551·509·39 : 551·524

Increase of random errors in temperature forecasts by numerical method and limiting period of reliable forecasts

Y. P. RAO and K. S. RAMAMURTI

Meteorological Office, Poona

(*Received 4 February 1969*)

ABSTRACT. It is shown that random errors of temperature observations increase in the course of forecasting temperature field by numerical method. The forecast equation under test is the advection equation for temperature with 'forward' differencing in time and 'centred' differencing in space, under an assumption that errors in temperature and wind at neighbouring grid-points are statistically uncorrelated with one another. An analytical expression has been derived for its growth as a function of wind speed, time step, grid distance, temperature gradient and random errors in wind and temperature. The analytical result has also been verified by numerical experiment. With a grid of 200 n. miles and time step of 20 min., in a wind regime of 20 kt, the standard deviation of initial random error doubles itself in about 12 days.

1. Introduction

Charney (1965) has established that the period of a reliable forecast may not exceed two weeks. He carried out numerical experiments with the model of Mintz (1964). The model equations were integrated for a period of 284 days for a constant position of the Sun. The r.m.s. temperature errors were calculated for an initial random error of 1°C introduced on the 234th day of the forecast. The increase in the errors was followed over a period of 30 days. By the end of the month the error had increased to 10°C. In his case the increase of random error with time was followed by computational experiment. The purpose of this note is to estimate the increase in error from general considerations.

2. Statistical method

Let a quantity C be calculated from measured or derived quantities A and B by the relationship —

$$C = A \times B$$

Knowing the standard deviations of A and B due to errors in measurement, it is desired to estimate the standard deviation of the consequent error in C. We assume that the errors in A and B are independent of one another. Let $\bar{A}$, $\bar{B}$ and $\bar{C}$ be the correct values of A, B and C and a, b and c their errors. σ is the standard deviation due to errors and is used in that sense throughout this note.

$$\begin{aligned}\sigma_c^2 &= \overline{\left[(\bar{A}+a)\,(\bar{B}+b) - \bar{A}\,\bar{B} \right]^2} \\ &= \overline{\left[\bar{A}b + \bar{B}a + ab \right]^2} \\ &= (\bar{A})^2\,\overline{b^2} + (\bar{B})^2\,\overline{a^2} + \overline{a^2b^2} + 2\bar{A}\bar{B}\,\overline{ab} + 2\,\bar{A}\,\overline{ab^2} + 2\bar{B}\,\overline{a^2b}\end{aligned}$$

As a and b are uncorrelated $\overline{ab} = 0$. Assuming that the frequency distributions of a and b are normal and independent of one another, the probability of simultaneous occurrence of certain values of a and b will be the product of the probabilities of their separate occurrences. Hence,

$$\overline{ab^2} = \int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} \frac{1}{\sigma_a\,\sigma_b\,2\pi}\,\exp\left[-\frac{a^2}{2\sigma_a{}^2}\right]\exp\left[-\frac{b^2}{2\sigma_b{}^2}\right] ab^2\,da\,db$$

$$= \frac{1}{2\pi\,\sigma_a\sigma_b}\int_{-\infty}^{+\infty}\exp\left[-\frac{a^2}{2\sigma_a^2}\right]a\,da\int_{-\infty}^{+\infty}\exp\left[-\frac{b^2}{2\sigma_b^2}\right]b^2\,db$$

$$= 0\times\sigma_b^2 = 0$$

Similarly $\overline{a^2 b} = 0$

$$\overline{a^2b^2} = \int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty}\frac{1}{2\pi\,\sigma_a\sigma_b}\exp\left[-\frac{a^2}{2\sigma_a^2}\right]\exp\left[-\frac{b^2}{2\sigma_b^2}\right]a^2b^2\,da\,db$$

$$= \int_{-\infty}^{+\infty}\frac{a^2}{\sigma_a\sqrt{2\pi}}\exp\left[-\frac{a^2}{2\sigma_a^2}\right]da\int_{-\infty}^{+\infty}\frac{b^2}{\sigma_b\sqrt{2\pi}}\exp\left[-\frac{b^2}{2\sigma_b^2}\right]db$$

$$= \sigma_a^2\,\sigma_b^2$$

*Hence $$\sigma_c^2 = (\overline{A})^2\sigma_b^2 + (\overline{B})^2\sigma_a^2 + \sigma_a^2\sigma_b^2 \qquad (1)$$

It is well known that if, $M = P + Q$:

$$\sigma_m^2 = \sigma_p^2 + \sigma_q^2 \qquad (2)$$

if p and q are uncorrelated. p, q and m are respectively errors of P, Q and M. By the use of equations (1) and (2) the total standard deviation due to errors can be computed for any equation.

3. Errors in temperature forecasts

Neglecting diabatic heating and assuming that motion is dry adiabatic, the rate of change of temperature is given by—

$$-\frac{\partial T}{\partial t} = u\frac{\partial T}{\partial x} + v\frac{\partial T}{\partial y} + w(\Gamma_d - \Gamma) \qquad (3)$$

Let G represent grid length in the horizontal (Fig. 1). $\partial T/\partial x$ at the point O is calculated between the values of T at points 1 and 2. $\partial T/\partial y$ is between the values at points 3 and 4. It will be assumed that errors in temperature or wind at points 1, 2, 3 and 4 are uncorrelated with one another.

$$\dagger\ \sigma^2_{\partial T/\partial x} = \frac{1}{(2G)^2}\left[\sigma^2_{T_1} + \sigma^2_{T_2}\right]$$

$$= \sigma_T^2/2G^2$$

as $\sigma^2_{T_1} = \sigma^2_{T_2} = \sigma_T^2$

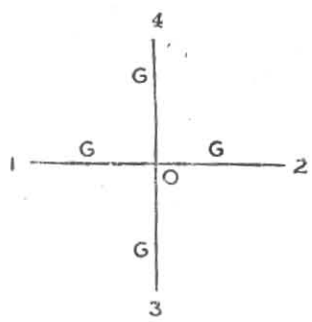

Fig. 1

By equation (1),

$$\sigma^2_{u.\partial T/\partial x} = \overline{u}^2\sigma^2_{\partial T/\partial x} + \frac{\overline{(\triangle T)_x^2}}{4G^2}\sigma_u^2 + \sigma^2_{\partial T/\partial x}\sigma_u^2$$

$$= \frac{u^2\sigma_T^2}{2G^2} + \frac{(\triangle T)_x^2}{4G^2}\sigma_u^2 + \frac{\sigma_T^2\sigma_u^2}{2G^2} \qquad (4)$$

$(\triangle T)_x$ is the temperature difference over the grid interval $2G$ in the X-direction.

---

* Shri K. N. Rao has pointed out that this formula is valid for a more generalized distribution as well

† σ is used with reference to errors in equations and does not refer to synoptic variations

Similarly,

$$\sigma^2_{v.\partial T/\partial y} = \frac{v^2 \sigma_T^2}{2\,G^2} + \frac{(\triangle T)_y^2}{4\,G^2}\,\sigma_v^2 + \frac{\sigma_T^2\,\sigma_v^2}{2\,G^2} \tag{5}$$

We may put $\sigma_u = \sigma_v = \sigma_U$ and $(\triangle T)_x = (\triangle T)_y = (\triangle T)$*

Hence,

$$\sigma^2_{u.\partial T/\partial x} = \frac{u^2 \sigma_T^2}{2\,G^2} + \frac{(\triangle T)^2 \sigma_U^2}{4\,G^2} + \frac{\sigma_T^2\,\sigma_U^2}{2\,G^2} \tag{6}$$

$$\sigma^2_{v.\partial T/\partial y} = \frac{v^2 \sigma_T^2}{2\,G^2} + \frac{(\triangle T)^2 \sigma_U^2}{4\,G^2} + \frac{\sigma_T^2\,\sigma_U^2}{2\,G^2} \tag{7}$$

$$\sigma^2_{\Gamma} = \sigma^2_{\partial T/\partial Z} = \tfrac{1}{2}\,\sigma_T^2 \,/\, G_Z^2$$

where G_Z is half the vertical interval used to calculate the lapse rate.

$$\begin{aligned}\sigma^2_{w(\Gamma_d-\Gamma)} &= w^2\,\sigma^2_{(\Gamma_d-\Gamma)} + (\Gamma_d-\Gamma)^2\,\sigma_w^2 + \sigma_w^2\,\sigma^2_{(\Gamma_d-\Gamma)} \\ &= \frac{w^2\,\sigma_T^2}{2\,G_Z^2} + (\Gamma_d-\Gamma)^2\,\sigma_w^2 + \frac{\sigma_w^2\,\sigma_T^2}{2\,G_Z^2} \end{aligned} \tag{8}$$

From equation (3),

$$\sigma^2_{\partial T/\partial t} = \sigma^2_{u.\partial T/\partial x} + \sigma^2_{v.\partial T/\partial y} + \sigma^2_{w(\Gamma_d-\Gamma)} \tag{9}$$

assuming no correlation between different terms in equation (3).

Substituting from equations (6), (7) and (8) in Eq. (9),

$$\begin{aligned}\sigma^2_{\partial T/\partial t} &= \frac{u^2 \sigma_T^2}{2\,G^2} + \frac{(\triangle T)^2 \sigma_U^2}{4\,G^2} + \frac{\sigma_T^2\,\sigma_U^2}{2\,G^2} + \\ &\quad + \frac{v^2 \sigma_T^2}{2\,G^2} + \frac{(\triangle T)^2 \sigma_U^2}{4\,G^2} + \frac{\sigma_T^2\,\sigma_U^2}{2\,G^2} + \\ &\quad + \frac{w^2\,\sigma_T^2}{2\,G_Z^2} + (\Gamma_d-\Gamma)^2\,\sigma_w^2 + \frac{\sigma_w^2\,\sigma_T^2}{2\,G_Z^2} \\ &= \frac{|\vec{V}|^2\,\sigma_T^2}{2\,G^2} + \frac{(\triangle T)^2\,\sigma_U^2}{2\,G^2} + \frac{\sigma_T^2\,\sigma_U^2}{G^2} + \\ &\quad + \frac{w^2\,\sigma_T^2}{2\,G_Z^2} + (\Gamma_d-\Gamma)^2\,\sigma_w^2 + \frac{\sigma_w^2\,\sigma_T^2}{2\,G_Z^2} \end{aligned} \tag{10}$$

where, $|\vec{V}|^2 = u^2 + v^2$

In areas of marked vertical ascent, terms involving w or σ_w may not be of much lower order of magnitude than those with $|\vec{V}|$ or σ_U, as $G_Z << G$ though $w << |\vec{V}|$. But in large areas $w \approx 0$ and we shall neglect the terms involving w or σ_w in equation (10). It should be remembered that in areas of convection $\sigma^2_{\partial T/\partial t}$ will be more than the estimate made with this simplification.

$$\sigma^2_{\partial T/\partial t} = \frac{|\vec{V}|^2\,\sigma_T^2}{2\,G^2} + \frac{(\triangle T)^2\,\sigma_U^2}{2\,G^2} + \frac{\sigma_T^2\,\sigma_U^2}{G^2} \tag{11}$$

*Formula can be developed without these simplifying assumptions; but it will be more elaborate

If T_0 is the initial temperature and T_1 after one computation with a time interval $\triangle t$,

$$T_1 = T_0 + \frac{\partial T}{\partial t} \triangle t$$

$$\sigma^2_{T_1} = \sigma^2_{T_0} + (\triangle t)^2 \sigma^2_{\partial T/\partial t} \tag{12}$$

Writing $\sigma^2_{T_0} = \sigma_T{}^2$ and substituting from (11) in equation (12),

$$\sigma^2_{T_1} = \sigma_T{}^2 + (\triangle t)^2 \left[\frac{|\vec{V}|^2 \sigma_T{}^2}{2G^2} + \frac{(\triangle T)^2 \sigma_U{}^2}{2G^2} + \frac{\sigma_T{}^2 \sigma_U{}^2}{G^2} \right]$$

$$= \sigma_T{}^2 \left[1 + \frac{(\triangle t)^2}{2G^2} \left(|\vec{V}|^2 + \frac{(\triangle T)^2 \sigma_U{}^2}{\sigma_T{}^2} + 2\sigma_U{}^2 \right) \right]$$

After the second operation in computation,

$$\sigma^2_{T_2} = \sigma^2_{T_1} \left[1 + \frac{(\triangle t)^2}{2G^2} \left(|\vec{V}_1|^2 + \frac{(\triangle T_1)^2 \sigma^2_{U_1}}{\sigma^2_{T_1}} + 2\sigma^2_{U_1} \right) \right]$$

$$= \sigma_T{}^2 \left[1 + \frac{(\triangle t)^2}{2G^2} \left(|\vec{V}|^2 + \frac{(\triangle T)^2 \sigma_U{}^2}{\sigma_T{}^2} + 2\sigma_U{}^2 \right) \right] \times$$

$$\times \left[1 + \frac{(\triangle t)^2}{2G^2} \left(|\vec{V}_1|^2 + \frac{(\triangle T_1)^2 \sigma^2_{U_1}}{\sigma^2_{T_1}} + 2\sigma^2_{U_1} \right) \right] \tag{13}$$

The subscript 1 indicates values prevailing after the first computation. A simplification may be made that $|\vec{V}_1| \approx |\vec{V}|$, $(\triangle T)_1 \approx (\triangle T)$ and $\sigma_{U_1} \approx \sigma_U$. We shall also substitute σ_T for σ_{T_1} in the second square bracket on the right hand side in equation (13). With these simplifications,

$$\sigma^2_{T_2} = \sigma_T{}^2 \left[1 + \frac{(\triangle t)^2}{2G^2} \left(|\vec{V}|^2 + \frac{(\triangle T)^2 \sigma_U{}^2}{\sigma_T{}^2} + 2\sigma_U{}^2 \right) \right]^2 \tag{14}$$

Similarly after n steps in computation —

$$\sigma^2_{T_n} = \sigma_T{}^2 \left[1 + \frac{(\triangle t)^2}{2G^2} \left(|\vec{V}|^2 + \frac{(\triangle T)^2 \sigma_U{}^2}{\sigma_T{}^2} + 2\sigma_U{}^2 \right) \right]^n \tag{15}$$

$$\text{If } \epsilon = \frac{(\triangle t)^2}{2G^2} \left(|\vec{V}|^2 + \frac{(\triangle T)^2 \sigma_U{}^2}{\sigma_T{}^2} + 2\sigma_U{}^2 \right) \tag{16}$$

$$\sigma^2_{T_n} = \sigma_T{}^2 (1+\epsilon)^n \tag{17}$$

A limiting period of reliable forecast may be defined as $N \triangle t$ when after N operations the standard deviation of the error, σ_{T_N} is a multiple K of the initial error. N is then given by —

$$\sigma^2_{T_N} = K^2 \sigma_T{}^2 = \sigma_T{}^2 (1+\epsilon)^N \tag{18}$$

$$N = 2 \log K / \log (1+\epsilon) \tag{19}$$

4. Discussion

4.1. The term $(\triangle t)^2 |\vec{V}|^2 / 2G^2$ in (15) is very similar to Courant, Freidrichs and Levi rule. The other terms in ϵ show the additional effect of initial errors in temperature and wind, and the gradient of temperature in the area. The time interval and grid distance also influence the accumulation of errors. Where geostrophic wind is used, it should be possible to express σ_U in terms of σ_T. At high wind speeds the last two terms in ϵ will be less important.

Usually ϵ will be less than $0{\cdot}1$ and equation (19) may be written as —

$$N = 2 \log_e 10 \log_{10} K / \epsilon = 4{\cdot}60 \log_{10} K / \epsilon$$

$$N \triangle t = 4{\cdot}60 (\log_{10} K)(\triangle t / \epsilon) \tag{20}$$

$N\triangle t$ is tabulated below for some plausible values of the parameters in ϵ.

Values of $N\triangle t$ (hrs)

$\sigma_T = 1°\text{C}$ $\quad \sigma_U = 5$ kt $\quad \triangle T = 5°\text{C}$ $\quad K = 2$

G (n.miles)	$\vert\vec{V}\vert$ (kt)	$\triangle t = 1$ hr: ϵ	$\triangle t = 1$ hr: $N\triangle t$ (hrs)	$\triangle t = 20$ minutes: ϵ	$\triangle t = 20$ minutes: $N\triangle t$ (hrs)
100	20	0·0537	26	0·0060	77
	50	0·1587	10	0·0176	26
200	20	0·0134	103	0·0015	308
	50	0·0397	35	0·0044	105

4.2. The above table shows the time taken for the initial root mean square error of 1°C in temperature to increase to 2°C in the course of computations. As ϵ is made smaller, forecasting range increases. With a grid of 200 n. miles and time step of 20 minutes it takes from 4 to 12 days depending upon the speed of the wind regime. This broadly corresponds to Charney's result. But as the horizontal grid length is decreased, the time step will have to be reduced to attain similar results. At higher wind speeds greater accuracy in wind and temperature measurements (as lower values of σ_T and σ_U) is less effective than smaller time steps in increasing the forecasting range. It may, therefore, suffice to keep the accuracy of measurements of wind and temperature within such limits as not to suppress the synoptic scale (grid-length) variations and decrease the time step to increase the limiting period of reliable forecast. Only errors arising from errors of observations have been discussed here.

4.3. In deriving equation (15), $\vec{V}$, $(\triangle T)$, σ_T and σ_U have been kept constant. This is to understand the multiplication of errors in different regimes. In any actual operation, random errors can be followed at selected points by applying the equation with appropriate values.

4.4. Some correlation may develop between errors at neighbouring grid-points. But the treatment given here may be regarded as a first approximation.

5. Cause of increase in errors

It is shown below that the increase of the initial random errors of temperature in the forecasted temperature values is not due to numerical method of integration but is inherent in the equation itself. For this purpose, we shall write equation (3) as —

$$-\frac{\partial T}{\partial t} = U\frac{\partial T}{\partial x} \tag{21}$$

Let us assume that U and $\partial T/\partial x$ are constants but $\partial T/\partial x$ is subject to an error because of random error in the measurement of temperature, T. Equation (21) can now be integrated over a time interval $n\triangle t$ analytically and

$$T_{(n\triangle t)} = T_0 - \left[U\frac{\partial T}{\partial x} \right] n.\triangle t \tag{22}$$

T_0 is the initial temperature and $T_{(n\triangle t)}$ after a time interval $n\triangle t$. As $\partial T/\partial x$ is subject to an error because of the random error in the observed value of T,

$$\sigma^2_{T_{(n\triangle t)_a}} = \sigma^2_{T_0} + (U\,n\triangle t)^2\,\sigma^2_{\partial T/\partial x} \tag{23}$$

Substituting
$$\sigma^2_{\partial T/\partial x} = \sigma_T^2 / 2G^2 \quad \text{and} \quad \sigma^2_{T_0} = \sigma_T^2$$

$$\begin{aligned}\sigma^2_{T_{(n\triangle t)_a}} &= \sigma_T^2 + (U\,n\triangle t)^2 \frac{\sigma_T^2}{2G^2} \\ &= \sigma_T^2 + \frac{U^2 (n\triangle t)^2 \sigma_T^2}{2G^2} \\ &= \sigma_T^2 \left(1 + \frac{n^2 (U\triangle t)^2}{2G^2} \right) \end{aligned} \tag{24}$$

This gives the error in the temperature obtained by the analytical solution. If the temperature had been calculated by n operations each of an interval $\triangle t$, the error in the final temperature given by equation (15) will be —

$$\sigma^2_{T_{(n\triangle t)_f}} = \sigma_T^2 \left(1 + \frac{(\triangle t)^2 U^2}{2G^2} \right)^n \tag{25}$$

Let us compare the errors in the two cases when $(\triangle t)^2 U^2 / 2G^2 = 0{\cdot}01$ and $n = 100$.

$$\begin{aligned}\sigma^2_{T_{(n\triangle t)_a}} &= \sigma_T^2 \left[1 + 0{\cdot}01\,(100)^2 \right] \\ &= 101\,\sigma_T^2 \\ \sigma^2_{T_{(n\triangle t)_f}} &= \sigma_T^2 \left[1 + 0{\cdot}01 \right]^{100} \\ &= 2{\cdot}7\,\sigma_T^2 \end{aligned}$$

Thus the error in the solution by the analytical method arrived at in one large time step is very much more than that computed by a large number of very small time steps. This shows that amplification of initial random error in the final product of integration is inherent in the equation itself.

6. Numerical experiment

A numerical computational experiment was tried to test whether the growth of errors is according to equation (17). Advection along a latitude circle was simulated by the equation $-\partial T/\partial t = U\,.\,\partial T/\partial x$. The temperature after one time step of $\triangle t$ is given by —

$$T_1(I) = T_0(I) - \frac{u\triangle t}{2G} \left[T\,(I+1) - T(I-1) \right]$$

I indicates the number of the grid point and $T_0\,(I)$ is the initial temperature and $T_1\,(I)$ after computation. For the next operation the computed $T_1\,(I)$ becomes $T_0\,(I)$. Twenty grid points at $G = 100$ n. miles and $u = 50$ kt and $\triangle t = 1$ hr were used. Initial temperature values were assigned to the twenty grid points such that their mean was zero while the standard deviation was 1°C as representing random errors. After ten operations the mean temperature at the grid points was zero while the standard deviation rose to 1·7°C, illustrating increase in random errors with repeated computations. The value calculated by using equation (17) with the appropriate value of ϵ (0·125 in this case) works out to 1·8°C, which tallies well within the above computed value.

REFERENCES

Charney, J. — 1965 — A global observation experiment. Rep. on the Int. Symp. on Large Scale Processes, Moscow.

Mintz, T. — 1964 — Very long-term Global Integration of the Primitive Equations of Atmospheric Motion. U.C.L.A. Dep. of Met., Contr. No. 111, July 1964.