A Theoretical Study of the Efficiency of the General Circulation

Lloyd L. Schulman

Department of Meteorology, Massachusetts Institute of Technology, Cambridge 02139

(Manuscript received 30 August 1976, in revised form 24 December 1976)

ABSTRACT

The hypothesis that the atmosphere may be constrained to operate at nearly maximum efficiency is examined. If atmospheric efficiency is defined as the ratio of the rate of production of kinetic energy to the rate at which solar energy reaches the top of the atmosphere, the problem becomes equivalent to finding the maximum rate at which diabatic heating generates available potential energy (APE), which can be estimated independently of any frictional processes. Since diabatic heating includes long- and shortwave radiative heating, the vertical flux of sensible heat by the small-scale eddies and the release of latent heat, this would entail finding the maximizing fields of temperature, water vapor, carbon dioxide, ozone, cloudiness and surface wind speed. By specifying the relative humidity to be constant and less than 100%, by ignoring ozone as an atmospheric constituent, and by using the observed mixing ratio of carbon dioxide as basic simplifying assumptions, the release of latent heat and clouds are eliminated, and for a specified solar forcing the efficiency becomes a function of the temperature field only.

Observational studies indicate that the actual rate of generation of APE is from 2–6 W m⁻², which corresponds to an atmospheric efficiency of about 1–2%. Experiments with a 5-level, 5-latitude model yield a maximum generation of APE near 12 W m⁻². A higher resolution 5-level, 9-latitude model leads to a maximum generation near 10 W m⁻². The corresponding maximizing temperature fields show many qualitative agreements with the observed zonally averaged temperature field, including horizontal temperature gradients whose magnitudes decrease with height and the absence of superadiabatic lapse rates. The results are relatively insensitive to relative humidity, albedo or surface wind speed, but do have a strong dependence on the sensible heat distribution scheme. These solutions suggest that the general circulation may indeed be operating at nearly its maximum efficiency.

1. Introduction

The atmosphere may be regarded as a heat engine driven by solar radiation, and as such possesses an energy cycle. Following Lorenz (1955, 1967) this basic energy cycle was shown to consist of a net generation (G) of available potential energy (APE) by diabatic heating, a net conversion (C) of APE into kinetic energy (KE) by adiabatic processes, and a net dissipation (D) of KE by friction. In the long term each of the three steps of the global energy cycle must take place at the same rate.

The rate at which the energy cycle proceeds must be related to the intensity of the incoming solar radiation and in fact must be less than the rate at which this solar energy enters the atmosphere. More specifically, the ratio of the rate at which KE is produced in the atmosphere, C, or equivalently G or D, to the rate at which the solar energy reaches the outer portions of the atmosphere, is a measure of the efficiency of the atmospheric heat engine.

As a basic quantity in atmospheric energetics many observational studies have been directed at determining the efficiency of the general circulation, by evaluating either G, C or D.

A direct evaluation of D is limited by a poor understanding of the dissipation processes. However, Brunt (1926), using an assumed surface wind distribution and an estimated global mean coefficient of eddy viscosity, obtained a value of 3 W m⁻² in the planetary boundary layer. He speculated that the remainder of the atmosphere should contribute an additional 2 W m⁻². Indirect attempts at estimating D as a residual term in the kinetic energy equation have been made, most notably by Kung (1966, 1967, 1969) and Holopainen (1963). Holopainen found a value of 10.4 W m⁻², while Kung’s most recent estimate, using one year of North American data at many levels, yields a total dissipation of 5.6 W m⁻².

A direct global estimate of C is also difficult to achieve, since it involves measurement of the vertical velocity, or equivalently, cross-isobar flow. Kung had to compute C in the course of determining D in his residual method, but the average of C over a limited region is not representative of a global estimate since its sign, unlike that of D, changes from region to region. Brunt’s estimate of D was also linked with C, since his
coefficient of eddy viscosity was based on observations of cross-isobar flow.

A direct evaluation of $G$, however, largely depends only upon the covariance of temperature and non-frictional heating within isobaric surfaces; APE is generated by a heating of the warmer regions and a cooling of the colder ones at the same elevation. The evaluation of $G$ therefore yields an estimate of the atmospheric efficiency which is virtually independent of friction, and offers the most reliable means of investigating the energy cycle.

Global computations by Newell et al. (1974) yield a value of near 5 W m$^{-2}$ for $G$, while Dutton and Johnson (1967), using temperature cross sections for a single meridian, found a value of 5.6 W m$^{-2}$. Oort (1964), in a comprehensive summary, chose a value of 2.3 W m$^{-2}$.

The solar constant is observed to be about 1400 W m$^{-2}$, which corresponds to an incoming flux of solar energy of approximately 350 W per square meter of the earth’s surface. Therefore from the various estimates of $G$, the atmospheric efficiency can be estimated to be about 1–2%.

An explanation of why the atmospheric efficiency should be as low as 1–2%, or why it should not be even lower, has been characterized by Lorenz (1967) as the fundamental theoretical problem of atmospheric energetics. This paper presents the major results of a thesis investigation by Schulman (1974) into this problem.

The only other theoretical investigation of the atmospheric efficiency, and which provided the impetus for the thesis research, was performed by Lorenz (1960). Lorenz sought the maximum possible value of atmospheric efficiency using a simple two-dimensional model. On the basis of his simple model, Lorenz speculated that the atmosphere might be constrained to operate at nearly maximum efficiency, and reasoned further that the more efficient circulation patterns might be favored over the less efficient ones.

Several other researchers have presented theories exploring the mechanics of mode choice in atmospheric flow from other viewpoints. Two of them have based their ideas on the maximization or minimization of some quantity. Dutton and Johnson (1967) proposed a theory based on the principle of least action. The action was defined to be the integral of the difference between the local kinetic and total potential energy densities. It was suggested that the natural atmospheric motions, subject to constraints such as the equation of continuity and the first law of thermodynamics, would minimize this integral.

Dutton (1973) studied mode selection in the atmosphere from the viewpoint of entropy. The central quantity in his theory was the total entropic energy, $S_0 - S$, where $S$ is the entropy of any natural state of the atmosphere and $S_0$ the entropy of its associated equilibrium or maximum entropy state. This is the motionless, hydrostatic and isothermal state toward which an atmosphere in isolation will naturally tend. Dutton suggested that atmospheric mode selection was related to the development of the entropy producing mechanisms which would most efficiently hold the total entropic energy to a minimum value, or in other words, minimize the departure of the atmosphere from its associated equilibrium state.

2. The hypothesis and the approach

An explanation of the rate of generation of APE, and hence an explanation of the atmospheric efficiency, requires an explanation of the atmospheric temperature distribution. Since the temperature field is influenced by the atmospheric motions, this is apparently no less a fundamental problem than explaining the general circulation itself.

The problem of explaining the atmospheric efficiency can be approached indirectly, however, by first considering the less complex problem suggested by the following simplified argument for a dry atmosphere. Since the generation of APE depends essentially on the horizontal covariance of heating and temperature, with no horizontal temperature contrast there would be no generation despite the strong heating and cooling which would then prevail. At the other extreme, if the horizontal temperature contrast were so great that the atmosphere was in radiative equilibrium, there would be no net heating and again no generation of APE. Therefore, since the temperature and heating are positively correlated, the maximum possible rate of generation of APE should accompany some intermediate temperature contrast. The important simplification brought about by seeking the maximum rate of generation of APE and the distribution of temperature with which it corresponds, is that this problem does not require a knowledge of the atmospheric motions leading to the temperature distribution.

This simplified approach is still quite complicated. The nonfrictional heating of the atmosphere is due to a combination of radiative heating, the vertical flux of sensible heat by small-scale eddies and the release of latent heat. Radiative heating depends not only on the temperature field, but also upon the distribution of clouds which reflect shortwave radiation and absorb longwave radiation, and upon the distribution of the constituents of the atmosphere which absorb long- and shortwave radiation. Similarly, the surface fluxes of sensible and latent heat depend upon the surface values of wind speed, water vapor mixing ratio and albedo, in addition to the surface temperature. These quantities are all affected by atmospheric motions.

Two ways of eliminating the effects of the atmospheric motions are immediately obvious. Either the combination of the fields of temperature, water vapor, carbon dioxide, ozone, cloudiness and surface wind
speed which maximizes the rate of generation of APE can be found, or certain simplifying assumptions involving the atmospheric quantities or heating processes can be introduced which will make them independent of the atmospheric motions.

Lorenz (1960) explored the second choice with his simple two-dimensional model and found a maximum atmospheric efficiency of about 2%. Based on those preliminary results, the hypothesis that the atmosphere may be constrained to operate at nearly maximum efficiency is examined.

The approach taken is to investigate a series of simple numerical models, increasing in sophistication and resolution, in which the rate of generation of APE is completely determined by the temperature field and the distribution of the incoming solar radiation at the top of the atmosphere, or more simply, the solar forcing. This can be accomplished by dealing with a hypothetical atmosphere whose major simplifications include specifying the relative humidity to be some constant value less than 100%, ignoring ozone as an atmospheric constituent and using the observed mixing ratio of carbon dioxide. In this manner the release of latent heat and clouds are eliminated, and the only variable absorption constituent in the radiative flux computations is solely a function of temperature. Ozone and carbon dioxide bear no such direct temperature dependence and are not nearly as important as water vapor in atmospheric heating. Sensible heating within this hypothetical atmosphere is simulated by the aerodynamic bulk formula, which relies on a surface temperature difference, and an eddy diffusion scheme in which the eddy coefficient depends only on the lapse rate of potential temperature. Other modeling assumptions, which are primarily necessary for simplicity and consistency of the atmospheric energetic processes, are presented in the following section.

The investigation of the hypothesis using this approach does not require reproducing the general circulation and its temperature fields. It only requires finding that temperature field which maximizes the efficiency of the general circulation.

The temperature field can be thought of as some mean state temperature field. Since the atmospheric efficiency can be expressed for a given solar forcing as a function of the temperature field only, and recalling that a maximum $G$ should accompany some temperature contrast weaker than that associated with radiative equilibrium, there should exist a temperature field for each solar forcing pattern for which the atmospheric efficiency is a maximum.

Furthermore, if it is assumed that the solar forcing is independent of longitude then the temperature distribution leading to the maximum atmospheric efficiency is likely to be independent of longitude also, and the same maximizing temperature field should exist at each longitude. If this is so, since in reality eddies do exist in a state compatible with the dynamic constraints, the atmospheric efficiency can be only nearly a maximum, since the atmospheric motions would require temperature differences within latitude circles. These differences would lower $G$ below its maximum value. A more general statement of this reasoning is that a zero solar forcing contrast in any horizontal direction should yield a zero contribution to the rate of generation. This investigation will only determine the two-dimensional maximizing temperature field without eddies, realizing that the generation accompanying the actual three-dimensional maximizing field should be somewhat lower.

The models need not be thought of as two-dimensional. For reasons which will become clear later, $G$ can be computed independently of the horizontal spatial distribution of the solar forcing. Although the vertical is represented by a space coordinate, the horizontal can be represented by a group of unordered locations, of which a space coordinate is a special ordered case. The model can therefore be considered to consist of an unordered ensemble of vertical columns of temperature, in which each column receives a given magnitude of solar radiation at its top, and a resultant $G$ can be computed. However, for ease of visualization the models will be treated as two-dimensional with latitude as the horizontal coordinate, and comparisons can be drawn between the maximizing temperature fields of the models and the observed zonally averaged temperature field. Since the atmosphere is not in steady state no specific observed temperature field can be directly compared with the maximizing fields of the models. However, as an illustration of the magnitude and general features of the observed zonally averaged temperature field the annual mean estimate of Oort and Rasmusson (1971) is presented in Fig. 1.
3. Modeling assumptions and techniques

a. Basic assumptions and constraints

As derived by Lorenz (1955) the rate of generation of APE per unit area of the earth's surface is given by

$$G = \int_{m} QNd\mu,$$

with

$$N = 1 - \left(\frac{p_R}{p}\right)^{\kappa},$$

where \(m\) is the mass of the atmosphere per unit area, \(Q\) the rate of diabatic heating per unit mass, \(\rho\) the pressure, \(p_R\) the reference pressure at a point (equal to the average pressure on the isentropic surface passing through that point), and \(\kappa\) a nondimensional constant \((1 - c_e/c_p)\) approximately equal to 0.286 for dry air. The parameters \(c_p\) and \(c_e\) are the specific heat capacities at constant pressure and volume, respectively. \(N\) may be thought of as the effectiveness of the heating at any point in generating APE.

Assuming hemispherical symmetry for the two-dimensional model,

$$G = \int_{0}^{\pi/2} \int_{0}^{\pi/2} -QN \cos \phi d\phi d\rho,$$

in which \(\phi\) is the latitude, \(\rho_0\) the surface pressure (assumed constant at 1000 mb) and \(g\) the acceleration of gravity.

If we let

$$s = \sin \phi$$

so that equal increments of \(s\) represent equal increments of surface area, then

$$G = \int_{0}^{\pi/2} \int_{0}^{\pi/2} -QNdsd\rho.$$

In addition to those regarding the constant value of relative humidity \((R)\) and the distribution of the other absorbing constituents, several assumptions must be made to ensure a consistent model for the evaluation of \(G\) from the field of temperature.

A uniform smooth surface will be taken as the lower boundary. There will be no heat storage or transport within the surface, so that a flux balance is necessary at each latitude. The surface will absorb all of the shortwave flux incident upon it and will be black in the infrared. The atmosphere will be taken as nonscattering, and as a result of the constant relative humidity, will be cloudless. A planetary albedo \((a)\), reflecting the incoming solar radiation at the top of the atmosphere, will be assigned as a function of latitude. As such it can be regarded strictly as a modification of the solar forcing.

Taking this viewpoint of albedo, there can be no absorption of the reflected shortwave radiation in those models for which it is applicable.

The temperature field, given at specified levels in the models, will be interpolated within the atmosphere by assuming that the potential temperature \((\theta)\) varies linearly in space. \(\theta\) is defined as a function of the pressure and temperature \((T)\) through Poisson's equation

$$\theta = T \left(\frac{p_R}{p}\right)^{\kappa}.$$}

This interpolation procedure was chosen to facilitate the computation of \(N\). The evaluation of the reference pressure \(p_R\) at a point with potential temperature \(\theta_p\) is necessary for the computation of \(N\) at that point. \(p_R\) is a measure of the fraction of the total mass of the atmosphere whose potential temperature exceeds \(\theta_p\).

That fraction is given by the ratio \(p_R/p_0\). The pressures at which the isentropic surface \(\theta_p\) intersect each latitude are easily determined. If the sum of those pressure intervals in which the potential temperatures exceed \(\theta_p\) is calculated for each latitude, and this sum denoted by \(P_I\), then

$$p_R = \int_{0}^{1} p_I ds.$$

In a completely statically stable atmosphere \(p_0\) becomes simply the average pressure of the surface of constant \(\theta_p\).

Only those temperature fields which satisfy two basic energetic constraints will be considered in seeking the field which maximizes the rate of generation of APE. The first constraint is that there must be no net gain of energy by the mean state model atmosphere, or

$$\int_{m} Qdm = 0.$$

With the assumption of a surface flux balance at each latitude this becomes equivalent to requiring that the net incoming and outgoing radiation integrated over the top of the atmosphere must also balance. The imposition of this constraint will be discussed as part of the numerical maximization scheme.

The second constraint, as in the real atmosphere, is that the model atmosphere must undergo no net change in entropy. Since frictional heating is positive everywhere, the nonfrictional heating must act to decrease the entropy, or

$$E = \int_{m} \frac{Q}{T} dm \leq 0.$$

This constraint is satisfied by a heating of the warmer regions and a cooling of the colder regions. Roughly speaking, such a distribution of temperature and heating
also leads to a positive rate of generation of APE, and in fact the constraint was satisfied by almost all the fields encountered in this investigation. However, it is possible to construct a model atmosphere for which both \( G \) and \( E \) are greater than zero, and so the constraint is still necessary. The constraint is imposed by rejecting from consideration those temperature fields for which \( E \) is positive.

b. Radiative and sensible heating schemes

The diabatic heating \( Q \) can be separated into three components

\[
Q = Q_t + Q_s + Q_d,
\]

where \( Q_t \) and \( Q_s \) are heating resulting from long- and shortwave radiative flux divergence, respectively; and \( Q_d \) is the sensible heating due to the small-scale turbulent transfers of heat from the surface to the atmosphere, and within the atmosphere.

The maximization scheme requires repeated computations of \( G \), making it imperative that fast but reliable methods for calculating the components of \( Q \) be used. Their calculations are simplified through the use of parameterizations and empirical formulas.

Fast but accurate computations of the longwave radiative fluxes in the models are obtained by using a modified version of the procedure suggested by Sasamori (1968). His program was originally developed for use in the NCAR general circulation model. Basically, his method consists of expressing the absorption functions of water vapor and carbon dioxide in Yamamoto's (1952) radiation chart as easily evaluated empirical formulas, which greatly simplifies the numerical integration of the radiative flux equations.

In a cloudless atmosphere the downward and upward longwave radiative fluxes at level \( p \) are given by

\[
F_i(p) = 4\sigma \int_0^{T(p)} \tilde{A}(u(T')) - u(T(p)) \, dT',
\]

\[
F_t(p) = \sigma T^4 + 4\sigma \int_{T(p)}^{T(p_0)} \tilde{A}(u(T')) - u(T') \, dT',
\]

where \( \tilde{A}(u, T) \) is the normalized absorptivity, \( u \) the path length of the absorbing constituents (water vapor and carbon dioxide), \( \sigma \) the Stefan-Boltzmann constant, and \( T_s \) the ground temperature as opposed to \( T(p_0) \) the temperature of the air a few meters above the ground. The transitional layer from \( T_s \) to \( T(p_0) \) is essentially transparent and contributes no flux. \( T_s \) will be discussed later in this section.

The normalized absorptivities are too complicated for easy evaluation. However, the mean absorptivities used in Yamamoto's chart change only slightly with temperature from around 200 to over 300 K. At extremely low temperatures the mean absorptivity increases rapidly with decreasing temperature. Using this information, and letting \( p_1 \) denote the level above which the path length changes very little (assumed less than 50 mb for the model), the radiative fluxes at level \( p \) may be rewritten:

\[
F_i(p) = 4\sigma \int_{T(p_1)}^{T(p)} \tilde{A}(u(T')) - u(T(p)) \, dT',
\]

\[
F_t(p) = \sigma T^4 + 4\sigma \int_{T(p_1)}^{T(p)} \tilde{A}(u(T')) - u(T') \, dT',
\]

where \( \tilde{A}(u) \) denotes the average absorption function \( \tilde{A}(u, T') \) for temperatures of approximately 200 to 320 K and \( \tilde{A}(u, T(p_1)) \) may be referred to as the isothermal absorptivity. These functions can be easily evaluated and are given by Sasamori (1968). For simplicity in choosing \( p_1 \) and computing \( \tilde{A} \), the temperature field is assumed constant above 100 mb.

The corresponding rate of heating due to the longwave flux divergence can be obtained from

\[
Q_i(p) = \frac{\partial F_i(p)}{\partial p},
\]

where \( F_i(p) \), the net longwave flux at level \( p \), is defined as

\[
F_i(p) = F_t(p) - F_i(p).
\]
The net longwave flux at the air-ground interface, written as \( F_{\text{le}} \), is equal to \( F_{\text{l}}(p_0) \).

The radiative flux equations are integrated using a trapezoidal method with unequal vertical temperature increments, and the net longwave fluxes are computed for each 100 mb interval. Further details of the numerical procedure are found in Schulman (1974).

A test calculation of the net radiative fluxes was made to compare the speed and accuracy of the modified Sasamori method with other numerical models of infrared cooling. The test was based on a hypothetical distribution of temperature, water vapor and carbon dioxide given at 20 pressure levels, which was previously used by Stone and Manabe (1968) to compare the longwave radiation scheme incorporated into the GFDL general circulation model, with the sophisticated but time-consuming method of Rogers and Walsh (1966). The Rogers and Walsh model, which includes a frequency integration over absorption bands and the influence of temperature on line intensity, is generally considered the most complete and reliable available.

A more recent version of this method, presented by Rogers in an unpublished report, was also used in the comparison. The results are shown in Fig. 2.

The modified Sasamori method compares favorably in accuracy with the other radiation schemes, despite the various simplifying assumptions and with only half the vertical resolution. The largest deviations from the more reliable computations occur between 200 and 600 mb, but are still less than 10%. Partly because of the fewer number of levels, Sasamori's method is also much faster. Computation times were given as approximately 1.5 s for the Rogers and Walsh scheme and 0.4 s for the GFDL scheme on the UNIVAC 1108 computer. The modified Sasamori scheme took only about 0.1 s on the comparable IBM 370/165 computer.

The absorption of solar radiation in the models is computed using a modified empirical formula of Mäge and Möller (1932). This method assumes a non-scattering cloudless atmosphere, and also neglects the absorption by carbon dioxide which, according to Roach (1961), accounts for only a very small portion of the atmospheric heating. A similar scheme was also adopted for use in an early version of the NCAR general circulation model (Kasahara and Washington, 1967).

Using this procedure, the flux of the direct solar radiation that is absorbed by water vapor in a vertical column extending from the top of the atmosphere to the level \( p \) on a given day and at a given latitude is obtained from (in units of W m\(^{-2}\))

\[
F_{\text{ar}}(p) = 120L\left[u_w(p)\ \text{sec}z\right]^{0.3} \cos z,
\]

where \( z \) is the instantaneous zenith angle of the sun, \( u_w \) the water vapor path length, \( L \) the fraction of the entire day with the sun above the horizon, and the overbar represents an integrated average over these daylight hours. The corresponding rate of heating due to the shortwave flux divergence can then be computed from

\[
Q_s(p) = \frac{\partial F_{\text{as}}(p)}{\partial p}.
\]

The two quantities \( \cos z \) and \( L \) can both be obtained from the relation

\[
\cos z = \sin \phi \sin \delta + \cos \phi \cos \delta \cos \beta,
\]

in which \( \phi \) is the latitude, \( \delta \) the solar declination and \( \beta \) the hour angle in radians (24 h corresponds to \( 2\pi \) rad). At sunrise or sunset \( \cos z = 0 \) and \( \beta = B \) is called the half-day length, but \( B \) can better be described as half the number of daylight hours. By integrating over the daylight hours

\[
\cos z = \int_{-B}^{B} (\sin \phi \sin \delta + \cos \phi \cos \delta \cos \beta) d\beta / \int_{-B}^{B} d\beta,
\]

which reduces to

\[
\cos z = \sin \phi \sin \delta + B^{-1} \cos \phi \cos \delta \sin B,
\]

and \( B \) can be computed from the nonaveraged equation with \( \cos z = 0 \) or

\[
\cos B = -\tan \phi \tan \delta \ (0 \leq \phi < \pi/2).
\]

From the definition of \( L \) we can then find

\[
L = \frac{B}{\pi}.
\]

As an estimate of the accuracy of this simple empirical scheme, computations of the solar heating were compared with those presented by Roach (1961), based upon the seasonal distributions of water vapor and carbon dioxide given in his paper. Roach used a fairly sophisticated procedure which included an integration of the laboratory compiled water vapor and carbon dioxide absorptivities of Howard et al. (1955) over wavelength. His daily averaged heating rates were obtained by integrating over the 24 h period in 30 min steps. The results for January are compared in Fig. 3.

Differences in the computed heating rates range between 0.1-0.2 K day\(^{-1}\), which is well within the error limits expected in any radiative flux calculation.

Computation of the incident solar flux, or solar forcing for the models, as a function of latitude and time of year is given by

\[
F_s(s) = S_0(L \cos z),
\]

where \( S_0 \) is the solar constant, taken as 1395 W m\(^{-2}\). For the main study results equinoctial forcing is assumed, so that \( \delta = 0 \), and \( L \cos z = (1/\pi) \cos \phi \).
After reflection at the top of the atmosphere the shortwave flux available for absorption becomes

\[ F_{sw}(s) = F_s(s)(1 - \alpha). \]

It follows that the shortwave flux absorbed by the ground can be written as

\[ F_{sg}(s) = F_{sw}(s) - F_{sw}(p_0). \]

With the development of notation for the longwave and shortwave fluxes the energetic constraint, as discussed in the previous section, which requires that the net flux exchange with space be zero, can now be formulated as

\[ H = \int (F_{sw} - F_{sw}(0)) ds = 0. \]

This constraint on allowable temperature fields stipulates that there be no net gain of energy by the model atmosphere.

Central to the radiative flux calculations is the computation of the water vapor path length. If the fields of temperature and relative humidity are known, the water vapor distribution and the resulting path lengths are completely determined. The saturation water vapor pressure \( e_s \) can be computed from the temperature field using a formula derived in unpublished notes by N. A. Phillips:

\[ e_s(T) = 6.11 \left( \frac{273.16}{T} \right)^{5.0066} \exp[24.84573(1 - 273.16/T)]. \]

Then, from the field of relative humidity the water vapor mixing ratio \( q \) is given by

\[ q(p, T) = \frac{e_s(T)}{p - e_s(T)} \]

and the pressure-corrected water vapor path length from the top of the atmosphere to level \( p \) in a vertical air column of unit cross section is approximated by

\[ u_w(p, T) = \frac{1}{g} \int_0^p \frac{1}{p} q dp, \]

based on the common assumption that the absorption lines change their width in linear proportion to the pressure.

The third component of the diabatic heating is the sensible heating. A surface flux of sensible heat is necessary to ensure a surface flux balance at each latitude, or

\[ F_{d_s}(s) = F_{sg}(s) - F_{d}(s), \]

where \( F_{d} \) is obtained from the bulk aerodynamic formula

\[ F_{d}(s) = \rho C_d u_e V_s [T_s - T(p_0)], \]

in which \( \rho_0 \) is the air density at 1000 mb, \( V_s \) the anemometer height wind speed and \( C_d \) the surface drag coefficient. Based on observations at the air-sea interface, Kraus (1972) estimated \( C_d \) to be \( 1.3 \times 10^{-3} \).

For a given atmospheric temperature profile, there is a unique value of \( T_s \) which will maintain the surface flux balance, since only \( F_{d_s} \) and \( F_{d} \) depend upon \( T_s \) and both increase with increasing \( T_s \). If \( T_{s(i)} \) is the \( i \)th estimate of \( T_s \) and

\[ y_i = F_{d_s} - F_{d}(T_{s(i)}) - F_{d}(T_{s(i)}), \]

then using the secant iterative method the proper \( T_s \) is chosen by

\[ T_{s(i+1)} = \frac{T_{s(i)} - y_{i-1}}{y_{i} - y_{i-1}}. \]

The sensible heat from the surface is distributed within the atmosphere by the small-scale turbulent eddies. This can only be considered a source of diabatic atmospheric heating with the assumption that the small-scale eddies, as distinguished from the large (synoptic) scale, are not regarded as part of the circulation. If only molecular scales of motion were not considered part of the circulation, conduction would lead to extremely large values of sensible heating in a very thin layer next to the ground.
The turbulent flux of sensible heat at a level $p$ within the atmosphere is parameterized by eddy diffusion, or

$$F_d(p) = g \rho^2 c_p K_H \frac{\partial \theta}{\partial p}, \quad 200 \text{ mb} < p < p_o,$$

with the assumption that

$$F_d(p) = 0, \quad p < 200 \text{ mb},$$

where $\rho$ is the atmospheric density at level $p$, and $K_H$ the eddy coefficient of sensible heat. The parameterizations of both $F_{\sigma_s}$ or $F_d(p_o)$ and $F_d(p)$ are similar to those used in the initial NCAR general circulation model.

The sensible heating at level $p$ can then be obtained from

$$Q_d(p) = g \frac{\partial F_d(p)}{\partial p}.$$

Little is known about the functional form of the dependence of $K_H$ on stability and height. However, $K_H$ should increase with decreasing stability to account for an upward heat flux in an atmosphere that is, on the average, stable. Deardorff (1967) has proposed an empirical formulation of $K_H$ which depends upon the vertical wind shear as well as the lapse rate and includes a correction to the lapse rate intended to account for the observed upward heat flux when $\partial \theta / \partial p$ is zero or slightly negative. However, since diabatic heating is solely a function of the temperature field in this study, the formulation of Fisher and Caplan (1963) has been adopted, i.e.,

$$K_H = K_H^* \exp \left( \epsilon \frac{\partial \theta}{\partial p} \right),$$

where $K_H^* = 15 \text{ m}^2 \text{s}^{-1}$ and $\epsilon = 130 \text{ m K}^{-1}$ are chosen as the values of the empirical constants.

An alternate formulation by Asai (1965), which will serve as a comparison, is given by

$$K_H = \frac{K_H^*}{1 - \epsilon \rho \frac{\partial \theta}{\partial p}} \leq 0.$$

In this case the empirical constants are set at $K_H^* = 10 \text{ m}^2 \text{s}^{-1}$ and $\epsilon = 500 \text{ m K}^{-1}$.

Palmén and Newton (1969) have crudely estimated the mean vertical distribution of the small-scale sensible heat flux during the winter for the region north of $32^\circ \text{N}$. Their flux estimate, which depicts a uniform heating up to about 500 mb and a cooling above, is presented in Fig. 4. They arbitrarily selected 600 mb as the level above which turbulent convection is negligible and where mechanically forced turbulence, leading to a downward heat flux, begins to dominate.

Since the temperature fields in this study represent mean state temperatures which are usually stable, the eddy diffusion schemes confine the bulk of the positive sensible heating to the lowest layer of the models. To investigate the influence of this apparent limitation on the maximizing fields of temperature, two alternate sensible heating schemes, which arbitrarily distribute a larger portion of the sensible heat through a greater

---

**Fig. 4.** Estimated mean small-scale vertical eddy sensible heat flux during winter for the region $32^\circ-90^\circ \text{N}$ by Palmén and Newton (1969).

**Fig. 5.** Hypothetical fields of $G$ and $H$ in two-dimensional temperature-phase space. Superimposed are vectors showing the changing temperature field during convergence to the maximum allowable $G$, which is represented by a small circle.
depth of the atmosphere, are the subject of experiments conducted in a later section.

c. Maximization scheme

An iterative “gradient” or “hill climbing” method in temperature-phase space is the basis of the numerical scheme used to find the maximizing temperature fields. The $n$-point temperature field can be represented in temperature-phase space by the $n$-dimensional vector

$$T_i, \; i = 1, 2, 3, \ldots, n.$$  

For each temperature field $T_i$, after the surface flux balance has been imposed and $T_e$ found for each latitude, values of $G$ and $H$ can be computed. A schematic example of hypothetical fields of $G$ and $H$ is given in Fig. 5 for two dimensions of the temperature-phase space.

A maximizing temperature field, subject to the constraint that the net flux exchange with space be zero, satisfies the equations

$$\frac{\partial G}{\partial T_i} - \lambda \frac{\partial H}{\partial T_i} = 0,$$  

where $\lambda$ is of the nature of a Lagrange multiplier and is determined in such a way that the constraint is satisfied. In the maximization procedure the initial temperature field is chosen so that $H=0$. Subsequent temperature fields in the iterative process need only satisfy the constraint that $H$ remains zero, or $\partial H/\partial t = 0$, where $t$ is just an iteration parameter. A convergence to the maximizing temperature field is attained by a series of successive approximations to the equations

$$\frac{\partial T_i}{\partial t} = k \left( \frac{\partial G}{\partial T_i} - \lambda \frac{\partial H}{\partial T_i} \right),$$  

in which $k$ is a constant chosen to keep the convergence stable, and $\lambda$ is found from

$$\frac{\partial H}{\partial t} = \sum_i \frac{\partial H}{\partial T_i} \frac{\partial T_i}{\partial t} = \sum_i \left( \frac{\partial H}{\partial T_i} \frac{\partial G}{\partial T_i} - \lambda \frac{\partial H}{\partial T_i} \frac{\partial H}{\partial T_i} \right) = 0$$  

so that

$$\lambda = \sum_i \frac{\partial G}{\partial T_i} \frac{\partial H}{\partial T_i} \left( \sum_i \frac{\partial H}{\partial T_i} \right)^{-2}.$$  

The computed changes to the temperature field are applied simultaneously at the completion of each iteration. Then the additional correction

$$\delta T_i = -\frac{H}{\sum_i \frac{\partial H}{\partial T_i}}$$  

is applied uniformly to this new temperature field at the end of each iteration to maintain $H=0$ more accurately. This is necessary since in general $H$ does not vary linearly in phase space. Hypothetical changes of the temperature field in two-dimensional phase space during the maximization procedure are also displayed in Fig. 5.

As mentioned previously, $t$ does not denote time, but only that the procedure is an iterative one. More specifically, for the $(m+1)$th approximation

$$\delta T_{i}^{m+1} = (k\delta t) \left( \frac{\partial G}{\partial T_i} \frac{\partial H}{\partial T_i} - \lambda \frac{\partial H}{\partial T_i} \frac{\partial H}{\partial T_i} \right).$$  

Values of $k\delta t$ less than 20 K$^2$ W$^{-1}$ m$^2$ were found to lead to convergence.

The “time differencing” is accomplished by an uncentered difference approximation. Lorenz’s (1971) $N$-cycle scheme was programmed into the maximizing procedure, but the convergence with one cycle was found to be as fast, or even faster, than the higher order schemes. One cycle is equivalent to the simple forward difference approximation. The space derivations $\partial G/\partial T_i$ and $\partial H/\partial T_i$ were evaluated using a centered difference approximation with $\Delta T$ equal to 0.01 K and 0.001 K. The results were the same for both temperature increments.

![Fig. 6. The 5-level, 5-latitude model. Independent parameters, other than the solar forcing, are shown on the left-hand side of each pressure level. Computed parameters are shown on the right. Solid lines denote the independent levels.](image-url)
4. The 5-level, 5-latitude model

a. The model

The 5-level, 5-latitude model was found to offer the best compromise between resolution and computer time. Most of the results of the study are obtained with this model.

Atmospheric heating in the model is accomplished by longwave and shortwave radiation and small-scale sensible heat fluxes. Results with shortwave heating will not be discussed until Section 4g. The five independent temperature levels are chosen as 100, 300, 500, 700 and 900 mb so as to coincide with the levels at which $Q$ and $N$ are computed. The isothermal layer extends from 100 mb to the top of the atmosphere. The five latitudes chosen afford full hemispheric coverage. They are represented by $s = 0.0, 0.25, 0.50, 0.75$ and 1.0 ($\phi = 0.0^\circ, 14.5^\circ, 30.0^\circ, 48.6^\circ$ and $90.0^\circ$). $G$ is then evaluated by the trapezoidal integration method which doubly weights the three interior latitudes. The full model is depicted in Fig. 6.

The horizontal resolution is temporarily increased in the computation of $N$. The potential temperatures at the five latitudes are linearly interpolated in the horizontal to a finer grid of the 41 latitudes $s = 0.0, 0.025, 0.050, \ldots 0.950, 0.975, 1.000$. With the additional nine latitudes between each of the independent latitudes the isentropic surfaces can be described more precisely, and better estimates of $\theta_R$ can be obtained. This is especially true when adiabatic or near adiabatic lapse rate conditions prevail at the independent latitudes.

b. Basic assumed values and schemes

The maximum value of $G$ is computed assuming relative humidity $R = 50\%$, planetary albedo $\alpha = 30\%$. 
surface wind speed $V_s = 10 \text{ m s}^{-1}$ and solar declination $\delta = 0$. $K_H$ is defined by the Fisher and Caplan formulation with $K_H^* = 15 \text{ m}^2 \text{s}^{-1}$ and $\epsilon = 130 \text{ m K}^{-1}$. The initial temperature field used in the maximization scheme and the corresponding fields of $\theta$, $N$ and $Q$ are given in Fig. 7. Somewhat resembling a very cold atmosphere, they satisfy the requirement $H = 0$ and yield an initial generation of $7.3 \text{ W m}^{-2}$.

Applying the maximization procedure to the initial temperature field a maximum generation of $12.2 \text{ W m}^{-2}$ was found. Its accompanying fields are shown in Fig. 8. The temperature field is smooth and has the same general magnitude and horizontal and vertical gradients as the observed zonally averaged field (see Fig. 1). The polar region lapse rates, as evidenced by the decreasing horizontal temperature gradient with height, are more stable than the equatorial lapse rates. Also, no superadiabatic lapse rates are present. When they were arbitrarily imposed, the sensible heating was invariably concentrated in the upper levels or regions of smaller $N$. Even with the increased radiational heating at the lower levels, negative generations or small positive generations resulted. These similarities appear to lend support to the hypothesos that the atmosphere may be constrained to operate at nearly its maximum efficiency.

Other observed features are not well represented. For example, the maximizing temperature field is too cold. However, these cold temperatures are not surprising in view of the $5-10 \text{ K}$ ground-air temperature difference necessary to supply the balancing sensible heat fluxes. With the increased outgoing longwave radiation due to the warm ground temperatures, colder atmospheric temperatures are needed to maintain $H = 0$. Still, the gross characteristics of the maximizing field are encouragingly similar to the observed zonal field.

During the maximization procedure a field yielding a $G$ of $11.9 \text{ W m}^{-2}$ after 100 iterations was thought to be the maximizing field. This field still manifested the reversal in the horizontal temperature gradient at 100
Table 1. Local contributions to the integrand QN by the maximizing fields of Fig. 8 and the resulting weighted latitudinal contributions. Units are $10^{-4}$ W kg$^{-1}$.

<table>
<thead>
<tr>
<th>$p$ (mb)</th>
<th>0.0</th>
<th>0.25</th>
<th>0.50</th>
<th>0.75</th>
<th>1.00</th>
</tr>
</thead>
<tbody>
<tr>
<td>100</td>
<td>-1.6</td>
<td>-3.7</td>
<td>2.2</td>
<td>3.9</td>
<td>3.7</td>
</tr>
<tr>
<td>300</td>
<td>-21.2</td>
<td>4.1</td>
<td>15.4</td>
<td>33.7</td>
<td>45.9</td>
</tr>
<tr>
<td>500</td>
<td>-23.1</td>
<td>-2.9</td>
<td>6.2</td>
<td>14.9</td>
<td>19.1</td>
</tr>
<tr>
<td>700</td>
<td>-22.2</td>
<td>-13.1</td>
<td>-2.1</td>
<td>7.8</td>
<td>9.7</td>
</tr>
<tr>
<td>900</td>
<td>127.9</td>
<td>70.7</td>
<td>34.2</td>
<td>-2.8</td>
<td>5.4</td>
</tr>
</tbody>
</table>

Weighted latitudinal contributions: 59.8 110.3 111.7 115.0 83.9

mb that is present in the initial field. However, after many more iterations the temperature gradient reversal slowly disappeared and the maximizing fields of Fig. 8 emerged. The only changes in the temperature field during the transition from 11.9 to 12.2 W m$^{-2}$ were in the 100 mb temperatures. The result of main importance, of course, is that the maximizing horizontal temperature gradient at 100 mb is very small, regardless of its direction. The direction is probably more a product of the modeling assumptions, i.e., an isothermal layer from 100 mb to the top of the atmosphere, than of the physical processes. This feature of the convergence scheme was not discovered until after many of the study computations had been completed. The other maximization experiments using this initial temperature field (with the exception of the experiment described in Section 4g) were carried out to only 100 iterations, and thus no conclusions should be implied about the resulting directions of the 100 mb temperature gradients.

The 0.3 W m$^{-2}$ increase in $G$ following the disappearance of the temperature gradient reversal is indicative of the small contribution to the generation from the 100 mb level. Local contributions to $G$ from other regions are shown in Table 1. The largest positive contributions are from 900 mb of the equatorial latitude and 300 mb of the polar latitude. However, when the individual contributions within each latitude are summed all latitudes contribute nearly equally.

The contributions to $G$ from the different diabatic heating sources can be seen from Fig. 9 which presents the components of the maximizing heating field. Infrared cooling causes a net destruction of APE, while sensible heating yields a net generation of APE. The large positive sensible heating in the regions of positive $\vec{N}$ and substantial negative sensible heating in the regions of negative $\vec{N}$ are the largest positive contributors. Longwave cooling in the tropical regions is the largest negative contributor.

The maximizing temperature field is somewhat too cold, and is a direct result of the high surface temperatures needed to supply the balancing sensible heat fluxes. The contribution to the outgoing longwave flux attributable solely to the surface temperature, $\sigma[T_0^4 - T(\rho_0)k]$, can be as large as 50 W m$^{-2}$, or almost 15% of the total outgoing longwave flux. The radiative and sensible heat fluxes at the surface and at the top of the atmosphere for the maximizing temperature field are given in Fig. 10, and the flux balances at both atmospheric boundaries can easily be seen. The excess of incoming solar radiation over the outgoing longwave radiation in the lower latitudes is smaller than observed in the real atmosphere because of the large surface contribution to the longwave flux. Fluxes of sensible heat are smaller than the longwave fluxes at all latitudes and in fact become negative near the pole where the incident solar flux goes to zero.

c. Different sensible heating schemes

Two alternate sensible heating schemes that arbitrarily distribute a larger portion of the sensible heat through a greater depth of the atmosphere are substituted for the eddy diffusion scheme in the maximization experiments. In this manner the sensitivity of the model results to the sensible heating scheme can be explored.

For a given temperature profile the sensible heat flux from the ground to the atmosphere, $F_{4p}$, is uniquely

![Fig. 9. The components of the maximizing heating field of Fig. 8.](image-url)
determined. The two alternate schemes assign a specific fraction of that heating to specific model layers. The first scheme, referred to hereafter as the \((\frac{3}{4}, \frac{1}{4})\) scheme, places \(\frac{1}{4}\) of the sensible heating in the 1000–800 mb layer and \(\frac{1}{4}\) in the 800–600 mb layer by designating

\[
F_d(800\text{ mb}) = 0.25 F_d(400\text{ mb}) \\
F_d(600\text{ mb}) = F_d(400\text{ mb}) \\
= F_d(200\text{ mb}) = F_d(0\text{ mb}) = 0
\]

The second scheme, referred to hereafter as the \((\frac{1}{2}, \frac{1}{3}, \frac{1}{6})\) scheme, distributes \(\frac{1}{4}\) of the sensible heating in the 1000–800 mb layer, \(\frac{1}{6}\) in the 800–600 mb layer, and \(\frac{1}{3}\) in the 600–400 mb layer by assigning

\[
F_d(800\text{ mb}) = 0.50 F_d(400\text{ mb}) \\
F_d(600\text{ mb}) = 0.17 F_d(400\text{ mb}) \\
F_d(400\text{ mb}) = F_d(200\text{ mb}) = F_d(0\text{ mb}) = 0
\]

Assigning the sensible heating to the upper level regions of lower \(N\) and eliminating the cooling by eddy diffusion in those regions lowers the rates of maximum generation.

![Fig. 10](image1.png)

**Fig. 10.** The radiative and sensible heat fluxes at the top of the atmosphere (top) and at the ground (bottom) associated with the maximizing fields of Fig. 8.

![Fig. 11](image2.png)

**Fig. 11.** The maximizing temperature fields obtained using the \((\frac{1}{4}, \frac{1}{4})\) and \((\frac{1}{2}, \frac{1}{3}, \frac{1}{6})\) sensible heating schemes in the 5-level, 5-latitude model.

Maximization experiments using the \((\frac{3}{4}, \frac{1}{4})\) and \((\frac{1}{2}, \frac{1}{3}, \frac{1}{6})\) sensible heating schemes were conducted with the 5-level, 5-latitude model and resulted in generations of only 6.5 and 5.6 W m\(^{-2}\), respectively. The associated temperature fields are given in Fig. 11.

The most notable changes brought about by the alternate sensible heating schemes occur below 500 mb. The increase in the lower level polar temperatures decreases the amount of positive sensible heating that must be distributed upward into the layers of large negative \(N\). The decrease in the lower level tropical temperatures then serves not only to maintain \(H = 0\), but also to increase the positive sensible heating in

<table>
<thead>
<tr>
<th>(V_0) (m s(^{-1}))</th>
<th>(G_{max}) (W m(^{-2}))</th>
</tr>
</thead>
<tbody>
<tr>
<td>5</td>
<td>10.4</td>
</tr>
<tr>
<td>10</td>
<td>12.2</td>
</tr>
<tr>
<td>20</td>
<td>13.1</td>
</tr>
<tr>
<td>40</td>
<td>14.0</td>
</tr>
</tbody>
</table>
regions of positive $N$. It is the reduced polar contribution resulting from a positive sensible heating in the regions of large negative $N$ and the absence of diffusive cooling in those same regions which is largely responsible for the decrease in $G$.

Several observations should be made. First, from a quantitative viewpoint, the magnitude of the maximum $G$ appears to be sensitive to the sensible heating scheme chosen, varying by over a factor of 2. However, taken as an efficiency the variation is less than 2%. Also it should be noted that the scheme assumed for the main study results yields the highest magnitude. In fact it can be inferred that the greater the volume over which the sensible heat is distributed, the lower the resulting maximum generation. Second, from a qualitative viewpoint, the maximizing temperature fields obtained with the alternate sensible heating schemes do not agree as well with the observed zonally averaged field in the lower levels as does the maximizing temperature field obtained with the eddy diffusion scheme. However, the accompanying $N$ fields do appear to be more realistic, especially with regard to the occurrence of the positive maximum in the middle levels of the tropics rather than at the surface.

\textit{d. Variations of the sensible heating parameters}

The eddy diffusion sensible heating scheme has used the Fisher and Caplan formulation of $K_H$ in the

\begin{table}[h]
\centering
\begin{tabular}{|c|c|}
\hline
$K_H^*$ & $G_{\text{max}}$ \\
(m$^2$s$^{-1}$) & (W m$^{-2}$) \\
\hline
3 & 7.6 \\
5 & 9.0 \\
10 & 10.1 \\
15 & 12.2 \\
45 & 23.5 \\
\hline
\end{tabular}
\caption{Variation of maximum $G$ with $K_H^*$ in Fisher and Caplan formulation of $K_H$ ($e=150$ m K$^{-1}$).}
\end{table}

\begin{figure}[h]
\centering
\includegraphics[width=\textwidth]{figure.png}
\caption{Different relative humidity fields used in the 5-level, 5-latitude model and the resulting maxima of $G$ in units of W m$^{-2}$.}
\end{figure}
computation of $F_d(p)$. The aerodynamic bulk formula has assumed $V_0 = 10\, \text{m s}^{-1}$ in the computation of $F_d(p_0)$. These parameters are varied to determine the sensitivity of the diffusion scheme to their values.

The surface wind speed cannot be increased arbitrarily. Since the rate at which kinetic energy is dissipated by friction should be correlated with $V_0$ and cannot exceed $G$, there should be some upper bound on the magnitude of $V_0$. Using classical Ekman theory (see Schulman, 1974) the dissipation per unit area in the friction layer can be written

$$D = (2\mu f p)^{1/2} (l + F) V_0^2,$$

in which $\mu$ is the coefficient of turbulent viscosity, $f$ the Coriolis parameter, and

$$l = \frac{\sin \psi}{\cos \psi - \sin \psi},$$

where $\psi$ is the angle between the surface wind and the geostrophic wind. With the values $\mu = 1\, \text{kg m}^{-1}\, \text{s}^{-1}$ (conservative estimate), $f = 10^{-4}\, \text{s}^{-1}$, $p = 1\, \text{kg m}^{-3}$ and $\psi = 25^\circ$

$$D = G = (23.3 \times 10^{-3}\, \text{W m}^{-2}\, \text{s}^2) V_0^2.$$

Letting $G = 25\, \text{W m}^{-2}$, the corresponding upper bound surface wind speed is $\sim 33\, \text{m s}^{-1}$. Ignoring dissipation in the free atmosphere and choosing $40\, \text{m s}^{-1}$ as the upper bound of $V_0$ for all generations below $25\, \text{W m}^{-2}$, maximization experiments were conducted and the results shown in Table 2.

The magnitude of maximum $G$ does not vary appreciably with $V_0$. It appears to increase only about 10% with each doubling of $V_0$. Increasing the ventilation for a given temperature profile decreases the value of $T_0$ necessary to provide the same surface sensible heat flux. However, since a smaller surface longwave flux is also associated with the decreased $T_0$, a larger value of $F_d(p_0)$ is needed to maintain the surface flux balance. It is this increased sensible heating accompanying the
larger values of $V_0$ which increases the generation rates. Smaller $T_p$ also reduce the surface contributions to the outgoing longwave flux so that a warmer atmosphere can exist without violating the requirement that $H = 0$.

The eddy conductivity $K_H$ can be varied by changing the values of $K^*_H$ or $\varepsilon$ in the Fisher and Caplan formulation or by using a different formulation. The results of maximization experiments conducted with different values of $K^*_H$ are presented in Table 3. Results using the Asai formulation of $K_H$ described in Section 3b are presented in Table 4.

Increasing $K^*_H$ above the assumed value of 15 m$^2$ s$^{-1}$ greatly affects the magnitude of the maximum $G$. Since $F_\delta(\rho)$ is always negative in a stable atmosphere, increasing the values of $K_H$ increases the sensible heat flux convergence in the 1000–800 mb surface layer and increases the divergence in the upper layers. For example, when $K^*_H$ was tripled, the sensible heating associated with the maximizing fields was almost 10 K day$^{-1}$ in the 1000–800 mb layer and $-4$ K day$^{-1}$ in the 400–200 mb layer. This resulted in a maximum generation of 23.5 W m$^{-2}$. On the other hand, decreasing $K^*_H$ reduced the values of the heating and cooling and somewhat lowered the maximum generations. The Asai formulation also generally led to lower rates of generation with its more moderate values of sensible heating.

c. Variation of the relative humidity field

Maximization experiments varying the field of relative humidity were conducted and the results are presented in Fig. 12. The magnitude of the maximum
generation increases with the relative humidity (see Fig. 12a) and the largest value occurs with \( R = 100\% \) everywhere (with no condensation). However, this \( G \) value is only 7\% higher than the 12.2 W m\(^{-2}\) maximum accompanying \( R = 50\% \) everywhere.

On closer inspection of Fig. 12b, it is the changes in the magnitude of the relative humidity in the middle and lower tropical levels to which \( G \) is most sensitive. Increasing \( R \) in this region increases the longwave counter radiation and thus decreases \( F_s(p_0) \). The surface flux balance must then be maintained by an increase in \( F_s(p_0) \), which leads to increased sensible heating in the 1000-800 mb layer of large positive \( N \). Increasing \( R \) also leads to increased infrared cooling, but almost all of the additional cooling occurs in the upper layers where smaller positive values of \( N \) prevail. The net effect is an increase in the rate of generation. However, since the magnitude of the increased sensible heating and infrared cooling is only on the order of 0.5 K day\(^{-1}\) for a fivefold increase in \( R \), the increases in generation are small.

\[ f. \text{ Variations of the albedo} \]

The incident solar forcing can be changed by varying the distribution of albedo with latitude. Maximization experiments exploring the effect of these changes on the maximum \( G \) were conducted and the results are shown in Fig. 13. As would be expected, the magnitude of the maximum generation increases as the solar forcing contrast with latitude increases. Raising the incoming solar flux where it is already large or lowering the incoming solar flux where it is already small increases the magnitude of the heating and cooling, respectively, in those regions and thus increases the generation of APE. This effect can best be seen by a comparison of the \( F_\text{s}(s) \) slopes for the 13.2, 16.2 and 17.8 W m\(^{-2}\) maxima in Fig. 13.

---

**Fig. 14.** The maximizing fields of \( T, N, Q_s \) and \( Q \) for the 5-level, S-latitude model with solar absorption.
Table 4. Variation of maximum $G$ with $\epsilon$ in Asai formulation of $K_H$ ($K_H' = 10$ m$^2$ s$^{-1}$).

<table>
<thead>
<tr>
<th>$\epsilon$ (m K$^{-1}$)</th>
<th>$G_{\text{max}}$ (W m$^{-2}$)</th>
</tr>
</thead>
<tbody>
<tr>
<td>500</td>
<td>8.9</td>
</tr>
<tr>
<td>1000</td>
<td>10.5</td>
</tr>
</tbody>
</table>

The rate of the maximum generation does not appear to be as sensitive to changes in the albedo distributions as it is to the sensible heating scheme. Over the range of realistic albedo distributions the results show the variation of the maximum $G$ to be less than 30%.

g. Results with solar absorption

Thus far only longwave and sensible heat fluxes have been included in determining the atmospheric heating. In this section the effects of solar absorption are added. Applying the maximization procedure to the initial temperature field in Fig. 7 yields a maximum generation of 11.5 W m$^{-2}$. Its associated fields, including the shortwave heating contribution to the total diabatic heating, are shown in Fig. 14. The maximizing temperature field is a few degrees warmer at the equator and a few degrees cooler at the pole than the maximizing field without solar absorption presented in Fig. 8. This increased horizontal temperature contrast serves to increase the magnitude of positive $N$ at the equatorial latitude where the additional heating from the solar absorption is taking place. The contrast is limited by the increased longwave cooling also accompanying the warmer equatorial temperatures.

Even with the increased contrast the magnitude of the maximum $G$ is still 0.7 W m$^{-2}$ lower with solar absorption because of the decreased net heating in the 1000–800 mb layer at the low latitudes. The depletion of the solar beam by water vapor decreases the value of $F_s(\phi)$ needed for the surface flux balance by as much as 35 W m$^{-2}$. The accompanying drop in sensible heating in the 1000–800 mb layer outweighs the shortwave heating in that region, and is the dominant effect in determining the total generation.

h. Different initial temperature fields

The identical initial temperature field was used in all the preceding maximization experiments. This initial field, however, is not radically different from the resulting maximizing field, and the possibility of higher maxima in other regions of the temperature-phase space should not be overlooked. Two other special initial temperature fields are used to investigate this possibility. The two fields chosen for this purpose are shown in Fig. 15. The first field is a mirror image about $s = 0.5$ of the original initial field presented in Fig. 7. The associated net heating at small $N$ and net cooling at large $N$ reduce the initial generation to 1.3 W m$^{-2}$. The second field consists of a representative mid-latitude profile at all latitudes so that there is no horizontal temperature contrast and $N$ is equal to zero everywhere. The initial generation is, of course, zero.

Maximization experiments with solar absorption were conducted with each of the two alternate initial temperature fields and convergence to nearly the same maximizing fields of Fig. 14 and its associated generation of 11.5 W m$^{-2}$ were attained. Other experiments with initial fields uniformly 20 K higher and lower than the original initial fields yielded the same results. This suggests that the previously computed maximizing fields for the 5-level, 5-latitude model are indeed the primary maxima.

5. A 5-level, 9-latitude model with solar absorption

A further refinement of the maximizing fields was achieved by increasing the horizontal resolution. Nine independent latitudes were specified at $s = 0.0, 0.125, 0.25, 0.375, 0.50, 0.625, 0.75, 0.875$ and 1.0 ($\phi = 0.0^\circ, 7.2^\circ, 14.5^\circ, 22.0^\circ, 30.0^\circ, 38.7^\circ, 48.6^\circ, 61.0^\circ$ and $90.0^\circ$). With the exception of the additional independent

![Fig. 15. Two alternate initial temperature fields used in the 5-level, 5-latitude model with solar absorption.](image-url)
latitudes the model is identical to the 5-level, 5-latitude model. However, since the computation time was nearly tripled only a few experiments were conducted.

The initial temperature field used in the 5-latitude model and presented in Fig. 7 is also used in the 9-latitude model. The corresponding initial generation, however, is 6.8 W m$^{-2}$, or 0.2 W m$^{-2}$ less than was found with the 5-latitude model. This decrease is simply a result of the changed resolution. The large positive contribution to the generation integral from the equator and pole is given decreased weight because of the additional independent latitudes. The values of $Q$ and $N$ at these intermediate latitudes do not have as high a covariance as at the equator and pole and thus lead to the difference in generation rates between the two models.

The moderating influence of the additional latitudes appears also to have affected the resulting maximum generation. Applying the maximization procedure to the initial temperature field and using the basic assumed values, a maximum generation of 10.3 W m$^{-2}$ was found. This is fully 1.2 W m$^{-2}$ lower than the maximum found with the 5-latitude model with solar absorption, and represents the best estimate of the maximum rate of generation of this study. However, the direction of the horizontal temperature gradient at 100 mb should not be considered significant since only 100 iterations were performed in the maximization procedure. The accompanying fields of $T$, $\theta$, $N$ and $Q$ are shown in Fig. 16.

The fields are very similar to those found with the 5-latitude model (Fig. 14) and possess the same general characteristics. Encouragingly, some of the features of the 9-latitude maximizing temperature field compare even more favorably with the observed zonally averaged field (Fig. 1) than do those found with the 5-latitude model. The most notable such feature is the smaller horizontal temperature gradient at the surface. This is brought about by an equatorial surface temperature 4 K cooler than in the 5-latitude maximizing field and
Fig. 17. The radiative and sensible heat fluxes at the top of the atmosphere (top) and at the ground (bottom) associated with the maximizing fields of Fig. 16.

surface temperatures from \( s = 0.125 \) to \( s = 1.0 \) that are progressively 1–10 K warmer. The decreased temperature gradient is most evident in the low latitudes. From the equator to \( s = 0.25 \) the surface temperature drops only 5 K, less than half of the 12 K drop found with the 5-latitude model. The observed mean field typically decreases 1–2 K over the same interval. A direct consequence of the smaller contrast is less extreme values of \( N \) near the equator and pole.

The 9-latitude model atmosphere, in general, is a few degrees warmer than previous models. However, because of the large outgoing longwave flux, it is still colder than the observed temperature field. The radiative and sensible heat fluxes describing the balances at the surface and 0 mb are presented in Fig. 17. The selective absorption of the incoming shortwave flux in low latitudes and the subsequent decreased sensible heating of the 1000–800 mb layer in those regions can easily be seen.

Other experiments using the Asai formulation of \( K_H \) and the \( (\bar{q}, \bar{q}) \) sensible heating scheme were conducted, and yielded maximizing fields similar, yet more refined, than with the 5-latitude model. The results of the 5-level, 9-latitude model give further reason for believing the hypothesis that the atmosphere may be constrained to operate at nearly maximum efficiency.

6. Summary and conclusions

This study has examined the problem of explaining the atmospheric efficiency, or equivalently, the rate of generation of APE. However, as seen earlier, a solution to this problem would require an explanation of the atmospheric temperature distribution, which is no less a problem than explaining the general circulation itself. This has been circumvented by considering the less complex problem of determining the maximum possible rate of generation of APE and its accompanying temperature field, which does not require a knowledge of the atmospheric motions leading to that temperature distribution. Using this approach a preliminary study by Lorenz (1960) has suggested that the atmosphere may already be operating at nearly its maximum efficiency. Because of the importance of such a conclusion a much more thorough investigation of this hypothesis has been made.

To make the problem tractable an atmosphere was assumed in which the relative humidity was constant and less than 100%, ozone was ignored as an atmospheric constituent and the observed distribution of carbon dioxide was used. These assumptions eliminated the release of latent heat and clouds, and for a specified solar forcing made the atmospheric efficiency a function of the temperature field only. Since the solar forcing does not vary with longitude it was argued that essentially two-dimensional models could be used to determine the maximum efficiency. It followed that the three-dimensional maximizing field with eddies would always lead to a somewhat lower efficiency.

Two simple two-dimensional numerical models yielded maximum efficiencies ranging from 2–5%. These computed maximum efficiencies are comparable to the actual atmospheric efficiency of approximately 1–2% indicated by observational studies. Most importantly, the features of the observed atmosphere and the maximizing fields of the models bore more than just a vague resemblance. Several prominent atmospheric features were duplicated. Foremost among these was the absence of superadiabatic lapse rates and a horizontal temperature gradient whose magnitude decreased with height, culminating in a small horizontal contrast near 100 mb. This resulted in a greater static stability in polar latitudes. With the highest resolution model the surface horizontal temperature contrast at low latitudes was also found to be small.

In addition to the horizontal and vertical gradients, the general magnitude of the maximizing temperature fields also agreed well with observations. They tended to be somewhat cold, but this was largely due to the modeling assumptions regarding the surface sensible heat flux. As would be hoped, the maximizing fields of
the most sophisticated and highest resolution model compared most favorably with the observed fields.

Many model parameters and schemes were varied to explore the sensitivity of the resulting maxima. The maxima were found to be relatively insensitive to variations of relative humidity, albedo or surface wind speed, but did display a strong dependence on the sensible heat distribution scheme. Yet, even the highest maxima attained with these variations corresponded to an efficiency of only about 5%. With the most realistic parameter values the maximum efficiency was more nearly 3%. Furthermore, when the sensible heating was arbitrarily distributed through a greater depth of the model, instead of being mainly confined to the lowest layer, the maximum efficiencies dropped to less than 2%.

These qualitative and quantitative results indicate that the atmosphere may be operating within approximately a factor of 2 of its maximum efficiency.

Other evidence supporting the hypothesis was furnished by the results of experiments which used different initial temperature fields in the maximization procedure. Several initial field exhibiting no features resembling the observed field were employed to test the possibility of larger maxima in other regions of the temperature-phase space. The generation rates of most of these fields were negative. However, in all cases convergence to the same maxima and their realistic temperature fields was found.

Assuming that the atmosphere is constrained to operate at nearly its maximum efficiency, the fundamental problem now reduces to an explanation of why this should be so. Lorenz (1960, 1967) has suggested that the less efficient circulation modes may be unstable and give way to the more efficient modes. Of course, the possibility exists that the atmosphere may maximize some other quantity closely related to the efficiency. Such a quantity might be more nearly maximized with the eddies present. These questions have not been addressed in this study.

The next logical extension of this research would seem to be some representation of the moist processes. An independent water vapor field which allows for condensation, the release of latent heat and clouds on all scales would certainly be difficult to model. Not only will it require new parameterization schemes, but the simultaneous computation of the maximizing fields of temperature and water vapor. However, it should add further worthwhile insight into the problems of atmospheric energetics. Perhaps such experiments might someday even be performed with a general circulation model.

Acknowledgments. I am indebted to Professor Edward N. Lorenz, who not only suggested the research undertaken, but also provided advice and constant encouragement until its completion.

I am additionally grateful to the National Center for Atmospheric Research at whose facilities much of the computing was accomplished.

This research was supported in part by the National Aeronautics and Space Administration under Grant NGR 22-009-727, and by a National Science Foundation Graduate Fellowship.

REFERENCES


