Information Content and Indirect Sensing Measurements

S. TWOMEY

Division of Radiophysics, CSIRO, Sydney, Australia

24 November 1969

ABSTRACT

Atmospheric transmission functions, being negative-exponential in character, are strongly interdependent and thereby limit the information content of indirect sensing even if a large number of measurements is made.

1. Introduction

With the advent of satellite observations increasing attention has been given to the indirect sensing of atmospheric variables such as ozone, temperature distributions, particulate content, etc., by observations which are usually optical in nature and which rely on differences in the transmission to the measuring instrument of energy emitted or scattered from different levels.

Although ambitious projections have been made, the literature shows that when feasible experiments have been analyzed quantitatively, it has not proved possible, on paper, to obtain *detailed* structure from such experiments. In fact, when realistic accuracies are postulated, a general suggestion seems to emerge that only a few (perhaps 3–6) independent inferences may be drawn about the distribution which the indirect sensing is intended to sense. Mateer (1965) showed that Umkehr measurements for determining ozone distributions from the ground (Götz, 1931) were interdependent to a very high degree and therefore provided only about four pieces of information even when 30 or more measurement points were obtained. Herman and Yarger (1969) investigated the derivation of ozone distributions from satellite measurements at various angles of observation (including full polarization measurements), and concluded that about 4–8 pieces of independent information could be obtained from measurements of high (but possible) accuracy.

In a different context, Twomey and Howell (1967) found that no more than about 3–4 inferences could be drawn from spectral transmission measurements in an aerosol the size distribution of which is being sought, while the writer has found about a fourfold degree of independence in sets of infrared atmospheric transmission functions.

This does not at all support the implication often encountered that more numerous measurements will automatically give more detail in atmospheric temperature profiles, ozone distributions, or other distributions being sensed by this kind of measurement.

The typical indirect sensing relies on differences in optical transmission to give different weightings to different levels of the atmosphere and thereby to enable the sorting out of contributions, emission or absorption, as the case may be, from different levels. Fundamentally, it is immaterial whether the differences in transmission are obtained by varying angle, wavelength, polarization or some combination of these or other variables.

In the present note it is hoped to show that a modest information content is a general consequence of the nature of transmission functions and that, therefore, the achievement of high-resolution, detailed distributions by remote sensing is limited, regardless of the precise nature of the sensing method or of the mathematical inversion applied to the data.

2. Analysis

Since transmission functions are quite generally decreasing exponentials or some combination of decreasing exponentials, the degree of interdependence existing among these functions is highly relevant to the ability of indirect sensing to obtain useful information about an unknown distribution. It may be noted at the outset that (quite unlike the sequence of functions e^{ix}, e^{2ix}, $\cdots$) the sequence e^{-x}, $e^{-2x}\cdots$, or the set of functions e^{-kx} for real positive k and x contains no mutually orthogonal functions, because

$$\int_a^b e^{-k_1x}e^{-k_2x}dx=(k_1+k_2)^{-1}[e^{-(k_1+k_2)a}-e^{-(k_1+k_2)b}]$$

cannot vanish for real positive k. There is, in other words, a considerable degree of interdependence between the exponential functions with real argument.

Consider a hypothetical experiment in which the value of

$$\int_0^1 e^{-kx}f(x)dx=y$$

can be measured for any desired values of k in order to estimate the unknown function $f(x)$. Since k is essentially an absorption parameter, it will be real and its possible range will be limited in any realizable physical experiment.

Although a very large number of k's can be packed into each decade, it is intuitively obvious and mathematically demonstrable that within the limits of error two measurements may be redundant if the values of k associated with them are too close. It is less obvious, but equally true, that if a number of values of k have been used, any further measurement using a value of k within or near the range spanned by the previous values may be redundant, in the sense that the result of the measurement could have been predicted to within the observational error from the other measurements, simply because e^{-kx} can be approximated quite accurately by a combination of exponentials of lower order. Useful information can be gained only when a new k is sufficiently far removed from the previously used values.

"Sufficient" and "useful" are qualitative terms only, but they can be made quantitative if one defines a useful measurement as one in which the unpredictable, nonredundant component in the measurement can equal the error "noise" component. Each measured y_m can be split up into 1) a component completely dependent on the preceding values $y_1, y_2 \cdots y_{m-1}$, 2) a component independent of the preceding y's, and 3) an error component. To achieve the division into the first two components, it is sufficient to express e^{-kx} as a sum of a dependent component formed by a linear combination of the preceding functions and a remainder which is orthogonal to them. This is just what is accomplished by the Gran-Schmidt orthogonalization procedure, which, from a set of arbitrary base functions such as e^{-k_1x}, e^{-k_2x}, $\cdots$, constructs a set of orthonormal functions $\phi_1(x)$, $\phi_2(x)$, $\cdots$, each $\phi_m(x)$ being a linear combination of the base functions up to the mth. The actual procedure is simple: the first ϕ, $\phi_1(x)$, is merely e^{-k_1x} normalized; the second ϕ is first derived in unnormalized form by taking out of e^{-k_2x} its component which is nonorthogonal to $\phi_1(x)$ and hence to e^{-k_1x}, i.e., before normalization $\phi_2(x)$ is

$$e^{-k_2x}-\phi_1(x)\int e^{-k_2x}\phi_1(x)dx;$$

after normalization of ϕ_2 one proceeds to ϕ_3 in a similar manner, and similarly derives ϕ_3, ϕ_4, etc. The orthonormal functions ϕ_1, ϕ_2, $\cdots$ provide a resolution of each original function such that

$$e^{-k_1x}=c_{11}\phi_1(x),$$
$$e^{-k_2x}=c_{21}\phi_1(x)+c_{22}\phi_2(x),$$
$$e^{-k_mx}=c_{m1}\phi_1(x)+\cdots+c_{mm}\phi_m(x).$$

If the square norm $\int[\ ]^2dx$ is accepted as a scalar measure of the size of a function, the size of e^{-k_mx} is given by $c_{m1}^2+c_{m2}^2+\cdots+c_{mm}^2$, while c_{mm}^2 gives the size of that part of e^{-k_mx} which is independent of the preceding exponentials.

To determine how closely values of k might usefully be spaced, a set of values was chosen for a quantity ϵ^2.

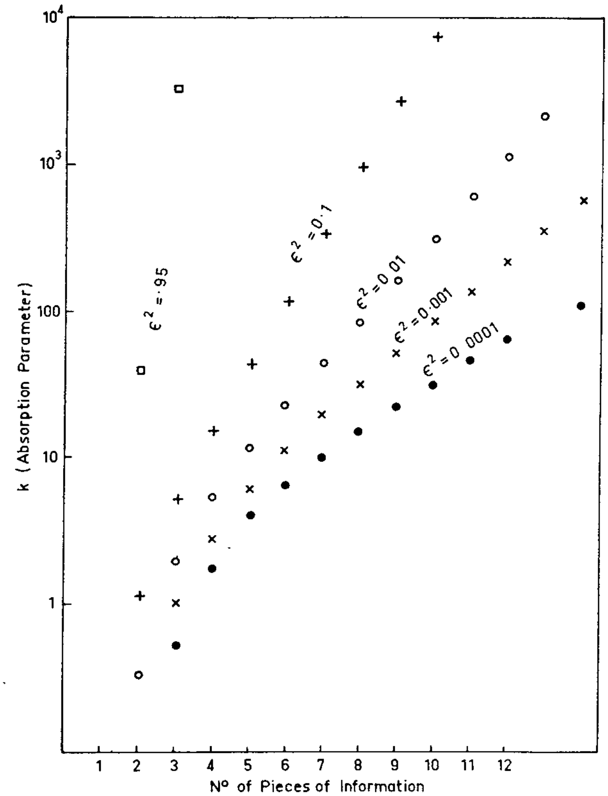

FIG. 1. Successive values of absorption parameter k needed to maintain a fixed degree of independence in each measurement for various ϵ^2, the relative (mean square) size of the independent component.

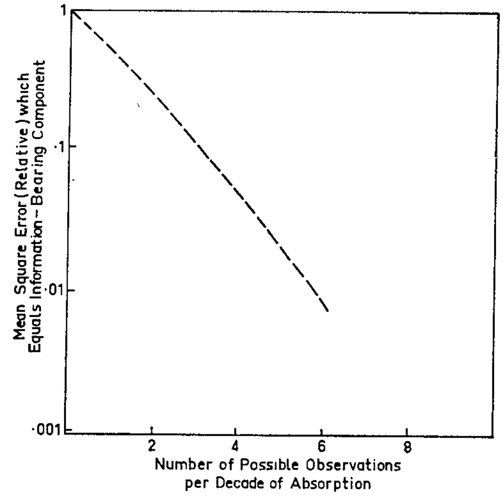

FIG. 2. Useful observations per decade for prescribed levels of independent information-bearing component.

For each value of ϵ^2 a set of values of k were derived by starting at $k_1=0$ and finding successive values k_2, k_3, etc., such that the square norm of the independent, informative component in any e^{-k_m} just equalled ϵ^2 times the square norm of the whole function, i.e., ϵ represents the rms magnitude of the fraction of e^{-kx} independent of the preceding exponentials.

3. Results

The procedure just described gave for each ϵ^2 a set of values for k which have been plotted in Fig. 1. The quantity ϵ can be interpreted in different ways but it can perhaps best be described as the relative accuracy of measurement needed to give at each new measurement a unit signal-to-noise ratio (i.e., a ratio of new information to error). As would be expected, a closer spacing of values of the absorption parameter k demands greater accuracy in the measurement to achieve a useful increase in information content.

The curves of Fig. 1 become very nearly linear beyond the first few values of k and the slope becomes a meaningful parameter; the inverse of the slope gives the number of additional independent pieces of information which are possible for each additional decade in k. A plot of the latter vs $\log\epsilon$ (Fig. 2) is also close to linear over the range considered here. Fig. 2 may be used to give a rough indication of how many (*at best*) useful observations can be made for each decade of absorption parameter. If, for example, an rms measurement accuracy of 3% was envisaged and an rms information error noise ratio of about 3:1 was deemed useful, then $\epsilon\sim0.01$ (giving about $3\frac{1}{2}$ pieces of information per decade).

Herman and Yarger's ozone results for a decimal absorption coefficient of 0.381 cm^{-1}, ozone amount

0.3 cm, and a nadir angle ranging up to $\cos^{-1} 0.01$ correspond to a maximum absorption parameter of about 30. For values of ϵ^2 of 0.01 to 0.001, Fig. 2 gives 6–8 pieces of information, which is very similar to the conclusions drawn by these authors from their eigenanalysis.

4. Conclusions

It is strongly suggested by these results that indirect sensing measurements which utilize differences in transmission to sort out contributions from various levels are fundamentally limited to a fairly small number of independent inferences concerning the unknown distribution.

The analysis shows that the inclusion of further measurements within a finite range of the absorption parameter becomes useless beyond a certain point since for observations of excellent accuracy three, or perhaps four, observations per decade are about the limit. Inclusion of more observations can be useful only if redundancy is desired.

REFERENCES

Götz, F. W., 1931: Zum Strahlungsklima der Spitzbergen Sommers. *Beitr. Geophys.*, **31**, 119.

Herman, B. M., and D. N. Yarger, 1969: Estimating the vertical atmospheric ozone distribution by inverting the radiative transfer equation for pure molecular scattering. *J. Atmos. Sci.*, **26**, 153–162.

Mateer, C. L., 1965: On the information content of Umkehr observations. *J. Atmos. Sci.*, **22**, 370–381.

Twomey, S., and H. B. Howell, 1967: Some aspects of the optical estimation of microstructure in cloud and fog. *App. Opt.*, **6**, 2125.