

# Analyzing time varying trends in stratospheric ozone time series using state space approach by M. Laine et al.

## Answers to referee comments, Referee #1

We thank both anonymous Referees for constructive commentary on our manuscript. We hope that we have successfully addressed all the relevant comments and suggestions.

As both Referees were inclined towards a major revision of the manuscript, and many of the concerns were about the presentation, we are writing the answers with the new revision in mind. First, we give a general overview of the changes. Then, we detail answers to the specific questions by the Referee.

First of all, we have simplified the text with respect to computational details on estimating the model parameters. We removed some of the technical details related to statistical computations from Section 2, especially subsections 2.3 and 2.4 for model parameter estimation and Markov chain Monte Carlo methods. A short description of the statistical calculation needed for DLM analysis is given in new Appendix A. Also, we moved the first five paragraphs of Section 3 to Section 2 as suggested by Reviewer #1. For the computational details, and for those interested in using or implementing the method we provide the full Matlab source code. We have expanded the web page that contains the computer code. We have included a dynamic linear model tutorial page that provides more details about the statistical theory and computation. This page or a pdf version of it could be included as supplementary material, also (<http://helios.fmi.fi/~lainema/dlm/dlmtut.html>).

The work on this paper has been done concurrently with the companion paper [Kyrölä], and some of the modelling decisions made for the companion paper have been reflected on the DLM approach. In particular, we decided to include autoregressive (AR) residual autocorrelation component in the model, also. Again, we see improvement in the dynamic linear model (DLM) approach over the multiple linear regression (MLR) approach, where an iterative procedure must be used and the uncertainty related to plug-in parameter values is not considered. In DLM, the AR coefficients and the innovation variance can be both estimated and accounted together with all the other model unknowns.

We have included ENSO proxy series by using the MEI index from NOAA, as suggested by Referee #1. We now utilise four proxy variables: Solar 10.7 cm radio flux, QBO 30 mb and 50 mb zonal wind indices, and MEI. For simplicity and for overall consistency, all the proxies are used in every fit even though not all of them are known to affect the ozone variability in all geographical regions.

Also, there has been some re-processing of the data compared to that used in the first version of the manuscript. The same data sets that are used in [Kyrölä] are used here.

After some thoughts, we decided to alter the mathematical notation in such a way that we use variable  $x$  for model state, instead of  $\theta$ . This might make the exposition more familiar to people working in geophysics. As a result of the shortening, some of the statistical notations were not needed any more.

The title with a spelling change and an added definite article is now: *Analysing time varying trends in stratospheric ozone time series using the state space approach*.

## Referee #1

The Referee has given many suggestions on improving the discussion paper. We hope that we have successfully addressed all the concerns below.

### General comments

The term “state space” is vague in my opinion.

The terminology used for dynamic regression varies in the literature as there are several fields that have used similar methods. Historically Kalman filtering and Bayesian analysis of dynamic linear models were developed separately although they are mathematically equivalent [see e.g. Särkkä]. Traditionally, statistical literature has used the term dynamic regression. However, [Durbin] which was our primary reference for implementation, talks about “state space methods”. We originally had the words "dynamic linear model" instead of "state space approach" in the title. Our motivation for the latter was that the state space approach is closely related to data assimilation and Kalman filter estimation of model states, which is more familiar to many researchers in geophysics. Also, the state space terminology emphasizes the representation of the process generating the observations as hidden model states. We particularly like these words by Referee #2:

By separating between an underlying “true” state and the projection of that state into observations with limited accuracy, and by assuming certain properties for the evolution of the underlying state, a different access to things like long-term variations becomes possible.

Back to Referee #1:

This paper seems to shy away from any results that are in contrast with its companion paper.

In our revision, we have tried to be more concrete and spelled out the findings more clearly. We have added a plot in the tropics at 20 km and 34 km for the individual altitudes, not the averaged altitudes used in the collected trend results plots. These plots show the difference of MLR results to DLM, and also the effect of the proxy variables. Related to effect of proxy variables, we have added a new Figure 4 showing relative contribution of different model components to the total variability on ozone. We now state more clearly the different estimated turnaround points.

The paper seems to be more of a ‘proof of concept’ with ad hoc choices of 0.06% of the mean level of the observations and a standard deviation of 80% (log-normal distribution).

For this revision, we performed sensitivity experiments with different prior specifications. In the original version of the manuscript, the priors were somewhat narrower. In the current version, we increased prior standard deviations in such a way that the priors are now relatively “weak” and non-informative, i.e. they allow the posterior distribution mostly depend on the observations and not on the priors. As the variance parameters to be estimated are positive, it is convenient to estimate them in logarithmic scale, having Gaussian prior for the logarithm and thus log-normal prior for the original value. The transformation of parameter does not affect the theoretical posterior obtained by MCMC sampling. If  $X$  is distributed as log-normal  $\log N(\mu, \sigma^2)$ , then  $\log(X)$  is Gaussian  $N(\mu, \sigma^2)$ . The mean of a log-normal distribution is approximately  $\log(\mu)$  and the parameter  $\sigma^2$  gives the order of magnitude in the deviation. For example standard deviation of  $\log N(0, 2^2)$  distribution is as large as 2926. The following table shows the prior distributions used in the final runs:

<i>estimated parameter</i>	<i>prior parameter log(<math>\mu</math>)</i>	<i>prior parameter <math>\sigma</math></i>
trend standard deviation, $\sigma_{\text{trend}}$	1/12 % of O <sub>3</sub> mean	1
seasonal standard deviation, $\sigma_{\text{seas}}$	1% of O <sub>3</sub> variability	2
autoregressive standard deviation, $\sigma_{\text{AR}}$	30% of O <sub>3</sub> variability	2
autoregressive parameter $\rho$	prior is Gaussian $N(0.45, 0.5^2)$ , truncated between [0,1]	

The process of building a statistical model is always a combination of explicit background / prior information, and trial and error. We hope that we have now explained the reasoning behind the choices done. In short, we wanted to extend MLR model so that the change points and other changes in the trend would not be prescribed by the model, but estimated from the observations. With DLM, we would also justify the use of simpler models, when appropriate, and see when they are not appropriate. There is probably no single method that automatically solves all trend analysis problems, although we feel that DLM is a good step toward such a method. Instead of allowing only one change point in the trend, we allow it to change smoothly if the data supports it. As seen from the residuals, our (still relatively simple) DLM model is well consistent with the data. In most of the cases, model residuals (defined by Kalman filter one step ahead prediction) are uncorrelated with unit variance. In some altitudes, a model with three seasonal components, instead of two, might have had better properties, but we decided to use a model that is as close to the original MLR model as possible for comparison purposes.

The inclusion of a solar term without the inclusion of more important proxies at 20 km could possibly lead to incorrect trends, partly because of the long period of a solar cycle.

For this manuscript, we included the ENSO proxy as fourth proxy variable. In both the companion paper [Kyrölä] and in the current manuscript, the coefficient for solar contribution was fixed for the whole period.

### Specific comments and questions:

These comment have all been taken into account. The ones that are in form of a direct question are answered below.

About the model residuals: Figure 1 has a residual Gaussian quantile-quantile plot with scaled model residuals (so called prediction residuals) on the y axis. The other panel has the estimated autocorrelation function. In this case the residuals are almost perfectly uncorrelated Gaussian with unit variance. So the answer is yes, residuals are Gaussian with unit variance.

We have used the word turning point rather informally to specify when the slope changes significantly. This is not necessarily the point where the slope turns positive. For example, in Figure 1 the instantaneous slope is positive only briefly near the year 2000 but there seems to be a turning point already at 1996. The same is also seen, if one studies the average decadal trend. We have deliberately kept from giving very accurate qualitative figures about change points, as one of our results is that they can depend on modelling choices, and thus would need more elaborate analysis. We have changed the text to better reflect the actual findings.

We decided to remove the RMSE, root mean standard error statistic from the plots. They were calculated as the scaled Kalman filter prediction residual standard error, with a value close to one suggesting a good fit with respect to the observation error and model error variances.

## Editorial comments

Again, we thank the Referee for careful reading of our manuscript and for the numerous editorial corrections. About the meaning of notations like  $p(y_t|\theta_t)$  vs.  $p(\theta_t|y_t)$  [in the new notation, these would be  $p(y_t|x_t)$  vs.  $p(x_t|y_t)$ ]. The first is the probability density function of observations at time  $t$  given (i.e. knowing) the values of model states at time  $t$ . It is also known as the likelihood function. The latter is the marginal posterior probability distributions of model states given the observations at time  $t$ . In our probabilistic settings, we have used terms probability density, probability distribution, and uncertainty distribution interchangeably. Now that we decided to go away with some of the computational and statistical details in the main text, the use of these notations was not strictly necessary, and we have replaced them with spelled out versions. The notation is used, and to some extent explained, in Appendix A and in the supplementary web page.

## References

[Durbin] Durbin, T. and Koopman, S.: Time Series Analysis by State Space Methods, Oxford Statistical Science Series, Oxford University Press, second edn., 2012.

[Kyrölä] E. Kyrölä et al.: Combined SAGE II-GOMOS ozone profile data set for 1984–2011 and trend analysis of the vertical distribution of ozone, Atmospheric Chemistry and Physics, 13, 10645–10658, <http://www.atmos-chem-phys.net/13/10645/2013/>, 2013.

[Särkkä] S. Särkkä: Bayesian Filtering and Smoothing. Cambridge University Press, 2013.