFILLED long-term Climate Data Record of surface soil moisture from merged multi-satellite observations (9.1). Dataset. TU Wien. <https://doi.org/10.48436/s5jq-rpd32>
- United Nations Environment Programme (2022) An eye on methane: International methane emissions observatory. Technical report
- Ran Y, Li X (2019) TANSAT: a new star in global carbon monitoring from China. *Sci Bull* 64(5):284–285
- Reichle RH, Koster RD (2003) Assessing the impact of horizontal error correlations in background fields on soil moisture estimation. *J Hydrometeorol* 4(6):1229–1242. [https://doi.org/10.1175/1525-7541\(2003\)004<1229:ATIOHE>2.0.CO;2](https://doi.org/10.1175/1525-7541(2003)004<1229:ATIOHE>2.0.CO;2)
- Rodell M, Houser P, Uea Jambor, Gottschalck J, Mitchell K, Meng C, Arsenault K, Cosgrove B, Radakovich J, Bosilovich M et al (2004) The global land data assimilation system. *Bull Am Meteorol Soc* 85(3):381–394. <https://doi.org/10.1175/BAMS-85-3-381>
- Roebeling RA, Bojinski S, Poli P, Schulz J (2025) On the determination of GCOS ECV product requirements for climate applications. *Bull Am Meteorol Soc*. <https://doi.org/10.1175/BAMS-D-24-0123.1>
- Sayer AM, Govaerts Y, Kolmonen P, Lipponen A, Luffarelli M, Mielonen T, Patadia F, Popp T, Povey AC, Stebel K, Witek ML (2020) A review and framework for the evaluation of pixel-level uncertainty estimates in satellite aerosol remote sensing. *Atmos Meas Tech*. <https://doi.org/10.5194/amt-13-373-2020>
- Secretariat U (2023) Technical Dialogue of the First Global Stocktake. Synthesis Report by the Co-facilitators on the Technical Dialogue, FCCC/SB/2023/9. UNFCCC
- Strobl PA, Woolliams ER, Molch K (2024) Lost in translation: the need for common vocabularies and an interoperable thesaurus in Earth observation sciences. *Surv Geophys*. <https://doi.org/10.1007/s10712-024-09854-8>
- Velde IA, Werf GR, Wees D, Schutgens NAJ, Vernooij R, Houweling S, Tonucci E, Chuvieco E, Randerson JT, Frey MM, Borsdorff T, Aben I (2024) Small fires, big impact: evaluating fire emission estimates in southern Africa using new satellite imagery of burned area and carbon monoxide. *Geophys Res Lett*. <https://doi.org/10.1029/2023GL106122>
- Wooster MJ, Roberts G, Freeborn PH, Xu W, Govaerts Y, Beeby R, He J, Lattanzio A, Fisher D, Mullen R (2015) LSA SAF METEOSAT FRP products - part 1: algorithms, product contents and analysis. *Atmos Chem Phys* 15:13217–13239
- Yang D, Boesch H, Liu Y, Somkuti P, Cai Z, Chen X, Noia A, Lin C, Lu N, Lyu D, Parker RJ, Tian L, Wang M, Webb A, Yao L, Yin Z, Zheng Y, Deutscher NM, Griffith DWT, Hase F, Kivi R, Morino I, Notholt J, Ohyama H, Pollard DF, Shiomi K, Sussmann R, Té Y, Velasco VA, Warneke T, Wunch D (2020) Toward high precision xCO₂ retrievals from TANSAT observations: Retrieval improvement and validation against TCCON measurements. *J Geophys Res Atmos* 125(22):2020–032794. <https://doi.org/10.1029/2020JD032794>
- Yang CX, Cagnazzo C, Artale V, Nardelli BB, Buontempo JC ad Busatto, Caporaso L, Cesarini C, Cionni I, Coll J, Crezee B, Cristofanelli P, Toma V, Essa YH, Eyring V, Fierli F, Grant L, Hassler B, Hirschi M, Huybrechts P, Le Merle E, Leonelli FE, Lin X, Madonna F, Mason E, Massonnet F, Marcos M, Marullo S, Müller B, Obregon A, Organelli E, Palacz A, Pascual A, Pisano A, Putero D, Rana A, Sanchez-Roman SA ad Seneviratne, Serva F, Storto A, Thiery W, Throne P, Van Tricht L, Verhaegen Y, Volpe G, Santoleri R (2022) Independent quality assessment of essential climate variables: lessons learnt from the Copernicus Climate Change Service. *BAMS* 103
- Yoshida Y, Kikuchi N, Morino I, Uchino O, Oshchepkov S, Bril A, Saeki T, Schutgens N, Toon G, Wunch D et al (2013) Improvement of the retrieval algorithm for GOSAT SWIR xCO₂ and xCH₄ and their validation using TCCON data. *Atmos Meas Tech* 6(6):1533–1547
- Zeng YJ, Su ZB, Barnpadimos I, Perrels A, Poli P, Boersma KF, Frey A, Ma XG, Bruin K, Goosen H, John VO, Roebeling R, Schulz J, Timmermans W (2019) Towards a traceable climate services: assessment of quality and usability of essential climate variables. *Remote Sens* 11(10)
- Zwieback S, Dorigo W, Wagner W (2013) Estimation of the temporal autocorrelation structure by the collocation technique with an emphasis on soil moisture studies. *Hydrol Sci J* 58(8):1729–1747
- Zwieback S, Colliander A, Cosh MH, Martínez-Fernández J, McNairn H, Starks PJ, Thibeault M, Berg A (2018) Estimating time-dependent vegetation biases in the SMAP soil moisture product. *Hydrol Earth Syst Sci* 22(8):4473–4489. <https://doi.org/10.5194/hess-22-4473-2018>

Authors and Affiliations

Claire E. Bulgin^{1,2}  · **Paul Green**³ · **Alexander Gruber**⁴ · **Claire Macintosh**⁵ · **Jonathan Mittaz**¹ · **Ashley Ramsay**³ · **Nick A. Rayner**⁶

✉ Claire E. Bulgin
c.e.bulgin@reading.ac.uk

Paul Green
paul.green@npl.co.uk

Alexander Gruber
alexander.gruber@geo.tuwien.ac.at

Claire Macintosh
claire.macintosh@esa.int

Jonathan Mittaz
j.mittaz@reading.ac.uk

Ashley Ramsay
ashley.ramsay@npl.co.uk

Nick A. Rayner
nick.rayner@metoffice.gov.uk

¹ Department of Meteorology, University of Reading, Reading RG6 6ET, UK

² National Centre for Earth Observation, University of Reading, Reading RG6 6ET, UK

³ Climate and Earth Observation Group, National Physical Laboratory, Teddington TW11 0LW, UK

⁴ Department of Geodesy and Geoinformation, Technische Universität Wien (TU Wien), 1040 Vienna, Austria

⁵ ESA Climate Office, European Space Agency, Harwell Campus, UK

⁶ Met Office Hadley Centre, Met Office, Exeter EX1 3PB, UK