

The Effect of Atmospheric Aerosols on Climate with Special Reference to Temperature near the Earth's Surface

J. MURRAY MITCHELL, JR.

Climatic Change Project, Environmental Data Service, NOAA, Silver Spring, Md.

(Manuscript received 16 December 1970, in revised form 23 February 1971)

ABSTRACT

A generalized model of the effect of an optically thin atmospheric aerosol on the terrestrial heat budget is proposed, and applied to the problem of estimating the impact of the aerosol on temperatures near the earth's surface. The distinction between warming and cooling near the surface attributable to the aerosol is found on the basis of this model to depend on whether the ratio of absorption a to backscatter b of incoming solar radiation by the aerosol is greater or less than the critical ratio

$$(a/b)_0 = C(1-A)(1-Ak) / [D(1+A) - C(1-A)],$$

where A is the surface albedo, C the fraction of sensible to total (sensible plus latent) solar heating of the surface, D the fraction of aerosol that is in convective contact with the surface, and k a multiple of b that measures the relative aerosol backscattering efficiency with respect to solar radiation reflected upward from the surface.

A distinction is drawn between a stratospheric aerosol ($D=0$) which generally cools the atmosphere near the surface, and a tropospheric aerosol ($D \rightarrow 1$) which may either cool or warm the atmosphere near the surface depending on various properties of the aerosol and of the surface itself. Over moist surfaces, such as vegetated areas and oceans, the critical ratio $(a/b)_0$ is of order 0.1. Over drier surfaces, such as deserts and urban areas, $(a/b)_0$ is of order unity. If the actual ratio a/b of most tropospheric aerosols is of order unity, as inferred by previous authors, then the dominant effect of such aerosols is warming except over deserts and urban areas where the effect is somewhat marginal between warming and cooling.

Further aerosol climatic effects are found likely to include a slight decrease of cloudiness and precipitation, and an increase of "planetary" albedo above the oceans, although not necessarily above the continents. Suggestions by several previous authors to the effect that the apparent worldwide cooling of climate in recent decades is attributable to large-scale increases of particulate pollution of the atmosphere by human activities are not supported by this analysis.

1. Introduction

In recent years the attention of meteorologists has increasingly focussed on atmospheric aerosols as an influencing factor in the radiation balance and thermal state of the earth-atmosphere system. The potential importance of atmospheric haze, dust and smoke in such respects has been recognized for a long time. The pursuit of quantitative studies of the aerosol problem has become powerfully motivated by accumulating evidence of a general buildup of atmospheric aerosol loading attributable to human activities, and by suggestions that this buildup may be causally related to the changes of climate now in progress around the world.

The relevance to climate of particulate pollution of the atmosphere by human activities is more or less obvious on a small geographical scale; thus, atmospheric dust and smoke around cities and industrial centers is often so heavy as to result in substantial reductions of solar energy reaching the ground. Less obvious, perhaps, is a connection with climate on a larger scale. Not only is the total aerosol loading confirmed to be several times

greater over urban areas than over rural areas (Peterson, 1969), but the total loading in rural areas themselves is found to have been measurably increasing in recent years, at least in the United States (Ludwig *et al.*, 1970). Moreover, indirect evidence can be cited of long-term increases in the so-called background levels of particulate loading over very large regions of the earth, both continental and oceanic. This evidence derives in part from dustfall measurements in alpine glaciers of Eastern Europe (see Bryson and Wendland, 1970), solar radiation records from such remote mountain stations as Mauna Loa, Hawaii, and Davos, Switzerland (Bryson and Wendland, 1970; McCormick and Ludwig, 1967), and electric conductivity of the atmosphere over the North Atlantic Ocean (Cobb and Wells, 1970). So far, at least, background increases over oceanic areas have not been detected outside the North Atlantic (e.g., Cobb and Wells); therefore, allusions one can find in the literature to "global" trends in aerosol (or atmospheric turbidity) may be an exaggeration in terms. The geographical scale of the problem is nevertheless enormous.

Turning to the presumed climatic effect of such large-scale background aerosol increases, a number of authors have called attention to the coincidence between these increases and a systematic decline of worldwide average temperature in the past two or three decades, and have considered the possibility of a causal connection between the two phenomena (McCormick and Ludwig, 1967; Bryson, 1968; Budyko, 1969; Bryson and Wendland, 1970; Mitchell, 1970).

It will serve as a convenient point of departure in this paper if we summarize here the basic rationale for attributing a climatic cooling effect to increases of atmospheric aerosol. To begin with, it is recognized that aerosol particles in the approximate size range 0.1–1 μ are of special relevance to the problem since these are relatively abundant in the atmosphere and are most effective in absorbing, scattering and attenuating shortwave solar radiation. It is assumed that particles in this size range have a negligibly small effect on long-wave terrestrial radiation in the atmosphere, and that larger particles which are capable of interacting more strongly with terrestrial radiation are by virtue of their smaller numbers also negligible in effect. It is further assumed, implicitly or otherwise, that the aerosol particles are much more efficient as scatterers of solar radiation than they are as absorbers. In the stated particle size range, scattering out of the direct solar beam will be primarily Mie scattering in forward (down-beam) directions. However, a relatively small fraction will be in backward (up-beam) directions, roughly 10–15% under average conditions (Charlson and Pilat, 1969; Budyko, 1969). This backward directed fraction, or *backscatter*, is principally returned to the upper atmosphere and to space, to which extent it represents a net loss of solar radiation into the region below the aerosol layer (nominally the earth's surface). The radiation loss, in turn, implies a lowered equilibrium temperature at or near the earth's surface.

With regard to the two basic assumptions underlying the above rationale, the neglect of longwave radiation effects can probably be accepted as valid to a first approximation (e.g., see Robinson, 1970). The neglect of absorption of solar radiation by the aerosol is much more difficult to accept however (Robinson, 1966, 1970; Lettau and Lettau, 1969; Charlson and Pilat, 1969). Inasmuch as absorption by the aerosol constitutes an atmospheric heat source that tends to offset the backscatter cooling effect, the *net* temperature response of near-surface atmospheric layers to the combined aerosol effects of backscatter and absorption is no longer so obviously one of cooling.

It is the purpose of this paper to outline the fundamental nature of the problem of aerosol effects on the thermal climate of the earth, taking into account not only the competing effects of aerosol absorption and backscatter on solar radiation, but also the vertical distribution of the aerosol together with certain properties of the earth's surface which have an equally im-

portant bearing on the problem. The analysis in this paper will not alone suffice to answer the question as to whether the observed increases of aerosol loading are a cause of the present-day cooling trend in climate, for that would require in addition a much better knowledge of the climatology of aerosol properties than we now possess. The analysis will, however, lead us to a generalized criterion for deciding whether the net thermal effect of an aerosol near the surface is one of warming or one of cooling. It then remains for actual observations on both natural and man-made aerosols to be tested against this criterion.

2. Partitioning of aerosol effects

At the outset it is necessary to consider how the radiative heating effects of an atmospheric aerosol are distributed, first with respect to their partitioning between various components of the terrestrial heat budget, then with respect to altitude, and finally with respect to synoptic meteorological factors.

a. Partitioning with respect to heat budget components

The anomalous insolation heating effects of an atmospheric aerosol are generally distributed between three physically distinct elements of the heat budget of the earth-atmosphere system. These¹ can be described as follows:

- δH_A Anomaly of (sensible) heating of the atmosphere within the aerosol layer, resulting from *in situ* absorption of insolation (positive).
- δH_S Anomaly of the sensible part of total heating at the earth's surface, resulting from the extinction of insolation by the overhead aerosol, which is associated with changes of temperature at the surface (negative).
- δH_L Anomaly of the latent part of total heating at the earth's surface, resulting from the extinction of insolation by the overhead aerosol, which is associated with evaporation of water (at constant temperature) from the surface and which is ultimately realized as sensible heat only at some remote point of condensation as clouds and precipitation (negative).

In a manner to be quantitatively specified later on, the sum of δH_S and δH_L depends on the albedo of the surface along with the total extinction of insolation by the aerosol, whereas the difference between these terms depends on the moistness of the surface.

b. Partitioning with respect to altitude

In general, the anomalous heating attributable to each of the three heat-budget components will be realized (as

¹ See the Appendix for a complete list of the symbols used in this paper.

sensible heating) in three more or less dissimilar ranges of altitude in the atmosphere (Fig. 1). The only component that is realized strictly at the earth's surface is the surface sensible heating δH_s . Both other components are realized above the surface, the surface latent heating (δH_L) in the range of altitudes occupied by precipitating clouds, and the atmospheric heating (δH_A) in the range of altitudes occupied by the aerosol itself. These two altitude ranges, in turn, differ in relation to each other depending on the type of aerosol involved. As to type of aerosol, a most important distinction (illustrated in Fig. 1) concerns the difference between a tropospheric aerosol maintained by relatively nearby surface emissions of particulates (both natural and man-made), and a stratospheric aerosol attributable, for example, to remote high-altitude injections of volcanic dust. In either case the atmospheric heating δH_A will tend to be distributed with altitude as the number concentration of the aerosol particles.

c. Partitioning with respect to synoptic pattern

The anomalous aerosol heating associated with each heat-budget component will be distributed differently in different regions of the synoptic weather regime that exists at any given time. For example, in the type of stagnant anticyclonic regime that is frequently associated in summer or fall with major air pollution episodes in the eastern United States, tropospheric aerosol loading becomes especially high in those areas where cloudiness and precipitation are minimal and insolation is strongest. It is reasonable to expect that the aerosol effects δH_s and δH_A would also be concentrated in those areas, but that the effect δH_L would be concentrated some hundreds of miles distant, and some hours or days later, where and when the surface moisture evaporated in the affected areas ultimately becomes involved in cloud formation and precipitation. Indeed, it is to be anticipated that under a wide range of meteorological conditions the latent heating component δH_L will affect atmospheric temperature (and atmospheric energetics in other respects as well) at different points in space and time than the sensible heating components δH_s and δH_A . To the extent that similar types of synoptic regimes tend to recur in preferred geographical locations the disengagement of δH_L from the other aerosol effects is justified in the context of synoptic climatology as well as of synoptic meteorology.

It should be emphasized that this view of the partitioning of aerosol thermal effects contrasts significantly with the assumptions implicit in earlier treatments of the problem. Hitherto, it has been tacitly assumed that the deficit of insolation heating at the ground due to aerosol attenuation is a deficit of sensible heating only. This is tantamount to assuming that $\delta H_L = 0$. In the following section we derive a simplified model of the effects of an aerosol, by means of which we may estimate the relative magnitudes of all three heat-budget com-

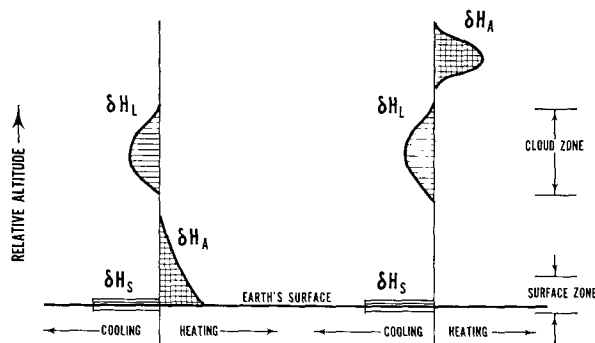


FIG. 1. Schematic distribution with altitude of the three basic aerosol heating effects defined in text, for contrasting aerosol types: tropospheric, left; stratospheric, right.

ponents as a function of relevant properties of both the aerosol and the underlying surface.

3. Aerosol-effects model

For the purposes of this paper we shall be concerned not so much with the total aerosol loading of the atmosphere as with an incremental part of the total loading, as for example a part attributable to human activities. Our approach will therefore be as follows. We begin by considering a "reference atmosphere" in which a significant aerosol content may already exist, along with gases such as water vapor and ozone which are selective absorbers of solar radiation. To compare with this reference atmosphere, we next consider an "aerosol atmosphere" which is physically identical in every respect except that it contains *additional* aerosol whose effects we wish to consider explicitly. Hereafter, in speaking of "aerosol" we generally refer to this increment of aerosol which distinguishes the "aerosol atmosphere" from the "reference atmosphere," and not necessarily to the total aerosol present. We characterize this added aerosol as a horizontally uniform, optically thin scattering and absorbing medium which occupies an atmospheric layer of arbitrary vertical thickness at an arbitrary altitude above the earth's surface.

Following Charlson and Pilat (1969) we invoke the Beer-Lambert law to describe the intensity of total solar radiation reaching the earth's surface in the presence of aerosol, expressed as a ratio to that reaching the surface in the absence of aerosol (i.e., under reference atmosphere conditions). In the notation of Charlson and Pilat this ratio can be written as

$$(S_t + S_{fs})/S_0 = \exp[-(b_{abs} + b_{bs})x],$$

where S_t and S_{fs} are respectively the transmitted (normal-incidence) and the forward-scattered (diffuse) contributions to the total radiation reaching the surface, S_0 is the total radiation reaching the surface in the absence of aerosol, b_{abs} and b_{bs} are the extinction coefficients due to absorption and to backscatter by the aerosol at solar wavelengths, and x is the effective path length of solar radiation through the aerosol layer which

depends in part on the vertical thickness of the aerosol layer and in part on the solar zenith angle.

For convenience we assume a fixed average path length \bar{x} appropriate to the average solar zenith angle in a specified latitude, season and diurnal period. That is to say, we define

$$a = b_{ab}\bar{x}, \quad b = b_{bs}\bar{x},$$

where a represents the total absorption of solar radiation attributable to the aerosol, and b the total backscatter attributable to the aerosol which escapes through the top of the aerosol layer.

Along the general lines of Charlson and Pilat we proceed to develop an expression for the incremental heating of the earth-atmosphere system, in terms of a and b , which explicitly recognizes the radiative effects of aerosol but not those of any other atmospheric constituent which is also present in the reference atmosphere. We distinguish between the total (sensible plus latent) heating at the earth's surface, denoted as H_{S+L} , and the (sensible only) heating within the atmosphere, denoted as H_A .

The surface heating H_{S+L} is composed of two parts, i.e.,

$$H_{S+L} = S_0 e^{-(a+b)}(1-A) + S_0 e^{-(a+b)} A(1-e^{-kb})(1-A). \quad (1a)$$

The first term on the right represents the heating of the surface by incoming solar radiation, given by the total solar radiation received at the surface in the presence of the aerosol, multiplied by the absorptance of the surface ($1-A$, where A is the surface albedo). The second term represents the heating of the surface by that portion of the incident solar radiation which is reflected by the surface and subsequently backscattered toward the

surface again by the aerosol. For later convenience we express the downward aerosol backscatter of the surface-reflected radiation by a constant (k) times the upward aerosol backscatter previously defined as b , where k is a nondimensional number of the order of unity.

The atmospheric heating H_A is likewise composed of two parts:

$$H_A = S_0(1-e^{-a}) + S_0 e^{-(a+b)} A e^{-kb}(1-e^{-a}). \quad (1b)$$

Here, the first term on the right represents the heating of the atmosphere by solar radiation that is absorbed by the aerosol on its way down through the aerosol layer toward the surface. The second term represents the additional heating of the atmosphere by solar radiation that is reflected upward again from the surface and absorbed by the aerosol on its second passage through the aerosol layer.²

We recall that our model is intended to refer to optically thin aerosols, for which values of both a and b are generally less than 10^{-1} . This allows us to invoke two simplifying assumptions that are amply justified for purposes of the generalized analysis to follow in this paper. First, we make use of the approximation $\exp(-a) \approx 1-a$, and likewise for b . Second, we neglect second-order terms in a and b . With these simplifications, (1a) and (1b) reduce to

$$\left. \begin{aligned} H_{S+L} &= S_0(1-A)(1-a-b+Abk) \\ H_A &= S_0(1+A)a \end{aligned} \right\}. \quad (2)$$

The surface heating H_{S+L} represents the sum of the surface sensible heating, denoted as H_S , and the surface latent heating, denoted as H_L . We let

$$\left. \begin{aligned} H_S &= CH_{S+L} \\ H_L &= (1-C)H_{S+L} \end{aligned} \right\}, \quad (3)$$

where C is assumed to depend only on the evaporability of the surface (neglecting advection and storage effects in the local heat balance), such that $0 < C < 1$. It should be noted that with this assumption $C = B/(B+1)$, where B is the ratio of sensible to latent heating, i.e., the Bowen ratio.

We take the parameter C to be a constant for a given location, time of year, type of surface, and moisture availability at the surface. This is consistent with usual assumptions involving the Bowen ratio, made either implicitly or explicitly in climatological approaches to the study of the terrestrial heat balance (see, for example, Budyko, 1958; Penman, 1963; Sellers, 1965). It reasonably implies that under otherwise identical circumstances any small change of radiant energy avail-

² The effects described by the final terms in both (1a) and (1b) were neglected by Charlson and Pilat. In the formulation of these terms shown here, it was arbitrarily assumed that downward backscattering of the surface-reflected solar radiation occurs before absorption of same by the aerosol. In the simplified version of the equations derived below, however, this arbitrariness is irrelevant.

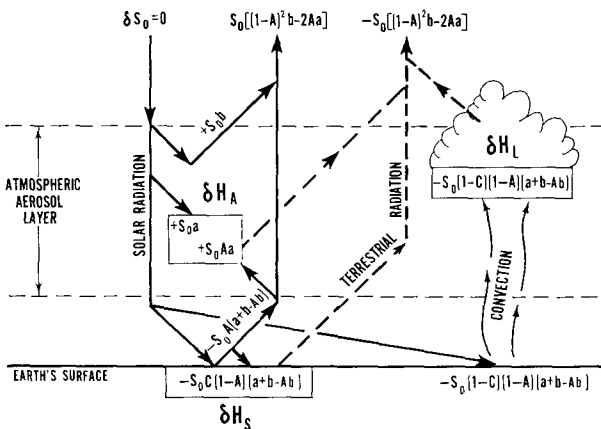


FIG. 2. Anomalous part of the terrestrial heat budget associated with atmospheric aerosol effects on solar radiation. All quantities are shown in terms of parameters of the simplified model developed in the text for the optically thin aerosol, for the special case $k=1$. Heavy solid lines refer to solar radiation, and heavy dashed lines to terrestrial (longwave) radiation. Associated arrows indicate direction of energy flow in total heat budget, to which aerosol effects may either add or subtract as indicated.

able to the surface is used in *unchanging proportions* to increase temperature and to evaporate water (or in the case of vegetated surfaces to evapotranspire water).

It remains for us to combine (2) and (3), and express the result in terms of the *anomalous* part of the heating attributable to the aerosol, in each of the three heat-budget categories. We thus write (2) for conditions applicable to the reference atmosphere, by setting $a=b=0$, and subtract this from (2) for conditions applicable to the aerosol atmosphere. Combination with (3) then yields the set of expressions

$$\left. \begin{aligned} \delta H_S &= -S_0 C(1-A)(a+b-Akb) \\ \delta H_L &= -S_0(1-C)(1-A)(a+b-Akb) \\ \delta H_A &= +S_0(1+A)a \end{aligned} \right\} \quad (4)$$

Eq. (4) constitutes the nucleus of the simplified aerosol effects model proposed in this paper; a schematic diagram of the model is given in Fig. 2 (for the special case $k=1$).

4. Climatological applications

In this section we apply the aerosol-effects model in a generalized analysis of the climatic impact of an atmospheric aerosol, with special emphasis on the effect of a tropospheric aerosol on average atmosphere temperatures near the earth's surface.

a. Evaluation of individual heat-budget effects of an aerosol

We begin by evaluating aerosol effects in terms of the basic heat-budget components as given by (4). For this purpose we treat the aerosol absorption and backscatter parameters (a and b) as variables, and solve (4) using values of surface albedo A and surface sensible heating index C which are representative of each of various types of surface as listed in Table 1. The results for

TABLE 1. Representative values of albedo and sensible-to-latent heating ratios of various types of surface.

Surface type	Surface albedo* <i>A</i>	Bowen ratio** <i>B</i>	Sensible heating index† <i>C</i>
Urban areas	0.20	4.0	0.80
Deserts	0.30	20.0	0.95
Prairies, grasslands, and farmlands (warm season)	0.20	0.67	0.40
Forests	0.16	0.43	0.30
Conifer	(0.14)	(0.50)	(0.33)
Deciduous	(0.18)	(0.33)	(0.25)
Oceans (mid-latitude)	0.08	0.10	0.09
Snowfields (stable)	0.70	0.10	0.09

* Based mainly on data of Geiger (1965) and Budyko (1958); also see Kondratyev (1969).

** Estimated from various sources, especially Budyko (1958) and Sellers (1965). Values for forests are due to Knoerr (personal communication, 1970).

† Defined as the ratio of sensible heating to total (sensible plus latent) heating, and derived from the Bowen ratio using the identity $C=B/(B+1)$.

three contrasting surface types (urban areas, vegetated areas and oceans) are illustrated in Fig. 3.

It can be seen in Fig. 3 that the relative magnitudes of the three aerosol heating components vary widely from one type of surface to another. In urban areas, for example, the latent heating deficit (δH_L) is small compared to the atmospheric heating (δH_A) and the surface cooling (δH_S), for most combinations of a and b . This follows from the relatively dry condition of urban surfaces. In ocean areas, on the other hand, much higher rates of surface evaporation occurring there result in a contrasting condition where the surface cooling (δH_S) is small compared to the other terms. The situation in vegetated areas is seen to be intermediate between these

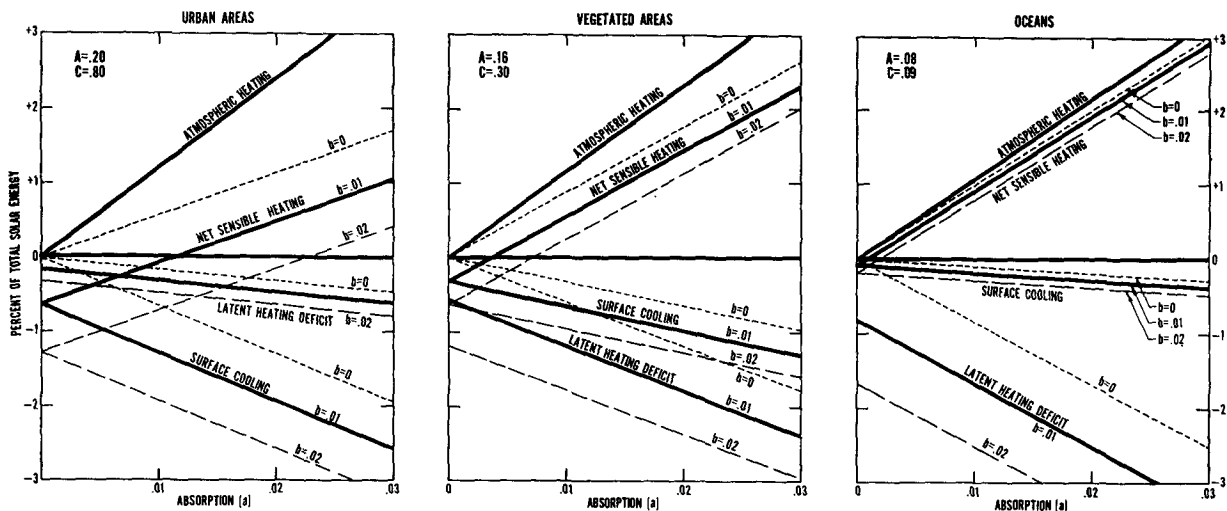


FIG. 3. Magnitude of various heating of aerosol effects (in percent of total insolation at the surface), as a function of aerosol absorption and for selected values of aerosol backscatter b . Each panel refers to a contrasting type of underlying surface, as indicated. Curves labelled "atmospheric heating" refer to δH_A in text; families of curves labelled "net sensible heating" refer to sum ($\delta H_A + \delta H_S$). Other families of parallel curves refer to surface effects. All curves are based on the assumption $k=1$.

two extremes; in this case, all three heating terms tend to be of comparable magnitudes for fixed values of a and b .

b. Net sensible heating

Focussing next on the relation of aerosol heat-budget effects on average temperature conditions in the lower atmosphere, it would seem appropriate as a first step to sum the two sensible heating components, δH_S and δH_A . We define this sum as the net sensible heating. This step is justified by the fact, noted in Section 2, that these components tend to be realized in the same synoptic locations as that of the aerosol itself, whereas the latent heating component tends to be realized (as sensible heat) in locations of active cloud and precipitation formation. For the present we neglect differences of atmospheric altitude at which the two sensible heating components are realized (δH_S at the earth's surface, and δH_A at some range of altitudes above the surface depending on the height and vertical distribution of the aerosol). We will return to the problem of aerosol altitude when we consider that part of the net sensible heating attributable to the aerosol which occurs sufficiently near the earth's surface to affect temperature near the surface.

The net sensible heating, as defined above, is shown along with the three basic heating components in Fig. 3, as a function of the aerosol absorption and backscatter intensities, and of the properties of the underlying surface.

An important fact revealed in Fig. 3 is that over relatively moist surfaces (most notably over oceans) the net

sensible heating effect of the aerosol is positive even under conditions where aerosol absorption is substantially smaller in magnitude than aerosol backscatter. This circumstance can be explained in the following way. That portion of the incident solar radiation which is denied to the earth's surface by absorption in the aerosol, and which goes entirely into sensible heating of the atmosphere (H_A), would in the absence of aerosol have gone only partially into sensible heating at the surface (H_S), with the remainder going into latent heating (H_L). Thus, if the surface is sufficiently moist, so that H_S is much smaller than H_L , it is possible for the net sensible heating of the combined earth-atmosphere system to be larger with aerosol than without aerosol, even when the aerosol absorption a is a small fraction of the aerosol backscatter b .

The relation of net sensible heating to surface type is shown in further detail in Fig. 4. Here the heating is plotted as a function of aerosol absorption, for a selected value of backscatter, for each of the six surface types listed in Table 1. A case of unusual interest shown in Fig. 4 is that of snowfields, where the net sensible heating is only slightly negative for zero absorption and increases especially rapidly with increasing absorption. This behavior is attributable to two circumstances. First, the strong reflection of solar radiation by the high-albedo snow surface contributes a relatively intense counter-stream of radiation available for absorption in the aerosol along with the incident radiation stream, leading to a very large warming component δH_A . Second, the radiation denied to the snow surface by extinction within the aerosol goes primarily toward reduced surface evaporation (or sublimation) rather than toward reduced surface sensible heating, leading to a very small cooling component δH_S .

c. Effect of aerosol altitude on net heating near surface

The local impact of the aerosol on the temperature near the earth's surface depends among other things on the altitude of the aerosol layer. If the aerosol is of a stratospheric form, confined to atmospheric layers many thousands of meters above the surface, then one would expect the temperature near the surface to be influenced only by the surface sensible heating component δH_S (refer to Fig. 1). Recalling that δH_S is a negative quantity, this always implies a surface cooling.

If, on the other hand, the aerosol is of a tropospheric form that is concentrated sufficiently near the earth's surface, then one would expect the temperature near the surface to be influenced by the sum of δH_S and δH_A , i.e., by what we have previously defined as the net sensible heating. This sum may be positive, zero or negative depending on various properties of the aerosol and the underlying surface. The surface temperature effect of a very low-altitude aerosol is therefore uncertain even as to sign unless those properties, described

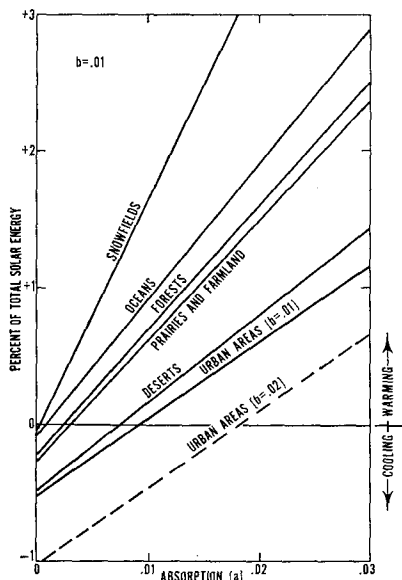


FIG. 4. Net sensible heating ($\delta H_A + \delta H_S$) as a function of aerosol absorption, in percent of total insolation, for each of the principal surface types in Table 1. All curves except the dashed curve are based on a common value of aerosol backscatter ($b=0.01$) with $k=1$.

by the parameters a , b , k , A and C in (4), can be quantitatively specified.

In the remainder of this paper, we shall be concerned primarily with a tropospheric aerosol, and with establishing what combinations of aerosol and surface properties distinguish warming and cooling effects specifically on atmospheric layers near the surface. In speaking categorically of a tropospheric aerosol, however, we must first of all recognize that such an aerosol is not necessarily concentrated "sufficiently near" the earth's surface that the assumptions of the immediately preceding paragraph are fully satisfied. This matter requires further discussion, as follows.

Based on a limited number of available observations, the *average* vertical distribution of a tropospheric aerosol is as indicated in Fig. 5. In all cases shown in the figure, the aerosol particle number densities tend to fall off with increasing altitude at a relatively rapid exponential rate. However, significant number densities remain in some cases (especially in summer-season data) at altitudes exceeding, say, 1 km; absorption of solar radiation at these higher altitudes does not necessarily contribute to a temperature increase in the vicinity of the earth's surface.

We allow for the above problem by defining a variant of the net sensible heating which represents only that part of the heating which occurs in the vicinity of the earth's surface, whatever the vertical distribution of the aerosol happens to be. We assume that this *surface-connected part* of the net sensible heating is given by

$$\delta H_D = \delta H_s + D\delta H_A, \tag{5}$$

where D is the fraction of aerosol which is contained in atmospheric layers below a certain critical altitude ($0 \leq D \leq 1$). This critical altitude, in turn, is appropriately defined as the maximum (daytime) altitude at which the air is in convectively active communication with the earth's surface. Thus, D becomes an additional parameter in the basic aerosol-effects model defined earlier. In specific meteorological circumstances, D is ideally to be evaluated from measurements of the vertical distribution of the aerosol in combination with upper air meteorological soundings to indicate the depth of convective mixing above the surface. In the absence of a climatology of D values, summarized by season and geographical location as required, we can gain at least some idea of the typical magnitude of D by reference to Fig. 5. For example, if we take 1 km as a reasonable average convective mixing depth, then Fig. 5 suggests that D ranges down to about 0.5 in summer and approaches unity in winter. Of course, the average convective mixing depth itself varies with season, being greater in summer than in winter. Therefore, if our concern is with the *total* tropospheric aerosol (to which Fig. 5 refers), we might adopt the value $D=0.75$ as a reasonable though tentative norm for such an aerosol in all seasons of the year.

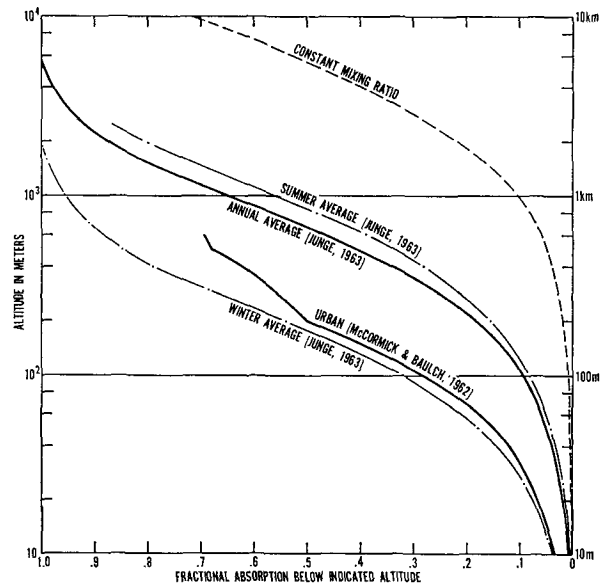


FIG. 5. Distribution with altitude of total tropospheric aerosol, based on averaged data from sources indicated and shown in terms of cumulative fraction of total aerosol particle number density below indicated altitude. Curves may alternatively be interpreted as showing fraction of δH_A below indicated altitude. Constant mixing ratio curve, shown for comparison, approximates the height distribution of a hypothetical aerosol that is thoroughly mixed into the entire atmosphere.

In summary, we propose to use (5) to define that portion of total net sensible heating due to aerosol which influences temperature near the earth's surface. The appropriate value of D to be used in (5) is

- $D=0$, for stratospheric aerosol
- $D=0.75$, for total tropospheric aerosol (tentative average value)
- $D=1$, for low-tropospheric aerosol (e.g., urban aerosol)

5. Aerosol effect on surface temperature: Warming or cooling?

In terms of our present model the net aerosol temperature effect near the earth's surface will be directly proportional to the surface-connected part of the net sensible heating δH_D , defined by (5). Substituting appropriate terms in (4) into (5), we find in particular that

$$\delta H_D = S_0 D(1+A)a - S_0 C(1-A)(a+b-Akb). \tag{6}$$

An especially useful application of (6) is that of determining what combinations of aerosol absorption, backscatter, and altitude will result in the condition $\delta H_D = 0$ which distinguishes a net warming effect from a net cooling effect near the surface. We accordingly set $\delta H_D = 0$ and solve (6) for the critical ratio of aerosol absorption to backscatter which corresponds to a zero

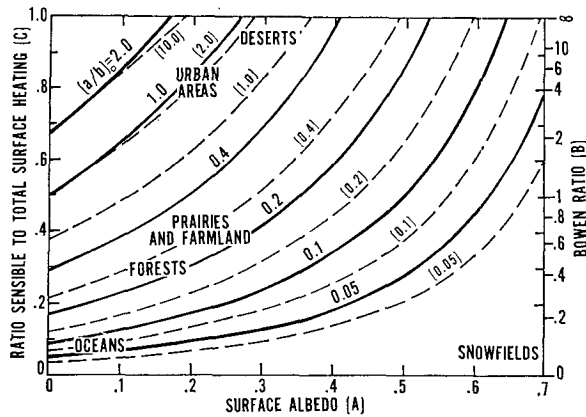


FIG. 6. Critical aerosol absorption-to-backscatter ratio, $(a/b)_0$, corresponding to zero net atmospheric heating effect near the surface, mapped in coordinates (A, C) that describe the underlying surface. Solid curves refer to special case $D=1$, and dashed curves to special case $D=0.75$, both applying to a tropospheric aerosol (see text). All curves are based on the assumption $k=1$.

temperature effect. This critical ratio is

$$\left(\frac{a}{b}\right)_0 = \frac{C(1-A)(1-Ak)}{D(1+A)-C(1-A)} \quad (7)$$

We note that this result is independent of the absolute magnitudes of absorption and backscatter, provided only that neither magnitude is so large as to invalidate the assumption of an optically thin aerosol that underlies our basic model. Actual ratios larger than this critical ratio are associated with a net warming near the surface, and smaller ratios are associated with a net cooling near the surface.³

We proceed to examine the implications of (7) from two different viewpoints, illustrated respectively in Figs. 6 and 7. From the first viewpoint, we investigate the dependence of the critical aerosol absorption-to-backscatter ratio $(a/b)_0$ on the nature of the underlying surface (as specified by arbitrary combinations of albedo A and sensible heating index C), given certain values of the aerosol altitude parameter D . From the second viewpoint, we investigate the dependence of $(a/b)_0$ on the aerosol altitude parameter D , given certain combinations of A and C appropriate to various types of underlying surface.

Specifically, in Fig. 6 we show solutions of (7) for $D=1.0$ and $D=0.75$, both applicable to tropospheric forms of aerosol. Each type of surface occupies a region of Fig. 6 determined by its coordinates in A, C space. The critical aerosol absorption-to-backscatter ratio applying to that region can be read from the figure. It is seen, for example, that for urban areas and deserts the distinction between an aerosol surface warming or cooling depends on whether the aerosol absorption-to-

backscatter ratio is respectively greater than, or less than, approximately unity. In the case of vegetated surfaces, the distinction between surface warming and cooling depends on whether this ratio exceeds a value of the order of 0.2-0.4. In the case of oceans (or other extensive water surfaces), the appropriate ratio is only about 0.1, while in the case of snowfields it is as small as 0.01.

In Fig. 7 we show solutions of (7) for selected surface conditions. In this case we assign values of A and C appropriate to each of the various types of surface, as indicated in Table 1, and investigate the relationship between the critical aerosol absorption-to-backscatter ratio and arbitrary values of D . This reveals (for each surface type) what relationship must exist between absorption, backscatter and aerosol altitude distribution in order to distinguish between a warming and a cooling effect near the surface. The figure also identifies a minimum magnitude of D , differing for each surface type, below which it is not possible *under any circumstances* for an aerosol to result in warming near the surface. This lower limit is prescribed by the fact that the denominator on the right side of (7) must be a positive quantity; for smaller D the denominator, and therefore $(a/b)_0$, become negative. In the latter event $\delta H_D < 0$ necessarily, and the only possible near-surface aerosol effect is that of cooling.

Since all evaluations of (7) thus far presented are based on the assumption that $k=1$, it is appropriate to consider to what extent the value of $(a/b)_0$ depends on

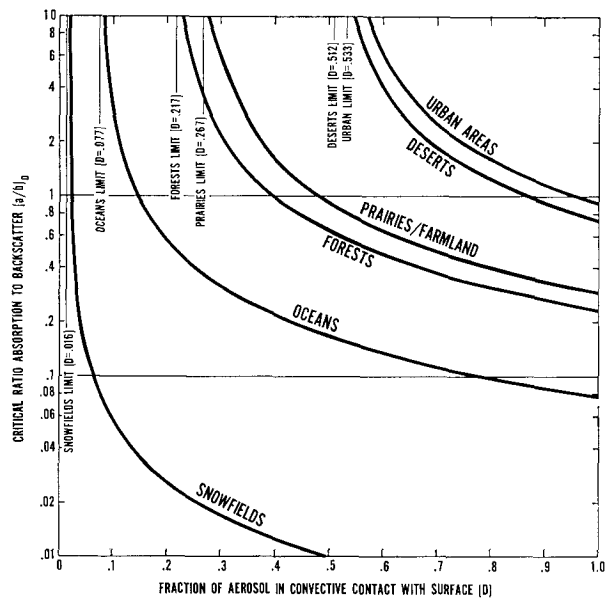


FIG. 7. Critical aerosol absorption-to-backscatter ratio, $(a/b)_0$, corresponding to zero net atmospheric heating effect near the surface, as a function of aerosol altitude parameter D for each of the principal surface types in Table 1. Vertical line segments identify limiting value of D for each surface type below which an aerosol cannot heat the atmosphere near the surface under any circumstances. All curves are based on the assumption $k=1$.

³ In the event that (7) yields a negative value of $(a/b)_0$, the condition $\delta H_D=0$ cannot be realized and the aerosol temperature effect near the surface is unconditionally one of cooling.

TABLE 2. Critical ratio $(a/b)_0$ of aerosol absorption to backscatter, corresponding to zero net atmospheric heating effect above various surfaces*, for selected values of D and k **

Surface type	$D=1.0$			$D=0.75$		
	$k=2.0$	$k=1.0$	$k=0.5$	$k=2.0$	$k=1.0$	$k=0.5$
Urban areas	0.68	0.91	1.03	1.48	1.97	2.21
Deserts	0.42	0.74	0.89	0.86	1.50	1.82
Prairies, grasslands and farmlands	0.26	0.29	0.33	0.39	0.44	0.50
Forests	0.19	0.23	0.26	0.28	0.34	0.38
Oceans (mid-latitude)	0.07	0.08	0.08	0.10	0.10	0.11
Snowfields (stable)	—	<0.01	0.01	—	<0.01	0.01

* Based on values of A and C for each surface in Table 1.

** Values of k applicable to real atmospheric conditions are considered unlikely to vary as widely from unity as either extreme shown here.

the choice of k . An indication of this is provided in Table 2.

6. Further climatic implications

a. Cloudiness and precipitation

In the foregoing analysis of aerosol effects on near-surface temperatures, we have pointed to the necessity of distinguishing between that part of the surface insolation heating which goes into sensible heating of the surface, and that part which goes into evaporating water from the surface. In contrast to the sensible heating, the part used for evaporation affects temperature at higher atmospheric altitudes (and generally in other synoptic locations as well) where clouds and precipitation occur. Moreover, we have recognized that atmospheric heating due to absorption of insolation by the aerosol is concentrated near the earth's surface only to the extent that the aerosol itself is concentrated near the surface.

The above considerations imply systematic effects of an atmospheric aerosol on cloudiness and precipitation.

First, if such an aerosol significantly alters the latent heating at the surface, and therefore the flux of water vapor into the atmosphere, it should also significantly alter total cloudiness and precipitation in those (downwind) areas where the water vapor in question becomes involved in precipitating weather systems. Since the effect of an aerosol on latent surface heating is always a negative quantity [refer to the expression for δH_L in (4)], the effect on cloudiness and precipitation is to decrease both. Typical magnitudes of δH_L can be estimated by reference to the "latent heating deficit" curves shown for different types of surfaces in Fig. 3. Over ocean surfaces, for example, it would appear that this deficit may easily reach 2–3% of the total insolation energy arriving at the sea surface, given what are thought to be typical values of total aerosol absorption and backscatter shown in the figure. Over other types of surfaces, the deficit would be somewhat less but still of the order of 1%. The associated deficits of downwind cloudiness and precipitation would presumably aggregate to similar amounts, depending, however, on such

other meteorological circumstances as the relative size of the areas of evaporation and precipitation.

Second, to whatever extent the aerosol is *not* concentrated near the earth's surface, atmospheric heating due to absorption of insolation by the aerosol occurs in upper atmospheric layers along with, or instead of, the near-surface layers. This circumstance augurs for an increased static stability of the lower atmosphere, as noted by Charlson and Pilat (1969), which under otherwise marginal conditions might further inhibit cloud and precipitation formation.

We conclude, from the thermodynamic rather than the cloud physics viewpoint, that an aerosol is likely to discourage cloud and precipitation formation for two reasons. On the other hand, the magnitude of any such effect is probably too small to be detectable in conventional climatological data on cloud cover and precipitation. In particular, it is probably too small to account for the apparent long-term decreases of precipitation in the northeastern United States in recent years which have been cited, for example, by Robinson (1970) as having a possible connection with long-term increases of atmospheric pollution in the area. The fact remains that a modest reduction of cloudiness and precipitation in, or downwind from, areas of increasing aerosol loading is predicted by the heat-budget arguments advanced in this paper. Further investigations of this and other possible hydrometeorological effects of an atmospheric aerosol are clearly warranted.

b. Albedo of the earth-atmosphere system

We observe that the sum of the three insolation aerosol heating effects,

$$\delta H = \delta H_A + \delta H_S + \delta H_L, \quad (8)$$

represents a net change in solar energy absorbed by the local earth-atmosphere system as a whole. This change in absorbed energy must be offset by an equivalent change in energy at solar wavelengths that is returned toward space by the system. With the aid of (4) it follows that the local albedo A' of the earth-atmosphere system, as sensed from above the aerosol layer, is

TABLE 3. Critical ratio $(a/b)'_0$ of aerosol absorption to backscatter, corresponding to zero net change of "planetary" albedo above various surfaces, for selected values of k .*

Surface type	Surface albedo A	k		
		2.0	1.0	0.5
Urban areas	0.20	1.2	1.6	1.8
Deserts	0.30	0.5	0.8	1.0
Prairies, grasslands and farmlands	0.20	1.2	1.6	1.8
Forests	0.16	1.8	2.2	2.4
Oceans (mid-latitude)	0.08	4.8	5.2	5.5
Snowfields (stable)	0.70	0	<0.1	0.1

* Values of k applicable to real atmospheric conditions are considered unlikely to vary as widely from unity as either extreme shown here.

altered by the presence of an aerosol to the extent that

$$\delta A' = -\delta H/S_0 = (1-A)(1-Ak)b - 2Aa. \quad (9)$$

It is instructive to inquire as to what relationship must exist between a and b in order that $\delta A' = 0$. This relationship is given by the ratio

$$(a/b)'_0 = (1-A)(1-Ak)/2A. \quad (10)$$

If we choose $k=1$ and neglect second-order terms in A , Eq. (10) reduces to the expression⁴ derived for a thin aerosol by Atwater (1970). We also note that (10) follows as a special case of (7) in which $C=D=1$.

Solutions of (10) are given in Table 3 under conditions of surface albedo appropriate to various types of surface, and for a range of values of k . If we arbitrarily assume that the actual absorption-to-backscatter ratio a/b of an atmospheric aerosol is not greater than unity, it can be seen from Table 3 that $a/b < (a/b)'_0$ with respect to all surfaces except snowfields and perhaps deserts. In that event, the aerosol effect is almost universally that of increasing the so-called "planetary" albedo, at least in areas relatively free of clouds. On the other hand, if we assume that the actual aerosol ratio a/b is only modestly larger than unity (i.e., $\gtrsim 2$), it can be inferred from the table that the aerosol effect is to decrease the planetary albedo everywhere except over the oceans. In the absence of better information on the optical properties of real aerosols, including those derived from human activities, we conclude that while most aerosols are almost certain to increase the planetary albedo over the oceans (and of the earth as a

⁴ Atwater's result for the case of an optically thin aerosol can be written in our notation as $(a/b)'_0 = (1-2A)/2A$. Unfortunately, the neglect of second-order terms in A in his derivation has the effect of appreciably underestimating the value of $(a/b)'_0$ when $A > 0.2$, and forces it to zero for $A \geq 0.5$. According to (10), $(a/b)'_0 > 0$ for all $Ak < 1$, which is generally the case except perhaps above a surface of exceptionally high albedo such as a fresh snow cover.

whole), the question is left open as to whether they increase or decrease the planetary albedo over the continents.

7. Concluding remarks

We have seen that the net effect of an atmospheric aerosol on temperatures near the surface may be one of either warming or cooling, depending on various properties of the aerosol and the underlying surface. In principle, the distinction between warming and cooling can be determined to a useful degree of approximation for real aerosols by comparing the observed aerosol absorption-to-backscatter ratio, a/b , with the critical absorption-to-backscatter ratio, $(a/b)'_0$, evaluated from Eq. (7). In practice, however, this determination is problematical because, of the various aerosol properties which are parameterized in this paper as a , b , k and D , none has been measured in the field except in relatively rare instances. Certainly none is known to an extent sufficient for purposes of establishing a reliable aerosol climatology, from which we may derive an "average" surface temperature effect of a given type of aerosol, over a given type of surface.

In view of the above-mentioned difficulties, we are unable to draw categorical conclusions as to whether the predominant effect of real-world aerosols is to warm or cool the atmosphere near the surface, except to say that the effect of high-altitude aerosols (for example, stratospheric dust veils of volcanic origin) is unconditionally one of cooling near the surface. In the case of tropospheric aerosols, however, it is to be noted that such estimates as exist of actual absorption-to-backscatter ratios suggest typical values the order of unity (Lettau and Lettau, 1969; Charlson and Pilat, 1969; Robinson, 1970). Such ratios are an order of magnitude or more larger than the critical absorption-to-backscatter ratios derived in this paper which apply to relatively moist types of surface such as oceans, vegetated areas and snowfields. We therefore maintain that, apart from arid regions and (on a local scale) urban areas, the predominant thermal effect of a tropospheric aerosol near the surface is highly likely to be one of warming rather than cooling. In arid and urban areas, the effect appears from our analysis to be rather marginal between warming and cooling, and more refined measurements on the properties of aerosols indigenous to such areas are necessary to resolve the matter.

If this tentative appraisal of the surface temperature effects of a tropospheric aerosol is borne out by further studies of aerosol properties, then the arguments noted

in the introduction to this paper supporting an inverse relation between aerosol loading and climatic temperature will require substantial qualification. With particular regard to the recent cooling trend of worldwide climate, the attribution of this cooling (or any significant part of it) to secular increases of atmospheric particles from human activities now appears unlikely, not merely on quantitative grounds (see Mitchell, 1970) but on qualitative grounds as well.

Indeed, long-term increases of particulate pollution of the atmosphere by man may serve to augment, rather than oppose, other warming effects of human activities, such as the increasing carbon dioxide content and direct thermal pollution of the atmosphere. In that event the observed climatic cooling of the past quarter century emerges more persuasively than ever as a natural geophysical phenomenon, with man the innocent bystander.

Acknowledgments. This paper is a direct outgrowth of the author's participation in the M.I.T. Summer Study on Critical Environmental Problems (SCEP), held July 1970 at Williamstown, Mass. He is indebted to a number of colleagues at SCEP, and particularly to Dr. William W. Kellogg, Chairman of the Work Group on Climatic Effects, for many stimulating discussions which, by underscoring the unsatisfactory state of our knowledge of the climatic significance of atmospheric aerosol, inspired this study. The author also acknowledges with thanks the helpful comments on the manuscript offered by Drs. James E. McDonald, James T. Peterson, Marshall A. Atwater, Robert J. Charlson and Helmut E. Landsberg.

APPENDIX

List of Symbols

H_L	latent part of total solar heating of earth's surface (associated with evaporation or evapotranspiration of water at the surface)
H_S	sensible part of total solar heating of earth's surface (associated with temperature changes at the surface)
H_{S+L}	total solar heating of earth's surface (H_S+H_L)
k	measure of total downward backscatter of surface-reflected solar radiation by aerosol, relative to AS_0 , expressed as a multiple of b
S_0	solar radiation reaching the vicinity of the earth's surface in the absence of aerosol
a/b	ratio of total absorption to total backscatter of solar radiation by aerosol
$(a/b)_0$	critical value of a/b for which $\delta H_D=0$
$(a/b)'_0$	critical value of a/b for which $\delta A'=0$
$\delta A'$	anomaly of A' attributable to aerosol
δH_A	anomaly of H_A attributable to aerosol (same as \bar{H}_A)
δH_D	anomaly of net sensible heating attributable to aerosol in vicinity of earth's surface ($\delta H_S+D\delta H_A$)
δH_L	anomaly of H_L attributable to aerosol
δH_S	anomaly of H_S attributable to aerosol
A	albedo of the earth's surface (fraction of incident solar radiation reflected by the surface)
A'	albedo of earth-atmosphere system (local "planetary" albedo), omitting the contribution of clouds
a	measure of total absorption of solar radiation by aerosol, relative to S_0
b	measure of total backscatter of solar radiation by aerosol that escapes through top of aerosol layer, relative to S_0
C	sensible heating of earth's surface by solar radiation, as a fraction of total (sensible plus latent) solar heating of surface
D	fraction of aerosol contained in convectively mixed region of the atmosphere in contact with the earth's surface
H_A	total heating of atmosphere associated with <i>in situ</i> absorption of solar radiation by aerosol

REFERENCES

- Atwater, M. A., 1970: Planetary albedo changes due to aerosols. *Science*, **170**, 64-66.
- Bryson, R. A., 1968: All other factors being constant . . . A reconciliation of several theories of climatic change. *Weatherwise*, **21**, 56-61.
- , and W. M. Wendland, 1970: Climatic effects of atmospheric pollution. *Global Effects of Environmental Pollution*, New York, Springer-Verlag/D. Reidel, 130-138.
- Budyko, M. I., 1958: *The Heat Balance of the Earth's Surface*. English translation, U. S. Department of Commerce, Office of Technical Services, Washington, D. C., 36-37, 62-63.
- , 1969: The effect of solar radiation variations on the climate of the earth. *Tellus*, **21**, 611-619.
- Charlson, R. J., and M. J. Pilat, 1969: Climate: The influence of aerosols. *J. Appl. Meteor.*, **8**, 1001-1002.
- Cobb, W. E., and H. E. Wells, 1970: The electrical conductivity of oceanic air and its correlation to global atmospheric pollution. *J. Atmos. Sci.*, **27**, 814-819.
- Geiger, R., 1965: *The Climate Near the Ground*, 4th ed. English translation, Harvard University Press, p. 15.
- Junge, C. E., 1963: *Air Chemistry and Radioactivity*. New York, Academic Press, **118**, 183-192.
- Kondratyev, K. Ya., 1969: *Radiation in the Atmosphere*. New York, Academic Press, 180-194, 411-452.
- Lettau, H., and K. Lettau, 1969: Shortwave radiation climatology. *Tellus*, **21**, 208-222.
- Ludwig, J. H., G. B. Morgan and T. B. McMullen, 1970: Trends in urban air quality. *Trans. Amer. Geophys. Union*, **51**, 468-475.
- McCormick, R. A., and D. M. Baulch, 1962: The variation with height of the dust loading over a city as determined from the atmospheric turbidity. *J. Air Pollution Control Assoc.*, **12**, 492-496.
- , and J. H. Ludwig, 1967: Climate modification by atmospheric aerosols. *Science*, **156**, 1358-1359.

- Mitchell, J. M., Jr., 1970: A preliminary evaluation of atmospheric pollution as a cause of the global temperature fluctuation of the past century. *Global Effects of Environmental Pollution*, New York, Springer-Verlag/D. Reidel, 139–155.
- Penman, H. L., 1963: *Vegetation and Hydrology*. England, Commonwealth Agricultural Bureaux, 34–35.
- Peterson, J. T., 1969: The climate of cities, a survey of recent literature. Natl. Air Pollution Control Admin., Publ. No. AP-59, 21–24.
- Robinson, G. D., 1963: Absorption of solar radiation by atmospheric aerosol as revealed by measurements at the ground. *Archiv. Meteor. Geophys. Bioklim.*, **B12**, 19–40.
- , 1966: Some determinations of atmospheric absorption by measurement of solar radiation from aircraft and at the surface. *Quart. J. Roy. Meteor. Soc.*, **92**, 263–269.
- , 1970: Long-term effects of air pollution—A survey. Unpubl. manuscript, Center for the Environment and Man, Inc., Hartford, 21–24.
- Sellers, W. D., 1965: *Physical Climatology*. University of Chicago Press, 103–105.