NOTES

The Extension of P. D. Thompson's Scheme to Multiple Times

JOHN M. LEWIS
NOAA/NESDIS Development Laboratory, Madison, WI 53706

LEE PANETTA[1]
Space Science and Engineering Center, University of Wisconsin, Madison 53706

8 January 1983 and 20 May 1983

ABSTRACT

P. D. Thompson devised a scheme to correct imperfect analyses of a conservative quantity at two observation times. His scheme has been extended to include a sequence of observation times. When the times are equally spaced, the governing adjustment equations simplify to an equation in one variable, a weighted average of the conservative property at the various times. The weights are found from Pascal's rule. The primary advantage of adding more observation times is to reduce the mean square error in the analyses. The limiting value of mean square error reduction is ½, ⅔, ¾, . . . , $(k-1)/k$ for 2, 3, 4, . . . , k times, respectively. The applicability of this method to adjustment of a sequence of mean temperature (thickness) fields from the VISSR Atmospheric Sounder (VAS) is discussed.

1. Introduction

Thompson (1969) used the variational approach to design an analysis–forecast scheme that can be used to adjust successive analyses that are clearly at odds with the prediction model. This procedure is designed to give perfect agreement at the time midway between two analyses, but with minimum adjustment of both analyses.

This approach has been adapted to the dynamical constraint of geostrophic potential vorticity conservation and used to adjust height analyses derived from the VISSR Atmospheric Sounder (VAS) which is to be carried on all GOES satellites during this decade (Lewis, 1982). In addition to producing analyses consistent with the dynamical law, a principal advantage is the ability to fill a data sparse area with extrapolated information from the neighboring time. Since the VAS instrument uses an infrared radiometer, the sparse data areas are those with clouds. Consequently, there is a distinct advantage to search a sequence of previous analyses to find clear areas and extrapolate to the present time. It is with this thought in mind that the extension of Thompson's approach to multiple times is investigated.

2. Mathematical development

Prior to discussing the mathematical details, a few words about the assumptions are appropriate. First, the geostrophic vorticity is assumed to be advected by an error-free geostrophic current, or stated alternatively, the errors in the advecting current are negligible in comparison to those in the vorticity. In practice, it is necessary to use a spatially and temporally smoothed current such as that advocated by Fjørtoft (1952). Second, all inconsistency between the analyses are assumed to reside in the observations and objective methods that are used to generate the analyses, i.e., the model is perfect.

To facilitate the description of the adjustment process, a schematic view of the sequential steps for a three-time version is shown in Fig. 1. In (a) three geopotential analyses separated by equal time intervals are shown. In (b) the Fjørtoft current is found by time and space smoothing of the three charts. In (c) the vorticities are superimposed over one of the steering currents. In (d) traces of vorticity along that steering current (ABC) are depicted. In (e) and (f) the adjustment process is depicted at the time(s) intermediate to the analysis times, a forecast from the earlier time and a hindcast from the latter time are compared. In this hypothetical case, the amplitude of the vorticity maximum is largest on the middle analysis. Also depicted in (e) are phase inconsistencies (→| |←) resulting from excessive separation of the maxima on the first two analyses, and too little separation between the maxima on the last two analyses. As expected intuitively, the least squares adjustment will decrease the amplitude on the middle analysis and raise the amplitude on the first and last analysis as shown in (f). The phase adjustments, de-

[1] Present affiliation: Forrestal Campus, Geophysical Fluid Dynamics Laboratory, Princeton University, Princeton, New Jersey 08540.

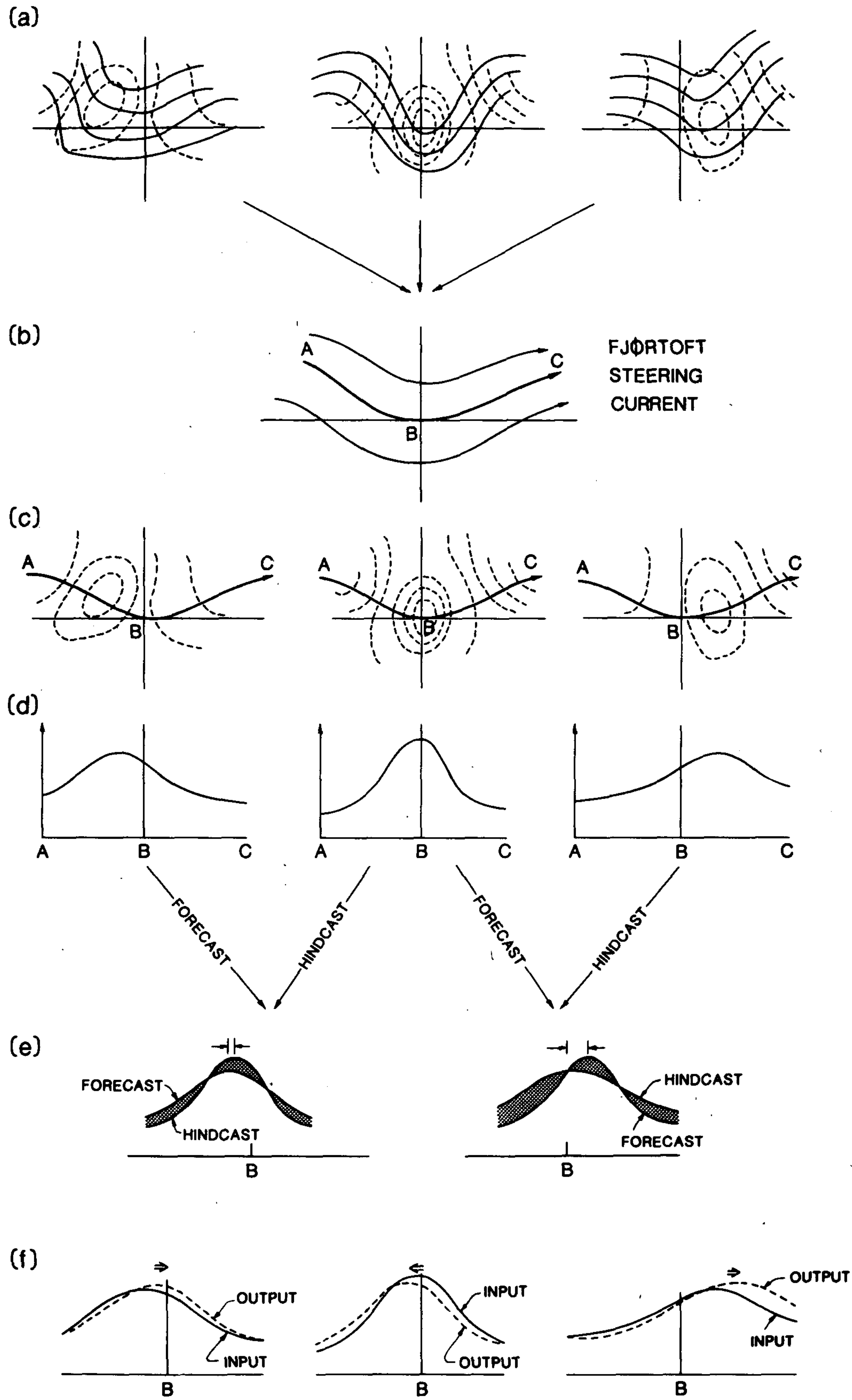

FIG. 1. Schematic view of the adjustment process at three times: (a) input heights, (b) construction of the steering current, (c) vorticity superimposed on the contour ABC, (d) traces of vorticity along ABC, (e) hindcast/forecast to intermediate times, and (f) adjusted vs. input vorticities.

noted by the arrow in (f), are such that the maxima is moved forward at the first time, backward at the second time and forward at the third time.

In the following equations the notation will follow Thompson's except that ζ will be used for vorticity (instead of Z), and instead of using $(n-1)$ and $(n+1)$ as subscripts to denote time, the subscripts 0, 2, 4, . . . , will be used to signify analyses at the first, second, third, . . . analysis time. Subscripts 1, 3, 5, · · · coincide with the times where the constraints are checked, i.e., the intermediate times. The list of symbols is found in Table 1. The constraint for the two time level problem is written as

$$f_1 = D_1 + \dot{S}_1 + E_1 = 0 \tag{1}$$

identical in form to Eq. (3) in Thompson. The adjustment is found by minimizing I:

$$I = \int (\delta_0^2 + \delta_2^2 + \lambda_1 f_1) ds, \tag{2}$$

where the integration is along the Fjørtoft steering current line, and λ_1 is the Lagrange multiplier. Following standard procedures in the calculus of variations (Lanczos, 1966), the equations governing the adjustments are:

$$\left.\begin{aligned} \delta_0 &= \tfrac{1}{2}(\dot{\lambda}_1 + \lambda_1) \\ \delta_2 &= \tfrac{1}{2}(\dot{\lambda}_1 - \lambda_1) \\ \ddot{\lambda}_1 - \lambda_1 + E_1 &= 0 \end{aligned}\right\} . \tag{3}$$

The last equation can also be written as

$$\ddot{S}_1 - S_1 + \dot{E}_1 = 0, \tag{4}$$

which is identical to Thompson's Eq. (4). Thompson used insight to simplify the problem at an early stage by adopting S_1 and D_1 as basic variables.

When three time levels are used, the constraints are

$$\left.\begin{aligned} f_1 = D_1 + \dot{S}_1 + E_1 = 0 \\ f_3 = D_3 + \dot{S}_3 + E_3 = 0 \end{aligned}\right\} , \tag{5}$$

and the functional takes the form

$$I = \int (\delta_0^2 + \delta_2^2 + \delta_4^2 + \lambda_1 f_1 + \lambda_3 f_3) ds. \tag{6}$$

In this case, the governing equations are

$$\left.\begin{aligned} 2\delta_0 &= \dot{\lambda}_1 + \lambda_1 \\ 2\delta_2 &= \dot{\lambda}_1 + \dot{\lambda}_3 + \lambda_3 - \lambda_1 \\ 2\delta_4 &= \dot{\lambda}_3 - \lambda_3 \\ \ddot{\lambda}_1 + \frac{\ddot{\lambda}_3}{2} + \dot{\lambda}_3 - \lambda_1 + \frac{\lambda_3}{2} + E_1 &= 0 \\ \ddot{\lambda}_3 + \frac{\ddot{\lambda}_1}{2} - \dot{\lambda}_1 - \lambda_3 + \frac{\lambda_1}{2} + E_3 &= 0 \end{aligned}\right\} . \tag{7}$$

Although the system of ordinary differential equations can be solved in this form, simplification can sometimes be achieved by combining the equations and isolating one variable. In this case, however, such a simplification isn't apparent by inspection. Recalling that Thompson was able to simplify the problem by adopting S_1 as the basic variable in the Euler–Lagrange (E–L) equation, the constraints are added to give

$$\delta_4 - \delta_0 + (\delta_0 + 2\delta_2 + \delta_4) + E_1 + E_3 = 0. \tag{8}$$

In this form, the weighted average $\delta_0 + 2\delta_2 + \delta_4 = \kappa$ suggests itself as the analog to $\delta_0 + \delta_2$ in Thompson's two time-level problem. Using this definition of κ in concert with the constraints yields

$$\left.\begin{aligned} 4\delta_0 &= \ddot{\kappa} + 2\dot{\kappa} + \kappa + 3E_1 + E_3 + (\dot{E}_1 + \dot{E}_3) \\ 4\delta_2 &= -\ddot{\kappa} + \kappa - E_1 + E_3 - (\dot{E}_1 + \dot{E}_3) \\ 4\delta_4 &= \ddot{\kappa} - 2\dot{\kappa} + \kappa - E_1 - 3E_3 + (\dot{E}_1 + \dot{E}_3) \end{aligned}\right\} . \tag{9}$$

The adjustment problem is significantly simplified in this form since the functional then reduces to the three squared discrepancies which are expressible in terms of κ and its derivatives. Forging through the minimization yields the following E–L equation:

$$\left.\begin{aligned} 3\ddddot{\kappa} - 6\ddot{\kappa} + 3\kappa + (E_1 - E_3) + 3(\ddot{E}_1 - \ddot{E}_3) \\ -7(\dot{E}_1 + \dot{E}_3) + 3(\dddot{E}_1 + \dddot{E}_3) = 0 \end{aligned}\right\} . \tag{10}$$

A solution to (10) requires four boundary conditions. Although the formal procedure of deriving the governing equation allows several options, we have chosen specification of the value of κ and its first derivative at each end of the streamline, or periodic conditions in the case of a closed contour.

Through the symmetry of the operators in the con-

TABLE 1. Nomenclature.

$\tilde{\zeta}$	input analysis	E_1	measure of inconsistency from first constraint
ζ	adjusted analysis	E_3	measure of inconsistency from second constraint
δ_0	$\zeta_0 - \tilde{\zeta}_0$, adjustment at earliest time	$\lvert\mathbf{V}\rvert$	speed of steering current
δ_2	$\zeta_2 - \tilde{\zeta}_2$, adjustment at middle time	$2\Delta t$	time interval between analyses
δ_4	$\zeta_4 - \tilde{\zeta}_4$, adjustment at latest time	l	$\lvert\mathbf{V}\rvert\Delta t$
S_1	$\delta_0 + \delta_2$	$(\dot{\ })$	derivative along streamline $[=l d/ds]$
S_3	$\delta_2 + \delta_4$	ϵ	original error (input-truth)
D_1	$\delta_2 - \delta_0$	$\hat{\epsilon}$	residual error (adjusted-truth)
D_3	$\delta_4 - \delta_2$	$\langle\ \rangle$	ensemble average

straints, the appropriate weighted average for more time levels can be found from the combinations:

$$\delta_0 + \delta_2$$

$$(\delta_0 + \delta_2) + (\delta_2 + \delta_4)$$

$$(\delta_0 + \delta_2) + 2(\delta_2 + \delta_4) + (\delta_4 + \delta_6)$$

$$(\delta_0 + \delta_2) + 3(\delta_2 + \delta_4) + 3(\delta_4 + \delta_6) + (\delta_6 + \delta_8)$$

or weighting coefficients

1:1

1:2:1

1:3:3:1

1:4:6:4:1

These are the coefficients of terms in the binomial expansion, $(p + q)^n$, and they have traditionally been generated by Pascal's rule and schematically pictured as Pascal's triangle (Cajori, 1980). For each added time level, the order of the governing adjustment equation increases by 2, but only even powered derivatives appear. This is a distinct advantage when the equation is discretized in space and solved iteratively by methods such as relaxation (Forsythe and Wasow, 1960).

3. Error reduction

Following Thompson's procedures for the analysis of error (Section 5 of his paper), some generalizations can be made. In particular, as $l \rightarrow 0$, the effect of the scheme is to smooth out the error by averaging, that is, to replace the error distribution ϵ_1 at time t_1, ϵ_2 at time $t_2, \ldots, \epsilon_k$ at time t_k with the error distribution

$$\hat{\epsilon}_j = \frac{1}{k}\sum_{i=1}^{k} \epsilon_i, \quad j = 1, 2, \ldots, k,$$

(notation has been altered slightly from Section 2, where only even numbers were used for analyses). Thompson mentioned this result for $k = 2$. The cumulative squared error over all analyses is $\sum_{i=1}^{k} \epsilon_i^2$ and the corresponding sum of residual squared error for the adjusted fields is

$$\sum_{i=1}^{k} \hat{\epsilon}_i^2 = k\left(\frac{1}{k}\sum_{i=1}^{k} \epsilon_i\right)^2 = \frac{1}{k}[\sum_{i=1}^{k} \epsilon_i^2 + 2\sum\sum_{i<j} \epsilon_i\epsilon_j]. \quad (11)$$

The difference between these cumulative mean square errors is

$$\frac{k-1}{k}\sum_{i=1}^{k} \epsilon_i^2 - \frac{2}{k}\sum\sum_{i<j} \epsilon_i\epsilon_j.$$

If now we regard the ϵ_i as being random variables, with ϵ_i and ϵ_j uncorrelated and with zero mean, the ensemble average or expected value (denoted by angle braces) of the difference is

$$\frac{k-1}{k}\sum_{i=1}^{k} \langle\epsilon_i^2\rangle - \frac{2}{k}\sum\sum_{i<j} \langle\epsilon_i\rangle\langle\epsilon_j\rangle$$

and the last term vanishes under the assumptions. So the error is reduced by a factor $(k - 1)/k$.

In practice, the vorticities are discretized along the contours and the reduction in mean square error is a function of the Courant number, $|\mathbf{V}|\Delta t/\Delta s$, and the wavelength of the event. Reference tests with analytic solutions to the linear advection equation can be used as an indication of the dependence of error reduction on this number. An ensemble of experiments were made for two cases: 1) the addition of spatially correlated error at two times and random error at the other, and 2) the addition of normally distributed random error at all three times.

Geopotential heights derived from the radiometric data from satellites have spatially correlated error (Hayden, 1977; Schlatter, 1981) and case 1) is meant to simulate the situation where rawinsonde derived heights are available at the first time and satellite derived heights are analyzed at the second and third times. For case 1), the amplitude of the vorticity signal is systematically underestimated at times 2 and 3. These amplitudes are drawn from the normal distribution with a mean of 0.9 and standard deviation of 0.1 with respect to the true vorticity amplitude. The amplitude of the vorticity distribution at time 1 is drawn from the normal distribution with mean of 1.0 and standard deviation of 0.1, again with respect to the true signal amplitude. The ensemble statistics stabilize well before the 100 trials used to generate the error reduction factor for each value of the Courant number. These results are displayed in Fig. 2b. As the Courant number approaches zero the reduction in mean square error approaches 55% at all three levels, less than the theoretical limit of ⅔ in the case of uncorrelated random error. The middle time definitely benefits from the availability of information on each side whereas the boundary times show the degrading effects of one-sidedness. The error reduction curve for the middle time crosses the Courant number axis at 3.75.

In case 2) normally distributed random noise is added at each time and the theoretical limit of reduction, ⅔, is reached when $|\mathbf{V}|\Delta t/\Delta s \rightarrow 0$ as shown in Fig. 2a. Again, the middle time shows substantially better reduction than the bounding times, crossing the Courant number axis at 2.25. These graphs pertain to a sinusoidal vorticity distribution with a wavelength of $16\Delta s$. As the wavelength decreases, the reduction in mean square error also decreases as expected from our knowledge of truncation error in finite-difference advection equations, i.e., truncation error is greater for short waves than long waves (Thompson, 1961).

The benefits gained by using three times instead of two are shown in Fig. 3. Normally distributed random error with the same characteristics is added to the distributions at each time. The residual errors after adjustment are averaged over the time levels and compared with the original mean square error to get the average reduction in error. There is an obvious benefit at low Courant numbers, but the results suggest that the critical value of the Courant number, beyond which there is no error reduction, is virtually the same in this averaged-sense.

4. Concluding remarks

The mechanics of including multiple times in P. D. Thompson's dynamical adjustment scheme has been developed. As expected, the procedures significantly simplify when the times are equally spaced. In such a case, the governing adjustment equations can be reduced to an ordinary differential equation in one variable, a weighted average of the conservative property at each time. The weights are the coefficients in the binomial expansion. The mean square errors in the analyses are reduced in proportion to the number of times used in the adjustment process. When the errors are uncorrelated, the limiting value of mean square error reduction is $(k - 1)/k$, where k is the number of times.

The adjustment of geopotential analyses from the VAS system in midlatitudes is suited to this procedure for the following reasons: 1) the conservation of geostrophic potential vorticity adequately describes the synoptic scale evolution in midlatitudes and depends only on the geopotential fields, and 2) the typical horizontal resolution for synoptic scale grids, ~300 km, coupled with the temporal spacing between analyses and speed of the steering currents give rise to Courant numbers <1.

As more times are used in the adjustment procedure, however, a point of diminishing returns can occur for two reasons: 1) the assumption of a non-adjustable steering current becomes less tenable as more times are used, and 2) mathematical difficulties accrue due to the increasing order of the governing equation and attendant boundary condition, especially when the analyses are conducted over limited areas. A time span of 9–12 h, 4 or 5 analysis times, respectively, would probably maximize the benefits of error reduction and minimize the objectionable aspects of strained physical assumption and mathematical drudgery.

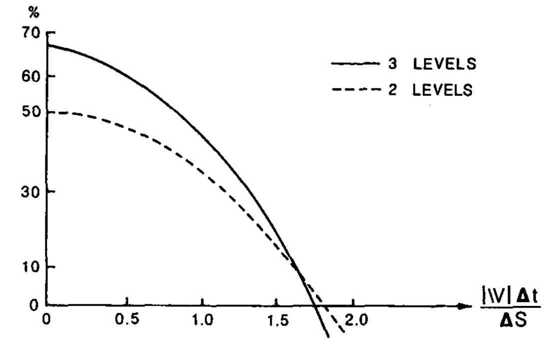

FIG. 3. Reduction in mean square error averaged over time levels. Solid curve indicates reduction for the case of three analysis periods whereas the dashed curve shows results when two analysis periods are used.

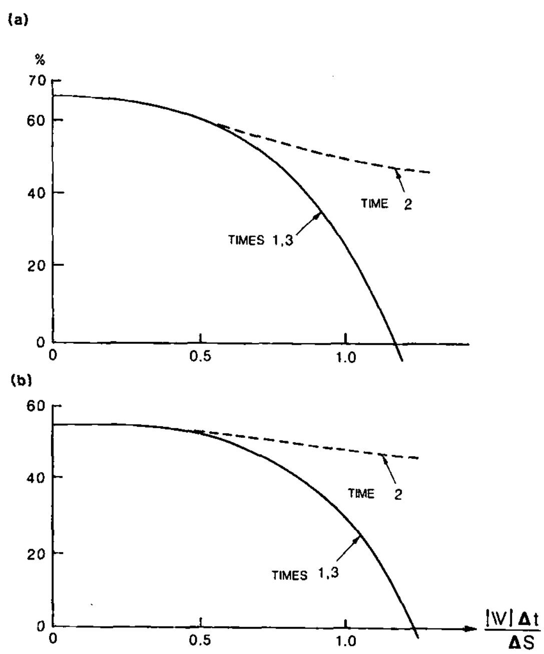

FIG. 2. Reduction in mean square error as a function of the Courant number: (a) random error added to vorticity at all three times and (b) systematic error added to vorticity distributions at second and third times while random error was added at the first time.

REFERENCES

Cajori, F., 1980: *A History of Mathematics*, 3rd ed. Chelsea Publishing, 524 pp.

Fjørtoft, R., 1952: On a numerical method of integrating the barotropic vorticity equation. *Tellus*, **4**, 179–194.

Forsythe, G. E., and W. R. Wasow, 1960: *Finite-Difference Methods for Partial Differential Equations*. Wiley and Sons, 444 pp.

Hayden, C. M., 1977: A summary of Nimbus-6 temperature retrieval accuracy statistics for DST-5 and DST-6. Third National NASA Administration Weather and Climate Program Science Review. NASA Conf. Pub. 2029, 271–276.

Lanczos, C., 1966: *The Variational Principles of Mechanics*, 4th ed. University of Toronto Press, 418 pp.

Lewis, J. M., 1982: Adaptation of P. D. Thompson's scheme to the constraint of potential vorticity conservation. *Mon. Wea. Rev.*, **110**, 1618–1634.

Schlatter, T. W., 1981: An assessment of operational TIROS-N temperature retrievals over the United States. *Mon. Wea. Rev.*, **109**, 110–119.

Thompson, P. D., 1961: *Numerical Weather Analysis and Prediction*, Macmillan, 157–161.

——, 1969: Reduction of analysis error through constraints of dynamical consistency. *J. Appl. Meteor.*, **8**, 738–742.