Article (refereed) - postprint

Van Oijen, Marcel; Cameron, David; Levy, Peter E.; Preston, Rory. 2017. **Correcting errors from spatial upscaling of nonlinear greenhouse gas flux models**.

Correcting errors from spatial upscaling of nonlinear greenhouse gas flux models

Marcel Van Oijen [1,*], *David Cameron* [1], *Peter Levy* [1], *Rory Preston* [1,2]
[1] *Centre for Ecology & Hydrology, Bush Estate, Penicuik EH26 0QB, UK*
[2] *University of Edinburgh, Edinburgh, UK*
[*] *mvano@ceh.ac.uk*

2017-04-11

Abstract

Ecological models are used to quantify processes over large regions. When the model is nonlinear and input variables are spatially averaged, the regional mean will be in error.

A formula for estimating the upscaling error can be derived from Taylor expansion of the model (Bresler and Dagan 1988). We test this for simple models under three different input distributions (Gaussian, exponential, lognormal). In several cases the formula is exact, in others it provides a reasonable approximation.

We then study models for emissions of methane, ammonia, and nitrous oxide across the UK. We scale from 1 x 1 km to 32 x 32 km. The UK-average upscaling errors are -12%, -48% and -3%, well estimated using the formula.

The formula is a useful tool for modellers desiring to correct upscaling error for their application. Calculation of second-order partial derivatives of model output is required, for which we provide R-code.

Keywords

Block support, Ecosystem modelling, Greenhouse gas emission, Point support, Spatial upscaling, Taylor expansion

1 Introduction

1.1 Ecosystem flux modelling and spatial upscaling errors

To estimate the biogenic greenhouse gas (GHG) balance of countries or larger areas, flux models are needed. Models have been developed that predict the magnitude of GHG-fluxes as a function of soil, weather and vegetation characteristics. For example, Cameron et al. (2013) used four process-based ecosystem models to quantify variation across Europe in fluxes of CO_2, N_2O, NO and H_2O between forest and atmosphere, for present and future climatic conditions. As a simpler example, Levy et al. (2012) estimated methane-emission from UK soils as a function of temperature and the soil's carbon

and water content. Most GHG flux models are spatially 1-dimensional: they simulate vertical fluxes but not horizontal ones such as flows of water or nutrients between neighbouring locations.

When ecosystem flux models are applied to countries or larger areas, the area is subdivided using a spatial grid, with each grid cell covering an area of multiple km^2. In contrast, the models themselves tend to be based on observations made at much finer scales such as individual crop fields, forest stands or experimental sites. Therefore, ecosystem flux models are said to have 'small-scale support', or even 'point support' (Heuvelink and Pebesma, 1999). This reflects the scale-dependence of ecological processes and variables which precludes models from being valid at all scales (Schweiger and Beierkuhnlein, 2016).

When using a small-scale support model to estimate GHG fluxes in a grid cell of many km^2, we should account for within-cell spatial heterogeneity. The mean output of a model f in a grid cell with spatially varying inputs x can be calculated as $\int_{-\infty}^{+\infty} f(x)\, p(x)\, dx$, where p(x) is the probability distribution of input values in the region. We shall refer to this integral more briefly as E[f(x)], using the expectation operator E[].

In practice, modellers rarely attempt to calculate this integral. Model output for a region, like a grid-cell, tends to be calculated by running the model just once, for the average value of the inputs in that region. For example, Van Oijen & Thomson (2010) calibrated their process-based model with data from forest stands in Wales and England, and went on to apply the model for the whole of the UK using a 20 x 20 km grid for the input variables. In such cases, what is being calculated for each grid cell is f(E[x]) rather than E[f(x)]. This causes error when the models are nonlinear in their inputs. We define the upscaling error Δ as:

$$\Delta = f(E[x]) - E[f(x)] \tag{1}$$

The erroneous assumption that the two expectations are the same (which would make Δ zero) has been called the *fallacy of averages* which according to Welsh et al. (1988) "is perhaps the most widespread statistical error in biology". However, upscaling error has in fact long been recognized as a problem by ecological and environmental modellers (e.g. Avissar and Pielke, 1989; Rastetter et al., 1992). In recent years, several studies have quantified the magnitude of the error caused by spatially aggregating model input data on soils or climate. Corstanje et al. (2008) showed for a model of NH_3-emission from soils that aggregating inputs within a range of 50 m led to 9% underestimation of emissions, and extending the range to 2 km increased the error to 20%. Corstanje and Lark (2008) showed that it was mainly aggregation of soil pH, which is highly spatially heterogeneous, that caused the scaling errors for their model. And more recently, several authors have quantified the impact of aggregating soil and/or atmospheric variables on the predictions of different crop models (e.g. Eyshi Rezaei et al., 2015; Hoffmann et al., 2016; Kuhnert et al., 2016; Zhao et al., 2015).

1.2 Approaches to reducing or correcting upscaling error

Several methods for reducing Δ have been proposed (Van Oijen et al., 2009). These include:

1. running the model for specific points (rather than for spatially averaged inputs), followed by geostatistical interpolation,
2. tiling and other methods for disaggregating inputs,

3. model reparameterisation, i.e. assigning 'effective parameter values' that allow the model to work with spatially averaged inputs without altering model structure.

Of these methods, the first seems not to have been used yet in ecological modelling, while input disaggregation has been attempted but with little success due to limited information on spatial heterogeneity (Corstanje et al., 2008). Reparameterisation has been used in hydrology and climate modelling with some success (e.g. Taylor, 1987), but is likely to fail with highly nonlinear models (Zhang, 2012). Overall, these methods for reducing Δ have rarely been used in model studies for upscaling GHG fluxes.

Instead of trying to reduce Δ, another approach is to estimate its magnitude and correct for it. For functions of one variable, we can do so based on the following approximation which can be found in statistical handbooks (e.g. Hines and Montgomery, 1990, Eq. 4-17):

$$E[f(x)] \approx f(E[x]) + \frac{1}{2} Var[x] f^{(2)}(E[x]), \tag{2}$$

where $f^{(2)}(E[x])$ is the second derivative of f evaluated at the mean of x, and Var[x] is the variance of x within the region. This formula is derived by applying the expectation operator to the second-order Taylor expansion of f(x). Combining this formula with Eq. (1) gives us an approximate formula for the upscaling error of models with one input variable:

$$\widehat{\Delta} = -\frac{1}{2} Var[x] f^{(2)}(E[x]), \tag{3}$$

where we use the ^-symbol to indicate that the formula provides an estimator for Δ, not Δ itself.

The $\widehat{\Delta}$-formula formalizes some common-sense notions about upscaling: it shows that the associated error is proportional to input variance, that it is zero when the model is linear ($f^{(2)}(x) = 0$) and negative when the model is convex ($f^{(2)}(x) > 0$).

The formula can be generalized to functions of multiple input variables, which are common in flux modelling. In such cases, the input x is not a scalar but a vector, and we need to use multivariate Taylor expansion (Bernardo and Smith, 1994). The result is again a simple approximative formula for Δ:

$$\widehat{\Delta} = -\frac{1}{2} tr(S\,H), \tag{4}$$

where S is the variance-covariance matrix of x, H is the Hessian matrix of second order partial derivatives of f(x), and tr denotes the trace (sum of diagonal elements) of the matrix product $S\,H$. We shall refer to Eq. (4) as "the $\widehat{\Delta}$-formula", with Eq. (3) representing a special case of it.

This multivariate Taylor-expansion based approach to correcting for upscaling error - albeit with inputs assumed independent - was introduced to crop modelling by Bresler & Dagan (1988, Eq. 23) who used it to study the impact of soil variability on crop yield. Band et al. (1991, Eq. 4) used the approach to speed up computation and to correct for errors in forest modelling at the scale of a watershed. Rastetter et al. (1992, Eq. 15) gave some examples of its use for simple models of light interception and light-use efficiency.

Despite these early applications, the method has not been used frequently in ecosystem modelling. The aim of this paper is to reconsider its applicability, in particular to large-scale GHG flux estimation. To do so we need to consider possible limitations to applicability of the $\widehat{\Delta}$-formula. We may expect that the formula is difficult to apply when:

- the model is high-dimensional, hampering the quantification of the S- and H-matrices,
- the model is not twice differentiable or its second-order partial derivatives are hard to calculate (e.g. for process-based models),
- the model is poorly approximated by a second-order Taylor expansion,
- information to apply the formula is lacking, e.g. knowledge about (co)variances of inputs.

These possible limitations to application of the $\widehat{\Delta}$-formula will be addressed in this paper.

1.3 Contents of the paper

We begin by listing a number of model types and input distributions for which the upscaling error Δ can be calculated exactly. Obviously, in such cases, there is no need to calculate its approximation with the $\widehat{\Delta}$-formula. However, we do so here to evaluate the formula. We show that the formula is exact (i.e. $\widehat{\Delta} = \Delta$) for some important cases, allowing full correction of model upscaling errors. In other cases, the formula provides only a partial correction. The analytical results that we tabulate may be of use to modellers who desire a quick estimate of the likely degree of upscaling error to be expected from their model.

After deriving these analytical results, we study scaling error for three non-linear models of GHG emission. These are: (1) the methane model of Levy et al. (2012), (2) the ammonia model of Denmead et al. (1982) and (3) the nitrous oxide model of Flechard et al. (2007). Each of the three models is written in the form of a single equation. We show that application of these models on a 32 x 32 km grid across the United Kingdom leads to significant upscaling error, but one that can at least partly be corrected for using the $\widehat{\Delta}$-formula. We conclude the study by briefly discussing how the $\widehat{\Delta}$-formula could be applied to more complex, high-dimensional functions such as process-based models.

2 Materials and Methods

2.1 Methane model

Levy et al's (2012) methane model is given by the following equation:

$$F_{CH_4} = 2.07 + 0.036\, C_s\, e^{0.094\, T_s}\, \theta^{4.77}, \tag{5}$$

where F_{CH_4} is the CH_4-emission from the soil (nmol m^{-2} s^{-1}), C_s is soil organic carbon (kg C m^{-2}), T_s is soil temperature ($^\circ C$) and θ is soil moisture as a fraction of volume (-).

2.2 Ammonia model

Denmead et al. (1982, see Potter et al. (2003)) proposed a model for the effects of soil conditions on the ratio of ammonia emission to available mineral nitrogen. This emission factor EF_{NH_3} (g N g^{-1} N) is written as follows:

$$EF_{NH_3} = \frac{1-\theta}{1+10^{\,0.09018+2729.92/(273.16+T_s)-1.3\,pH}}\,, \tag{6}$$

where θ and T_s are as above and pH (-) measures soil acidity.

2.3 Nitrous oxide model

Flechard et al. (2007) proposed a model for emissions of nitrous oxide in response to fertilization. They write the emission factor EF_{N_2O}, which is the ratio of N_2O-emission to fertilization in g N g^{-1} N, as a function of precipitation (P; mm $month^{-1}$), soil temperature (T_s; ($^\circ C$)), and soil moisture (θ; -):

$$EF_{N_2O} = 0.01\,exp\left(-5.52 + 0.01P + 0.18T_s + \frac{2.4}{\left(\frac{\theta-0.75}{0.15}\right)^6+1}\right). \tag{7}$$

The nitrous oxide model was derived by regression on data from 40 fertilization events across seven countries, with P ranging from 0 to 279 mm $month^{-1}$), T_s from 1.0 to 24.8 ($^\circ C$)) and θ from 0.27 to 0.89 (Flechard et al., 2007). We truncated our environmental data when they exceeded the regression-intervals.

2.4 Input data for the modelling

The methane model requires data on soil temperature, carbon content and water content. Here we use data for June 1999 taken from the CHESS dataset on the server of the Centre for Ecology and Hydrology (https://eip.ceh.ac.uk/chess), for a 1 x 1 km grid across the United kingdom covering 217,369 cells. Soil temperature is not in the dataset, and we use air temperature instead. Soil water content is also not provided, and is estimated from soil carbon content as $\theta = 0.97/(1+exp(1-C_s/20))$, which implies high water levels in carbon-rich organic soils. This empirical function is not intended to yield a reliable map for water contents in UK soils, its sole purpose is to allow application of the model so we can estimate its upscaling error Δ, and assess the quality of the approximation provided by the $\widehat{\Delta}$-formula. Fig. 1 shows UK-wide maps for the three input variables.

The same data on soil temperature and water content are used by the ammonia and nitrous oxide models. The ammonia model additionally requires data on soil pH which were retrieved from the Harmonized World Soil Database (HWSD), and the nitrous oxide model requires data on monthly rainfall which were taken from the CHESS dataset. Both variables were retrieved at 1 x 1 km resolution.

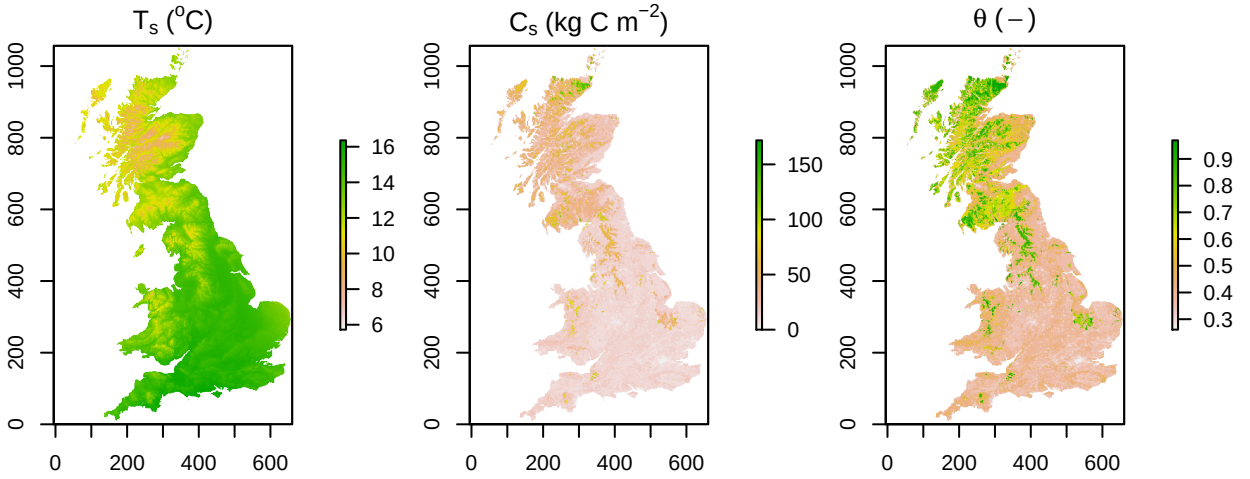

Figure 1: Input maps at 1 x 1 km resolution for June 1999.

Next we created low-resolution input maps by aggregating the 1 x 1 km cells into 32 x 32 km blocks. For each of the 322 blocks, we not only calculated the mean input values across its 1024 cells (or fewer in case some cells covered water bodies), but also their variances (Fig. 2) and covariances.

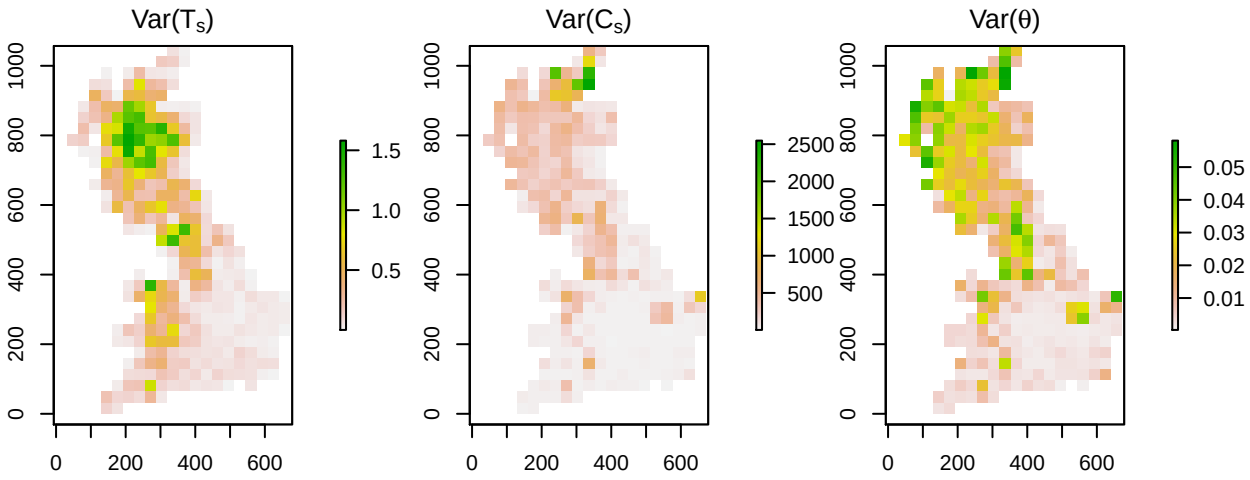

Figure 2: Variances of inputs within 32 x 32 km blocks.

2.5 Calculating actual (Δ) and approximated ($\widehat{\Delta}$) upscaling error

The methane model was applied to each 1 x 1 km cell using the high-resolution input maps (Fig. 1), and mean outputs were calculated for each 32 x 32 km block, thus providing values for E[f(x)]. Then the same calculations were carried out with the models for nitrous oxide and ammonia. The three models were also run using the mean inputs for each block to provide values for f(E[x]). Together, these results were used to calculate each model's spatial upscaling error $\Delta = f(E[x]) - E[f(x)]$, for each of the 322 blocks.

Besides the actual upscaling error Δ, we calculated the predicted upscaling error $\widehat{\Delta}$ given by Eq. (4) using the R-code for calculating each block's Hessian matrix (H) shown in the Appendix. Variance-

covariance matrices (S) were calculated for each block from the input values of its constituent cells, as described above.

3 Results

Upscaling error depends on the distribution of input values in a region. To increase the readability of the equations to follow, we shall use the symbol μ for the mean of the input distribution, which can be a scalar or a vector depending on the dimensionality of model input. We write σ^2 for the variance of univariate input distributions, and Σ for the variance-covariance matrix of multivariate distributions.

3.1 Analytical modelling: functions of scalar input

In this section and the next, we show that upscaling error Δ can be quantified exactly for a number of simple models and input distributions, and we compare such exact solutions with the values estimated by means of the $\widehat{\Delta}$-formula. We first examine models with univariate input. We already saw that for linear models, $f(x) = ax + b$, upscaling error Δ and its estimator $\widehat{\Delta}$ are both zero. Upscaling error is not zero for the simple quadratic function $f(x) = x^2$. For that case, Eq. (1) shows that $\Delta = \mu^2 - E[x^2]$. We can simplify that by noting that the variance of any distribution equals $\sigma^2 = E[x^2] - E[x]^2 = E[x^2] - \mu^2$, so Δ is equal to $-\sigma^2$. Upscaling error for $f(x) = x^2$ thus is equal to minus its variance, and this is true irrespective of the type of distribution for x. And the $\widehat{\Delta}$-formula gives the same result because $f^{(2)}(\mu) = 2$, so $\widehat{\Delta} = -\frac{1}{2}\sigma^2 \times 2 = -\sigma^2$. In other words, for the quadratic function, the estimate of upscaling error provided by the $\widehat{\Delta}$-formula is not just an approximation, it is the exact value.

Using a similar derivation, we can generalise this result to find that $\Delta = \widehat{\Delta} = -a\,\sigma^2$ for $f(x) = a\,x^2 + b\,x + c$, for any values of a, b and c, and for any type of distribution for x.

It is possible to generalise to higher powers of x than 2, but for such models the type of input distribution does affect Δ as well as the quality of the $\widehat{\Delta}$-estimator. We shall limit ourselves to discussing calculations of Δ and $\widehat{\Delta}$ for $f(x) = x^3$ and $f(x) = x^4$, under three common distributions $p(x)$: Gaussian, exponential and lognormal.

Expectation values for x^3 and x^4 can be found by means of moment generating functions (mgf). The mgf for any distribution p(x) is defined as $M_{p(x)}(t) = E[e^{t\,x}]$. This expectation exists for most common distributions except the lognormal one. The property of mgf's that we need here is that the expectation of x^n under any distribution can be found as the nth-order derivative of the mgf for that distribution, evaluated at t=0. In formula-form: $E[x^n] = M^{(n)}_{p(x)}(0)$.

For Gaussian distributions, the mgf is $M_G(t) = e^{t\mu + \frac{1}{2}\sigma^2 t^2}$, and for exponential distributions, $M_e(t) = 1/(1 - \mu t)$. Taking 3rd- and 4th-order derivatives and setting t to 0 shows that for Gaussian distributions, $E[x^3] = \mu^3 + 3\mu\sigma^2$ and $E[x^4] = \mu^4 + 6\mu^2\sigma^2 + 3\sigma^4$. And for inputs with an exponential distribution (for which $\mu = \sigma$), we find that $E[x^3] = 6\mu^3$ and $E[x^4] = 24\mu^4$. For lognormal distributions we cannot use mgf's but integrals over x can be calculated directly, leading to $E[x^3] = e^{3\mu_l + 4.5\sigma_l^2}$ and $E[x^4] = e^{4\mu_l + 8\sigma_l^2}$, where μ_l and σ_l are mean and standard deviation of ln(x) rather than of x itself. With these results, we can now calculate the actual upscaling error Δ for $f(x) = x^3$ or x^4 under the three distributions (Gaussian, exponential, lognormal), as shown in

Table 1: Upscaling error for functions $f(x)$ when input is distributed as $p(x)$ with mean μ and standard deviation σ. Δ: true upscaling error; $\widehat{\Delta}$: upscaling error approximated using Eq. (3); μ_l and σ_l: mean and standard deviation of $ln(x)$.

$f(x)$	$p(x)$	Δ	$\widehat{\Delta}$
x	ALL	0	0
x^2	ALL	$-\sigma^2$	$-\sigma^2$
ax^2+bx+c	ALL	$-a\sigma^2$	$-a\sigma^2$
x^3	Gaussian	$-3\mu\sigma^2$	$-3\mu\sigma^2$
	exponential	$-5\mu^3$	$-3\mu\sigma^2 = -3\mu^3$
	lognormal	$\mu^3 - e^{3\mu_l+4.5\sigma_l^2}$	$-3\mu\sigma^2$
x^4	Gaussian	$-6\mu^2\sigma^2 - 3\sigma^4$	$-6\mu^2\sigma^2$
	exponential	$-23\mu^4$	$-6\mu^2\sigma^2 = -6\mu^4$
	lognormal	$\mu^4 - e^{4\mu_l+8\sigma_l^2}$	$-6\mu^2\sigma^2$

Table 1. In contrast to the actual upscaling error, our estimator for upscaling error, $\widehat{\Delta}$, does not distinguish between the distributions, but is always $-\frac{1}{2}\sigma^2 f^{(2)}(\mu)$. This evaluates to $\widehat{\Delta} = -3\sigma^2\mu$ for $f(x) = x^3$ and $\widehat{\Delta} = -6\sigma^2\mu^2$ for $f(x) = x^4$. We summarise our analytical results so far in Table 1.

3.2 Analytical modelling: functions of vector input

We now turn to functions f(x) where x is an n-dimensional vector with elements $x_1, .., x_n$. A simple case is that of so-called quadratic forms, which can be written as $f(x) = x^T \Lambda x$, where x^T is the transpose of the column vector x and Λ is a symmetric matrix of coefficients. For example, if n=2 and Λ has a first row of (5,1) and a second row (1,7), then $f(x) = f(x_1, x_2) = 5x_1^2 + 2x_1x_2 + 7x_2^2$. For any multivariate distribution p(x) with mean μ and covariance matrix Σ, the expectation of the quadratic form is:

$$E[f(x)] = E[x^T \Lambda x] = tr(\Lambda\Sigma) + \mu^T \Lambda \mu. \tag{8}$$

In contrast, $f(E[x]) = \mu^T \Lambda \mu$. So the error Δ is simply $-tr(\Lambda\Sigma)$. To test how well this result could have been predicted, we now need to use the multivariate form of the error-estimation formula, $\widehat{\Delta} = -\frac{1}{2} tr(\Sigma H)$, where H is the Hessian of $f(x)$. For the bivariate quadratic form, the Hessian of $f(x) = f(x_1, x_2)$ is:

$$H = 2 \begin{pmatrix} \Lambda_{11} & \Lambda_{12} \\ \Lambda_{12} & \Lambda_{22} \end{pmatrix} = 2\Lambda. \tag{9}$$

Therefore, the predicted upscaling error becomes $\widehat{\Delta} = -tr(\Sigma\Lambda)$, and we thus find that the upscaling error formula also gives the correct result for quadratic forms.

We leave our list of analytical examples at this point, except to note that actual and estimated upscaling errors $(\Delta, \widehat{\Delta})$ for any function f(x) that can be written as a sum of other functions, e.g. $f(x, y) = g(x) + h(y)$, are simply the sum of the errors for g(x) and h(y). This follows directly from the fact that the expectation operator E[] is linear.

3.3 GHG emission modelling

Methane emissions as predicted by the Levy et al. (2012) model for a 1 x 1 km grid across the UK are shown in the left panel of Fig. 3. When we scale-up correctly to 32 x 32 km, by averaging the model output values for each block, we get the result shown in the middle panel. The result of incorrect upscaling, by first averaging the inputs for each block and then applying the model to the mean input, is shown in the right panel.

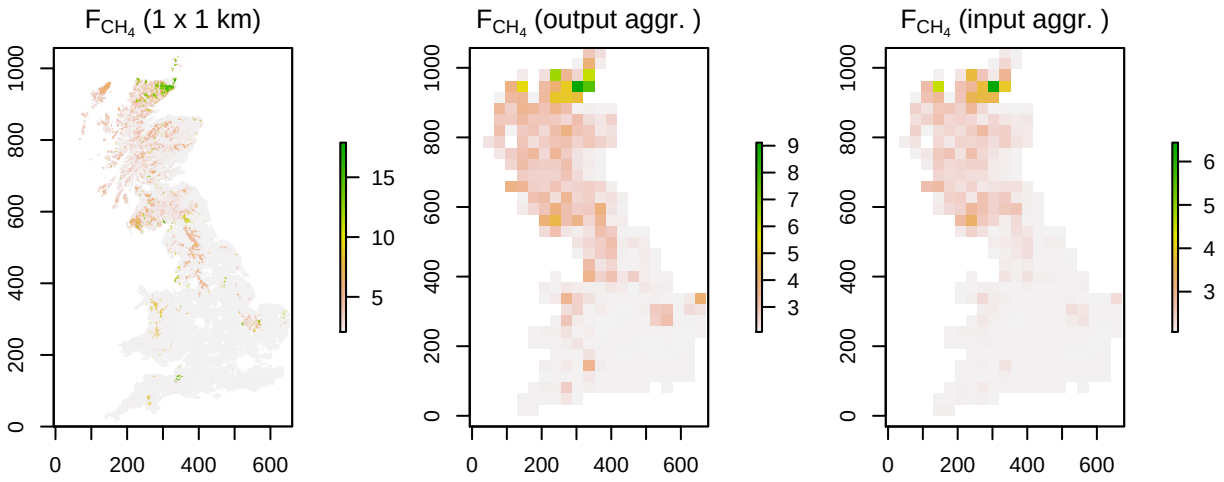

Figure 3: Outputs of methane model (nmol m^{-2} s^{-1}). From left to right: 1 x 1 km, 32 x 32 km (after output aggregation), 32 x 32 km (after input aggregation).

The upscaling error Δ is calculated by subtracting the correctly upscaled model results (Fig. 3, middle) from the incorrect upscaling (Fig. 3, right), and is mapped in Fig. 4, left panel. The upscaling errors that were predicted by the $\widehat{\Delta}$-formula applied to the inputs and their (co)variances are shown in the middle panel of the same figure. The quality of the error-prediction can be evaluated from the scatterplot on the right showing $\widehat{\Delta}$ vs. Δ for all blocks.

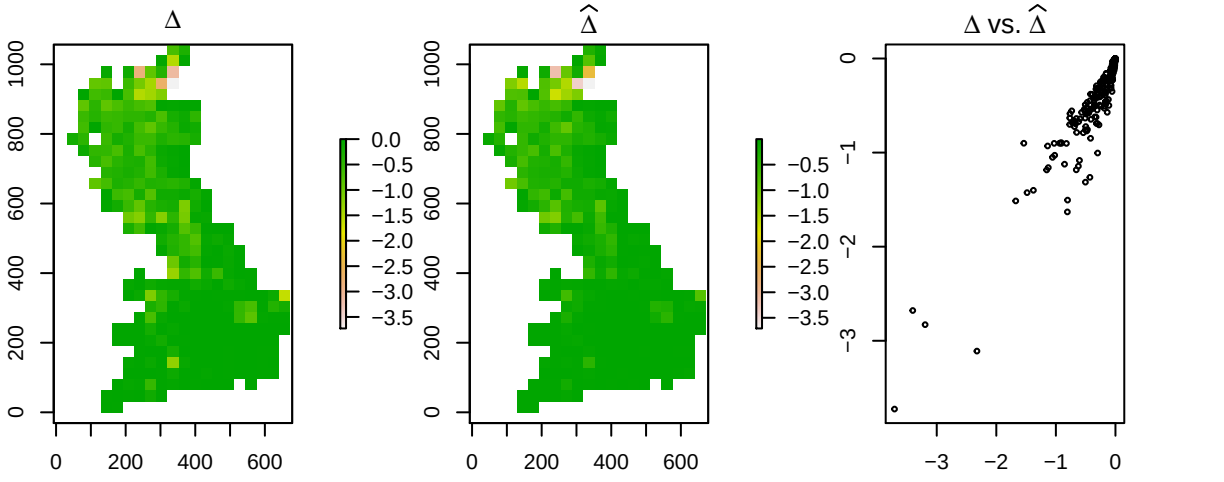

Figure 4: Error analysis for the methane model (nmol m^{-2} s^{-1}).

The same calculations were carried out for the ammonia and nitrous oxide models. We summarise

Table 2: Average GHG model outputs for blocks of 32 x 32 km across the UK (n=322).

Model	UK mean	UK mean after input aggregation	UK mean after input aggregation and error correction
Methane (F_{CH_4}; nmol m^{-2} s^{-1})	2.54	2.25 (-12%)	2.48 (-3%)
Ammonia (EF_{NH_3}; mg N g^{-1} N)	37.0	19.3 (-48%)	33.8 (-9%)
Nitrous oxide (EF_{N_2O}; mg N g^{-1} N)	2.80	2.71 (-3%)	2.87 (+2%)

the results for the three models, i.e. the average emissions across 32 x 32 km blocks, in Table 2.

4 Discussion

4.1 Need and scope for correcting upscaling errors

Ecological modelling is increasingly a problem of big data. Networks of sensors for environmental monitoring retrieve data from multiple sites across countries, and modern methods of data assimilation are employed to feed the measurement information into models. Most of the data pertain to small areas at or around the locations of the sensors, and therefore models that are calibrated using such data have small spatial support. Despite this, the models are required to generate country-wide maps of environmental and ecological variables even though some model input variables may not be available at high spatial resolution for the whole area of interest. A model could, for example, depend on data for atmospheric pollutant deposition which EMEP provides for Europe at a 50 x 50 km resolution (http://www.emep.int/mscw/mscw_data.html). The resulting problem of model spatial interpolation or upscaling is therefore of increasing importance, in particular for nonlinear models. Whenever upscaling is done using a grid-based approach, the correction-formula studied here can be used. The resolution of the grid does not matter: as long as the spatial support of the model is less than the size of a grid-cell, error is introduced and can be estimated using the $\widehat{\Delta}$-formula. Despite the fact that our GHG examples were for upscaling from 1 x 1 km to 32 x 32 km, the formalism applies to upscaling in general, irrespective of the source- and target-resolutions. The $\widehat{\Delta}$-formula only considers the (co)variances of input variables, not their spatial extent.

In fact, the formula can also be applied to upscaling in time rather than space. When, for example, an environmental model has been calibrated using daily values of input-output relationships, then upscaling error will arise if monthly inputs are used instead (Nonhebel, 1994). Soil conditions are unlikely to change rapidly over time but, depending on the source- and target resolutions of the scaling, weather conditions may have similar within-block variance in time and space. Wirtz (2000) used Taylor-expansion based error correction with a complex model for the dynamics of canopy photosynthesis, allowing the use of large timesteps in order to reduce computation time.

4.2 Adequacy of the $\widehat{\Delta}$-formula

We showed that for several analytical models, the formula provides exact estimates of spatial upscaling error. This included models with polynomial scalar inputs up to order two ($f(x) = ax^2 + bx + c$) and models with vector inputs that follow a quadratic form ($f(x) = x^T \Lambda x$). So for these models, upscaling error is independent of the type of distribution of inputs - we only need to know the (co)variances - and can be quantified and corrected. This is also possible for more complex models that are written as the sum of such models.

We showed that the $\widehat{\Delta}$ formula also gives the correct estimate of upscaling error for $f(x) = x^3$ when the distribution of x is Gaussian (Table 1). If the distribution is exponential, then the formula underestimates the error by 40%, but the error will be small if the mean of x is close to zero. For lognormal distributions the upscaling error and quality of its estimator depend on both the mean and the variance of the inputs - and can only be studied on a case-by-case basis. For $f(x) = x^4$ we see slightly worse results (Table 1). If the distribution of x is Gaussian and the variance is smaller than the mean, then the $\widehat{\Delta}$-formula provides a good estimate of error. If the distribution is exponential, then the error-estimate is only 6/23 of the true value. For lognormal distributions, again no general statement can be made.

The analysis of the quadratic forms suggests that the results for polynomial functions of scalar input tend to hold also for polynomial functions with vector input of the same order. This was confirmed by the results for Levy et al's (2012) methane model, which includes a dependency on water content to the power 4.77 (Eq. (5)). According to the model, the average methane emission for the 217,369 grid cells of 1 x 1 km was 2.55 nmol m^{-2} s^{-1} (Fig. 3, left panel). A similar average emission value of 2.54 nmol m^{-2} s^{-1} was found after correct upscaling: aggregating the cells into 322 blocks of 32 x 32 km and averaging model outputs per block (Fig. 3, middle panel). In contrast, when block-emissions were calculated by first averaging inputs and then running the model only once per block, an average emission of 2.25 nmol m^{-2} s^{-1} was found, i.e. an error of -12% (Fig. 3, right panel). This error could be predicted well (Fig. 4, right panel), and if we had corrected each input-averaged block prediction, average emission would have been estimated as 2.48 nmol m^{-2} s^{-1}, thus reducing the error for the UK-mean to -3%.

The need and scope for error correction in the methane model would have been even stronger if we had limited the study to the blocks with the largest emissions. For the 55 blocks with emissions larger than 3 nmol m^{-2} s^{-1}, upscaling error from input-aggregation was -27% but after correction using the $\widehat{\Delta}$-formula this was again reduced to only -3%.

We found an even larger average upscaling error for the ammonia model, -48%, which was reduced to -9% by application of the $\widehat{\Delta}$-formula to each 32 x 32 km block. For the nitrous oxide model, average upscaling error was small, only -3%. In this case, the formula overestimated the error, so the slightly underestimated emission was changed to an overestimate of 2%. We analysed this case further and found that the key problem was the complicated effect of soil water content on N_2O-emission. Using the code in the Appendix, we calculated the second derivative of EF_{N_2O} as a function of soil water content and found it to vary strongly, as shown in Figure 5, left panel. We see that the second derivative has a local maximum at $\theta \approx 0.60$ and a local minimum at $\theta \approx 0.65$. So the upscaling error for any block of 32 x 32 km with a mean water content close to either of these extremes is likely to lead to under- resp. overestimation of Δ (because of the minus-sign in the $\widehat{\Delta}$-formula). We see this confirmed in Figure 5, right panel, where for each of the 322 blocks, the error in $\widehat{\Delta}$ is plotted against the block average of θ.

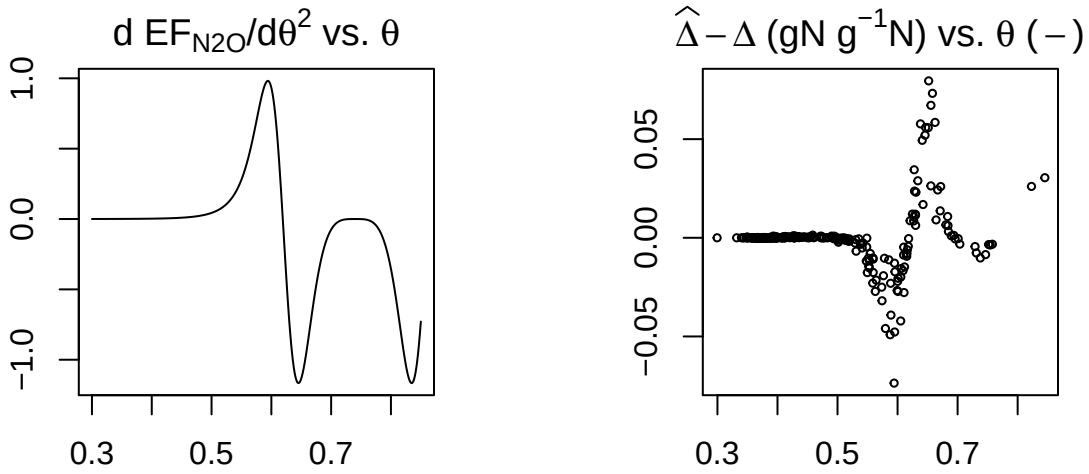

Figure 5: Left: second derivative of EF_{N2O} as a function of θ for P = 50 mm month^{-1} and $T_s = 10$ $^{\circ}C$. Right: overestimation of upscaling error for the N_2O-model in the 322 blocks of 32 x 32 km.

Overall, we find that the $\widehat{\Delta}$-formula in many cases provides the correct value or a close approximation to the true upscaling error. Where it does not do so, as in the case of exponentially distributed inputs, the error-estimates tend to be too low, but still usable for a partial correction of the model results.

We found that the degree of upscaling error was difficult to predict from a quick inspection of model structure and input distributions. The three GHG models all were nonlinear functions of three input variables, with temperature and soil water content being common to the models. Despite this superficial similarity, the models differed strongly in UK-wide upscaling error, from -3% for EF_{N_2O} via -12% for F_{CH_4} to -48% for EF_{NH_3}. The fact that these differences were predicted and corrected well by the $\widehat{\Delta}$-formula (Table 2) shows that the upscaling errors were largely accounted for by differences in the covariance matrices S and Hessian matrices H. We recommend that modellers calculate these matrices, and $\widehat{\Delta}$, at the time of any new model application that involves spatial upscaling.

4.3 Possible limitations to the use of the $\widehat{\Delta}$-formula

In the Introduction we mentioned four possible reasons for the minimal uptake of the $\widehat{\Delta}$-formula by the ecological modelling community, and we discuss these now.

The first problem arises when the model is high-dimensional, hampering quantification of the S- and H-matrices. In such cases, it may be necessary to reduce the dimensionality of the model by identifying the dominant inputs and fitting a response surface to them. Corstanje et al. (2008) found, for process-based models of NH_3 volatilization, that only a subset of the model's input variables caused errors when spatially aggregated. Procedures such as response-surface fitting are likely to have a smoothing effect, thereby reducing the absolute magnitude of the second derivatives. That in itself would already reduce upscaling error, which could be reduced further using the $\widehat{\Delta}$-formula. Such a combination of dimensionality-reduction and error-correction deserves closer study.

A second issue is that many ecological models are not twice differentiable (e.g. if they contain discrete response thresholds) or have second-order partial derivatives that are hard to calculate

(e.g. in the case of process-based models). However, there is no need to find analytical solutions for the second-order derivatives as the Hessian can be quantified numerically. For that purpose, we implemented algorithms in R, which can be found in the Appendix.

Thirdly, the model may be poorly approximated by a second-order Taylor expansion - on which the $\widehat{\Delta}$-formula is based. This is a valid objection to the approach, but in practice the formula may still provide useful results as we showed here for $f(x) = x^3$, $f(x) = x^4$ and the GHG models. It may be prudent though, when using a new model, to evaluate the behaviour of the $\widehat{\Delta}$-formula using a range of measured or virtual data, as we did in this paper.

Finally, information may be lacking about the variances and covariances of inputs. If all we have is information about the mean value of inputs in a region, then the variance-covariance matrix itself needs to be estimated before we can estimate upscaling error. For some input variables, this may be done using proxy-variables. For example, high-resolution maps of elevation are available for many regions, and the variance of temperature (in $°C$) can often be approximated well as 0.006^2 times the variance of elevation (in m). And meteorologists have estimated spatial heterogeneity in surface roughness from the regional density of edges bounding different land-use types, for which good maps are available (Ridder, 2003). Where no proxy can be found, it may be possible to derive estimates by comparing the target-region with similar but better-studied regions elsewhere. For soil variables we could use commonly observed empirical variograms that relate variance between points to distance (e.g. Webster and Oliver, 1990). For weather variables, the variances and covariances built into weather generators may be suitable (e.g. Yiou, 2014). Whenever such indirect information is used, it will be important to quantify the resulting uncertainty about the covariance matrix S, which will propagate linearly to uncertainty about Δ itself. In contrast, the Hessian matrix H can be determined exactly for deterministic environmental models.

4.4 Extending the accuracy and use of the $\widehat{\Delta}$-formula

We now briefly discuss the scope for improving the $\widehat{\Delta}$-formula itself. One obvious way to do so would be to truncate the Taylor expansion for the model function at a higher order than two. This was already proposed by Rastetter et al. (1992). Doing this might prevent the overshoot in error-correction that we found for the nitrous oxide model (Table 2, Fig. 5). However, higher-order truncation would exacerbate the problem of assembling the information needed for error-estimation. Not only would we then need to quantify at least the third-order partial derivatives of the model, but we would also need to find estimates for the higher central moments of the input distribution beyond the variance.

On the other hand, it is possible to extend our approach by not only providing estimates for E[f(x)] but also for the variance Var[f(x)]. Using the same Taylor-expansion based approach, Bresler and Dagan (1988) derived an approximative formula for Var[f(x)]. That may be useful when we want to estimate not only the correct mean model output for a region, but also what fraction of the region has output-values that exceed a given threshold.

5 Conclusions

- Spatial aggregation of the inputs of nonlinear models causes upscaling error.

- The upscaling error can be estimated using the $\widehat{\Delta}$ formula, known since the 1980s but rarely used in ecological and environmental modelling.
- $\widehat{\Delta}$ provides the exact value of upscaling error for some analytical models and input distributions, and good approximations for others.
- For Levy et al's (2012) methane model, average error could be reduced from -12% to -3% for the whole of the UK, and from -27% to -3% for high-emission areas. For Denmead et al's (1982) ammonia model, average error was -48% resp. -9% before and after error correction. For the nitrous oxide model of Flechard et al. (2007), the errors were smaller and swapped sign from -3 to +2%.
- The $\widehat{\Delta}$ formula requires information on input variances and covariances; where this is lacking, proxy variables or comparisons with better-studied regions may be needed.
- The formula is difficult to apply when models are high-dimensional; the possibility to reduce their dimensionality by fitting response-surfaces to key input variables needs to be examined.

Acknowledgements

This work was supported by the UK's Natural Environment Research Council through projects GREENHOUSE and MACSUR. We thank our colleagues in both projects for fruitful discussions and for data used in the GHG emission modelling.

Appendix

5.1 R-code for the Hessian

The following code can be used with any model f to calculate its Hessian around an input vector x (typically the grid-cell mean). The code consists of three functions. The first function calculates first-order derivatives (slopes). It is called by the second function, which calculates second-order derivatives (curvatures) and is itself called by the third function to populate the Hessian matrix.

```
df_dxi <- function( f, x, i=1, dxi=1.E-2 ) {
  dimf   <- length(x)
  dx_i   <- rep(0,dimf) ; dx_i[i] <- dxi
  xiplus <- x + dx_i
  df_xi  <-( f(xiplus) - f(x) ) / dxi
  return(df_xi)
}

d2f_dxixj <- function( f, x, i=1, j=2, dxi=1.E-2, dxj=1.E-2 ) {
  dimf      <- length(x)
  dx_j      <- rep(0,dimf) ; dx_j[j] <- dxj
  xjplus    <- x + dx_j
  d2f_dxixj <- ( df_dxi(f,xjplus,i,dxi) - df_dxi(f,x,i,dxi) ) / dxj
  return(d2f_dxixj)
}
```

```
Hf <- function( f, x, dx=rep(1.E-2,length(x)) ) {
  dimf <- length(x)
  H    <- matrix(0,dimf,dimf)
  for(i in 1:dimf) {
    for(j in 1:dimf) {
      dxi<- dx[i] ; dxj <- dx[j]
      H[i,j] <- d2f_dxixj( f, x=x, i=i, j=j, dxi=dxi, dxj=dxj )
    }
  }
  return(H)
}
```

5.2 R-code for $\widehat{\Delta}$

Direct implementation in R of Eq. (4) for the estimation of Δ cannot be done with R's base functions as they do not include a function for calculating the trace of a matrix. To avoid having to install an R-package (such as 'matrixcalc') that does contain a trace function, our code implements the following formula which for symmetric matrices such as S is mathematically equivalent to Eq. (4):

$$\widehat{\Delta} = -\frac{1}{2}\sum_{i,j} S(i,j)H(i,j).$$

The code that implements this formula is:

```
Delta_hat <- function( f, x, S, dx=rep(1.E-2,length(x)) ) {
  Delta_hat <- -0.5 * sum( S * Hf(f=f,x=x,dx=dx) )
  return(Delta_hat)
}
```

References

Avissar, R., Pielke, R.A., 1989. A Parameterization of Heterogeneous Land Surfaces for Atmospheric Numerical Models and Its Impact on Regional Meteorology. Monthly Weather Review 117, 2113–2136. doi:10.1175/1520-0493(1989)117<2113:APOHLS>2.0.CO;2

Band, L.E., Peterson, D.L., Running, S.W., Coughlan, J., Lammers, R., Dungan, J., Nemani, R., 1991. Forest ecosystem processes at the watershed scale: Basis for distributed simulation. Ecological Modelling 56, 171–196. doi:10.1016/0304-3800(91)90199-B

Bernardo, J.M., Smith, A.F.M., 1994. Bayesian Theory. Wiley.

Bresler, E., Dagan, G., 1988. Variability of yield of an irrigated crop and its causes: 1. Statement of the problem and methodology. Water Resources Research 24, 381–387. doi:10.1029/WR024i003p00381

Cameron, D.R., Van Oijen, M., Werner, C., Butterbach-Bahl, K., Grote, R., Haas, E., Heuvelink,

G.B.M., Kiese, R., Kros, J., Kuhnert, M., Leip, A., Reinds, G.J., Reuter, H.I., Schelhaas, M.J., De Vries, W., Yeluripati, J., 2013. Environmental change impacts on the C- and N-cycle of European forests: A model comparison study. Biogeosciences 10, 1751–1773. doi:10.5194/bg-10-1751-2013

Corstanje, R., Kirk, G.J.D., Lark, R.M., 2008. The behaviour of soil process models of ammonia volatilization at contrasting spatial scales. European Journal of Soil Science 59, 1271–1283. doi:10.1111/j.1365-2389.2008.01086.x

Corstanje, R., Lark, R.M., 2008. On effective linearity of soil process models. European Journal of Soil Science 59, 990–999. doi:10.1111/j.1365-2389.2008.01046.x

Denmead, O.T., Freney, J.R., Simpson, J.R., 1982. Dynamics of ammonia volatilization during furrow irrigation of maize. Soil Science Society of America Journal 46, 149. doi:10.2136/sssaj1982.03615995004600010028x

Eyshi Rezaei, E., Siebert, S., Ewert, F., 2015. Impact of data resolution on heat and drought stress simulated for winter wheat in Germany. European Journal of Agronomy 65, 69–82. doi:10.1016/j.eja.2015.02.003

Flechard, C.R., Ambus, P., Skiba, U., Rees, R.M., Hensen, A., van Amstel, A., Dasselaar, A. van den P.-v., Soussana, J.F., Jones, M., Clifton-Brown, J., Raschi, A., Horvath, L., Neftel, A., Jocher, M., Ammann, C., Leifeld, J., Fuhrer, J., Calanca, P., Thalman, E., Pilegaard, K., Di Marco, C., Campbell, C., Nemitz, E., Hargreaves, K.J., Levy, P.E., Ball, B.C., Jones, S.K., van de Bulk, W.C.M., Groot, T., Blom, M., Domingues, R., Kasper, G., Allard, V., Ceschia, E., Cellier, P., Laville, P., Henault, C., Bizouard, F., Abdalla, M., Williams, M., Baronti, S., Berretti, F., Grosz, B., 2007. Effects of climate and management intensity on nitrous oxide emissions in grassland systems across Europe. Agriculture, Ecosystems & Environment, The greenhouse gas balance of grasslands in europe 121, 135–152. doi:10.1016/j.agee.2006.12.024

Heuvelink, G.B.M., Pebesma, E.J., 1999. Spatial aggregation and soil process modelling. Geoderma 89, 47–65. doi:10.1016/S0016-7061(98)00077-9

Hines, W., Montgomery, D., 1990. Probability and Statistics in Engineering and Management Science, 3rd ed. Wiley, New York.

Hoffmann, H., Zhao, G., Asseng, S., Bindi, M., Biernath, C., Constantin, J., Coucheney, E., Dechow, R., Doro, L., Eckersten, H., Gaiser, T., Grosz, B., Heinlein, F., Kassie, B.T., Kersebaum, K.-C., Klein, C., Kuhnert, M., Lewan, E., Moriondo, M., Nendel, C., Priesack, E., Raynal, H., Roggero, P.P., Rötter, R.P., Siebert, S., Specka, X., Tao, F., Teixeira, E., Trombi, G., Wallach, D., Weihermüller, L., Yeluripati, J., Ewert, F., 2016. Impact of Spatial Soil and Climate Input Data Aggregation on Regional Yield Simulations. PLOS ONE 11, e0151782. doi:10.1371/journal.pone.0151782

Kuhnert, M., Yeluripati, J., Smith, P., Hoffmann, H., Van Oijen, M., Constantin, J., Dechow, R., Eckersten, H., Gaiser, T., Grosz, B., Haas, E., Kersebaum, K.-C., Kiese, R., Klatt, S., Lewan, E., Nendel, C., Raynal, H., Sosa, C., Specka, X., Teixeira, E., Wang, E., Weihermüller, L., Zhao, G., Zhao, Z., Ogle, Ewert, F., 2016. Impact analysis of climate data aggregation at different spatial scales on simulated Net Primary Productivity for croplands. European Journal of Agronomy in press.

Levy, P.E., Burden, A., Cooper, M.D.A., Dinsmore, K.J., Drewer, J., Evans, C., Fowler, D., Gaiawyn, J., Gray, A., Jones, S.K., Jones, T., McNamara, N.P., Mills, R., Ostle, N., Sheppard, L.J., Skiba, U., Sowerby, A., Ward, S.E., Zieliński, P., 2012. Methane emissions from soils: Synthesis and analysis of

a large UK data set. Global Change Biology 18, 1657–1669. doi:10.1111/j.1365-2486.2011.02616.x

Nonhebel, S., 1994. The effects of use of average instead of daily weather data in crop growth simulation models. Agricultural Systems 44, 377–396.

Potter, C., Klooster, S., Krauter, C., 2003. Regional modeling of ammonia emissions from native soil sources in California. Earth Interactions 7, 1–28.

Rastetter, E.B., King, A.W., Cosby, B.J., Hornberger, G.M., O'Neill, R.V., Hobbie, J.E., 1992. Aggregating Fine-Scale Ecological Knowledge to Model Coarser-Scale Attributes of Ecosystems. Ecological Applications 2, 55. doi:10.2307/1941889

Ridder, K.D., 2003. Horizontal Scale of Roughness Variations for Realistic Landscapes. Boundary-Layer Meteorology 109, 49–57. doi:10.1023/A:1025407819823

Schweiger, A.H., Beierkuhnlein, C., 2016. Scale dependence of temperature as an abiotic driver of species' distributions. Global Ecology and Biogeography 25, 1013–1021. doi:10.1111/geb.12463

Taylor, P.A., 1987. Comments and further analysis on effective roughness lengths for use in numerical three-dimensional models. Boundary-Layer Meteorology 39, 403–418. doi:10.1007/BF00125144

Van Oijen, M., Thomson, A., 2010. Toward Bayesian uncertainty quantification for forestry models used in the United Kingdom Greenhouse Gas Inventory for land use, land use change, and forestry. Climatic Change 103, 55–67. doi:10.1007/s10584-010-9917-3

Van Oijen, M., Thomson, A., Ewert, F., 2009. Spatial upscaling of process-based vegetation models: An overview of common methods and a case-study for the UK, in: StatGIS. Milos, Greece, p. 6.

Webster, R., Oliver, M.A., 1990. Statistical Methods in Soil and Land Resource Survey. Oxford University Press.

Welsh, A., Peterson, A., Altmann, S., 1988. The fallacy of averages. The American Naturalist 132, 277–288.

Wirtz, K.W., 2000. Second order up-scaling: Theory and an exercise with a complex photosynthesis model. Ecological Modelling 126, 59–71. doi:10.1016/S0304-3800(99)00174-X

Yiou, P., 2014. AnaWEGE: A weather generator based on analogues of atmospheric circulation. Geosci. Model Dev. 7, 531–543. doi:10.5194/gmd-7-531-2014

Zhang, W., 2012. Spatial upscaling of a forest ecosystem model (MSc-thesis). University of Edinburgh.

Zhao, G., Siebert, S., Enders, A., Rezaei, E.E., Yan, C., Ewert, F., 2015. Demand for multi-scale weather data for regional crop modeling. Agricultural and Forest Meteorology 200, 156–171. doi:10.1016/j.agrformet.2014.09.026