Spatial Statistics 18 (2016) 194–220

Contents lists available at ScienceDirect

Spatial Statistics

journal homepage: www.elsevier.com/locate/spasta

Non-Gaussian bivariate modelling with application to atmospheric trace-gas inversion

Andrew Zammit-Mangion [a,*], Noel Cressie [a], Anita L. Ganesan [b]

[a] National Institute for Applied Statistics Research Australia (NIASRA), School of Mathematics and Applied Statistics (SMAS), University of Wollongong, Northfields Avenue, Wollongong, NSW 2522, Australia

[b] School of Geographical Sciences, University of Bristol, University Road, Bristol, BS8 1SS, UK

ARTICLE INFO

Article history:
Received 19 January 2016
Accepted 13 June 2016
Available online 23 June 2016

Keywords:
Bivariate spatial model
Conditional multivariate model
Methane emissions
Multivariate geostatistics
Trans-Gaussian model
Box–Cox transformation

ABSTRACT

Atmospheric trace-gas inversion is the procedure by which the sources and sinks of a trace gas are identified from observations of its mole fraction at isolated locations in space and time. This is inherently a spatio-temporal bivariate inversion problem, since the mole-fraction field evolves in space and time and the flux is also spatio-temporally distributed. Further, the bivariate model is likely to be non-Gaussian since the flux field is rarely Gaussian. Here, we use conditioning to construct a non-Gaussian bivariate model, and we describe some of its properties through auto- and cross-cumulant functions. A bivariate non-Gaussian, specifically trans-Gaussian, model is then achieved through the use of Box–Cox transformations, and we facilitate Bayesian inference by approximating the likelihood in a hierarchical framework. Trace-gas inversion, especially at high spatial resolution, is frequently highly sensitive to prior specification. Therefore, unlike conventional approaches, we assimilate trace-gas inventory information with the observational data at the parameter layer, thus shifting prior sensitivity from the inventory itself to its spatial characteristics (e.g., its spatial length scale). We demonstrate the approach in controlled-experiment studies of methane inversion, using fluxes extracted from inventories of the UK and Ireland and of Northern Australia.

* Corresponding author.
E-mail addresses: azm@uow.edu.au (A. Zammit-Mangion), ncressie@uow.edu.au (N. Cressie), anita.ganesan@bristol.ac.uk (A.L. Ganesan).

http://dx.doi.org/10.1016/j.spasta.2016.06.005
2211-6753/

1. Introduction

Atmospheric trace-gas inversion is the procedure by which flux fields (gas sinks and sources) are identified from gas mole-fraction observations. Unlike conventional problems in spatial statistics, the spatial field of principal interest, the flux field, is rarely directly observed. Instead, satellite and surface gas-concentration instruments are sensitive to gas particles after they have been transported over potentially large distances over time. Spatio-temporal statistical methodology is well placed to obtain global space–time maps of mole fractions of an atmospheric constituent from irregular observations (see, for example, Cressie et al., 2010; Cameletti et al., 2013; Lindgren et al., 2011); however, the inversion of these maps in order to pinpoint the sources and sinks of the gas is a much more difficult problem. Its solution is key to effective policy implementation with regard to greenhouse gas emissions and climate change (Edenhofer et al., 2014).

Current approaches to trace-gas inversion build on the data-assimilation framework described, for instance, in Tarantola (2005), where prior beliefs on a flux field are updated with observations, to produce posterior beliefs (see, for example, Rigby et al., 2011; Stohl et al., 2009). The prior distribution is usually formulated to have prior expectation equal to values in an *inventory*, a flux database constructed from auxiliary activity data (e.g., vehicles per unit area) and emission factors associated with the emission sector (e.g., carbon dioxide emissions per vehicle), while the prior covariance is usually constructed using values calculated from the inventory. Reasonable prior marginal variances are typically assumed; for example, a prior density function at each location may be chosen such that the area under the curve between 0.5 and 1.5 times the value of the inventory at that location is around 68% (Ganesan et al., 2014). Frequently, a diagonal structure is imposed on the prior covariance for all locations. It is also usually assumed that the prior expectation and prior covariance completely specify the prior distribution.

These current approaches can be critiqued due to their use of inappropriate or inflexible models. First, atmospheric trace-gas inversion is inherently a bivariate spatio-temporal problem, where observations are not readings of fluxes, but rather readings of a second, mole-fraction, field. The mole-fraction field is generated by the underlying flux field and by meteorology, which determines transport of the trace gas. Making this distinction allows one to attribute uncertainties appropriately, either to instrumentation error or to imprecise mole-fraction field modelling (e.g., due to linearisation of the flux–mole-fraction mapping or imprecise specification of transport modelling/boundary conditions). Second, a Gaussian model is often inappropriate for the flux field (Ganesan et al., 2014; Miller et al., 2014), which can be inherently non-negative and which, as seen from inventories, may exhibit skewness and higher-order features. Third, spatial correlations in prior error covariances, even when using regressors, should be assumed since errors in the inventories are likely to be spatially correlated. Fourth, several works claim that inventories are highly inaccurate in certain regions (e.g., Lunt et al., 2015). Consequently, there is a drive to divert from use of the inventories in the model, and to take a data-driven approach to flux inversion, even when sub-national resolution is required and so prior information is especially important.

These reservations were first discussed in Zammit-Mangion et al. (2015), where the authors constructed a hierarchical, lognormal bivariate spatio-temporal model for flux inversion of methane in the UK and Ireland. In that article, an empirical hierarchical modelling approach (Cressie and Wikle, 2011, Section 2.1) was adopted, where numerous parameters were estimated offline prior to drawing samples from the empirical posterior distribution of the flux field. In this article, we extend the modelling and inferential approach in Zammit-Mangion et al. (2015) in two ways. First, we relax the assumption of lognormality by defining a more general non-Gaussian model, and subsequently we model the flux field as a trans-Gaussian (here a Box–Cox) spatial process (De Oliveira et al., 1997). The class of Box–Cox spatial processes includes the lognormal spatial process as a special case. Second, following a likelihood approximation, we adopt a fully Bayesian approach to flux-field inversion that naturally propagates variability in the parameters (parameters that were estimated offline in Zammit-Mangion et al., 2015) to our inferences on the flux field. In order to reduce reliance on the flux inventory, which is based on emission factors that are not precisely known, we do not assume it to be the prior mean of the flux field. Instead, we take a new approach by assuming that it is an independent realisation of the flux field that we wish to infer. Hence, we assume that the inventory