**Atmospheric
Measurement
Techniques**
Discussions

Interactive comment on "Spatial mapping of ground-based observations of total ozone" *by* K.-L. Chang et al.

Anonymous Referee #2

Received and published: 19 June 2015

The paper aims to evaluate the performance of two different approaches to spatial estimation of global ozone concentrations. As already noted by one reviewer, the comparison however doesn't make it clear what is actually being compared, as the two main approaches used differ not only in model structure, but also in the techniques used for parameter estimation. Furthermore, there is a pervasive confusion about the difference between *models*, *estimation concepts*, and *computational methods*. The SPDE and chordal distance covariance models used in the paper are *models*, whereas *kriging* is an estimation concept, with associated *computation methods* that only differ in the details between the two models; the SPDE calculations use precision matrices for basis expantion weights, and the covariance calculation use more traditional covariance matrix expressions. However, for sufficiently high resolution for the finite element

construction for the SPDE model, and for covariance models equal to the covariance of solutions to the SPDE, the results should differ only in numerical details, as the calculations fundamentally target the same *kriging* estimate of the spatial field. This mixup between concepts unfortunately runs through the entire paper, with *SPDE vs kriging* instead of the perhaps more appropriate *SPDE vs chordal covariance model*, and the use of different statistical estimation techniques for the two models (Bayesian inference vs cross-validation) makes it very difficult to tell what the comparison results actually demonstrate.

Specific comments

1. p3970, l16: The sentence *Statistical models assumes that the unknown function is a realisation of a Gaussian random spatial process.* is incorrect. There are plenty of non-Gaussian spatial models, including discrete valued Markov models, point process models, and transformed Gaussian models to mention just a few. The sentence following it also confuses the model with the method; a "mean field" is naturally estimated as a part of the estimation process if it is part of the statistical model; kriging is just a word used for optimal least squares estimation of the spacial process, and in the completely Gaussian case equivalent to the conditional expectation of the field given the data.

2. p3970, l20. Again, there is confusion between methods and models. Kriging itself has no problem with non-stationarity. That is all down to the *model*. It's true that it's difficult to construct general non-stationary covariance functions, but that is largely unrelated to the kriging method as such.

3. p3970, l24. The year for Lindgren, Rue, and Lindström is 2011, not 2010. (This is an error in the reference list on page 3988, so correcting it there would likely fix all the references to that paper.)

4. p3972, Section 2.2. Here, there is a false opposite between the SPDE approach and spatial kriging. The Gaussian process in the spatial model, $Z(s)$, can be defined as a realisation of the process $X(s)$ from Section 2.1. The model for the observations exists independently of the kriging *method*, which can be applied to any model of this type; the practical details are in how the kriging estimate is *computed*, not in the model structure statement itself.

5. p3973, l13. *"exactly the same covariance function"* appears to contradict other statements in the paper about which models were used. For example, if the SPDE models used $\alpha = 2$, then the smoothness on the sphere (a 2-manifold) is $\nu = 1$ (so the statement on page 3973, line 15, is incorrect). But on page 3974, line 8, it's stated that the covariance based calculations were done for a model with $\nu = 20$, which is virtually indistingusishable from a squared exponential (or Gaussian) covariance, and very different from $\nu = 1$.

6. p3973, line 15. $\nu = 0.5$ should most likely be $\nu = 1$ (see previous comment). The dimension d in the relation $\alpha = \nu + d/2$ is the dimension of $\mathbb{R}^d$ for a regular Maérn model, and the relevant dimension when solving the SPDE on the 2-manifold that is the surface of the globe is $d = 2$, not 3. Further, in this paper, ν may have been fixed to 1 for the SPDE models, but the general SPDE/GMRF models have no such restriction.

7. p3973, l 20. *model set-ups for both SPDE and kriging*; again, confused comparison. An actual name for the covariance specified model is needed, as kriging is used for *both* that model and for the model based on an SPDE.

8. p3973, l25 to p3974, l1. *spherical harmonics for the expansions of κ and τ [...] (the default choice in the R-INLA package).*
The R-INLA package does not have any default basis functions for non-stationary models (it does have helper functions for some commonly used basis functions that the user can choose to use).

9. p3974, l15-20. Here, it's unclear what is really being compared. How were the parameters in the chordal distance covariance model estimated? Are the differences due to differences in statistical estimation techniques, e.g. with some models estimated with cross validation and others estimated with Bayesian inference? If a) the models are fundamentally different (different ν) and b) estimated with different methods, one shouldn't expect the results to be comparable, as one combination of model and method is likely to be better than others (Even if none of them perfectly matches the data). How were the priors chosen for the weights for the basis functions for $\log \kappa$ and $\log \tau$ in the SPDE model?

10. p3977, l6-8. Again, *kriging* here should really be replaced by a name for the covariance based model! Further, is it possible that the unstable predictions are due to the ν-estimation? We are not really in the infill asymptotic domain here, so reliable estimation of ν is very difficult.

11. p3977, l8. *variations.* should be *variation.*

12. p3980, l2. The passage *underestimation and disappear of estimated annual cycle* is clearly incorrect, but I'm not sure what it is supposed to say. Perhaps *disappearance of the* was intended, but that would still leave a strange statement.

13. p3981, l7. The statement *SPDE approach is more robust than kriging against incomplete information* is surprising, even when replacing the word *kriging* with *the chordal distance covariance model*. The non-stationary SPDE model has many more parameters than the explicit covariance model. Is the result that that model is more adaptive to the data? Unfortunately, since the models also apear to differ greatly in terms of smoothness (ν) and were estimated using very different estimation techniques, I don't think one can draw any strong conclusions about *which* of those differences are more or less important to the results.

14. p3983, l20. Why is this model completely different to the model stated in Section

2.2? Did the results using SPDE models not include any fixed/mean effects in the modelling? That could further explain some of the differences in the results.

15. p3984, l16-17. It is true that identifying both the mean and covariance is impossible based on a single realisasion, without further modelling assumptions. However, that is completely unrelated to the statements about the $\kappa(s)$ and $\tau(s)$ functions. Perhaps a section completing the sentence fragment *To avoid this identifiability problem.* is missing from the manuscript?

16. p3985, l15-17. In the sententence *The main limitation is that R-INLA provides* $0 < \alpha \leq 2$ *case (though* $0 < \alpha < 2$ *not as extensively tested).* there is a *the* missing after *provides*, and *case* should be *cases*. The statement in brackets appears to be a direct quote from Lindgren and Rue (2015), *Bayesian Spatial Modelling with R-INLA*, Journal of Statistical Software, 63(19) (http://www.jstatsoft.org/v63/i19), but that paper has not been referenced (it's quite possible that an unpublished draft of that paper was used when preparing this paper).

17. p3985, l17-18. See the earlier comment about the fact that the sphere is a 2-manifold, so the relevant dimension for the smoothness relationship is 2, not 3.

Interactive comment on Atmos. Meas. Tech. Discuss., 8, 3967, 2015.