Technical note: Facilitating the use of low-cost methane (CH\(_4\)) sensors in flux chambers – calibration, data processing, and an open-source make-it-yourself logger

David Bastviken, Jonatan Nygren, Jonathan Schenk, Roser Parellada Massana, and Nguyen Thanh Duc
Department of Thematic Studies – Environmental Change, Linköping University, 58183 Linköping, Sweden

Correspondence: David Bastviken (david.bastviken@liu.se)

Received: 19 December 2019 – Discussion started: 20 January 2020
Revised: 31 May 2020 – Accepted: 9 June 2020 – Published: 14 July 2020

Abstract. A major bottleneck regarding the efforts to better quantify greenhouse gas fluxes, map sources and sinks, and understand flux regulation is the shortage of low-cost and accurate-enough measurement methods. The studies of methane (CH\(_4\)) – a long-lived greenhouse gas increasing rapidly but irregularly in the atmosphere for unclear reasons, and with poorly understood source–sink attribution – suffer from such method limitations. This study presents new calibration and data processing approaches for use of a low-cost CH\(_4\) sensor in flux chambers. Results show that the change in relative CH\(_4\) levels can be determined at rather high accuracy in the 2–700 ppm mole fraction range, with modest efforts of collecting reference samples in situ and without continuous access to expensive reference instruments. This opens possibilities for more affordable and time-effective measurements of CH\(_4\) in flux chambers. To facilitate such measurements, we also provide a description for building and using an Arduino logger for CH\(_4\), carbon dioxide (CO\(_2\)), relative humidity, and temperature.

1 Introduction

Methane (CH\(_4\)) is the second most important of the long-lived greenhouse gases (GHGs). Its global warming potential per mass on a 100-year time horizon (GWP\(_{100}\)) is 28–34 times greater than the GWP\(_{100}\) for carbon dioxide (CO\(_2\)), and its relative increase in the atmosphere since 1750 has been much greater than for other GHGs (e.g., Myhre et al., 2013). The atmospheric CH\(_4\) originates from multiple sources including incomplete combustion, handling natural gas or biogas, or microbial CH\(_4\) production in agriculture, ruminant digestive tracts, and other anaerobic environments such as wetlands and lakes – the microbial CH\(_4\) accounting for approximately two-thirds of the total emissions (Saunois et al., 2016). The high diversity of sources, many yielding fluxes that have high spatiotemporal variability, makes it difficult to quantify fluxes and understand flux regulation without a large number of local measurements. At the same time, common methods to measure fluxes rely on expensive equipment or labor-demanding procedures. Consequently, the CH\(_4\) flux from various sources is poorly constrained. This is exemplified by the discovery of inland waters and flooded forests as two large global CH\(_4\) sources during the last decade (Bastviken et al., 2011; Pangala et al., 2017). Greater availability of measurement approaches that are inexpensive enough to allow many measurements and assessment of both spatial and temporal variability simultaneously would greatly improve our ability to assess landscape CH\(_4\) fluxes and flux regulation.

There is substantial interest in sensitive, small, and affordable CH\(_4\) sensors, but so far the commercially available low-cost CH\(_4\) sensors were typically developed for explosion warning systems and thereby for high concentrations (mole fractions at percent levels). CH\(_4\) detection at percent levels is of high interest for environmental research, including the measurements of CH\(_4\) ebullition, and for such applications cost-efficient sensor applications have been presented (e.g., Maher et al., 2019). For measurements of other types of CH\(_4\) fluxes, sensors with robust and reliable detection at lower levels (mole fractions in the parts per million range) are needed. Previous attempts to use and calibrate such sensors at parts per million (ppm) levels have been promising (Eugster and
Kling, 2012) but also reported remaining challenges, and the use of these sensors in environmental research or monitoring has not yet become widespread. The direct monitoring of atmospheric CH$_4$ mole fractions to resolve fluxes, demanding fast and accurate detection of changes on the order of 10 ppb, still represents a challenge for low-cost sensors. However, relevant mole fraction ranges for flux chamber studies (2–1000 ppm depending on environment, chamber type, and deployment times) appear within reach.

One commercially available low-cost sensor type, showing promising performance in previous studies, is represented by the TGS 2600 tin dioxide (SnO$_2$) semiconductor sensor family made by Figaro. This type of sensor has been evaluated multiple times at CH$_4$ mole fractions near ambient background air (from 1.8 to 9 ppm; different ranges in different studies (Eugster and Kling, 2012; Casey et al., 2019; Collier-Oxandale et al., 2018; van den Bossche et al., 2017). Given their low cost, they performed surprisingly well under nonsulfidic conditions (H$_2$S may interfere with the sensors), although it was challenging to generate calibration models with $R^2 > 0.8$, and the reported interferences from relative humidity (RH) and temperature ($T$) were large (van den Bossche et al., 2017). We here evaluate one member of this sensor family for a larger CH$_4$ range (2–719 ppm), selected to be appropriate for use in automated and manual flux chambers. We propose further development of the equations suggested by the manufacturer for data processing, and we provide guidance on how to address the sensor response to humidity ($H$), RH, and $T$ in flux chamber applications. We also describe a simple CH$_4$–CO$_2$–RH–$T$ logger based on the evaluated sensors, an Arduino microcontroller, and a corresponding logger shield.

2 Methods

2.1 The CH$_4$ sensor

The sensor used in this study is the Figaro NGM2611-E13, which is a factory pre-calibrated module based on the Figaro TGS 2611-E00. The factory calibration is made at 5000 ppm, 20°C, and 65 % RH. The CH$_4$ mole fraction in the factory calibration is not relevant for applications near atmospheric concentration, but the NGM2611-E13 is compact and ready to use, facilitating its integration with data loggers and equipment for flux measurements (e.g., automated flux chambers; Duc et al., 2013; Thanh Duc et al., 2020). The detection range given by the manufacturer is 500–10000 ppm, but the sensor has been used successfully for measuring indoor ambient concentrations of methane (2–9 ppm) (van den Bossche et al., 2017). The potential of another similar sensor, the Figaro TGS 2600, for atmospheric concentration monitoring has been investigated (Eugster and Kling, 2012; Collier-Oxandale et al., 2018; Eugster et al., 2020). The main difference of the TGS 2611-E00, compared to the TGS 2600, is the presence of a filter that reduces the interference of other combustible gases with the sensor, making it more selective towards CH$_4$ (Figaro Product Information TGS 2611, 2020). The TGS 2611-E00 is also more than 10 times cheaper than the sensor used in Duc et al. (2013) and its detection range is wider, allowing for reliable measurements of concentration above 1000 ppm, which makes the sensor potentially useful in both low- and high-emitting environments.

2.2 Calibration setup

The sensor evaluation setup was designed to resemble real measurement conditions in floating flux chambers in aquatic environments. The sensors were placed in the headspace of a plastic bucket positioned upside down on a water surface in a tank. We used a 7 L plastic bucket in which we located 20 TGS 2611-E13 sensors connected to electronic circuitry and a sensor signal logging system described in detail separately (Thanh Duc et al., 2020). The chamber headspace was continuously pumped from the chamber, through the measurement cell of an ultra portable greenhouse gas analyzer (UGGA, Los Gatos Research), and then back to the chamber. The UGGA served as a reference instrument for CH$_4$. The air $T$ and RH inside the chamber were measured with 10 K33-ELG CO$_2$ sensors (Senseair) which have an accuracy of ±0.4 °C and ±3 % RH (Bastviken et al., 2015). The large number of K33-ELG sensors was due to a separate test of wireless data transfer (outside the scope of this work), and one K33-ELG sensor would have been enough for this CH$_4$ sensor study. The entire installation was placed in a climate room to allow for varying $T$, and thereby also absolute humidity ($H$) in the chamber headspace. $T$ and $H$ covary under field conditions in measurements near moist surfaces. Therefore although $T$ and $H$ were not controlled independently, their variability under this calibration setup was reflecting chamber headspace conditions under in situ field conditions.

The CH$_4$ concentration in the chamber was changed by direct injections of methane into the chamber by syringe via a tube. The CH$_4$ concentrations during the calibration experiments ranged from 2 to 719 ppm. We performed multiple separate calibration experiments at different $T$ and RH levels ranging from 10 to 42 °C and 18 %–70 %. At temperatures below 20 °C the RH was usually 50 %–70 %, while at temperatures > 20 °C, RH ranged from 18 % to 60 %. The highest absolute water vapor mole fraction was 35 000 ppm H$_2$O. Values were recorded once per minute. $T$ and RH values from the K33-ELG sensors were averaged among all sensors (because all sensors were in the same chamber and we could not link specific K33-ELG sensors to specific CH$_4$ sensors).

The response time to changing chamber headspace CH$_4$ levels differed between the sensors situated in the chamber (responding rapidly) and the UGGA (delayed response time due to the residence time of the measurement cell and tub-
The second step (Step 2) regards calculation of CH₄ mole fractions from Rₛ/R₀. Several models were tested, where the CH₄ mole fractions were estimated as a function of Rₛ/R₀, H, T, and a constant to consider offsets that may differ among sensors. We tried several linear and power functions. In line with viewing the sensor surface as an active site where CH₄ and H₂O compete for space, the H effect was in some models represented as an interaction with the sensor response.

In all above cases, models were generated by curve fitting in Python using the scipy.optimize curve_fit function. Predicted CH₄ mole fractions were evaluated by comparison with mole fractions independently measured by the UGGA. The specific model equations are provided in Tables 1 and 2. We tested models using RH or H (which was calculated from RH and T (Vaisala_Technical_Report, 2013). Each evaluation included a combination of both steps above and generated one set of fitted parameters per sensor used, including the parameters for Steps 1 and 2.
2.3.2 Simplified calibration approaches without dedicated calibration experiment data (Approaches II and III)

The model combinations from Steps 1 and 2 above that generated the best fit with the minimum number of parameters were selected for tests of two simplified calibration approaches. In Approach II we tested if model parameters in Step 2 can be predicted from parameters derived in Step 1, hypothesizing that the derived model parameters in both Step 1 and Step 2 reflect the sensor capacity to respond to CH₄ and humidity levels as well as the individual sensor offset. If correct, the parameters in Step 1 should be correlated with parameters in Step 2. If this correlation is strong enough, it may be possible to predict parameters in Step 2 from parameters in Step 1, which can be derived from measurements at background air concentrations under the natural variation in humidity (e.g., the diel variability), as a part of the regular measurements, preferentially using data when the atmospheric boundary layer is well mixed (e.g., windy conditions). Under such conditions atmospheric background CH₄ concentrations can be relatively accurately assumed. Hence this Approach II would not require access to sensor calibration chambers, nor expensive reference gas analyzers, which in turn would make sensor measurements available much more broadly. To test this approach, we searched for the best possible regression equations to predict Step 2 parameters from Step 1 parameters, and then we used these equations to estimate CH₄ mole fractions and compared this with the UGGA reference measurements.

In Approach III we evaluated if reasonable accurate Step 1 and 2 equations can be derived from the combination of (i) minimum background atmospheric level Vₐ, at different humidity and (ii) a limited number of randomly collected independent manual flux chamber samples. If so, a few manual samples during the regular measurements could replace tedious dedicated calibration experiments. To test this approach the calibration data for each sensor were subsampled randomly and these random subset data were combined with the minimum Vₐ data to derive calibration parameters as done in Approach I. Using these parameters, the CH₄ mole fractions for all of the calibration data were estimated and compared with observed values. Monte Carlo simulations were run to test effects of the number of random reference samples (1–50) and the methane concentration ranges (3–500 or 3–50 ppm, respectively) in the subset data.

2.3.3 A low-cost Arduino-based CH₄–CO₂–RH–T logger

To facilitate use of the sensors and our results, we also gathered instructions for how to build a logger for CH₄, CO₂, RH, T measurements, using the CH₄ sensor tested here, and the Senseair K33 ELG CO₂–RH–T sensor described elsewhere (Bastviken et al., 2015), a supplementary DHT22 sensor for RH and T, an Arduino controller unit, and an Adafruit Arduino compatible logger shield with a real-time clock (input voltage 7–12 V; 10 bit resolution; Fig. 1). This development was based on sensor specifications and the open-source knowledge generously shared on the internet by the Arduino user community. The full description of this logger unit is found in the Supplement.

3 Results and discussion

The results of different Step 1 and Step 2 calibration equations are shown in Tables 1 and 2. The models including H were equal or superior to models using RH. This is reason-
Table 2. Model results for Step 2 of the data evaluation, i.e., the determination of methane (CH$_4$) mole fractions (ppm) from the sensor response expressed as $R$ (corresponding to $R_S/R_0$) using different calibration models. (RH), $H$, and $T$ as defined in Table 1. The model parameters $a$, $b$, $c$, $d$, $e$, $f$, and $K$ are constants for each sensor derived by curve fitting. The models were evaluated via a linear regression of observed versus predicted CH$_4$ mole fractions, where $k$ and $M$ are the slope and the intercept, respectively. SE denotes standard error, $R^2$ the adjusted coefficient of determination (mean and minimum to maximum for the 20 sensors tested), and RMSE the root-mean-square error (ppm). The table shows the most successful subset of all models tested. $N = 619–930$ per sensor in total and 203–313 for the data subset with CH$_4$ mole fractions < 50 ppm. See text for details.

| Model No. | Equation | $V_0$ | CH$_4$ | $k$ | Observed = $k$ · predicted + $M$ |
|-----------|----------|-------|-------|-----|---------------------------------
| 1         | $CH_4 = aR + b(RH) + cT + K$ | V1 | 719 | 0.024–0.036 | $k$ (min–max) | $M$ (min–max) | $R^2$ | RMSE (min–max) |
|           |          |      |       | 1.006–0.010 | $-3.2 \times 10^{-7}$ to $3.5 \times 10^{-7}$ | 5.8–8.2 | 0.58 | 117 (104–127) |
| 2         | $CH_4 = aR^b + c(RH)^d + eT^f + K$ | V1 | 719 | 0.003–0.006 | $-8.8 \times 10^{-7}$ to $4.2 \times 10^{-7}$ | 1.6–2.6 | 0.96 | 35.9 (32–45) |
| 3         | $CH_4 = aR^b + c(RH)(aR^b) + dT(aR^b) + K$ | V1 | 719 | 0.002–0.003 | $-4.3 \times 10^{-7}$ to $3.2 \times 10^{-7}$ | 0.90 | 1.00 | 11.4 (9–16) |
| 4         | $CH_4 = aR^b + cH(aR^b) + dT(aR^b) + K$ | V1 | 719 | 0.001–0.003 | $-3.3 \times 10^{-7}$ to $4.1 \times 10^{-7}$ | 0.87 | 1.00 | 10.6 (8–16) |
| 5         | As No. 4 | V2 | 719 | 0.001–0.003 | $-4.1 \times 10^{-7}$ to $3.6 \times 10^{-7}$ | 0.82 | 1.00 | 9.8 (8–15) |
| 6         | As No. 4 | V3 | 719 | 0.001–0.003 | $-2.2 \times 10^{-7}$ to $2.8 \times 10^{-7}$ | 0.82 | 0.99 | 9.8 (8–14) |
| 7         | As No. 4 | V4a | 719 | 0.001–0.003 | $-5.6 \times 10^{-7}$ to $1.3 \times 10^{-7}$ | 0.84 | 1.00 | 10.4 (8–15) |
| 8         | $CH_4 = aR^b + cH(aR^b) + K$ | V4a | 719 | 0.001–0.003 | $-7.8 \times 10^{-7}$ to $1.4 \times 10^{-6}$ | 0.84 | 0.99 | 10.4 (8–15) |
| 9a        | As No. 8 with equation V4b to determine $V_0$. | V4b | 719 | 0.007–0.014 | $-4.1 \times 10^{-8}$ to $8.1 \times 10^{-8}$ | 0.33 | 0.98 | 2.1 (2–3) |
| 9b        | As No. 9a with lower max CH$_4$ level. | V4b | 50 | 0.001–0.012 | $-108$ to $1.1$ | 0.39 | 1.00 | 74 (18–150) |
| 10a       | As No 8. Parameters $a$, $b$, $c$, and $K$ estimated from relationships with parameters in V4b; see text. | V4b | 719 | 0.001–0.024 | $-122$ to $1.9$ | 0.43 | 0.99 | 88 (20–154) |
| 10b       | As No 8. Parameters $a$, $b$, $c$, and $K$ estimated from relationships with parameters in V5b; see text. | V5b | 719 | 0.006–0.021 | $-51$ to $-14$ | 0.30 | 0.98 | 28 (21–35) |
| 10c       | As No 10a with lower max CH$_4$ level. | V4b | 50 | 0.006–0.021 | $-6.5$ to $7.1$ | 0.41 | 1.00 | 13 (8.8–20) |
| 11a       | As No 8. Parameters $a$, $b$, $c$, and $K$, derived from six to seven minimum $V_0$ values at different $H$ and 20 samples at random CH$_4$ mole fractions between 3 and 500 ppm. | V4b | 719 | 0.002–0.004 | $-0.7$ to $0.5$ | 0.17 | 0.97 | 2.5 (2–3) |
| 11b       | As No 11a with the 20 random samples at CH$_4$ mole fractions between 3 and 500 ppm. | V4b | 50 | 0.008–0.017 | $-0.7$ to $0.5$ | 0.17–0.41 | 0.97 | 2.5 (2–3) |

---

$^a$ Minimum and maximum mean intercepts for the group of 20 sensors. The confidence interval around the mean intercept was ±1.1 ppm in Model 7 (having the lowest RMSE).

$^b$ Monte Carlo simulations with 1000 runs generating random data subsets used for deriving the model parameter ranges.
able because it is the absolute water molecule abundance that influences the sensor response. Hence, models using $H$ were prioritized. In Approach I, several Step 1 models, including a constant minimum $V_L$, and power, linear, and Michaelis–Menten-based equations gave similar $R^2$ (0.85 to 0.9) and root-mean-square error (RMSE) when comparing predicted versus observed results (Table 1). The effect of $T$ appeared negligible compared to $H$, which may be related to the built-in heating of the active sensor surface (280 mW; this heating was focused on a small part of the sensor, and no self-heating around sensors was detected). It is possible that the Michaelis–Menten equation is superior over the full theoretically possible $H$ range. However, under our experiment conditions, covering normal field $H$ levels, the combination of best fit and minimum number of parameters in Step 1 was found for a simple linear equation with $H$ (Model V4 in Table 1), which was used for later tests of Approaches II and III.

The tests of different equations in Approach I, Step 2, showed that power relationships with $H$ and $T$ represented as interactions with the sensor response performed best (Table 2, Model $\geq$ 4). With the exception of Models 10a–c, all these models had in the regression of observed versus predicted a slope and intercept that were statistically indistinguishable from 1 and 0, respectively ($p < 0.05$), and an $R^2$ of 0.98–1.00 (Table 2, Figs. 2, S1, and S2). Again, $T$ had a marginal effect and $H$ was clearly most important. Hence, while Model 7 including $T$ in Table 2 had the lowest RMSE (9.8), Model 8 represented a good compromise between minimum number of parameters and low RMSE (10.4) and was used in Approaches II and III. The nonlinear response of the sensor yielded a stronger and more coherent response at low CH$_4$ levels, and a large part of the uncertainty was generated at the higher CH$_4$ levels in the studied range (Figs. 2, S1, and S2). Near the atmospheric background at 2 ppm, the confidence interval for individual sensor response was on the order of ±1.1 ppm (Model 7 having the lowest RMSE). Hence, the presented calibration equations have a limited accuracy in terms of absolute CH$_4$ mole fractions and are not optimized for high-precision measurements at atmospheric background levels (as shown by SE for the model intercept corresponding to 0.16–1 ppm; Table 2). However, high $R^2$ and low SE for the slope of several models indicate that the relative change of CH$_4$ levels over time, which is the core of flux chamber measurements, can be assessed efficiently with the sensors if calibrated properly (Table 2).

Approach II, deriving all calibration equations from a small set of minimum $V_L$ values using Models V4 or V5 (Table 1) and 10a–c (Table 2), generated substantially greater RMSE. Most of this RMSE change was due to less accurate prediction of the intercept. The $R^2$ and slope standard error range remained similar to the other models (Table 2), but the actual slope values could deviate substantially from 1 and varied considerably among sensors (in contrast to the models for all other approaches always having slopes close to 1 and similar among sensors; Figs. 2, S1, and S2). Thereby, Approach II could lead to a large bias in absolute mole fractions. This crude generation of calibration equations may be adequate primarily for assessing relative change over time measured by the same sensor, and cross comparisons among sensors should be avoided when using this approach. Examples of equations for the parameter estimation in Approach II are provided in Table S1 in the Supplement. Applying Approach II to a smaller concentration range yielded a considerably lower RMSE (Table 2, Model 10c).

Approach III (Model 11a and 11b in Table 2) showed that as few as 10–20 reference samples, collected at random occasions during actual measurements, could substantially reduce the RMSE of the calibration models, reaching close to the lowest levels based on the 619–930 measurements and the full range up to 719 ppm in Approach I (Table 2 Model 11a; Fig. S3). The concentrations of the reference samples did not appear important for the RMSE within a given specific data range. However, simulations using data for CH$_4$ mole fractions below 50 ppm only generated much lower RMSE than using all data (Table 2, Model 11b). This supports the conclusion that the sensors are more sensitive and give a stronger relative response in the low part of the studied concentration range.

An overview of approaches to derive calibration models for this type is shown in Table S2. The challenges found regarding monitoring of background atmospheric levels were confirmed by our study, while use for relative changes of greater magnitudes in flux chambers appears promising based on our results, also with a simplified calibration (Approach III). As a general note for all approaches when used...
under variable environmental conditions, the best precision may be achieved by using absolute partial pressure units for both CH₄ and humidity, thereby compensating for variability in atmospheric pressure. In addition, for use above the maximum absolute humidity in this study, extra care is advised to check sensor capacity and behavior. An example of real field data is provided in Fig. 3. This example illustrates that the direct sensor response is heavily influenced by humidity and that the calibration equations shown in Tables 1 and 2 are needed to reveal the CH₄ part of the sensor signal.

Further, Fig. 3 illustrates that sensors can be highly useful in very variable environments with rapid changes in humidity and when using data loggers with 10 bit resolution only. Optimal calibration procedures in stable environments and with higher logger resolution would likely indicate better sensor performance, while we deliberately focused on calibration procedures closely linked to field use with inexpensive equipment to provide information of relevance for as many conditions as possible.

Long-term stability of the sensors was not addressed here but is important for environmental use. Other studies of the same type of sensors show promising results. For example, van den Bossche et al. (2017) found no drift over 31 d. Eugster et al. (2020) using the similar Figaro TGS 2600 sensor for outdoor measurements over 7 years concluded that the drift was on the order of 4–6 ppb yr⁻¹. This suggests that the sensor drift is modest even when exposed to variable weather over long time. However, it is possible that the drift can be faster under some conditions, and regular drift checks are therefore advised.

4 Conclusions

The main conclusions can be summarized as follows.

– The tested CH₄ sensors are suitable for use in flux chamber applications if there are simultaneous measurements of relative humidity and temperature.

– Sensor-specific calibration is required.

– Occasional independent reference samples during regular measurements can be an alternative to designated calibration experiments. Background atmospheric levels in combination on the order of 10–20 in situ reference samples at other CH₄ levels can yield rather accurate calibration models.
– For the highest accuracy regarding absolute CH$_4$ concentrations, careful designated calibration experiments covering relevant environmental conditions are needed.

– These results, together with the increased accessibility of low-cost sensors and data logger systems (one example described in the Supplement), open supplementary paths toward improved capacity for greenhouse gas measurements in both nature and society.

**Code and data availability.** Python code for data evaluation and the calibration experiment data are available from the main author upon request. Please note that both the code and the data are specific for the experimental setup. The Python code needs modifications for use with other data, and the CH$_4$ sensor data cannot represent results from other sensors as sensor-specific calibration is needed.


**Supplement.** The supplement related to this article is available online at: https://doi.org/10.5194/bg-17-3659-2020-supplement.

**Author contributions.** DB and NTD designed and supervised the study. NTD designed and built the experimental setup, the sensor units, and the data logging system. JN and NTD performed the calibration experiment. The analysis of the sensor data was led by DB with contributions from NTD, JS, and JN. DB and NTD developed the logger units presented in the Supplement. RP helped to build and test logger units and drafted a user manual. DB wrote the first complete draft of the manuscript and led the manuscript development, with contributions from NTD, JS, RP, and JN.

**Competing interests.** The authors declare that they have no conflict of interest.

**Acknowledgements.** We thank Sivakiruthika Balathandayuthabani, Magnus Gälfalk, Balathandayuthabani Panneer Selvam, Gustav Pajala, David Rudberg, Henrique Oliveira Sawakuchi, Anna Sieczko, Jimmy Sjögren, and Ingrid Sundgren, for stimulating discussions regarding sensor use.

**Financial support.** This research was supported by the European Research Council (ERC) under the European Union’s Horizon 2020 Research and Innovation program (grant agreement no. 725546; METLAKE), the Swedish Research Council VR (grant no. 2016-04829), and FORMAS (grant no. 2018-01794).

**Review statement.** This paper was edited by Ji-Hyung Park and reviewed by three anonymous referees.

**References**


