



Mapping spatio-temporal variables: The impact of the time-averaging window width on the spatial accuracy

Yuval*, David M. Broday, Yohay Carmel

Department of Civil and Environmental Engineering, Technion, Israel Institute of Technology, Haifa 3200, Israel

Received 29 October 2004; received in revised form 8 February 2005; accepted 25 February 2005

Abstract

Spatial mapping of variables that vary in space and time is a common procedure in many research fields. Very often it is of interest to map the time-average or time-integration of the variable over the whole period of interest. Normally, such a map is produced by spatially interpolating the whole period averages of the observed data. An alternative option is to first spatially interpolate narrow time slice averages of the variable and then sum the resultant maps. This paper discusses the latter option, and the accuracy of the spatio-temporal variable interpolation as a function of the width of the time-averaging window. Theoretically, using a linear and data-value independent operator to interpolate a complete data set (i.e. without missing data), the accuracy is independent of the width of the time-averaging window. However, using a nonlinear or a data-value dependent interpolation operator, and/or in the presence of missing data, the accuracy of the interpolation can vary with the averaging window width. The concept is demonstrated using a set of half-hourly SO₂ concentrations measured at 20 monitoring stations in Haifa Bay area, Israel, during the years 1996–2002. Cross-validated interpolation accuracy measures calculated for this data set vary significantly with the time-averaging window width, showing a clear minimum at daily averaging. The results and their general implications for the interpolation of spatio-temporal variables are discussed.

© 2005 Elsevier Ltd. All rights reserved.

Keywords: Mapping accuracy; Spatial interpolation; Time-averaging window

1. Introduction

Often, the time-aggregation of spatio-temporal variables, i.e. their temporal average or time-integration, is of interest. For example, climatologists are usually interested in meteorological variables averaged at a time resolution coarser than a season (Fasullo, 2004; Lucero and Rodrigues, 2004; Sherwood, 2000; Skirvin et al., 2003) and epidemiologists often look at the exposure to air pollutants (usually estimated as the time-integration

of the pollutant concentration) over long periods in the order of years (Lall et al., 2004; Liblik et al., 2003; Samet et al., 2000). Since researchers using time-aggregated data usually feel comfortable with the data's temporal resolution, they mainly focus on the interpolation of the spatially sparse observations to a fine regular grid (De Cesare et al., 2001). Thus, it is not unusual that spatial maps of precipitation (e.g. Doggett et al., 2004; Karnieli, 1990) or exposure to air pollution are generated by spatially interpolating annually, or multi-annual averages of the observed variables (Nikiforov et al., 1998; Wong et al., 2004).

An alternative option for mapping time-aggregated variables is to first compute averages of the data records

*Corresponding author. Tel.: +972 4 8292676;
fax: +972 4 8292606.

E-mail address: lavyuy@tx.technion.ac.il (Yuval).

over narrow time slices, interpolate these time slice averages to full grid maps and then sum, or average, the maps to obtain the spatial time-integration or time-averaged maps for the whole period. By the definition of a linear operator (Weidmann, 1980, p. 50), a linear and data-value independent interpolation, based on a complete set of data points, yields the same spatial map regardless of the width of the time-averaging window applied to the observed data prior to the interpolation. However, using a data-value dependent or a nonlinear interpolation operator, the temporal resolution of the data to which averaging is applied might determine the accuracy of the time-aggregated map. Moreover, almost inevitably, data sets are missing many data points due to instrumentation failure, scheduled maintenance, electricity outages, etc. (e.g. Carrol et al., 1997; Sherwood, 2001). The accuracy of time-aggregated maps, derived using different averaging window widths, might be influenced by the way those missing data points are handled regardless of the method used for carrying out the interpolation. Sherwood (2000) found large differences between spatial maps of atmospheric tides using the narrowest and the widest possible time-averaging window widths. Jeffrey et al. (2001) discussed the relative advantages of using daily and monthly data for spatial interpolation of precipitation in Australia. Bearing in mind the various time scales on which environmental phenomena vary, a systematic study of the effects of time-averaging on the accuracy of the interpolation maps is required.

Two spatial interpolation methods are considered here, the inverse distance weighting (IDW) and kriging. The IDW is a linear interpolation operator (Isaaks and Srivastava, 1989) which depends only on the fixed distances between the data collection locations. It is independent of the actual data values on which it operates and thus easy to implement in an automatic manner for the purpose of carrying out the thousands of sequential interpolations required for this study. The data-value independency and the pure linearity of IDW enabled isolating the effects of missing data on the variation in mapping accuracy from the effects of the nonlinearity and data-value dependency. The kriging optimal interpolation (Cressie, 1993; Isaaks and Srivastava, 1989) has long been established in many fields as the mainstream technique for spatial interpolations (Goovaerts, 1997). Normally, kriging interpolation is manually performed by a skilled user who determines the values of the required input parameters. In our case however, such an interactive procedure is not a viable option. Hence, an unsupervised kriging scheme was developed to carry out the many sequential interpolations in an objective and efficient manner.

Using the IDW and kriging interpolation methods, this work studies the impact of the temporal resolution at which time-averaging is applied on the spatial

accuracy of mapping. A data set of half-hourly SO₂ concentrations collected at 20 monitoring stations in the Haifa Bay area during the years 1996–2002 is used for demonstrations. A full cross-validated procedure tests the accuracy of the interpolation maps as a function of the width of the time-averaging window which was used to produce them. Corresponding full grid maps are produced to qualitatively visualise the significance of the variations in the cross-validated accuracy measures. The implications of the finding for interpolation of other spatio-temporal variables are discussed. The paper begins with an exposition of the main theme and the methodologies used to study it. The SO₂ data set is presented in the next section and is followed by the results and the discussion sections.

2. Methodology

2.1. Interpolation of time-aggregated data

Consider a variable $V(\mathbf{x}, t)$, where \mathbf{x} are the 2D or 3D spatial coordinates and t is time. For simplicity, in the following we take \mathbf{x} as 2D but the concept can be readily adapted to three spatial dimensions. In each of the spatial locations \mathbf{x}_j , $j = 1, \dots, M$, the value of V is sampled at N regularly spaced time points, t_i , $i = 1, \dots, N$, to yield a data set $V(\mathbf{x}_j, t_i)$; $j = 1, \dots, M$; $i = 1, \dots, N$.

A spatial time-aggregated map of $V(\mathbf{x}, t)$ at a fine grid is normally produced by first averaging $V(\mathbf{x}_j, t)$ over the whole period $\tau = N\delta t$, where $\delta t = t_i - t_{i-1}$,

$$\bar{V}(\mathbf{x}_j, \tau = N\delta t) = \frac{1}{N} \sum_{i=1}^N V(\mathbf{x}_j, t_i), \quad j = 1, \dots, M. \quad (1)$$

In the case of Eq. (1), the time-averaging window width, W , equals the total number of time points in the data records. The interpolation of the M averages $\bar{V}(\mathbf{x}_j, \tau = N\delta t)$ to the fine grid can be carried out through

$$\bar{V}_{\mathcal{F}}(\mathbf{x}, \tau = N\delta t) = \mathcal{F}[\bar{V}(\mathbf{x}_j, \tau = N\delta t)], \quad j = 1, \dots, M, \quad (2)$$

where $\bar{V}_{\mathcal{F}}(\mathbf{x}, \tau = N\delta t)$ is the full grid interpolated map of the whole period average of $V(\mathbf{x}_j, t)$, and \mathcal{F} is a spatial interpolation operator. Multiplication of $\bar{V}_{\mathcal{F}}(\mathbf{x}, \tau = N\delta t)$ by N yields the estimated time-integration of $V(\mathbf{x}, t)$ over the full study domain. Note that in this case, the complete spatio-temporal field representation of $V(\mathbf{x}, t)$ is given by N identical maps $\bar{V}_{\mathcal{F}}(\mathbf{x}, \tau = N\delta t)$ assigned to each of the time points $1, \dots, N$. This points to a clear disadvantage of this approach, where information about the temporal variability of the complete spatio-temporal field is lost, with possible adverse effects on the spatial mapping of the time-aggregation of the variable.

An alternative approach considers N/W time slices of the sampled data, each spanning a time window $\tau = W\delta t$, where W is in the range of $1 \leq W \leq N$ time points and such that N/W is an integer. The averages of these time slices are given by

$$\bar{V}(\mathbf{x}_j, \tau_k) = \frac{1}{W} \sum_{i=(k-1)W+1}^{kW} V(\mathbf{x}_j, t_i),$$

$$j = 1, \dots, M, k = 1, \dots, N/W, \tag{3}$$

where τ_k is the time period of the k th time slice. The times slice averages $\bar{V}(\mathbf{x}_j, \tau_k)$ can be interpolated to the fine grid through

$$\bar{V}_{\mathcal{F}}(\mathbf{x}, \tau_k) = \mathcal{F}[\bar{V}(\mathbf{x}_j, \tau_k)],$$

$$j = 1, \dots, M, k = 1, \dots, N/W. \tag{4}$$

The final interpolation map of the time-averaged $V(\mathbf{x}, t)$, based on time-averaging window width W , is obtained by

$$\bar{V}_{\mathcal{F}}(\mathbf{x}, \tau = W\delta t) = \frac{1}{N} \sum_{k=1}^{N/W} \bar{V}_{\mathcal{F}}(\mathbf{x}, \tau_k). \tag{5}$$

Multiplication of $\bar{V}_{\mathcal{F}}(\mathbf{x}, \tau = W\delta t)$ by N yields a time-integrated estimation map of $V(\mathbf{x}, t)$.

Note that the second approach generalises the first one, with the procedure in Eqs. (3)–(5) degenerating to that described by Eqs. (1) and (2) when $W = N$. It will be shown later that using a refined temporal resolution of time-averaging can result in substantial benefits for the spatial mapping. An apparent drawback of using a narrow W , especially for data sets consisting of a large number of time points, is the additional computation time required to carry out many more spatial interpolations.

2.2. Estimating the accuracy of the spatial interpolation

There will always be some level of discrepancy between real values of a spatial variable and their corresponding interpolated surrogates. Interpolating a time-aggregated variable, the discrepancy can be minimised using $\bar{V}_{\mathcal{F}}(\mathbf{x}, \tau = W_{\text{opt}}\delta t)$, where W_{opt} is the width of an optimal time-averaging window. Estimating the accuracy of the mapping for the purpose of choosing the W_{opt} should be carried out by cross-validating results against the actual observed values (Stone, 1974; Diem, 2003). We used a leave-one-out cross-validation procedure, which is the limiting case and probably best form of the Jackknife validation (Miller, 1974). A description of its implementation in our case follows.

The cross-validated interpolation of the data averages at the k th time slice is given by

$$\bar{V}_{\mathcal{F}}(\mathbf{x}_j, \tau_k) = \tilde{\mathcal{F}}[\bar{V}(\mathbf{x}_j, \tau_k)],$$

$$j = 1, \dots, M, k = 1, \dots, N/W, \tag{6}$$

where the operator $\tilde{\mathcal{F}}$ produces an interpolated value at each of the observation locations using data only from the other $M - 1$ locations. The operation $\tilde{\mathcal{F}}$ results in a set of M interpolation values at the data sampling locations, each of which produced using only the sampled data at the other locations. As such, this set simulates interpolated values at locations where data were not sampled. The cross-validated time-average of $V(\mathbf{x}_j, t)$ over the whole period is given by

$$\bar{V}_{\tilde{\mathcal{F}}}(\mathbf{x}_j, \tau = W\delta t) = \frac{1}{N} \sum_{k=1}^{N/W} \bar{V}_{\tilde{\mathcal{F}}}(\mathbf{x}_j, \tau_k),$$

$$j = 1, \dots, M, k = 1, \dots, N/W \tag{7}$$

and its multiplication by N yields the corresponding time-integration. Note that the procedure described above implements a true cross-validation. In the case of kriging interpolation, the semivariogram is computed separately for each of the cross-validated interpolation points. This is in contrast to the common kriging cross-validation procedure which recalculates the kriging weights for each cross-validated point but utilises the same semivariogram, computed using all of the data points.

A quantitative estimation of the relative fidelity by which the full grid interpolation $\bar{V}_{\mathcal{F}}(\mathbf{x}, \tau)$ represents the corresponding true $\bar{V}(\mathbf{x})$ can be made by various measures of the difference between the cross-validated $\bar{V}_{\tilde{\mathcal{F}}}(\mathbf{x}_j, \tau)$ values and the true values at the observation locations $\bar{V}(\mathbf{x}_j)$. Following Willmott (1982), we compared the means and standard deviations of the observed values and their cross-validated counterparts calculated for each time-averaging window width, in addition to the corresponding root mean square error (RMSE) and the index of agreement (IA). The RMSE summarises the mean difference in units of the observed and interpolated values. The IA is a nondimensional relative bounded measure in the range $[0, 1]$ recommended by Willmott (1982) for comparisons between model results and actual values. Full grid interpolation maps, $\bar{V}_{\mathcal{F}}(\mathbf{x}, \tau = W\delta t)$, were produced to obtain an insight into how, and to what extent, the differences in the cross-validated accuracy measures are manifested in spatial maps produced using different time-averaging window widths.

2.3. Interpolation methods

IDW is the first interpolation method used in this study. Detailed discussion of the method is given by Isaaks and Srivastava (1989). The IDW interpolation estimate is a linear combination of the observed values, inversely weighted by the distances of the observation locations from the interpolation point. The IDW operator depends only on these distances and is independent of the observed values on which it operates.

In the context of this paper, this means independence of W , provided that no data are missing from the data set. The IDW operator can be readily applied to a series of data time slices in an automatic manner. As such, it serves as a benchmark to the more complicated unsupervised kriging interpolation described below. The linearity and the time slice independence of the IDW interpolation is exploited in this study to gain insight into the results achieved by the data-value dependent and slightly nonlinear kriging method.

Kriging is an optimal spatial interpolation in the sense of minimising the squared interpolation error. The ordinary kriging interpolation estimator used in this paper is a linear combination of the observed values, weighted by a set of optimal weights derived using an empirical semivariogram. The semivariogram measures the data variance as a function of spatial lag and provides information about the spatial autocorrelation in the data (Diem, 2003). Through its dependence on the semivariogram, the kriging operator depends on the observed values on which it operates. This introduces some level of nonlinearity in the kriging operator. It also means that the operator varies with the time window width, and hence the time-averaged spatial maps produced by Eq. (5) or (7) depend on the data division to time slices and on the width of the averaging window W which was used to create them. In spite of being relatively computationally intensive, the kriging method became very popular in many fields due to its theoretical reasoning and successful practical applications (e.g. Matías et al., 2004; Cattle et al., 2002). An excellent complete derivation and discussion of the method is given by Cressie (1993).

In most applications, kriging interpolation is performed interactively. A manual selection of various parameters by a user determines the interpolation map. Obviously, the thousands of sequential kriging applications required for this study cannot be carried out manually. To facilitate a sequential unsupervised application of the kriging interpolation, we first note that in all applications of Eq. (4) or (6) the sampling locations are identical for all the different data time slices. This enabled choosing in advance for all the thousand interpolations, a single set of geography-related parameters such as anisotropy angle and ratios, the number of distance lags, and the range up to which the model semivariogram should be fitted. The model semivariogram itself was selected from a list of conditionally negative defined functions as the one best fitting the empirical semivariogram (Cressie, 1993; Zimmerman et al., 1999). The nugget parameter was set to zero in all cases due to the relatively small sampling errors (see next section). In some instances, mainly when the range of the data values was very narrow, a decreasing linear function best fitted the empirical semivariogram. In such cases, a nugget model semivariogram was automatically

enforced, rendering all the interpolated values equal to their mean value. To ensure robustness of the unsupervised kriging, the quality of the interpolations was assessed at each time slice by the RMSE between the real values and the corresponding cross-validated interpolations. Various parameters were adjusted by experimentation until the unsupervised kriging achieved errors within reasonable predefined bounds at every single time slice. As a final check, unsupervised full grid interpolations of a few hundred randomly selected time slices were compared to the corresponding results obtained by a manual interactive process. In the great majority of cases, the differences were hardly noticeable.

3. Data

Concentrations of SO_2 are routinely collected at stations of the air quality monitoring network in Haifa Bay area, Israel. Half-hourly data from 1 January 1996 to 31 December 2002 were available for the study. The monitoring system includes 16 stations that operated since 1996 and four additional stations that started to report at the beginnings of 1997, 1998, 2000 and 2002. Fifteen of the stations are maintained by the Haifa District Municipal Association for the Environment and the rest are maintained by the Israel Electric Corporation, whose local power station is a major source of SO_2 . As depicted in Fig. 1, the monitoring stations are heterogeneously distributed in the study area in a pattern dictated by the population distribution and the municipal borders. The sharp topographic variations (see Fig. 1) cause complex regional wind patterns which are locally affected by the presence of many narrow ravines. These serve as air pollution conduits and may

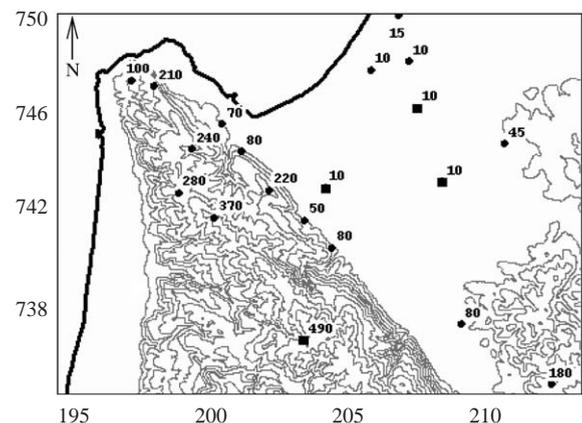


Fig. 1. Study area. Shown are the Haifa Bay area shoreline (thick line), elevation contours in 50m spacing, and the locations of the monitoring stations and their elevations. Marked with large squares are the four stations that started to report during the study period.

Table 1
Missing data points sorted into groups of missing data period lengths

	1	2–4	5–14	15–49	50–199	200–499	500–999	1000–4999
1996	405	11,340	1494	3256	4993	3102	637	10,550
1997	658	11,857	1081	3467	5903	3756	2140	8749
1998	539	12,262	914	1535	6368	3519	3166	3951
1999	821	11,670	772	2135	4127	3664	1172	4653
2000	3073	6915	2590	7287	5245	2752	2074	12,173
2001	3875	7879	1187	1274	1971	2223	1659	1068
2002	4212	7902	2343	5909	3965	2240	3227	0
1996–2002	13,583	69,825	10,381	24,863	32,572	21,256	14,075	41,144

The table gives the total number of missing data of each group for each year separately and for the whole study period. Only data from the 16 stations that reported throughout the study period were considered.

be responsible for large horizontal SO₂ concentration gradients. The instrumentation in all the stations is periodically calibrated. The typical observation error is estimated at 1–2%, with an upper limit the smaller between 5% of the data values and 20 µg m⁻³. The average SO₂ values at the stations vary between 4.0 and 14.9 µg m⁻³ but confirmed values of up to 2691 µg m⁻³ were measured, and values in the hundreds of micrograms per cubic metre are common.

On an average, 10% of the data are missing in each station, but this varies between stations and throughout the years. As can be seen in Table 1, most common are missing data periods of 1–4 time points (0.5–2.0 h). However, continuous periods of missing data of up to three months (4300 time points) exist and contribute a large fraction of the total number of missing data points. Table 1 was constructed using only data of the 16 stations that operated throughout the whole study period. For improved spatial exposure coverage, one may want to use the data from the additional four stations. In that case, SO₂ concentrations at these locations in the years prior to their establishment should be regarded as missing data points as well, with the consequence of elevating the proportion of missing data points to 18% of the complete 20 stations data set.

Except when mentioned otherwise, missing data were not explicitly filled. To handle missing data points, the summation of time series in Eqs. (1), (3), (5), and (7) was always carried out by averaging the existing (i.e. actually measured) values and multiplying the average by the number of time points in the series. While this operation is identical to simple summation if no missing data exist, it implicitly assigns the whole series' mean value to data points that are missing. In the case of long periods of missing data, the true mean of the missing data period may be very different from the mean of the whole series and hence the calculated sum might be far from the true one. Attempts to fill the missing data gaps using linear and nonlinear regression were only very marginally effective for missing data periods longer than 48 h.

Neither other possible methods (e.g. Little and Rubin, 2002; Schafer, 1997) can be expected to assign perfect substitutes, especially for data gaps during long time periods. Thus, some effect of the missing data on the mapping accuracy is unavoidable. In order to explore it in a simple setting, none of the advanced data filling methods was used.

4. Results

The cross-validation procedure described by Eqs. (6)–(7) was applied to the SO₂ data set using time-averaging window widths of $W = 122,640, 17,520, 4380, 1460, 336, 48, 12,$ and 6 time points, which correspond to the whole period (7 years), yearly, seasonal, monthly, weekly, daily, 6-h, and 3-h averaging, respectively. Table 2 compares the observed and estimated cross-validated means and standard deviations of the exposure to SO₂ at the monitoring stations. The kriging mean estimates at all the time-averaging window widths are slightly higher than the observed mean. The IDW estimates are distributed around the observed value. However, the standard deviation of the observed exposure is almost double that of the cross-validated estimates, pointing to a lack of a dynamic range in the exposure estimated by both kriging and IDW.

Fig. 2 shows the RMSE and the IA between the actual exposure to SO₂ at the monitoring stations and the estimated exposures. Both the RMSE and the IA are given as functions of the width of the averaging window W . Results are given for the cases of IDW and kriging as the interpolation operators, using both the full 20 stations set and the partial set of 16 stations that operated throughout the whole study period. Examining the upper panel of Fig. 2, we first note that the RMSE of both the kriging and IDW results vary as a function of the averaging window width. Moreover, using any time-averaging window width resulted in the RMSE being lower than that of full period averaging, with the

Table 2
Comparison of observed and estimated statistics

	3 hour	6 hour	Day	Week	Month	Season	Year	Period
Observed mean	1.742	1.742	1.742	1.742	1.742	1.742	1.742	1.742
Estimated mean, kriging	1.770	1.759	1.759	1.763	1.756	1.753	1.763	1.832
Estimated mean, IDW	1.729	1.727	1.737	1.754	1.760	1.762	1.771	1.790
Observed Std	0.685	0.685	0.685	0.685	0.685	0.685	0.685	0.685
Estimated Std, kriging	0.328	0.328	0.334	0.341	0.329	0.336	0.352	0.258
Estimated Std, IDW	0.359	0.359	0.356	0.361	0.368	0.366	0.372	0.337

The estimated statistics refer to the cross-validated interpolations carried out using data from all the 20 monitoring stations through Eq. (7), using IDW or kriging as the interpolation operators. All values are given in $10^9 \mu\text{g m}^{-3}\text{s}$.

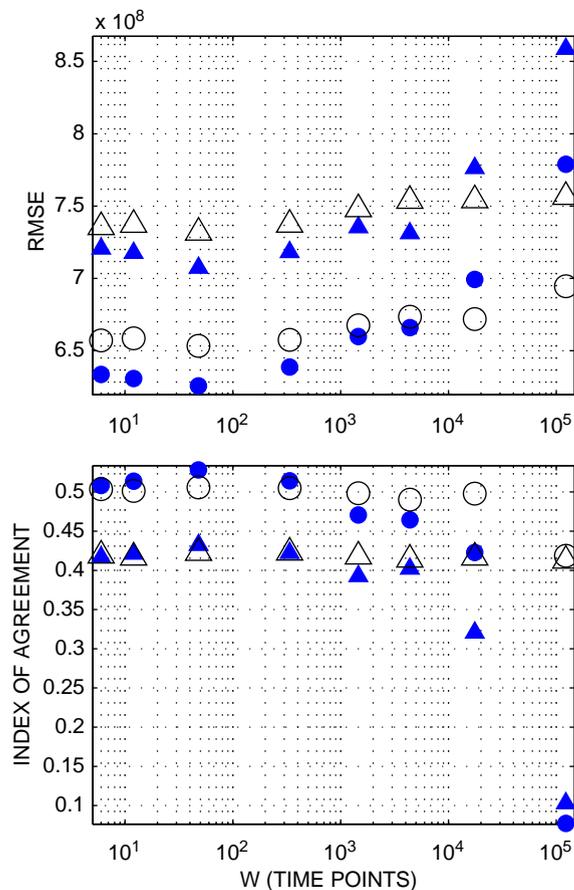


Fig. 2. RMSE (in $10^9 \mu\text{g m}^{-3}\text{s}$) and IA between the true exposure at the monitoring stations and the corresponding cross-validated exposure, calculated as a function of the time-averaging window width W . Window widths are given in time points units that correspond to 3-h, 6-h, daily, weekly, monthly, seasonally, yearly, and the whole period (7 years) averaging. Circles denote results obtained using the complete 20 stations data set. Triangles denote using only data from the 16 monitoring stations that operated throughout the whole study period. Results using kriging and IDW interpolations are denoted by filled and empty symbols, respectively.

minimum error achieved at $W = 48$ time points. The IA variations are smaller, especially in the IDW results, but still the maximum agreement is achieved using $W = 48$. This implies that for our SO_2 data set, the most accurate exposure estimation is achieved using daily averaging. For a complete data set (no missing data), using a linear interpolation operator like IDW, the exposure values and thus the interpolation accuracy measures should not vary with the time-averaging window width. Hence, the existence of missing data in the SO_2 is probably the cause for any variations noted in the IDW results. The results obtained using kriging interpolation show much more variability, pointing to the important role of the data-value dependency and the slight nonlinearity of kriging in determining the interpolation accuracy as a function of W .

A second point to note in Fig. 2 is that for both interpolation methods the exposure estimates obtained using data from all the 20 stations (circles) have a lower RMSE and a higher IA than those obtained using the partial set of 16 stations (triangles). This is not surprising since spatial interpolations tend to improve as a result of better spatial coverage of observations. However, in our case, the additional monitoring stations supply additional observed information only during part of the study period. Errors due to inaccurately filling data in these stations for the periods before their establishment partially cancel the benefits of using additional stations when a whole period time-averaging window width is used. However, employing narrower time-averaging window widths eliminates this problem and enables better use of the improved spatial coverage provided by the additional stations.

The last important point that Fig. 2 depicts is the superior performance of kriging compared to that of the IDW interpolation. However, note that more accurate exposure estimation is obtained using IDW interpolation at the wide time-averaging window widths. Only employment of narrow averaging window widths enabled realising the advantage of kriging and achieving the best overall estimation.

To appreciate the significance of the variations in the cross-validated RMSE and IA for full grid mapping, maps produced using different time-averaging window widths can be examined and compared with the actual exposure at the monitoring stations. Fig. 3a shows the full grid map obtained through Eqs. (3)–(5) using kriging interpolations and the optimal window width of $W = 48$ time points. In comparison, Fig. 3b has been obtained using the same interpolation method through Eqs. (1) and (2), i.e. using a whole period averaging window width. The two full grid exposure maps were

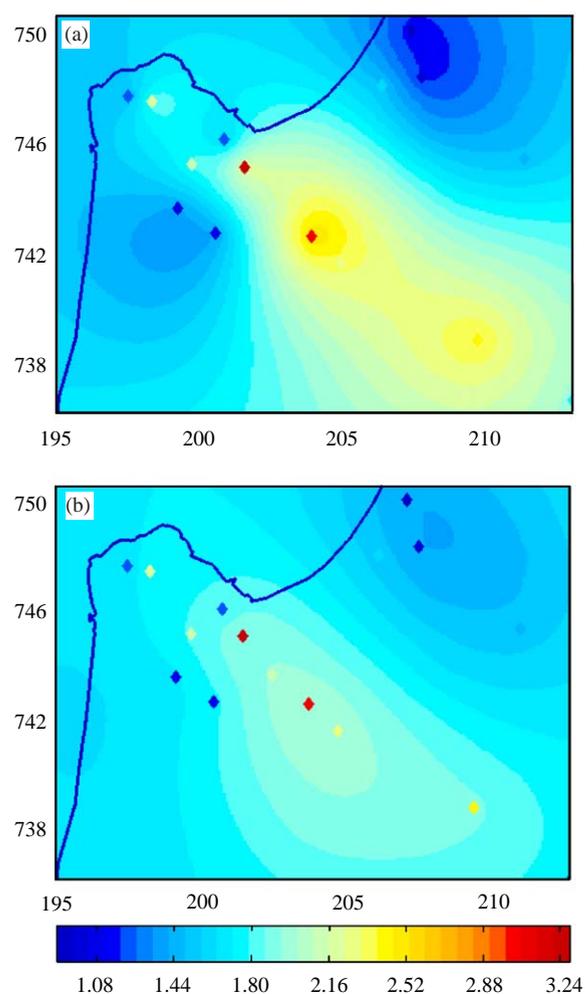


Fig. 3. Actual exposure to SO₂ at the 16 monitoring stations that operated since 1996, superimposed on the estimated spatial exposure based on interpolation of the monitoring data from all the 20 stations to a fine regular grid. (a) Fine grid interpolation carried out using daily averages of the data. (b) Fine grid interpolation of the whole period (7 years) data averages. The two full grid exposure maps were identically colour-scaled with the colours at the extreme edges of the scale corresponding to the minimum and maximum actual exposure values. Exposure values are given in $10^9 \mu\text{g m}^{-3} \text{ s}$.

produced using data from all the 20 monitoring stations. They were identically colour-scaled with the colours at the extreme edges of the scale corresponding to the minimum and maximum actual exposure values at the 16 monitoring stations that operated since 1996. These actual exposure values are superimposed on both maps for comparison. The exposure at the four additional stations are not shown due to lack of data from the periods before they started to operate.

The interpolated exposure values in Fig. 3a show a dynamic range which is wider than that of the corresponding values in Fig. 3b, and is closer to the dynamic range of the actual exposure at the monitoring locations. The topography in the Haifa Bay area is complex (Fig. 1) and results in large anisotropic horizontal gradients of SO₂ concentrations. Accordingly, the spatial resolution of the exposure map in Fig. 3a is finer, and better reproduces the actual exposure values in the stations. Thus, the spatial exposure estimates in Fig. 3a are visually more plausible than the exposure estimates in Fig. 3b, as well as possessing the corresponding better cross-validated RMSE and IA, as shown in Fig. 2.

5. Discussion

It has been recently shown that the apparently innocent technical issue of time-averaging can have a significant impact on the prediction of time series using nonlinear predicting models (Yuval and Hsieh, 2002) and on the analysis of regime structures of variables that have non-Gaussian temporal distribution (Teng and Monahan, 2004). This paper studied the impact that time-averaging have on the accuracy of time-aggregated spatial mapping of spatio-temporal variables. The results given above clearly demonstrate the possible benefits of spatial mapping which is based on narrow time-averaging window widths.

The processes of testing for the optimal averaging window width and producing the corresponding best full grid interpolation maps might be computationally expensive if the data records are long and the optimal time-averaging window is narrow. A practically important issue is therefore the trade-off between computing time and interpolation accuracy. Another important issue is the choice of method to carry out the interpolations. The results of this study suggest that the kriging method can have an advantage over the IDW interpolation but that this advantage might be realised only while using relatively narrow time-averaging window widths. Implementing kriging in the processes of Eqs. (3)–(5) and (6)–(7) is not straightforward and requires a reliable and robust scheme to carry out unsupervised interpolations. This study has shown that development of such a scheme is feasible.

An interesting question regards the mechanisms that generate the noted differences in the mapping accuracy and how they affect its variability as a function of the time-averaging window width. In particular, it is important to understand in which cases mapping of time-aggregated data can be expected to benefit from using a narrow time-averaging window width, and whether the width of the optimal averaging window can be related to any known parameter of the data. This study points to the presence of missing data as one cause for the noted variation in mapping accuracy. Filling missing data, especially during long time periods, is prone to some level of inaccuracy even when a sophisticated data imputation scheme is used (Little and Rubin, 2002). Thus, the mapping accuracy is always detrimentally affected, to a certain degree, in the presence of missing data. Using narrow time-averaging window widths, the deleterious effect of the missing data is minimised. The effect that the time-averaging had on the accuracy of mapping using the kriging method was much larger than the effect on mapping using IDW (see Fig. 2), suggesting that the kriging method can benefit more from using narrow time-averaging window widths due to its nonlinearity and/or data-value dependency. Studies comparing the performance of the kriging and IDW interpolation methods disagree about which method is more accurate (Zimmerman et al., 1999). The results of this paper suggest that the conclusions drawn from such studies which use time-aggregated variables may be biased by the selected width of the data's time-averaging window.

Analysis of geostatistical variables that show pronounced spatio-temporal behaviour is a developing research field. Recent publications (e.g. Carrol et al., 1997; Christakos and Vyas, 1998; De Cesare et al., 2001; Kyriakidis and Journel, 2001) proposed promising statistical methods for modelling complete spatio-temporal fields. Improved spatial maps of time-aggregated variables will probably result as by-products of such processes. However, there are still many theoretical and practical problems that need addressing (De Cesare et al., 2001) before these methods enter the arsenal of the geostatistics practitioner. This paper addresses an issue concerning the accuracy of mapping using the mainstream data interpolation methods. We believe that for the foreseeable future adopting our conclusions can lead to generation of more accurate maps of time-aggregated variables.

Acknowledgements

The SO₂ data were kindly provided by the Haifa District Municipal Association for the Environment and by the Israel Electric Corporation. This work was partially supported by the Urban-Exposure consortium

agreement, Contract EVK4-CT-2002-00090, under the 5th Framework Program of the Research Directorate-General of the European Commission, by the Israeli Ministry of Science and Technology, and by an Israeli Council for Higher Education scholarship to the first author. We thank Dr. Ronit Nirel for her useful comments on an initial version of the manuscript.

References

- Carrol, R.J., Chen, R., George, E.I., Li, T.H., Newton, H., Schmiediche, H., Wang, N., 1997. Ozone exposure and population density in Harris County Texas. *Journal of the American Statistical Association* 92, 392–404.
- Cattle, J.A., McBratney, A.B., Minasny, B., 2002. Kriging method evaluation for assessing the spatial distribution of urban soils lead contamination. *Journal of Environmental Quality* 5, 1576–1588.
- Christakos, G., Vyas, V.M., 1998. A composite space/time approach to studying ozone distribution over eastern United States. *Atmospheric Environment* 32, 2845–2857.
- Cressie, N.A.C., 1993. *Statistics for Spatial Data*, Revised ed. Wiley, New York.
- De Cesare, L., Myers, D.E., Posa, D., 2001. Estimating and modeling space-time correlation structures. *Statistics and Probability Letters* 52 (1), 9–14.
- Diem, J.E., 2003. A critical examination of ozone mapping from a spatial-scale perspective. *Environmental Pollution* 125, 369–383.
- Doggett, M., Daly, C., Smith, J., Gibson, W., Taylor, G., Johnson G., Pasteris, P., 2004. High-resolution 1971–2000 mean monthly temperature maps for the western United States. In: *Proceedings of the 14th AMS Conference on Applied Climatology, 84th AMS Annual Meeting Combined Preprints*, American Meteorological Society, Seattle, WA, 13–16 January 2004, Paper 4.3, CD-ROM.
- Fasullo, J., 2004. Biennial characteristics of Indian monsoon rainfall. *Journal of Climate* 17, 1981–2972.
- Goovaerts, P., 1997. *Geostatistics for Natural Resources Evaluation*. Oxford University Press, New York.
- Isaaks, E.H., Srivastava, R.M., 1989. *An Introduction to Applied Geostatistics*. Oxford University Press, New York.
- Jeffrey, S.J., Carter, J.O., Moodie, K.B., Beswick, A.R., 2001. Using spatial interpolation to construct a comprehensive archive of Australian climate data. *Environmental Modelling and Software* 16, 309–330.
- Karnieli, A., 1990. Application of kriging technique to areal precipitation mapping in Arizona. *GeoJournal* 22 (4), 391–398.
- Kyriakidis, P.A., Journel, A.G., 2001. Stochastic modeling of atmospheric pollution: a spatial time-series framework. Part I: methodology. *Atmospheric Environment* 35, 2331–2337.
- Lall, R., Kendall, M., Ito, K., Thurston, G.D., 2004. Estimation of historical PM_{2.5} exposures for health effects assessment. *Atmospheric Environment* 38, 5217–5226.
- Liblik, V., Pensa, M., Rätsep, A., 2003. Air pollution zones and harmful pollution levels of alkaline dust for plants. *Water Air and Soil Pollution: Focus* 3, 199–209.
- Little, R.J.A., Rubin, D.B., 2002. *Statistical Analysis with Missing Data*. Wiley, New York.

- Lucero, O.A., Rodríguez, N.C., 2004. Evolution of spatial patterns of subdecadal signals in annual rainfall in Southern South America and Southern and Central North America. *Atmospheric Research* 70, 147–169.
- Matías, J.M., Vaamonde, A., Taboada, J., González-Manteiga, W., 2004. Comparison of kriging and neural networks with application to exploitation of a slate mine. *Mathematical Geology* 36, 463–486.
- Miller, R.J., 1974. The Jackknife—a review. *Biometrika* 61, 1–15.
- Nikiforov, S.V., Aggarwal, M., Nadas, A., Kinney, P.L., 1998. Methods for spatial interpolation of long-term ozone concentrations. *Journal of Exposure Analysis and Environmental Epidemiology* 8, 465–481.
- Samet, J.M., Dominici, F., Curriero, F.C., Coursac, I., Zeger, S.L., 2000. Fine particulate air pollution and mortality in 20 US cities 1987–1994. *The New England Journal of Medicine* 343, 1742–1749.
- Schafer, J.L., 1997. *Analysis of Incomplete Multivariate Data*. Chapman & Hall, London.
- Sherwood, S., 2000. Climate signal mapping and an application to atmospheric tides. *Geophysical Research Letters* 27, 3525–3528.
- Sherwood, S., 2001. Climate signals from station arrays with missing data, and an application to winds. *Journal of Geophysical Research* 105, 29,489–29,500.
- Skirvin, S.M., Marsh, S.E., McClaran, M.P., 2003. Climate spatial variability and data resolution in a semi-arid watershed, south-eastern Arizona. *Journal of Arid Environments* 54, 667–686.
- Stone, M., 1974. Cross-validators choice and assessment of statistical predictions (with discussion). *Journal of the Royal Statistical Society B* 36, 111–147.
- Teng, Q., Monahan, A.H., 2004. Effects of time averaging on climate regimes. *Geophysical Research Letters* 31.
- Weidmann, J., 1980. *Linear Operators in Hilbert Spaces*. Springer, New York.
- Willmott, C.J., 1982. Some comments on the evaluation of model performance. *Bulletin of the American Meteorological Society* 6, 1309–1313.
- Wong, D.W., Yuan, L., Perlin, S.A., 2004. Comparison of spatial interpolation methods for the estimation of air quality data. *Journal of Exposure Analysis and Environmental Epidemiology* 14, 404–415.
- Yuval, Hsieh, W., 2002. The impact of time-averaging on the detectability of nonlinear empirical relations. *Quarterly Journal of the Royal Meteorological Society* 128, 1609–1622.
- Zimmerman, D., Pavlik, C., Ruggles, A., Armstrong, M.P., 1999. An experimental comparison of ordinary kriging and universal kriging and inverse distance weighting. *Mathematical Geology* 31, 375–390.