SMOOTHING AND FILTERING OF TIME SERIES AND SPACE FIELDS

J. Leith Holloway, Jr.

U. S. Weather Bureau, Washington D.C.

1. Introduction

Time and space smoothing are widely used in the study of geophysical problems. For example, the drawing of smooth isolines through data plotted on a map of some geophysical variable is a type of space smoothing. The computation of consecutive monthly mean values of a series of measurements and the use of long-lag instruments for suppressing rapid fluctuations in readings are examples of time smoothing. The purpose of this paper is to give the reader a better understanding of just what these smoothing methods really accomplish so as to provide a rational basis for the selection of any particular smoothing method. Smoothing will be shown to be a special type of filtering, and the analysis will therefore be extended to cover numerical filters of all types.

Much of the information in this paper is not new, but is scattered rather widely in the mathematical, statistical, and scientific literature. An attempt is made here to combine all pertinent information on smoothing and filtering in a concise, simple, understandable form for the benefit of geophysicists concerned with these problems. Since this paper is primarily a survey of the field, rigorous mathematical derivations of the equations are not given.

2. Time Smoothing and Filtering

In statistical work a series of data arranged chronologically is commonly called a time series. For example, a series of daily mean temperatures in order of date is a time series. Generally the data in time series are equally spaced in time, and therefore all time series discussed in this paper will have equally spaced data. The variations in the data with time may be relatively smooth and orderly or may be rather complex and without apparent pattern.

It is convenient to consider that the variations in time are produced by superimposed sinusoidal waves of various amplitudes, frequencies, and phases. The Fourier theorem states that no matter how complicated the fluctuations in the data may be they can be accounted for by the superposition of a number of simple component sinusoidal waves [1]. The amplitudes, frequencies, and phases of these waves are generally changing constantly with time. The exception to this is where fixed cyclic processes such as diurnal and annual influences tend to induce waves of constant frequency (one cycle per day or per year) into the data.

The purpose of time smoothing is to attenuate the amplitudes of high-frequency waves in the data without significantly affecting the low-frequency components. The attenuation is roughly proportional to frequency. Above some high frequency, depending upon the properties of the smoothing method used, the attenuation is complete for all practical purposes. The assumption upon which the use of smoothing is justified is that high-frequency oscillations in the data are either random error ("noise") or are of no significance to the particular type of evaluation of the data to be carried out after the smoothing. Smoothing thus enables one to concentrate on the low frequencies without the distraction caused by high-frequency noise and other irrelevant fluctuations. A smoothed value of an observation in a time series is merely an estimate of what the value would be if noise and other undesired high frequencies were not present in the series.

Smoothing is a special case of the broader general process of filtering, a concept brought into the field of time series analysis from electrical engineering. An electrical filter, such as a simple resistance-capacitance network, separates various sinusoidal components of an "electrical time series" according to frequency. An electrical analog of a numerical time series is a continuously varying voltage or current—usually referred to as a "signal" in electrical engineering. Electrical filters can be designed to pass only low frequencies of the signal while attenuating or eliminating high frequencies. This type of filter is commonly called a "low-pass filter." Thus smoothing of a numerical time series is analogous to low-pass filtering of an electrical signal. Filters can also be designed to pass only high frequencies and attenuate low frequencies ("high-pass filter"). Finally other filters pass a band of intermediate frequencies ("band-pass filter") and attenuate both very low and very high frequencies in the input signal. Numerical band-pass and high-pass filters will be described in Section 7.

3. Equalization, Pre-emphasis, and Inverse Smoothing

Two other terms which may be borrowed from electrical engineering and applied to time series analysis are "equalization" and "pre-emphasis." Equalization is the process of restoring the original balance of amplitudes of sinusoidal waves in a signal which has previously been altered by filtering or pre-emphasis. Pre-emphasis is the amplification before transmission of certain bands of frequencies in a signal above a standard amplification. For example, the high audio-frequencies are pre-emphasized before transmission by a frequency-modulation transmitter to override high audio-frequency noise introduced by the transmitter, the atmosphere, and the receiver. This can be done since high audio-frequencies are generally of low amplitude in the original signal. At the receiver the correct balance of frequencies in the audio signal is restored by an equalization filter ("*de*-emphasis" filter) which in this case is merely a low-pass filter having an attenuation which is the inverse of the frequency characteristics of the pre-emphasis at the transmitter.

A second example of electrical equalization is the compensation for the attenuation of high frequencies in a long transmission line. In this case the line itself acts as a low-pass filter. The original balance of frequencies in the signal is approximately restored at the end of the line by an equalization amplifier which has greater amplification at high than at low frequencies. As long as some fraction of the signal at a given frequency is received above the noise in the line, the original relative amplitude at this frequency can be approximately restored. Of course, in order to equalize properly it is necessary to know how much the line attenuates the signal at each frequency.

It follows that it should be possible to equalize a time series previously smoothed provided that the characteristics of the smoothing were known. Later in this paper, a method for performing this equalization numerically will be given. The term "inverse smoothing" will be used for the numerical equalization used for accentuating high frequencies in respect to low frequencies in order to restore the original balance of frequencies to a smoothed time series. A good deal of work on inverse smoothing already has been done in several scientific fields, notably in radio-astronomy [2–10].

4. Smoothing and Filtering Functions

Smoothing of a time series is performed by a type of numerical or mathematical operator which will be termed a "smoothing function" in this paper. Generally the smoothing function consists of a series of fractional values, called weights. In this paper, the term "filtering function" will be applied to operators which perform filtering of time series other than smoothing. Filtering functions also generally consist of various weights similar to those of smoothing functions. The weights determine in what proportion each observation in the time series contributes to the estimate of the smoothed or filtered value. In the process of smoothing or filtering, successive observations in a time series are cumulatively multiplied by these weights.

The smoothed or filtered value corresponding to the observation x_t in the time series is computed from observations x_{t-n} through x_{t+m} by the following linear equation:

$$\bar{x}_t = \sum_{k=-n}^{m} w_k x_{t+k} = w_{-n}x_{t-n} + \cdots + w_{-1}x_{t-1} + w_0 x_t + w_1 x_{t+1} + \cdots + w_m x_{t+m} \tag{4.1}$$

where w_k is a particular weight in the smoothing or filtering function. The weight w_0, which is multiplied by the observation x_t, will be termed the principal weight in this paper. The principal weight is the central one in the case of the equally-weighted running mean where all weights are identical and equal to $1/N$ where N is the number of observations used in computing the mean.

The use of smoothing and filtering functions is illustrated in Fig. 1. The weights of the smoothing function in the block shown in this figure are cumulatively cross-multiplied by the adjacent values in the time series and the resulting product is entered opposite the time series value multiplied by the principal weight. Then the smoothing function is moved down one time increment (data interval) along the time series and the cross-multiplication is repeated to obtain a second smoothed value. This process is repeated until the lowest weight in the smoothing function reaches the end of the series.

The sum of the weights of a smoothing or filtering function determines the ratio of the mean of the smoothed or filtered series to the mean of the original series assuming that these means are computed over periods long enough to insure stable results. In smoothing it is generally desired to leave the mean of the series unchanged, and consequently the sum of the weights of most smoothing functions is made equal to unity. With some filters to be discussed later it is not necessary to preserve the mean of the series, and in these cases the sum of the weights may be different from unity.

For theoretical evaluation of smoothing and filtering functions it is often useful to consider that these functions are continuous rather than composed of discrete weights. The continuous analytic form of the smoothing or filtering function will be some function $w(t)$ describing the envelope of the discrete weights. The time origin here is the time of the principal weight; namely, the time of the observation for which a smoothed value is being computed. The area under the continuous

		Time Series			Time Series		
Smoothing Function	.016 ×	28			28		
	+.094 ×	23		.016 ×	23		
	+.234 ×	21		+.094 ×	21		Smoothed
Principal → Weight	+.312 ×	24 →	22.3	+.234 ×	24	22.3	Values
	+.234 ×	22		+.312 ×	22 →	21.4	
	+.094 ×	20		+.234 ×	20		
	+.016 ×	17		+.094 ×	17		
		18		+.016 ×	18		
		29			29		
		35			35		

FIG. 1. Illustration of smoothing a time series by means of a typical smoothing function.

smoothing or filtering function corresponds to the sum of the weights of a discrete-valued function. Thus the ratio R of the mean of the filtered series to that of the original series is

$$R = \int_{-\infty}^{\infty} w(t)\, dt \tag{4.2}$$

in the case of continuous functions.

5. Frequency Response of Smoothing Functions and Other Filters

The ratio of the amplitude of a wave of a given frequency in the time series after filtering to the original amplitude before filtering is the frequency response of the filter at this frequency. The frequency response is a function of frequency. For example, the value of the frequency response of a smoothing function is near unity at low frequencies and near zero at high frequencies.

The frequency response of a mathematical filter such as a smoothing function can be derived by determining the effect of this filter on a unit-amplitude sinusoidal wave of frequency f. This wave may be represented by a unit vector rotating about the origin with angular velocity of $2\pi f$ in the complex plane. The projections of this unit vector on the real and imaginary axes define the phase of the unit amplitude wave at any time. The projection of this vector on the real axis is $\cos(2\pi ft)$, and on the imaginary axis, $\sin(2\pi ft)$, where t is time.[1] After smoothing or some other form of filtering, the modified wave can be represented by another vector rotating about the origin in the complex plane with the same angular velocity but having a magnitude and phase angle generally different from that of the original unit vector. The amplitude of the smoothed or filtered wave and its phase angle at any time is determined by the instantaneous projections of this modified vector on the real and imaginary axes. These projections can be shown to be merely the weighted mean of the projections of the original unit vector on these axes averaged over the interval during which the smoothing or filtering function is operating, hereafter to be referred to as the filtering interval.[2] The weighted mean projection of the unit vector on the real axis averaged over the filtering interval is obtained by summing the products of each weight and the corresponding value of $\cos(2\pi ft)$ at time t. The weighted mean projection on the imaginary axis is obtained in a similar manner but with the sine substituted for the cosine. In the case of continuous analytic forms of the smoothing and filtering functions these mean projections must be computed by integration.

The ratio of the magnitude of the modified vector to the magnitude of the unit vector (unity) is the frequency response of the filter. Thus, the magnitude of the modified vector is the frequency response. This magnitude is the absolute value of the complex quantity $R(f)$ given by

$$\begin{aligned} R(f) &= \int_{-\infty}^{\infty} w(t) \exp(2\pi i f t)\, dt \\ &= \int_{-\infty}^{\infty} w(t) \cos(2\pi f t)\, dt + i \int_{-\infty}^{\infty} w(t) \sin(2\pi f t)\, dt \end{aligned} \tag{5.1}$$

The cosine term of (5.1) is a weighted time mean projection of the unit vector on the real axis, and the sine term is the simultaneous weighted time mean projection of this vector on the imaginary axis. This equation will be recognized as the inverse Fourier transform of $w(t)$. Therefore, the inverse Fourier transform of the smoothing or filtering function is the filter's frequency response function. The absolute value of $R(f)$ is the

[1] The arguments in this paragraph are valid for any time origin in addition to that used in the last section.

[2] For example, the filtering interval in Fig. 1 is seven data intervals.

square root of the sum of the squares of the real and imaginary parts of $R(f)$; namely

$$|R(f)| = ([Re\{R(f)\}]^2 + [Im\{R(f)\}]^2)^{1/2} \tag{5.2}$$

The angle between the original and the modified vector is the phase shift which the filtering function $w(t)$ produces at frequency f. This angle ϕ is given by

$$\phi = \tan^{-1} [Im\{R(f)\}/Re\{R(f)\}] \tag{5.3}$$

For smoothing and filtering functions having $(n + m + 1)$ discrete weights the frequency response is computed by the following form of equation (5.1)

$$R(f) = \sum_{k=-n}^{m} w_k \cos (2\pi fk) + i \sum_{k=-n}^{m} w_k \sin (2\pi fk) \tag{5.4}$$

where units of frequency f are cycles per data interval. The absolute value of $R(f)$ obtained from the above equation is again computed by equation (5.2).

It is desirable that smoothing and filtering functions not shift the phase of waves of any frequency. The phase shift angle can be made equal to zero by requiring that the numerator of the argument in equation (5.3) (namely, the specified weighted time mean projection of the unit vector on the imaginary axis) be zero. This in turn can be accomplished by requiring that the function $w(t)$ be even (namely, that $w(-t) = w(t)$),[3] for if $w(t)$ is even, the terms containing the sines in equations (5.1) and (5.4) are zero, and $R(f)$ is a pure real quantity computed by

$$R(f) = \int_{-\infty}^{\infty} w(t) \cos (2\pi ft)\, dt = 2 \int_0^{\infty} w(t) \cos (2\pi ft)\, dt \tag{5.5}$$

for continuous $w(t)$ functions or

$$R(f) = \sum_{k=-n}^{n} w_k \cos (2\pi fk) = w_0 + 2 \sum_{k=1}^{n} w_k \cos (2\pi fk) \tag{5.6}$$

for smoothing and filtering functions having $(2n + 1)$ discrete weights.

The frequency response of an equally-weighted running mean of $(2n + 1)$ consecutive terms may be computed from equation (5.6) be-

[3] In the case of discrete-weighted smoothing and filtering functions this condition is that $w_{-k} = w_k$. The smoothing function in Fig. 1 is even.

cause $w_{-k} = w_k$; in fact, in this case every weight $w_k = 1/(2n + 1)$. Equation (5.6) gives

$$(5.7)\quad \begin{aligned} R(f) &= w_0 + 2w_1 \cos(2\pi f\Delta t) + 2w_2 \cos(4\pi f\Delta t) + \cdots \\ &\qquad\qquad + 2w_n \cos(2n\pi f\Delta t) \\ &= (2n+1)^{-1}[1 + 2\cos(2\pi f\Delta t) + 2\cos(4\pi f\Delta t) \\ &\qquad\qquad + \cdots + 2\cos(2n\pi f\Delta t)] \end{aligned}$$

where Δt is the data interval. The use of Δt here avoids the requirement that the units of frequency be cycles per data interval; the units of frequency in (5.7) are thus cycles per particular unit of time used.

A convenient approximation of the frequency response of the equally-weighted running mean can be computed from equation (5.5) by using the analytic form of the envelope of the weights, $w(t)$; namely,

$$(5.8)\quad w(t) = \begin{cases} 1/T, & |t| \le T/2 \\ 0, & |t| > T/2 \end{cases}$$

where T is the filtering interval. Equation (5.5) gives

$$(5.9)\quad R(f) = 2\int_0^{T/2} T^{-1} \cos(2\pi f t)\, dt = (\pi f T)^{-1} \sin(\pi f T)$$

Equation (5.9) gives a very accurate approximation of the frequency response of equally-weighted running means.[4]

The exact frequency response of an equally-weighted running mean having five weights of one-fifth each is shown in Fig. 2 as a dotted line. This response is computed from equation (5.7) assuming a data interval of $T/5$ so as to make the filtering interval equal to T. For comparison, the approximate frequency response for this type of running mean computed from equation (5.9) is shown in this figure as a solid line. Notice that the agreement between these two curves is quite good; the agreement would be even better with running means of more than five terms.

The physical meaning of the negative response at some frequencies in Fig. 2, is that the input waves of these frequencies are reversed in polarity in addition to being changed in amplitude. A reversal of polarity of a wave means that its maxima are changed into minima and vice versa.[5] Positive and negative values of the frequency response above the frequency of the first zero response point are undesirable because they will introduce many unwanted and misleading high-frequency ripples into the smoothed output.

A method for suppressing or eliminating these undesirable responses

[4] By L'Hopital's rule this function has the value unity at $f = 0$.

[5] Reversal of the polarity of a wave corresponds to a reversal of the direction of rotation of the vector representing this wave. This is equivalent to a 180-degree phase shift of the wave.

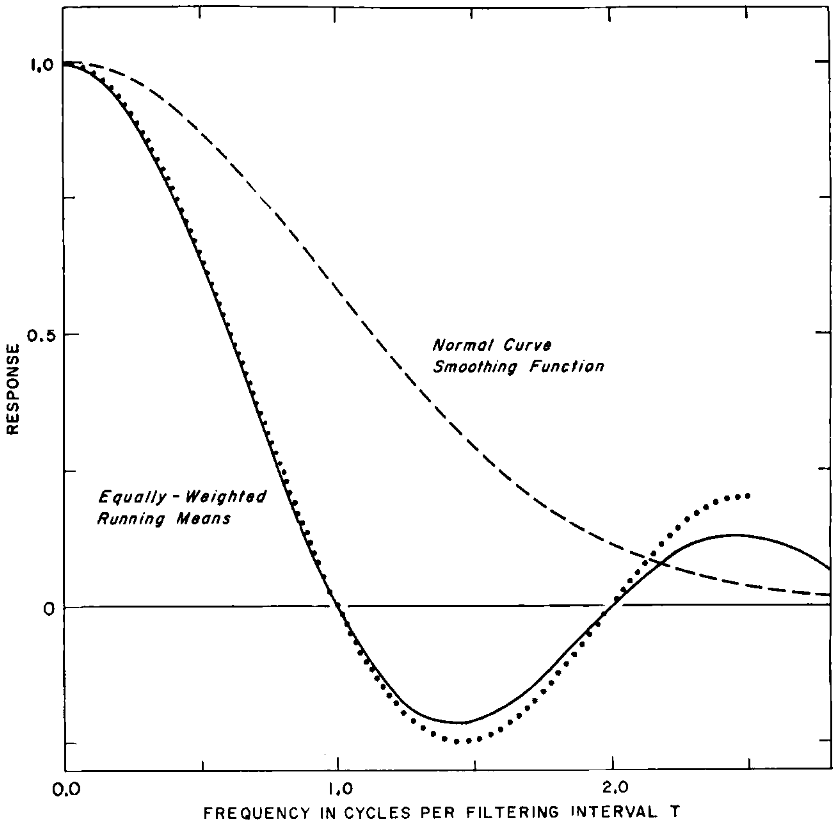


FIG. 2. Frequency responses of equally-weighted running means and of a normal probability curve smoothing function. Solid line is response of the running mean computed from equation (5.9). Dotted line is response of the five-term running mean computed from equation (5.7). Dashed line is response of normal curve smoothing function having $\sigma = T/6$ computed from equation (5.11).

is to provide a smoothing function having weights decreasing in magnitude outward from the principal weight. For example, the smoothing function weights may be made proportional to the ordinates of the normal probability curve. A continuous analytic form of this smoothing function is

$$w(t) = (2\pi\sigma^2)^{-1/2} \exp(-t^2/2\sigma^2) \tag{5.10}$$

The area under the above function is unity so that the mean of the smoothed series will be conserved. The frequency response of this smoothing function is obtained from equation (5.5); namely,

$$R(f) = \exp(-2\pi^2\sigma^2 f^2) \tag{5.11}$$

The frequency response given by equation (5.11) for the normal curve smoothing function decreases smoothly with increasing frequency and asymptotically approaches zero. Thus it avoids the negative values of response exhibited by the equally-weighted running mean. Although zero response is theoretically never reached, for practical purposes this smoothing function has a finite "cutoff frequency." The "cutoff frequency" of a smoothing function will be defined as the lowest frequency at which the response reaches zero for all practical purposes and remains zero for all higher frequencies. Thus in cases like this where the response function approaches zero asymptotically the cutoff frequency can be taken as that frequency where the response drops to some arbitrarily chosen low value, say one per cent. The cutoff frequency of the normal curve smoothing function is controlled by the value of the parameter σ in equation (5.10).

The response of the normal curve smoothing function having a filtering interval of T is plotted in Fig. 2 as a dashed line for comparison with that of the equally-weighted running mean having the same filtering interval. The filtering interval of a normal curve smoothing function is taken to be 6σ, for beyond 3σ from the origin the normal curve ordinates have negligible value.

The equations in this section may also be used to compute the frequency response of an exponential smoothing function; namely, one having the analytic form:

$$w(t) = \begin{cases} 0, & t > 0 \\ \lambda^{-1}e^{t/\lambda}, & t \leq 0 \end{cases} \tag{5.12}$$

where λ is the so-called time constant or lag coefficient. The area under this function is unity so that the mean of the series is unaffected by the smoothing. This type of smoothing is that which is performed by physical instruments which are viscous-damped and have constant λ and by simple, two-element resistance-capacitance low-pass electrical filters. An example of an instrument which smooths in this manner is the simple mercury-in-glass thermometer. Since this smoothing function defined above is continuous and not even,[6] equation (5.1) must be used for determining its frequency response. This equation gives

$$R(f) = (1 + 4\pi^2f^2\lambda^2)^{-1} - (2\pi if\lambda)(1 + 4\pi^2f^2\lambda^2)^{-1} \tag{5.13}$$

and from (5.2),[7]

$$|R(f)| = (1 + 4\pi^2f^2\lambda^2)^{-1/2} \tag{5.14}$$

[6] This smoothing function has zero values for $t > 0$ since no physical instrument can take future variations into account in its smoothing.

[7] Middleton and Spilhaus obtain this same result by means of differential equations [11].

Since the imaginary part of $R(f)$ is not zero owing to the unsymmetrical distribution of weighting about the origin, the smoothed series has phase error which is computed from equation (5.3) and given below

$$\phi = \tan^{-1}(-2\pi f\lambda) \tag{5.15}$$

The frequency response and phase shift for exponential smoothing is shown in Fig. 3 where λ equals unity. Notice that this frequency response also approaches zero asymptotically.

Smoothed data are often obtained by deliberately increasing the viscous damping of the measuring instrument [12] or by use of simple resistance-capacitance electrical filters. The fact that this type of smoothing introduces phase error into the data suggests that a better smoothing

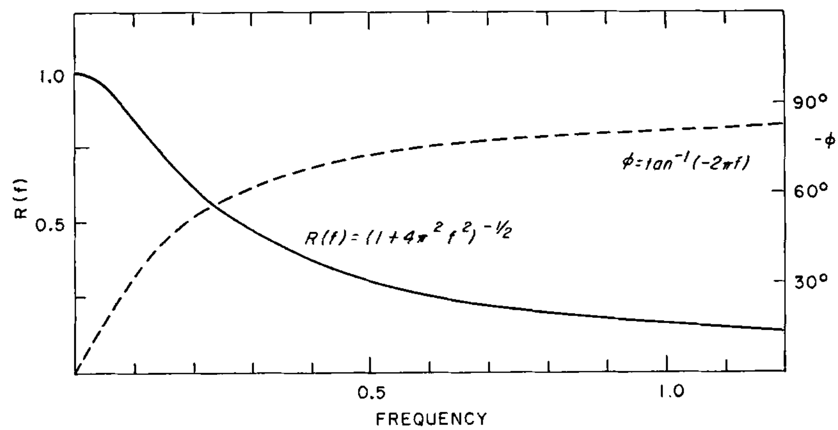


FIG. 3. Frequency response $R(f)$ and phase shift ϕ of an exponential smoothing function having a lag coefficient of unity.

procedure would be to obtain the data from the fastest response instrument available and to smooth these data later by means of smoothing functions having no phase error, such as the normal curve smoothing function. However, no matter how fast its response is, any physical instrument smooths the data to some extent. It is only necessary that the time constant of the instrument be short in comparison with the wavelength of variations in the data whose amplitudes and *phases* are desired to be recorded accurately.

The results of smoothing a time series by the three methods described above are compared in Fig. 4 where the original unsmoothed series is shown as a solid line and equally-weighted running means of five consecutive observations are connected by a dotted line. The dashed line in this figure shows the series smoothed by a discrete-weighted approximation of a normal curve smoothing function having $\sigma = 5/6$ data interval; the weights of this function are 0.03, 0.23, 0.48, 0.23, and 0.03 in that order. The frequency responses of the running means and the normal curve smoothing function are given by the dotted and dashed lines, respectively, in Fig. 2 when the filtering interval is taken to be 5 data intervals. The dot-dashed line in Fig. 4 represents the series smoothed by a discrete-weighted approximation of an exponential smoothing function having $\lambda = 2.5$ data intervals. The frequency response of this exponential filter versus frequency in cycles per data interval is obtained from Fig. 3 by multiplying the values on the abscissa by $\lambda^{-1} = 0.4$. The filtering intervals and σ's and λ's of these three smoothing functions were chosen so that the degree of smoothing of each of the three methods would be about

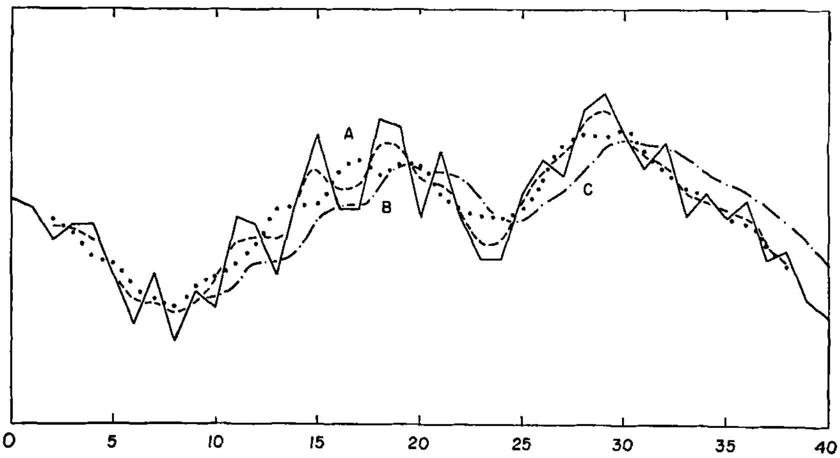


FIG. 4. Time series (solid line) and the same series smoothed by means of an equally-weighted running mean (dotted curve), by a normal curve smoothing function (dashed curve) and by exponential smoothing (dot-dashed curve).

the same. Fig. 4 illustrates the unfortunate polarity reversals effected by equally-weighted running means at some frequencies (for example, at point A) and the phase shift produced by exponential smoothing (at points B and C). The normal curve smoothing does not exhibit either of these two shortcomings.

The frequency response of many other smoothing and filtering functions may be determined from the equations in this section. When the exact response function of a particular discrete-weighted smoothing or filtering function is desired, equation (5.4) or (5.6) must be used. However, as illustrated in this section, equations (5.1) and (5.5) may be used for determining estimates of the response of such functions having a relatively large number of weights, provided the envelopes of these weights can be expressed in simple integratable analytic forms. If the smoothing or filtering function is intrinsically continuous (as is the case with exponential smoothing by an electrical filter, for example), either equation (5.1) or (5.5) will be required for computing the exact frequency response function. Other methods for computing frequency response functions will be discussed later in this paper.

6. Design of Smoothing Functions and Filters with Specified Frequency Response

The procedure in the last section may be reversed and a smoothing or filtering function $w(t)$ having a specified frequency response function of $R(f)$ may be obtained by solving the integral equation (5.1). This solution is[8]

$$w(t) = \int_{-\infty}^{\infty} R(f) \exp(-2\pi i f t)\, df \tag{6.1}$$

For an even response function $R(f)$, equation (6.1) reduces to

$$w(t) = 2\int_0^{\infty} R(f) \cos(2\pi f t)\, df \tag{6.2}$$

These equations may be recognized as those for Fourier transforms of $R(f)$.

An example of the use of equation (6.2) is the determination of the smoothing function having a flat response of unity out to some cutoff frequency f_c and zero response beyond; namely,

$$R(f) = \begin{cases} 1, & 0 \le f \le f_c \\ 0, & f > f_c \end{cases} \tag{6.3}$$

Use of equation (6.2) gives

$$w(t) = 2\int_0^{f_c} \cos(2\pi f t)\, df = (\pi t)^{-1} \sin(2\pi f_c t) \tag{6.4}$$

This smoothing function is a damped wave extending forward and backward in time from the origin. However, because the damping is rather slow, this function will often be impractical to use, since it will extend over so much of the series to be smoothed. The function can, of

[8] A wave of negative frequency is merely one of the corresponding positive frequency with reversed polarity. Ordinarily frequency is not thought of as having negative values; the above definition is given only to clarify equations such as (6.1) where frequency is allowed to take on negative values. From the above it is clear that at least the absolute value of all frequency response functions must always be even functions, for it would be contradictory for a filter to have a different response to a negative frequency than to the corresponding positive one. When $|R(f)|$ is not even, $w(t)$ will generally be a complex function (composed of both real and imaginary parts) and therefore will have no physical meaning as a filtering function.

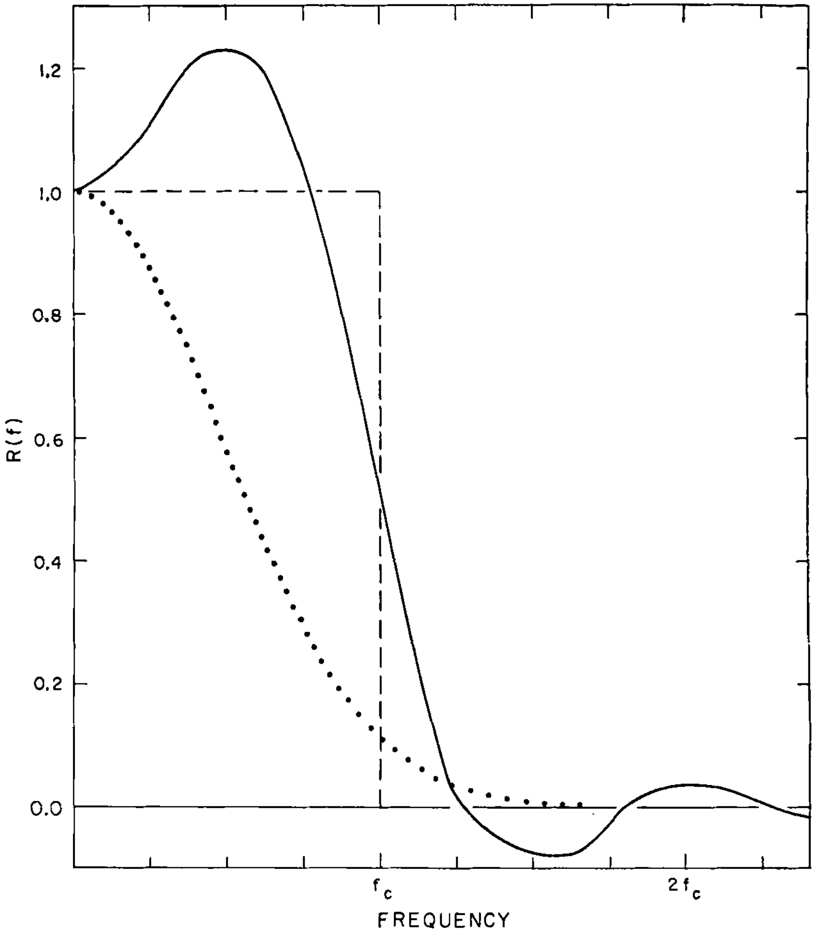


FIG. 5. Theoretical frequency response of a smoothing function having the shape of the damped wave defined by equation (6.4) (dashed line), and the actual response of this function truncated beyond the first negative lobes on either side of the central positive lobe (solid curve). The dotted curve is the response of the normal curve smoothing function having $\sigma = \frac{1}{3f_c}$.

course, be taken to be zero (truncated) at some convenient distance on each side of the origin, but this alters the response in an undesirable way, and the closer to the origin it is truncated, the less desirable the response becomes. For example, the frequency response of this function truncated at $t = \pm 1/f_c$ is shown in Fig. 5 as a solid line. This condition specifies two negative lobes on each side of the central positive weights. The desired frequency response specified by (6.3) is shown in this figure as a dashed line. Notice that the truncation causes the actual frequency response of the smoothing function to differ considerably from the desired response at most frequencies. In fact, there is actually amplification at intermediate frequencies and undesirable negative response at certain higher frequencies. For comparison, the frequency response of a normal curve smoothing function having $\sigma = \frac{1}{3f_c}$ is also shown in Fig. 5 as a dotted line. The filtering intervals of these two smoothing functions are essentially the same. It is seen that the cutoff frequency of this normal curve smoothing function as defined in the last section is lower than that of this truncated version of the function designed to have sharp cutoff characteristics.

It should be mentioned here that in performing mathematical filtering it is tacitly assumed that the periodicities present at the time for which the filtered variable in the time series is being estimated are unchanged in amplitude and phase during the filtering interval. Thus, it is advisable to have this filtering interval as short as possible so as to have this assumption reasonably justified. Smoothing functions having negative weights beyond the positive central values do stretch this assumption rather far owing to their longer filtering intervals for given cutoff frequencies.

7. High-Pass and Band-Pass Filtering Functions

Earlier in this paper electrical band-pass and high-pass filters were described. The same type of frequency separation can be accomplished numerically by a modification of the smoothing procedures. If smoothed values are subtracted from the corresponding values in the original unsmoothed time series, only high frequencies will remain; thus, this operation is equivalent to high-pass filtering. If well-smoothed values in a time series are subtracted from values smoothed to a lesser extent, only intermediate frequencies will remain, for the high frequencies will have been smoothed out and the low frequencies will have been subtracted out of the original series; this operation then is equivalent to band-pass filtering. Therefore, by use of these methods the oscillations in a time series can be separated into three time series each containing a particular band of frequencies—high, intermediate, and low.

For example, let x_t represent the observation in a time series at time t, and $\bar{x}_t$ and $\bar{\bar{x}}_t$ be the smoothed values computed by normal curve smoothing functions having $\sigma = \frac{1}{2}$ day and $\sigma = 5$ days, respectively (see equation (5.10) for a definition of σ). Only low frequencies of the original data will appear in the time series of $\bar{\bar{x}}_t$ values. The intermediate frequencies will appear in the series resulting from the subtraction of $\bar{\bar{x}}_t$ from $\bar{x}_t$, and only the high frequencies will remain in the series computed by subtracting $\bar{x}_t$ from x_t.

The frequency response of the intermediate frequency band-pass filter is given by

$$R(f)_{\bar{x}-\bar{\bar{x}}} = R(f)_{\bar{x}} - R(f)_{\bar{\bar{x}}} \tag{7.1}$$

where the $R(f)$'s are frequency responses and the subscripts indicate the filters to which they correspond. The frequency response of the high-pass filter is

$$R(f)_{x-\bar{x}} = 1 - R(f)_{\bar{x}} \tag{7.2}$$

The frequency responses of these filters in the above example are shown in Fig. 6. Notice that there is considerable overlap in the response

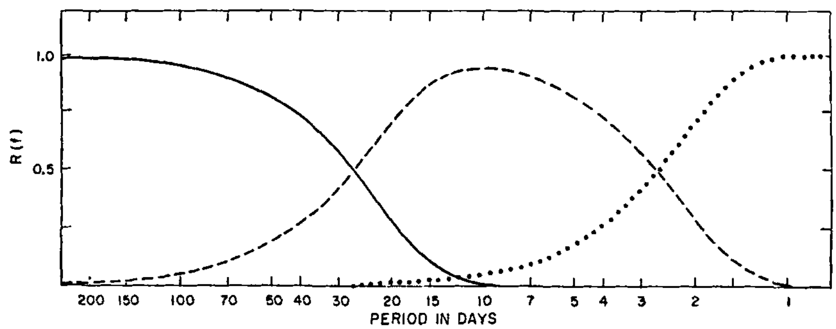


FIG. 6. Frequency responses of a low-pass filter (solid curve), a band-pass filter (dashed curve), and a high-pass filter (dotted curve) generated by two normal curve smoothing functions having $\sigma = \frac{1}{2}$ day and $\sigma = 5$ days, respectively.

curves of the three filters, and the transition between the appearance of a wave in the output of one filter and in the next filter is smooth with a uniform change of the frequency of the wave. If $\bar{x}_t$ and $\bar{\bar{x}}_t$ had been computed by equally-weighted running means instead of by normal curve smoothing functions, this transition would have been less smooth and the response of each filter would have been more irregular owing to the negative response characteristics of running means to certain frequencies.

To illustrate this filtering technique a series of twice daily barometric pressures at the Washington National Airport for spring 1956 are filtered by means of the filters described above. The pressures at the National Airport at 0100 and 1300 EST are plotted in Fig. 7(A), and a solid line is drawn connecting them. These values are smoothed by the normal curve smoothing function having $\sigma = 5$ days, and the resulting smoothed series is represented by a dashed line in Fig. 7(A). The weights for a discrete-valued approximation of this smoothing function are given in Table I. The process of smoothing individual half-day observations by this smoothing function would be laborious. However, it is only necessary to compute these well-smoothed values for every tenth observation (every fifth day) because the resulting smoothed series contains no waves of shorter period than ten days. Intermediate smoothed values needed as the subtrahends of the band-pass filter can be obtained by graphical interpolation. The series resulting from the band-pass and high-pass filtering are shown in

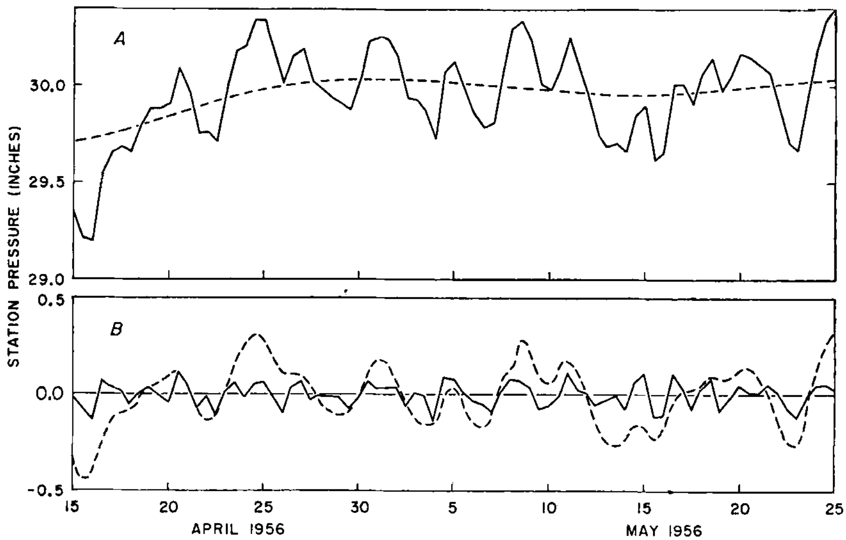


FIG. 7. Illustration of low-pass, band-pass, and high-pass filtering of station pressures at the Washington National Airport. A. In this the original pressures are indicated by the solid line and the output of the low-pass filter is dashed. B. In this the solid line is the output of the high-pass filter and the dashed line that of the band-pass filter.

Fig. 7(B) as dashed and solid lines, respectively. The weights of the normal curve smoothing function having $\sigma = \frac{1}{2}$ day used in computing the lesser smoothed values are given in Table I.

The low-pass filtered series in Fig. 7(A) appears to have a low amplitude wave of about 30 days period. Since the band-pass filter has appreciable response to waves of this period, this wave also appears weakly in the output of this filter along with waves of much shorter period. Likewise there are also waves which appear in the output of both the band-pass and high-pass filters. Thus, the partition of the waves according to period is not perfect, but it is adequate to facilitate greatly the analysis of complicated fluctuations in data such as are exhibited by the original pressure series in Fig. 7(A).

TABLE I. Weights for discrete-valued approximations of normal curve smoothing functions having $\sigma = 5$ days and ½ day.*

$\sigma = 5$ days	$\sigma = ½$ day
0.001	
0.001	0.004
0.001	0.054
0.001	0.242
0.002	0.400 (Principal weight)
0.002	0.242
0.003	0.054
0.004	0.004
0.004	
0.005	
0.007	
0.008	
0.010	
0.011	
0.013	
0.015	
0.017	
0.020	
0.022	
0.024	
0.027	
0.029	
0.031	
0.033	
0.035	
0.037	
0.038	
0.039	
0.040	
0.040 (Principal weight)	
0.040	
0.039	
0.038	
etc.†	

* Values are rounded to a few significant places for convenience in computation.

† Rest of the weights of this smoothing function are identical with the ones above the principal weight but in reverse order.

Similar sets of filters with different frequency ranges could be designed for studying other scales of atmospheric phenomena such as atmospheric turbulence. For example, Panofsky [13] used this type of filtering for making a crude spectral analysis of wind fluctuations at the Brookhaven National Laboratory.

If desired, the operation of smoothing and subtracting may be combined into a single filtering function. The appropriate weights for the high-pass filter are the negative of the weights for the associated smoothing function except for the principal weight which is one minus the principal weight of the smoothing function. For example, a high-pass filter generated by the subtraction of values smoothed by a smoothing function having weights of $\frac{1}{16}$, $\frac{1}{4}$, $\frac{3}{8}$, $\frac{1}{4}$, and $\frac{1}{16}$ in that order (where $\frac{3}{8}$ is the principal weight) would have the weights: $-\frac{1}{16}$, $-\frac{1}{4}$, $+\frac{5}{8}$, $-\frac{1}{4}$, and $-\frac{1}{16}$. Notice that the sum of the weights of a high-pass filter such as this will be zero because it does not pass the mean of the original series (a component of zero frequency). Therefore, the output of this high-pass filter will be a series having a mean of zero.

Two smoothing functions are used in the derivation of the band-pass filter. The weights for the band-pass filtering function are merely the weights of the smoothing function having the lowest cutoff frequency subtracted from the corresponding weights of the other smoothing function. This filter also does not pass the mean of the original time series.

8. Elementary Smoothing and Filtering Functions

A useful approach to the design of smoothing functions and other mathematical filters is to build up filtering functions from elementary filtering functions consisting of only three weights each. The two outer weights of the elementary filtering functions are made equal,[9] and the sum of the weights is made equal to unity for smoothing functions and zero for filters not passing the mean of the original series.

The frequency response of an elementary smoothing or filtering function is obtained from equation (8.1) below, which is derived from equation (5.6)

$$R(f) = w_0 + 2w_1 \cos(2\pi f \Delta t) \tag{8.1}$$

where w_0 is the central (and principal) weight, w_1 is the value of each of the two outer weights, and Δt is the data interval.

If one elementary filtering function provides insufficient filtering, a time series may be successively[10] smoothed or filtered by this same elementary filtering function until the desired smoothing or filtering is

[9] Each elementary filtering function is made even (namely, $w_{-1} = w_1$) so that it will produce no phase error and so that more complicated filters built up from a combination of these elementary filters will also be even and therefore produce no phase error.

[10] By "successive" filtering here is meant the repeated filtering of the entire series by the same or different filter—not the normal application of the filter to the series centered on each successive observation which is required for computing the filtered series.

accomplished. Also different elementary filtering functions may be used on succeeding filterings. It can be shown that the frequency response of the sequential application of these elementary filtering functions is the product of all the responses of the individual elementary filtering functions used; namely,

$$R(f)_R = \prod_{k=1}^{M} R(f)_k \tag{8.2}$$
$$= R(f)_1 \cdot R(f)_2 \cdot \cdots \cdot R(f)_M$$

By use of this equation it is possible to specify certain features of the resultant frequency response $R(f)_R$ desired at given frequencies and to solve for the various weights required in the elementary filtering functions to be used to give this response. This type of procedure has been successfully used by Shuman [14].

Once the elementary filtering functions have been designed, a composite filtering function can be generated which will circumvent the operation of sequential application of each elementary filtering function and give the resultant filtered series in one step. This is accomplished by first computing the cumulative cross-products of the weights of the first and second elementary filtering functions at various lags. Then the resulting series of weights is multiplied in a similar manner by the third elementary filtering function and so on until all the elementary filtering functions have been multiplied as many times as they would have been used in the sequential filtering operation. The final resulting series of weights is the composite filtering function. For example, when two three-weight elementary filtering functions having principal weights of w_0 and W_0 and outer weights of w_1 and W_1, respectively, are combined, the composite filtering functions will have weights of w_1W_1, $w_0W_1 + w_1W_0$, $2w_1W_1 + w_0W_0$, $w_0W_1 + w_1W_0$, and w_1W_1 in that order with the principal weight being $2w_1W_1 + w_0W_0$. After combining a number of elementary filtering functions, often many of the outer weights of the composite functions will become insignificantly small and may be neglected. However, it will be necessary to adjust the remaining weights in the composite function to insure that their sum is correct (namely, unity in the case of a smoothing function).

A method of smoothing that has been discussed by Brooks and Carruthers [15] is to compute running means of pairs of observations in a time series and in turn to take running means of pairs of these smoothed values and so on until the time series is smoothed sufficiently. This procedure may be considered as the repeated application of a two-weight elementary smoothing function having weights of one-half each. Composite smoothing functions generated solely from this elementary smoothing function will have weights proportional to the familiar symmetrical binominal coefficients; namely, the coefficients of the expansion of $(p + q)^n$.

For example, the composite smoothing function generated from two of these two-weight elementary smoothing functions has weights of $\frac{1}{4}$, $\frac{1}{2}$, and $\frac{1}{4}$ in that order, these weights being proportional to the coefficients (1, 2, and 1) of the second power expansion of $(p + q)$. This particular three-weight elementary smoothing function will be referred to hereafter as the elementary binomial smoothing function.

It is well known that the envelope of the coefficients in the expansion of $(p + q)$ approaches the shape of the normal probability curve as a limit as the exponent increases. Therefore, the above procedure of taking repeated running means of pairs of observations in a time series approximates the use of a normal curve smoothing function when the number of successive smoothings is large; this corresponds to the use of a large number of weights in the composite smoothing function.

The frequency response of the two-weight elementary smoothing function may be computed from equation (8.1) by setting $w_0 = 0$ and $w_1 = \frac{1}{2}$. However, the data interval to be used in this equation, $\Delta t'$, will obviously be one-half the actual data interval Δt of the time series. This substitution into equation (8.1) gives

$$R(f) = 0 + 2(0.5) \cos (2\pi f \Delta t') = \cos (\pi f \Delta t) \tag{8.3}$$

From equation (8.2) the frequency response of the elementary binomial smoothing function will be the square of the result in equation (8.3); namely,

$$R(f) = \cos^2 (\pi f \Delta t) \tag{8.4}$$

The above result can also be determined directly from equation (8.1) by setting $w_0 = \frac{1}{2}$ and $w_1 = \frac{1}{4}$. In general, the binomial smoothing function having weights proportional to the coefficients of $(p + q)^n$ has a frequency response of

$$R(f) = \cos^n (\pi f \Delta t) \tag{8.5}$$

The derivation of equation (8.5) assumes that the sum of the weights of the binomial smoothing function is made equal to unity as is done with all smoothing functions considered in this paper.

Some workers smooth time series by successive use of running means of more than two terms each. Brooks and Carruthers [15] give the following general expression for the frequency response of the operation of M

successive smoothings by the equally-weighted running means of N terms each

$$R(f) = \left[\frac{\sin (N\pi f\Delta t)}{N \sin (\pi f \Delta t)}\right]^M \tag{8.6}$$

With N set equal to two, this expression is equivalent to equation (8.5).

9. Design of Inverse Smoothing Functions

An application of the use of elementary filtering functions to the design of filters is the computation of weights for inverse smoothing functions. A composite inverse smoothing function can be generated from elementary inverse smoothing functions in the same way that composite smoothing functions can be generated from elementary smoothing functions; that is, cumulative cross-products of the weights of the inverse smoothing functions involved are computed for various lags. Unfortunately, it is not possible to design a three-weight inverse smoothing function which has exactly the inverse or reciprocal frequency response of the elementary binomial smoothing function. If this were possible, the design of inverse smoothing functions for correcting for binomial and normal curve type smoothing would be facilitated.

However, the three-weight elementary inverse smoothing function having a central weight of 2 and outer weights of -0.5 has a frequency response which is roughly the reciprocal of the response of the elementary binomial smoothing function (having weights of $\frac{1}{4}$, $\frac{1}{2}$, and $\frac{1}{4}$). The frequency response of this inverse smoothing function as computed by equation (8.1) is

$$R(f) = 2 - \cos (2\pi f \Delta t) \tag{9.1}$$

This frequency response equals the reciprocal of the response of the elementary binomial smoothing function at $f = 0$ where $R(f) = 1$ and at $f = \frac{1}{4\Delta t}$ where $R(f) = 2$. However, the response is three at $f = \frac{1}{2\Delta t}$ as compared with a value of infinity for the reciprocal of the zero response of the elementary binomial smoothing function at this frequency.

It is impossible to restore a wave to a time series by inverse smoothing if this wave was completely smoothed out previously. Some residual amplitude must remain in the smoothed series at each frequency to be restored, and this residual amplitude has to be greater than the noise for equalization to be effected with relative freedom from errors owing to noise. For example, rounding of smoothed values would introduce a type of random error, called "quantizing error," into the smoothed series. High-frequency noise is amplified in addition to the true high-frequency fluctuations in the time series, and this noise can "drown out" not only true high frequencies but also all other components besides if the amplification at these high frequencies is too great. Overemphasis of noise or random nonsignificant variation in the time series to be equalized creates a type of mathematical instability which plagues many attempts at inverse smoothing. Thus, in designing inverse smoothing functions a compromise must be made between having a stable output and having a response nearly equal to the inverse of that of the original smoothing. An amplification of three at $f = \frac{1}{2\Delta t}$ instead of infinity in equation (9.1) is an example of just such a compromise.

By combining two of the above elementary inverse smoothing functions (having weights of -0.5, $+2.0$, and -0.5) a composite function can be designed to reverse the smoothing by a binomial smoothing function generated from two elementary binomial smoothing functions; namely, the function having weights of $\frac{1}{16}$, $\frac{1}{4}$, $\frac{3}{8}$, $\frac{1}{4}$, and $\frac{1}{16}$. Although this inverse smoothing function would provide the correct restoration of waves of frequency equal to $\frac{1}{4\Delta t}$, it would overemphasize oscillations of some of the lower frequencies by as much as 40%. In order to eliminate this excessive overemphasis of low frequencies the composite inverse smoothing function should instead be generated from two elementary three-weight functions each having a central weight w_0 of 1.7 and outer weights w_1 of -0.35. These weights were obtained empirically. This composite function would have the five weights: $+0.12$, -1.19, $+3.14$, -1.19, and $+0.12$ (with $+3.14$ being the principal weight), and its frequency response would be computed according to equation (8.2) from the square of the right-hand side of equation (8.1); namely,

$$(9.2) \qquad \begin{aligned} R(f) &= w_0^2 + 4w_0w_1 \cos(2\pi f\Delta t) + 4w_1^2 \cos^2(2\pi f\Delta t) \\ &= 2.89 - 2.38 \cos(2\pi f\Delta t) + 0.49 \cos^2(2\pi f\Delta t) \end{aligned}$$

At low frequencies, this response is approximately the inverse of that of the binomial smoothing function considered above. The frequency response of the latter is

$$(9.3) \qquad R(f) = \cos^4(\pi f\Delta t)$$

from equations (8.2) and (8.4) or (8.5). These frequency responses are compared in Fig. 8 to show the extent to which this type of inverse smoothing can restore the original balance of frequencies in the smoothed time series. A curve representing the product of the frequency response of the inverse smoothing function considered and that of the binomial smoothing function at each frequency is also included in this figure.

Ideally this product should be a straight line of unit response. Departure from this ideal product shows the extent to which this form of inverse smoothing fails to compensate for the previous binomial smoothing.

Figure 9 showing three curves illustrates how well inverse smoothing

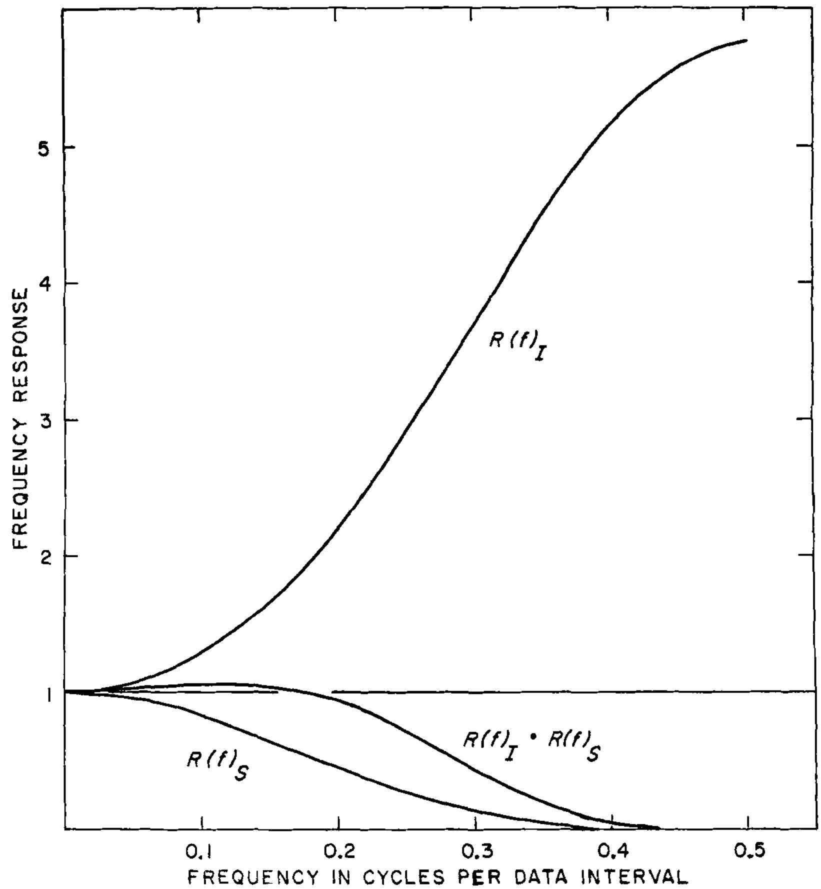


FIG. 8. Frequency response of a five-weight binomial smoothing function, $R(f)_S$, and that of an inverse smoothing function designed to reverse this binomial smoothing, $R(f)_I$. The product of these two responses is also shown.

performs on an actual time series. The solid line in this figure is the original time series; the dashed line is this series smoothed by the five-weight binomial smoothing function considered above; and the dotted line is the smoothed series inversely smoothed using the five-weight function described above. The departure of the dotted line from the solid line indicates the limitations to applying inverse smoothing to this type of smoothed series. At point A in Fig. 9 where the original series contains primarily low frequencies there is nearly perfect restoration of the original series by the inverse smoothing. On the other hand, the inverse smoothing does not restore the high-frequency fluctuation at B in the original series because this component was completely smoothed out by the previous binomial smoothing.

The examples in this section illustrate how inverse smoothing functions may be designed to compensate for two types of binomial smoothing. These procedures may also be applied to the problem of reversing many other forms of smoothing. However, serious error can be made when inverse smoothing is applied to a series which has been smoothed by a method having appreciable negative response beyond the first cutoff frequency as is the case with equally-weighted running means. As

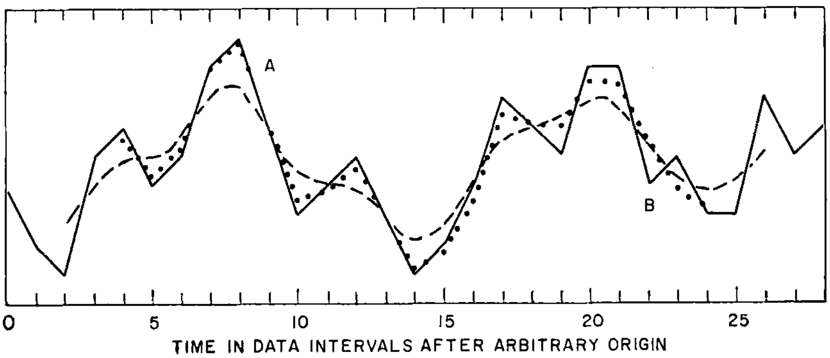


FIG. 9. Time series (solid line), same series smoothed by a five-weight binomia smoothing function (dashed line), and the smoothed series inversely smoothed (dotted line).

has been mentioned in Section 5 such a smoothing function will introduce spurious fluctuations into the smoothed output, and these fluctuations will generally be amplified by the inverse smoothing process. A good way of remedying this condition is to smooth the series again by the same smoothing method applied originally. This will restore the original polarity to all the waves in the series. Then this double smoothing may be partially reversed by an inverse smoothing function similar to those described earlier in this section. In this case inverse smoothing can at best only partially restore the original series, for no procedure can restore waves previously completely smoothed out. Smoothing methods having negative responses generally completely smooth out waves of a number of relatively low frequencies. The fact that smoothing by equally-weighted running means is not very reversible is another good argument against using this type of smoothing.

The concept of reversibility occurs in the field of thermodynamics in connection with entropy change. The more reversible a thermodynamical process is the less entropy is increased during this process. The concept of entropy, which is a measure of the disorder or randomness in a system, has been extended to statistical analysis by information theory where the negative of entropy change is taken as a measure of the information change occurring in statistical data. The foregoing implies that the less reversible a smoothing or filtering process is the more entropy is created, and therefore the more information is lost in the operation. No entropy would be created, and thus no information would be lost by perfectly reversible smoothing or filtering.

In order for smoothing to be perfectly reversible several conditions must be fulfilled. First, the original smoothing cannot have completely smoothed out any wave of finite frequency. This is true in the case of exponential and normal curve smoothing. Second, the smoothed series must be continuous—not a series of discrete values. Third, the series must be completely free from noise. Finally, waves of all frequencies must remain unchanged in phase and amplitude during the filtering interval of the original smoothing function. This last condition is fulfilled in general only when the filtering interval is infinitesimal as is the case when the series is smoothed by the derivative method described in Section 11 of this paper. These last two conditions are so stringent that for all practical purposes no smoothing operation is completely reversible, although some (for example exponential and normal curve smoothing) are more reversible than others.

10. Design of Pre-emphasis Filters

It may be desirable to amplify high frequencies in a time series more than the low frequencies for reasons other than for correcting for previous smoothing. For example, in statistical spectral analysis it is often desirable to pre-emphasize high-frequency components before computing the autocorrelation function of a series. This type of high-frequency pre-emphasis has been named "pre-whitening" by Tukey because it tends to equalize amplitudes at all frequencies [16]. This pre-emphasis prevents a great deal of instability in making spectral estimates when originally the amplitudes of waves of low frequency in a series are much greater than those of high frequency. After the spectral estimates have been obtained, they are then corrected by a factor which is the inverse of the frequency response of the pre-emphasis operation applied to the original series.[11]

[11] The square of the frequency response is required for making this correction when the relative variance rather than the relative amplitude is computed in the spectral analysis.

An inverse smoothing function similar to those described in the last section may be used for this type of pre-emphasis. However, in spectral analysis work a simpler two-weight filter is generally used which has a principal weight w_0 of unity and a weight w_{-1} of minus b. By means of this two-weight operator the pre-emphasized variable x_t' is computed from time series values x_t and x_{t-1} as follows:

$$x_t' = x_t - bx_{t-1} \tag{10.1}$$

where the subscripts refer to the times of the observations. The weight b is generally made equal to 0.75 in studies involving atmospheric turbulence. This choice of b makes the sum of the weights of this filter equal to one-quarter. Therefore, the mean of the resulting pre-emphasized series will be only one-fourth of that of the original series. However, in spectral analysis work the original series is usually made up of deviations from a mean and therefore has a mean of zero. Consequently, the effect of the pre-emphasis on the mean of the series is of no concern.

Because this two-weight pre-emphasis filter is not an even function, equations (5.2) and (5.4) must be employed for determining its frequency response; namely,

$$R(f) = [1 - b\cos(-2\pi f\Delta t)] + i[-b\sin(-2\pi f\Delta t)] \tag{10.2}$$

and

$$|R(f)| = [1 - 2b\cos(2\pi f\Delta t) + b^2]^{1/2}$$

This pre-emphasis filter also introduces phase error which may be computed from equation (5.3). However, this error is of no significance unless co- and quadrature spectra are to be computed between the pre-emphasized series and another series which has either not been pre-emphasized or has been done so differently.

A special case of this type of pre-emphasis is where b in equation (10.1) equals unity [17]. This filter is a finite difference approximation of the derivative of a time series. The response of this filter is given by equation (10.2); namely,

$$|R(f)| = [2 - 2\cos(2\pi f\Delta t)]^{1/2} = 2\sin(\pi f\Delta t) \tag{10.3}$$

When $\pi f\Delta t$ is small,

$$|R(f)| \approx 2\pi f\Delta t \tag{10.4}$$

The phase shift of this filter would be given by equation (5.3) as follows:

$$\begin{aligned}\phi &= \tan^{-1}\left[\frac{\sin(2\pi f\Delta t)}{1 - \cos(2\pi f\Delta t)}\right] = \tan^{-1}[\cot(\pi f\Delta t)] \\ &= \pi/2 - \pi f\Delta t = 90^\circ(1 - 2f\Delta t)\end{aligned} \tag{10.5}$$

In the limit as Δt goes to zero, ϕ becomes 90° which is the phase shift accomplished by differentiating any sine curve.

An exactly opposite type of pre-emphasis also used in spectral analysis work consists of taking consecutive means instead of consecutive instantaneous values of a variable which is continuously recorded such as wind [16, 18]. These means are usually obtained by eye from a graphical record. This procedure emphasizes low frequencies and is used to reduce errors resulting from "aliasing"—the process by which high-frequency waves appear as lower frequencies in a time series having a data interval too long to portray these shorter wavelengths. The spectral estimates obtained from the series generated in this manner are corrected by the inverse of the frequency response of the equally-weighted mean (computed by equation (5.9)).

11. Filtering by Means of Derivatives of Time Series

Additional insight into the operation of smoothing and filtering may be gained by considering what effects are produced on the time series by adding specified fractions of time derivatives of the series to the series itself. Consider a sine wave of angular frequency $\omega = 2\pi f$, amplitude a, and phase ϕ. A time series having only this one component would be defined as

$$x(t) = a \sin (\omega t + \phi) \tag{11.1}$$

The second, fourth, sixth, and $2n$th time derivaties of this series would be:

$$\begin{aligned} x(t)^{ii} &= d^2x/dt^2 = -a\omega^2 \sin (\omega t + \phi) \\ x(t)^{iv} &= d^4x/dt^4 = a\omega^4 \sin (\omega t + \phi) \\ x(t)^{vi} &= d^6x/dt^6 = -a\omega^6 \sin (\omega t + \phi) \\ x(t)^{2n} &= d^{2n}x/dt^{2n} = (-1)^n a\omega^{2n} \sin (\omega t + \phi) \end{aligned} \tag{11.2}$$

From the derivatives given above one can form a power series which defines a new modified time series as follows:

$$\bar{x}(t) = x(t) + k_2 x(t)^{ii} + k_4 x(t)^{iv} + \cdots + k_{2n} x(t)^{2n} \tag{11.3}$$

Substituting values in equations (11.1) and (11.2) into equation (11.3) one obtains

$$\bar{x}(t) = [1 - k_2\omega^2 + k_4\omega^4 - k_6\omega^6 + \cdots + (-1)^n k_{2n}\omega^{2n}] a \sin (\omega t + \phi) \tag{11.4}$$

Notice that the power series in brackets in equation (11.4) is actually the frequency response for the operation defined in equation (11.3) since it is the factor by which the original sine wave is multiplied in order to obtain the new filtered wave; namely,

$$R(\omega) = 1 - k_2\omega^2 + k_4\omega^4 - k_6\omega^6 + \cdots + (-1)^n k_{2n}\omega^{2n} \tag{11.5}$$

If in equation (11.5)

$$\begin{aligned} k_2 &= c^2, & k_6 &= c^6/3! = c^6/6, \\ k_4 &= c^4/2! = c^4/2, & k_{2n} &= c^{2n}/n! \end{aligned} \tag{11.6}$$

this frequency response reduces to exp $(-c^2\omega^2)$ which is the response of a normal curve smoothing function having $\sigma = c\sqrt{2}$. Thus it follows that the addition of these fractions of each even time derivative to the original series is equivalent to normal curve smoothing of the series.

The technique described here would be an alternate method for designing a filter with specified frequency response. The k's in equation (11.5) can be determined so that this power series would be equal to the desired frequency response. Many desired frequency responses may be expressed as power series, and in such cases the determination of the k's would amount merely to equating the coefficients in two power series. However, in practical application of this method only a small number of derivatives can be used, so that the actual frequency response of the derived filter is not exactly the same as the desired frequency response expanded into the complete power series. Furthermore, if the derivatives are computed by the method of finite differences, additional discrepancies between the actual and the desired frequency response will result owing to the tendency for the finite difference method to underestimate derivatives.

This method is especially valuable for the design of inverse smoothing functions. For example, suppose that the original smoothing had a frequency response of

$$R(\omega) = \exp(-c^2\omega^2) \tag{11.7}$$

The frequency response of the inverse smoothing function required to reverse this smoothing should be the reciprocal of the above $R(\omega)$ or

$$\begin{aligned} R'(\omega) = [R(\omega)]^{-1} = \exp(c^2\omega^2) = 1 + c^2\omega^2 + (2!)^{-1}c^4\omega^4 \\ + \cdots + (n!)^{-1}c^{2n}\omega^{2n} \end{aligned} \tag{11.8}$$

Equating coefficients in equation (11.8) with the corresponding ones in the power series in equation (11.5) one obtains the following values of the k's:

$$\begin{aligned} k_2 &= -c^2, & k_6 &= -c^6/6, \\ k_4 &= c^4/2, & k_{2n} &= (-1)^n c^{2n}/n! \end{aligned} \tag{11.9}$$

These k's are identical with those in (11.6) except that alternate ones have negative signs. When these values of k are substituted into equation (11.3), the required inverse smoothing operator is obtained.

That derivatives can be used for inverse smoothing could have been inferred from the fact that smoothing itself is a form of integration over time and that the inverse of integration is differentiation. The most straightforward application of differentiation to inverse smoothing is that of the equalization of a time series smoothed by an exponential process such as is performed by an instrument with a constant lag coefficient.[12] For such an instrument the following well-known differential equation holds:

$$(x - \bar{x}) = \lambda(d\bar{x}/dt) \qquad \text{or} \qquad \bar{x} = x - \lambda(d\bar{x}/dt) \tag{11.10}$$

where $\bar{x}$ is the smoothed value of the variable x in the time series and λ is the lag coefficient. If equation (11.10) is solved for x, one obtains the so-called inverted lag equation after McDonald [19]

$$x = \bar{x} + \lambda(d\bar{x}/dt) \tag{11.11}$$

From equation (11.11) it follows that perfect restoration of the original time series is theoretically possible given the smoothed values without noise, by differentiating the smoothed values with respect to time, multiplying this derivative by λ, and adding this result to the original smoothed values. A finite difference form of equation (11.11) is

$$x_t = \bar{x}_t + \lambda(\bar{x}_t - \bar{x}_{t-1}) = (1 + \lambda)\bar{x}_t - \lambda\bar{x}_{t-1} \tag{11.12}$$

where λ is in units of Δt, the data interval. Since the time derivative in equation (11.12) is estimated by the method of finite differences, it is necessary that the data interval be a small fraction of the lag coefficient, say smaller than one-fifth, in order for the results to be accurate.

If the output of the instrument is electrical, the above equalization may be done by means of a resistance-capacitance network as shown by Hall [8]. This circuit performs the required differentiation electrically. A practical limitation to this method is that the lag coefficient is seldom constant but is a complicated function of many variables. Secondly, the record must be relatively free from noise for accurate results to be obtained by this procedure.

12. Space Smoothing and Filtering

So far in this paper only smoothing and filtering of time series has been dealt with. The methods of time smoothing and filtering may, nevertheless, apply equally as well to space smoothing and filtering in one dimension, except, that the terms "wavelength" and "wave number" (number of waves in a given distance) are used instead of period and frequency,

[12] See Section 5 for a further discussion of this type of smoothing.

respectively. Furthermore, these concepts may be extended to space smoothing and filtering in two dimensions. In space smoothing in one dimension, the term wave number response may be substituted for frequency response. However, with two-dimensional space smoothing the concept of wave number or wavelength is not very meaningful unless the waves are essentially one-dimensional. The term "scale size" is perhaps better to use in this case. Scale size will be defined as the average

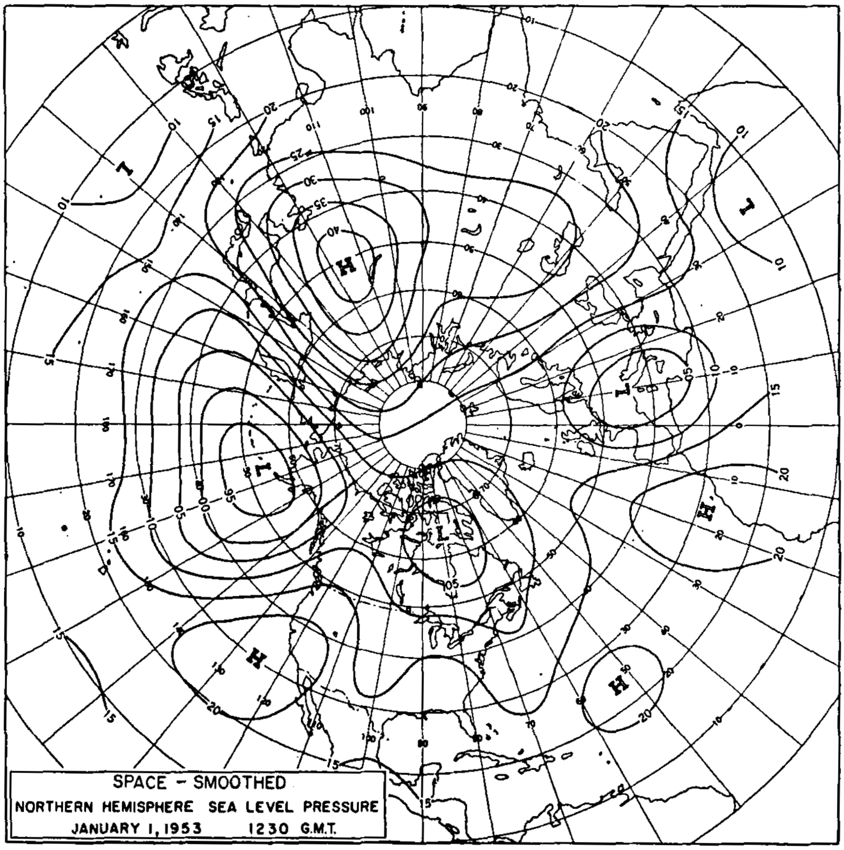
$$\begin{array}{ccc} & \frac{1}{3} & \\ \frac{1}{3} & & \frac{1}{3} \end{array} \qquad \begin{array}{ccccc} & \frac{1}{9} & & \frac{1}{9} & \\ \frac{1}{9} & & \frac{3}{9} & & \frac{1}{9} \\ & \frac{1}{9} & & \frac{1}{9} & \end{array} \qquad \begin{array}{ccccccccc} & & \frac{1}{81} & & \frac{2}{81} & & \frac{1}{81} & & \\ & \frac{2}{81} & & \frac{8}{81} & & \frac{8}{81} & & \frac{2}{81} & \\ \frac{1}{81} & & \frac{8}{81} & & \frac{15}{81} & & \frac{8}{81} & & \frac{1}{81} \\ & \frac{2}{81} & & \frac{8}{81} & & \frac{8}{81} & & \frac{2}{81} & \\ & & \frac{1}{81} & & \frac{2}{81} & & \frac{1}{81} & & \end{array}$$

$$\begin{array}{cc} \frac{1}{4} & \frac{1}{4} \\ \frac{1}{4} & \frac{1}{4} \end{array} \qquad \begin{array}{ccc} \frac{1}{16} & \frac{2}{16} & \frac{1}{16} \\ \frac{2}{16} & \frac{4}{16} & \frac{2}{16} \\ \frac{1}{16} & \frac{2}{16} & \frac{1}{16} \end{array} \qquad \begin{array}{ccccc} \frac{1}{256} & \frac{4}{256} & \frac{6}{256} & \frac{4}{256} & \frac{1}{256} \\ \frac{4}{256} & \frac{16}{256} & \frac{24}{256} & \frac{16}{256} & \frac{4}{256} \\ \frac{6}{256} & \frac{24}{256} & \frac{36}{256} & \frac{24}{256} & \frac{6}{256} \\ \frac{4}{256} & \frac{16}{256} & \frac{24}{256} & \frac{16}{256} & \frac{4}{256} \\ \frac{1}{256} & \frac{4}{256} & \frac{6}{256} & \frac{4}{256} & \frac{1}{256} \end{array}$$

FIG. 10. Generation of composite space smoothing functions from the elementary space smoothing functions in the first column. The functions in the second column result from one iteration of those in the first; those in the third column, from one iteration of those in the second.

distance between adjacent centers of high values or between adjacent centers of low values. Space smoothing will decrease the ranges between values in high and low centers of small scale without much affecting these ranges between large-scale centers.

Space smoothing and filtering functions are generally isotropic; namely, they are functions only of the radial distance from the origin. The function will usually have a maximum at the origin and decrease with radial distance outward from this origin. The scale response of two-dimensional smoothing and filtering functions will be comparable to the wavelength response of a one-dimensional filter which is a cross-section of this two-dimensional filter. Discrete-weighted space filters will have weights proportional to the ordinates of continuous two-dimensional space filtering functions. The sum of these weights is made equal to unity



(A)

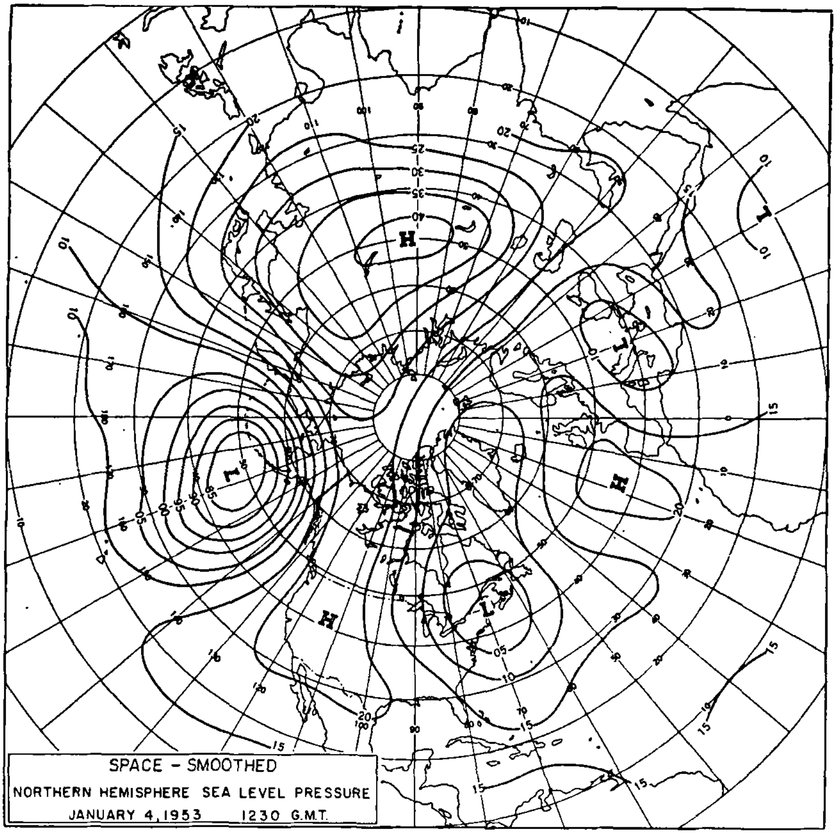
FIG. 11 (A,B,C,D). Space-smoothed Northern Hemisphere sea-level pressure maps.

for smoothing functions and zero for filters not passing the mean of the field.

As with most one-dimensional filters, the central weight of a two-dimensional filter is the principal weight. If no weight is centrally located, the space-smoothed value corresponds to the location of the center of gravity of the filter's weights.

It is conceivable that nonisotropic space smoothing may be desirable for some purposes. For example, one may wish to suppress north-south fluctuations more than those in the east-west direction. This could be accomplished with a space smoothing function whose weights decrease to zero faster in the east-west direction than to the north and south.

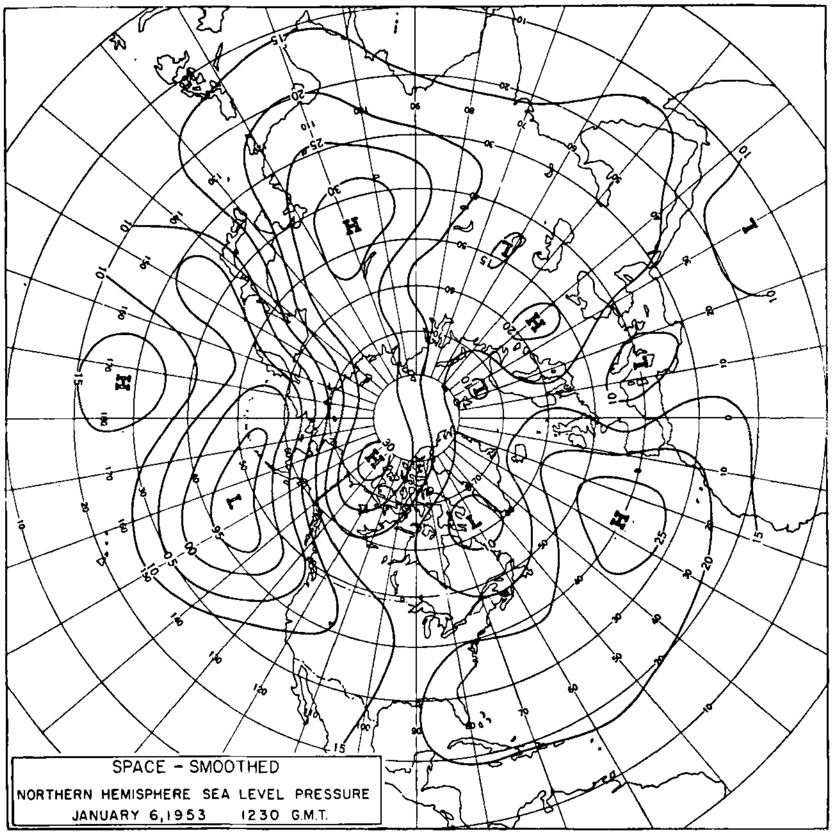
Composite space smoothing and filtering functions may be generated



(B)

FIG. 11. (*Continued*).

from elementary two-dimensional functions just as composite one-dimensional smoothing and filtering functions were built up from elementary one-dimensional functions earlier in this paper. Such elementary functions could consist of either an array of three weights of one-third each at the corners of an equilateral triangle or an array of four weights of one-fourth each at the corners of a square. Iterative cumulative multiplication of these elementary space functions at various relative positions creates composite smoothing functions whose weights approach the circularly-symmetric bi-variate normal distribution (with standard deviation the same in all directions from the origin) as is illustrated in Fig. 10. The bi-variate normal smoothing function has the same desirable characteristics in two dimensions that the normal curve smoothing function has in one dimension. One such desirable property is that the bi-



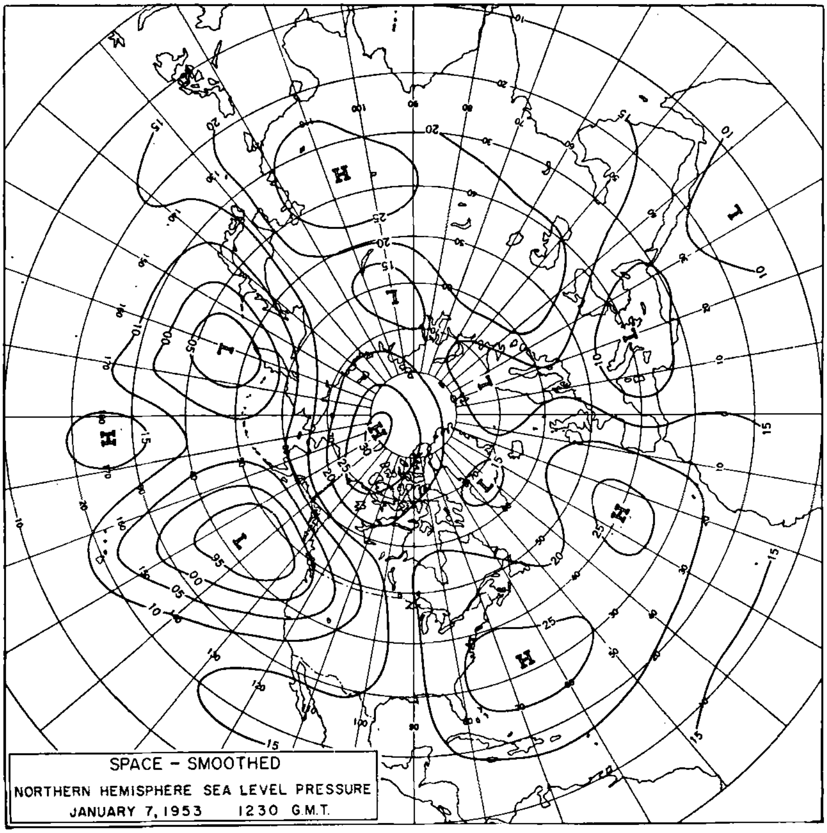
(C)

FIG. 11. (*Continued*).

variate normal smoothing function will not reverse polarities of fluctuations of any scale whereas, for example, an equally-weighted space smoothing function will reverse the polarity of features at some scale sizes.

The familiar Fjørtoft method of space smoothing is merely the single application of the square elementary space smoothing function described above [20]. Therefore, by smoothing twice by the Fjørtoft method a first approximation is made to circular bi-variate normal smoothing illustrated in Fig. 10.

The use of space smoothing is illustrated in Fig. 11 showing four Northern Hemisphere sea-level pressure maps smoothed by a space smoothing function very similar to the one in the upper right of Fig. 10. The distance between the grid points on the triangular grid used averaged



(D)

FIG. 11. (*Continued*).

about 500 miles. This smoothing almost completely attenuates features having a scale size of about 1500 miles, but it retains 4000-mile features at about 75% of their original amplitude. These smoothed maps thus display only large-scale pressure systems which are of particular interest to extended weather forecasters.

The small-scale features of the pressure field can be isolated by subtracting the smoothed map from the original unsmoothed pressures. This operation constitutes the space analog of the high-pass filter of time series terminology. An example of this process is shown in Fig. 12 where the

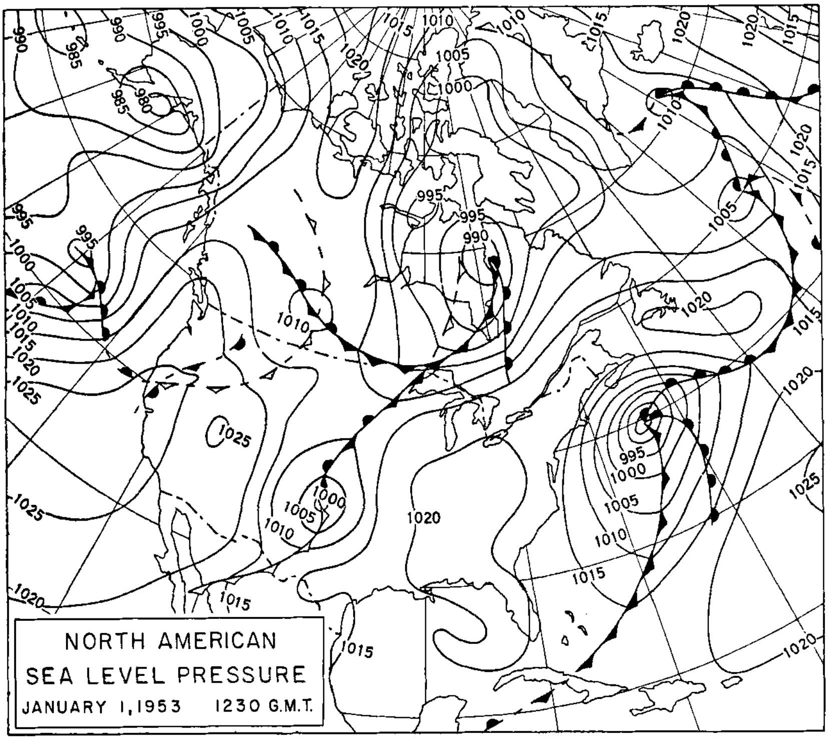


FIG. 12. Original North American sea-level pressures (left) and small-scale pressures (right) obtained by subtracting the space-smoothed map from the original map.

small-scale patterns of the January 1, 1953 North American surface map are isolated by subtraction of the smoothed map for this date obtained previously. This procedure essentially eliminates systems of greater extent than about 4000 miles. The original map is also shown in Fig. 12 for comparison. The filtered map has the same general appearance as the original map but gives slightly more emphasis to certain smaller-scale features. The fact that these maps look very similar suggests that in viewing a weather map we naturally concentrate more on small-scale features than on the large-scale ones and thus do high-pass filtering in our "mind's eye."

ACKNOWLEDGMENTS

The writer wishes to acknowledge the helpful suggestions and assistance given during the preparation of this paper by the following people: Messrs. F. Hall, G. W. Brier, I. Enger, R. A. McCormick, and J. E. Caskey, Jr., of the U. S. Weather Bureau, Col. P. D. Thompson of the Joint Numerical Weather Prediction Unit and Dr. Max A. Woodbury of New York University.

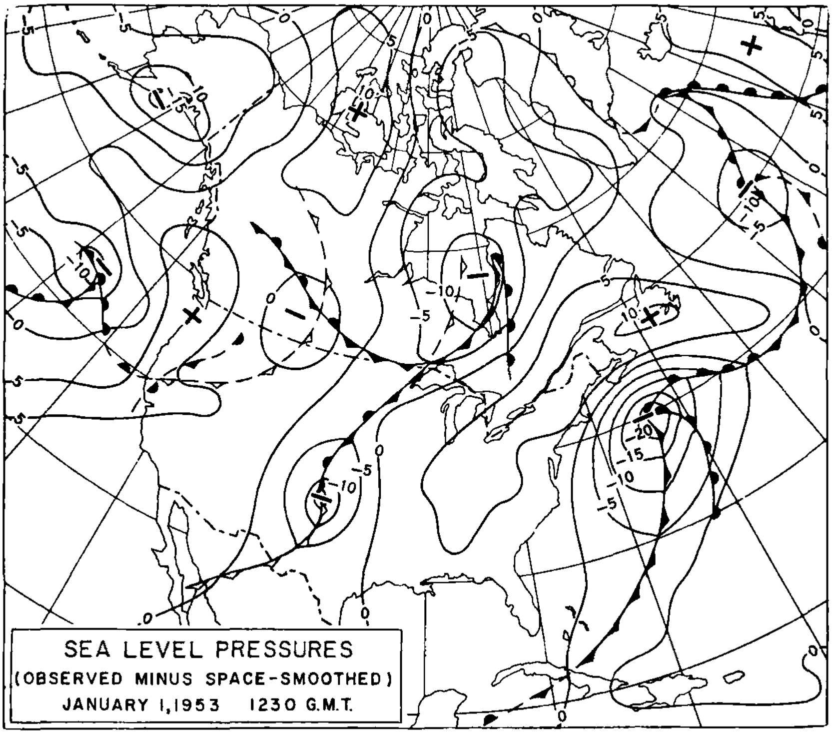


FIG. 12. (*Continued*).

LIST OF SYMBOLS

a	amplitude of wave
b	weight in pre-emphasis filter
c	constant
f	frequency
f_c	cutoff frequency
i	square root of minus one
$Im\{y\}$	imaginary part of y
k	summation variable
k_n	nth coefficient
M	number of times a series is smoothed or filtered
m	number of weights in smoothing or filtering function after principal weight
N	number of terms in an equally-weighted running mean
n	number of weights in smoothing or filtering function before principal weight; arbitrary exponent
p, q	arbitrary variables
R	ratio of the mean of the filtered series to that of the original series
$R(f)$	frequency response function; frequency response when it is a pure real quantity
$\lvert R(f)\rvert$	absolute value of frequency response function; frequency response when $R(f)$ is complex

$R(f)_R$	resultant frequency response
$R(f)_y$	frequency response of filter y
$R(\omega)$	frequency response as a function of ω, the angular frequency
$Re\{y\}$	real part of y
T	filtering interval
t	time
Δt	data interval
W_k, w_k	kth weight of a smoothing or filtering function
W_0, w_0	principal weight of a smoothing or filtering function
$w(t)$	smoothing or filtering function
x_t	unsmoothed or unfiltered discrete variable at time t
$\bar{x}_t$	smoothed or filtered discrete variable at time t
$\bar{\bar{x}}_t$	well-smoothed discrete variable at time t
x_t'	pre-emphasized discrete variable at time t
$x(t)$, x	continuous function of time
$\bar{x}(t)$, $\bar{x}$	smoothed continuous function of time
$x(t)^i$	ith time derivative of $x(t)$
λ	lag coefficient
σ	normal curve dispersion parameter
ϕ	phase angle
ω	angular velocity or angular frequency $= 2\pi f$

References

1. Carslaw, H. S. (1930). "Introduction to the Theory of Fourier's Series and Integrals." Macmillan, New York.
2. Bracewell, R. N. (1955). Correcting for Gaussian aerial smoothing. *Australian J. Physics* **8**, 54–60.
3. Bracewell, R. N. (1955). A method of correcting the broadening of X-ray line profiles. *Australian J. Physics* **8**, 61–67.
4. Bracewell, R. N. (1955). Correcting for running means by successive substitutions. *Australian J. Physics* **8**, 329–334.
5. Bracewell, R. N. (1955). Simple graphical method of correcting for instrumental broadening. *J. Opt. Soc. Amer.* **45**, 873–876.
6. Bracewell, R. N., and Roberts, J. A. (1954). Aerial smoothing in radio astronomy. *Australian J. Physics* **7**, 615–640.
7. Burr, E. J. (1955). Sharpening of observational data in two dimensions. *Australian J. Physics* **8**, 30–53.
8. Hall, F. (1950). Communication theory applied to meteorological measurements. *J. Meteorol.* **7**, 121–129.
9. Kovasznay, L. S. G., and Joseph, H. M. (1953). Processing of two-dimensional patterns by scanning techniques. *Science* **118**, 475–477.
10. Kovasznay, L. S. G., and Joseph, H. M. (1955). Image Processing. *Proc. I.R.E.* **43**, 560–570.
11. Middleton, W. E. K., and Spilhaus, A. F. (1953). "Meteorological Instruments," 3rd rev. ed. Univ. of Toronto Press, Toronto.
12. Amble, O. (1953). A smoothing technique for pressure maps. *Bull. Am. Meteorol. Soc.* **34**, 293–297.
13. Panofsky, H. A. (1953). The variation of the turbulence spectrum with height under superadiabatic conditions. *Quart. J. Roy. Meteorol. Soc.* **79**, 150–153.
14. Shuman, F. G. (1955). A method of designing finite-difference smoothing operators

to meet specifications. Technical Memorandum No. 7, Joint Numerical Weather Prediction Unit, Washington, D. C. To be published in *Monthly Weather Rev.* **85.**

15. Brooks, C. E. P., and Carruthers, N. (1953). "Handbook of Statistical Methods in Meteorology," H.M.S.O., London.
16. Gifford, F., Jr. (1955). A simultaneous Lagrangian-Eulerian turbulence experiment. *Monthly Weather Rev.* **83,** 293–301.
17. Dedebant, G., and Machado, E. A. M. (1955). Effectos de ciertos filtros sobre la correlatión. *Meteoros* **5,** 163–176.
18. Griffith, H. L., Panofsky, H. A., and Van der Hoven, I. (1956). Power-spectrum analysis over large ranges of frequency. *J. Meteorol.* **13,** 279–282.
19. McDonald, J. E. (1952). Lag effects in the measurement of turbulent temperature fluctuations. Scientific Report No. 1, Contract No. AF19(122)-440, Iowa State College, Ames, Iowa.
20. Fjørtoft, R. (1952). On a numerical method of integrating the barotropic vorticity equation. *Tellus* **4,** 179–194.

General References

Berry, F. A., Haggard, W. H., and Wolff, P. M. (1954). Description of Contour Patterns at 500 mb. Bureau of Aeronautics Project AROWA, U. S. Naval Air Station, Norfolk, Va.

Goldman, S. (1953). "Information Theory." Prentice Hall, New York.

Jeffreys, H., and Jeffreys, B. (1950). "Methods in Mathematical Physics." Cambridge Univ. Press, London and New York.

Wiener, N. (1949). "Extrapolation, Interpolation, and Smoothing of Stationary Time Series." Wiley, New York.