Improving a plot-scale methane emission model and its performance at a Northeastern Siberian tundra site

Y. Mi¹, J. van Huissteden¹, F. J. W. Parmentier², A. Gallagher¹, A. Budishchev¹, C. T. Berridge¹, and A. J. Dolman¹

¹Department of Earth and Life Sciences, Earth and Climate Cluster, VU University Amsterdam, De Boelelaan 1085, 1081 HV Amsterdam, the Netherlands
²Department of Physical Geography and Ecosystem Science, Lund University, Sölvegatan 12, 223 62 Lund, Sweden

Received: 19 October 2013 – Accepted: 28 November 2013 – Published: 19 December 2013

Correspondence to: Y. Mi (y.mi@vu.nl)

Published by Copernicus Publications on behalf of the European Geosciences Union.
Abstract

In order to better address the feedbacks between climate and wetland methane (CH$_4$) emissions, we tested several mechanistic improvements to the wetland CH$_4$ emission model Peatland-VU with a longer Arctic dataset than any other model: (1) inclusion of an improved hydrological module; (2) incorporation of a gross primary productivity (GPP) module; (3) a more realistic soil-freezing scheme.

A long time series of field measurements (2003–2010) from a tundra site in Northeastern Siberia is used to validate the model, and the Generalized Likelihood Uncertainty Estimation (GLUE) methodology is used to test the sensitivity of model parameters.

Peatland-VU is able to capture both the annual magnitude and seasonal variations of the CH$_4$ flux, water table position and soil thermal properties. However, detailed daily variations are difficult to evaluate due to data limitation. Improvements due to the inclusion of a GPP module are less than anticipated, although this component is likely to become more important at larger spatial scales because the module can accommodate the variations in vegetation traits better than at plot-scale.

Sensitivity experiments suggest that the methane production rate factor, the methane plant oxidation parameter, the reference temperature for temperature dependent decomposition, and the methane plant transport rate factor are the most important parameters affecting the data fit, regardless of vegetation type. Both wet and dry vegetation cover are sensitive to the minimum water table level, in addition to the runoff threshold and open water correction factor and the subsurface water evaporation and evapotranspiration correction factors, respectively.

These results shed light on model parameterization and future improvement of CH$_4$ modelling. However, high spatial variability of CH$_4$ emissions within similar vegetation/soil units and data quality prove to impose severe limits on model testing and improvement.
1 Introduction

Northern boreal and Arctic permafrost regions contain a large quantity of climate vulnerable carbon (Zimov, 2006; Koven et al., 2011; Hugelius et al., 2013). Peatlands are a common feature of this region and cover $3.556 \times 10^6$ km$^2$, or approximately 19% of the northern circumpolar permafrost region (Tarnocai et al., 2009; van Huissteden and Dolman, 2012).

Methane (CH$_4$) emissions from peat soils are strongly linked to the atmospheric CH$_4$ concentration (Yu et al., 2013; IPCC, 2007; Umezawa et al., 2012). These emissions are the net result of a balance between CH$_4$ production by methanogenic microorganisms within anaerobic soil and CH$_4$ oxidation by methanotrophic microorganisms in aerated soil and in plants (van Huissteden et al., 2009). These processes are controlled by water table position, soil temperature, methane transport pathways, and substrate availability and quality (Walter and Heimann, 2000).

Key drivers of methanogenesis and oxidation are predicted to change (IPCC, 2007). The feedbacks between climate and peatland CH$_4$ emission, however, are complex. Precisely how climate change will impact the northern high latitudes is not fully understood. Increases in air temperatures and/or precipitation levels may result in an increase in active layer depth, anoxic soils conditions and raised soil temperatures, therefore producing elevated CH$_4$ emissions. Alternatively, CH$_4$ emissions may decrease if permafrost degradation improves soil drainage (van Huissteden et al., 2011). The latter is particularly true in discontinuous permafrost regions (Smith, 2005).

To constrain these uncertainties, several research stations around the Arctic have started to measure methane fluxes and associated parameters. However, field measurements from the northern high latitudes are spatially limited due to the vast size and remoteness of this region, the extreme climate and logistical difficulties. Process-based computer simulations could bridge the gap between the field situation and the global knowledge level, and increase our understanding of these feedbacks. Ultimately,
this process enables us to make observationally justifiable projections of future climate via the up-scaling of available field data.

Several process-based models have been developed to simulate CH$_4$ fluxes at different spatial scales (Liebner et al., 2011; Cao et al., 1996; Christensen et al., 1996; Bekki and Law, 1997; Wania et al., 2009, 2010; Zhuang et al., 2004). Compared with large-scale models, site-scale models have the advantage of using site-specific chemical and physical properties and plant data to parameterize the model, thus enabling an interactive process representation of soil, biosphere and atmosphere. In addition, they can be validated against field measurements under a variety of conditions (Granberg et al., 2001; Zhang et al., 2002; Arah and Stephen, 1998; van Huissteden et al., 2006; Comer et al., 2000).

Conversely, large-scale modelling of CH$_4$ fluxes always requires aggregated and simplified information on vegetation and soil, which are more difficult to parameterize or validate. However, these models, particularly when coupled to climate models, are highly important in understanding the feedbacks between climate change and CH$_4$ emissions (Gedney, 2004; Kaplan, 2002; Chen and Prinn, 2006; Nicolsky et al., 2007).

Large-scale models are usually up-scaled from small-scale models (Walter et al., 2001; Segers and Leffelaar, 2001). If a parameter has a strong influence on the modelled fluxes on the plot-scale, it is then likely that it also has a large influence in an up-scaled version of this model. Depending on model structure, this may hold also for other models that use the same or similar parameters (van Huissteden et al., 2009). Careful parameter sensitivity analysis and analysis of the effects of model structure is therefore necessary to reduce modelling uncertainty and eliminate the parameters that do not contribute significantly to model accuracy in order to eliminate redundant computational enterprise.

In this study we test the updated plot-scale version of the methane emission model Peatland-VU on the Kytalyk tundra site in Northeastern Siberia, which has a long data series (2003–2010). The new model version (compared to van Huissteden et al., 2006) has been improved by: (1) including a hydrological module to dynamically calculate the
water table when actual water table data is not available, as this is a key factor controlling the environment for CH$_4$ production; (2) adopting a GPP module to simulate gross primary productivity for each vegetation type, which affects the substrate availability, and (3) improving the soil-freezing scheme with a more realistic calculation of the soil thermal conductivity.

We employ the GLUE (Generalized Likelihood Uncertainty Estimation) methodology (Lamb et al., 1998; Beven, 2004; van Huissteden et al., 2009) to test the model sensitivity with validation data from on-site chamber measurements. We also test the effects of changes in model structure, such as the addition of the GPP module.

2 Study area and methods

2.1 Study area

The study area is located in the Kytalyk nature reserve (70°49’ N, 147°29’ E) in Sakha Republic, Northeastern Siberia. This is a continuous permafrost zone with a permafrost thickness in excess of 300 m (Fig. 1).

The climate of this region is continental. Climatic records from the nearest weather station (Chokurdakh 70°37’ N, 147°53’ E, elevation 48 m, Köppen climate classification ET, polar tundra, data from 1961 to 1990) show a mean annual air temperature of −14.3°C, a July average of 9.5°C and a January average of −34.6°C. The mean annual precipitation is 232 mm, approximately half of which falls as rain during the growing season whilst the rest falls as snow. Although this amount of precipitation is similar to the yearly total in semi-arid areas, total evaporation is much lower, thus the soil remains very wet and the wilting point is normally not exceeded (Parmentier et al., 2011).

Summer temperatures are highly variable due to the large contrast between winds from the North and South: northern winds blow cold air from the East Siberian Sea while southern winds bring hot summer air from the Siberian interior. The wind direction can change over hours, and therefore affects both air and soil surface temperature.
Through this effect on temperature, wind direction has an indirect influence on the CH$_4$ flux.

Snowmelt occurs between mid May and the start of June, while bud break begins at the end of June or early July. The growing season usually stops at the end of August or early in September, when temperatures begin to drop below 0°C and sunlight diminishes.

We include two morphological units in this study: the frequently inundated river floodplain and a river terrace with tundra vegetation. Floodplain sediments consist of silt or silty peat in the back swamps, and silt and fine sand in the levees. Mineral soils dominate, but soils below dense sedge vegetation have a thin (<10 cm) organic horizon. Active layer thickness ranges from 25 to 55 cm, averaging 42 cm, and is thickest below the levees. The river terrace soil consists of silt overlain by 15 to 30 cm of peat. The active layer in dry areas ranges from 12 to 28 cm, while it is markedly thicker in wet areas (22 to 50 cm) (van Huissteden et al., 2005).

The floodplain levees are overgrown with tall Salix shrubs, while the back swamps are dominated by meadows with grasses and sedges, ranging from low grass (Arc-tophila fulva) in the lowest, wettest, parts to tall grasses and sedges (Carex arctisberica, glacialis, Eriophorum angustifolium) in the slightly less wet parts. The tundra terrace shows a larger diversity of vegetation types: dry shrubs such as Betula nana and Salix pulchra in dryer regions (polygon rims, low palsas), moist tundra with Eriophorum vaginatum tussocks in less well drained areas, and wet areas with Sphagnum sp. and Carex sp. and some Eriophorum angustifolium.

The water table depth is spatially heterogeneous. In the sedge dominated wet depressions it could be 20 cm above the soil surface, while in the shrub dominated dry areas, the soil may be unsaturated down to the top of the permafrost. Although the water table depth is highly heterogeneous, soil moisture changes are less abrupt and most soils remain moist during the growing season.

Depending on geomorphology, soil moisture and vegetation, the measurements are categorized into floodplain dry (FD) and wet (FW1, FW2) groups and river terrace dry...
(TD) and wet (TW1, TW3, TW4) groups, where “wet” is defined as largely water saturated soils, with a water table not lower than 5 cm below the surface. These locations are shown in (Fig. 1).

FD is associated with floodplain levees, dominated by tall Salix shrub. FW1 and FW2 are floodplain back swamps with different vegetation. TD represents higher elevations in the micromorphology associated with permafrost and high soil ice content (low palsas and ridges along ice wedges), with thin active layers and unsaturated soil horizons above the permafrost. TW1 is situated in diffuse drainage lines consisting of interconnected depressions and characterized by dense and species-poor Carex eriophorum vegetation. TW3 and TW4 are associated with ice wedge polygon centers, which show more species-rich vegetation with Sphagnum sp., sedges and Potentilla palustris.

2.2 Model description

Peatland-VU is a process-based model designed to simulate methane (CH$_4$) and carbon dioxide (CO$_2$) flux from a column of peat soil with unit surface area (van Huissteden et al., 2006). In this study we focus on improving the performance of the CH$_4$ flux module and applying it to permafrost peat soils.

The CH$_4$ flux depends on CH$_4$ production in the anaerobic soil zone, consumption by methanotrophic bacteria in the aerated zone and the different transport pathways to the atmosphere (Walter and Heimann, 2000). The model subdivides the soil column into fifteen layers of equal thickness (0.1 m) and calculates the flux rate of each layer, before integrating over all layers to obtain the total flux:

$$\frac{\partial}{\partial t} D_{CH_4}(t,z) = -\frac{\partial}{\partial z} F_{\text{diff}}(t,z) + Q_{\text{eb}}(t,z) + F_{\text{pl}}(t,z) + R_{\text{pr}}(t,z) + R_{\text{ox}}(t,z)$$ (1)

where $D_{CH_4}(t,z)$ is the CH$_4$ concentration at time $t$ and depth $z$, $F_{\text{diff}}$ is the diffusive flux, $Q_{\text{eb}}$ and $F_{\text{pl}}$ represent ebullition and plant transport, $R_{\text{pr}}$ and $R_{\text{ox}}$ are the CH$_4$ sources and sinks due to CH$_4$ production and oxidation.
CH_4 production is temperature dependent and linearly related to substrate availability (Eq. 2), where R_0 (µM h^{-1}) is a constant rate factor, C_{fresh} is the carbon concentration in the fresh soil organic matter (SOM) reservoirs, T(t,z) is the soil temperature at depth z and time t, and T_{ref} is a reference temperature, approximately the yearly mean soil temperature below the water table.

\[ R_{pr}(t,z) = R_0 \cdot C_{fresh} \cdot Q_{10}^{(T(t,z)-T_{ref})/10} \] (2)

The soil temperature gradient is calculated using a thermal diffusion equation. The thermal diffusivity is estimated from the volumetric heat capacity and thermal conductivity. An improved method for estimating soil thermal conductivity from generic soil composition data (Balland and Arp, 2005) has replaced the calculation in the previous version of Peatland-VU, which was based on Williams and Smith (1991). This adjustment allows the model to simulate a more realistic active layer depth due to a better approximation of the thermal conductivity of the frozen soil.

The substrate pool is the sum of several labile SOM classes (manure, root exudates, dead roots and litter, dead microbes) and stable SOM (peat, humic matter). The labile SOM reservoirs are replenished by gross primary production (GPP) and manure addition (the latter is used in managed wetlands and not included in the experiments described here). In the previous version, GPP is estimated from the temperature of the topmost ten centimeters soil or fed by another model. We incorporate a GPP module to Peatland-VU in this study, which has two options: method (i), adopted from Lund–Potsdam–Jena Dynamic Global Vegetation Model (LPJ) (Sitch et al., 2003; Haxeltine et al., 1996), within which GPP is dependent on vegetation structures and phenology, driven by input of climatology, soil type and atmospheric CO_2 concentration, and method (ii), a simpler set of equations based on Shaver et al. (2007), within which GPP is primarily related to temperature, light and leaf area index (LAI).

CH_4 oxidation is temperature sensitive and its calculations depend on the CH_4 concentration at each time-step (Eq. 3), where K_m (µM) and V_{max} (µM h^{-1}) are the
Michaelis–Menten constants. \( Q_{10,ox} \) determines the temperature sensitivity of the process.

\[
R_{ox}(t, z) = -\frac{V_{\text{max}} \cdot C_{\text{CH}_4}(t, z)}{K_m + C_{\text{CH}_4}(t, z) \cdot Q_{10,ox}} \cdot \frac{T(t, z) - T_{\text{ref}}}{T_{10}}
\]  

(3)

The anaerobic-aerated zone boundary is defined by water table, which in the previous version of Peatland-VU is prescribed. We add a hydrological module based on Granberg et al. (1999) and Yurova et al. (2007) to simulate dynamic water table positions. This has been extended to include saturated zone water transport dependent on water head, distance to nearest drainage line, and saturated hydraulic conductivity.

Finally, The model simulates CH\(_4\) transport in three ways: transport by diffusion above and below water table, transport by bubble formation (ebullition) below water table, and transport through plants.

### 2.3 Measurements

Methane flux data are available for Kytalyk for every year between 2003 to 2010, recorded using round static chambers attached to an Innova 1312 photoacoustic gas analyzer (see van Huissteden et al., 2005).

During the campaign days, the flux data were collected from 53 different chamber placements that are representative of the measurement groups mentioned above. Accompanying each flux measurement, active layer depth and water table level was measured manually next to the chamber collar. These data are compared with model simulations. Despite that all measurements have been carefully screened, uncertainties in the data are inevitable and may lead to poor model-data fits.

Key sources of uncertainties include induced ebullition during measurements or possible chamber leakage, the CH\(_4\) flux calculation method (the change in CH\(_4\) concentration within the chamber is presumed to be linear, which in fact is not the case (Conen and Smith, 1998; Forbrich et al., 2010; Levy et al., 2011), the spatial and temporal variability (which is difficult to capture due to the limited distribution of measurement spots...
and measurement frequency), and the measurements error (e.g. when the measuring plot is ponded, we take the vertical distance between the water surface and soil surface as water table level. However, it is hard to separate the soil surface from the vegetation roots, particularly with the presence of tall sedges or peat moss hummocks).

### 2.4 Model set-up

The parameters of the soil physics module, CH$_4$ module and primary production and soil organic mater production module are described in detail in van Huissteden et al. (2006, 2009); their values are set up in accordance with the vegetation types.

The new parameters for the water table and GPP modules are listed in Table 1. These values are optimized to on-site measurements or estimates from the literature. For instance, soil texture information (e.g. organic matter content, bulk density), permeability and soil water retention curves are derived from laboratory experiments on soil samples from the Kytalyk site.

The hydrological parameters are as follows: $W_{\text{min}}$ is the lowest water table level of the modelling period, in meters. Negative values denote subsurface water tables. $E_{\text{WT}}$ is a correction factor that reduces evaporation if the water table is below the surface. FW2 has a slightly higher $E_{\text{WT}}$ value due to the shallowest water table level. $Z_{\text{runoff}}$ is the threshold value above which a ponded water layer produces runoff, depending on topography condition. The floodplain backswamp groups FW1 and FW2 can hold a small amount of water above the ground surface before generating runoff, while the low depression groups TW1 and TW4 have much higher $Z_{\text{runoff}}$ values. $E_o$ and $E_{\text{veg}}$ are evaporation correction factors for open water and vegetation properties respectively. $K_{\text{sat}}$ is the horizontal hydraulic conductivity of saturated soil. $D_D$ and $D_L$ are the distance to the nearest drainage line and the water level limit above which the ground water starts to drain.

Major parameters of the GPP module are: $K_{\text{Beer}}$, Beer’s law molar extinction coefficient in unit ground area per unit leaf area, $F_{\text{par}}$, fraction of photosynthetically active radiation, $P_{\text{maxL}}$, Light-saturated photosynthetic rate per leaf area, which is high (∼ 20)
for tussocks dominated group TD2 and low (≈ 14) for *Betula nana* dominant TD1, and $F_{\text{phe}}$, the phenology characteristics, including the base for calculating the heat sum (growing degree days), $F_{\text{phe1}}$, the heat sum when maximum leaf area index (LAI) is reached, $F_{\text{phe2}}$, and the maximum LAI, $F_{\text{phe3}}$.

Snow depth, evapotranspiration and precipitation data, essential to drive the model, are collected from the meteorological tower of the study site and Chokurdakh weather station, situated 30 km southwest of the former. The latter of which is used only to gap-fill the in situ observations. To complete the evapotranspiration dataset, we use the FAO (Food and Agriculture Organization of the United Nations) potential evaporation calculator ETo (Raes, 2012); a formula based on air temperature, wind speed and solar radiation.

For the earlier years there are no or fragmentary data from the research sites during the winter months. For the winter of 2011 and 2012, we have complete air temperature data sets from both the research site and the Chokurdakh weather station; by comparing the wintertime data for both stations, it proves that the winter air temperatures at the research site are on average 0.96°C lower than those at the Chokurdakh station. Our site is located north of Chokurdagh, and a heat island effect for the Chokurdagh weather station cannot be excluded. We prepared two air temperature datasets to test to what extent the model is sensitive to the air temperature input change. The first one is a two-meter air temperature record from Chokurdakh, augmented with local site data when available (in summer) and the second one is obtained by subtracting 1°C from the dataset one for the winter data.

### 2.5 GLUE method

In order to assess the influence of input parameters, which are difficult to quantify, we use the Monte Carlo based GLUE methodology (Beven, 2009). For each parameter, the value range is predefined. A set of parameters is selected randomly within their own ranges to complete a model run.
For each group, three thousand Monte Carlo simulations were run to test the performances of the hydrology and CH$_4$ flux calculations separately. The CH$_4$ module test involves three model structure assessments: (a) switch off the GPP module, (b) GPP calculated according to method (i), and (c) GPP calculated according to method (ii). We pre-ran one thousand simulations to rule out the parameters that do not have significant influences on the model output, and incorporating the results from van Huissteden et al. (2009), we select the parameters to be tested.

The results of each model run are compared with site measurements, and evaluated by objective function values. van Huissteden et al. (2009) explain each objective function and its applicable scope. We choose the Nash–Sutcliffe efficiency coefficient to assess the model performance, which is essentially a measure of how good the model performs in predicting the data with respect to an estimate of the fluxes based on the data average: 1 indicates a perfect simulation-data fit; values close to or below 0 indicate an error variance of the same magnitude, or larger, than the variance of the observations:

$$E = 1 - \frac{\sigma_e^2}{\sigma_o^2}$$  \hspace{1cm} (4)

$$\sigma_e^2 = \frac{1}{T-1} \sum_{t=1}^{T} (\hat{y}_t - y_t)$$  \hspace{1cm} (5)

where $E$ is the Nash–Sutcliffe efficiency, $\sigma_e^2$ is the error variance, $\sigma_o^2$ is the variance of the observations, $\hat{y}_t$ is the predicted value at time $t$, and $y_t$ is the observed value.

### 3 Results

The model is run over an eight year period, from 2003 to 2010. In general it performs well, agreeing with site measurements for magnitude, seasonal pattern of the soil thermal properties, water table level and CH$_4$ flux. Due to the low measurement frequency,
in comparison to the model time step (one day), the exact amplitude of data and model output cannot be compared.

3.1 Active layer thickness

Peatland-VU calculates soil temperature and active layer depth. We present the results of the tundra wet group, TW1. The other groups show similar behaviors. Summer campaign data suggest an average maximum active layer depth of 50–60 cm (van Huissteden et al., 2005), however, the measurements are sparse and do not cover the whole thawing season.

Forcing the model by dataset one results in an overestimation of the active layer depth by 125 % (not shown). Using dataset two as the model input improves the simulated active layer thickness considerably; the bias is 12 %. The maximum depth of thaw is about 80 cm, attained in autumn. For most years there is an exact match between modelled and observed active layer thickness (Fig. 2).

3.2 Water table level module

Figures 3–5 illustrate the relationship between measured and modelled water table levels. We present the results from one dry group (FD), one floodplain wet group (FW2) and one tundra wet group (TW1), as these are based on the largest number of measurements. The model shows plausible yearly cycles for all groups. However, since the spatial and temporal variability in the data is high, the coefficient of determination ($R^2$) values of all the groups are low (0.13 for FD, 0.15 for FW2, and 0.23 for TW1).

The model failed in reproducing the water table values for the extremely wet year 2007 for both the wet groups, FW2 and TW1, since during this year FW2 was influenced by prolonged river flooding, also impeding drainage from the nearby TW1 sites. However, we do not expect this has a considerable effect on CH$_4$ production because when the water table is above soil surface, the soil profile is already wholly saturated, creating an optimal anaerobic environment for methanogenesis, therefore the change
of ponded water level does not affect this anaerobic condition. In this case, methane emissions are mostly controlled by changes in the soil temperature and organic substrate.

Despite realistic trends and temporal variations, some simulated water table positions are significantly deeper than the observed ones, for instance, in the year 2005 for the FD group, 2005 for FW2 and 2004 and 2009 for TW1. This can be partly explained by the limited number of measurement locations and high data uncertainty. In addition, this underperformance can also be due to the physical structure of model. This one-dimensional hydrological module does not consider an upstream runoff water source, therefore water table rises due to runoff input from neighboring areas are not captured by the model.

3.3 CH$_4$ flux

We use the summed total CH$_4$ flux of the three transport pathways (diffusion, ebullition and plant-mediated) from model simulations to compare with each of the flux chamber measurements. Figures 6 and 7 present the comparison between simulated and observed (with ±1 standard error bars) from groups FW2 and TW1. The rest of the groups yield similar results (not shown).

The modelled magnitude of the total CH$_4$ flux generally agrees well with the data and the simulated seasonal pattern is clear. The $R^2$ values of the available data are low for all groups, regardless of whether the GPP module is on, or which calculations are used. However, this mismatch should not compromise the capability of the model to investigate the spatial and temporal CH$_4$ flux patterns since the model still explains part of the variance of the data, as indicated by the Nash–Sutcliffe efficiency in the sensitivity analysis section.

Measured fluxes reveal a considerable range among groups, from less than 0 (FD) to more than 100 (FW1) mg m$^{-2}$ h$^{-1}$. Encouragingly, the model reproduced this variation adequately. The lowest emissions are from the tundra dry group TD, with values around 0, while the highest one originates from FW1 during the summer of 2006, with a value 20018
of 77.18 mg m$^{-2}$ h$^{-1}$. Mean daily fluxes during growing seasons produced by Peatland-VU range from 0.05 mg m$^{-2}$ h$^{-1}$ (TD) to 15.1 mg m$^{-2}$ h$^{-1}$ (FW1).

Some of the data from the dry groups show small negative fluxes, suggesting that CH$_4$ is taken up from the oxic topsoil or atmosphere by methanotrophic microorganisms. The model represents this mechanism, but does not reproduce these fluxes very well, although it simulates low to negative fluxes. These very small fluxes also have a very large measurement uncertainty which makes the assessment of the model performance impossible.

Figure 8 shows the simulated CH$_4$ emissions (GPP module off), water table position, active layer depth and observed air temperature of TW1, 2008; the other years yield similar trends (not shown). The emissions begin in early summer, with three peak emissions on 29 June, 17 July and 8 August. All of them are related to high water table positions and rapid increases of temperature. The first peak drops along with the temperature, although the water table stays high. The later two peaks follow directly the decline of water level, which develops a large oxic zone where CH$_4$ is consumed. The importance of water level is also confirmed from the inter-annual emission variations of TW1, whose high 2007 emissions correspond to the high water level during the most productive time (June to August).

### 3.4 Model sensitivity analysis

By applying the GLUE methodology, we randomly selected three thousand combinations of parameters for the water table module and the same amount of parameter sets for the CH$_4$ flux module.

The water table simulation tests show that most of the parameter sensitivities are group-dependent, except for $W_{\text{min}}$, which is highly sensitive at all groups. The dry groups are also sensitive to $E_{\text{WT}}$, while the wet groups are strongly affected by $Z_{\text{runoff}}$ and $E_0$. The vegetation correction factor for evaporation, $E_{\text{veg}}$, is highly important to
both the dry groups and the wet groups, FW1 and TW1. Conversely, the drainage distance, $D_D$, does not have a large influence on any group.

Figure 9 shows the deviations of cumulative distributions of the parameters, between the top 100 model performances (blue lines) and all three thousand simulations (red lines), group TW1. A large deviation of the red and blue lines indicates a high sensitivity of the model to the parameter. There is a clear sensitivity to the parameters of $E_o$, $E_{\text{veg}}$, $W_{\text{min}}$, $Z_{\text{runoff}}$, and $D_L$. Model responses to the rest parameter changes are negligible.

For the CH$_4$ testing we did not include all the available field data, only years with relatively high quality data that covered the majority of the summer period were included. The floodplain data for 2007 was excluded due to the unrepresentativeness of those particular measurements for the floodplain as a whole (F. J. W. Parmentier, personal communication, 2013).

The distributions of all groups indicate that the CH$_4$ production rate factor, $R_0$, the CH$_4$ plant oxidation parameter, $F_{\text{ox}}$, the reference temperature for temperature dependent decomposition, $T_{\text{ref}}$, and the CH$_4$ plant transport rate factor, $F_{\text{plant}}$, contribute significantly to the best model-data fits. The comparison of different groups shows the parameter sensitivity between geomorphology and vegetation.

When the GPP module is switched off, the model performs better over high $R_0$ value ranges (0.35 to 0.50) for groups FW1 and TW3 in comparison with lower values, and conversely for the remaining groups. As for $F_{\text{ox}}$, groups FW2 and TW4 perform better over lower values (0.1 to 0.5) than the other groups (not shown).

The improvement created by the addition of the GPP module is generally insignificant. The higher objective function results for groups FW2, TW1 and TW4 indicate a better model performance with the GPP module on. However, turning off the GPP module gives a better performance for FW1 and TW3 groups (Table 2). Emissions from FW1 tend to be underestimated, especially for the large values, while fluxes from TW4 are overestimated, particularly for low values.

Most of the parameters in both GPP calculation methods show little sensitivity to the total emissions except a plant phenology factor, number of growing degree days
(heat sum), $F_{\text{phe1}}$. This suggests that a complicated GPP module is not necessary for plot-scale CH$_4$ emission modeling, instead meteorological inputs can be used as a surrogate and the effect of photosynthesis can be parameterized in the subroutines.

The two GPP calculation methods do not give different results in most cases except for FW2 where method (i) gives a much higher value than method (ii). The maximum objective function values are listed in Table 2, FW2 runs with GPP module on and all the TW4 simulations have passed the $F$ test $p = 0.1$ probability limit (NS $> 0.3212$).

Same as in Fig. 9, Figs. 10–12 show the deviations of cumulative distributions of the parameters, between the top 100 model performances (blue lines) and all three thousand simulations (red lines), with and without GPP functions for group TW1.

We also tested the sensitivity of CH$_4$ fluxes to the calculated soil temperature and active layer thickness. Forcing the model by the second air temperature data set (1 °C lower than the data from Chokurdakh weather station in winter time), a much better match for the active layer depth was achieved. However, this does not give a better model-data fit with respect to the CH$_4$ fluxes. With the lower winter air temperature input, the best model fit resulting from the GLUE analysis is slightly lower than that with the higher air temperature input, although the number of good-fitting model runs is higher. This holds for experiments with and without the GPP module. Well-performed model runs using the higher air temperature time series have a lower $Q_{10}$ (2–3.5) value and a higher $T_{\text{ref}}$ (15–18 °C) compared to runs using the lower air temperature time series, which have a wider range of $Q_{10}$ values (2–5) and a lower $T_{\text{ref}}$ (5–15 °C).

## 4 Discussion

The model-data comparison shows that the modelled active layer depth and soil temperature is very sensitive to soil surface temperature input. Comparing the air temperature data sets from the research site and Chokurdakh station (winter 2011 to 2012) shows the air temperatures at the research site are on average 0.96 °C lower than...
those at the Chokurdakh station. Subsequently, Subtracting 1°C from the winter air temperature improves the modelled active layer thickness considerably.

These experiments show the importance of modelling the temperature gradient between the soil surface and top of the canopy for accurate modelling of active layer thickness. Importantly, it shows a strong sensitivity of the model to small temperature differences resulting from spatial separation between modelled site and temperature data collection.

The new water table module allows Peatland-VU to simulate more realistic conditions for CH$_4$ production when actual water table data is not available; an essential capability when modelling this data-sparse biome. Long-term tests show that Peatland-VU is able to reproduce the magnitude, spatial and temporal pattern of the water table, even though the spatial variation between floodplain wet groups and tundra dry groups can be large. Nonetheless, detailed daily variations are difficult to evaluate due to limited data on this scale.

The model failed to reproduce the water table values for the wet groups in 2007; an extremely wet year. This is likely due to the assignment of a specific runoff threshold value in the water table algorithm for each group. This was implemented according to our general knowledge of the normal yearly flood situation, over which we assume the surface water is drained as runoff. This approach ignores flooding in situations where water input occurs from upstream, such as in TW1 and FW groups (the TW1 groups also act as drainage channels in very wet conditions). Moreover, in 2007, spring snowmelt, high summer precipitation and poor drainage conditions on the floodplain exacerbated the runoff, which is not reflected in the model structure. This a justifiable omission as the depth of excess flood water does not change the fact that the subsurface is still anoxic and the organisms behave similarly, thus carrying forward negligible changes in methane fluxes regarding the influences of water table position.

In general, the model captured the magnitude and seasonal pattern of the CH$_4$ flux and the disparities between wet and dry groups. The emission fluctuations due to water table level, temperature and active layer thickness are also reflected well in the simu-
lations. However, model-data comparisons give low $R^2$ values, although the Nash–Sutcliffe objective function values are still positive, indicating that the model captures at least part of the variance of the data. This is partly due to the uncertainties in measurements. There is in some cases considerable uncertainty in the exact location of soil surface and water level, in particular for sedge and Sphagnum vegetation (see Sect. 2). The CH$_4$ flux measurements are laced with an immeasurable uncertainty due to induced ebullition during measurement, in particular large fluxes could have been influenced although care has been taken to scrutinize the data (see Sect. 2). Furthermore, this could also be due to the fact that Peatland-VU is a one-dimensional model, which is appropriate to reproduce the important processes from a characteristic soil profile and vegetation unit, while the measurement plots show a very high spatial variability of measured fluxes, even within one vegetation unit, reflected by the error bars on the measurements in Figs. 6 and 7. This variability and uncertainty limits model data comparison, but this mismatch should not compromise the capability of the model to investigate the spatial and temporal CH$_4$ flux patterns.

The GLUE model sensitivity experiment on the water table parameters shows that most of the sensitivity to these parameters is group-dependent. The dry groups are sensitive to the minimum ground water table level, $W_{\text{min}}$, and the subsurface water evaporation correction factor, $E_{\text{WT}}$, while the wet groups are sensitive to $W_{\text{min}}$, the runoff threshold, $Z_{\text{runoff}}$, and the open water correction factor for evaporation, $E_0$. This is in accordance with the fact that the dry groups in general have a lower water table position (below soil surface) compared with the wet groups (ponded).

The results of the CH$_4$ flux sensitivity experiment suggest that the CH$_4$ production rate, $R_0$, the CH$_4$ plant oxidation parameter, $F_{\text{ox}}$, the reference temperature for temperature dependent decomposition, $T_{\text{ref}}$, and the CH$_4$ plant transport rate factor, $F_{\text{plant}}$ are the most important factors affecting the data fit of all groups. From the model runs with higher objective function values, the parameter range differences in vegetation types are found. For example, the tall sedges dominating group FW2 and TW4, generally have low $F_{\text{ox}}$ values, in accordance with high vascular plants methane transport ability
(Bubier et al., 1995; Joabsson and Christensen, 2001). While for the Swamp group, FW1, and Sphagnum group, TW3, $R_0$ values are higher, which suggests that the highs and lows of $\text{CH}_4$ production rate control the flux variability.

The sensitivity experiments for the inclusion of GPP module give higher objective function results for groups FW2, FD, TW1 and TW4, but the results are reversed for FW1 and TW3 groups. We attribute this to the fact that FW1 and TW3 are groups with a very low vegetation canopy, with moss-dominated vegetations and low vegetation total biomass, under which conditions the photosynthesis is driven more by soil water and temperature than by light. Similar results can be found from recent studies (Zona et al., 2011; Street et al., 2012). The results from two GPP calculation methods do not differ significantly except for FW2. Most of the parameters in both GPP calculation methods are not sensitive at any group except a plant phenology factor, $F_{\text{phe1}}$. Therefore we conclude that a GPP module is not essential in process based methane emission modelling, although the model generally performs slightly better with one of the GPP models that we tested, compared to a simple approach based on soil temperature. However, in this study, the module is only run on a plot-scale. On the small scale, a model can be fine-tuned to local conditions, making a GPP module less useful. On the larger scale, however, variations in vegetation traits are to be expected and fine-tuning to one particular site will lead to poor model performance. In such cases, a GPP module may accommodate these spatial variations better, and be of additional benefit to the modeling effort.

The model experiments also show the effect of variations between modelled and measured soil temperature on the modelled fluxes. Deviations in the modelled soil temperature do affect the $\text{CH}_4$ result, but may be compensated by changes in the $Q_{10}$ and $T_{\text{ref}}$ parameters. The model may be improved by adding a vegetation canopy temperature gradient model. However, derivation of soil surface temperature from air temperature based on empirical relations derived from measurement data can also be applied.
There are a number of published plot-scale models attempting to simulate CH$_4$ emissions from wetlands and permafrost tundra available for comparison, especially when it comes to model performance. Segers and Leffelaar (2001) apply a process-based model at the Nieuwkoopse Plassen (west-Netherlands, non-permafrost), Although their modelled CH$_4$ fluxes have the same order of magnitude as the measurements, the model fails in capturing the seasonal pattern. Zhang et al. (2002) present the Wetland-DNDC model and report the simulations of three North America sites, whose $R^2$ values range from 0.03 to 0.76. Wania et al. (2010) parameterizes LPJ-WHYMe at three permafrost sites, BOREAS, Abisko and Ruoergai. The simulated daily flux compared well with data for Ruoergai site, however, the model-data fits are poor for the other two sites. Tang et al. (2010) test three process-based models of different complexities at two Michigan peatlands (non-permafrost), Hollow Bog and Big Cassandra Bog. The simulated fluxes from the latter site are less agreeable than the first one, with the highest $R^2$ values of 0.31 and 0.60 respectively.

The generally low $R^2$ values may suggest that this type of methane emission model performs poorly. However, our analysis shows that this may not be only a matter of the model performance, but also uncertainty in the data in the shape of measurement errors or high spatial variability of fluxes within the same vegetation units. In our case, the model-data fit indicated by $R^2$ is low. However, the Nash–Sutcliffe objective function values still indicate that the model performs better than a flux estimate based on data averages and variance. The model ineffectively captures high emission peaks and negative emissions, and these are also the measurements with the highest uncertainty. Nevertheless, seasonal changes are modelled correctly. One can also question the use of various objective functions to judge model performance. Here, we used the Nash–Sutcliffe efficiency, which is essentially a measure of how good the model performs in predicting the data with respect to an estimate of the fluxes based on the data average, and the $R^2$ goodness-of-fit, which assesses a more exact point-by-point model-data agreement. The high spatial and temporal variability in the data severely limits the use-
fulness of the latter. Other choices of objective functions, which may be better suited to this type of data, need to be considered in the future.

In addition, the length of the data time series and model runs may be a factor resulting in a poorer model-data agreement. Methane models are typically tested with few years of data or less, (e.g. van Huissteden et al., 2006, 2009), for which a good data-model fit is more easily obtainable. As shown by this study and van Huissteden et al. (2009), longer data series may reslut in a poorer fit. We can only speculate on the causes. However, longer data series may contain more inhomogeneities due to measurement methods, instrumental drift not accounted for by calibration, and natural year to year variability (Mastepanov et al., 2013).

5 Conclusions

Our model testing shows that Peatland-VU is able to capture both the annual magnitude and seasonal variations of the CH$_4$ flux, water table position and soil thermal properties. However, detailed daily variations are difficult to evaluate due to the large uncertainty, mainly caused by spatial variation, in the data. The testing of two different models for vegetation primary production, shows that improvements can be gained by adding primary production modelling, although the improvement is not very large, and for differences in model structure, marginal only. However, with a primary production module, our model could also be coupled into climate models and dynamic vegetation models in order to better explain spatial and temporal variations in CH$_4$ emissions from northern permafrost, and to predict responses under future change scenarios. However, the micro-topographical features within one wetland or even one vegetation type can vary widely, which control the hydrology and biogeochemical processes and therefore influence the CH$_4$ fluxes, therefore careful parameterization is needed when upscaling. The large uncertainty in the data due to high spatial variability limits model-data comparison. It may be useful to experiment with objective functions that can behave better under such a situation than the often used $R^2$ measure.
Acknowledgements. The research leading to these results has received funding from the (European Community’s) Seventh Framework Programme (FP7 2007–2013) under grant agreement no. (238366).

References


van Huissteden, J., van den Bos, R., and Marticorena, A. I.: Modelling the effect of water-table management on CO₂ and CH₄ fluxes from peat soils, Neth. J. Geosci., 85, 3–18, 2006. 20008, 20011, 20014, 20026


Table 1. Parameters in the hydrology module and GPP module and their ranges.

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Description</th>
<th>Units</th>
<th>Range</th>
</tr>
</thead>
<tbody>
<tr>
<td>$W_{\text{min}}$</td>
<td>Deepest water table position</td>
<td>m</td>
<td>−0.1 to 0.5</td>
</tr>
<tr>
<td>$E_{\text{WT}}$</td>
<td>Evaporation correction factor for ground water</td>
<td>n/a</td>
<td>1 to 12</td>
</tr>
<tr>
<td>$Z_{\text{runoff}}$</td>
<td>Ponded water depth limit</td>
<td>n/a</td>
<td>0.0 to 0.2</td>
</tr>
<tr>
<td>$E_{\text{o}}$</td>
<td>Evaporation correction factor for open water</td>
<td>n/a</td>
<td>0.1 to 1.5</td>
</tr>
<tr>
<td>$E_{\text{veg}}$</td>
<td>Evaporation correction factor for crop/vegetation</td>
<td>n/a</td>
<td>2.0 to 8.0</td>
</tr>
<tr>
<td>$K_{\text{sat}}$</td>
<td>Horizontal saturated hydraulic conductivity</td>
<td>m day$^{-1}$</td>
<td>0.001 to 0.1</td>
</tr>
<tr>
<td>$D_{\text{D}}$</td>
<td>Down-slope drainage distance</td>
<td>m</td>
<td>2 to 25</td>
</tr>
<tr>
<td>$D_{\text{L}}$</td>
<td>Down-slope drainage water level limit</td>
<td>m</td>
<td>−0.8 to 0.0</td>
</tr>
<tr>
<td>$K_{\text{Beer}}$</td>
<td>Beer's law extinction coefficient</td>
<td>m$^{-2}$ ground m$^{-2}$ leaf</td>
<td>0.4 to 0.9</td>
</tr>
<tr>
<td>$P_{\text{maxL}}$</td>
<td>Light-saturated photosynthetic rate per leaf area</td>
<td>μmol m$^{-2}$ leaf s$^{-1}$</td>
<td>13 to 20</td>
</tr>
<tr>
<td>$F_{\text{phe1}}$</td>
<td>Base for calculating heat sum (growing degree days)</td>
<td>n/a</td>
<td>0.7 to 2.0</td>
</tr>
<tr>
<td>$F_{\text{phe2}}$</td>
<td>Heat sum when maximum LAI is reached</td>
<td>n/a</td>
<td>0.7 to 2.0</td>
</tr>
<tr>
<td>$F_{\text{phe3}}$</td>
<td>Maximum LAI</td>
<td>n/a</td>
<td>0.7 to 2.0</td>
</tr>
</tbody>
</table>
Table 2. The maximum objective function values of GLUE run, with GPP module off, GPP calculated by Method (i) and GPP calculated by Method (ii).

<table>
<thead>
<tr>
<th></th>
<th>FW1</th>
<th>FW2</th>
<th>FD</th>
<th>TW1</th>
<th>TW3</th>
<th>TW4</th>
<th>TD1</th>
<th>TD2</th>
</tr>
</thead>
<tbody>
<tr>
<td>None</td>
<td>0.28</td>
<td>0.26</td>
<td>−0.19</td>
<td>−0.92</td>
<td>−0.09</td>
<td>0.34</td>
<td>−3.32</td>
<td>−0.91</td>
</tr>
<tr>
<td>Method (i)</td>
<td>0.27</td>
<td>0.71</td>
<td>−0.04</td>
<td>0.28</td>
<td>−0.17</td>
<td>0.38</td>
<td>−3.32</td>
<td>−0.91</td>
</tr>
<tr>
<td>Method (ii)</td>
<td>0.20</td>
<td>0.45</td>
<td>0.10</td>
<td>0.31</td>
<td>−0.15</td>
<td>0.39</td>
<td>−3.32</td>
<td>−0.91</td>
</tr>
</tbody>
</table>
Fig. 1. Location of the Kytalyk site and the distribution of different measurement groups. (Left) Circumpolar permafrost map (Rekacewicz, 1998). (Right) Landscape of Kytalyk site (van Huissteden et al., 2005).
Fig. 2. Soil temperature and active layer depth produced by Peatland-VU, tundra wet (TW1) group, year 2004 to 2010. The blue vertical lines on top of the figures indicate the measured active layer thickness, with the length of the line indicating one standard deviation. The black lines show the modeled active layer boundaries.
Fig. 3. Water table produced by Peatland-VU compared with field data, floodplain dry group (FD), 2003 to 2010.
Fig. 4. Water table produced by Peatland-VU compared with field data, floodplain wet group (FW2), 2003 to 2010.
Fig. 5. Water table produced by Peatland-VU compared with field data, floodplain dry group (TW1), 2003 to 2010.
Fig. 6. \( \text{CH}_4 \) flux produced by Peatland-VU compared with field data, floodplain wet (FW2) group, 2003 to 2010. For the measurements (red), the standard error of the flux is shown by an error bar.

\[ \text{CH}_4 \text{ (mg m}^{-2} \text{ h}^{-1}) \]

\[ \text{Year} \]

2003 2004 2005 2006 2007 2008 2009 2010

-10 0 10 20 30 40 50 60 70

GPP off
GPP method (i)
GPP method (ii)
Fig. 7. CH$_4$ flux produced by Peatland-VU compared with field data, tundra wet (TW1) group, 2003 to 2010. For the measurements (red), the standard error of the flux is shown by an error bar.
Fig. 8. CH$_4$ flux (GPP module off), water table level and active layer thickness produced by Peatland-VU and the observed air temperature, tundra wet (TW1) group, 2008.
Fig. 9. Deviations of cumulative distributions of parameters in hydrology module (TW1). (Blue) top 100 model performance. (Red) all 3000 model runs. x-axis denotes the parameter range and y-axis denotes the cumulated parameter proportion. A large deviation of the red and blue lines indicates a high sensitivity of the model to the parameter. \( E_0 \), evaporation correction factor for open water; \( E_{\text{veg}} \), evaporation correction factor for crop/vegetation; \( E_{\text{WT}} \), evaporation correction factor for ground water; \( W_{\text{min}} \), deepest water table position; \( Z_{\text{runoff}} \), ponded water depth limit; \( K_{\text{sat}} \), horizontal saturated hydraulic conductivity; \( D_D \), down-slope drainage distance and \( D_L \), down-slope drainage water level limit.
Fig. 10. Deviations of cumulative distributions of parameters in CH₄ module (TW1), GPP function off. (Blue) top 100 model performance. (Red) all 3000 model runs. x-axis denotes the parameter range and y-axis denotes the cumulated parameter proportion. A large deviation of the red and blue lines indicates a high sensitivity of the model to the parameter. $R_0$, CH₄ production rate factor; $F_{\text{plant}}$, CH₄ plant transport rate factor; $F_{\text{ox}}$, CH₄ plant oxidation factor; $F_{\text{exu}}$, plant root exudate factor and $T_{\text{ref}}$, reference temperature for decomposition.
Improving a plot-scale methane emission model and its performance

Y. Mi et al.

Fig. 11. Deviations of cumulative distributions of parameters in CH$_4$ module (TW1), GPP calculated by method (i). (Blue) top 100 model performance. (Red) all 3000 model runs. x-axis denotes the parameter range and y-axis denotes the cumulated parameter proportion. A large deviation of the red and blue lines indicates a high sensitivity of the model to the parameter. $R_0$, CH$_4$ production rate factor; $F_{\text{plant}}$, CH$_4$ plant transport rate factor; $F_{\text{ox}}$, CH$_4$ plant oxidation factor; $Q_{10}$, $Q_{10}$ factor; $F_{\text{phe1}}$, plant pheology factor, base for calculating heat sum; $F_{\text{phe2}}$, plant pheology factor, heat sum when maximum LAI is reached; $F_{\text{phe3}}$, plant pheology factor, maximum LAI, $F_{\text{exu}}$, plant root exudate factor and $T_{\text{ref}}$, reference temperature for decomposition.
Fig. 12. Deviations of cumulative distributions of parameters in CH\textsubscript{4} module (TW1), GPP calculated by method (ii). (Blue) top 100 model performance. (Red) all 3000 model runs. X-axis denote parameter range, Y-axis denote the cumulated parameter proportion. A large deviation of the red and blue lines indicates a high sensitivity of the model to the parameter. \(R_0\), CH\textsubscript{4} production rate factor; \(F_{\text{plant}}\), CH\textsubscript{4} plant transport rate factor; \(F_{\text{ox}}\), CH\textsubscript{4} plant oxidation factor; \(Q_{10}\), \(Q_{10}\) factor; \(F_{\text{phe1}}\), plant pheology factor, base for calculating heat sum; \(F_{\text{phe2}}\), plant pheology factor, heat sum when maximum LAI is reached; \(F_{\text{phe3}}\), plant pheology factor, maximum LAI, \(F_{\text{exu}}\), plant root exudate factor, \(T_{\text{ref}}\), reference temperature for decomposition and \(K_{\text{Beer}}\), Beer’s law extinction coefficient.