Spatial Prediction of \(\text{N}_2\text{O}\) Emissions in Pasture: A Bayesian Model Averaging Analysis

Xiaodong Huang\textsuperscript{1,2,*}, Peter Grace\textsuperscript{2}, Wenbiao Hu\textsuperscript{3}, David Rowlings\textsuperscript{2}, Kerrie Mengersen\textsuperscript{1}

\textsuperscript{1}Mathematical Sciences, Queensland University of Technology, Brisbane, Australia, \textsuperscript{2}Institute of Sustainable Resources, Queensland University of Technology, Brisbane, Australia, \textsuperscript{3}School of Population Health, The University of Queensland, Brisbane, Australia

Abstract

Nitrous oxide (\(\text{N}_2\text{O}\)) is one of the greenhouse gases that can contribute to global warming. Spatial variability of \(\text{N}_2\text{O}\) can lead to large uncertainties in prediction. However, previous studies have often ignored the spatial dependency to quantify the \(\text{N}_2\text{O}\) – environmental factors relationships. Few researches have examined the impacts of various spatial correlation structures (e.g., independence, distance-based and neighbourhood based) on spatial prediction of \(\text{N}_2\text{O}\) emissions. This study aimed to assess the impact of three spatial correlation structures on spatial predictions and calibrate the spatial prediction using Bayesian model averaging (BMA) based on replicated, irregular point-referenced data. The data were measured in 17 chambers randomly placed across a 271 m\(^2\) field between October 2007 and September 2008 in the southeast of Australia.

We used a Bayesian geostatistical model and a Bayesian spatial conditional autoregressive (CAR) model to investigate and accommodate spatial dependency, and to estimate the effects of environmental variables on \(\text{N}_2\text{O}\) emissions across the study site. We compared these with a Bayesian regression model with independent errors. The three approaches resulted in different derived maps of spatial prediction of \(\text{N}_2\text{O}\) emissions. We found that incorporating spatial dependency in the model not only substantially improved predictions of \(\text{N}_2\text{O}\) emission from soil, but also better quantified uncertainties of soil parameters in the study. The hybrid model structure obtained by BMA improved the accuracy of spatial prediction of \(\text{N}_2\text{O}\) emissions across this study region.

Introduction

Soils have been considered as an important source for nitrous oxide (\(\text{N}_2\text{O}\)), a well-known greenhouse gas [1]. \(\text{N}_2\text{O}\) fluxes often exhibit spatial autocorrelation at multiple scales due to the distribution of soil properties and topography. It is difficult to precisely estimate annual \(\text{N}_2\text{O}\) emissions at a field scale level because of high spatial variability within the field [2]. In light of these large uncertainties in prediction, spatial variation should be an explicit consideration in any analysis of \(\text{N}_2\text{O}\) emissions [3–5].

To date, the relationship between \(\text{N}_2\text{O}\) emissions and environmental covariates has largely been quantified by aggregating over all sites and assuming independent observations in multiple linear regression models. However, the presence of spatial correlation can render these models invalid since they can lead to biased estimates and incorrect inferences [6,7].

In the past decade, a variety of models that take into account the spatial nature of data have been developed [8,9] and are widely applied in ecology, epidemiology, economics and so on. These models can help to better identify and explore influential factors and guide more informed inferences, as well as improve further experimental design in order to obtain more precise estimates [10].

Bayesian spatial conditional autoregressive (CAR) models are appropriate for all locations that have a similar size and are regularly arranged [11], whereas geostatistical models are more suitable for spatial data with unidentified neighbourhoods [7]. Most published research on the comparison of spatial models has been based on areal data with identified neighbours or point data with a regular sampling pattern [7–9,12,13]. However, differences between the CAR model and geostatistical model with respect to parameter estimation and predicted spatial distribution based on point-referenced data with an irregular sampling interval and undetermined boundaries are not well understood.

One concern with spatial models is that different representations of the spatial correlation based on the same dataset might give different estimated effect sizes, inferences about significant parameters or estimated error structures [7,9]. Many candidate spatial correlation structures are available in spatial analysis. It is often difficult to determine the best spatial correlation structure based on standard information-based criteria. However, Bayesian model averaging (BMA) can take account of such model uncertainty and provide better average predictive performance [14,15]. For example, Boone and Bullock [16] used BMA to pool information from four spatial candidate structures in the analysis of a loblolly pine dataset.

In this study, we consider three spatial correlation structures (independence, distance-based and neighbourhood-based) in spatial analyses of \(\text{N}_2\text{O}\) emission for point data obtained from irregular sampling intervals in pasture. All models are developed under a hierarchical Bayesian inferential framework. Key attri-
buties of Bayesian approaches are the use of probability for quantifying uncertainty in inferences, formal accommodation of parameter uncertainty [17], and flexibility of model description [9,18]. The deviance information criterion (DIC) is used to compare the various models [19] and provide weights for BMA [20]. The aims of this study are to assess the effects of various spatial dependencies on spatial prediction, to calibrate spatial predictions of \( N_2O \) by BMA across the study region based on the environment-\( N_2O \) relationships obtained from the three models.

**Materials and Methods**

**Study Site and Data Collection**

The study site is located at Moooloolah (26°39’40” S., 152°56’23”E.) on the Sunshine Coast in Australia. \( N_2O \) (ug \( N_2O \)-N m\(^{-2}\) hr\(^{-1}\)) emissions and 7 potential independent variables (gravimetric soil moisture (%), soil temperature (°C), soil \( \text{NO}_3^- \) concentration (Kg N ha\(^{-1}\)), soil pH; soil sand, silt and clay content (%)) were measured at 17 chambers randomly placed across a 271 m\(^2\) subtropical pasture at monthly intervals between October 2007 and September 2008.

The pasture was a mixture of the tropical grass *Setaria spicata* and the legumes Silverleaf Desmodium (*Desmodium uncinatum*) and White Clover (*Trifolium repens*). No nitrogenous fertilizer had been applied to the pasture site for over 20 years. The soil was classified according to the Australian Soil Classification as a Haplic, Eutrophic, Black Dermosol [21] and had a bulk density of 1.0 g cm\(^{-3}\) (0–10 cm) and an organic carbon content of 2.8%. Average soil texture across the site was classified as a loam.

The closed static chamber technique was used for measurements of \( N_2O \) emissions. Chambers were 200 mm high (diameter 200 mm) inserted 100 mm into the soil, allowing a headspace of 80–100 mm. Chambers remained in situ throughout the length of the experiment. Chambers were closed for one h and sampled using 12 ml evacuated glass vials (Exetainer; Labco, High Wycombe, Buckinghamshire, UK) at zero (0) min and 60 min. Full details of chamber method and site climate are described in Rowlings et al. [22].

**Statistical Analysis**

The observed data can be defined as point-referenced data [23]. In order to assess the effects of different covariance structures on estimates of spatial variation in \( N_2O \) fluxes and compare the estimated parameters among three Bayesian spatial models, we used a Thiessen-polygon approach to convert point-referenced data to areal data. This method creates a polygon enclosing each original point, such that each point has its own polygon. The defined boundaries of the Thiessen polygons can be used to establish a neighbourhood weight matrix for each data point [24].

In this study, we focused on a linear regression model with three different correlation structures: 1) independent model (no spatial correlation structure), 2) geostatistical (EXP) model (spatial correlation described as the exponential decay function of the distance between pairs of points), and 3) conditional autoregressive model (spatial correlation described as first-order neighbourhood). Although other spatial correlation models, for example, simultaneous autoregressive model (SAR) and geostatistical models with other Matérn correlation functions are available, CAR and EXP models are most commonly used in practice [23].

In all of the following models, let \( y_{ir} \) be the observed \( N_2O \) fluxes at location \( i \) for replicate \( r \) \((i=1,…,Q; r=1,…,M; Q=17, M=1)\). The vector \( Y_i=[y_{i1},y_{i2},…,y_{iM}] \) represents \( N_2O \) fluxes at the \( i \)th location. Let \( X_{ir} \) be a vector of length \( K=7 \), representing the covariates comprising soil moisture, soil temperature, \( \text{NO}_3^- \), soil texture (including sand, silt and clay) and soil pH at the \( i \)th site for replicate \( r \). Measurements of \( N_2O \) fluxes exhibited skewness, so were log-transformed to better approximate a normal distribution.

**Bayesian Linear Regression Model (Independent Structure)**

In this first model, we assumed that locations were independent and that \( N_2O \) emissions were affected by the nominated covariates independently at each location so that:

\[
y_{ir} = \beta_0 + \sum_{k=1}^{7} \beta_k x_{irk} + \epsilon_{ir}
\]

where \( \beta_k \) are the regression coefficients and \( \epsilon_{ir} \sim N(0,\sigma^2) \) is the residual under the independence and normality assumptions [25]. In a Bayesian framework, the posterior distribution for the parameters of interest is thus given by:

\[
p(\beta,\sigma^2|Y) \propto p(Y|\beta,\sigma^2)p(\beta)P(\sigma^2)
\]

where \( p(Y|\beta,\sigma^2)=N(Y|\beta,\sigma^2) \). Diffuse priors were imposed for the regression parameters, so that \( \beta \sim N(0,10.66) \) and \( \sigma \sim U(0,5) \).

**Bayesian Geostatistical Model (EXP Model)**

The second model considered is an extension of the normal linear regression described above, with an additional term to account for spatial correlation between the experimental sites. The additional term is modelled as a random effect with the variance reflecting the spatial correlation. Letting \( s=(s_i; i=1,…,17) \) be the vector of site-specific spatial Gaussian random effects, equation (1) is extended as follows:

\[
y_{ir} = \beta_0 + \sum_{k=1}^{7} \beta_k x_{irk} + s_i + \epsilon_{ir}
\]

\[
Y((\beta,\sigma^2,S) \sim N(X\beta+S,\sigma^2 I)
\]

\[
S(\sigma^2,0) \sim N(0,\sigma^2 \Phi(0))
\]

\[
\Phi_{ij}=f(d_{ij},0,\delta), i,j=1,…,17
\]

\[
f(d_{ij},0,\delta)=\exp[-(\delta d_{ij})^\theta], 0<\delta<2; \theta>0
\]

Here, \( \epsilon_{ir} \sim N(0,\sigma^2) \) is a spatially uncorrelated error term; \( I \) is an \( r \times r \) identity matrix and \( S \sim N(0,\sigma^2 \Phi(0)) \) is assumed to be a stationary, isotropic Gaussian process with mean zero and correlation matrix \( \Phi \) with elements \( \Phi_{ij}=f(d_{ij},0,\delta) \) between \( s_i \) and \( s_j \) [23]. The pairwise correlations \( \Phi_{ij} \) are usually described as a parametric function of the distance \( d_{ij} \) between each pair of sites \( i \) and \( j \). The exponential decay function \( f(d_{ij},0,\delta)=\exp[-(\delta d_{ij})^\theta] \) [17] is the most popular. Here \( \theta \) is the rate of decrease in spatial correlation per unit of distance, with a large value of \( \theta \) indicating that the spatial correlation decreases rapidly [6]. The prior distribution for \( \theta \) was specified as Uniform with lower and upper
bounds corresponding to a correlation of 0.05 the maximum distance (25.35 m) and minimum distance (0.75 m), respectively, between any pair of locations across the study site [26]. Covariate coefficients were modelled with diffuse normal prior distributions \( \beta \sim N(0.0, 1.0E6) \). The parameter \( \delta \) controls the amount of spatial smoothing. Thomas et al. [27] advise a value of \( \delta = 1 \). The standard deviation \( \sigma \) was described by a uniform prior \( \sigma \sim U(0,5) \).

**Bayesian Spatial Intrinsic Conditional Autoregressive Model (CAR)**

The third model considered employed a different representation of the spatial nature of the data. Here equation (1) is extended as follows:

\[
y_{i} = \beta_{0} + \sum_{k=1}^{7} \beta_{k} x_{ik} + u_{i} + \epsilon_{i}
\]

\[
Y(\beta, \sigma^2, U) \sim N(X\beta + U, \sigma^2 I)
\]

\[
u_{i}|u_{-i} \sim N(\sum_{j=1}^{Q} w_{ij} u_{j} - \frac{\sigma^2}{\sum_{j=1}^{Q} w_{ij}^2})
\]

Here \( \epsilon_{i} \sim N(0, \sigma^2) \) is the within-site residual variation. A conditional autoregressive (CAR) model was used to describe the spatial component. This is represented by the term \( U \), with elements \( u \), denoting the local dependence at site \( i \) as a function of the site’s neighbours \( u_{i} \), where \( u_{i} = [u_{1i}, u_{2i}, \ldots, u_{ni}, u_{ni+1}, \ldots, u_{Q}]^{T} \) [28]. The local neighbourhood relationship is represented by a symmetric \( n \times n \) matrix \( W \) of spatial weights with elements \( w_{ij} \), and \( \sum_{j=1}^{Q} w_{ij} \). This representation allows a great deal of flexibility in describing the spatial correlation. For example the spatial neighbourhood may be specified only as first-order neighbourhood for each site, in which case \( w_{ij} = 1 \) if sites \( i \) and \( j \) share a boundary, and zero otherwise. As before, all covariate coefficients had diffuse normal priors, given by \( \beta \sim N(0.0, 1.0E6) \), and \( \sigma_{u} \) and \( \sigma \) had uniform priors, \( \sigma_{u} \sim U(0,10) \) and \( \sigma \sim U(0,5) \).

**Bayesian Model Averaging (BMA)**

Bayesian model averaging can account for model uncertainty by taking a weighted average of models over a given model space [14]. Let \( \Delta \equiv 1 \) model structures \( M_{l} \) with parameter set \( \pi_{l} \)based on data \( D \). Let \( \Delta \) be the quantity of interest; this could represent, for example, the posterior predictive distribution of \( y \). Hence the posterior distribution of \( \Delta \) given data \( D \) is [14]:

\[
p(\Delta|D) = \sum_{l=1}^{L} p(\Delta|M_{l}|D)p(M_{l}|D)
\]

The posterior probability for \( M_{l} \) is given by:

\[
p(M_{l}|D) = \frac{p(D|M_{l})p(M_{l})}{\sum_{l=1}^{L} p(D|M_{l})p(M_{l})}
\]

where \( p(D|\pi_{l}, M_{l}) = \int p(D|\pi_{l}, M_{l})p(\pi_{l}|M_{l})d\pi_{l} \).

Here, \( p(D|\pi_{l}, M_{l}) \) is the marginal likelihood of the data given model \( M_{l} \) and \( p(\pi_{l}|M_{l}) \) is the prior density of \( \pi_{l} \) given model \( M_{l} \). \( p(M_{l}) \) is the prior probability for model \( M_{l} \) when \( M_{l} \) is regarded as the true model [14]. A Laplace approximation, typically the Bayesian information criterion (BIC) [29] can be used to approximate \( p(D|\pi_{l}, M_{l}) \) [14,30,31]:

\[
\log\{p(D|\pi_{l}, M_{l})\} \approx \log\{p(D|\pi_{l}, M_{l})\} - d_{l}\log(n)
\]

\[
BIC = -2\log\{p(D|\pi_{l}, M_{l})\} + d_{l}\log(n)
\]

Here \( \log\{p(D|\pi_{l}, M_{l})\} \) is the maximized log-likelihood of model \( l \), which estimates goodness of fit; \( d_{l} \) is the number of parameters in model \( l \), and \( n \) is the sample size. In the absence of other information, it is common to assume equal prior model probabilities \( p(M_{l}) \) for the candidate models [16,31]. Hence the BMA weights are approximately:

\[
w_{l} = \exp(-0.5 \ast BIC)
\]

The posterior probability for \( M_{l} \) is calculated as

\[
p(M_{l}|D) = \frac{w_{l}}{\sum_{l=1}^{L} w_{l}}
\]

Other information criterion can be used instead of the BIC. For example Akaike’s information criterion \( (AIC = -2\log\{p(D|\pi_{l}, M_{l})\} + 2d_{l}) \) [32] was suggested by Jackson et al. [31]. In the present study, candidate models were compared and combined using the deviance information criterion (DIC) [19]. The DIC is based on the posterior expectation of the deviance \( D \) and the effective number of parameters \( p_{D} \) in the model, and is expressed as:

\[
DIC = D_{P} + D
\]

Deviance is defined as \( -2\log\{p(D|\pi_{l}, M_{l})\} \) where \( p_{D} \) is the difference between the expected deviance and the deviance value for the posterior expectation. The DIC is easily computed from the samples generated through MCMC [23]. A smaller DIC value indicates a better model fit, accounting for model parsimony. In the BMA analysis, we let \( p(M_{l}) \) be 1/3, indicating no prior preference for any of the three correlation structures considered in this study.

**Bayesian Analysis and Spatial Interpolation**

Markov chain Monte Carlo (MCMC) was used to obtain distributions and corresponding posterior structures of means, standard deviations and quantiles for parameters of interest. Convergence was assessed by checking the trace and the autocorrelation plots for the sample of each chain [33]. For each model we ran a single MCMC chain for 150,000 iterations, discarding the first 50,000 iterations as burn-in. The MCMC
analysis was undertaken using WinBUGS software version 1.4 [34].

The posterior predictions of N2O obtained from the three models and hybrid model developed by BMA were mapped across the study site using GS+ software [35]. If input values are available across the study site, the model can be used to provide predictions between the experimental locations. In our case, these values were not available, so kriging was used for spatial interpolation of the predicted N2O values.

Results

Summary statistics for observed N2O and covariates were provided in Table 1. The overall means were 27.4 µg N2O-N m⁻² hr⁻¹, 19.0 Kg N ha⁻¹, 35.6%, 22.2°C and 5.5 for N2O, NO3⁻, soil moisture, soil temperature and soil pH under 17 sampling chambers, respectively. N2O tended to be more variable. For soil texture across 17 chambers, the overall means were 18.4%, 44.3% and 37.3%, with range 9.7 to 23.3, 34.4 to 60.9 and 22.8 to 50.9% for soil clay, soil silt and soil sand, respectively. Soil clay had lower percentages in soil texture.

The DIC values, measuring goodness of fit of each model, are shown in Table 2. The DIC was obviously smaller for the CAR and EXP models compared with the independent model, indicating the value of including spatial dependency in describing the N2O emissions in this dataset. The DIC values were similar for the CAR model and EXP model, indicating little difference in overall goodness of fit between the two representations of spatial variation. The results also showed that there were 2.7%, 1.8% and 1.4% of observed values that did not fall within the 95% posterior predictive intervals for the linear regression model, CAR and EXP model, respectively. The sum of the squared residuals from the geostatistical model was 317.12, while the sum of the squared residuals from the CAR model was 319.5.

Table 2 shows the posterior means and 95% credible intervals (CI) of parameters for the three models. Soil moisture and soil temperature had a substantive positive relationship with N2O emissions in all three models. Only the two spatial models showed a negative relationship between N2O emissions and NO3⁻ in the presence of the other variables in the model. Soil pH and soil texture, such as clay, silt and sand, were not substantial influential factors for N2O emissions in the three models in this study.

The spatial patterns of predicted N2O using the CAR and EXP models were similar to the observed spatial pattern, particularly the CAR model [Figure 1]. However, there were slight errors for classifications of areas into different emission level groups for the two spatial models, particularly the EXP model. The results of the independent model could not match the observed spatial distribution of N2O emission.

Figure 2 shows the distributions of the posterior means of the spatial variation in N2O emissions which were obtained using the two spatial models and BMA model. The three maps of posterior spatial variation show similar patterns. However, the CAR model displayed slightly larger areas for high or low spatial variation in N2O than those from the EXP model [Figure 2].

The CAR model had the highest posterior probability of 5.43E-1, whereas the independent model had a negligible probability of 0.000E-1. The EXP model had a posterior probability of 4.56E-1. The map of the averaged posterior predicted N2O emissions across the three structures was much more similar to the map of observed N2O emissions comparison to the maps of the CAR, EXP and independent models [Figure 1]. The map of averaged spatial predictions of N2O displayed better performance in high or low emission areas than that obtained from the EXP model and also it improved the accuracy of the spatial prediction of N2O on left side of the map compared with the CAR model. There were also slight changes in the map of the distribution of averaged posterior spatial variation [Figure 2].

Discussion

The CAR and EXP models are popular approaches for describing spatially correlated data and are widely used in many areas of scientific research. In this case study, we applied these two models and a baseline model that ignored the spatial correlation altogether. In order to gain some insight into the effects of different assumed spatial correlation structures on parameter estimation and spatial prediction of N2O emission for the same point data on an irregular grid, and to account for the uncertainty in evaluating spatial variability of N2O emissions using Bayesian model averaging.

All three models identified soil temperature and soil moisture as potentially important influential factors positively associated with N2O emissions. This is supported by a large body of previous research [5,36–42]. In this study, the average of soil moisture was around 36% in the pasture. Nitrification occurs when soil water-filled pore space is <60% [43]. Our result supported that increasing soil moisture and soil temperature increased N2O emissions via the nitrification pathway [36,39,44]. Only the CAR and EXP models yielded a significant negative coefficient for NO3⁻. The inverse relationships between N2O and NO3⁻ from nitrification have been found in grass-clover pasture and laboratory study [36,45]. The results showed that allowing for spatial dependence in the model affected not only the scale of posterior mean but also affected the determination of significant factors in the model for the pasture data. Moreover, the different spatial correlation structures in the models resulted in differences in the magnitude of the corresponding coefficients. Hence, the selection of an appropriate model structure is a critical step [9].

The sum of the squared residuals showed that the geostatistical model was slightly better than the CAR model. However, the plots of spatial interpolation of the predicted N2O values by kriging showed that the geostatistical model tended to oversmooth high N2O emission areas in comparison to the results of CAR model. We found that the predicted N2O values of the locations which were close to the highest emission site were underestimated by the geostatistical model in comparison to the CAR model and the observed data. The tendency of the EXP model to oversmooth is supported by Best [8]. On the other hand, mapping the spatial prediction of interest is often an important aim of developing a spatial model. Spatial interpolation is a straightforward approach

| Variables | Mean  | SD   | Minimum | Maximum |
|-----------|-------|------|---------|---------|
| N2O (µg N2O-N m⁻² hr⁻¹) | 27.2  | 39.4 | 0.0     | 280.4   |
| NO3⁻ (Kg N ha⁻¹) | 18.98 | 14.1 | 0.0     | 90.34   |
| Gravimetric soil moisture (%) | 35.57 | 9.27 | 9.65    | 70      |
| Soil temperature (°C) | 22.16 | 3.07 | 14.8    | 27.3    |
| pH | 5.47  | 0.29 | 5.2     | 6.4     |
| Sand (%) | 37.25 | 7.74 | 22.2    | 50.89   |
| Silt (%) | 44.34 | 7.2  | 34.44   | 60.94   |
| Clay (%) | 18.4  | 3.07 | 9.65    | 23.34   |

doi:10.1371/journal.pone.0065039.t001
Table 2. Posterior means and 95% credible intervals of parameters for three models for pasture.

| Parameter          | CAR model          | EXP model          | Independent model |
|--------------------|--------------------|--------------------|-------------------|
| $b_0$              | $-49.2 (-100.689)$ | $162.6 (-1728-2049)$ | $-5.84 (-4.62-30.2)$ |
| $b_{soil moisture}$| $0.055 (0.034-0.075)$ | $0.054 (0.034-0.075)$ | $0.039 (0.02-0.06)$ |
| $b_{soil temperature}$ | $0.16 (0.1-0.22)$ | $0.16 (0.1-0.22)$ | $0.15 (0.08-0.21)$ |
| $b_{NO3}^{-2}$     | $0.018 (-0.031-0.004)$ | $0.017 (-0.031-0.003)$ | $-0.006 (-0.02-0.008)$ |
| $\beta_{pH}$      | $0.4 (-0.73-1.55)$ | $0.32 (-1.19-1.84)$ | $0.21 (-0.42-0.91)$ |
| $\beta_{sand}$    | $0.45 (-0.11-1.0)$ | $-1.67 (-20.54-17.22)$ | $0.023 (-0.35-0.37)$ |
| $\beta_{silt}$    | $0.46 (-0.096-0.978)$ | $-1.66 (-20.52-17.24)$ | $0.033 (-0.34-0.38)$ |
| $\beta_{clay}$    | $0.4 (-0.16-0.94)$ | $-1.7 (-20.57-17.21)$ | $0.004 (-0.37-0.36)$ |
| $\sigma^2$        | $1.59 (1.31-1.92)$ | $1.59 (1.30-1.93)$ | $2.0 (1.66-2.43)$ |
| $\sigma_w^2, \sigma_s^2$ | $1.78 (0.46-4.81)$ | $0.76 (0.25-1.84)$ | $2.0 (1.66-2.43)$ |
| $DIC$              | 745.75             | 746.1              | 786.69            |

doi:10.1371/journal.pone.0065039.t002

Figure 1. Maps of observed and posterior mean Ln(N$_2$O) (ug N$_2$O-N m$^{-2}$ hr$^{-1}$) from the CAR, EXP, BMA and linear regression models across the study site in pasture.
doi:10.1371/journal.pone.0065039.g001
Spatial Prediction of N₂O Emissions

In this research, it is acknowledged that spatial analysis not only improves prediction but also highlights clustering and probabilistic uncertainties. The maps of the spatial distribution and the spatial variation in N₂O emissions may help to guide cultivation practices and determine emission reduction strategies. More attention should still be paid to how to select appropriate spatial correlation structures. This justified model averaging was a potentially useful method to take account of uncertainty of different spatial correlation structures and could improve the accuracy of spatial prediction of N₂O emissions. The best spatial correlation structure was unclear based on the DIC values associated with the CAR and EXP models in this study. The two spatial models showed differences in spatial prediction of N₂O distribution. This justified model averaging across the three structures via BMA. In this article, the posterior model probability was approximated by the DIC, which can be considered as a Bayesian analogue of the AIC suitable for hierarchical models with random effects [19].

Conclusions

Our study showed that soil temperature, soil moisture and NO₃⁻ were important influential factors in N₂O emissions in pasture across this study region. High emission areas were accompanied by high uncertainties after taking soil moisture, soil temperature, NO₃⁻, soil pH and soil texture into account. It was important to incorporate spatial dependency in the model when quantifying the relationship between N₂O emissions and environmental factors for this pasture. Allowing for spatial dependency in the model could yield accurate spatial prediction of N₂O. The CAR and EXP models yielded similar parameter estimates based on point-referenced data with irregular sampling intervals. The CAR model was better at capturing high N₂O emissions areas. The hybrid model structure obtained by BMA could improve the accuracy of mapping the spatial prediction of N₂O emissions in pasture across this study region. The maps of spatial distribution and spatial variation in N₂O can improve subsequent experiment design and investigate unknown influential environmental covariates in a study site.

Acknowledgments

We thank the Institute of Sustainable Resources (ISR), Queensland University of Technology (QUT) for assistance on data collection.

Author Contributions

Conceived and designed the experiments: XH KM PG. Performed the experiments: XH KM PG DR. Analyzed the data: XH KM WH. Contributed reagents/materials/analysis tools: XH KM PG WH DR. Wrote the paper: XH KM PG DR WH.
References

1. Bouwman AF (1998) Nitrogen oxides and tropical agriculture. Nature 392: 866–867.
2. Saggar S, Andrews RM, Tate KR, Hedley CB, Redda NJ, et al. (2004) Modelling nitrous oxide emissions from dairy-grazed pasture. Nutr Cycl Agroecosys 68: 243–253.
3. Bouwman AF, Vanderhoek KW, Olivier JGJ (1993) Uncertainties in the global source distribution of nitrous oxide. J Geophys Res Atmos 100: 2785–2800.
4. Bouwman AF (1996) Direct emission of nitrous oxide from agriculturals soils. Nutr Cycl Agroecosys 46: 53–70.
5. Dalal R, Wang W, Robertson G, Parton W (2003) Nitrous oxide emission from Australian agricultural lands and mitigation options: a review. Soil Res 41: 65–195.
6. Cressie N (1993) Statistics for Spatial Data. 2nd ed. New York: Wiley.
7. Keitt TH, Bjornstad OT, Dixon PH, Citron-Pousty S (2002) Accounting for spatial pattern when modelling organism-environment interactions. Ecography 25: 616–625.
8. Best N, Richardson S (2005) A comparison of Bayesian spatial models for disease mapping. Stat methods med res 14: 35–59.
9. Dormann CF, McPherson JM, Araújo MB, Bivand R, Bolliger J, et al. (2007) Methods to account for spatial autocorrelation in the analysis of species distributional data: a review. Ecography 30: 609–628.
10. Huang XD, Grace P, Mengersen K, Weier K (2011) Spatio-temporal variation in soil derived nitrous oxide emissions under sugarcane. Sci Total Environ 409: 4572–4578.
11. Kelsall J, Wakefield J (1999) Discussion of “Bayesian models for spatially correlated disease and exposure data”, by Best et al. In: Bernardos J, Berger J, Dawid A, Smith A, editors. Bayesian Statistics 6. Oxford: Oxford University Press.
12. Wall MM (2004) A close look at the spatial structure implied by the CAR and SAR models. J Stat Plan Infer 121: 311–324.
13. Goovaerts P, Gebreab S (2008) How does Poison Kriging compare to the popular BYM model for mapping disease risks? International journal of health geographics 7:6 doi:10.1186/1476-072X-7-6.
14. Hoeting JA, Madigan D, Raftery AE, Volinsky CT (1999) Bayesian Model Averaging A Tutorial. Stat Sci 14: 382–417.
15. Draper D (1995) Assessment and propagation of model uncertainty. J Roy. Statist. Soc. B 57: 45–97.
16. Boone EL, Bullock BP (2008) Spatial correlation matrix selection using Bayesian model averaging to characterize inter-tree competition inlobolly pine trees J Appl Stat 35: 967–977.
17. Diggle PJ, Tawn JA, Moyeed RA (1998) Model-based geostatistics. ApplStatist 47: 229–255.
18. Geldman A, Carlin JB, Stern HS, Rubin DB (2004) Bayesian data analysis. New York: Chapman and Hall/CRC.
19. Spiegelhalter D, Best NG, Carlin BP, van der Linde A (2002) Bayesian Measures of Model Complexity and Fit (with discussion). J R Stat Soc B Sta 64: 583–639.
20. Borgonovi E (2006) Measuring uncertainty importance: investigation and comparison of alternative approaches. Risk Analysis 26: 1349–1361.
21. Isbell RF (2002) The Australian soil classification. Collingwood: Csiro Press.
22. Rowlings DW, Grace PR, Kiese R, Weier KL (2012) Environmental factors controlling temporal and spatial variability in the soil-atmosphere exchange of CO2, CH4 and N2O from an Australian subtropical rainforest. Global Change Biology 18: 726–738.
23. Banerjee S, Carlin BP, Gelfand AE (2004) Hierarchical Modeling and Analysis for Spatial Data. USA: CHAPMAN&HALL/CRC.
24. Anselin L (2003) GeoDa 0.9 User’s Guide. Spatial Analysis Laboratory. Available: www.unc.edu/~emch/gislab/geoda93.pdf. Accessed 2011 Apr 18.
25. Wackerly D, Mendellhall III W, Scheaffer R (2002) Mathematical Statistics with Applications, Sixth Edition. USA: Duxbury.
26. Thomas A, Best N, Arnold R, Spiegelhalter D (2002) GeoBUGS User Manual Version 1.1. Available: wwwmrc-bsuc- macuk/bugs/winbugs/manual.pdf. Accessed 2011 May 2.
27. Thomas A, Best N, Lunn D, Arnold R, DS (2004) GeoBUGS User Manual Version 1.2. Available: wwwmrc- bsucmacuk/bugs/winbugs/geobugs12manualpdf. Accessed 2011 Apr 26.
28. Besag J, York J, Mollié A (1991) Bayesian image restoration, with two applications in spatial statistics. Ann I Stat Math 43: 1–59.
29. Schwarz G (1976) Estimating the dimension of a model. Ann Stat 6: 461–464.
30. Clyde M (2000) Model uncertainty and health effect studies for particulate matter. Environmentalmetrics 11: 745–763.
31. Jackson CH, Thompson SG, Sharples LD (2009) Accounting for uncertainty in health economic decision models by using model averaging. J R Stat Soc A Sta 172: 303–404.
32. Akaike H (1973) Information theory and an extension of the maximum likelihood. 2nd International Symposium on Information Theory. Budapest: Akademia Kiado. 267–281.
33. Gelman A, Rubin DB (1992) Inference from iterative simulation using multiple sequences. Stat Sci 7: 457–472.
34. Spiegelhalter D, Thomas A, Best N, Lunn D (2003) WinBUGS User Manual, Version 1.4. Available: wwwmrc-bsucmacuk/bugs/winbugs/manualpdf. Accessed 2010 Jun 12.
35. Robertson GP (2008) Geostatistics for the Environmental Sciences. Michigan USA: Gamma Design Software.
36. Goodroad LL, Keeney DR (1984) Nitrous oxide production in aerobic soils under varying pH, temperature and water content. Soil Biol Biochem 16: 39–43.
37. Williams E, Hutchinson G, Fehsenfeld F (1992) NOx and N2O emissions from soil. Global Biogeochem Cycles 6: 351–358.
38. Kamp T, Steindl H, Hanschel R, Biese F, Munch J (1998) Nitrous oxide emissions from a fallow and wheat field as affected by increased soil temperatures. Biol Fer Soils 27: 307–314.
39. Maag M, Vinther FP (1996) Nitrous oxide emission by nitrification and denitrification in different soil types and at different soil moisture contents and temperatures. Appl Soil Ecol 5: 5–14.
40. Yan Y, Sha L, Cao M, Zheng Z, Tang J, et al. (2008) Fluxes of CH4 and N2O from soil under a tropical seasonal rain forest in Xishuangbanna, Southwest China. J Environ Sci 20: 207–215.
41. Machlerlert S, Dye N, Goulding K, Whitehead P (2004) Nitrous oxide emissions from two riparian ecosystems: key controlling variables. Water, Air, Soil Poll 4: 427–436.
42. Huang X, Grace P, Rowlings D, Mengersen K (2013) A flexible Bayesian model for describing temporal variability of N2O emissions from an Australian pasture. Sci Total Environ 434: 206–210.
43. Davidson E (1991) Fluxes of nitrous oxide and nitric oxide from terrestrial ecosystems. In: Rogers J, Whitman W, editors. Microbial Production and Consumption of Greenhouse Gases: Methane, Nitrogen Oxides, and Halomethanes. Washington, DC: Am Soc Microbiol.
44. Huang X, Grace P, Weier K, Mengersen K (2012) Nitrous oxide emissions from subtropical horticultural soils: a time series analysis. Soil Research 50: 596–606.
45. Ambus P (2003) Relationship between gross nitrogen cycling and nitrous oxide emission in grass-clover pasture. Nutr Cycl Agroecosys 72: 189–199.