Biogeosciences

Open Access

Testing the applicability of neural networks as a gap-filling method using CH$_4$ flux data from high latitude wetlands

S. Dengel[1], D. Zona[2,3], T. Sachs[4], M. Aurela[5], M. Jammet[6], F. J. W. Parmentier[7], W. Oechel[3], and T. Vesala[1]

[1]University of Helsinki, Department of Physics, P.O. Box 48, 00014 University of Helsinki, Finland
[2]University of Sheffield, Department of Animal and Plant Sciences, Western Bank, Sheffield S10 2TN, UK
[3]San Diego State University, San Diego, CA. 92182, USA
[4]Helmholtz Centre Potsdam, GFZ German Research Centre for Geosciences, 14473 Potsdam, Germany
[5]Finnish Meteorological Institute, P.O. Box 503, 00101, Helsinki, Finland
[6]Center for Permafrost, Department of Geosciences and Natural Resources Management, University of Copenhagen, Øster Voldgade 10, 1350 Copenhagen K, Denmark
[7]Department of Physical Geography and Ecosystem Science, Lund University, Sölvegatan 12, 223 62 Lund, Sweden

Correspondence to: S. Dengel (sigrid.dengel@helsinki.fi)

Abstract. Since the advancement in CH$_4$ gas analyser technology and its applicability to eddy covariance flux measurements, monitoring of CH$_4$ emissions is becoming more widespread. In order to accurately determine the greenhouse gas balance, high quality gap-free data is required. Currently there is still no consensus on CH$_4$ gap-filling methods, and methods applied are still study-dependent and often carried out on low resolution, daily data.

In the current study, we applied artificial neural networks to six distinctively different CH$_4$ time series from high latitudes, explain the method and test its functionality. We discuss the applicability of neural networks in CH$_4$ flux studies, the advantages and disadvantages of this method, and what information we were able to extract from such models.

Three different approaches were tested by including drivers such as air and soil temperature, barometric air pressure, solar radiation, wind direction (indicator of source location) and in addition the lagged effect of water table depth and precipitation. In keeping with the principle of parsimony, we included up to five of these variables traditionally measured at CH$_4$ flux measurement sites. Fuzzy sets were included representing the seasonal change and time of day. High Pearson correlation coefficients (r) of up to 0.97 achieved in the final analysis are indicative for the high performance of neural networks and their applicability as a gap-filling method for CH$_4$ flux data time series. This novel approach which we show to be appropriate for CH$_4$ fluxes is a step towards standardising CH$_4$ gap-filling protocols.

1 Introduction

Methane is one of the most important long-lived greenhouse gases, second only to CO$_2$ (IPCC, 2007), with natural wetlands thought to be the biggest individual source (IPCC, 2007; EPA, 2010). Since the advancement in CH$_4$ gas analyser technology and its applicability to eddy covariance flux measurements (Hendriks et al., 2008; Eugster and Plüss, 2010; Dengel et al., 2011; McDermitt et al., 2011; Peltola et al., 2013), monitoring of CH$_4$ emissions is becoming more widespread in northern regions (Mastepanov et al., 2008; Sachs et al., 2008; Zona et al., 2009; Sturtevant et al., 2012). These measurements contribute to a better understanding of the greenhouse gas balance of the Arctic and subarctic. In order to accurately estimate annual greenhouse gas budgets, time series of high quality gap-free data are required (Falge et al., 2001; Rinne et al., 2007).

Currently there is no consensus on CH$_4$ gap-filling methods. Several studies (Zona et al., 2009; Gažovič et al., 2010; Sturtevant et al., 2012) did not apply any gap-filling to their CH$_4$ flux data. Studies where gap-filling was applied were site dependent and often applied to low resolution, daily

mean values (Hargreaves et al., 2001; Rinne et al., 2007; Ri-utta et al., 2007; Jackowicz-Korczyński et al., 2010; Long et al., 2010; Tagesson et al., 2012), while Wille et al. (2008), Parmentier et al. (2011a) and Forbrich et al. (2012) employed a model in order to recover missing data in their daily, 3 h and 30 min mean data, respectively.

Hargreaves et al. (2001), Rinne et al. (2007), Long et al. (2010) and Tagesson et al. (2012) identified a non-linear relationship between CH_4 flux and peat temperature at depths of 0–10 cm, 35 cm, and 50 cm in subarctic ecosystems, respectively. During extended periods where no dependency on peat temperature was found, Rinne et al. (2007) and Tagesson et al. (2012) applied a simple interpolation to gap-fill these data sets. In addition, Tagesson et al. (2012) applied an exponential regression between half-hourly CH_4 fluxes and friction velocity measured after the soil was completely frozen. No dependency on water table position was found by the two studies mentioned above. A similarly simple peat temperature relationship with CH_4 emissions was also found by Zona et al. (2009) and Jackowicz-Korczyński et al. (2010). Wille et al. (2008) and Sachs et al. (2008) found strong relationships between CH_4 flux, friction velocity and soil temperature at a depth of 20 and 10 cm, respectively. Some of the above mentioned studies considered non-linear relationships to gap-fill their daily averaged CH_4 fluxes, while Parmentier et al. (2011a) provide a method for gap-filling of higher resolution (3 h) data. This method applied a gap-filling model that includes the attenuating effect of atmospheric stability on flux measurements, where methane production was related to soil temperature and water table level. Recently Forbrich et al. (2011) tested various models where peat temperatures at various depths, water table level, barometric pressure and friction velocity were integrated in order to gap-fill their time series. Furthermore, large uncertainties in applied methods do still exist with no common protocol on missing data recovery of CH_4 eddy covariance flux data.

The application of neural networks (Jain et al., 1996; Svozil et al., 1997; Elizondo and Góngora, 2005; Saxén and Pettersson, 2006) for data recovery and gap-filling (Aubinet et al., 2000; Gorban et al., 2002; Papale and Valentini, 2003; Ooba et al., 2006; Moffat et al., 2007 and Schmidt et al., 2008) has proven to be a very reliable tool in several scientific disciplines (Gardner and Dorling, 1998, 1999; Lek and Guégan, 1999; Lee and Jeng, 2002; Toptygin et al., 2005). In atmospheric sciences (Gardner and Dorling, 1998; Toptygin et al., 2005; Chattopadhyay G. and Chattopadhyay S., 2008), applying neural networks in forecasting has become a standard application. Neural networks have the reputation of being a "black box" where transparency is limited in most cases (Elizondo and Góngora, 2005). This partly results from a neural network's high capacity in training itself where coefficients are distributed through fitted weights and spread across several layers to accurately reproduce a given data set. In the current study, we discuss the applicability of neural networks to gap-fill CH_4 flux data from northern high lati-
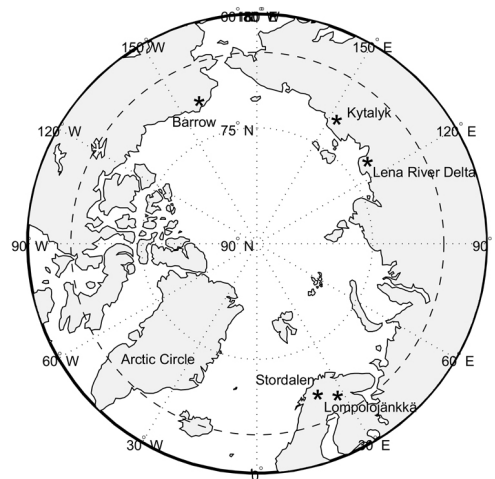

Fig. 1. Map of the Arctic and subarctic region, with Stordalen (Sweden), Lompolojänkkä (Finland), both Siberian sites Lena River Delta and Kytalyk in Russia, and Barrow (Alaska) marked with an asterisk (*) at their respective location.

tude ecosystems (wet sedge tundra, sedge fen and polygonal tundra), some driver dynamics, the advantages and disadvantages of this method, and what information can be extracted from such models.

Since CH_4 is the second most potent, long-lived greenhouse gas in the atmosphere (IPCC, 2007), it is becoming increasingly important to introduce a method which is capable of dealing with such high resolution data combined with auxiliary measurements, and which is easy to implement across a variety of ecosystems. Regarding Arctic and subarctic regions, it is very important to work with time series where data gaps have been filled using reliable methods in order to accurately determine CH_4 emissions, potential annual budgets, and prediction of future emissions under a changing climate (Anisimov, 2007; IPCC, 2007). The data sets introduced in the current study were chosen, as they show distinctive differences in their emission patterns and originate from high latitude ecosystems (Fig. 1), to assure the broad applicability of the introduced methods and results.

The aim of the current study is to test and compare three different neural network approaches as a gap-filling method for high resolution, 30 min and 1 h methane (CH_4) eddy covariance flux data from Arctic and subarctic ecosystems. Included are a limited number of standard meteorological variables that are measured at all sites that act as drivers for methane emissions in such dynamic ecosystems. This novel approach in CH_4 studies is a first step towards standardising CH_4 gap-filling and a contribution to standardising CH_4 measurement protocols.

2 Materials and methods

2.1 Methane flux and meteorological data

The CH_4 eddy covariance flux data used in the current study originates from five distinctively different northern ecosystems (Fig. 1): the subarctic sites of Stordalen (68°20′ N, 19°03′ E, northern Sweden), a mixed mire (Johansson et al., 2006) and Lompolojänkkä a nutrient-rich sedge fen located in the aapa mire region of north-western Finland (67°59′ N, 24°12′ E) (Aurela et al., 2009), and the tundra sites underlain by permanent permafrost: Samoylov Island in the southern central Lena River Delta (72°22′ N, 126°30′ E) (Sachs et al., 2008, 2010), Kytalyk (70°49′ N, 147°29′ E) (Parmentier et al., 2011a, b) and Barrow, a wet sedge tundra in the northern part of the Arctic Coastal plain (71°17′ N, 156°36′ W) (Zona et al., 2009, 2012).

The CH_4 fluxes were measured by the eddy covariance (EC) method (Baldocchi, 2003). Instrumentation used in these six studies were the three-dimensional sonic anemometer R3-50 (Gill Instruments Ltd., Lymington, Hampshire, England) coupled with a closed path Fast Greenhouse Gas Analyser (FGGA, Los Gatos Research, Mountain View, California, USA) in Stordalen; the USA-1 (METEK, Elmshorn, Germany) three-axis sonic anemometer/thermometer and the closed-path DLT-100 fast response CH_4 gas analyser (Los Gatos Research, Mountain View, California, USA) in Lompolojänkkä, and the three-dimensional Solent R3 sonic anemometer (Gill Instruments Ltd., Lymington, Hampshire, UK) and the TGA 100 tunable diode laser spectrometer (Campbell Scientific Ltd., USA) in the Lena River Delta. At the Kytalyk site, a three-dimensional Solent R3-50 sonic anemometer (Gill Instruments Ltd., Lymington, Hampshire, UK) and a closed-path DLT-100 fast response CH4 gas analyser (Los Gatos Research, Mountain View, California, USA) were used in both years, while a WindMasterPro sonic anemometer (Gill Instruments Ltd., Lymington, Hampshire, UK) and the closed-path DLT-100 fast response CH4 gas analyser (Los Gatos Research, Mountain View, California, USA) was used in Barrow. The reader is advised to consult Jackowitz-Korczynski et al. (2010), Aurela et al. (2009), Sachs et al. (2008), Parmentier et al. (2011a, b) and Zona et al. (2009), for more details about the sites, measurements and further instrumentation.

All five sites recorded standard meteorological variables, such as air temperature, solar radiation, soil temperature at various depths, wind speed and wind direction (for the current study wind speed has been decomposed into its horizontal (along wind u) and perpendicular (across wind v) components). Furthermore, CH_4 eddy covariance flux data from Lompolojänkkä were u^* filtered, using $0.1\,\mathrm{m\,s^{-1}}$ as a threshold. Here, fluxes were binned according to u^* and tested on two soil temperature ranges resulting in the threshold value mentioned above. At the Barrow site a threshold of $0.1\,\mathrm{m\,s^{-1}}$ (Zona et al., 2012) was also applied. The data from Kytalyk were filtered for occurrences of high atmospheric stability, prior to including it in the current study. Wille et al. (2008) have shown that u^* itself can enhance methane fluxes and it is therefore useful to obtain also data points at low u^*, as long as storage effects are not an issue. Storage did not represent an issue at the Kytalyk site, and after filtering for high stability no significant effect of u^* was seen. The reader is advised to consult Parmentier et al. (2011a, b) for more background information on the processing of these fluxes. Data introduced in the current study was not previously gap-filled at 30 min and 1 h (Lena River Delta) resolution. In order to integrate the lagged effect of the water table depth (WTD) and that of precipitation as a potential input variable, we included both variables as they were, and in addition lagged precipitation by one and WTD by 12 days, as has been identified by Kettunen et al. (1996) and Suyker et al. (1996).

To give an overview of the data coverage and availability for the current study, and the representation of day and night (07:00–18:30 and 19:00–06:30) respectively, required for the current application (efficient network training) relevant information has been listed in Table 1.

2.2 Artificial neural networks

The topology of a simple multi-layer, feed-forward neural network includes non-linear elements (neurons) that are arranged in successive layers (Fig. 2). The information flows unidirectionally, from the input (covariates) layer to the output (response) layer, through the hidden layer(s) (Jain et al., 1996; Svozil et al., 1997; Elizondo and Góngora, 2005; Saxén and Pettersson, 2006).

In the initial phase, a set of input and target data is used for training and presented to the network many times (also known as iterations). A training data set should have sufficient data to be representative of the overall data set, meaning the whole range of meteorological and flux variability should be available for training (including emission events so that the network can learn such conditions) (Moffat et al., 2010; Papale, 2012). Furthermore, the cross-correlative (Guan et al. 2007) and cross-dependency nature of climatic variables should be taken into account when choosing the appropriate input variables, as some add only little extra information to the network (Moffat et al., 2010). Training is carried out by constantly adjusting the fitted weights so that the network output matches the target data. In order to test the trained network, a new set of input data is fed into the network and the desired output compared with those predicted by the network. The agreement or disagreement of these two data sets is an indication of the performance of the neural network model. A chosen error function measures the difference between predicted and observed output.

One of the drawbacks of neural networks is the non-uniqueness of the global minimum (Hammerstrom, 1993; Nguyen and Chan, 2004) which changes as each training run achieves different weights and results (it is important to find

Table 1. Overview over the data coverage and availability in the current study. All values are 30 min values apart from Lena River Delta which has hourly data. All numbers, except percentage values, are given in number of data rows. Daytime and night-time data cover the time periods between 07:00–18:30 and 19:00–06:30, respectively.

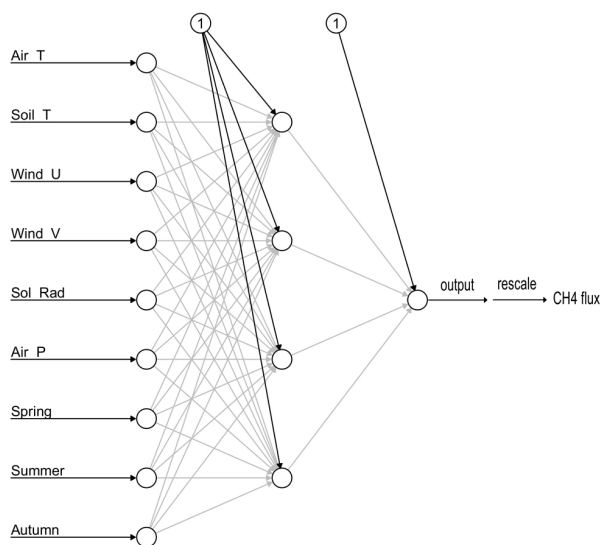| Site	Days (n)	100 % coverage	Actual coverage	Coverage (%)	Coverage (days)	Daytime (n)	Daytime (%)	Night (n)	Night (%)
Stordalen	84	4032	1633	41	34.0	901	55	732	45
Lompolojänkkä	181	8688	3223	37	67.1	1820	56	1258	40
Lena R. delta	94	2256	1508	67	62.8	607	40	763	51
Kytalyk 2008	15	720	540	75	11.3	266	49	274	51
Kytalyk 2009	30	1440	1275	89	26.6	624	49	651	51
Barrow	46	2208	1625	74	33.9	784	48	841	52

Fig. 2. The architecture of the "seasonal" neural network topology used in the current study. Input variables (left side of the network) are fed into the network with weights fitted (along grey arrows) with information flowing unidirectionally to the nodes (marked as circles) within the hidden layer, where a bias (offset) (marked as 1) is added (along black arrows). Here a sigmoid function (activation function) is applied to the weighted sum, leading further to the next layer, the "output" layer where a new set of weights is distributed, together with a bias and the sigmoid activation function before making an estimate for the output value. As the output still has a range of 0–1, it is rescaled prior to replacing missing data values. Actual fitted weights and biases are removed from the graph for clarity. Please see Table 1 for a definition of the input variables.

a set of weights that processes data accurately enough for the application). Another issue with neural networks is the possibility of under- or over-fitting of networks (Hansen and Salamon, 1990; Jain et al., 1996; Svozil et al., 1997). This can happen when data used to train the network is not representative enough for the entire observation span, if the number of hidden layers or neurons is not correct, if the global minimum is overshot or when the network learns the train-

ing pattern well but is underperforming during its validation (poor generalisation) (Jain et al., 1996; Gardner and Dorling, 1998; Nguyen and Chan, 2004; Wang et al., 2005; Saxén and Pettersson, 2006; Stathakis, 2009). To prevent this from happening, one can remove redundant input data (Gunaratnam et al., 2003; Saxén and Pettersson, 2006), reduce or increase the number of neurons in the network and use the appropriate generalisation such as early stopping (Hansen and Salamon, 1990; Amari et al., 1997; Svozil et al., 1997; Wang et al., 2005) or another stopping criteria (Günther and Fritsch, 2010; Fritsch and Günther, 2012).

2.3 Pre-processing of data

Recently, several studies (Zhang and Qi, 2005; Klevecka and Lelis, 2009) pointed out an ongoing debate on whether data should be de-seasonalised prior to applying neural networks. Nelson et al. (1999) showed better results for de-seasonalised time series, while others (Sharda and Patil, 1992; Franses and Draisma, 1997) found that neural networks are able to model seasonality directly and prior de-seasonalisation is not necessary. Regarding gap-filling of atmospheric trace gas fluxes (van Wijk and Bouten, 1999; Aubinet et al., 2000; Carrara et al., 2003; Papale and Valentini, 2003; Ryan et al., 2004; Ooba et al., 2006 and Schmidt et al., 2008), no de-seasonalisation of data was carried out prior to applying neural networks. As artificial neural networks have the ability to deal with complex data sets and as no de-seasonalisation of data was carried out in previous gap-filling studies (see above references) we decided to proceed in the same manner.

Table 1 gives an overview of the data availability and distribution used in the current study. All sites are located within the Arctic Circle and are experiencing the polar day (~ 3 months of no darkness over the summer period). As u^* filtering has been applied to two sites only (very good night-time data coverage, nonetheless), data were not split into the two categories day/night as required for CO_2 gap-filling studies. Furthermore, by including all available data (see effective data coverage in days in Table 1) we avoided a further reduction in number of data points and risking working with models that lack in power or representativeness. Also, grouping

data by different meteorological conditions would lead here to a subjective choice as not every emission event, or weather condition has the same effect on methane fluxes. One of the reasons could be lagged effects of some of the drivers (only the lagged effect of water table depth and precipitation are discussed here) that need further investigation.

We are dealing with seasonal and diurnal data that experience regular and predictable changes. In order to add this seasonal and diurnal effect to their neural network, Papale and Valentini (2003) introduced several fuzzy data sets reflecting the diurnal and seasonal variation to reduce the linear cumulative numerical weight of time in relation to other variables. Adding this type of input to neural networks does not always increase the neural network performance, and has been shown by Schmidt et al. (2008) to have sometimes little effect. Our analysis, however, showed an increase in network performance when including these fuzzy sets. Therefore, four fuzzy sets representing the diurnal effect and three (spring, summer and autumn) mirroring seasonal variations were included in the current study. Winter as a fuzzy set has been excluded as none of the used flux data sets extended into the winter period. Furthermore, these seasonal fuzzy sets were adjusted to correspond to the seasons in northern latitudes (onset of spring is later than in temperate regions, for example). We kept the models simple following the principle of parsimony (Beck, 1943; Bugmann and Martin, 1995) and the quality assurance standards highlighted in Moffat et al. (2010). The predictive ability of a model initially increases with complexity but they do also have the tendency to decline once a model becomes too complicated (Bugmann and Martin, 1995).

The meteorological, soil, and CH$_4$ flux data as well as the fuzzy data sets consist of different magnitudes and units. In order to generalise the data, we have scaled all data from 0 to 1 as has been previously applied by van Wijk and Bouten (1999), Papale and Valentini (2003), Nguyen and Chan (2004) and Moffat et al. (2010). Furthermore, the range between 0 and 1 is also necessary as we are applying a sigmoid activation function (Cybenko, 1989), which has a range of 0–1. By scaling the data that we are feeding into the network, all data is being treated equally and weights can be distributed evenly. A sigmoid function was also used in the output layer (Fig. 2), as has been previously applied by Papale and Valentini (2003).

Three approaches were tested here: two including meteorological variables recorded at all five sites ("seasonal" and "diurnal") and a third, a more thermo-hydrological approach ("lagged"), where the lag effects of water table depth (WTD) and precipitation were incorporated at some of the sites.

2.4 Applying artificial neural networks to data

Introducing the neural network topology (Fig. 2) used in the current study, input variables (left side of the network) are fed into the network with weights fitted and spread across the two layers with information flowing unidirectionally (grey arrows) to the 4 nodes (marked as circles) within the hidden layer, where a bias (offset) (marked as 1) is added.

In this case, the underlying function is simply written as

$$o = f\left(b + \sum_{j=1}^{n} x_i w_{ij}\right) \qquad (1)$$

where n is the number of hidden neurons (4 in our case), x_i represents the input variables (x_1, …, x_9 – in our case), w_{ij} denotes the fitted weights (w_1, …, w_4 – for each neuron) attached to each input variable and b denotes the bias (or offset) that is added to the weighted sum prior to applying the sigmoid activation function (f), leading further to the next layer, where a new set of weights are distributed, together with a bias and the sigmoid activation function before making an estimation of the output values (o). These output values also have a range of 0–1 which were then rescaled to their appropriate physical unit (nmol m^{-2} s^{-1}).

As we are testing and comparing different artificial neural networks as a gap-filling method for CH$_4$ eddy covariance flux data, we utilised the artificial gap length scenarios introduced in Moffat et al. (2007, Appendix A). These scenarios served in extracting the same data rows ($\sim 10\%$) from all data sets which were utilised for model performance testing, necessary for reliable model comparison and evaluation. Moffat et al. (2007) generated secondary data sets by flagging 10 % of the data as artificial gaps. Ten data sets were generated for each gap scenario, each having a different temporal shift in gap distribution. We have chosen three scenarios per gap length (Figs. 3–5). These gap lengths represent very short gaps (V, 1–3) of random 30 min values; short gaps (S, 1–3) of random 4 h gaps, medium (M, 1–3) of 1.5 days, long (L, 1–3) of 12 full days and a mixed scenario (X, 1–3), representing a mix of the above mentioned gap lengths, resulting in 15 scenarios per data set. These scenario labels 1–3 do not correspond to the consecutive scenario numbering introduced in Moffat et al. (2007), but are chosen for simplicity. Scenarios were chosen in such a way that the maximal existent data coverage is achieved, together with an appropriate gap length. There were cases where artificial gap data points coincided with already existing gaps, resulting in a non-uniform length of data pairs used as test data sets. Nevertheless, each scenario extended the already existing gaps by a further 8–14 %. The artificial gap scenarios introduced in Moffat et al. (2007) are for 30 min resolution files which were adjusted for the Lena River Delta data set, where a 30 min artificial gap was applied to the respective hour value. The remaining available data were used for training, avoiding any further reduction of the number of data rows out of reasons mentioned in the previous section.

Several learning algorithms are available for neural network training. In the current study we applied the resilient backpropagation algorithm (Riedmiller, 1994). It is a first-order optimisation algorithm that acts on each weight

separately. It modifies the weights in order to find a local minimum of the error function. The weights are modified going in the opposite direction of the partial derivatives until a local minimum is reached (if the partial derivative is negative, the weight is increased; if the partial derivative is positive, the weight is decreased). This ensures that a local minimum is reached, leading to an efficient and transparent adaptation process (Riedmiller, 1994; Günther and Fritsch, 2010). The resilient backpropagation algorithm has been chosen as it appears to be the fastest in training and most consistent learning algorithm in several studies (Schiffmann et al., 1994; Treadgold and Gedeon, 1996; Kişi and Uncouğlu, 2005). Furthermore, this learning algorithm does limit the size of the weights, which are generally also linked to over-fitting of a model.

In order to test the network's performance, various error functions can be applied. We chose the sum of squared errors (SSE), as previously used by Moffat et al. (2010) and assembled the neural network by implementing the "neuralnet" package (Günther and Fritsch, 2010; Fritsch and Günther, 2012) in R statistical language (R Development Core Team, 2013). Here we applied the built-in training function and modified it accordingly to suit our purposes (maximum number of iterations and threshold value for the partial derivatives of the error function). Moffat (2010) employed a similar method but used the root mean square error (RMSE) as an error function. The process stops when this training error levels off and all partial derivatives of the error function reach the pre-specified threshold value of 0.01 (1 %) that acts as a stopping criteria.

One of the advantages of applying the "neuralnet" package is the possibility to choose this integer specifying the threshold as a stopping criteria (Günther and Fritsch, 2010) that should be achieved during the training phase, along with the maximum number of iterations (per repetition) the network should carry out in order to fulfil our requirements (convergence of the network and finding the local minimum). The trained network is then tested on the extracted $\sim 10\,\%$ of data (artificial gaps) by applying the test function (compute) within the "neuralnet" package. If the error function is equivalent to the negative log-likelihood function then the Akaike information criterion (AIC) can be used to avoid an overfitting of the trained network (F. Günther, personal communication, 2013).

The data were not split into the traditional three-way cross-validation data sets: training, validation and testing subsets (where a model is trained on a training data set with a smaller validation data set that is periodically passed through the network, before being tested on the independent test data). One reason is the use of the applied package, but also to keep the training and testing of the network transparent. Another reason is the applicability of the method to short time series. Would a short time series, such as Kytalyk (2008), be split into three data sets then the resulting analysis would lack any power and representativeness. This way all six data sets were

treated in the same manner. In order to ensure the reliability of all trained models an additional test across sites has been carried out. Moffat (2010) showed that model reliability can also be checked via cross-validation using data from another site.

A reliable model does require a good generalisation (also to avoid overtraining/over-fitting), which was taken into account in the current study. Moffat (2010) does list the main points of having the right number of neurons, stopping the training as soon as the error levels off, pruning the nodes (neurons) and penalising large weights by regularisation. When applying the resilient backpropagation learning algorithm weights are limited internally in size (Riedmiller, 1994). Verifying the reliability of a model not only on its test data but also via cross-validation with data from another site does also contribute to a good generalisation of a model. In order to ensure the reliability of our networks and their applicability in CH_4 studies (also regarding drivers/input parameters) site combinations were chosen according to their flux range and site properties.

There is currently no consensus in the scientific community on the number of neurons that should be used (Svozil et al., 1997; Saxén and Pettersson, 2006; Stathakis, 2009) when applying neural networks to data series. In order to apply the appropriate number of neurons, 25 repetitions were run for a selection of neurons (1–12) to help in choosing the appropriate number of neurons (Järvi et al., 2012) to be applied within the hidden layer of our networks. Furthermore, utilising the integrated AIC's calculation did also support the choice in applied number of neurons. The AIC value reflects the overall fit of a model and it is used to avoid an overfitting of the trained model. While by itself the AIC does not give much information, it only becomes applicable when it is compared to the AIC value of a series of models in which the same data sets (observations) were used. It does also help to determine which model is most parsimonious and an information criterion commonly used for model selection (Hurvich and Tsai, 1989; Burnham and Anderson, 2004). Once no further improvement could be realised in terms of goodness of fit and the networks were able to predict CH_4 fluxes relatively accurately (Pearson's correlation coefficient (r) and the RMSE achieved when testing the trained networks (data not shown)), we decided to proceed further using four neurons within the hidden layer. All networks (15 scenarios per site for all tree approaches) were trained again 50 times and the same error values calculated again (Figs. 3–5).

2.5 Statistical analysis

In order to examine all input variables and their effect on methane fluxes, we applied a simple stepwise regression (a combination of backward elimination and forward selection) in R (R Development Core Team, 2013), in order to search for the best predictors or combinations of predictors from among all available 30 min and 1 h (Lena River Delta)

resolution data. Following the principle of parsimony, we decided on a few meteorological variables and four fuzzy sets representing time of day (morning, afternoon, evening and night) and three representing the seasonal variation (spring, summer and autumn) (Fig. 2, Table 2). This selection helped to prune the network by avoiding insignificant input data (Gunaratnam et al., 2003; Saxén and Pettersson, 2006) and to avoid cross-correlations (to some extent) between input variables, adding little extra information to the network.

We have tested three approaches by using three different sets of input data: the most simple approach included solar radiation as a time of day indicator and the three seasonal fuzzy sets ("seasonal", Table 2), a second approach in which solar radiation has been removed and replaced by the time of day fuzzy sets ("diurnal", Table 2), while a third, a more thermo-hydrological approach is tested by integrating the lagged effect of precipitation and WTD which was applied to four out of six data sets ("lagged", Table 2).

For the first two approaches, we chose those variables that appeared important and available in all data sets. Furthermore it did help to standardise the method and make it applicable to all six different data sets in the same way. The study by Wille et al. (2008) has shown that u^* can enhance CH_4 emissions. Therefore, adding u^* as an input variable where data has previously been u^* filtered, would lead to uncertainties, as u^* filtered data do not provide the information necessary for the network to train and learn such conditions in order to predict CH_4 fluxes occurring under similar conditions.

Precipitation and water table depth can also act as CH_4 drivers (Whalen and Reeburgh, 1992; Roulet et al., 1992; Christensen, 1993; van Huissteden et al., 2008), and can additionally also have a lagged effect on CH_4 emissions (Windsor et al., 1993; Bubier et al., 1995; Kettunen et al., 1996; Suyker et al., 1996). In order to integrate these two hydrological variables, data from Kytalyk were tested by using the current corresponding precipitation values as well as precipitation values lagged by one day, as has been identified in the study conducted by Kettunen et al. (1996). Kettunen et al. (1996) and Suyker et al. (1996) identified a distinctive 12 day lag when investigating the water table fluctuation effect on CH_4 fluxes from boreal wetlands. Stordalen and Lompolojänkkä (both subarctic) were used to test this approach. Barrow was tested by integrating other thermo-hydrological variables, such as soil moisture, vapour pressure deficit (VPD) and soil heat flux as input variables. Lena River Delta data were used to test relative humidity in addition to barometric air pressure as an indicator for atmospheric changes.

In order to investigate the performance of all the networks of the three approaches we estimated the mean r (at 95 % confidence level), AIC and root mean square error (RMSE) as a goodness of fit indicator of the measured and predicted fluxes. The simple RMSE indicates the range of the error for each scenario and each site (shown in their true physical

unit of $nmol\,m^{-2}s^{-1}$). The resulting r values for the training (grey full circles) and test (black full circles) are shown in Figs. 3–5 (lower panel), while the RMSE for the test data are shown as box-plots in the upper panel of these figures. AIC values, necessary for model selection and an indicator for model performance (the lower the value, the better the model) for all six sites, scenarios and model approaches are visualised in Fig. 6. The Pearson correlation coefficient was also utilised as goodness of fit in the cross-validation analysis (Fig. 7).

3 Results

We applied artificial neural networks to six different CH_4 flux data sets originating from subarctic and Arctic regions of Fennoscandia, Siberia and Alaska. We have tested three approaches by using three different sets of input data. These input variables (listed in Table 2) were, amongst others, air temperature (Air T), soil temperature at the depth of 10 cm (Soil T), wind direction (decomposed wind speed into along (u) and across wind (v)), barometric air pressure (Air P) and the fuzzy transformation of the time of day and seasonal variations.

In the first step of the analysis 1–12 neurons were applied to both standardised approaches in order to find the optimal number of neurons, as has been introduced in Järvi et al. (2012) prior to deciding on the final number of neurons included in the hidden layer. The training distribution showed an increase in correlation coefficient value with each added neuron (data not shown here), while some of the test results showed no improvement with increase in neurons added. The distribution indicated that from four neurons onwards no real improvement was visible, be it for short, long or mixed gap length scenarios. This is also confirmed by the lack of statistical significance at the 95 % confidence level (notched Tukey box plots) beyond four neurons leading to the assumption that four neurons were ideal within the hidden layer to be included in this study for all three approaches. The integrated AIC calculations did also support this approach.

The output from the first approach (Table 2) (seasonal effect) is visualised in Fig. 3, showing the r coefficients achieved during the training (in grey) and test phases (in black), by applying four neurons only. In the upper panel the RMSE is given in true physical units ($nmol\,m^{-2}\,s^{-1}$) for the test outcome only. Missing values indicate the lack of data or inappropriate length of data coverage or a distribution outside the visualised plot margins.

Figure 4 shows the outcome from the second approach (diurnal effect) where solar radiation, as an indicator of time of day, was replaced by the four fuzzy sets representing the diurnal effect of the data. Here again, the performance of the network at each site and their error values become visible and comparable. The third approach (thermo-hydrological or lagged effects) in which different input variables were chosen

Table 2. Input variables used in three different neural network gap filling approaches. The input variables listed are: air temperature (Air T), soil temperature at the depth of 10 cm (Soil T), wind direction (Wind U and Wind V (horizontal and perpendicular)), solar radiation (Sol Rad, substituted with photosynthetic active radiation where not available), barometric air pressure (Air P), relative humidity (RH), water table depth (WTD), precipitation, soil heat flux, soil moisture, soil temperature of the polygon rim (Rim temp) and vapour pressure deficit (VPD). Furthermore, included are the fuzzy transformation of the seasonal variation and time of day represented by spring, summer and autumn, as well as the four time periods morning, afternoon, evening and night.

| All sites (approach 1) "seasonal"	All sites (approach 2) "diurnal"	Stordalen (approach 3) "lagged"	Lompolojänkkä (approach 3) "lagged"	Lena River Delta (approach 3) "lagged"	Kytalyk (approach 3) "lagged"	Barrow (approach 3) "lagged"
Air T	Air T	Air T	Air T	Air T	Air T	Air T
Soil T	Soil T	Soil T	Soil T	Soil T	Soil T	Soil T
Air P	Air P	Air P	Air P	Air P	Air P	Air P
Wind U	Wind U	Wind U	RH	Rim temp	Wind U	Soil heat flux
Wind V	Wind V	Wind V	WTD	RH	Wind V	Soil moisture
Sol rad	–	WTD	WTD (12 days)	RH (1 day)	Precipitation	VPD
–	–	WTD (12 days)	–	–	Prec. (1 day)	–
–	–	–	–	–	–	–
Spring	Morning	Spring	Spring	Spring	Spring	Spring
Summer	Afternoon	Summer	Summer	Summer	Summer	Summer
Autumn	Evening	Autumn	Autumn	Autumn	Autumn	Autumn
–	Night	–	–	–	–	–
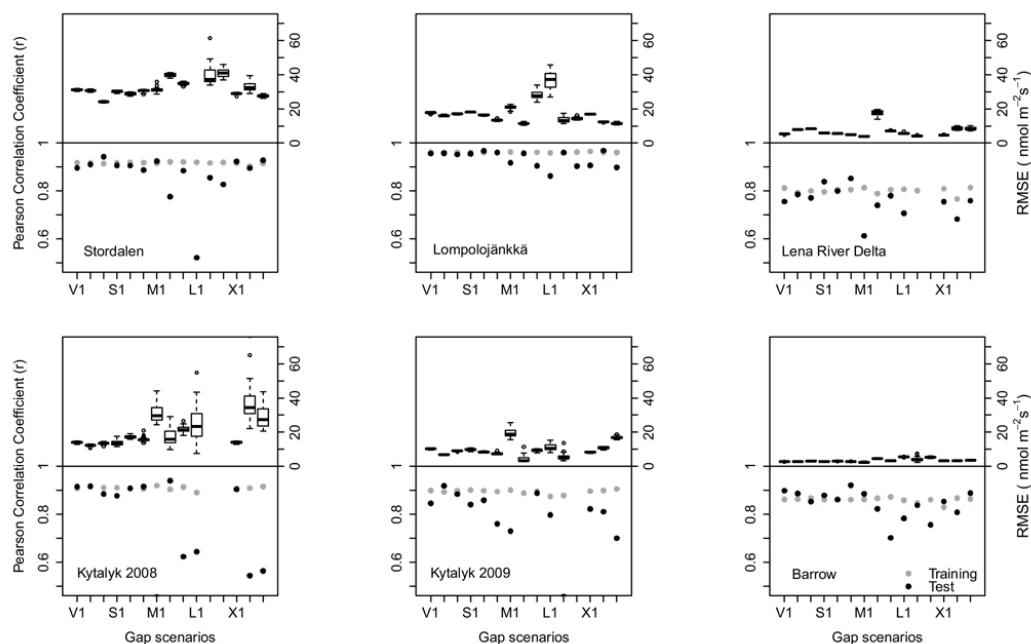
Fig. 3. Output from the "seasonal" approach. Summarised are the mean Pearson correlation coefficients (left-side scale) of 25 repetitions and the respective RMSE in true physical unit (nmol m^{-2} s^{-1}) as a measure of the network performance at each site (right-side scale). x axes labels represent very short gaps (V, 1–3) of random 30 min values; short gaps (S, 1–3) of random 4 h gaps, medium (M, 1–3) of 1.5 days, long (L, 1–3) of 12 full days and a mixed scenario (X, 1–3), representing a mix of the above mentioned gap lengths, resulting in 15 scenarios per data set. The scenario labels 1–3 do not correspond to the consecutive scenario numbering introduced in Moffat et al. (2007).

for each site is illustrated in Fig. 5. The error values for Stordalen, Lompolojänkkä and Kytalyk (2008) (Figs. 3–5) appear much higher than those of the remaining three results; they also have the highest flux ranges (see Fig. 8) in our

study. In order to visualise the model performances (overall fit of the model) we included the AIC values as an aid for model selection (Fig. 6) for each scenario. The "seasonal" and "diurnal" model included the same number of input vari-
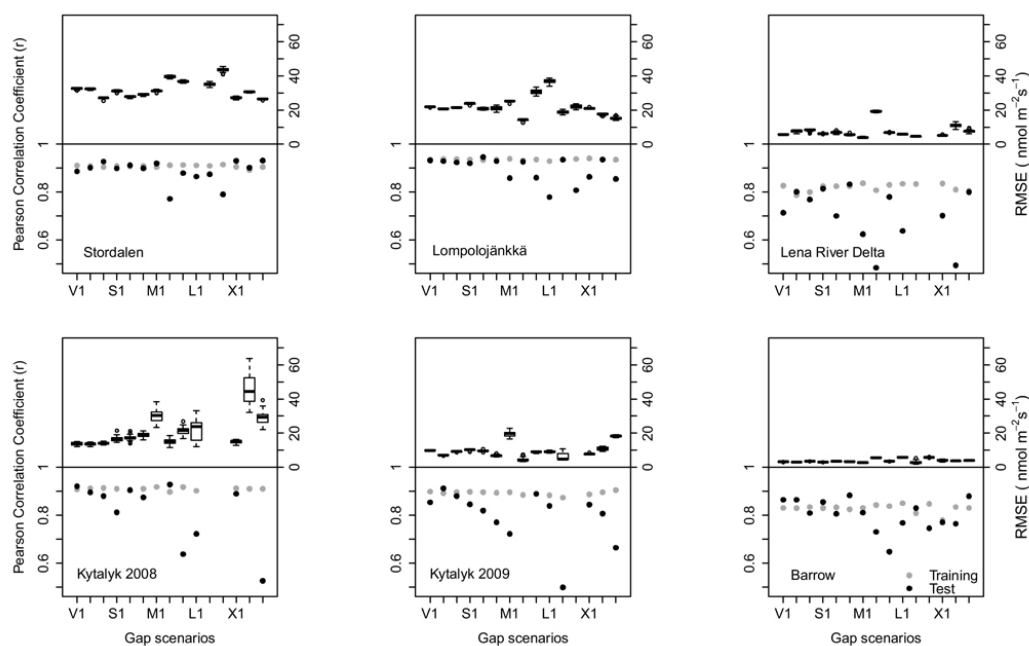
Fig. 4. Output from the "diurnal" approach. Summarised are the mean Pearson correlation coefficients (left-side scale) of 25 repetitions and the respective RMSE in true physical unit ($\text{nmol m}^{-2}\,\text{s}^{-1}$) as a measure of the network performance at each site (right-side scale). x axes labels represent very short gaps (V, 1–3) of random 30 min values; short gaps (S, 1–3) of random 4 h gaps, medium (M, 1–3) of 1.5 days, long (L, 1–3) of 12 full days and a mixed scenario (X, 1–3), representing a mix of the above mentioned gap lengths, resulting in 15 scenarios per data set. The scenario labels 1–3 do not correspond to the consecutive scenario numbering introduced in Moffat et al. (2007).

ables and same number of neurons. Their values appeared the lowest across all five sites with the "seasonal" model showing the lowest values for all scenarios in five out of six data sets. The AIC values for the "lagged" model for Stordalen and Kytalyk show the highest values. This information criterion is a measure of the overall fit but does also indicate how parsimonious a model is. The "lagged" models for these two sites do include more input variables and they are therefore "penalised" accordingly (part of the AIC principle), (Hurvich and Tsai, 1989). Therefore, the "seasonal" model was chosen to visualise the results from our additional model reliability test carried out for all six trained networks (Fig. 7). None of these results point to any degradation of the network performance or over-fitting of the trained networks.

In order to visualise the performance of the artificial neural networks applied to the "seasonal" data, we illustrated the goodness of fit of the predicted and actually measured CH_4 flux data (test data) for all three mixed scenarios showing their distribution along the ideal 1 : 1 regression line in Fig. 8. High Pearson correlation coefficients of up to 0.97 were achieved when testing the network performance. The r values for Kytalyk (2008) show the lowest correlation values of 0.61, 0.69 and 0.92. The lowest values do still pass the 95 % confidence level (54 degrees of freedom).

4 Discussion

Artificial neural networks that have previously been successfully implemented as a gap-filling method (Falge et al., 2001, Moffat et al., 2007) for carbon dioxide flux time series (Aubinet et al., 2000; Carrara et al., 2003; Papale and Valentini, 2003 and Schmidt et al., 2008) have been described as a robust, reliable and versatile tool. Nevertheless, their application is time consuming, particularly in finding the appropriate input variables, the appropriate number of hidden layers, and neurons/nodes within these layers, as well as the choice of training and test data sets (data rows). Furthermore, the global minimum (Hammerstrom, 1993; Nguyen and Chan, 2004) is not unique and changes with each training run because every training run/repetition (a run includes many iterations) achieves different fitted weights and results (it is important to find a set of weights that processes data accurately enough for the application).

In the current study, we tested the applicability of neural networks as a gap-filling tool for methane flux data and also made an attempt to standardise the method by including the same input variables for all data sets and using the same number of neurons within the hidden layer for each data scenario. In order to test their applicability, we applied the method to various ecosystems by including six distinctively different data sets from high latitudes, one showing
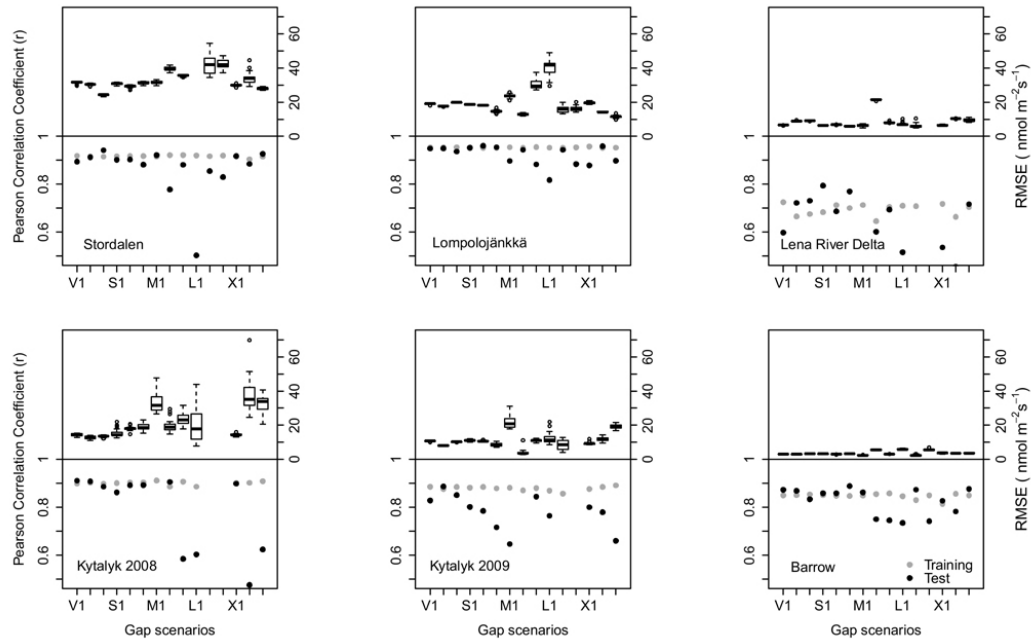
Fig. 5. Output from the "lagged" approach. Summarised are the mean Pearson correlation coefficients (left-side scale) of 25 repetitions and the respective RMSE in true physical unit (nmol m^{-2} s^{-1}) as a measure of the network performance at each site (right-side scale). x axes labels represent very short gaps (V, 1–3) of random 30 min values; short gaps (S, 1–3) of random 4 h gaps, medium (M, 1–3) of 1.5 days, long (L, 1–3) of 12 full days and a mixed scenario (X, 1–3), representing a mix of the above mentioned gap lengths, resulting in 15 scenarios per data set. The scenario labels 1–3 do not correspond to the consecutive scenario numbering introduced in Moffat et al. (2007).
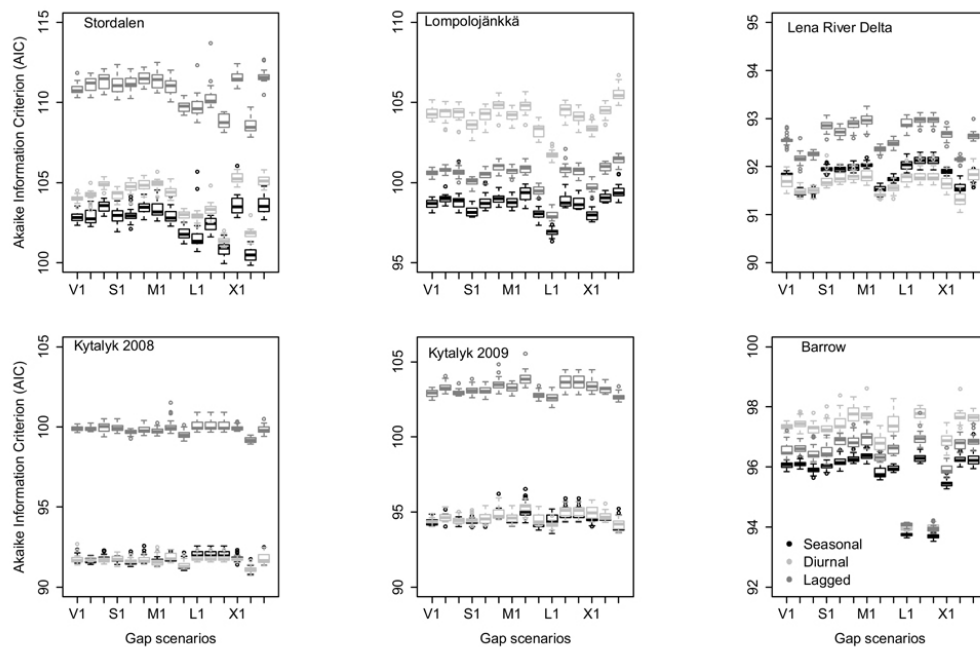

Fig. 6. Distribution of the AIC values across all gap scenarios for all sites and model approaches. These values represent the outcome from the models presented in the previous figures (Figs. 3–5). Illustrated are the results from the "seasonal" model in black, the "diurnal" model in light grey and the "lagged" model in dark grey. Following the AIC principle it becomes visible that the "seasonal" fitted model appears to be the preferred across all 15 scenarios.

diurnal and seasonal variation, one only seasonal while three data sets do not show any diurnal or seasonal variation. The sixth data set (Stordalen) reflected its position on the shores of Villasjön Lake in its emission patterns. Three different approaches have been introduced here: two including the same input variables across all sites (in an attempt to standardise the method), while a third approach included different input variables at each site and is also taking the lagged effect of water table depth and precipitation into account. Furthermore, we tested the reliability of our models by testing our trained network with data from another site with similar ranges in fluxes or site properties.

The chosen input variables (Table 2; commonly recorded meteorological parameters that act as CH_4 drivers and wind direction) as well as fuzzy sets representing time of day and seasonal changes appear to be the right choice. On the one hand adding more input variables would not comply with the simple model approach, while, on the other hand, adding more input variables would introduce the risk of cross-dependency and to some extent cross-correlation (Guan et al., 2007) between input variables, as input variables should remain as independent as possible as described in Moffat et al. (2010). Reducing the number of chosen input variables any further would not comply anymore with the principle of parsimony, of keeping a model simple but not too simple, leading to an underperformance of the network as predicted fluxes did not reach the full range which means the networks were underestimating the fluxes. The chosen meteorological variables included in the current study belong to the main drivers as shown in previous studies (see Introduction). Hydrological properties, such as precipitation and water table depth, can have a lagged effect on methane emissions (Windsor et al., 1993; Bubier et al., 1995; Kettunen et al., 1996; Suyker et al., 1996). Kettunen et al. (1996) carried out a cross-correlation study looking at the different lag effects for both variables at a boreal mire complex (six sites). They identified a significant lag of one day for both their boreal flark and hummock sites, indicating that precipitation increased emissions throughout the summer. Furthermore, Kettunen et al. (1996) also identified a 12 day lag with WTD at both these sites. Here again, a rise in water table was reflected in CH_4 emissions. Suyker et al. (1996) also identified a distinctive 12 day lag when investigating the water table fluctuation effect on the midday CH_4 fluxes from their boreal fen.

Water table position is not always recorded, or not recorded with the same time resolution as the CH_4 fluxes. Nevertheless two of the selected sites (Stordalen and Lompolojänkkä) do record water table depth continuously and were included in the current study. Results show that their network performance remained very high showing these two lagged variables to be reliable input variables in the current study. Including the lagged effect of precipitation did slightly decrease the performance of the Kytalyk (2008) network (Fig. 5). A reason for the decrease could be the data set

lengths or the fact that the lagged effect of precipitation was not present.

The performance of the mixed scenarios (representation of the most realistic gap scenario in flux data) originally chosen by Moffat et al. (2007) as a crosscheck of the other four different gap scenarios and also included in the current study show the same, in some cases better, results than some of the individual scenarios (see Figs. 3–5) themselves. The Lena River Delta appears to display the lowest mean correlation coefficient values. The extreme discrepancy in case of the Lena River Delta results are due to low CH_4 fluxes recorded at the site with few significant emission events resulting in much higher fluxes. Furthermore the correlation coefficient values for the test (artificial gaps) data sets showed sometimes higher values than those achieved in the training phase (Figs. 3–5). This could be due to the fact that Barrow (for example) experienced a diurnal trend and the data composing the test data set (artificial gaps) did not include any specific events. In case of the Lena River Delta data set, the artificially mixed gaps included two little events that the "seasonal" network was capable to predict (Fig. 8). This performance could be credited to the fact that the network has learnt about such events from similar conditions during the training phase (indication of good representativeness of the overall data as part of the training data set). The introduced lag effect for this site (Fig. 5) did slightly reduce the network performance, showing low correlation values with values still within acceptable significance level margins, nonetheless. This reduction in network performance could be credited maybe to the choice of relative humidity as an input variable in this study.

Some of the Pearson correlation coefficients achieved in the current study appear low (Figs. 3–8), compared to those achieved for CO_2 fluxes when applying the same method (Moffat et al., 2007). Much higher correlation coefficients ($r > 0.95$) were achieved in the current study when comparing trained data versus actual measured data, but resulted frequently in no acceptable values when testing the network performance. Outliers were introduced in places where there were no high or low fluxes. Such results could also be due to existing and gap scenario data distribution, as artificial gaps coincided with existing gaps, reducing the number of testing data points. None of the data sets included an entire gap free 12 day period that could have been used as a classical "long gap" test data set, as introduced in Moffat et al. (2007). Furthermore it is to be expected that predicting CH_4 emission events is more complex than predicting CO_2 fluxes that undergo a regularity and predictability in respiration or photosynthetic uptake, respectively. Integrating the additional test did help in testing the reliability of our trained networks beyond the simple testing using test data originating from the same overall data set. These results do show that, though methane does behave in a different manner than CO_2, it is still possible to apply this type of cross-validation method to such fluxes as common drivers do act similarly at
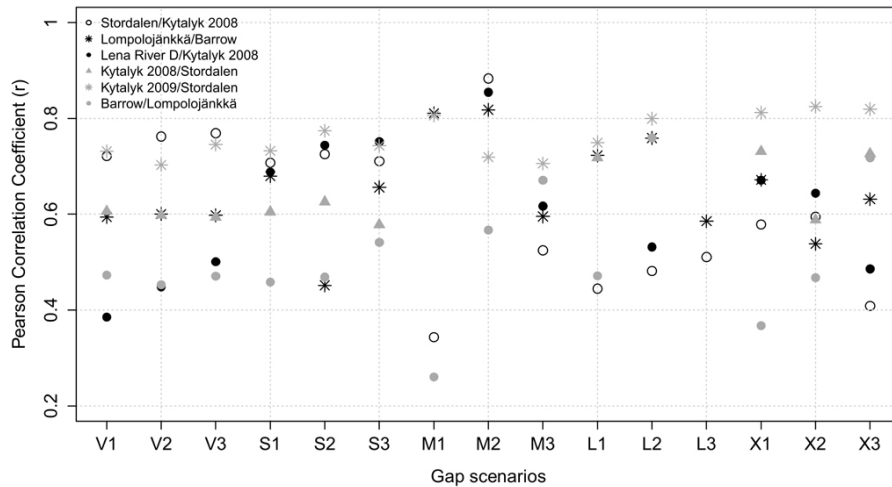
Fig. 7. Results from the additional tests carried out, testing each "seasonal" trained network with data from another site that showed a similar range in fluxes or similar site properties. Site names on the left side in the legend indicate the name of the trained network, while the site name on the right side indicate the site name whose data has been used to test the reliability of the model.
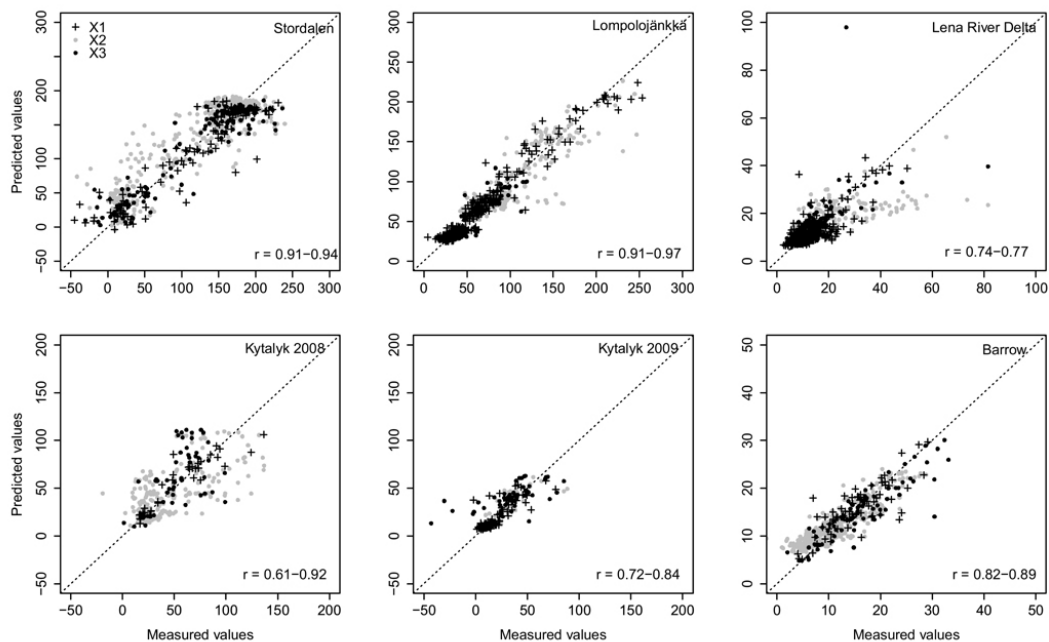

Fig. 8. Scatter plots showing the distribution of the actually measured CH_4 flux values against the mean predicted CH_4 values (25 runs) for each of the "seasonal" mixed scenarios (X, 1–3), representing a mix of gap lengths, for all six data sets, and their distribution along the ideal $1:1$ regression line. All units are in nmol m^{-2} s^{-1}. Included is also the range of the achieved mean r coefficients.

all sites. These results are also indicative of the robustness of the models introduced in the current study.

Two other important aspects in the current study are the length of the time series (Table 1), and thus consequently the sufficiency of available training data, and secondly the gap length in the existing time series. Time series vary between 15 days (Kytalyk, 2008) and 181 days (Lompolojänkkä),

with Lompolojänkkä having the highest proportion of gaps in the CH_4 fluxes. Regarding sufficiency of available training data, we believe that by utilising the method described here, we kept the method transparent, avoided working with models lacking in power, and making this method applicable to short and longer time series. Moffat et al. (2010) showed that neural networks are actually also applicable to one day

of incomplete data, when applied to CO_2 flux data. However, none of the time series included in the current study were gap free or included an entire year's worth of data. Therefore, no bias error to indicate the bias induced on the annual sums/budgets, as referred to by Moffat et al. (2007), was calculated.

Nevertheless, neural networks show excellent performance (Figs. 3–8), which proves that this standardised method is easy to be implemented and applicable to many different ecosystems in the northern latitudes (Fig. 1). We have reliably reproduced and predicted methane fluxes. Furthermore, we were able to incorporate successfully the lagged effect of hydrological properties such as water table depth at two sites and precipitation at one site (Kytalyk). We find artificial neural networks to be recommendable as a reliable and robust gap-filling method for high resolution CH_4 flux data originating from various ecosystems at high latitudes, as estimated annual budgets rely on accurate gap-free or gap-filled data.

5 Challenges and recommendations

A peculiar characteristic of CH_4 is its higher emission variability than CO_2 fluxes, often connected to specific events such as those visible in the Lena River Delta data set. In our case, the network was able to reproduce these events (Figs. 6–8). In case such events are triggered by other drivers or physical forcing (not included as input variables in neural networks), predicted values do diverge from the actually measured values. Precipitation and water table depth can affect CH_4 fluxes (Windsor et al., 1993; Bubier et al., 1995; Kettunen et al., 1996; Suyker et al., 1996) with each rain event being different in intensity and length. The same can be said of water table depths whose rise and fall are not equally predictable after each such event also behaving differently at each site. This lagged effect has, so far, only been confirmed in in situ soil chamber studies (Windsor et al., 1993; Kettunen et al., 1996), laboratory experiments (Funk et al., 1994; Kettunen et al., 1999) and midday CH_4 flux values, measured with the eddy covariance method (Suyker et al., 1996) but not in high resolution CH_4 eddy covariance flux studies. In addition, further uncertainties in predicting/estimating accurate CH_4 emissions do exist regarding the insufficiently understood "pressure pumping effect" (Zamolodchikov et al., 2003; Takle et al., 2004) and friction velocity (u^* – correction), a parameter known to act as a driver for CH_4 emissions (Wille et al., 2010) but also used as a filtering criterion for low turbulence, both affecting methane emissions. These factors might have an influence on how far neural networks are reproducing and predicting CH_4 fluxes accurately.

In order to evaluate, carry out and apply standardised CH_4 flux measurements, data post-processing and gap-filling methods in the future, standardised protocols and required auxiliary measurements are needed to be implemented. First steps towards such procedures have already been initiated during the international ESF (European Science Foundation) explanatory (Germany, April 2012) and FLUXNET CH_4 (Finland, September 2012) workshops.

References

Amari, S., Murata, N, Müller, K.-R., Finke, M., and Yang, M. H.: Asymptotic statistical theory of overtraining and cross-validation, IEEE T. Neural Networ., 8, 985–996, 1997.

Anisimov, O. A.: Potential feedback of thawing permafrost to the global climate system through methane emission, Environ. Res. Lett., 2, 045016, doi:10.1088/1748-9326/2/4/045016, 2007.

Aubinet, M., Grelle, A., Ibrom, A., Rannik, U., Moncrieff, J., Foken, T., Kowalski, A. S., Martin, P. H., Berbigier, P., Bernhofer, C., Clement, R., Elbers, J., Granier, A., Grunwald, T., Morgenstern, K., Pilegaard, K., Rebmann, C., Snijders, W., Valentini, R., and Vesala, T.: Estimates of the annual net carbon and water exchange of forests: The EUROFLUX methodology, Adv. Ecol. Res., 30, 113–175, 2000.

Aurela, M., Lohila, A., Tuovinen, J. P., Hatakka, J., Riutta, T., and Laurila, T.: Carbon dioxide exchange on a northern boreal fen., Boreal Environ. Res., 14, 699–710, 2009.

Baldocchi, D. D.: Assessing the eddy covariance technique for evaluating carbon dioxide exchange rates of ecosystems, past, present and future, Glob. Change Biol., 9, 479–492, 2003.

Beck, L. W.: The Principle of Parsimony in Empirical Science, J. Phil., 40, 617–633, 1943.

Bubier, J. L., Moore, T. R., and Juggins, S.: Predicting Methane Emission from Bryophyte Distribution in Northern Canadian Peatlands, Ecology, 76, 677–693, 1995.

Bugmann, H. and Martin, P.: How Physics and Biology matter in forest gap models, Clim. Change, 29, 251–257, 1995.

Burnham, K. P. and Anderson, D. R.: Multimodel Inference, Understanding AIC and BIC in Model Selection, Sociological methods and research, 33, 261–304, 2004.

Carrara, A., Kowalski, A. S., Neirynck, J., Janssens, I. A., Yuste, J. C., and Ceulemans, R.: Net ecosystem CO_2 exchange of mixed forest in Belgium over 5 years, Agr. Forest Meteorol., 119, 209–227, 2003.

Chattopadhyay, S. and Chattopadhyay, G.: Identification of the best hidden layer size for three-layered neural net in predicting monsoon rainfall in India, J. Hydroinform., 10, 181–188, doi:10.2166/hydro.2008.017, 2008.

Christensen, T. R.: Methane emission from Arctic tundra, Biogeochemistry, 21, 117–139, 1993.

Cybenko, G.: Approximation by superpositions of a sigmoidal function, Math. Control Signal., 2, 303–314, 1989.

Dengel, S., Levy, P. E., Grace, J., Jones, S. K., and Skiba, U.: Methane emissions from sheep pasture, measured with an open-path eddy covariance system, Glob. Change Biol., 17, 3524–3533, 2011.

Elizondo, D. and Góngora, M.: Current Trends on Knowledge Extraction and Neural Networks, Artificial Neural Networks: Formal Models and Their Applications – ICANN 2005, edited by: Duch, W., Kacprzyk, J., Oja, E., and Zadrozny, S., Springer Berlin, Heidelberg, 752–752, 2005.

EPA: Methane and Nitrous Oxide Emissions from Natural Sources, US, Environmental Protection Agency, Washington, DC, USA, 2010.

Eugster, W. and Plüss, P.: A fault-tolerant eddy covariance system for measuring CH_4 fluxes, Agr. Forest Meteorol., 150, 841–851, 2010.

Falge, E., Baldocchi, D. D., Olson, R., Anthoni, P., Aubinet, M., Bernhofer, C., Burba, G., Ceulemans, R., Clement, R., Dolman, H., Granier, A., Gross, P., Grünwald, T., Hollinger, D., Jensen, N.-O., Katul, G., Keronen, P., Kowalski, A., Lai, C.-T., Law, B. E., Meyers, T., Moncrieff, J., Moors, E., Munger, J. W., Pilegaard, K., Rannik, Ü., Rebmann, C., Suyker, A., Tenhunen, J., Tu, K., Verma, S., Vesala, T., Wilson, K., and Wofsy, S.: Gap filling strategies for defensible annual sums of net ecosystem exchange, Agr. Forest Meteorol., 107, 43–69, 2001.

Forbrich, I., Gazovic, M., Kutzbach, L., Wille, C., Wolf, U., Becker, T., Schreiber, P., and Wilmking, M.: Cross-evaluation of measurements of peatland methane emissions on microform and ecosystem scales using high-resolution landcover classification and source weight modelling, Agr. Forest Meteorol., 151, 864–874, 2011.

Franses, P. H. and Draisma, G.: Recognizing changing seasonal patterns using artificial neural networks, J. Econometrics, 81, 273–280, 1997.

Fritsch, S. and Günther, F.: neuralnet, Training of Neural Networks, R package version 1.32, http://CRAN.R-project.org/package=neuralnet (05 April 2013), 2012.

Funk, D. W., Pullman, E. R., Peterson, K. M., Crill, P. M., and Billings, W. D.: Influence of water table on carbon dioxide, carbon monoxide, and methane fluxes from Taiga Bog microcosms, Global Biogeochem. Cy., 8, 271–278, 1994.

Gardner, M. W. and Dorling, S. R.: Artificial neural networks (the multilayer perceptron) – a review of applications in the atmospheric sciences, Atmos. Environ., 32, 2627–2636, 1998.

Gardner, M. W. and Dorling, S. R.: Neural network modelling and prediction of hourly NO_x and NO_2 concentrations in urban air in London, Atmos. Environ., 33, 709–719, 1999.

Gazovic, M., Kutzbach, L., Schreiber, P., Wille, C., and Wilmking, M.: Diurnal dynamics of CH_4 from a boreal peatland during snowmelt, Tellus B, 62, 133–139, 2010.

Gorban, A., Rossiev, A., Makarenko, N., Kuandykov, N., and Dergachev, V.: Recovering data gaps through neural network methods, International Journal of Geomagnetism and Aeronomy, 3, 191–197, 2002.

Günther, F. and Fritsch, S.: neuralnet, Training of neural networks, R Journal, 2, 30–38, 2010.

Guan, L., Yang, J., and Bell, J. M.: Cross-correlations between weather variables in Australia, Build. Environ., 42, 1054–1070, 2007.

Gunaratnam, D. J., Degroff, T., and Gero, J. S.: Improving neural network models of physical systems through dimensional analysis, Appl. Soft Comput., 2, 283–296, 2003.

Hammerstrom, D.: Working with neural networks, Spectrum, IEEE, 30, 46–53, 1993.

Hansen, L. K. and Salamon, P.: Neural network ensembles, Pattern Analysis and Machine Intelligence, IEEE Transactions on, 12, 993–1001, 1990.

Hargreaves, K. J., Fowler, D., Pitcairn, C. E. R., and Aurela, M.: Annual methane emission from Finnish mires estimated from eddy covariance campaign measurements, Theor. Appl. Climatol., 70, 203–213, 2001.

Hecht-Nielsen, R.: Kolmogorov's mapping neural network existence theorem, P. 1st Int. Conf. Neural Network., 3, 11–14, 1987.

Hendriks, D. M. D., Dolman, A. J., van der Molen, M. K., and van Huissteden, J.: A compact and stable eddy covariance set-up for methane measurements using off-axis integrated cavity output spectroscopy, Atmos. Chem. Phys., 8, 431–443, doi:10.5194/acp-8-431-2008, 2008.

Hurvich, C. M. and Tsai, C. L.: Regression and Time Series Model Selection in Small Samples, Biometrika, 76, 297–307, 1989.

IPCC: Summary for Policymakers, in: Climate Change 2007, The Physical Science Basis, Contribution of Working Group I to the Fourth Assessment Report of the Intergovernmental Panel on Climate Change, edited by: Solomon, S., Qin, D., Manning, M., Chen, Z., Marquis, M., Averyt, K. B., Tignor, M. and Miller, H. L., Cambridge University Press, Cambridge, United Kingdom and New York, NY, USA, 2007.

Jackowicz-Korczyński, M., Christensen, T. R., Bäckstrand, K., Crill, P., Friborg, T., Mastepanov, M., and Ström, L.: Annual cycle of methane emission from a subarctic peatland, J. Geophys. Res., 115, G02009, doi:10.1029/2008JG000913, 2010.

Jain, A. K., Mao, J., and Mohiuddin, K. M.: Artificial Neural Networks, A Tutorial. Computer, 29, 31–44, 1996.

Järvi, L., Nordbo, A., Junninen, H., Riikonen, A., Moilanen, J., Nikinmaa, E., and Vesala, T.: Seasonal and annual variation of carbon dioxide surface fluxes in Helsinki, Finland, in 2006–2010, Atmos. Chem. Phys., 12, 8475–8489, doi:10.5194/acp-12-8475-2012, 2012.

Johansson, T., Malmer, N., Crill, P. M., Friborg, T., Åkerman, J. H., Mastepanov, M., and Christensen, T. R.: Decadal vegetation changes in a northern peatland, greenhouse gas fluxes and net radiative forcing. Glob. Change Biol., 12, 2352–2369, 2006.

Kettunen, A., Kaitala, V., Alm, J., Silvola, J., Nykänen, H., and Martikainen, P. J.: Cross-correlation analysis of the dynamics of methane emissions from a boreal peatland, Global Biogeochem. Cy., 10, 457–471, 1996.

Kettunen, A., Kaitala, V., Lehtinen, A., Lohila, A., Alm, A., Silvola, J., and Martikainen, P. J.: Methane production and oxidation potentials in relation to water table fluctuations in two boreal mires, Soil Biol. Biochem., 31, 1741–1749, 1999.

Kişi, Ő. and Oncuoğlu, E. : Comparison of three back-propagation training algorithms for two case studies, Indian J. Eng. Mater. S., 12, 434–442, 2005.

Klevecka, I., and Lelis, J.: Pre-Processing of Input Data of Neural Networks: The Case of Forecasting Telecommunication Network Traffic, Telektronikk, 3, 168–178, 2008.

Lee, T. L. and Jeng, D. S.: Application of artificial neural networks in tide-forecasting, Ocean Eng., 29, 1003–1022, 2002.

Lek, S. and Guégan, J. F.: Artificial neural networks as a tool in ecological modelling, an introduction, Ecol. Model., 120, 65–73, 1999.

Long, K. D. and Flanagan, L. B.: Diurnal and seasonal variation in methane emissions in a northern Canadian peatland measured by eddy covariance, Glob. Change Biolo., 16, 2420–2435, 2010.

Mastepanov, M., Sigsgaard, C., Dlugokencky, E. J., Houweling, S., Ström, L., Tamstorf, M. P., and Christensen, T. R.: Large tundra methane burst during onset of freezing, Nature, 456, 628–630, 2008.

McDermitt, D., Burba, G., Xu, L., Anderson, T., Komissarov, A., Riensche, B., Schedlbauer, J., Starr, G., Zona, D., Oechel, W., Oberbauer, S., and Hastings, S.: A new low-power, open-path instrument for measuring methane flux by eddy covariance, Appl. Phys. B-Lasers O., 102, 1–15, 2011.

Moffat, A. M.: A new methodology to interpret high resolution measurements of net carbon fluxes between the terrestrial ecosystems and the atmosphere, Doctoral thesis, Friedrich Schiller University, Jena, 2010.

Moffat, A. M., Papale, D., Reichstein, M., Hollinger, D. Y, Richardson, A. D., Barre, A. G., Beckstein, C., Braswell, B. H., Churkina, G., Desai, A. R., Falge, E., Gove, J. H., Heimann, M., Hui, D., Jarvis, A. J., Kattge, J., Noormets, A., and Stauch, V. J.: Comprehensive comparison of gap-filling techniques for eddy covariance net carbon fluxes, Agr. Forest Meteorol., 147, 209–232, 2007.

Moffat, A. M., Beckstein, C., Churkina, G., Mund, M., and Heimann, M.: Characterization of ecosystem responses to climatic controls using artificial neural networks. Glob. Change Biol., 16, 2737–2749, 2010.

Nelson, M., Hill, T., Remus, W., and O'Connor, M.: Time series forecasting using neural networks: should the data be deseasonalized first?, J. Forecasting, 18, 359–367, 1999.

Nguyen, H. and. Chan, C.: Multiple neural networks for a long term time series forecast, Neural Comput. Appl., 13, 90–98, 2004.

Ooba, M. and Hirano, T.: Comparisons of gap-filling methods for carbon flux dataset, A combination of a genetic algorithm and an artificial neural network, Ecol. Model., 198, 473–486, 2006.

Papale, D.: Data Gap Filling, edited by: Aubinet, M., Vesala, T., and Papale, D., Eddy Covariance, A Practical Guide to Measurement and Data Analysis, Springer Atmospheric Sciences, 159–172, 2012.

Papale, D. and Valentini, R.: A new assessment of European forests carbon exchanges by eddy fluxes and artificial neural network spatialization, Glob. Change Biol., 9, 525–535, 2003.

Parmentier, F. J. W., van Huissteden, J., van der Molen, M. K., Schaepman-Strub, G., Karsanaev, S. A., Maximov, T. C., and Dolman, A. J.: Spatial and temporal dynamics in eddy covariance observations of methane fluxes at a tundra site in northeastern Siberia, J. Geophys. Res.-Biogeo., 116, G03016, doi:10.1029/2010JG001637, 2011a.

Parmentier, F. J. W., van der Molen, M. K., van Huissteden, J., Karsanaev, S. A., Kononov, A. V., Suzdalov, D. A., Maximov, T. C., and Dolman, A. J.: Longer growing seasons do not increase net carbon uptake in the northeastern Siberian tundra, J. Geophys. Res.-Biogeo., 116, G04013, doi:10.1029/2011JG001653, 2011b.

Peltola, O., Mammarella, I., Haapanala, S., Burba, G., and Vesala, T.: Field intercomparison of four methane gas analyzers suitable for eddy covariance flux measurements, Biogeosciences, 10, 3749–3765, doi:10.5194/bg-10-3749-2013, 2013.

R Core Team: R: A language and environment for statistical computing, R Foundation for Statistical Computing, Vienna, Austria, http://www.R-project.org/ (last access: 05.09.2013), 2013.

Riedmiller, M.: Rprop-description and implementation details, Technical Report, University of Karlsruhe, Germany, 1994.

Rinne, J., Riutta, T., Pihlatie, M., Aurela, M, Haapanala, S., Tuovinen, J.-P., Tuittila, E.-S., and Vesala, T.: Annual cycle of methane emission from a boreal fen measured by the eddy covariance technique, Tellus B, 59, 449–457, 2007.

Riutta, T., Laine, J., Tuittila, E.-S.: Spatial variation in plant community functions regulates carbon gas dynamics in a boreal fen ecosystem. Tellus B, 59, 838–852, 2007.

Roulet, N., Moore, T. I. M., Bubier, J., and Lafleur, P.: Northern fens, methane flux and climatic change, Tellus B, 44, 100–105, 1992.

Ryan, M., Müller, C., Di, H. J., and Cameron, K. C.: The use of artificial neural networks (ANNs) to simulate N_2O emissions from a temperate grassland ecosystem, Ecol. Model., 175, 189–194, 2004.

Sachs, T., Giebels, M., Boike, J., and Kutzbach, L.: Environmental controls on CH4 emission from polygonal tundra on the microsite scale in the Lena river Delta, Siberia, J. Geophys. Res., 113, G00A03, doi:10.1029/2007JG000505, 2008.

Saxén, H. and Pettersson, F.: Method for the selection of inputs and structure of feedforward neural networks, Compu. Chem. Eng., 30, 1038–1045, 2006.

Schiffmann, W., Joost, M., and Werner, R.: Optimization of the backpropagation algorithm for training multilayer perceptrons, Technical report, University of Koblenz, Insitute of Physics, 1994.

Schmidt, A., Wrzesinsky, T., and Klemm, O.: Gap Filling and Quality Assessment of CO_2 and Water Vapour Fluxes above an Urban Area with Radial Basis Function Neural Networks, Bound-Lay Meteorol., 126, 389–413, 2008.

Sharda, R. and Patil, R. B.: Connectionist approach to time series prediction: an empirical test, J. Intell. Manuf., 3, 317–323, 1992.

Stathakis, D.: How many hidden layers and nodes?, Int. J. Remote Sens., 30, 2133–2147, 2009.

Sturtevant, C. S., Oechel, W. C., Zona, D., Kim, Y., and Emerson, C. E.: Soil moisture control over autumn season methane flux, Arctic Coastal Plain of Alaska, Biogeosciences, 9, 1423–1440, doi:10.5194/bg-9-1423-2012, 2012.

Suyker, A. E., Verma, S. B., Clement, R. J., and Billesbach D. P.: Methane flux in a boreal fen: Season-long measurement by eddy correlation, J. Geophys. Res., 101, 28637–28647, doi:10.1029/96JD02751, 1996.

Svozil, D., Kvasnicka, V., and Pospichal, J.: Introduction to multilayer feed-forward neural networks, Chemometr. Intell. Lab., 39, 43–62, 1997.

Tagesson, T., Mölder, M., Mastepanov, M., Sigsgaard, C., and Tamstorf, M. P.: Land-atmosphere exchange of methane from soil

thawing to soil freezing in a high-Arctic wet tundra ecosystem. Glob. Change Biol., 18, 1928–1940, 2012.

Takle, E. S, Massman, W. J., Brandle, J. R., Schmidt, R. A., Zhou, X., Litvina, I. V., Garcia, R., Doyle, G., and Rice, C. W.: Influence of high-frequency ambient pressure pumping on carbon dioxide efflux from soil. Agr. Forest Meteorol., 24, 193–206, 2004.

Toptygin, A. Y., Gribanov, K. G., Imasua, R., Bleuten, W., and Zakharovet, W. I.: Seasonal methane content in atmosphere of the permafrost boundary zone in Western Siberia determined from IMG/ADEOS and AIRS/AQUA data, Honolulu, HI, USA, SPIE, 2005.

Treadgold, N. K. and Gedeon, T. D.: The Sarprop Algorithm, A Simulated Annealing Enhancement To Resilient Back Propagation, Proceedings International Panel Conference on Soft and Intelligent Computing, 1996.

van Huissteden, J., Maximov, T. C., Kononov, A. V., and Dolman, A. J.: Summer soil CH_4 emission and uptake in taiga forest near Yakutsk, Eastern Siberia, Agr. Forest Meteorol., 148, 2006–2012, 2008.

van Wijk, M. T. and Bouten, W.: Water and carbon fluxes above European coniferous forests modelled with artificial neural networks, Ecol. Model., 120, 181–197, 1999.

Wang, W., van Gelder, P. H. A. J. M., and Vrijling, J. K.: Some Issues About the Generalization of Neural Networks for Time Series Prediction Artificial Neural Networks: Formal Models and Their Applications – ICANN 2005, edited by: Duch, W., Kacprzyk, J., Oja, E., and Zadrozny, S., Springer Berlin, Heidelberg, 749–749, 2005.

Whalen, S. C. and Reeburgh, W. S.: Interannual variations in tundra methane emission: A 4 year time series at fixed sites, Global Biogeochem. Cy., 6, 139–159, 1992.

Wille, C., Kutzbach, L., Sachs, T., Wagner, D., and Pfeiffer, E-M.: Methane emission from Siberian arctic polygonal tundra, eddy covariance measurements and modelling, Glob. Change Biol., 14, 1395–1408, 2008.

Windsor, J., Moore, T. R., and Roulet, N. T.: Episodic fluxes of methane from subarctic fens, Can. J. Soil Sci., 72, 441–452, 1993.

Zamolodchikov, D. G., Karelin, D. V., Ivaschenko, A. I., Oechel, W. C., and Hastings, S. J.: CO_2 flux measurements in Russian Far East tundra using eddy covariance and closed chamber techniques, Tellus B, 55, 879–892, 2003.

Zhang, G. P. and Qi, M.: Neural network forecasting for seasonal and trend time series, Eur. J. Oper. Res., 160, 501–514, 2005.

Zona, D., Oechel, W. C., Kochendorfer, J., Paw, U. K. T., and Salyuket, A. U.: Methane fluxes during the initiation of a large-scale water table manipulation experiment in the Alaskan Arctic tundra, Global Biogeochem. Cy., 23, GB2013, doi:10.1029/2009GB003487, 2009.

Zona, D., Lipson, D. A., Paw, U K. T., Oberbauer, S. F., Olivas, P., Gioli, B., and Oechel, W. C.: Increased CO_2 loss from vegetated drained lake tundra ecosystems due to flooding, Global Biogeochem. Cy., 26, GB2004, doi:10.1029/2011GB004037, 2012.