CoCO2
Prototype system for a Copernicus CO_2 service

Report on the proposed EQC tool

Frédéric Chevallier (CEA) and
Nicolas Bousserez (ECMWF)

With help from Richard Engelen, Paul Palmer
and Marko Scholze

coco2-project.eu

D6.7 Report on the proposed EQC tool

Dissemination Level:	Public
Author(s):	Frédéric Chevallier (CEA) and Nicolas Bousserez (ECMWF)
Date:	13/12/2023
Version:	1.0
Contractual Delivery Date:	31/12/2023
Work Package/ Task:	WP6/ T6.6
Document Owner:	ECMWF
Contributors:	Dominik Brunner, Richard Engelen, Andrea Kaiser-Weiss, Paul Palmer and Marko Scholze
Status:	Final

CoCO2: Prototype system for a Copernicus CO_2 service

Coordination and Support Action (CSA)
H2020-IBA-SPACE-CHE2-2019 Copernicus evolution – Research activities in support of a European operational monitoring support capacity for fossil CO2 emissions

Project Coordinator:	Dr Richard Engelen (ECMWF)
Project Start Date:	01/01/2021
Project Duration:	36 months

Published by the CoCO2 Consortium

Contact:
ECMWF, Shinfield Park, Reading, RG2 9AX,
richard.engelen@ecmwf.int

The CoCO2 project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 958927.

Table of Contents

1 Introduction

The European Commission is establishing an operational observation-based anthropogenic CO_2 emissions monitoring and verification support capacity (CO2MVS) as part of its Copernicus Earth Observation programme. Initially focussed on carbon dioxide (CO_2), as its name suggests, the CO2MVS is being progressively extended to methane (CH_4). Demonstrator systems for this CO2MVS are being developed in the Prototype System for a Copernicus CO_2 service (CoCO2) project.

In its information to the 1st Global Stocktake entitled *Data contribution of the European CoCO2 project to the first Global StockTake* (GST, Deliverable D6.5[1]), CoCO2 introduced data from five demonstrator systems, each exploring and prefiguring different aspects of the future CO2MVS: three fossil fuel CO_2 emissions estimates differing by their geographical coverage, a global CO_2 emissions estimate for the Agriculture, Forestry and Other Land Use sector (AFOLU), and a CH_4 emission estimate from the extended Copernicus Atmosphere Monitoring Service global monitoring system.

A *Functional Requirements Specification Document* (Deliverable D6.4[2]) described the codes that were used to generate this data, in response to the user requirements identified by CoCO2 in Deliverable 6.3 “User Requirement Document”[3], had been addressed from a technical point of view by each demonstrator. A *Fitness for Purpose Document* (Deliverable D6.6[4]) discussed how these user requirements had been addressed from a scientific point of view. In this CoCO2 action for the GST, user requirements (for the GST) were placed at the beginning and end of data production. Similarly, the data produced by the future CO2MVS will be evaluated primarily with respect to the motivation behind generating those data.

This report goes one step further in the performance assessment of the future CO2MVS by synthesising different strategies for an Evaluation and Quality Control (EQC) tool. It builds on existing approaches in the CAMS operational service, on the experience acquired in D6.5 for the 1st GST with real data, and on the experience gained in the atmospheric inversion community in general (e.g., Michalak et al. 2017). The proposed evaluation methodology is generic enough to accommodate the heterogeneous nature of the posterior emission products considered for the CO2MVS (global, regional, direct mass-balance or Bayesian transport model inversions, etc.). It can address observation-based statistical validation methods as well as data assimilation sensitivity methods using simulated observations. The objective metrics developed (Sections 2-6) can be complemented by additional evaluations by experts and users (see Sections 7 and 8, respectively). Each section is divided into a general presentation and a discussion of operational aspects. The assessment of the EQC process itself is not discussed but it is obvious that the EQC framework of the CO2MVS will evolve over time in order to increase robustness.

2 Internal diagnostics

2.1 General

The initial step in evaluating the CO2MVS is the verification of the processing chain itself, i.e. the inspection and control of the process that generated the product distributed to users.

All processing approaches considered for the CO2MVS involve the minimization of a cost function that quantifies the fit of the solution to assimilated observations and, for Bayesian approaches, to

[1] https://www.coco2-project.eu/node/364
[2] https://www.coco2-project.eu/node/369
[3] https://www.coco2-project.eu/node/331
[4] https://www.coco2-project.eu/node/370

prior data. Checking that the corresponding cost function has indeed been minimised to the expected extent therefore constitutes an obvious first check of the health of the inversion process.

The cost function aggregates the deviations of the solution from all the observed and a priori data that have been used: it is also interesting to decompose it in order to check the evolution of the cost function for different types of data (a priori, observations, types of observations, etc.), different geographical domains or different temporal domains.

For optimal Bayesian systems, the value of the minimised cost function itself is significant, but only under restricted statistical assumptions which may not be valid (Chevallier 2017): achieving an ideal value of the cost function is neither necessary nor sufficient for the solution to be optimal (Talagrand et al. 2014).

The CO2MVS will gather a large diversity of prior and observed data with different update timescales: estimates of the CO_2 emissions from biomass fires are available in near-real-time through CAMS, official CO_2 column retrievals from NASA's second Orbiting Carbon Observatory (OCO-2) are available about two months after real time, while retrievals from Japan's Greenhouse Gas Observing SATellite (GOSAT) are produced by CAMS within a couple of days; http://carbonmonitor.org updates its global estimate of the CO_2 emissions from the use of fossil fuels and the production of cement every month covering a period that includes the previous month; the NOAA's "near-real time" database of CO_2 measurements is actually released about three times per year (https://gml.noaa.gov/ccgg/obspack/release_notes.html#obspack_co2_1_NRT, last access 12 December 2023), etc. Quality assurance also involves this data collection and data update. Systematically checking the input data flows for each inversion, so that each of its input data is up-to-date at its own cadence, will be necessary but demanding for products generated with low latency.

The CO2MVS will also assemble a large diversity of physical models, statistical models and processing steps that all contribute to the final products. Quality assurance involves regular evaluations of each major component that participates in the service. The evaluation covers both accuracy and efficiency: are computing resources appropriately spent to provide the best data? A modular structure for the CO2MVS processes in which components and process tasks are grouped into clear modules that can be tested outside of the larger system and possibly replaced, will greatly facilitate this assessment. The Community Inversion Framework (Berchet et al., 2021) that has been built on *elementary transformations* or *plugins* illustrates such a strategy.

2.2 Operational aspect

An operational chain needs to include such internal diagnostics so that unexpected system behaviours can be automatically detected. Furthermore, the individual components of the chain need to be tested when a new version becomes available. These checks require sufficient allocation of human resources.

3 Comparison to independent data

3.1 General

The second step in evaluating the CO2MVS is the verification of its products against independent information. We distinguish here two types of independent information.

The first type relates to information of the same nature as the CO2MVS products: CO_2 or CH_4 emission and absorptions at given time and spatial scales. It forms the most obvious validation dataset. Ideally, we wish to validate the CO2MVS products against very accurate measurements, like those made directly in smokestacks, or those made with the eddy-covariance technique either from towers at plot scale or from aircraft at regional scale (e.g., Lauvaux et al. 2009). However, not all the CO2MVS products may reach the high spatiotemporal resolution of these measurements and some aggregation of the measurements may be needed (e.g., Broquet et al 2013). The strong multi-scale heterogeneity of the CO_2 or CH_4 emission and absorptions cannot be underestimated and aggregation is a challenge in itself. At coarse scale (e.g., national and annual), inventory data may also provide interesting references with uneven accuracy, higher for fossil fuel emissions, lower for AFOLU emissions and absorptions.

The second type of independent information concerns observations of a nature different from that of the CO2MVS products. The most obvious ones are atmospheric observations (air-sample measurements in the boundary-layer or in the upper atmosphere, remote-sensing data from the ground or from space) since they will constitute a major part of the input data. Their comparison to the model outputs implies that the corresponding observation operators are available. Some of these data may have been specifically retained as a fraction of the CO2MVS input data for cross-validation purposes, or may have come from other data streams.

In both cases, the uncertainty of independent data can be significant (remote sensing data, inventory data, etc.) and must be considered when interpreting their differences with CO2MVS products. Likewise, CO2MVS products have to be accompanied by uncertainty statistics. Ultimately, the differences between CO2MVS products and independent data should be consistent with the different error models involved.

The distribution of independent data in time is irregular: aircraft measurement campaigns may increase the data volume for a few weeks only, while all measurement programmes have a start, an end and possible interruptions in-between[5]. Some of the independent data are also reprocessed over time, due to changes in the calibration scale or improvements in their own processing chains. Last, emission and absorption anomalies, like during the Coronavirus recessions (e.g., Chevallier et al. 2020), are rare events by definition and the evolution of the CO2MVS skill during these extreme events is of particular interest. For all these reasons, the CO2MVS will have to monitor the performance of its successive versions for extended periods, at least a decade, with repeated reprocessing.

3.2 Operational aspect

Some independent data can be monitored at marginal cost by the CO2MVS operational chain, such as operational unassimilated in situ measurements, but most arrive at an irregular pace and must be downloaded manually. This activity fits well in reanalysis mode. It can also be carried out offline to evaluate past operational production, supported by associated human resources.

4 Forecast vs. analyses vs. observations

4.1 General

The quality of the CO2MVS products can be evaluated by assessing the added value of the inversion in the atmospheric forecast. In that framework, the atmospheric transport model is run from initial conditions and emissions are generated from three configurations: no assimilation, atmospheric-state-only analysis and emission inversion experiments (with emission scaling factors kept constant in forecast mode). The performance of the forecast is then quantified based on model comparison with observations of the same nature as the ones assimilated in the inversion system but associated with future times and therefore independent (e.g., McNorton et al., 2022). Additional cross-validation experiments wherein a subset of observations within the assimilation window are left for evaluation of the posterior product can also be carried out for further quality control and validation.

[5] see, e.g., the interruption of the Mauna Loa record in 2022-2023, https://www.noaa.gov/news-release/broken-record-atmospheric-carbon-dioxide-levels-jump-again

4.2 Operational aspect

This evaluation approach would be typically used in a reanalysis mode to assess the overall performance of the inversion system and to optimise the parameters of the prior error covariance matrix (i.e., variances, spatial and temporal correlation length scales) for selected periods before running the full year of joint state/emission reanalysis. In that context, the three configurations (control, atmospheric-state-only analysis and emission inversion) could be run in parallel for one month in each season and the 5-day forecasts used for the observation-based evaluation of the product.

5 OSSEs, uncertainty estimation, sensitivity studies

5.1 General

Observing System Simulation Experiments (OSSEs) consist of assimilating synthetic observations generated from some known atmospheric state or/and parameters (e.g., emissions) considered as the ‘truth’. While many OSSEs are performed using three simulation experiments (i.e., control, nature run and inversion), a Monte-Carlo approach similar to the Ensemble of Data Assimilation system used at ECMWF or at LSCE (Chevallier et al., 2007) can provide useful statistical information.

In this idealised framework, the performance of the inversion system based on comparison with the true state and parameters can be evaluated under different assumptions. For instance, one can assume that all hyperparameters of the prior error covariance (variances, spatial and temporal correlation length scales) are known, in which case the OSSE can be used to compute the formal posterior error covariance matrix. One can also evaluate the performance of the system when such input parameters are mis-prescribed or the prior mean is biased with respect to the truth. In that case, the prior error covariance matrix and/or the prior mean used in the inversion system are different from the error covariance and mean used to sample the prior in the ensemble of inversions. The same can apply for the observational error term, for which, e.g., a bias in the model can be introduced to additional sensitivity analysis.

5.2 Operational aspect

The OSSE framework would be used to provide information on uncertainties that are not accounted for in the formal ensemble-based (e.g., Monte-Carlo) posterior covariance estimation, namely, uncertainties arising from mis-prescribed input parameters in the inversion. These sensitivity experiments would be conducted in reanalysis mode for selected periods to help characterise the associated average uncertainties. The OSSE framework may also be used in the context of further developments of the CO2MVS in order to test changes in the system and their impact on the quality of the product before operational use.

6 Comparison between inversion products

6.1 General

Intercomparison exercises, prominently within the framework of the Atmospheric Tracer Transport Model Intercomparison Project (TransCom), but also in other contexts such as the activities of the OCO-2 science team or the Global Carbon project, have played a key role over the past two decades in identifying both robustness and weaknesses of existing inversion products in general, debugging some participating systems, sharing expertise, stimulating scientific and technical developments, and promoting atmospheric inversion.

Usually, all inversions are treated in this context under the assumption that they can be equally reliable, independent from the sophistication of the systems or the skill of their components, even though this assumption may not be fair to, e.g., the new CO2MVS. Moving away from this consensual approach is not simple, but recent efforts introduced quality criteria in order to exclude suspicious submissions (e.g., Friedlingstein et al. 2022) or to weight all submissions individually based on some criterion (Cressie et al. 2022).

6.2 Operational aspect

Intercomparison exercises come with associated constraints. Intercomparison exercises usually involve a protocol that imposes some restriction on the inversion set-up or on product format for participants. For instance, the successive versions of the OCO-2 Model Intercomparison Project imposed the choice of the assimilated data and that of the prior CO_2 fossil fuel emissions (Byrne et al. 2023 and references therein). They also imposed a format for the submission of inversion data. Participating in such occasional intercomparisons therefore necessitates dedicated staff and computing resources, which is not obvious in an operational service. At a minimum, it would be interesting to support possible European efforts towards an inversion benchmark, which could involve national efforts and ICOS efforts. Deliverable D5.6 "Quantification of uncertainty ranges from European multi-model inversions and ways to benchmark inversion systems"[6] proposes to set up a framework/tool on a European level to facilitate a quasi-operational comparison of inversion products.

Additionally, intercomparison exercises may only be one aspect of a larger and operational effort where obtaining the multi-model ensemble mean is among the main objectives, as is the case for instance in the CAMS European air quality forecasts. Such a motivation may emerge from the Global Greenhouse Gas Watch, a large greenhouse gas monitoring infrastructure that the World Meteorological Organization (WMO) is setting up (https://wmo.int/news/media-centre/world-meteorological-congress-approves-global-greenhouse-gas-watch, last access: 12/12/2023).

7 Expert knowledge

7.1 General

By providing the geographical distribution of greenhouse gas emission and uptake, the CO2MVS will address diverse emission and absorption processes, on land and oceans, mainly natural or resulting from human activity. There should be a specific scientific or expert community behind each process to ensure the credibility of the CO2MVS products. The EQC should therefore involve a broad spectrum of expertise to contribute to the EQC, drawing early attention to features of CO2MVS products that may be considered less or not realistic throughout production.

This strategic feedback should still be taken with caution as opinions among experts may vary and controversies sometimes arise (e.g., Fan et al. 1998, Reuter et al. 2017, Wang et al. 2022a, Wang et al. 2022b). The conclusions of the EQC should therefore be based on concrete evidence gathered by the experts, rather than on their opinions.

7.2 Operational aspect

Expert group members should devote sufficient time to analysing relevant CO2MVS results frequently, like every month. This duty may need to be formalised in a contract.

[6] https://www.google.com/url?q=https://confluence.ecmwf.int/download/attachments/340761784/CoCO2-D5-6-V2-0.docx?version%3D2%26modificationDate%3D1699866817521%26api%3Dv2&sa=D&source=docs&ust=1700904396846843&usg=AOvVaw3y8Q3OebmkFEoPUlm2_VE0

8 User uptake and feedback

8.1 General

Ultimately, the CO2MVS must satisfy its users in general (see, e.g., WMO 2005 for public weather services) and its EQC must be driven by users, within its programmatic objectives, as much as possible. This user-oriented feedback will ensure that the objective figures provided by the technical validation and evaluation methods described above are relevant to users and correspond to their assessment (e.g., WMO 2000 for public weather services). More than that, the technical validation and evaluation methods must ultimately serve user assessment rather than only providing quantitative skill targets.

The existing CAMS framework for user interaction with workshops, user support, a User Requirements Data Base (URDB), online user forums and its national collaboration programme illustrates such an approach. The CO2MVS follows it as well, with the CoCO2 loop D6.3 → D6.4 → D6.5 → D6.6 → D6.3, D6.3 being user requirements. The novel nature of CO2MVS may delay user feedback, as the number of users giving proper feedback may not be large during the first years. But every opportunity should be taken to identify directly from users those aspects of the product whose quality is not sufficient, so that corrective measures can be implemented where possible. Significant progress identified internally on the quality of products and services must also be confirmed by users, which implies a certain monitoring over time.

8.2 Operational aspect

The time frame for collecting user feedback is much slower than the time frame for data production. However, users may focus on specific events as they occur (e.g., the impact of an on-going climate anomaly on the carbon cycle) and put pressure on operations to achieve specific results or quality requirements. As has been the case with CAMS when large climate anomalies occur, staff must be prepared to respond.

9 Acknowledgments

This report benefitted from comments and suggestions from Richard Engelen, Paul Palmer and Marko Scholze. We thank them for their time and their insight.

10 References

Berchet, A., Sollum, E., Thompson, R. L., Pison, I., Thanwerdas, J., Broquet, G., Chevallier, F., Aalto, T., Berchet, A., Bergamaschi, P., Brunner, D., Engelen, R., Fortems-Cheiney, A., Gerbig, C., Groot Zwaaftink, C. D., Haussaire, J.-M., Henne, S., Houweling, S., Karstens, U., Kutsch, W. L., Luijkx, I. T., Monteil, G., Palmer, P. I., van Peet, J. C. A., Peters, W., Peylin, P., Potier, E., Rödenbeck, C., Saunois, M., Scholze, M., Tsuruta, A., and Zhao, Y.: The Community Inversion Framework v1.0: a unified system for atmospheric inversion studies, Geosci. Model Dev., 14, 5331–5354, https://doi.org/10.5194/gmd-14-5331-2021, 2021.

Broquet, G., Chevallier, F., Bréon, F.-M., Kadygrov, N., Alemanno, M., Apadula, F., Hammer, S., Haszpra, L., Meinhardt, F., Morguí, J. A., Necki, J., Piacentino, S., Ramonet, M., Schmidt, M., Thompson, R. L., Vermeulen, A. T., Yver, C., and Ciais, P.: Regional inversion of CO2 ecosystem fluxes from atmospheric measurements: reliability of the uncertainty estimates, Atmos. Chem. Phys., 13, 9039–9056, https://doi.org/10.5194/acp-13-9039-2013, 2013.

Byrne, B., Baker, D. F., Basu, S., Bertolacci, M., Bowman, K. W., Carroll, D., Chatterjee, A., Chevallier, F., Ciais, P., Cressie, N., Crisp, D., Crowell, S., Deng, F., Deng, Z., Deutscher, N. M., Dubey, M. K., Feng, S., García, O. E., Griffith, D. W. T., Herkommer, B., Hu, L., Jacobson, A. R., Janardanan, R., Jeong, S., Johnson, M. S., Jones, D. B. A., Kivi, R., Liu, J., Liu, Z., Maksyutov, S., Miller, J. B., Miller, S. M., Morino, I., Notholt, J., Oda, T., O'Dell, C. W., Oh, Y.-S., Ohyama, H., Patra, P. K., Peiro, H., Petri, C., Philip, S., Pollard, D. F., Poulter, B., Remaud, M., Schuh, A., Sha, M. K., Shiomi, K., Strong, K., Sweeney, C., Té, Y., Tian, H., Velazco, V. A., Vrekoussis, M., Warneke, T., Worden, J. R., Wunch, D., Yao, Y., Yun, J., Zammit-Mangion, A., and Zeng, N.: National CO_2 budgets (2015–2020) inferred from atmospheric CO_2 observations in support of the global stocktake, Earth Syst. Sci. Data, 15, 963–1004, https://doi.org/10.5194/essd-15-963-2023, 2023.

Chevallier, F.: Impact of correlated observation errors on inverted CO2 surface fluxes from OCO measurements, Geophys. Res. Lett., 34, L24804, https://doi.org/10.1029/2007GL030463, 2007.

Chevallier, F., F.-M. Bréon, and P. J. Rayner, 2007: The contribution of the Orbiting Carbon Observatory to the estimation of CO2 sources and sinks: Theoretical study in a variational data assimilation framework. J. Geophys. Res., 112, D09307, doi:10.1029/2006JD007375

Chevallier, F., Zheng, B., Broquet, G., Ciais, P., Liu, Z., Davis, S. J., et al. (2020). Local anomalies in the column-averaged dry air mole fractions of carbon dioxide across the globe during the first months of the coronavirus recession. *Geophysical Research Letters*, 47, e2020GL090244. https://doi.org/10.1029/2020GL090244

Cressie, N., Bertolacci, M., & Zammit-Mangion, A. (2022). From many to one: Consensus inference in a MIP. *Geophysical Research Letters*, 49, e2022GL098277. https://doi.org/10.1029/2022GL098277

Fan, S., Gloor, M., Mahlman, J., Pacala, S., Sarmiento, J., Takahashi, T., and Tans, P., 1998. A large terrestrial carbon sink in North America implied by atmospheric and oceanic carbon dioxide data and models, Science, 282, 442–446, https://doi.org/10.1126/science.282.5388.442

Friedlingstein, P., O'Sullivan, M., Jones, M. W., Andrew, R. M., Gregor, L., Hauck, J., Le Quéré, C., Luijkx, I. T., Olsen, A., Peters, G. P., Peters, W., Pongratz, J., Schwingshackl, C., Sitch, S., Canadell, J. G., Ciais, P., Jackson, R. B., Alin, S. R., Alkama, R., Arneth, A., Arora, V. K., Bates, N. R., Becker, M., Bellouin, N., Bittig, H. C., Bopp, L., Chevallier, F., Chini, L. P., Cronin, M., Evans, W., Falk, S., Feely, R. A., Gasser, T., Gehlen, M., Gkritzalis, T., Gloege, L., Grassi, G., Gruber, N., Gürses, Ö., Harris, I., Hefner, M., Houghton, R. A., Hurtt, G. C., Iida, Y., Ilyina, T., Jain, A. K., Jersild, A., Kadono, K., Kato, E., Kennedy, D., Klein Goldewijk, K., Knauer, J., Korsbakken, J. I., Landschützer, P., Lefèvre, N., Lindsay, K., Liu, J., Liu, Z., Marland, G., Mayot, N., McGrath, M. J., Metzl, N., Monacci, N. M., Munro, D. R., Nakaoka, S.-I., Niwa, Y., O'Brien, K., Ono, T., Palmer, P. I., Pan, N., Pierrot, D., Pocock, K., Poulter, B., Resplandy, L., Robertson, E., Rödenbeck, C., Rodriguez, C., Rosan, T. M., Schwinger, J., Séférian, R., Shutler, J. D., Skjelvan, I., Steinhoff, T., Sun, Q., Sutton, A. J., Sweeney, C., Takao, S., Tanhua, T., Tans, P. P., Tian, X., Tian, H., Tilbrook, B., Tsujino, H., Tubiello, F., van der Werf, G. R., Walker, A. P., Wanninkhof, R., Whitehead, C., Willstrand Wranne, A., Wright, R., Yuan, W., Yue, C., Yue, X., Zaehle, S., Zeng, J., and Zheng, B.: Global Carbon Budget 2022, Earth Syst. Sci. Data, 14, 4811–4900, https://doi.org/10.5194/essd-14-4811-2022, 2022.

Lauvaux, T., Gioli, B., Sarrat, C., Rayner, P., Ciais, P., Chevallier, F., Noilhan, J., Miglietta, F., Brunet, Y., Ceschia, E., Dolman, H., Elbers, J., Gerbig, C., Hutjes, R., Jarosz, N., Legain, D., and Uliasz, M.: Bridging the gap between atmospheric concentrations and local ecosystem measurements, Geophys. Res. Lett., 36, L19809, https://doi.org/10.1029/2009GL039574, 2009

McNorton, J., Bousserez, N., Agustí-Panareda, A., Balsamo, G., Cantarello, L., Engelen, R., Huijnen, V., Inness, A., Kipling, Z., Parrington, M., and Ribas, R.: Quantification of methane emissions from hotspots and during COVID-19 using a global atmospheric inversion, Atmos. Chem. Phys., 22, 5961–5981, https://doi.org/10.5194/acp-22-5961-2022, 2022.

Michalak, A. M., Randazzo, N. A., and Chevallier, F.: Diagnostic methods for atmospheric inversions of long-lived greenhouse gases, Atmos. Chem. Phys., 17, 7405–7421, https://doi.org/10.5194/acp-17-7405-2017, 2017.

Reuter, M., M. Buchwitz, M. Hilker, J. Heymann, H. Bovensmann, J. P. Burrows, S. Houweling, Y. Y. Liu, R. Nassar, F. Chevallier, P. Ciais, J. Marshall and M. Reichstein, 2017: How much CO_2 is taken up by the European terrestrial biosphere? Bull. Amer. Meteo. Soc., 98, 665–671, doi:10.1175/BAMS-D-15-00310.1.

Talagrand, O.: Errors. A posteriori diagnostics, in Advanced Data Assimilation for Geosciences Lecture Notes of the Les Houches School of Physics: Special Issue, June 2012, edited by: Blayo, É., Bocquet, M., Cosme, E., and Cugliandolo, L. F., Oxford University Press, 608 pp., 2014.

Wang, Y., Wang, X., Wang, K., Chevallier, F., Zhu, D., Lian, J., He, Y., Tian, H., Li, J., Zhu, J., Jeong, S., Canadell, J.G., The size of the land carbon sink in China. *Nature* 603, E7–E9 (2022a). https://doi.org/10.1038/s41586-021-04255-y

Wang, J., Feng, L., Palmer, P.I. *et al.* Reply to: The size of the land carbon sink in China. *Nature* **603**, E10–E12 (2022b). https://doi.org/10.1038/s41586-021-04256-x

WMO (2000), Guidelines on performance assessment of public weather services, TD 1023, https://library.wmo.int/idurl/4/44889

WMO (2005), Guidelines on Quality Management Procedures and Practices for Public Weather Services, TD 1256, https://library.wmo.int/idurl/4/30442

Document History

Version	Author(s)	Date	Changes
1.0	All authors	27/11/2023	Initial version
1.1	All authors	13/12/2023	Address reviews

Internal Review History

Internal Reviewers	Date	Comments
Dominik Brunner	09/12/2023	Only small comments. Overall the report is very clear, comprehensive, and to the point.
Andrea Kaiser-Weiss	11/1/2023	Nice document, all aspects covered, easy to read and well referenced. Only when touching the topic of users it could be improved by being more specific.

This publication reflects the views only of the author, and the Commission cannot be held responsible for any use which may be made of the information contained therein.