**SOP 7 Data reporting and archiving for dissolved CH4 and N2O**

Authors so far: Damian Arevalo1, Annette Kock1, Cara C. M. Manning2, Tobias Steinhoff1 ...

1GEOMAR Helmholtz Centre for Ocean Research Kiel, Research Division 2: Marine Biogeochemistry

2Plymouth Marine Laboratory, Plymouth, UK

**Section 1. Background**

The first six SOPs contained in this Best Practice Guide for CH4 and N2O outline quality control and quality assessment activities designed to improve and standardize measurements of dissolved CH4 and N2O. This seventh SOP builds on the initial six documents to highlight good practices in data reporting for dissolved CH4 and N2O. Although this SOP is the final Chapter in this Best Practice Guide, it should be read prior to the collection and analysis of samples. This is because data reporting is intimately associated with SOP procedures including sample analysis, calibration, internal controls, and consensus material (Figure 1).



Figure 1. The SOPs described in this Best Practice Guide for dissolved CH4 and N2O are connected to each other. Underpinning all of these activities is high quality data reporting which should be useful, rigorous, and standardized. The schematic is adapted from Bushnell et al. (2019)

Achieving high quality data reporting is critical to acquire accurate spatial and temporal distributions of CH4 and N2O in the marine environment. However until now, there has not been any community agreed upon reporting procedures for CH4 and N2O datasets. Nationally funded data repositories provide a general list of parameters that should accompany any dataset, but there has not been specific consideration of what needs to be reported for CH4 and N2O. As a result, there is a large range in the amount of information submitted alongside published CH4 and N2O datasets. This SOP outlines the data reporting necessary to help establish a coordinated observing program for CH4 and N2O. Whilst it is specific for CH4 and N2O, it builds on published recommendations for sustained oceanographic observations such as hydrographic surveys (Hood et al, 2010), and also for specific parameters such as trace elements (Worsfold et al., 2019), CO2 (Wanninkhof et al., 2019) and ocean acidification experiments (Riebesell et al., 2011).

**Section 2. General background on data archiving**

There has been large advances in the establishment of public data repositories in alignment with the FAIR (Findable, Accessible, Interoperable, and Reusable) Guiding Principles (Wilkinson et al., 2016; <https://www.go-fair.org/fair-principles>). Most funding agencies have mandatory policies for the archiving of data in these repositories and certain scientific publishers also recognize their value (Brewer, 2017)⁠. All CH4 and N2O data must be archived to a domain-specific open access repository e.g. BCO-DMO or PANGAEA that is aligned with the FAIR Principles and capable of issuing a digital object identifier. Once the domain repository is selected, the repository can be consulted to determine the specific metadata and data formatting requirements. This should be done as early as possible in the research program to ensure that all relevant data is retained and to facilitate the data submission process. To guarantee that the shared data provide a maximum additional benefit to the scientific community, scientists should take a number of principles into account upon data archiving. For CH4 and N2O, data must thus be:

(i) archived in an open-access, searchable domain-specific repository with a globally unique and persistent identifier.

(ii) archived in an easily accessible data format (see e.g. recommendations of ONRL (Hook, Vannan, Beaty, Cook, & Wilson, 2010)⁠), which can be read and imported using non-proprietary software.

(iii) published with a clear and accessible data usage license.

(iv) archived with a rich metadata description as described below in Sections 3 and 4 for discrete and continual measurements, respectively.

**Section 3. Data reporting procedures for discrete measurements**

The required information when reporting discrete measurements of CH4 and N2O datasets is listed below.

a) Position, date and time. Position data must be given in decimal degrees. The time format must enable unambiguous identification of day/month.

b) Sampling depth. The precision of the position, depth and time data must be sufficient to reflect the temporal and spatial resolution of the dataset, but also be able to capture the temporal and spatial variability that can be expected in the respective dataset. Geospatial and temporal accuracy and precision may be of higher importance in coastal areas than in the open ocean, and for near-surface measurements than for deep samples. Scientists should be aware of the spatial and temporal representativeness of their respective measurements and choose the precision accordingly.

c) Sample identifier. If N2O and CH4 samples have been collected, but not archived together with other parameters, the same identifier should be used for all parameters taken during the same campaign (e.g. the same combination of CTD cast number and bottle number).

d) Water temperature and salinity. These data are needed for the calculation of N2O and CH4 solubilities. N2O and CH4 data must be reported together with corresponding water temperature and salinity data in units of in situ temperature and practical salinity (the measured variables) rather than conservative temperature and absolute salinity (the calculated variables) (McDougall & Barker, 2011). The precision of the temperature and salinity data must be sufficient that the calculation of the solubility does not introduce significant additional uncertainty to the N2O and CH4 data.

e) N2O and/or CH4 concentrations. N2O and CH4 data must be submitted in specified units e.g. moles kg-1. If quality flags and/or uncertainty ranges for individual N2O and CH4 data are provided, the metadata information must include a detailed definition of these. If partial pressure or mole fraction of N2O and CH4 are recorded, the data must include additional information if the data are reported in wet or dry conditions.

f) Atmospheric N2O and CH4 measurements. If available, atmospheric data can also be included e.g. if the atmospheric measurements have been matched to the corresponding seawater measurements. In this case, the metadata description must include all of the metadata described below for atmospheric data, as well as the way the atmospheric measurement data were matched to the oceanic data.

g) Ancillary meteorological data. It may be useful to include ancillary meteorological data (e.g. wind speed, relative humidity, air pressure) into the oceanic data spreadsheet, e.g. if the data are used in a publication for calculating sea-air gas fluxes or atmospheric N2O or CH4 data are provided. Their metadata description must include information on the analysis method, the sampling height and sampling period and the way the atmospheric measurement data were matched to the oceanic data.

h) Ancillary seawater parameters. If ancillary parameters (e.g. oxygen, nutrients) are submitted together with the N2O and CH4 dataset, their metadata description should include information on the analysis method, the precision and accuracy of the data. If the data originates from additional laboratories, the appropriate permission and acknowledgements should be included.

i) Method used. Reference to a published method, including any modifications.

j) Precision. The analytical precision is a commonly reported parameter, provided by the mean and standard deviation of a selected number of samples. This can be obtained by analysis of samples simultaneously or over the course of the entire sample run. Knowing the precision for both of these activities is useful

k) Accuracy. Quantifying the analytical accuracy is not straightforward as there is no dissolved standard or reference material for CH4 and N2O. However, Chapter 3 outlines how air-equilibrated seawater can be used to provide an estimate of analytical accuracy. While this is most applicable for low nanomolar CH4 and N2O concentrations, it can still provide a reference value for elevated concentrations.

l) Calibration scale. Analytical instrumentation used to quantify CH4 and N2O is calibrated using gaseous standards. Information of the standards used should be provided and cross-calibration to a primary standard (a gaseous standard with high certainty) referenced when this has been achieved. Details of the calibration curve, including any assessment.

m) Sample storage. Sample storage can be an inevitable part of fieldwork involving dissolved CH4 and N2O. It was shown in Chapter 1 that sample storage of CH4 samples can impact the integrity of the final concentrations. Therefore, details of sample storage should be provided and an estimate of its impact provided.

n) Cross-referencing to other datasets. There now exist several time-series datasets and repeat hydrographic surveys and discrete expeditions. Cross-over stations help take into account different sampling and analytical procedures. Could also include atmospheric or underway measurements.

o) Consensus material. One of the best assessments of analytical bias is through the measurement of certified reference material alongside the analysis of samples. Examples of where this has been successfully employed in oceanography include DIC and GEOTRACES. However, there no reference material available for CH4 and N2O, and even if planned consensus material is made available it can only be used to assess analytical bias not stored for analysis alongside batches of samples. This situation increases the importance of providing information about any accompany analysis of internal controls.

**Section 4. Data reporting procedures for underway measurements**

The required information when reporting underway measurements of dissolved CH4 and N2O concentration data is listed below.

a) Dataset title

b) Version

c) Platform type (e.g. ship, mooring)

d) Platform name

e) Expedition name

f) Campaign dates

g) Dataset creator (Name, Organisation, Email)

h) PI contact details (Name, Organisation, Email)

i) Host data centre (including link)

j) Relevant publication (if available, including doi)

k) List of variables included in the data set (using SI units)

* Minimum metadata: position (latitude and longitude), UTC date and time
* Minimum ancillary data: equilibration temperature and pressure, *in situ* temperature, *in situ* salinity, atmospheric pressure (normalised to sea level)
* Recommended ancillary data: true wind speed and direction (with height of the anemometer/sensor), ship speed, oxygen concentration (in µmol kg-1)
* CH4 / N2O units nmol kg-1

l) Information of sample collection

* Location and depth of water intake
* Location and height of air intake
* Location and type meteorological sensors (e.g. height above sea-level)

m) Information on measurement technique(s)

* Publication describing the method used
* If unavailable, include:
* Analytical Instrument (Manufacturer/Model)
* Measurement interval
* Measurement uncertainty (mole fractions)
* List of calibration gases (indicating traceability to NOAA/WMO scale)
* Calibration uncertainty
* Mention of solubility function used (literature reference)
* Uncertainty of ancillary data (temperature, pressure, salinity)

n) Other remarks

**Section 5. The MEMENTO database**

MEMENTO has been initiated to collect global oceanic N2O and CH4 data, and to create a harmonized global CH4 and N2O dataset that is freely available for all interested users (Kock, Bange). MEMENTO is hosted at GEOMAR, Helmholtz Centre for Ocean Research Kiel, and administered by the GEOMAR Digital Research Services Centre. The database is accessible through the MEMENTO webpage (<https://memento.geomar.de/database>). Data submission to MEMENTO can be performed by submitting individual datasets using the MEMENTO data template, by submitting individual files in other formats. Further details on the required information that should be contained in the data submission can be found at the MEMENTO webpages. When using the data submission template of MEMENTO, the users are asked to provide all essential information and ancillary variables to fulfill the recommendations of this SOP. With the permission of the data originators, data submissions to MEMENTO will be forwarded to PANGAEA (<https://www.pangaea.de/>). This guarantees the long-term archive of the data and provides the datasets with a DOI. Data ownership will remain with the data originators. Alternatively, data originators can forward the information on their data submission to their national or other research-programme-specific databases and the MEMENTO data administrators will extract this dataset and transfer it to MEMENTO.

**References**

Berg, J. (2017). Editorial retraction. *Science*, *356*(6340), 812. <https://doi.org/10.1126/science.aan5763>

Brewer, P. (2017). Do you expect me to just give away my data? *Eos*, *98*. https://doi.org/10.1029/2018EO081175

Bushnell M, Waldmann C, Seitz S, Buckley E, Tamburri M, Hermes J, Henslop E and Lara-Lopez A (2019) Quality Assurance of Oceanographic Observations: Standards and Guidance Adopted by an International Partnership. Front. Mar. Sci. 6:706. doi: 10.3389/fmars.2019.00706

Hood, E. M., Sabine, C. L., and Sloyan, B. M. (2010). The GO-SHIP Repeat Hydrography Manual: A Collection of Expert Reports and Guidelines, IOCCP Report Number 14, ICPO Publication Series Number 134. http://www.go-ship.org/HydroMan.html

Hook, L. A., Vannan, S. K. S., Beaty, T. W., Cook, R. B., & Wilson, B. E. (2010). *Best Practices for Preparing Environmental Data Sets to Share and Archive*. Oak Ridge, TN, U.S.A.: Oak Ridge National Laboratory Distributed Active Archive Center. https://doi.org/10.3334/ORNLDAAC/BestPractices-2010

Riebesell, U., Fabry, V.J., Hansson, L. and Gattuso, J.P., 2011. Guide to best practices for ocean acidification research and data reporting. Office for Official Publications of the European Communities.

Wilkinson, M. D., Dumontier, M., Aalbersberg, Ij. J., Appleton, G., Axton, M., Baak, A., Blomberg, N., Boiten, J.-W., da Silva Santos, L. B., Bourne, P. E., Bouwman, J., Brookes, A. J., Clark, T., Crosas, M., Dillo, I., Dumon, O., Edmunds, S., Evelo, C. T., Finkers, R., Gonzalez-Beltran, A., Gray, A. J. G., Groth, P., Goble, C., Grethe, J. S., Heringa, J., ’t Hoen, P. A. C., Hooft, R., Kuhn, T., Kok, R., Kok, J., Lusher, S. J., Martone, M. E., Mons, A., Packer, A. L., Persson, B., Rocca-Serra, P., Roos, M., van Schaik, R., Sansone, S.-A., Schultes, E., Sengstag, T., Slater, T., Strawn, G., Swertz, M. A., Thompson, M., van der Lei, J., van Mulligen, E., Velterop, J., Waagmeester, A., Wittenburg, P., Wolstencroft, K., Zhao, J. and Mons, B.: The FAIR Guiding Principles for scientific data management and stewardship, Sci. Data, 3, 160018 [online] Available from: https://doi.org/10.1038/sdata.2016.18, 2016.

Worsfold PJ, Achterberg EP, Birchill AJ, Clough R, Leito I, Lohan MC, Milne A and Ussher SJ (2019) Estimating Uncertainties in Oceanographic Trace Element Measurements. Front. Mar. Sci. 5:515. doi: 10.3389/fmars.2018.00515