GLOBAL MAPPING OF OZONE DISTRIBUTION USING SOFT INFORMATION AND DATA FROM TOMS AND SBUV INSTRUMENTS ON THE NIMBUS 7 SATELLITE

George Christakos, Alexander Kolovos, Marc L. Serre, and Chandra Abhishek
Center for the Advanced Study of the Environment (CASE), University of North Carolina-Chapel Hill, NC

F. Vukovich
Science Applications International Corp. (SAIC), Raleigh, NC

Abstract: Modern spatiotemporal geostatistics can be used to efficiently assimilate salient and of varying uncertainty physical knowledge bases about atmospheric ozone, in order to generate and update realistic pictures of ozone distribution across space and time. We use the BME method which manages to eschew the restrictive assumptions of competitive techniques (linearity, normality, physical model-free, overparameterization, etc.). In addition, BME assimilates uncertain measurements and secondary (soft) information in terms of total ozone-tropopause pressure empirical equations, thus producing accurate predictions of the ozone values at unsampled locations in space. By analyzing and processing data sets generated by different measuring instruments on board the Nimbus 7 satellite, the BME-generated composite space/time maps are more informative and accurate than those obtained by traditional data analysis techniques.

Key words: Ozone, TOMS, SBUV, BME, modern spatiotemporal Geostatistics.

1. INTRODUCTION

Ozone (O_3) is a very reactive gas present in the stratosphere (90% of the total atmospheric O_3) and the troposphere. The ozone distribution shows a considerable variability across space and time, with a global average of 300 DU (corresponding to a 3mm column of O_3 at standard temperature and surface pressure). In addition its natural variability, the different levels of accuracy of the algorithms used to generate data from the measuring instruments introduce major sources of uncertainty in spatiotemporal O_3 modelling. It is therefore very important to obtain a realistic picture of O_3 distribution across space and time [1]. This work demonstrates advanced spatiotemporal modelling techniques that integrate the various knowledge bases available about O_3 (data collected at sparse SBUV measurement points, uncertain evidence, and secondary physical information) to predict the distribution of atmospheric O_3 concentration values at unsampled locations in a mathematically rigorous and physically meaningful manner.

2. OZONE MEASUREMENT INSTRUMENTS

In this work we study the distribution of total ozone values across space. The term "total column ozone" refers to the amount of O_3 in a column of air of unit area from the surface to the top of the atmosphere, and is estimated by measuring backscattered radiances of incoming solar radiation at wavelengths between 312 and 340 nm, as ahown in [2] and [3]. Two main groups of instruments are used to measure O_3 concentrations, as follows [4]: (a) Total ozone mapping spectrometers ($TOMS$), which are instruments generating measurements of the total column ozone in the atmosphere at different angles sideways from the path the satellite. Using $TOMS$ instruments, total O_3 maps can be produced once a day. (b) The solar backscatter ultraviolet ($SBUV$) instruments, which measure the ozone column separately in each of 12 superimposed atmospheric layers ([3] and [4]). Using $SBUV$ instruments a global ozone map can be generated within 7 days, approximately. Both the $TOMS$ and the $SBUV$ instruments were on board the Nimbus 7 Spacecraft and the relevant datasets cover several years. However, there are some concerns that the $SBUV$-based data may be less accurate than the $TOMS$-based data. The present study used data obtained by the instruments onboard the Nimbus-7 satellite. For illustration , the locations of the $TOMS$ measurements of total O_3 obtained on July 6, 1988 are indicated in Fig. 1 (small crosses). The locations of the $SBUV$ measurements during the same day are also shown in Fig. 1 (circles).

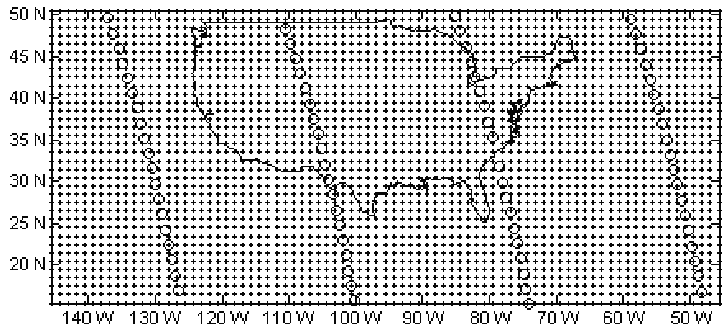

Figure -1. Grid coverage of satellite ozone measurements on July 6, 1988, for the *TOMS* (plus markers) and *SBUV* (circles) instruments.

3. SPATIOTEMPORAL MODELLING AND MAPPING OF OZONE DISTRIBUTION

3.1 A Review of Modern Spatiotemporal Geostatistics

Modern Spatiotemporal Geostatistics [5] provides a powerful framework for generation of informative maps of natural processes across space and time by accounting for general knowledge to define a space of plausible events and then restricting this space to be consistent with available site-specific knowledge. A spatiotemporal random field (S/TRF) $X(\boldsymbol{p})$ is used for a mathematically rigorous and physically meaningful representation of the distribution of O_3 concentrations across space and time ([6], [7]). The vector $\boldsymbol{p}=(\boldsymbol{s},t)$ defines a point in the space $\boldsymbol{s}$ and time t domain. Given certain general knowledge about the entire O_3 field and a set of site-specific data $c_{\text{data}}(c_1,..,c_m)$ at points $\boldsymbol{p}_{\text{data}}=(\boldsymbol{p}_1,...,\boldsymbol{p}_m)$, the O_3 studies are generally concerned with the estimation of the spatiotemporal O_3 distribution at a network of points $\boldsymbol{p}_{\boldsymbol{k}}=(\boldsymbol{p}_{k_1},...,\boldsymbol{p}_{k_\ell})$.

In the context of *MSG* one seeks to derive the probability density functions (PDF) $f_{KB}(c_{\boldsymbol{k}})$ that characterizes $X(\boldsymbol{p})$ at every point of the mapping grid in light of the physical knowledge sources considered. The principle of maximum expected information is applied upon general knowledge bases (denoted by G-KB) such as physical laws, governing relationships, primitive equations, and space/time statistical moments. This results in an intermediate corresponding *PDF* f_G, which is then conditioned with the available specificatory KB (S -KB). We adopt the widely used *BME* (Bayesian Maximum Entropy) technique of *MSG*, which employs the Bayesian conditionalization rule (bc) to yield updated (also called integration or posterior) *PDF*s f_K^{bc} that are consistent with the S –KB available, where $K = G \cup S$. The O_3 estimates $\hat{c}_{\boldsymbol{k}} = (\hat{c}_{k_1}, \ldots, \hat{c}_{k_\ell})$ at any set of grid points $\boldsymbol{p_k}$ are derived from the *PDF* at $\boldsymbol{p_k}$ by means of a suitable and flexible criterion, depending on the study goals (e.g., the most probable O_3 estimates, optimization of some cost function, etc.). In the following, we consider a *BME*-based numerical experiment presented in sections 3.3-3.6, and we focus our attention on the subregion of Fig. 1 shown in Fig. 2.

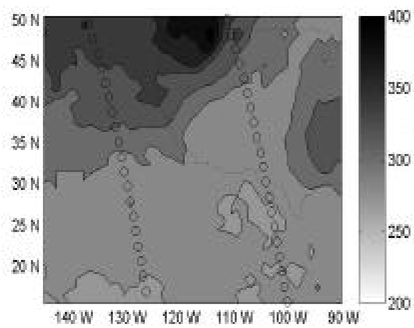

Figure –2 . Actual distribution of total ozone (in *DU*) obtained from the *TOMS* instrument on July 6, 1988.

3.2 Spatial Correlation of Total Ozone Distribution

Using the entire *TOMS* data set, in Figure 2 we show the actual map of total ozone, TO_3, for the western part of US as a reference. In line with the *S/TRF* representation mentioned previously, the TO_3 distribution is modeled by the spatial random field

$$TO_3(\boldsymbol{s}) = \overline{TO_3}(\boldsymbol{s}) + X(\boldsymbol{s}), \tag{1}$$

where $\boldsymbol{s}$ is a spatial location vector, $\overline{TO_3}(\boldsymbol{s})$ is the spatial trend of TO_3, and $X(\boldsymbol{s})$ is a zero mean spatially homogeneous random field of ozone fluctuations across space. Given the TO_3, the $\overline{TO_3}(\boldsymbol{s})$ is extracted from the data with an exponential filter (see *BMElib* in [8]), and the residual ozone

$X(\boldsymbol{s})$ hard data value is calculated from Eq. (1). The covariance of $X(\boldsymbol{s})$ is modelled by the following equation, which is assimilated in the G-KB of BME:

$$c_x(r_{ij}) = c_1 e^{-3r_{ij}/a_1} + c_2 e^{-3r_{ij}^2/a_2^2}, \tag{2}$$

where $r_{ij} = |\boldsymbol{s}_i - \boldsymbol{s}_j|$ is the spatial distance between any pair of locations in the atmosphere. Model (2) is fitted to the experimental covariance values obtained from c_{hard}, so that $c_1 = 75$ (DU^2), $a_1 = 15$ (degrees), $c_2 = 75$ (DU^2), $a_2 = 9$ (degrees). Each component of Eq. (12) accounts for half of the total variance of 150 (DU^2).

3.3 Generating Soft Information for Atmospheric Ozone

The atmospheric pressure at a given height is given by (see, e.g., [9])

$$P = P_0\, e^{-H/H_0}, \tag{3}$$

where H is the height above the surface and H_0 is called the scale height of the atmosphere (approximately 7 Km or 4.3 miles). The tropopause height H_t is monitored by collecting the pressure P_t at H_t. P_t-files are usually obtained using observations and a model mapping global distribution. In the present analysis, the necessary P_t-files were provided by the Langley Research Center. For each value of P_t we derive a soft *PDF* representing the probabilistic distribution of the total ozone TO_3 values, which provides the physical basis for producing the soft information to be used by *BME* (see Approach 2 in Section 3.5 below), based on the relationship

$$TO_3 = a_0 + a_1 \log P_t. \tag{4}$$

In the above, $a_0 = TO_{3,0} + aH_{t,0} - a\,H_0 \log P_0$ and $a_1 = a\,H_0$ can be estimated by experimental data fitting. We have considered the zero subscript parameters to correspond to some initial state values. The a_0 and a_1 are viewed as random variables representing uncertainty sources.

For numerical illustration, Fig. 3 depicts a typical scatter plot of TO_3 vs. P_t experimental values at concurrent points (shown with plus markers).

The general behavior of the physical TO_3 - P_t relationship (4) is represented well by the dotted line, for which a_0 and a_1 have been best fit to the experimental data. However, due to the stochastic nature of Eq. (4), to each P_t-value corresponds an uncertain TO_3-value. In order to obtain the probabilistic (soft) information representing the TO_3 uncertainty, we first divide the data into classes of contiguous non overlapping P_t-intervals. Then, for each class of P_t-values we derive the experimental mean and variance of the corresponding TO_3-values, as well as theis *PDF*s. Some of these *PDF* associated with three selected P_t-classes are plotted in Fig. 3, for illustration. Using this procedure, we can assign a TO_3 probability datum to each P_t data point, thus representing the uncertainty in the TO_3-values.

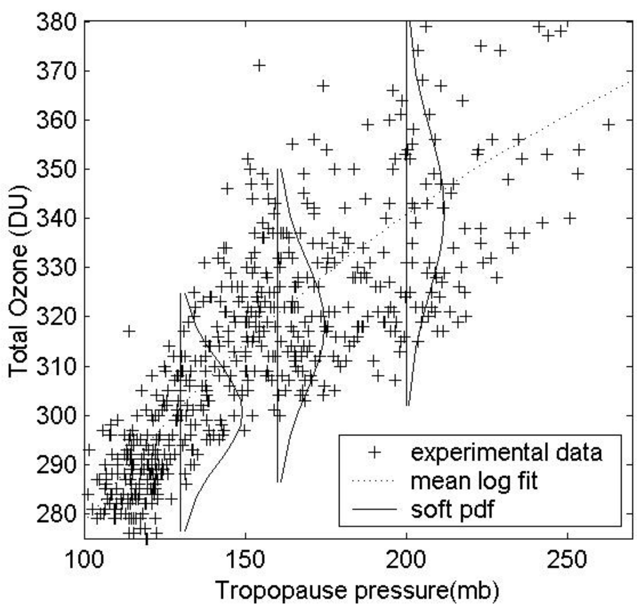

Figure -3. Scatter plot of total ozone measurements vs. tropopause pressure. A physical equation is fitted to the data from which soft PDF can be derived.

3.4 Approach 1

In the context of Approach 1 we assumed that the site-specific KB, S, consists solely of the hard TO_3 data set at the *SBUV* measurement points (circles in Fig. 1). In this case, a spatial regression-based technique (also known as simple spatial kriging) can be used to estimate TO_3 in the remaining area (i.e., at all points shown with small crosses in Fig. 1). It is noteworthy that the simple kriging technique can be derived as a limiting case of the general *BME* approach under the S-KB restrictive conditions described above. The corresponding TO_3 map is shown in Fig. 4a. A comparison with the actual map of Fig. 2 shows poor estimation accuracy away from the hard data points (circles). Additionally, there is a discontinuity in the distribution of the estimated TO_3 values along the axis inbetween the satellite paths (Fig. 4a). This is rather an artifact of Approach 1 not referring to a realistic scenario. The TO_3 estimation error standard deviations (s_e) for Approach 1 is given by

$$s_e(\mathbf{s}_k) = [c_x(0) - \textstyle\sum_i^M l_i c_x(r_{ik})]^{1/2}, \qquad (5)$$

where M is the number of TO_3 data used in estimating the TO_3 value at the grid point $\mathbf{s}_k$, and l_i are the estimation weights calculated from the kriging system ([10]). The s_e-map associated with the TO_3 map of Fig 4a is plotted in Fig. 4b.

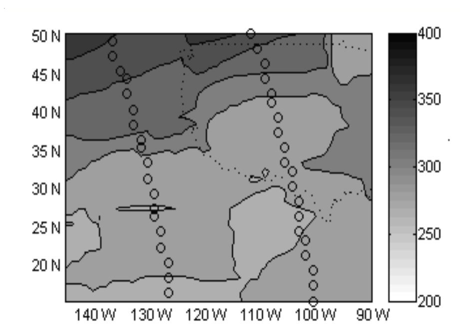

Figure -4a.. Kriging map of total ozone estimates (in *DU*) using only hard data at the *SBUV* points (shown in circles).

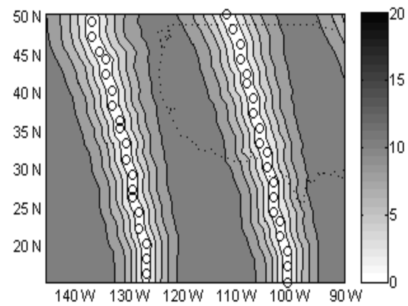

Figure -4b. The associated map of estimation error standard deviations for Fig. 4a.

3.5 Approach 2

In Approach 2 we use site-specific soft information in addition to the hard data set. As was discussed above (Section 3.3), the soft information in the form of local PDFs is derived from a physical equation relating TO_3 and P_t measurements obtained independently from the *TOMS* data. In this case, the resulting *BMEmean* map is plotted in Fig 5a.

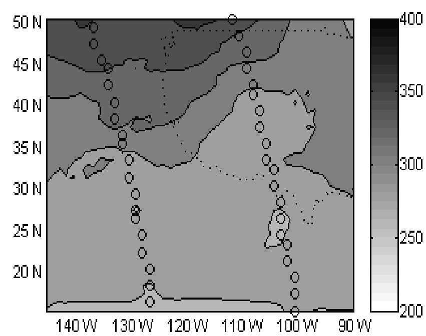

Figure –5a. Map of the *BME* estimates of total ozone (in *DU*) using both hard data (at *SBUV* points, shown in circles) and soft information.

3.6 Some comparisons

As can be seen by comparing the *BMEmean* map (Fig. 5a) with the actual map of Fig. 2, Approach 2 leads to a noticeable improvement in TO_3 estimation across space. This map is more realistic as it does not suffer from the artifact observed in Fig. 4 The associated map of estimation error standard deviation (s_K) values is obtained using the expression

$$s_K(s_k) = [\int dc_k (c_k - \overline{x_k})^2 f_K^{bc}(c_k)]^{1/2}, \qquad (6)$$

and is shown in Fig. 5b (the mean value of the TO_3 fluctuation at the estimation point $\boldsymbol{p_k}$ is $\overline{x_k} = 0$, in this case). Note the substantial accuracy improvement where the mapping error decreases from a maximum of about 15 DU (Fig 4b) down to as low as 5 DU (Fig 5b).

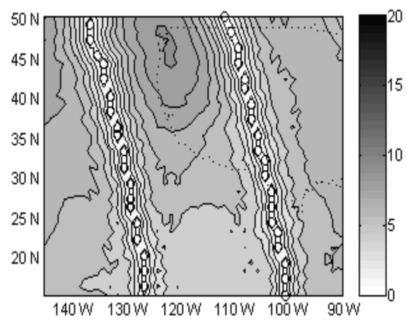

Figure –5b. The associated map of estimation error standard deviations for Fig. 5a.

The kriging standard deviation (s_e) considered in Approach 1 has been the subject of some criticism (see, e.g., [11]) being independent of the data values. The *BME* standard deviation (s_K), on the other hand, depends on the specific data set considered. The s_K can provide an adequate estimation error assessment when the shape of the *PDF* is not very complicated, otherwise a realistic assessment of the mapping error can be achieved using *BME* confidence sets, etc.

Furthermore, we calculated the differences between the estimated TO_3 values (Fig 4a and 5a) and the actual values (Fig 3) at all data points for which TO_3 values are available from *TOMS* (small crosses in Fig. 1). The histograms of the estimation errors are shown in Fig. 6 for Approach 1 (dotted line) and for Approach 2 (plain line). Clearly, the former has a sharper peak around zero estimation error, which implies that accounting for the physical equation the *BME* map of Approach 2 produced more accurate TO_3 estimates at a much higher frequency than Approach 1. In addition, the mean square error (*MSE*), i.e. the average of the squared estimation errors, drops from 106.5 DU^2 (Approach 1) down to 79.1 DU^2 (Approach 2) -- an improvement of 26% in accuracy. Another measure of error indicating bias is the mean error (*ME*), i.e. the plain average of estimation errors. As is shown in Fig. 6, the *ME* is equal to -3 DU when only hard data are used (indicating a slight bias), whereas the *ME* value drops to -0.8 DU when both hard and soft data are used.

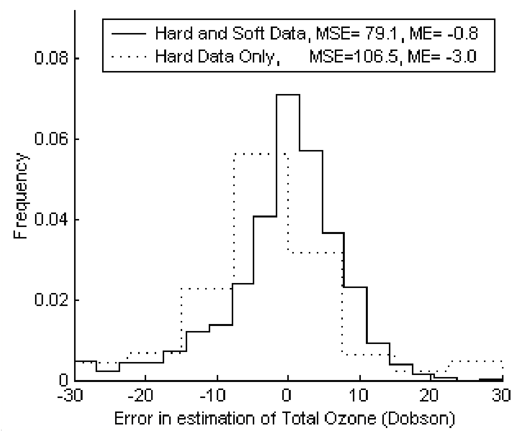

Figure –6. Frequency istribution of spatial estimation errors of total ozone mapping obtained by *BME* (plain line) and spatial kriging (dotted line).

4. CONCLUSIONS

We demonstrated the usefulness and practicality of *MSG* techniques to assimilate data from various information sources (different instruments, empirical laws, etc.) and generate accurate maps of total ozone in the atmosphere. The *BME* technique integrates sparse data obtained at the locations of the *SBUV* measurements with physical knowledge bases, as well as soft data obtained from the total ozone-tropopause pressure analysis involving an empirical physical equation. Soft data is critical information that is rigorously assimilated by the *BME* method, thus yielding more accurate maps of total ozone than other techniques currently in use. Future work will extend the numerical analysis to the use of *SBUV* data sets to construct maps of the ozone profile throughout the Earth. Such maps will be important tools in the dynamic monitoring of ozone's distribution in the atmosphere, which has considerable financial, social, ecological and human health implications.

ACKNOWLEDGEMENTS

This work has been supported by grants from the National Aeronautics and Space Administration (60-00RFQ041), the National Institute of Environmental Health Sciences (P42-ES05948 and P30-ES10126), and the Army Research Office (DAAG55-98-1-0289).

REFERENCES

1. Herman, J.R., P.K. Bhartia, O. Torres, C. Hsu, C. Seftor and E. Celarier, 1997. "Global Distribution of UV-Absorbing Aerosols From Nimbus-7/TOMS Data", *J. Geophys. Res*., 102, 16, 911-16, 922.
2. Bhartia, P.K., S. Taylor, R. D. McPeters, and C. Wellemeyer, 1995. "Applications of the Langley Plot Method to the Calibration of SBUV Instrument on Nimbus-7 Satellite", *J. Geophys. Res.*, 100, 2997-3004.
3. Bhartia, P.K., R.D. McPeters, C.L. Mateer, L.E. Flynn, and C. Wellemeyer, 1996. "Algorithm for the estimation of vertical ozone profile from the backscattered ultraviolet (BUV) technique", *J. Geophys. Res.*, 101, 18793-18806, 1996.
4. Aikin, A., M. D. Binder, E. Celarier, D. Considine, E. Cordero, M.F. Crum, E. Fleming, B. Gage, W. Heaps, E. Hilsenrath, S. Hollandsworth, D. Larko, G. Morris, P. A. Newman, R. Stolarski, R.M. Todaro, and C. Weaver, 2002. *Stratospheric Ozone Electronic Textbook.* Goddard Space Flight Center, Atmospheric Chemistry and Dynamics Branch (Code 916). Web page: http://see.gsfc.nasa.gov/edu/SEES/strat/class/S_class.htm
5. Christakos, G., 2000, *Modern Spatiotemporal Geostatistics*, Oxford University Press, New York, NY (3rd reprint, 2001).
6. Christakos, G. and V. Vyas, 1998. "A composite spatiotemporal study of ozone distribution over eastern United States". *Atmospheric Environment*, 32(16), 2845-2857.
7. Christakos, G. and A. Kolovos, 1999. A study of the spatiotemporal health impacts of the ozone exposure, *J. Exposure Anal. and Env. Epidemiol.*, 9 (1999), pp. 322-335.
8. Christakos, G., P. Bogaert, and M.L. Serre, 2002. *Temporal GIS: Advanced Functions for Field-Based Applications*, Springer-Verlag, New York, N.Y. (CD Rom included).
9. Wallace, J.M. and P.V. Hobbs, 1977. *Atmospheric Sciences-An Introductory Survey*, Acad. Press, San Diego, CA.
10. Stein, M. L., 1999. *Interpolation of Spatial Data.* Springer-Verlag, New York, N.Y.
11. Goovaerts, P., 1997. *Geostatistics for Natural Resources Evaluation.* Oxford Univ. Press, New York, N.Y.