SIAM J. SCI. COMPUT.
Vol. 18, No. 3, pp. 949–956, May 1997

015

TIMELY COMMUNICATION

The SIAM Journal on Scientific Computing occasionally publishes "Timely Communications," short articles whose timely content warrants an accelerated review and publication schedule. These articles are expected to significantly benefit the research community and usually appear in the journal within six months of acceptance. Submissions of Timely Communications should not exceed six journal pages.

MULTISCALE ALGORITHM FOR ATMOSPHERIC DATA ASSIMILATION*

ACHI BRANDT† AND LEONID YU. ZASLAVSKY‡

Abstract. We propose a novel multiscale algorithm for the problem of model assimilation of data. The algorithm allows one to efficiently perform optimal statistical interpolation of observed data from a given forecast w^f and vector of observations w^o. The core of the new approach is a combination of two multiscale tools: a multiresolution iterative process and a multigrid fast-summation technique. Our approach allows efficient computations related to global filtering and interpolation of the observations, particularly between data-rich and data-sparse areas.

In this paper, we describe an iterative process based on a multiresolution simultaneous displacement technique and a localized variational calculation of iteration parameters. We explain how this process can be efficiently combined with the multigrid fast-summation procedure.

Key words. data assimilation, iterative, multilevel, multiscale, multigrid, multiresolution

AMS subject classifications. 65R10, 65R20, 65U05, 86A10

PII. S106482759528942X

1. Introduction.

Atmospheric data assimilation technologies provide estimates of the state of the geophysical system using its dynamical model and observational data measured in different places and at different times. Modern three- and four-dimensional data assimilation technologies lead to more complete and more accurate estimates than estimates achieved from separate analysis of sets of observations. They provide mechanisms for filtering and interpolating the observations and for transporting the information from data-rich to data-sparse areas [1], [2], [3], [4], [5]. The necessity of processing a large amount of information in a complicated way makes data assimilation computationally demanding. Novel computational techniques have to be developed for this challenging area.

*Received by the editors July 24, 1995; accepted for publication (in revised form) July 30, 1996. A preliminary version of this paper appeared in the Proc. Seventh Copper Mountain Conference on Multigrid Methods, Copper Mountain, CO, 1995 vol. CP3339, NASA, Hampton, VA, 1996, pp. 87–96; *Multiscale Algorithm for Atmospheric Data Assimilation. Part* I. *Multiscale Iterative Process*, Tech. report CS95-09, Weizmann Institute of Science, Rehovot, Israel, 1995. This research was supported in part by NASA contract NAS5-32339 through Ecodynamics Research Associates, Inc., grant F49620-92-J-0439 from the United States Air Force, and by the K. F. Gauss Minerva Center for Scientific Computation at the Weizmann Institute of Science.

http://www.siam.org/journals/sisc/18-3/28942.html

†Department of Applied Mathematics and Computer Science, Weizmann Institute of Science, Rehovot, 76100, Israel.

‡Department of Applied Mathematics and Computer Science, Weizmann Institute of Science, Rehovot, 76100, Israel. Current address: Courant Institute of Mathematical Sciences, New York University, 251 Mercer Street, New York, NY, 10012 (zasl@crick.biomath.nyu.edu). The work of this author was supported by a Sir Charles Clore Post-Doctoral Fellowship.

In this paper, we propose a novel multiscale approach for the problem of optimal statistical interpolation of observed data. At the initial stage, we consider the case of univariate analysis of single-level radiosonde height data. The problem of optimal statistical interpolation is related to calculating the best estimated state of the atmosphere w^a ("analyzed state"), from the forecast w^f, and the set of observations w^o. Mathematically, the best estimated state is determined by the formula [4], [6], [7]

$$w^a = w^f + P^f H^T (H P^f H^T + R)^{-1} (w^o + H w^f). \tag{1}$$

Typically, the forecast field w^f is defined on a regular spherical grid, while the set of observations w^o is defined on an irregular network of observation points. H is an interpolation operator from the regular grid to the observation network, P^f is the forecast error covariance matrix, and R is the observation error covariance matrix. The observation error covariance matrix R is assumed to be diagonal, with $R_{ii} = (\sigma_i^o)^2$. The forecast error covariance function $P^f(x_1, x_2)$ is defined for any pair of points x_1 and x_2 on the sphere by the formula $P^f(x_1, x_2) = \sigma^f(x_1)\mu(x_1, x_2)\sigma^f(x_2)$, where the forecast error correlation function $\mu(x_1, x_2)$ is described as a smooth decreasing function of the distance between the points x_1 and x_2 [7]. The matrices P^f and μ are the restrictions of functions $P^f(x_1, x_2)$ and $\mu(x_1, x_2)$ on the regular latitude–longitude grid.

In the data assimilation system GEOS-1/DAS currently used in the Data Assimilation Office, NASA/Goddard Space Flight Center, the equivalent of the equation

$$(H P^f H^T + R) y = g$$

is solved with a conjugate gradient algorithm that is carefully preconditioned [7]. The preconditioner is based on solving the localized system in subdomains. As long as the number of measurements and the grid size of the forecast system are moderate, such a technique gives an appropriate result. (Note that the algorithm based on the local preconditioning is, in a wide sense, a two-scale algorithm.) However, for the much larger sets of measurements and the much more detailed forecasts to be used in the near future, more advanced computational techniques have to be developed. Indeed, if the subdomains contain large amounts of observations, preconditioning becomes too expensive. On the other hand, if the subdomains are too small, the preconditioner is poor and the iterative process converges slowly. This effect can be overcome if *a sequence of spatial scales* is used in the iterative process, instead of only *local* and *global* scales as in the procedure used in the current GEOS-1/DAS algorithm [7]. Thus, it is quite natural to develop a novel fast computational procedure for assimilation of data using *multiscale* computational approaches.

The optimal statistical interpolation procedure (1) is multiscaled in *two stages* as follows. A multiresolution approach allows us to construct an iterative process with a high convergence rate for the equation

$$(H P^f H^T + R) y = g, \tag{2}$$

where $g = w^o - H w^f$. Multigrid fast-summation techniques [8], [9], [10] make it possible to quickly evaluate the discrete integral transform

$$z = P^f H^T y. \tag{3}$$

The combination of these two multiscale techniques is the basic idea behind the novel multiscale approach for assimilation of data proposed here.

We have to emphasize that standard multiscale techniques, such as the multigrid methods [11], [12], [13], [14] and the wavelet methods [15], [16], cannot be applied to the problem of data assimilation in a straightforward way. As we have already mentioned, a global data assimilation system has to provide a reasonable mechanism for information transport between data-rich and data-sparse areas. This strong inhomogeneity in the observation network cannot be treated in a straightforward way either by multigrid or by wavelet methods. However, the proposed technique, based on combining a fast multiresolution iterative process with an efficient multigrid fast-summation procedure, successfully treats this inhomogeneity.

This paper is mostly devoted to constructing a multiresolution iterative algorithm for solving (2), which efficiently treats different scales. Accordingly, in the computer experiments, all summations have been performed in a straightforward way. An efficient multigrid fast-summation procedure has been developed by A. Brandt and A. A. Lubrecht[1] (see [8], [9] for more details). A comprehensive review of fast-summation algorithms is given in the paper by L. Greengard [10].

Without loss of generality, equation (2) can be replaced by the system of equations

$$\sum_j \tilde{P}^f_{ij} y_j + R_{ii} y_i = g_i, \tag{4}$$

where $\tilde{P}^f_{ij}$ is assumed to be of the form $\tilde{P}^f_{ij} = \tilde{\sigma}^f_i \mu_{ij} \tilde{\sigma}^f_j$ for the respective ith and jth observation points x_i and x_j, $\mu_{ij} = \mu(x_i, x_j)$, and $\tilde{\sigma}^f_i = (H\sigma^f)_i$. (While the matrix P is defined for the points of the regular grid and interpolated to the observation network using the operator H, the matrix $\tilde{P}$ is defined by the same formula directly on the observation network.) Indeed, the term $HPH^T y - \tilde{P}y$ may be treated as an additional source in the right-hand side. This small term is nonprincipal at all scales and can easily be taken into account in iterations.

Since we want to deal explicitly with the smoothness properties of the kernel μ_{ij}, we write equation (4) in the form

$$\sum_j \mu_{ij} u_j + \left(\frac{\sigma^o_i}{\tilde{\sigma}^f_i} \right)^2 u_i = f_i, \tag{5}$$

where $u_i = \tilde{\sigma}^f_i y_i$, and $f_i = (\tilde{\sigma}^f_i)^{-1} g_i$, or in the matrix form

$$Au = f, \tag{6}$$

where A is symmetric and is a presumably positive definite matrix with elements

$$A_{ij} = \mu_{ij} + \left(\frac{\sigma^o_i}{\tilde{\sigma}^f_i} \right)^2 \delta_{ij},$$

and δ_{ij} is discrete Dirac delta function.

[1]The fast-summation algorithm developed by A. Brandt and A. A. Lubrecht is based on the following idea. Summation $w(x_i) = \sum_j K(x_i, x_j) u(x_j)$ with the kernel $K(x_i, x_j)$, which is smooth enough, can be performed on the coarse grid in the following way: the function $u(x_i)$ is averaged from the fine to the coarse grid, summation is performed on the coarse grid using the coarse grid approximation of the smooth kernel, and the result is interpolated from the coarse to the fine grid. In a more general case, the kernel $K(x_i, x_j)$ is decomposed in two parts: a local part that is efficiently evaluated at the fine grid and a smooth part that is evaluated on the coarse grid. In practice, a sequence of grids is used.

2. General strategy. It is important to understand why many common iterative processes, such as Jacobi, Gauss–Seidel, or conjugate gradients, converge slowly when applied to equation (6). Consider, for example, the simplest iterative process

$$u^{(n+1)} = u^{(n)} + \omega r^{(n)}, \tag{7}$$

where the residual

$$r^{(n)} = f - Au^{(n)},$$

and the parameter $\omega \approx (\rho(A))^{-1}$ with $\rho(A)$ being the spectral radius of the operator A.

The process (7) effectively reduces the error components corresponding to the large eigenvalues λ_l of the operator A such that

$$\omega\lambda_l \sim 1,$$

while the error components corresponding to the small eigenvalues λ_s, for which

$$\omega\lambda_s \ll 1,$$

are reduced slowly [17]. Since the summation $\sum_j \mu_{ij}u_j$ in (5) is made with a smooth kernel, eigenvectors of A corresponding to the large eigenvalues are (mostly) spatially smooth, and eigenvectors of A corresponding to small eigenvalues are oscillatory in space. Therefore, one cannot define one particular value of ω that would give an essential reduction of all spectral error components.

The effect described above is well studied for the case when (6) is obtained as a grid approximation of the continuous integral equation. A few multiscale techniques based on multigrid [8], [9] and wavelet [16] approaches were developed in the 1990s and successfully applied to a range of problems. Unfortunately, these techniques cannot be applied to the considered problem in a straightforward way. As we mentioned above, the global data assimilation system has to carefully transport information between data-rich and data-sparse areas. Neither standard multigrid nor wavelet methods can efficiently treat strong inhomogeneities of the observation network. However, a novel multiscale technique that we have developed can overcome this difficulty.

The central idea of the multiresolution approach developed below is to filter sequentially spectral components of $r^{(n)}$ and to choose for each of them a value of the iteration parameter that gives an essential reduction of the corresponding error component.

The major particular difficulty that has been overcome successfully in this work is how to define variable pass spatial filters $\mathcal{F}_h$ depending on the scale parameter h for a field defined on a very inhomogeneous network. An appropriate filter will be described in section 3.

When some component $\mathcal{F}_h r^{(n)}$ of the residual $r^{(n)}$ has been filtered, one should next calculate the correction vector. A simple way to do this is to use a *scalar* iteration parameter ω_h, i.e., to calculate the correction as $\omega_h \mathcal{F}_h r^{(n)}$. Then, the modified iterative process (7) can be written as

$$u^{(n+1)} = u^{(n)} + \omega_h^{(n)} \mathcal{F}_h r^{(n)}, \tag{8}$$

where the iteration parameter depends on the scale h in some way. An intrinsic disadvantage of schemes like (8) is that one *global* iteration parameter $\omega_h^{(n)}$ is determined for the entire domain. The optimal correction at a spatial point x_i should, however,

depend only on the residual values at points located at most a few h from x_i. Therefore, in section 4, we construct a procedure for calculating an iteration parameter $\omega_{h,i}^{(n)}$ for each point x_i *locally*, using only the values of the residual components in some area around it. This means that the iterative process that we construct can be written as

$$u^{(n+1)} = u^{(n)} + \Omega_h^{(n)} \mathcal{F}_h r^{(n)}, \tag{9}$$

where

$$\Omega_h = \operatorname{diag}(\omega_{h,i}^{(n)}).$$

We discuss the structure of the multiscale iterative cycle in section 5.

3. Spatial filter. In this section, we construct a filter applicable to functions defined on a very irregular discrete network. Obviously, we want our filter to work like a usual spectral high-pass filter in the data dense regions. What do we want to get in regions of sparse data? Suppose we have an observation point s that is separated from other points by distance

$$d_s = \min_{p \neq s} \operatorname{dist}(s, p),$$

where $\operatorname{dist}(s, p)$ is the three-dimensional distance between the points s and p. We would like to take into account the residual component r_s only on the scales h, which are large enough ($h \simeq d_s$ and $h > d_s$), and neglect it on the scales h, which are small compared with d_s.

We define the filter $\mathcal{F}_h$ that satisfies these requirements by the formula

$$(\mathcal{F}_h r)_i = r_i - \gamma_i \sum_j r_j \exp\left(-\frac{1}{2}\frac{(\operatorname{dist}(i,j))^2}{h^2}\right), \tag{10}$$

where h is the current scale, and the parameter γ_i is defined by the formula

$$\gamma_i = \left(\sum_j \exp\left(-\frac{1}{2}\frac{(\operatorname{dist}(i,j))^2}{h^2}\right)\right)^{-1}.$$

This filter can be efficiently calculated using the fast-summation procedure [8], [9].

4. Calculation of Ω_h. The scalar iteration parameter $\omega_h^{(n)}$ in (8) can be determined from the variational condition of minimizing the Euclidean norm of the scale-h component of the new residual $r_h^{(n+1)}$:

$$\min_{\omega_h^{(n)}} ||\mathcal{F}_h(r^{(n)} - \omega_h^{(n)} A \mathcal{F}_h r^{(n)})||^2. \tag{11}$$

The Euclidean norm on the observation network and corresponding inner product are defined by the respective formulas

$$||u||^2 = \sum_i |u_i|^2,$$

$$(u, v) = \sum_i u_i v_i,$$

where the summation is made over the observation points. Condition (11) leads to the formula

$$\omega_h^{(n)} = \frac{(p,q)}{(q,q)}, \tag{12}$$

where

$$p = \mathcal{F}_h r^{(n)},$$
$$q = \mathcal{F}_h A p.$$

An intrinsic disadvantage of formula (12) is that it is global. In order to *localize* it, we use a family of weighted Euclidean norms. Let us introduce a family of weighted inner products

$$(u,v)_{H,i} = \sum_j u_j v_j \exp\left(-\frac{1}{2}\frac{(\mathrm{dist}(i,j))^2}{H^2}\right),$$

where the summation is made over the observation points and $H > 0$ is the scale parameter. An appropriate choice of the matrix $\Omega_h^{(n)}$ in (9) is

$$\Omega_h^{(n)} = \mathrm{diag}(\omega_{h,i}^{(n)}),$$

where

$$\omega_{h,i}^{(n)} = \frac{(p^{(n)}, q^{(n)})_{3h,i}}{(q^{(n)}, q^{(n)})_{3h,i}}, \tag{13}$$
$$p^{(n)} = \mathcal{F}_h r^{(n)},$$
$$q^{(n)} = \mathcal{F}_h A p^{(n)}.$$

Note that both the numerator and the denominator of (13) can be efficiently evaluated by the fast-summation procedure [8], [9].

5. Structure of the multiscale iterative cycle. In order to define the order of the multiresolution iterations, some spatial scale h has to be prescribed for each iteration. The current spatial scale is determined by the formula $h = H \cdot 2^{1-\nu(n)}$, where H is the largest scale and $\nu(n)$ is an integer parameter prescribed to the nth iteration. The multiscale iterative algorithm can be written as follows:

DO n = 1, NITER

$r^{(n)} = f - Au^{(n)}$	*Residual calculation*
IF $\nu(n) > 0$ THEN	
$h = H \cdot 2^{1-\nu(n)}$	*Definition of the current scale*
$p^{(n)} = \mathcal{F}_h r^{(n)}$	*Filtering*
$q^{(n)} = \mathcal{F}_h A p^{(n)}$	
$\omega_{h,i}^{(n)} = \frac{(p^{(n)},q^{(n)})_{3h,i}}{(q^{(n)},q^{(n)})_{3h,i}}$	*Calculation of the iteration parameters*
$\Omega_h^{(n)} = \mathrm{diag}(\omega_{h,i}^{(n)})$	
$u^{(n+1)} = u^{(n)} + \Omega_h^{(n)} p^{(n)}$	*Calculation of the new approximation to the solution*

ELSE
$p^{(n)} = r^{(n)}$ *Filtering is not used*
$q^{(n)} = Ap^{(n)}$
$\omega_h^{(n)} = \frac{(p^{(n)}, q^{(n)})}{(q^{(n)}, q^{(n)})}$ *Calculation of the iteration parameter*
$u^{(n+1)} = u^{(n)} + \omega_h^{(n)} p^{(n)}$ *Calculation of the new approximation to the solution*
ENDIF
ENDDO

For our initial tests we have used the standard $V(2,2)$ cycle with three iterations at the zeroth level [11], [12]. This means that the function $\nu(n)$ is periodic: $\nu(LC + k) = \nu(k)$ for any $k > 0$, and

$$\nu(n) = \begin{cases} NLVL - k + 1 & \text{if } n = 2 \cdot k - 1 + l, \quad k = 1, 2, \ldots, NLVL, \quad l = 0, 1; \\ 0 & \text{if } n = 2 \cdot NLVL + l, \quad l = 1, 2, 3; \\ k & \text{if } n = 2 \cdot NLVL + 2 + 2k + l, \quad k = 1, 2, \ldots, NLVL, \\ & \qquad l = 0, 1, \end{cases}$$

where $NLVL$ is the finest level number and $LC = 4 \cdot NLVL + 3$.

6. Numerical results. At this initial stage of the work, we have performed all of the numerical tests with unilevel (500 hPa) radiosonde height data. The locations of 715 radiosonde stations, residuals $w^o - Hw^f$ computed with a six-hour forecast from GEOS-1/DAS and radiosonde observations at 500 hPa, and values of σ^o and σ^f, were obtained from the Data Assimilation Office, NASA/Goddard Space Flight Center. Observation error variances σ_i^o were taken to be equal to 14.6 m for all radiosonde stations. Forecast error variances σ_i^f vary from point to point, ranging from 18 m to 35 m. The forecast correlation function is modeled by the formula

$$\mu(x_1, x_2) = \left(1 + \frac{(\text{dist}(x_1, x_2))^2}{L^2}\right)^{-1.208},$$

where $\text{dist}(x_1, x_2)$ is the three-dimensional distance between points x_1 and x_2 and L is the correlation distance $L = 951$ km.

We made our experiments with $NLVL = 5$. The scales that were used are shown in Table 1. The results of our experiments are shown in Table 2.

7. Discussion. The computational approach described in this paper leads toward a fast and efficient solver for the atmospheric data assimilation problem. It has been shown that the multiresolution iterative algorithm can provide a fast convergent solver. As long as the number of measurements is moderate (less than a few thousand, say), this algorithm by itself is already effective enough and it does have significant computational advantages in comparison with the current GEOS-1/DAS solver based on the preconditioned conjugate gradient method. Moreover, the full potential of the new approach can be used effectively for much larger sets of measurements. In this case, the developed multiresolution iterative process has to be combined with the multigrid fast-summation procedure.

Acknowledgments. This paper presents part of a joint work with S. Cohn, G. Gaspari, and A. da Silva of the Data Assimilation Office, NASA/Goddard Space Flight Center. The authors thank them for raising their interest in the problem

TABLE 1
Scale structure.

Level number	1	2	3	4	5
Scale h, km	10 000	5 000	2 500	1 250	625

TABLE 2
Convergence of the iterative procedure.

Multiscale cycle	L_2 norm of the residual	Rate of decrease of the norm
Initial	2.5510^{+1}	
1	6.8610^{-1}	0.027
2	3.7610^{-2}	0.054
3	2.0510^{-3}	0.054
4	8.8710^{-5}	0.043

of atmospheric data assimilation and providing the observation data and model parameters. The authors are also indebted to A. Kheifets, I. Rivin, M. Shapiro, and C. H. Venner for helpful discussions.

REFERENCES

[1] A. BRANDT AND L. YU. ZASLAVSKY, *Multilevel algorithm for atmospheric assimilation*, in Seventh Copper Mountain Conference on Multigrid Methods, Copper Mountain, CO, N. D. Melson, T. A. Manteuffel, S. F. McCormick, and C. C. Douglas, eds., vol. CP 3339, NASA, Hampton, VA, 1996, pp. 87–96.

[2] A. BRANDT AND L. YU. ZASLAVSKY, *Multiscale Algorithm for Atmospheric Data Assimilation. Part* I. *Multiscale Iterative Process*, Tech. report CS95-09, Weizmann Institute of Science, Rehovot, Israel, 1995.

[3] NATIONAL RESEARCH COUNCIL, *Four Dimensional Model Assimilation of Data. A Strategy for the Earth System Sciences*, National Academy Press, Washington, DC, 1991.

[4] R. DALEY, *Atmospheric Data Analysis*, Cambridge University Press, London, 1991.

[5] S. D. SCHUBERT, R. D. ROOD, AND J. PFAENDTER, *An assimilated dataset for earth science applications*, Bull. Amer. Meteor. Soc., 74 (1993), pp. 2331–2342.

[6] W. E. BAKER, S. C. BLOOM, J. S. WOOLLEN, M. S. NESTLER, E. BRIN, T. W. SCHLATTER, AND G. W. BRANSTATOR, *Experiments with a three-dimensional statistical objective analysis scheme using FGGE data*, Mon. Wea. Rev., 115 (1987), pp. 273–296.

[7] J. PFAENDTER, S. BLOOM, D. LAMICH, M. SEABLOOM, M. SIENKIEWICZ, J. STOBIE, AND A. DA SILVA, *Documentation on the Goddard Earth Observing System (GEOS) Data Assimilation System - Version* 1, NASA Tech. Memorandum 104606, Vol. 4, NASA, Goddard Space Flight Center, Greenbelt, MD, 1995.

[8] A. BRANDT AND A. A. LUBRECHT, *Multilevel multi-integration and fast solution of integral equations*, J. Comput. Phys., 90 (1990), pp. 348–370.

[9] A. BRANDT, *Multilevel algorithms of integral transforms and particle integrations with oscillatory kernels*, Comput. Phys. Comm., 65 (1991), pp. 24–38.

[10] L. GREENGARD, *Fast algorithms for classical physics*, Science, 265 (1994), pp. 903–914.

[11] A. BRANDT, *Multi-level adaptive solutions to boundary-value problems*, Math. Comp., 31 (1977), pp. 333–390.

[12] W. HACKBUSCH, *Multi-Grid Methods and Applications*, Springer-Verlag, Berlin, 1985.

[13] J. H. BRAMBLE, *Multigrid Methods*, Pitman Research Notes in Mathematics Series, Longman Scientific and Technical, Essex, U.K., 1993.

[14] P. WESSELING, *An Introduction to Multigrid Methods*, John Wiley and Sons, Chichester, U.K., 1992.

[15] I. DAUBECHIES, *Ten Lectures on Wavelets*, SIAM, Philadelphia, PA, 1992.

[16] B. ALPERT, G. BEYLKIN, R. COIFMAN, AND V. ROKHLIN, *Wavelet-like bases for the fast solution of second-kind integral equations*, SIAM J. Sci. Comput., 14 (1993), pp. 159–184.

[17] R. S. VARGA, *Matrix Iterative Analysis*, Prentice–Hall, Englewood Cliffs, NJ, 1963.