Studies of Large Scale Atmospheric Energetics

Department of Meteorology
The University of Wisconsin-Madison
Madison, Wisconsin 53706

Contributions by

B. R. Bullock
R. L. Gall
D. R. Johnson
R. C. Lo
P. J. Schmidt

L. H. Horn and D. R. Johnson, Joint Principal Investigators

ANNUAL REPORT

The research in this document has been supported
by the National Environmental Satellite Center of the
Environmental Science Services Administration under Grant E-8-69-(G)

December 1969

Scanner's note:

This page is blank.

TABLE OF CONTENTS

Scanner's note:

This page is blank.

INTRODUCTION

The research presented in this annual report consists of continuing studies involving the large scale energetics of the atmosphere and the development of a statistical model to infer cloud distribution from satellite information. Because of a delay in the publication of this report, two of the papers, "The Available Potential Energy of Storms" and "The Generation of Available Potential Energy by Latent Heat Release in a Mid-Latitude Cyclone" that appear in this report have already been published in the journals. They are reprinted here with the permission of Journal of Atmospheric Sciences and Monthly Weather Review, respectively.

The studies were initiated under WBG Grant-52 and have been completed with the partial support of ESSA Grant E-8-69(G) and NSF Grant 10997 from the Atmospheric Sciences Section, National Science Administration. Nearly all of the studies were also partially supported by the Wisconsin Alumni Research Foundation and National Science Foundation through the Wisconsin Research Committee in the form of a grant for computational funds and use of computer facilities.

Lyle H. Horn
Donald R. Johnson
Joint Principal Investigators

Madison, Wisconsin
December, 1969

Scanner's note:

This page is blank.

Reprinted from JOURNAL OF ATMOSPHERIC SCIENCES, Vol. 27, No. 5, August 1970, pp. 727–741
American Meteorological Society
Printed in U. S. A.

The Available Potential Energy of Storms[1]

DONALD R. JOHNSON

Dept. of Meteorology, University of Wisconsin, Madison

(Manuscript received 28 November 1969)

The storm's available potential energy and its rate of change are derived for a vertically walled volume encircling the storm and extending from the surface to the top of the atmosphere. The rate of change includes explicit expressions for the generation of the storm's available potential energy, for its conversion to kinetic energy, and for its change through boundary work and energy flux. The theoretical results in isentropic coordinates show that it is not desirable to conduct available potential studies in a quasi-hydrostatic atmosphere for regions of limited vertical extent. The results also show the difficulty of inferring kinetic energy change from a total potential energy budget of a limited atmospheric domain. Opposite time rates of change for the storm's total and available potential energy are allowed through boundary processes for frictionless isentropic flows within mechanically open regimes. The relation of the available potential energy of storms to the available potential energy of the atmosphere is also established.

Several recent diagnostic studies of the generation of the storm's available potential energy by individual diabatic components are summarized for the hurricane and extratropical cyclone. The amount of available potential energy generated within the storm is a significant fraction of the rate of its kinetic energy production.

1. Introduction

In his classic paper, "On the Energy of Storms," Margules (1903) developed the concept of available potential energy in a search for the source of kinetic energy within storms. Margules confined his attention to the smaller quasi-horizontal scales of atmospheric motion and concluded that the storm's kinetic energy is derived from the release of internal and gravitational potential energy within the disturbance and the region surrounding the maximum storm winds. In the same paper he admitted bewilderment over the question of why a store of available potential energy rises without being immediately released by virtue of its existence in an atmosphere in motion. However, he avoided this question in seeking the source of the storm's kinetic energy by assuming an initial stage generalized from observations within a fixed volume surrounding the disturbance.

Lorenz (1955) applied Margules' concept in his development of zonal and eddy components of available potential energy to explain the maintenance of the general circulation. Within this framework, the generation of zonal available potential energy by diabatic processes, its conversion into eddy available potential energy by the action of the waves in the westerlies, and the realization of the kinetic energy at the eddy scale by the sinking of colder air and rising of warmer air at the same latitude are emphasized. The kinetic energy continually being created is used to offset the ever-present frictional dissipation.

In many diagnostic studies the eddy scale within this framework is often identified with the storms that Margules studied. However, since the horizontally invariant, statically stable density stratification of the reference atmosphere used in the definition of available potential energy is based upon an adiabatic redistribution of the entire atmosphere, investigators, including the author, have found it impossible to rigorously define and study the changes of available potential energy of any particular storm, scale and/or region. Since reliable global data are still unavailable, all diagnostic studies of available potential energy involving limited regions of the atmosphere are studies of the "contributions to the global budget" whether it is the generation, conversion or amount of available potential energy. Furthermore, in global studies the energy processes within the smaller scales are subjected to severe averaging because of their small horizontal extent. Thus, there are diagnostic limitations associated with the global concepts in studying the available potential energy within the storm. In a comparative sense it is within storms that not only is the genesis and frictional dissipation of kinetic energy a maximum, but both the horizontal and vertical differential heating attain extreme values. The probability also exists that the generation of energy available for conversion to kinetic energy within the storm is also a relative extremum.

Following the development of Dutton and Johnson (1967), the concept of available potential energy of the storm and/or limited regions of the atmosphere is developed in this paper. The energetics of the domain of interest, an open system, will be explicitly separated from the global domain. The interactions with the larger domain occur through energy advection by mass transport and pressure work at the lateral boundaries. In so doing, the available potential energy of the storm

[1] This paper was partially completed while the author was a visiting associate professor with the Department of Meteorology, Pennsylvania State University, during the academic year 1968–69.

can now be defined and studied within the framework of Margules' original concept.

After the theory of open systems and diagnostics results for several storms are presented, the relationship between the available potential of the storm and the global concept is developed. This relationship shows that the available potential energy of the global concept is the sum of the storm concept of available potential energy from all the open regions that together comprise the entire atmosphere plus a component that is due to thermodynamic contrasts between regions. From one viewpoint, this represents a scale decomposition of the global available potential energy. Diagnostic results from such an approach may be useful for the study of energy processes of various scales as well as scale interactions if the kinetic energy of the atmosphere is correspondingly partitioned. However, the basic development for the available potential and kinetic energy of the open system presented in this study is based on fundamental energy principles which are independent of scale decomposition. Thus, the theoretical and diagnostic results are valid for open systems and are independent of the global concept.

The necessity of another derivation of available potential energy with some modifications from that of Dutton and Johnson (1967) is to rigorously prove that the available potential energy theory of an open system can be developed from Margules' definition and to provide a firm basis for several diagnostic studies already completed at the University of Wisconsin. Smith (1969) formulated the energy budget equations of the contribution of available potential energy from a fixed volume to the global amount. However, since he developed the equations in pressure coordinates he was not able to isolate the time dependency of a local reference state and was forced to use estimates of the global reference state from other time periods. The problems with the translation of the storms through the budget volume, with boundary fluxes, and the generation estimates that were encountered in the interpretation of the time series of the energy statistics (Smith and Horn, 1969), revealed that the optimum approach would be to isolate the storm as a separate entity. Generation estimates already completed by Anthes and Johnson (1968) for a hurricane and by Bullock and Johnson (1970), and Gall and Johnson (1969) for an extratropical cyclone using the concept of the storm's available potential energy have been revealing. Their estimates show that the *in situ* generation for the storm's available potential energy is sufficient to offset a significant portion of the frictional dissipation within the storm. In addition, Hahn and Horn (1969) have completed a study of the evolution of the storm's efficiency factor distribution which suggests that, due to the changing mass potential temperature distribution, the effects of the various diabatic components will change with the growth, maturation and decay of the storm.

2. The concept of the available potential energy of a storm

Following Margules, the storm's available potential energy is defined to be the difference between the sum of the internal and potential energies within a volume encircling the storm and the sum of these energies which would exist after an isentropic redistribution to a horizontally invariant, statically stable density stratification in the absence of lateral boundary work and advection. The volume encircling the storm must extend from the earth's surface to the top of the atmosphere and its lateral boundaries must be vertical. The dimensional constraints of the vertical boundaries are required implicitly from the hydrostatic specification of the mass distribution of the hypothetical stratified atmosphere. The stratified state will be called the reference state. No restrictions are imposed on the cross-sectional area of the region. For the conceptual development in this paper the cross-sectional area may be assumed to be a circular area surrounding and moving with the center of the storm. The theory for the open system allows relative mass transport through the lateral boundaries. However, as Margules notes, thermal and mechanical insulation for the open system, and isentropic motion within the volume, must be assumed in order to define a unique amount of energy in the reference state. The absence of lateral boundary work and advection satisfies the constraints of mechanical and thermal insulation.

All of the assumptions used to define the reference state energy pose difficult conceptual ideas if one imagines that kinetic energy is realized only when the atmosphere tends toward the reference state when these conditions are satisfied. However, as Dutton and Johnson (1967) note, there is no reason to expect that the atmosphere on a rotating earth in the presence of heating tries to reach a hypothetical flat reference state, and the theory is useful for studying the processes which maintain the available potential and kinetic energy supply against the ever-present frictional destruction of organized motion.

The mathematical theory for the concept of the available potential energy of a storm follows the development of Dutton and Johnson hereafter referred to as D & J. In this paper, identical assumptions, fundamental relations and properties of the reference state are employed (see Sections 2.2 and 2.3 of D & J). D & J's development is based on the condition that the mass-potential temperature distribution is preserved in the hypothetical flow toward the reference state. This relation is expressed by

$$\beta=\bar{m}(\theta_T,\theta)=m_r(\theta_T,\theta), \tag{1}$$

where the mass distribution function for the storm region is defined by

$$\bar{m}(\theta_T,\theta)=\int_{\theta}^{\theta_T}\overline{\rho J}\,d\theta, \tag{2}$$

$$m_r(\theta_T,\theta)=\int_{\theta}^{\theta_T}\rho_r J_r d\theta, \tag{3}$$

where θ is potential temperature, ρ density, and J the transformation Jacobian, $\partial h/\partial\theta$, between the height h of the isentropic surface and potential temperature. The subscript r denotes variables of the hypothetical reference state. The overbar denotes the ratio of the value of an integral over an isentropic surface within the region of interest to its projected horizontal cross-sectional area. From (2) and (3) the conservation constraint of the mass potential temperature distribution also requires, on each isentropic surface, that

$$\rho_r J_r=\overline{\rho J}. \tag{4}$$

By definition, the available potential energy of the storm is

$$A=\pi-\pi_r, \tag{5}$$

where the total potential energy, internal plus gravitational, for the natural and reference states of the region of interest are, respectively,

$$\pi=\int_{V_\theta}(e+\phi)\rho J dV_\theta, \tag{6}$$

$$\pi_r=\int_{V_\theta}(e_r+\phi_r)\rho_r J_r dV_\theta. \tag{7}$$

The specific internal and gravitational potential energies are e and ϕ, and V_θ is the volume of integration in isentropic coordinates encircling the storm and extending vertically from the lowest observed potential temperature θ_0 to θ_T equal to infinity. The incremental volume dV_θ is $dxdyd\theta$. Through the definition of hydrostatic balance for the reference state and the mass-potential temperature relation (1), D & J have shown that the variables of the reference state are

$$p_r=g\beta, \tag{8}$$

$$T_r=(g/p_{00})^\kappa\theta\beta^\kappa, \tag{9}$$

$$\phi_r=-c_p(g/p_{00})^\kappa\int_{\theta_0}^{\theta}\theta\frac{\partial\beta^\kappa}{\partial\theta}d\theta, \tag{10}$$

where κ is the ratio of R, the gas constant to c_p, the specific heat at constant pressure, g is the acceleration of gravity and p_{00} is 1000 mb.

These relations and (4) are sufficient to completely specify the reference state total potential energy (7) from the mass-potential temperature distribution of the actual atmosphere. In (2.48) of their monograph, D & J have shown that

$$\pi_r=c_p(g/p_{00})^\kappa\int_{V_\theta}\theta\beta^\kappa\rho J dV_\theta. \tag{11}$$

From an equivalent expression for the total potential energy [Eq. (2.11) in D & J], and Eqs. (5) and (11), the storm's available potential energy is

$$A=\frac{c_p}{p_{00}^\kappa}\int_{V_\theta}\theta(p^\kappa-p_r^\kappa)\rho J dV_\theta+\int_{V_\theta}\frac{\chi}{g}\phi\rho J dV_\theta, \tag{12}$$

where χ is the hydrostatic defect defined in D & J by

$$\chi=-\frac{1}{\rho}\frac{\partial p}{\partial z}+g. \tag{13}$$

Under the assumption of local hydrostatic balance, D & J show that (12) becomes

$$A=[c_p/gp_{00}^\kappa(1+\kappa)]\int_0^\infty\int_\sigma(p^{1+\kappa}-\bar{p}^{1+\kappa})d\sigma d\theta, \tag{14}$$

which is the equivalent to Lorenz' expression for the available potential energy. The horizontal surface area of V_θ is σ.

3. The rate of change of the storm's total potential energy

The total potential energy change of the moving volume is determined through a generalization of Leibnitz's rule for the differentation of integrals. In general, let the integral over the volume be defined by

$$F(t)=\int_{V_\theta}f\rho J dV_\theta, \tag{15}$$

where f is a specific quantity, for example, energy per unit mass. The time rate of change of $F(t)$ is

$$\frac{d}{dt}F(t)=\int_{V_\theta}\frac{\partial}{\partial t_\theta}(f\rho J)+\nabla_\theta\cdot(f\rho J\mathbf{W})+\frac{\partial}{\partial\theta}\left(f\rho J\frac{d\theta_B}{dt}\right)dV_\theta, \tag{16}$$

where $\mathbf{W}$ is the velocity of the lateral boundary of the moving volume and $d\theta_B/dt$ represents the movement of the vertical boundary with respect to the vertical coordinate θ. The first term represents changes inside the region while the last two terms are associated with the divergence of the region of interest. Use of the equation of continuity,

$$\frac{\partial}{\partial t_\theta}\rho J+\nabla_\theta\cdot(\rho J\mathbf{u})+\frac{\partial}{\partial\theta}\left(\rho J\frac{d\theta}{dt}\right)=0, \tag{17}$$

allows the alternate form for (16),

$$\frac{d}{dt}F(t)=\int_{V_\theta}\rho J\frac{df}{dt}+\nabla_\theta\cdot[f\rho J(\mathbf{W}-\mathbf{u})]+\frac{\partial}{\partial\theta}\left[f\rho J\left(\frac{d\theta_B}{dt}-\frac{d\theta}{dt}\right)\right]dV_\theta. \tag{18}$$

The moving volume considered in this study is bounded by vertical walls and includes the entire vertical extent of the atmosphere. The condition that the lower boundary becomes a material surface is insured by equating $(d\theta_B/dt)$ to $(d\theta/dt)$ (x,y,θ_s, T), where θ_s is the observed surface potential temperature. At the upper boundary, $\rho J \rightarrow 0$ and $d\theta_B/dt$ is equated to $d\theta_T/dt$. From these conditions the storm volume is closed vertically. Changes in $F(t)$ are due to individual changes within the volume and relative advection through the lateral boundary.

From (6) and (18), the equation for the change of total potential energy is

$$\frac{d\pi}{dt}=\int_{V_\theta}\rho J\frac{d}{dt}(e+\phi)+\nabla_\theta\cdot[(e+\phi)\rho J(\mathbf{W}-\mathbf{u})]dV_\theta. \quad (19)$$

Now combination with the first law of thermodynamics and a rearrangement yields

$$\frac{d\pi}{dt}=\int_{V_\theta}\{\rho JQ_m+J(\mathbf{u}\cdot\nabla p+\rho gw)-J\nabla\cdot(p\mathbf{u}) +\nabla_\theta\cdot[(e+\phi)\rho J(\mathbf{W}-\mathbf{u})]\}dV_\theta. \quad (20)$$

Use of the hydrostatic defect, the definition for the Montgomery streamfunction ψ, and transformation of the last two terms to surface integrals yields

$$\frac{d\pi}{dt}=\int_V\rho JQ_m+\rho J[\mathbf{u}\cdot\nabla_\theta\psi+\chi(w-\mathbf{u}\cdot\nabla_\theta h)]dV_\theta -\int_{\theta_0}^{\theta_T}\oint\rho J[RT\mathbf{u}+(e+\phi)(\mathbf{u}-\mathbf{W})]\cdot\mathbf{n}dLd\theta, \quad (21)$$

where $\mathbf{n}$ is the outer unit normal vector to the lateral boundary. In the transformation of the pressure work boundary integral, the third right-hand term of (20), the boundary conditions that $p=0$ at θ_T and $v_n=0$ at the earth's surface were used.

The change of total potential energy, for a volume defined in isentropic coordinates, (21), is due to energy production by diabatic processes, its transformation to and from kinetic energy, pressure work, and relative advection of internal and gravitational potential energy through the lateral boundaries. The pressure-work term does not include the velocity of the boundary because there is no expansion of the fluid associated with the mathematical aspect of the moving boundary.

4. The rate of change of the storm's reference state internal and gravitational potential energies

In the derivation of the rate of change, the total potential energy of the reference state is divided into the basic forms. Even though the ratio between the internal and gravitational potential energy is always unique due to the hydrostatic condition defined for the reference state, this separation will later provide insight into the vertical propagation of energy in an atmosphere continually adjusting to restore hydrostatic balance. From (7), the reference state total potential energy is

$$\pi_r=E_r+\Phi_r, \quad (22)$$

where the reference state internal and gravitational potential energies are

$$E_r=\int_{V_\theta}e_r\rho_rJ_rdV_\theta=\int_{V_\theta}e_r\rho JdV_\theta, \quad (23)$$

$$\Phi_r=\int_{V_\theta}\phi_r\rho_rJ_rdV_\theta=\int_{V_\theta}\phi_r\rho JdV_\theta. \quad (24)$$

The alternate expressions are obtained from use of (4) and deletion of the averaging operator after substitution. Now the latter forms of (23) and (24) are identical to (15). Thus, employing (18) for the direct differentiation, the rate of change of the storm's reference state internal energy is

$$\frac{dE_r}{dt}=\int_{V_\theta}\rho J\frac{de_r}{dt}+\nabla_\theta\cdot[e_r\rho J(\mathbf{W}-\mathbf{u})]dV_\theta, \quad (25)$$

where the boundary conditions are identical to those employed for (19) and by which the volume is closed vertically and includes the entire vertical extent of the atmosphere. By using Leibnitz's rule for the differentiation of the specific internal energy, $e_r(\theta, t)$, the equation of continuity, the divergence theorem and a rearrangement, (25) becomes

$$\frac{dE_r}{dt}=c_v(g/p_{00})^\kappa\int_{V_\theta}\left[\beta^\kappa\rho J\overline{\frac{d\theta}{dt}}-\frac{c_v}{c_p}\frac{\partial\phi_r}{\partial\theta}\int_\theta^{\theta_T}\overline{\nabla_\nu\cdot\rho J\mathbf{u}}d\nu\right]dV_\theta +\int_{\theta_0}^{\theta_T}\oint e_r\rho J(\mathbf{W}-\mathbf{u})\cdot\mathbf{n}dLd\theta, \quad (26)$$

where ν is the variable of integration representing θ. From an integration by parts and use of the divergence theorem, the rate of change of the storm's reference state internal energy is

$$\frac{dE_r}{dt}=c_v\int_{V_\theta}\left(\frac{p_r}{p_{00}}\right)^\kappa\rho J\overline{\frac{d\theta}{dt}}dV_\theta-\frac{c_v}{c_p}\int_{\theta_0}^{\theta_T}\oint\phi_r\rho J\mathbf{u}\cdot\mathbf{n}dLd\theta -\int_{\theta_0}^{\theta_T}\oint e_r\rho J(\mathbf{u}-\mathbf{W})\cdot\mathbf{n}dLd\theta. \quad (27)$$

Through the definitions for the reference state variables, (8) thru (10), (27) is completely specified by variables in the natural state of the atmosphere. The three terms represent the change of the reference state internal energy by heating and by relative mass flux weighted by the gravitational and internal reference state energy distributions. The second term will combine with a similar term in the equation for the gravitational potential energy change to define changes of total

potential energy due to pressure work and an exchange between the gravitational, potential and internal energies.

In like manner from (18) and (24), the rate of change of the storm's reference state gravitational potential energy is

$$\frac{d\Phi_r}{dt}=\int_{V_\theta}\rho J\frac{d\phi_r}{dt}+\nabla_\theta\cdot[\phi_r\rho J(\mathbf{W}-\mathbf{u})]dV_\theta. \quad (28)$$

By using Leibnitz's rule for the differentation of the specific gravitational potential energy, $\phi_r(\theta,t)$, and the equation of continuity, (28) becomes

$$\frac{d\Phi_r}{dt}=R(g/p_{00})^\kappa\int_{V_\theta}\overline{\rho J}\left\{\int_{\theta_0}^{\theta}\varphi\frac{\partial}{\partial\varphi}\left[\beta^{\kappa-1}\int_{\theta}^{\theta_T}\overline{\nabla_r\cdot(\rho J\mathbf{u})}d\nu\right]\right.$$
$$\left.+\beta^{\kappa-1}\overline{\rho J\frac{d\theta}{dt}}d\varphi\right\}+\nabla\cdot[\phi_r\rho J(\mathbf{W}-\mathbf{u})]dV_\theta, \quad (29)$$

where φ is a variable of integration representing θ. After an integration by parts and a rearrangement, we find that

$$\frac{d\Phi_r}{dt}=R\int_{V_\theta}\left(\frac{p_r}{p_{00}}\right)^\kappa\overline{\rho J\frac{d\theta}{dt}}dV_\theta-\int_{\theta_0}^{\theta_T}\oint\rho J$$
$$\times\left[\left(RT_r-\frac{c_v}{c_p}\phi_r\right)\mathbf{u}+\phi_r(\mathbf{u}-\mathbf{W})\right]\cdot\mathbf{n}dLd\theta. \quad (30)$$

Combining (27) and (30), the rate of change of the reference state total potential energy becomes

$$\frac{d\pi_r}{dt}=\int_{V_\theta}c_p\left(\frac{p_r}{p_{00}}\right)^\kappa\overline{\rho J\frac{d\theta}{dt}}dV_\theta-\int_{\theta_0}^{\theta_T}\oint RT_r\rho J\mathbf{u}\cdot\mathbf{n}dLd\theta$$
$$+\int_{\theta_0}^{\theta_T}\oint(e_r+\phi_r)\rho J(\mathbf{W}-\mathbf{u})\cdot\mathbf{n}dLd\theta. \quad (31)$$

A more detailed derivation of the time rate of change of the reference state energies is presented by Johnson (1969). *

In a comparison of the equations for the change of the storm's actual and reference state total potential energies, (21) and (31), corresponding terms exist in both equations except for the expression for the conversion of total potential energy to kinetic energy. Interestingly, (21) reduces to (31) in hydrostatic and flat barotropic regimes of the atmosphere.

The nonconservative processes, diabatic heating, boundary work and relative advection, that produce changes in the storm's actual total potential energy will generally produce changes in the reference state total potential energy. If the lateral boundaries are extended to include the entire atmosphere, the boundary terms vanish, and (31) corresponds to Eq. (2.75) of D & J in which diabatic heating is the sole process changing the reference state total potential energy. Because the boundary terms are usually nonzero, the result emphasizes that boundary processes must be considered in the definition of an available potential energy budget of an open system.

5. The rate of change of the storm's available potential and kinetic energies

The rate of change of the storm's kinetic energy, defined by

$$K=\int_{V_\theta}\rho Jk dV_\theta, \quad (32)$$

is

$$\frac{dK}{dt}=\int_{V_\theta}\rho J\frac{dk}{dt}dV_\theta+\int_{\theta_0}^{\theta_T}\oint\rho Jk(\mathbf{W}-\mathbf{u})\cdot\mathbf{n}dLd\theta, \quad (33)$$

where k is the specific kinetic energy, $(\mathbf{u}\cdot\mathbf{u}/2)$. From the definition of the available potential energy by (5), its time rate of change is the difference between (21) and (31), which, when combined with (33), produces the budget equations for the storm's available potential and kinetic energy:

$$\frac{dA}{dt}=G(A)-C(A,K)+B(A)+W(A), \quad (34)$$

$$\frac{dK}{dt}=C(A,K)+B(K)-D(K). \quad (35)$$

The generation is

$$G(A)=\int_{V_\theta}\left[1-\left(\frac{p_r}{p}\right)^\kappa\right]\rho JQ_m dV_\theta, \quad (36)$$

and the conversion of available potential energy to kinetic energy is

$$C(A,K)=-\int_{V_\theta}[\mathbf{u}\cdot\nabla_\theta\psi+\chi(w-\mathbf{u}\cdot\nabla_\theta h)]\rho J dV_\theta. \quad (37)$$

The three new boundary terms are the relative advection of energy into the moving volume

$$B(A)=-\int_{\theta_0}^{\theta_T}\oint[(e-e_r)+(\phi-\phi_r)]$$
$$\times\rho J(\mathbf{u}-\mathbf{W})\cdot\mathbf{n}dLd\theta, \quad (38)$$

the boundary pressure work

$$W(A)=-\int_{\theta_0}^{\theta_T}\oint\left[\frac{R}{c_v}(e-e_r)\right]\rho J\mathbf{u}\cdot\mathbf{n}dLd\theta, \quad (39)$$

*Derivation is presented in Appendix A, pp. 16 to 20.

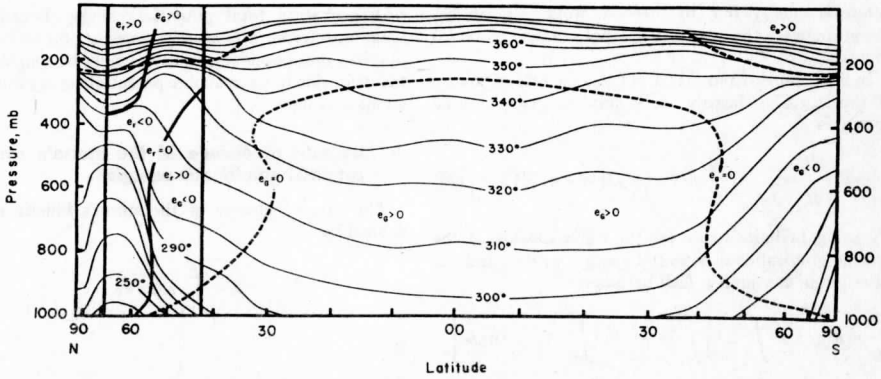


FIG. 1. A meridional distribution of potential temperature along 75W for 27 February 1958 portraying the zero isopleths of the efficiency factor for the global (e_G) and storm (e_r) concept of available potential energy. The zero isopleths for the storm (heavy lines) and the global (dashed lines) efficiency factors separate the positive and negative regions. The vertical lines delineate the boundary of the storm volume. Areas on the cross section are proportional to the global distribution of mass.

and the relative advection of kinetic energy

$$B(K)=-\int_{\theta_0}^{\theta_T}\oint k\rho J(\mathbf{u}-\mathbf{W})\cdot\mathbf{n}dLd\theta. \tag{40}$$

The total dissipation of kinetic energy is

$$D(K)=-\int_{V_\theta}\rho J\mathbf{u}\cdot\mathbf{F}dV_\theta, \tag{41}$$

where the vector **F** denotes the specific frictional force.

Eqs. (34)–(41) along with (8)–(10) define the rate of change of available potential energy and the kinetic energy from the atmospheric structure and processes occurring in the storm's volume. The expressions are valid for nonhydrostatic atmospheres but not valid for atmospheres with local static instabilities. The more general relation for an atmosphere with local static instabilities is developed in Section 2.6 of D & J. While the proof will not be repeated, this development applies and the storm's available potential energy and its rate of change may be defined even in the presence of local static instabilities.

In the generation integral G, the factor

$$[1-(p_r/p)^\kappa]=\epsilon,$$

commonly called the efficiency factor, weighs the heat addition to determine the local contribution to the generation of available potential energy. Its physical basis is founded in the condition that diabatic heating may change the total potential energy and the reference states by different amounts. For example, infrared cooling may generate available potential energy by decreasing the total potential energy of the actual atmosphere less than that of the reference atmosphere. A positive generation of available potential energy within the storm results from net heating at higher pressures and cooling at lower pressures with respect to the storm's reference pressure on an isentropic surface, a result which is consistent with Sandström's theorem (Godske *et al.*, 1957).

The generation for the storm becomes equivalent to the global expression when the lateral boundaries are extended to include the entire atmosphere. In general, however, the reference pressure of the open system is different from its global distribution. An important facet of this result is that there will be heating distributions which generate storm available potential energy, but destroy global available potential energy. In Fig. 1 an idealized meridional potential temperature distribution is portrayed. The vertical lines denote boundaries for the open system. The heavy isopleths represent the zero efficiencies, the one for the storm, the other for the entire atmosphere. For heat addition in the lower right portion of the storm volume, a region with positive storm and negative global efficiencies, the storm's available potential energy is increased but the global available potential energy is decreased. Cooling in the same region decreases the storm's available potential energy and increases the global available potential energy. A general relationship between generation for the storm and for the global scale will be derived after the relationship of the storm and the global definition of available potential energy is later established, and general conclusions should not be made from this idealized situation. However, the illustration reveals that localized heat addition within the open system in conjunction with the larger scale baroclinic structure of the circumpolar vortex of the global scale may generate and concentrate the energy available for conversion to the kinetic energy of storms.

The expression $C(A,K)$, designated to be the conversion of available potential energy to kinetic energy,

is the precise definition for the volume integral of the local production of specific kinetic energy. The physical basis for this distinction lies in the equation of motion itself, in that kinetic energy changes in frictionless regions occur when the velocity and the net absolute specific force, i.e., the sum of the pressure gradient and gravity forces, have common components. The local production of kinetic energy is not necessarily realized from the store of available potential energy within the volume, but in an open system may be associated with both boundary processes and *in situ* generation. Even within a closed system the diabatic processes and the production of kinetic energy may be linked in a manner that preserves the mass-potential temperature distribution and thus the available potential energy. In general, the implication that a positive $C(A,K)$ implies conversion of *in situ* available potential energy must be erroneous and the source of the kinetic energy may only be determined in the mean by solving integrals (34) and (35). Apart from this classic problem of the conversion between internal and kinetic energies in open systems, the integral $C(A,K)$ must be positive to offset the kinetic energy dissipation unless the storm's kinetic energy increase from the relative advection of kinetic energy into the volume is equal to or larger than the integral $D(K)$. There seems little doubt from the general circulation studies summarized by Starr (1958) and Palmén and Newton (1969) that the conversion to and dissipation of kinetic energy are maximized within the storm.

The three boundary terms (38), (39) and (40) involve the exchange of energy by relative advection or by boundary work. For energy studies of the entire atmosphere, the terms vanish and the energy equations reduce to those derived by D & J. Changes in the storm's available potential energy associated with the boundary terms $B(A)$ and $W(A)$ are from the condition that, in general, the rate of change of the actual and reference state total potential energy by boundary effects will differ. For simplicity assume the storm center is stationary ($\mathbf{W}=0$) and let us consider several hypothetical examples.

In the first example assume that the density stratification is horizontally invariant and that the region expands, i.e., $\oint \mathbf{u}\cdot\mathbf{n}dL>0$. By (21) the total potential energy decreases. However, from the relative advection of energy (38) and the boundary pressure work (39), no change of available potential energy occurs because the changes of total potential energies of the actual and reference atmosphere are identical. In general, it is clear that changes in the storm's available potential energy are not uniquely related to changes of its total potential energy.

In the second example, an expansion process by divergence is contrasted with contraction by local mass convergence. Let the expansion be due to velocity divergence just inside the boundary. After a finite time, the actual atmosphere just inside the boundary will contract vertically in its attempt to restore hydrostatic balance and depart from a flat horizontal state. Thus, available potential energy within the volume increases while the total potential energy decreases. In contrast, let the total potential energy increase by velocity convergence just inside the boundary. The actual atmosphere near the boundary expands vertically to restore hydrostatic balance and again departs from the flat reference state. Both the total and available potential energy increase. Thus, within the region, available potential energy is produced by both boundary effects, even though total potential energy changes were opposite. This hypothetical illustration is analogous to Lorenz's (1955) example of how local cooling or heating in an initally flat hydrostatic atmosphere will create energy available for conversion to kinetic energy, but produce opposite changes of total potential energy. Both examples illustrate the virtue of the theory of available potential energy in attempting to isolate the processes which lead to the production of the storm's kinetic energy within open systems. However, available potential energy theory neither negates nor, in general, replaces the theory of total potential energy. In general, both budgets of the two forms of energy must be satisfied.

Boundary pressure work $W(A)$ and the relative advection of available potential energy $B(A)$ must be important. The baroclinic mechanism of the release of existing available potential energy at the eddy scale is now well established (Starr, 1958), particularly during cyclogenesis when the temperature wave lags the pressure wave in the middle latitudes. If the storm were self-sustaining during its entire history, in that *in situ* generation exactly balanced frictional dissipation, boundary effects should be negligible. However, for the kinetic energy dissipation to be offset by the existing available potential energy of the global scale or eddies with wavelengths larger than the diameter of the region, the large-scale energy must be transferred into the region of the storm by boundary processes, since the existing available potential energy of the storm is insufficient to sustain the high frictional dissipation rate for any substantial period of the storm. Thus, the boundary terms $B(A)$ and $W(A)$ are measures of the source of the available potential energy of the open system from its environment.

Computationally, the boundary integrals $B(A)$ and $W(A)$ enjoy a significant advantage over corresponding boundary integrals $B(E+\Phi)$ and $W(E)$ in the total potential energy budget. Because the characteristic magnitude of $(e+\phi)$ is approximately 10^3 times the magnitude of k and the relative error of the boundary mass transports is approximately 10%, the error in estimating $B(E+\Phi)$ is larger than either $B(K)$ or $C(A,K)$. Thus, reliable balance of the integral relations cannot be obtained for any one storm. However, for the storm energy budget $|(e-e_r)+(\phi-\phi_r)|\approx|k|$. Thus, the relative error of $B(A)$ and $W(A)$ should be the same order as the error of $B(K)$ and $C(A,K)$ and reliable

balance between the integrals for the storm energy may possibly be obtained. However, a detailed analysis will show that the reliability of the statistics will depend on the thermodynamic and momentum structure for each storm due to the nonlinearity in the energy equations.

The other boundary term $B(K)$ is identified with the transfer of existing kinetic energy into or from the volume. The importance of this transfer in cyclogenesis is indicated in theoretical and diagnostic results. Eliasen (1960) studied the combined baroclinic and barotropic instability of frontal waves and concluded that in one particular case two-thirds of the wave energy was supplied by the potential energy and one-third by the kinetic energy of the mean state. Recently in a diagnostic study Sechrist (1968) has verified the importance of the advection of existing kinetic energy of the mean flow into the region of cyclogenesis. Simultaneously, the kinetic energy is converted to available potential energy within the region of the storm by an indirect circulation in the rear of the developing wave which increases the baroclinity of the upper troposphere front.

Sechrist's results within the framework of energetics agree with the earlier work of Newton (1954), Sanders (1955), Reed (1955) and Danielsen (1966). With respect to frontogenesis, Reed concluded that the increase in horizontal temperature gradient was due to the steepening of the slope of the isentropic surface as a result of the indirect solenoidal circulation.

Besides its obvious importance in frontogenesis and cyclogenesis, the condition of the steepening of the isentropic surfaces and the intensification of the horizontal temperature gradient serves two purposes within the energetics framework. The first is to increase the amount of kinetic energy which may be stored in quasi-geostrophic balance within the confines of the storm volume. Through the thermal wind relation the geostrophic wind is

$$\mathbf{u}_g=\mathbf{u}_{g0}+\frac{c_p}{fp_{00}{}^\kappa}\int_{\theta_0}^{\theta}\mathbf{k}\times\nabla_\theta p^\kappa d\theta, \tag{42}$$

and the geostrophic kinetic energy of the storm volume is

$$K_g=\frac{1}{2}\int_{V_\theta}\rho J_\theta\left\{u_{g0}{}^2+\frac{2c_p}{fp_{00}{}^\kappa}\mathbf{u}_{g0}\cdot\left[\mathbf{k}\times\nabla_\theta\left(\int_{\theta_0}^{\theta}p^\kappa d\nu\right)\right]+\left[\frac{c_p}{fp_{00}{}^\kappa}\left|\nabla_\theta\left(\int_{\theta_0}^{\theta}p^\kappa d\nu\right)\right|\right]^2\right\}dV_\theta, \tag{43}$$

where $\mathbf{u}_{g0}$ is the surface geostrophic velocity. The component of geostrophic kinetic energy of the storm volume coupled with the sloping isentropes will be primarily represented by the last term of (43), and steeping of the isentropic surfaces by frontogenesis would make this term more positive. In addition, an observed characteristic of the aging of storms is for the axis of the baroclinic zone of the entire storm to be tilted from the horizontal into the vertical. Thus, through changes of its large-scale baroclinic distribution, the storm has the potential of bringing into quasi-geostrophic balance large amounts of kinetic energy created within the system from either adiabatic, diabatic or boundary processes.

The second effect of the indirect circulation in the developing stage is to create a more favorable efficiency factor distribution for the generation of storm available potential energy by increasing the magnitudes of both the positive and negative efficiencies within the storm volume. In a flat barotropic region there is no generation for either heating or cooling. In contrast, the more baroclinic a region becomes, the larger the magnitude of the generation of storm available potential energy for a given distribution of heating. The frontogenesis caused by the indirect circulation of the incoming jet implies that the generation by a diabatic process may become a relative extremum, being positive (negative) for heating (cooling) at high pressure and cooling (heating) at low pressure. Thus, for the quasi-geostrophic regions one may conclude that the storm generation of available potential energy is coupled strongly with the presence of the meandering baroclinic zone of the polar front and jet streams.

Sechrist (1968) also found that the trajectories of the low troposphere subtropical air ahead of the wave displayed large accelerations into and upward in the developing vortex—a kinetic energy generating process from the conversion of available potential energy. The likelihood that a portion of the kinetic energy produced within the ascending subtropical trajectories is from an *in situ* diabatic energy source may be inferred, noting that if the heating distribution and efficiency distribution are in phase the relative generation extremum will be positive. Bullock and Johnson's (1970) results, although for another storm, confirm the importance of the *in situ* generation by the latent heat component for the mature and early occluded stages. One may speculate that as the system occludes and as the latent heat released in the cold vortex exceeds that in the warm sector, the generation should decrease and possibly become negative.

The implication of the various results are that baroclinic and barotropic development as well as diabatic processes are all important because of their simultaneous occurrence in the life cycle of the storm. A diagnostic study using the framework developed in this paper holds the promise of studying these components of the storm's energy exchange and providing insight into the processes of fronto- and cyclogenesis and fronto- and cyclolysis.

6. Diagnostic results

In D & J the emphasis on the importance of the various diabatic components in eddy generation, although speculative, was supported primarily from physical considerations of the organization of the

TABLE 1. Total and component generation estimates of the available potential energy of storms (W m^{-2}). Bracketed values are estimates using the isobaric approximations.

	Latent heat	Sensible heat	Infrared	Shortwave	Total	Boundary frictional dissipation
Hurricane Hilda[1]						
Anthes and Johnson (1968)						
Mature, 1200 GMT, 1/10/64	2.5		0.6	0.2	3.3	—
Extratropical cyclone[2]						
Bullock and Johnson (1970)						
Incipient, 1200 GMT, 25/3/64	0.9	—	0.9 (0.0)	—	1.8	3.1
Mature, 0000 GMT, 26/3/64	8.0	—	1.2 (0.0)	—	9.2	5.1
Occluded, 1200 GMT, 26/3/64	6.0	—	1.1 (0.0)	—	7.1	9.1
Extratropical cyclone[2]						
Gall and Johnson (1969)						
Mature, 1200 GMT, 14/2/68						
Upper estimate		6.1 (−7.6)				7.1
Lower estimate		3.1				

[1] Radius of storm area 1000 km.
[2] Radius of storm area 1700 km.

diabatic heating at the secondary scale. In their summary of eddy generation, they suggest that under suitable configurations all of the diabatic components can serve to generate available potential energy in a disturbance. At other stages the diabatic heating fields may destroy available potential energy.

The example of the storm generation and global destruction of available potential energy serves to illustrate the danger of attempting to infer the importance of diabatic processes in the thermodynamics of the storm scale if the global concept is utilized. However, the presence of the baroclinicity of the circumpolar vortex in the middle latitudes means that a favorable efficiency factor distribution within the storm volume is frequently found and if a localized heat source is favorably located, a source of storm available potential energy appears.

In the diagnostic work at the University of Wisconsin our objective has been to test the hypothesis that the generation at the secondary scale embedded within the long-wave regime of the circumpolar vortex is a significant source of energy for the development and maintenance of the storm's kinetic energy. Final conclusions concerning the importance of diabatic processes in the thermodynamics and mechanics of the secondary scale must be based on a complete solution of the mass, momentum and energy relations. However, the results to date, which are now summarized, provide support for the hypothesis and more extensive diagnostic studies are in progress.

For our first diagnostic study of storm generation, the hurricane was selected. Anthes and Johnson's (1968) generation results for the four primary diabatic components are presented in Table 1. Of the average generation of 3.3 W m^{-2} for a 1000 km radial volume, 77% was generated by latent heating, 17% by IR cooling and 6% by direct solar absorption. Although the direct generation by sensible heating at the earth's interface was a negligible positive quantity, this diabatic process plays a very important role in the maintainance of the deep cumulus convection through which the latent heat is released. The total generation of 10.3×10^{12} W for the storm compares favorably with Hawkins and Rubsam's (1968) estimate of 8.6×10^{12} W for the kinetic energy production, a result supporting the concept of the local storm generation.

In Bullock and Johnson's (1970) study of a moderately intense extratropical cyclone only the generation by latent heat and IR emission were estimated for three different stages. In addition, the frictional dissipation of the planetary boundary layer for a 1700-km radial volume was estimated for comparison. Although all generation estimates are positive in Table 1, they vary significantly during the cyclogenetic, mature and early occluded stages. In the cyclogenetic stage the average boundary layer frictional dissipation of 3.1 W m^{-2} exceeds the combined generation of 1.8 W m^{-2} that is equally divided between the IR and latent heat components. However, within the succeeding 12-hr period of development the latent heat component increases to 8.0 W m^{-2}, the infrared becomes 1.2 W m^{-2} and their combined generation is 9.2 W m^{-2}. The structure of the storm's efficiency factors indicates that the generation by solar absorption and sensible heating at the earth's interface is also positive. However, magnitude considerations indicate that the generation by these two components is probably less than that by the IR component. In the occluded stage the latent and IR generation are 6.0 and 1.1 W m^{-2}, respectively, and the total generation of 7.1 W m^{-2} is now less than the boundary dissipation of 9.1 W m^{-2}.

The increase of the generation from 1.8 W m^{-2} for the incipient to 9.2 W m^{-2} for the mature stage and the excess of the generation over the frictional dissipation of 5.1 W m^{-2} in the mature stage strongly suggest that diabatic processes are important in cyclone development. The *in situ* generation is an additional source of energy to that realized through wave instability. In the

occluded stage the sink of kinetic energy by frictional dissipation exceeds the storm generation, and the cyclone will die unless boundary processes provide a source of available potential or kinetic energy for the storm.

In a study of the evolution of the efficiency factor distribution for a cyclone situated over the United States between 10 and 12 March 1962, Hahn and Horn's (1969) results indicate that the relative importance of the four generation components varies with the age of the storm. In the cyclogenetic stage, the magnitude of the positive and negative efficiency factors of the low and high troposphere, respectively, are at their maximum. Thus, in the incipient stage sensible heat addition near the surface and cooling aloft, a static stability decreasing process, are more effective in generating the storm's available potential energy than later in the storm development. As the storm matures, the magnitudes of the low and high tropospheric efficiencies are reduced somewhat. The isopleths of the efficiencies become more vertically orientated due to the vertical development of the vortex with its baroclinic support and generation by horizontal differential heating becomes more important. Such an efficiency structure enhances the generation by the latent heat release in the cyclone's warm sector. Later in the cyclone's life, the decrease of baroclinity by the occlusion process and the establishment of a cold core vortex leads to a reduction of the magnitude of the efficiency factors for the storm and a general disorganization of their distribution. Thus, in the occlusion process the effect that heating or cooling becomes less effective in the generation, and the condition that the advection of latent energy into the storm volume is cut off, mean that the *in situ* source of storm energy by diabatic processes decreases after the storm matures. The generation may even become negative if the primary region of latent heat release shifts from the anticyclonic side of the upper tropospheric jet core, a region of positive efficiencies, to the cold core vortex center of the cyclone, a region of negative efficiencies.

In D & J's review of the theory of available potential energy, sensible heating at the earth's surface was identified as a likely mechanism of the generation of available potential energy in the cyclogenetic regions over the oceans just east of the Asian and North American continents. Synoptic experience indicates that explosive cyclogenesis in these situations is associated with optimum distributions of baroclinicity and vorticity and with strong sensible heating of the cold polar air flowing offshore from the continents over the warm ocean surface. Gall and Johnson (1969) confirmed that sensible heating of the cold polar air in the mature stage of a cyclone in the North Atlantic generates available potential energy. Due to our lack of knowledge concerning the vertical redistribution of the sensible heat flux, a lower bound estimate of 3.1 W m^{-2} was made using a uniform heating distribution within the adiabatic layer of several hundred millibars vertical extent and an upper bound estimate of 6.14 W m^{-2} was made by adding the sensible heat to the surface layer. While the average frictional dissipation of 7.09 W m^{-2} exceeds the sensible heat generation, at this stage it is likely that the total generation, which would include the latent heat component, exceeds the boundary layer dissipation.

Additional indirect evidence of the importance of diabatic effects is the recent success in predicting the birth of second and third generation cyclones by the GFDL general circulation model (Miyakoda *et al.*, 1967). This model includes the processes of radiative transfer, latent heat release, and turbulent exchange of sensible and latent heat with the lower boundary. The accurate prediction of East Coast cyclogenesis reported by Miyakoda *et al.* must be associated with the improved modeling of the diabatic processes which generate available potential energy within the storm scale.

Included in Table 1 are several estimates of the IR and sensible heat generation (in brackets) using Lorenz's (1955) approximate expression based on the isobaric covariance of temperature and heating. From his approximate form he suggested that the eddy generation is negative in view of the probable warming of cold air masses and cooling of warm air masses in middle latitudes. Recently, by inference from the approximate expression, Palmén and Newton (1969) in their discussion of the role of latent and sensible heating within the extratropical cyclone argue that in the East Coast cyclogenetic regions the latent heat component creates and sensible heat destroys eddy available potential energy. In Table 1, the approximate estimates for the sensible heat component are consistent with the inference that the heating of cold air masses destroys available potential energy. However, the exact generation estimates show that both sensible heat and infrared emission generate available potential energy. The contrast in the results, exact and approximate, illustrates the danger of using expressions for either inference or diagnostic computations in situations where the assumptions utilized in deriving the approximations are violated. Approximate generation estimates for the components of sensible heating and infrared emission, as emphasized by D & J and also Johnson (1967), will be erroneous. Such results have lead many investigators to erroneous conclusions concerning the importance of the eddy generation within the storm scale.

All of the results from the exact expression support the hypothesis that the *in situ* generation at the storm scale is sufficient 1) to offset a significant portion of the storm's kinetic energy dissipation, 2) to imply that during certain stages the storm may be energetically self-sufficient, and 3) to suggest that diabatic heating is an important factor in determining the storm's transient behavior during its growth, maturation and decay.

Although the comparative significance of diabatic, adiabatic and boundary energy sources for the extra-

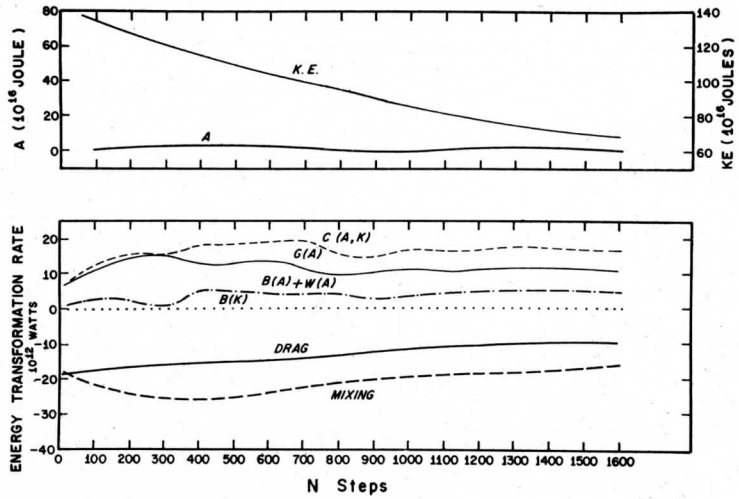


FIG. 2. The time distribution of the kinetic and available potential energy budgets with the components of their rate of change from an isentropic model of the hurricane (from Anthes, 1970). The conversion of available potential energy is $C(A,K)$, the generation is $G(A)$, the boundary advection of available potential and kinetic energy are $B(A)$ plus $W(A)$ and $B(K)$, and the dissipation is $D(K)$. The dissipation of kinetic energy is divided into two components, boundary layer drag and internal mixing.

tropical cyclone's budget has not been ascertained by this diagnostic model, some insight is provided by the results from Anthes' (1970) numerical model of the hurricane in isentropic coordinates. Fig. 2 portrays the time variation of the generation, the conversion of available potential to kinetic energy, the boundary fluxes of both the available potential and kinetic energies, and the frictional dissipation due to drag and internal mixing as the mass, momentum and energy come into steady state for the given thermal forcing. For the steady state, both $G(A)$ and $B(A)$ plus $W(A)$ are sources of available potential energy to offset the frictional sink and advection of kinetic energy from the storm volume of 500-km radial extent. The conversion of available potential to kinetic energy is equal to both the sources and sinks. When the radius of the storm volume was increased to 1000 km, and the source and sink of internal energy within the volume were nearly equal, the steady-state profiles of the energy terms showed that the generation of available potential energy within the volume was sufficient to balance the sinks of energy by frictional dissipation and outward advection of kinetic energy. Although mass was allowed to flow through the lateral boundaries, the boundary advection of available potential energy $B(A)$ decreased to near zero. Thus, the thermodynamic source of energy within the large volume was sufficient to sustain the hurricane, a result consistent with Anthes and Johnson's (1968) conclusions for hurricane Hilda. The contrast in the estimates of $B(A)$ plus $W(A)$ and $G(A)$ for the two volumes suggests that for a volume of sufficient size the inward flux of energy available for conversion to kinetic energy may vanish. A disadvantage of this condition is that the generation and the boundary integrals become functions of the size of the volume. However, this apparent problem is actually an advantage because by varying the radial extent of the volume, the interaction of the storm with its environment as a function of scale may be studied. Thus, the thermodynamic and mechanical spheres of influence of the environment on the storm may be determined.

7. The relationship of the storm's available potential energy to global available potential energy

In the energy equations for the storm, the three boundary terms represent the interaction between the open system and the atmospheric region exterior to the storm's volume. From the principles of energy conservation, it is important to establish the relation of the storm's available potential energy with energy of the exterior region and to show that the storm's available potential energy is an integral part.

The global available potential energy given by the difference of the total potential energy of the actual and global reference states is

$$A_G=\int_{V_\theta}[(e-e_G)+(\phi-\phi_G)]\rho J dV_\theta, \qquad (44)$$

where the subscript G denotes properties or quantities of the global reference state. Now V_θ is the entire atmospheric volume. The global available potential energy is divided into contributions from n different regions, i.e.,

$$A_G=\sum_{i=1}^{n} A_i, \tag{45}$$

where

$$A_i=\int_{V_{\theta i}} [(e-e_G)+(\phi-\phi_G)]\rho J dV_\theta. \tag{46}$$

Of the n regions, $(n-1)$ contain storms while the nth region is the portion of the atmosphere exterior to the $(n-1)$ storms.

Within the ith region, the contribution to global available potential energy can now be partitioned into a baroclinic and a barotropic component of available potential energy. The baroclinic contribution of the ith region (or the storm's available potential energy) is

$$A_{S_i}=\int_{V_{\theta i}} [(e-e_{r_i})+(\phi-\phi_{r_i})]\rho J dV_\theta, \tag{47}$$

and the barotropic contribution

$$A_{B_i}=\int_{V_{\theta i}} [(e_{r_i}-e_G)+(\phi_{r_i}-\phi_G)]\rho J dV_\theta, \tag{48}$$

both of which transform [after (12)] to

$$A_{S_i}=\frac{c_p}{p_{00}{}^\kappa}\int_{V_{\theta i}} \theta(p^\kappa-p_{r_i}{}^\kappa)\rho J dV_\theta+\int_{V_{\theta i}} \frac{\chi}{g}\phi\rho J dV_\theta, \tag{49}$$

$$A_{B_i}=\frac{c_p}{p_{00}{}^\kappa}\int_{V_{\theta i}} \theta(p_{r_i}{}^\kappa-p_G{}^\kappa)\rho J dV_\theta. \tag{50}$$

From these definitions for the partitioning of the global available potential energy into barotropic and baroclinic components from each of the regions, the time rate of change of the contributions are defined using the prior results.

The rate of change of the ith baroclinic contribution, i.e., the available potential energy of the storm, is

$$\frac{d}{dt}A_{S_i}=G(A_{S_i})-C(A_{S_i},K)+B(A_{S_i})+W(A_{S_i}), \tag{51}$$

where the right-hand terms are, respectively, identical to the right-hand terms of (34) and defined in (36)–(39). The rate of change of the ith barotropic contribution is

$$\frac{d}{dt}A_{B_i}=G(A_{B_i})+B(A_{B_i})+W(A_{B_i}), \tag{52}$$

where the barotropic generation contribution of the ith region can be expressed as

$$G(A_{B_i})=\int_{V_{\theta i}} [(p_{r_i}{}^\kappa-p_G{}^\kappa)/p^\kappa]\rho J Q_m dV_\theta. \tag{53}$$

The barotropic interaction associated with the mass advection of reference energies for the ith region is

$$B(A_{B_i})=-\int_{\theta_0}^{\theta_T}\oint (e_{r_i}+\phi_{r_i})\rho J(\mathbf{u}-\mathbf{W})\cdot\mathbf{n}dLd\theta, \tag{54}$$

and that of the barotropic interaction associated with pressure work for the ith region as

$$W(A_{B_i})=-\frac{R}{c_v}\int_{\theta_0}^{\theta_T}\oint e_r\rho J\mathbf{u}\cdot\mathbf{n}dLd\theta. \tag{55}$$

The interpretation of these two terms is difficult if one insists from the concept of available potential energy that the atmosphere tends to the horizontally stratified state within each storm region and then eventually to a flat global state. Zero-order discontinuities would be introduced. Such a condition is never realized by the nature of the continuum of the fluid. As D & J state, the concept that by virtue of the existence of available potential energy the atmosphere tends to the flat reference state is misleading. The atmosphere first of all tends to quasi-hydrostatic balance and in extratropical latitudes also tends to the geostrophic state. The existence of these two states of balance requires baroclinicity to provide the support through the thermal wind to maintain a geostrophic component of the momentum. Thus, both available potential and kinetic energy must be present at all times. The flat reference state of each region as well as the global state is a hypothetical state and in all probability is not a dynamically realizable state for a rotating atmosphere responding internally to thermal forcing, friction and boundary processes.

The form of (48) shows that the energy of the barotropic state is given by a weighted mass distribution in the volume where the weighting factor in the square brackets of the integrand is a function of potential temperature. The density distribution remains continuous within each region and across boundaries, while the weight function changes only at each boundary and remains constant within each region. Thus, for isentropic flow, changes in the barotropic component of available energy in each region are determined by boundary processes. An extremely interesting facet of this concept is the condition that the barotropic component is unavailable for conversion to kinetic energy for any scale within each open region. In contrast, the concept of the global available potential energy states that both the sum of the baroclinic and barotropic con-

tribution from all regions are available for conversion to kinetic energy. Mathematically, the fact that from the global concept more energy is available for conversion to kinetic energy than from the sum of the baroclinic components is due to the condition that additional constraints are introduced in the mass potential temperature distribution by the definition of local states. Thus, in the absence of boundary processes there is less storm than global energy available for conversion to kinetic energy unless p_{r_i} of all the regions is equal to p_G.

In the physical sense, the barotropic component of each region becomes available only by an interaction with an adjacent region through boundary processes in association with the scale-wise exchange of energy, e.g., the potential energy exchange between the storm and scales larger than the storm volume, but within which the storm is embedded. In the situation where the areal extent of the important energy transformations and the kinetic energy production is by scales of motion larger than the area of an open system, the mass redistribution associated with the large-scale release of available potential energy will be through changes in the barotropic component of adjacent regions and boundary fluxes. In this sense the barotropic component due to the contrast between regions is the available potential energy of scales larger than the open region, and the baroclinic component due to contrast within regions is the available potential energy of all scales equal to or smaller than the region. As the energy leaves one region either by advection by or pressure work, the factor which determines the energy available for conversion to kinetic energy and the unavailable portion is the weighting of the mass advection in the integrands of Eqs. (38), (39), (54) and (55). Only when the reference states of adjacent regions are equal will the ratio of the baroclinic to the barotropic portion be preserved through interaction at the common boundary. Vertical cross sections of the weighting factors in the square brackets of (38) and (39) along the boundary of the storm should reveal the portion of the total potential energy that becomes available for conversion to kinetic energy through boundary processes and reveal the influence of larger scales. In the same manner cross sections of the efficiency factors reveal the portion of the total heating that generates available potential energy.

8. Vertical energy exchange through hydrostatic adjustment

Recently, Wiin-Nielsen (1968) has emphasized the difference in principle between the energetics of non-hydrostatic and hydrostatic atmospheres. Because the hypothetical reference state is hydrostatic by definition, the equations for the change of the reference state internal and gravitational energies elucidate why the available potential energy of a region of the atmosphere with limited vertical extent cannot be defined and how changes in internal energy and gravitational potential energy propagate vertically in a quasi-hydrostatic medium as hydrostatic balance is maintained.

The equations were developed for any arbitrary vertically walled region. Thus, let us consider a hydrostatic and horizontally stratified region of the atmosphere defined in a hypothetical moving volume with incremental horizontal area δA and assume that the motion is isentropic with boundary velocity $\mathbf{W}(x, y, \theta, t_0) = \mathbf{U}(x, y, \theta, t_0) \neq 0$. The condition from $\mathbf{W}=\mathbf{W}(\theta)$ that the lateral walls at t_0 will not remain vertical is unimportant in considering the underlying process of the internal-gravitational potential energy adjustment associated with boundary work. Under the simplified conditions the change of internal energy from (27) is representative for a thin vertical column of the real atmosphere and reduces to

$$\frac{dE}{dt} = -\frac{c_v}{c_p}\int_{\theta_0}^{\theta_T} \phi \oint_{\delta A} \rho J \mathbf{u}\cdot\mathbf{n} dL d\theta$$

$$= -\frac{c_v}{c_p}\int_{\theta_0}^{\theta_T} \phi \oint_{\delta A} \nabla_\theta\cdot(\rho J\mathbf{u}) dA d\theta. \qquad (56)$$

Within the column the internal energy change associated with lateral pressure work is equal to the product of the gravitational potential energy and the divergence of the mass transport or alternatively momentum. Momentum divergence near the earth's boundary does not contribute significantly to internal energy changes through pressure work because ϕ is near zero and $p(\theta)$ on all surfaces above the region of divergence is preserved in the hydrostatic adjustment. In contrast, for a region high in the atmosphere where ϕ is large, momentum divergence reduces $p(\theta)$ on all underlying isentropic surfaces and causes a decrease in the internal energy in the column below the level of divergence. As $T(\theta)$ decreases, the hypothetical atmosphere contracts to maintain hydrostatic balance and a change in its gravitational potential energy occurs. The rate of change of gravitational potential energy from (30) is

$$\frac{d\phi}{dt} = -\int_{\theta_0}^{\theta_T}\left[\left(\frac{R}{c_v}e - \frac{c_v}{c_p}\phi\right)\oint \rho J\mathbf{u}\cdot\mathbf{n} dL\right]d\theta,$$

$$= -\int_{\theta_0}^{\theta_T}\left[\left(\frac{R}{c_v}e - \frac{c_v}{c_p}\phi\right)\int_{\delta A} \nabla_\theta\cdot(\rho J\mathbf{u}) dA\right]d\theta. \qquad (57)$$

The last term of (57) is equal and opposite to the internal energy change in (56). Their sum yields

$$\frac{d\pi}{dt} = -\int_{\theta_0}^{\theta_T}\left[\frac{R}{c_v}e\int_{\delta A} \nabla_\theta\cdot(\rho J\mathbf{u}) dA\right]d\theta, \qquad (58)$$

which is the result for total potential energy change by pressure work in a barotropic region. Within the basic energy equations, the pressure work associated with expansion or contraction only produces an internal energy change. However, in an atmosphere which adjusts to maintain hydrostatic balance, changes must occur for both internal and gravitational potential energy to maintain this characteristic ratio and a readjustment occurs through the entire column. It is only because the time scale of the hydrostatic adjustment is minute compared to the time scale of the large organized scales that the energetics of the quasi-hydrostatic atmosphere are effectively equivalent to the energetics of a hypothetical hydrostatic atmosphere. The atmosphere is never in exact hydrostatic balance and an assumption of balance does not alter the mode of energy exchange for the real atmosphere. However, these equations illustrate the changes between internal and gravitational potential energy in a quasi-hydrostatic and compressible atmosphere that are forced by momentum divergence through pressure work.

In view of this mode of energy propagation in the vertical direction, not only is it impossible to define available potential energy for regions of the atmosphere with limited vertical extent (Van Mieghem, 1967), but it is also impossible to determine its rate of change. For example, significant changes of geopotential energy in the stratosphere will be forced by momentum convergence in the lower troposphere. The tropospheric momentum convergence (or divergence) changes the gravitational potential energy and the reference state gravitational potential energies of the stratosphere by differing amounts. Thus, in the physical sense, changes of the stratospheric component of available potential energy will be coupled and likely strongly forced by tropospheric events, a conclusion which agrees with the diagnostic results of Oort (1964). Whether the component of the available potential of the upper stratospheric and mesospheric regions is strongly coupled with tropospheric changes depends on whether or not the vertical energy exchange associated with a form of Dines compensation is restricted to the region below 25 km. It is certain, in the mathematical sense, that it is impossible to formulate a kinetic-available potential energy conservation principle for frictionless isentropic motion for any region of the atmosphere limited in its vertical extent. Thus, the presence or lack of coupling must depend on the nature of the momentum distribution within the atmosphere, a problem which Charney and Drazin (1961) considered for the quasi-geostrophic component of the planetary disturbances. Since Eqs. (32)–(34) show that the energy exchange occurs primarily through the irrotational component of the momentum distribution, their conclusions concerning the lack of the escape of tropospheric energy into the upper atmosphere rest primarily on the conditions of quasi-horizontal orientation for isentropic surfaces, horizonally invariant static stability, hydrostatic balance, and geostrophy of the medium (Charney, 1948)

9. Conclusions

The question whether or not *in situ* diabatic processes are important in the thermodynamics of the storm has remained unresolved since Espy (1841) emphasized the importance of latent heat release in the cyclone. With the introduction of baroclinic instability theory, the pendulum of thought has been swayed to the viewpoint that the storms of the extratropics are manifestations of the adiabatic release of available potential energy generated by diabatic processes at the zonal scale. However, the statistical evidence of Petterssen (1950) showing that cyclogenetic areas are usually geographical regions where either sources of sensible or latent heat exist has not been satisfactorily explained. Within the vorticity framework, Petterssen (1956) has emphasized the effects of heat sources and sinks intensifying either cyclonic or anticyclonic circulations. Within the energy framework, Palmén (1958), Danard (1964, 1966) and Palmén and Newton (1969) among others have emphasized the internal production of kinetic energy due primarily to the latent heat release in cyclones. The diagnostic results for the generation of available potential energy summarized in this study support the argument that the *in situ* source of energy by diabatic processes is important in explaining the maintenance of the storm's kinetic energy. Still a complete theory of the birth, the maturation and the decay of a storm must involve the simultaneous solution of the mass, momentum and energy relations.

Possibly the two important facets of this study are the development of the theory of available potential energy for open systems and from it the concept of a storm efficiency structure connected with the meandering baroclinic zone embedded in the long waves. In general, the storm generation is positive, by heating at high pressure and cooling at low pressure, only in the presence of a favorable baroclinic distribution. A favorable distribution of baroclinity and of heating and cooling will tend to increase the number of solenoids in the plane transverse to the major axis of the core of momentum. Thus, for any initially balanced distribution of mass and momentum on a rotating sphere, through the solenoidal form of Bjerknes' circulation theorem (Godske *et al.*, 1957), we see that the baroclinity and pattern of heating and cooling of the secondary circulation that generates available potential energy will also force a transverse circulation that increases the kinetic energy of the basic current. Diagnostic evidence had already established that strong small-scale vorticity advection associated with a polar jet and its thermal support is important in cyclogenesis. The theory of storm available potential energy points out that an *in*

situ source of available potential energy by diabatic processes can only be realized in the presence of the local baroclinity, a local condition created by the propagating polar jet core in its adjustment to quasi-geostrophic balance. Preliminary diagnostic results show that the generation is positive and significant at the secondary scale. When these results within the framework of energetics are coupled with the effects of thermodynamic processes in the mass and momentum relations, a theory for the extratropical cyclone that emphasizes the joint importance of adiabatic and diabatic processes will likely be established.

Acknowledgments. I express my appreciation to Prof. Frank Sechrist for many stimulating conversations on the energetics of extratropical cyclones, to Prof. Lyle Horn for continued cooperation over the several years of research in atmospheric energetics, to Dr. Richard Anthes for permission to present his results shown in Fig. 2, and to Dr. Ben Bullock and Mr. Robert Gall at the University of Wisconsin who contributed results on the generation of available potential energy. I gratefully acknowledge the visiting appointment for the academic year 1968–69 in the Department of Meteorology of Pennsylvania State University where I was provided the opportunity for reflection and writing. The research was jointly supported by the National Environmental Satellite Center, Environmental Science Services Administration, under ESSA Grants WBG-52 and E-8-69(G), and the Atmospheric Sciences Section, National Science Foundation, under NSF Grant GA-10997 to the University of Wisconsin and NSF Grant GA-1595 to The Pennsylvania State University.

REFERENCES

Anthes, R. A., 1970: A diagnostic model of the tropical cyclone in isentropic coordinates. Ph.D. thesis, University of Wisconsin, Madison.

——, and D. R. Johnson, 1968: Generation of available potential energy in Hurricane Hilda (1964). *Mon. Wea. Rev.*, **96**, 291–302.

Bullock, B. R., and D. R. Johnson, 1970: The generation of available potential energy by latent heat release in a mid-latitude cyclone. *Mon. Wea. Rev.* (in press).

Charney, J. G., 1948: On the scale of atmospheric motions. *Geofys. Publiskasjoner*, **17**, No. 2, 17 pp.

——, and P. G. Drazin, 1961: Propagation of planetary scale disturbances from the lower into the upper atmosphere. *J. Geophys. Res.*, **66**, 83–110.

Danard, M. B., 1964: On the influence of released latent heat on cyclone development. *J. Appl. Meteor.*, **3**, 27–37.

——, 1966: On the contribution of released latent heat to changes in available potential energy. *J. Appl. Meteor.*, **5**, 81–84.

Danielsen, E., 1966: Research in four dimensional diagnostics of cyclonic storm cloud systems. Sci. Rept. No. 1, Contract No. AF 19(628)-4762, The Pennsylvania State University, 53 pp.

Dutton, J. A., and D R. Johnson, 1967: The theory of available potential energy and a variational approach to atmospheric energetics. *Advances in Geophysics*, Vol. 12, New York, Academic Press, 333–436.

Eliasen, E., 1960: On the initial development of frontal waves. *Publ. Danski Meteor. Inst., Meddelser*, No. 13.

Espy, J. P., 1841: *Philosophy of Storms*. Boston, Mass., 552 pp.

Gall, R. L., and D. R. Johnson, 1969: The generation of available potential energy by sensible heating: A case study. Ann. Rept., ESSA Grant E-8-69(G), Dept. of Meteorology, University of Wisconsin, Madison.

Godske, C. L., J. Bergeron, J. Bjerknes and R. C. Bundgaard, 1957: *Dynamic Meteorology and Weather Forecasting*. Amer. Meteor. Soc., Boston, and Carnegie Institution of Washington, Washington, D. C.

Hahn, D., and L. H. Horn, 1969: The generation of available potential energy in a mid-latitude cyclone. Final Rept. ESSA Grant WBG 52, Dept. of Meteorology, University of Wisconsin, Madison.

Hawkins, H. F., and D. T. Rubsam, 1968: Hurricane Hilda, 1964; II. Structure and budgets of the hurricane on October 1, 1964. *Mon. Wea. Rev.*, **96**, 617–636.

Johnson, Donald R., 1967: The role of terrestrial radiation in the generation of zonal and eddy available potential energy. *Tellus*, **19**, 517–539.

——, 1969: The available potential energy of storms. Ann. Rept., ESSA Grant E-8-69(G), Dept. of Meteorology, University of Wisconsin, Madison.

Lorenz, E. N., 1955: Available potential energy and the maintenance of the general circulation. *Tellus*, **7**, 157–167.

Margules, M., 1903: Uber die Energie der Stürme. *Jahrb. Zentralanst. Meteor.*, 1–26. (Transl. by C. Abbe, 1910: *The Mechanics of the Earth's Atmosphere*. 3rd Coll., Smithsonian Inst., Washington, D. C., 553–595.)

Miyakoda, K., J. Smagorinsky, R. F. Strickler and G. D. Hembree, 1967: Experimental extended prediction with a nine level hemispheric model. *Mon. Wea. Rev.*, **97**, 1–76.

Newton, C., 1954: Frontogenesis and frontolysis as a three dimensional process. *J. Meteor.*, **11**, 449–461

Oort, A. H., 1964: On the energetics of the mean and eddy circulations in the lower stratosphere. *Tellus*, **16**, 309–327.

Palmén, E., 1958: Vertical circulation and release of kinetic energy during the development of hurricane Hazel into an extratropical storm. *Tellus*, **10**, 1–23.

——, and C. W. Newton, 1969: *Atmospheric Circulation Systems*. New York, Academic Press, 603 pp.

Petterssen, S., 1950: Some aspects of the general circulation of the atmosphere. *Cent. Proc., Roy. Meteor. Soc.*, 120–155.

——, 1956: *Weather Analysis and Forecasting*, Vol. 1. New York, McGraw-Hill, 428 pp.

Reed, R., 1955: A study of a characteristic type of upper level frontogenesis. *J. Meteor.*, **12**, 226–237.

Sanders, F., 1955: An investigation of the structure and dynamics of an intense surface frontal zone. *J. Meteor.*, **12**, 542–552.

Sechrist, F. S., 1968: Energy conversions in a developing cyclone. Ph.D. thesis, Dept. of Meteorology, Pennsylvania State University.

Smith, P. J., 1969: On the contribution of a limited region to the global budget. *Tellus*, **21**, 202–207.

——, and L. H. Horn, 1969: A computational study of the energetics of a limited region of the atmosphere. *Tellus*, **21**, 193–201.

Starr, V. P., 1958: What constitutes our new outlook on the general circulation? *J. Meteor. Soc. Japan, Ser. II*, **36**, 167–173.

Van Mieghem, J., 1967: Energy transport across internal boundaries. *Beit. Phys. Atmos.*, **40**, 1–6.

Wiin-Nielsen, A., 1968: On the intensity of the general circulation of the atmosphere. *Rev. Geophys.*, **6**, 559–579.

APPENDIX A: A derivation of the time rate of change of reference state energies

In this appendix an extended derivation of the energies of the reference state is presented for the interested researcher.

The rate of change of the specific internal energy of the reference state from its definition and (9) is

$$\frac{de_r}{dt} = c_v\left(\frac{g}{p_{00}}\right)^{\kappa}\left[\beta^{\kappa}\frac{d\theta}{dt} + \kappa\theta\beta^{\kappa-1}\left(\frac{\partial\beta}{\partial t}_{\theta} + \frac{d\theta}{dt}\frac{\partial\beta}{\partial\theta}\right)\right] \tag{A-1}$$

where the condition that $\beta = \beta(\theta, t)$ is used. From (1), (2), the equation of continuity and an integration the time rate of change of β is

$$\frac{\partial\beta}{\partial t}_{\theta} = -\int_{\theta}^{\theta_T} \nabla_{\vartheta}\cdot\overline{\rho J\underline{u}}\, d\vartheta + \overline{\rho J\frac{d\theta}{dt}} \tag{A-2}$$

where ϑ is a variable of integration representing θ and the condition that $\rho J\frac{d\theta}{dt}$ is zero at θ_T is applied. The combination of (A-1) and (A-2) with (25) yields

$$\frac{dE_r}{dt} = c_v\left(\frac{g}{p_{00}}\right)^{\kappa}\int_{V_\theta}\rho J\left\{\beta^{\kappa}\frac{d\theta}{dt} + \kappa\theta\beta^{\kappa-1}\left[-\int_{\theta}^{\theta_T}\nabla_{\vartheta}\cdot\overline{\rho J\underline{u}}\, d\vartheta + \overline{\rho J\frac{d\theta}{dt}} + \frac{d\theta}{dt}\frac{\partial\beta}{\partial\theta}\right]\right\}dV_\theta$$

$$+\int_{V_\theta}\nabla_{\theta}\cdot[e_r\rho J(\underline{W}-\underline{u})]dV_\theta \tag{A-3}$$

From the relations

$$\frac{\partial\beta}{\partial\theta} = -\overline{\rho J} \tag{A-4}$$

$$\kappa\theta\,\overline{\rho J}\beta^{\kappa-1} = \theta\frac{\partial\beta^{\kappa}}{\partial\theta} = -\frac{1}{c_p}\left(\frac{p_{00}}{g}\right)^{\kappa}\frac{\partial\varphi_r}{\partial\theta} \tag{A-5}$$

and an area averaging within the storm volume (A-3) reduces to

$$\frac{dE_r}{dt} = c_v \int_{V_\theta} \left(\frac{g}{p_{00}}\beta\right)^\kappa \rho J \frac{d\theta}{dt} - c_p^{-1}\left(\frac{\partial \varphi_r}{\partial \theta} \int_\theta^{\theta_T} \nabla_\theta \cdot \overline{\rho J \underline{u}}\, d\vartheta\right) dV_\theta$$

$$+ \int_{V_\theta} \nabla_\theta \cdot [e_r \rho J(\underline{W} - \underline{u})] dV_\theta \qquad \text{(A-6)}$$

An integration by parts and use of the definition of p_r from (8) yields

$$\frac{dE_r}{dt} = c_v \int_{V_\theta} \left(\frac{p_r}{p_{00}}\right)^\kappa \overline{\rho J \frac{d\theta}{dt}} - c_p^{-1}\left[\frac{\partial}{\partial \theta}\left(\varphi_r \int_\theta^{\theta_T} \nabla_\vartheta \cdot \rho J \underline{u}\, d\vartheta\right) + \phi_r \nabla_\theta \cdot \overline{\rho J \underline{u}}\right] dV_\theta$$

$$+ \int \nabla_\theta \cdot [e_r \rho J(\underline{W} - \underline{u})] dV_\theta \qquad \text{(A-7)}$$

The second right-hand term integrates to the vertical boundaries and vanishes from the condition that ϕ_r is zero at θ and $\int_\theta^{\theta_T} \nabla_\vartheta \cdot \overline{\rho J u}\, d\vartheta$ is zero at θ_T. Finally, from the divergence theorem, the time rate of change of the internal energy of the reference state becomes

$$\frac{dE_r}{dt} = c_v \int_{V_\theta} \left(\frac{p_r}{p_{00}}\right)^\kappa \overline{\rho J \frac{d\theta}{dt}} - \int_\theta^{\theta_T} \oint \rho J \left[\frac{c_v}{c_p}\phi_r \underline{u} + e_r(\underline{u} - \underline{W})\right] \cdot \underline{n}\, dL d\theta \qquad \text{(A-8)}$$

The rate of change of specific gravitational potential energy using Leibnitz's rule to differentiate $\phi_r(\theta, t)$ from (10) is

$$\frac{d\phi_r}{dt} = -c_p \left(\frac{g}{p_{00}}\right)^\kappa \left\{ \int_\theta^{\theta_T} \theta \frac{\partial^2 \beta^\kappa}{\partial \theta \partial t} d\theta - \theta \frac{\partial \beta^\kappa}{\partial \theta} \frac{d\theta}{dt} \Big]_{\theta_0}^{\theta} \right\} \qquad \text{(A-9)}$$

The use of (A-2) and (A-4) yields

$$\frac{d\phi_r}{dt} = R\left(\frac{g}{p_{00}}\right)^\kappa \left\{ \int_{\theta_0}^{\theta} \theta \frac{\partial}{\partial \theta}\left\{\beta^{\kappa-1} \int_\theta^{\theta_T} \nabla_\vartheta \cdot \overline{\rho \underline{J} \underline{u}} d\vartheta\right\} - \theta \frac{\partial}{\partial \theta}\left(\beta^{\kappa-1} \overline{\rho J \frac{d\theta}{dt}}\right) d\theta \right.$$

$$\left. + \theta \beta^{\kappa-1} \overline{\rho J \frac{d\theta}{dt}} \Big]_{\theta_0}^{\theta} \right\} \qquad \text{(A-10)}$$

From an integration by parts of the second right-hand term, the time rate of change of specific internal energy becomes

$$\frac{d\varphi_r}{dt} = R\left(\frac{g}{p_{00}}\right)^{\kappa}\left\{\int_{\theta_0}^{\theta}\theta\frac{\partial}{\partial\theta}\left(\beta^{\kappa-1}\int_{\theta}^{\theta_T}\nabla_{\vartheta}\cdot\overline{\rho J\underline{u}}\,d\vartheta\right)+\beta^{\kappa-1}\overline{\rho J\frac{d\theta}{dt}}\,d\theta\right.$$

$$\left.-\theta\beta^{\kappa-1}\left[\overline{\rho J}\frac{d\theta}{dt}-\overline{\rho J\frac{d\theta}{dt}}\right]_{\theta_0}^{\theta}\right\} \qquad \text{(A-11)}$$

The velocity of the lower boundary of the reference atmosphere is defined by

$$\frac{d\theta_0}{dt} = \left[\overline{\rho J\frac{d\theta}{dt}}\Big/\overline{\rho J}\right]_{\theta_0} \qquad \text{(A-12)}$$

i.e., it is determined by area averaged vertical mass transport through θ_0 in the actual atmosphere. From the use of the lower boundary condition, (A-11) reduces to

$$\frac{d\phi_r}{dt} = R\left(\frac{g}{p_{00}}\right)^{\kappa}\left[\int_{\theta_0}^{\theta}\theta\frac{\partial}{\partial\theta}\left(\beta^{\kappa-1}\int_{\theta}^{\theta_T}\nabla_{\vartheta}\cdot\overline{\rho J\underline{u}}\,d\vartheta\right)+\beta^{\kappa-1}\overline{\rho J\frac{d\theta}{dt}}\right.$$

$$\left.-\theta\beta^{\kappa-1}\left(\overline{\rho J}\frac{d\theta}{dt}-\overline{\rho J\frac{d\theta}{dt}}\right)\right] \qquad \text{(A-13)}$$

The substitution of (A-13) into the time rate of change of gravitational potential energy of the reference state, from (28), yields

$$\frac{d\Phi_r}{dt} = R\left(\frac{g}{p_{00}}\right)^{\kappa}\int_{V_\theta}\overline{\rho J}\left\{\int_{\theta_0}^{\theta}\left[\varphi\frac{\partial}{\partial\varphi}\left(\beta^{\kappa-1}\int_{\theta}^{\theta_T}\nabla_{\vartheta}\cdot\overline{\rho J\underline{u}}\,d\vartheta\right)+\beta^{\kappa-1}\overline{\rho J\frac{d\theta}{dt}}\right]d\varphi\right\}dV_\theta$$

$$+\int_{V_\theta}\nabla_\theta\cdot\left[\phi_r\,\rho J(\underline{W}-\underline{u})\right]dV_\theta \qquad \text{(A-14)}$$

where both φ and ϑ are variables of integration representing θ. The last two terms of (A-13) summed to zero by area averaging within the storm volume.

Through the use of (A-4) and an integration by parts, we find that

$$\frac{d\Phi_r}{dt} = -R\left(\frac{g}{p_{00}}\right)^{\kappa} \int_{V_\theta} \frac{\partial}{\partial\theta}\left\{\beta \int_{\theta_0}^{\theta} \left[\varphi \frac{\partial}{\partial\varphi}\left(\beta^{\kappa-1} \int_{\theta}^{\theta_T} \nabla_\vartheta \cdot \overline{\rho J \underline{u}}\, d\vartheta\right) + \beta^{\kappa-1} \overline{\rho J \frac{d\theta}{dt}}\, d\varphi\right]\right\}$$

$$-\left\{\beta^{\kappa} \overline{\rho J \frac{d\theta}{dt}} + \theta\beta \frac{\partial}{\partial\theta}\left(\beta^{\kappa-1} \int_{\theta}^{\theta_T} \nabla_\vartheta \cdot \overline{\rho J \underline{u}}\, d\vartheta\right)\right\} + \int_{V_\theta} \nabla_\theta \cdot \phi_r \rho J(\underline{u} - \underline{W}) dV_\theta \qquad \text{(A-15)}$$

The first right-hand term integrates to the vertical boundaries and vanishes from the conditions that $\beta(\theta_T)$ is zero and the indefinite integral with respect to φ is zero when evaluated at θ_0. The differentiation of the third right-hand term and use of (8), (9) and (A-5) yields

$$\frac{d\Phi_r}{dt} = \int_{V_\theta} R\left(\frac{p_r}{p_{00}}\right)^{\kappa} \overline{\rho J \frac{d\theta}{dt}} - RT_r \nabla_\theta \cdot \overline{\rho J \underline{u}} - \frac{c_v}{c_p}\left(\frac{\partial \varphi_r}{\partial\theta} \int_{\theta}^{\theta_T} \nabla_\vartheta \cdot \overline{\rho J \underline{u}}\, d\vartheta\right)$$

$$+ \nabla_\theta \cdot [\phi_r \rho J(\underline{W} - \underline{u})] dV_\theta \qquad \text{(A-16)}$$

Again through an integration by parts, (A-16) becomes

$$\frac{d\Phi_r}{dt} = \int_{V_\theta} R\left(\frac{p_r}{p_{00}}\right)^{\kappa} \overline{\rho J \frac{d\theta}{dt}} - RT_r \nabla_\theta \cdot \overline{\rho J \underline{u}} - \frac{c_v}{c_p}\left\{\frac{\partial}{\partial\theta}\left[\phi_r \int_{\theta}^{\theta_T} \nabla_\vartheta \cdot \overline{\rho J \underline{u}}\, d\vartheta\right.\right.$$

$$\left.\left. + \phi_r \nabla_\theta \cdot \overline{\rho J \underline{u}}\right\}\right] + \nabla_\theta \cdot [\varphi_r \rho J(\underline{W} - \underline{u})] dV_\theta \qquad \text{(A-17)}$$

Through vertical integration and an application of boundary conditions, the third right-hand term is zero. Finally, from a rearrangement and use of the divergence theorem, the rate of change of the specific gravitational potential energy becomes

$$\frac{d\Phi_r}{dt} = R \int_{V_\theta} \left(\frac{p_r}{p_{00}}\right)^{\kappa} \overline{\rho J \frac{d\theta}{dt}} dV_\theta - \int_{\theta_0}^{\theta_T} \oint \rho J\left[\left(\frac{R}{c_v} e_r - \frac{c_v}{c_p} \phi_r\right) \underline{u}\right.$$

$$\left. + \phi_r(\underline{u} - \underline{W})\right] \cdot \underline{n}\, dL\, d\theta \qquad \text{(A-18)}$$

Scanner's note:

This page is blank.

Reprinted from MONTHLY WEATHER REVIEW, Vol. 99, No. 1, Jan. 1971, pp. 1-14.

UDC 551.515.11:551.511.33

THE GENERATION OF AVAILABLE POTENTIAL ENERGY BY LATENT HEAT RELEASE IN A MID-LATITUDE CYCLONE

BEN R. BULLOCK[1] and DONALD R. JOHNSON

Department of Meteorology, The University of Wisconsin, Madison, Wis.

ABSTRACT

The theory of available potential energy applied to a "translating" atmospheric volume is used to estimate the generation of available potential energy for the cyclogenetic, mature, and occluding stages of a mid-latitude cyclone. Primary attention is focused upon the importance of the diabatic process of latent heat release in generating the storm's available potential energy. In addition, preliminary estimates for the process of infrared cooling are presented.

Total latent heat release is determined from observed precipitation rates. Three different models for the vertical distribution of latent heat release together with the storm's structure provided by isentropic analyses are utilized to estimate the contribution by latent heat release to the available potential energy of the disturbance. For each stage of the storm, variations in the generation estimates between models were extremely small. For the cyclogenetic, mature, and occluding stages, generation estimates of approximately 1, 8, and 6 W m^{-2}, respectively, reflected changes in the horizontal distribution of precipitation about the storm.

Two simplified cooling distributions were assumed to evaluate the importance of infrared cooling in the generation. The first was one of uniform cooling at the rate of 1.4°C day^{-1} thoughout the volume, and the second was one in which the clear air was cooled at a greater rate than the cloudy air. Positive generation estimates on the order of 1 to 2 W m^{-2} resulted from these calculations.

The results of this study indicate that the diabatic process of latent heat release is very likely an important factor in the subsequent behavior of the system. It is speculated that an energy supply of this magnitude, available for immediate conversion to kinetic energy, is sufficient to offset a major portion of the storm's frictional dissipation. Generation estimates for the process of infrared cooling, while less reliable than those for latent heat release, indicate that this process also contributes to the storm's available potential energy supply.

1. INTRODUCTION

In contrast to earlier theories of the general circulation envisioning one or more meridional cells, atmospheric scientists have rejected this simplified view and emphasized the importance of the atmosphere's quasi-horizontal motion in satisfying the required heat and momentum transport for the global scale. The general circulation theory formulated during the last two decades and summarized by Lorenz (1967) is based upon evidence that the "nonzonal inequalities" are the principal site of the conversion of atmospheric potential energy to kinetic energy (Starr 1958). According to this theory, the east-west thermal contrast within the migratory systems represents available potential energy that is converted to kinetic energy through the process of rising warm air and sinking cold air.

The importance of the adiabatic conversion of available potential energy to kinetic energy within the cyclones and anticyclones is a central feature of this theory. Recently, however, the importance of diabatic effects, in particular latent heat release, within the cyclonic circulations is being realized. Investigations by Aubert (1957), Eliassen (1962), Rao (1966), and others have emphasized the contribution of released latent heat to the vertical motion field, while Danard (1964, 1966) has shown that this process contributes to available potential energy production and an increased rate of energy conversion during the mature stage of a cyclone. These results support Petterssen's (1960) belief that it is not reasonable to treat the mobile weather systems as manifestations of adiabatic processes while simultaneously searching for energy sources and sinks to account for the general circulation of which these systems are an integral part.

[1] Now affiliated with the Department of Earth Science, Edinboro State College, Edinboro, Pa.

To provide a better understanding of the importance of diabatic processes within the energetics of cyclones, this study focuses primary attention on the distribution of released latent heat in producing energy available for conversion to kinetic energy within the scale of these systems. Brief consideration is also given to its production by the process of infrared emission.

The generation of available potential energy by the release of latent heat is studied using models approximating the vertical heating distribution. The total latent heat release in the vertical is determined from the surface precipitation rate, while the distribution of heating is specified by relating variations in the condensation rate to models of vertical motion. These heating models and the atmospheric structure provided by isentropic analyses are used to estimate the generation of available potential energy for the developing, mature, and early occluding stages of a mid-latitude cyclone. The generation by infrared radiative processes is obtained from assumed infrared cooling distributions. This approach provides an estimate of the in situ generation of the storm's available potential energy. The boundary layer frictional dissipation is also estimated to contrast the relative importance of the storm's energy source by diabatic processes with its sink by frictional effects.

2. BASIC ENERGY EQUATIONS

Following Margules (1903), Lorenz (1955) defined available potential energy as the sum of the internal and gravitational potential energy in the actual atmosphere minus the sum of these energies after an adiabatic mass redistribution of the global atmosphere to a horizontal, statically stable density stratification. The atmospheric configuration achieved by this hypothetical redistribution is commonly called the reference state.

Unfortunately, available potential energy theory based upon a global reference state does not provide a firm basis for energy studies of smaller scales of atmospheric motion. Because the most intense conversion of available potential energy to kinetic energy occurs in the migratory cyclones and anticyclones of middle latitudes, budget studies of this scale of motion are necessary for a proper understanding of their role in the general circulation. The theoretical basis for such studies has been provided recently by Johnson (1970), who has developed budget equations for the available potential energy of a storm by defining a reference state for a volume surrounding and moving with the disturbance. Within this volume, the storm's available potential energy is

$$\overline{A}=\overline{\pi}-\overline{\pi}_r \tag{1}$$

where π and π_r are respectively the total potential energy (internal plus gravitational) in the existing atmosphere and the reference state. The overbar represents the area average of a quantity, and the subscript r represents the value of a quantity in the reference state.

The time rate of change of the storm's available potential energy is

$$\frac{d\overline{A}}{dt}=\frac{d\overline{\pi}}{dt}-\frac{d\overline{\pi}_r}{dt} \tag{2}$$

that shows the dependence on the changes of the actual and reference state energies in the translating volume. From Johnson (1970), the rate of change of the storm's available potential energy is

$$\frac{d\overline{A}}{dt}=\overline{G}-\overline{C}(A, K)+\overline{B} \tag{3}$$

where G is the generation of available potential energy by diabatic processes, $C(A, K)$ is the rate of conversion of available potential energy to kinetic energy, and B is the change of available potential energy produced by pressure work on and relative advection through the boundary of the translating volume. The major differences between the available potential energy expressions for a global or storm budget are the boundary term and the redefined reference state. If the volume is extended to include the entire atmosphere, the boundary term vanishes, and the generation and conversion correspond to the global expressions derived by Dutton and Johnson (1967). In this study, only the generation of the storm's available potential energy is determined.

In isentropic coordinates, the generation of the storm's available potential energy is

$$\overline{G}=\int_{\theta_0}^{\theta_T}\overline{\epsilon\rho J_\theta Q_m}d\theta \tag{4}$$

where ρ is density, Q_m is the rate of specific heat addition, θ_0 and θ_T are the atmosphere's minimum and maximum potential temperatures, respectively. J_θ is the transformation Jacobian $|\partial z/\partial\theta|$. The efficiency factor in eq (4) is

$$\epsilon=1-\left(\frac{p_r}{p}\right)^\kappa \tag{5}$$

where κ is the ratio of the gas constant R to the specific heat at constant pressure c_p and p_r is the reference pressure of an isentropic surface. Under the hydrostatic assumption, the reference pressure p_r may be evaluated by area-averaging the atmospheric pressure on an isentropic surface (Dutton and Johnson 1967).

In diagnostic energy studies of the effects of diabatic heating, the efficiency factor is an important element. It is a measure of the condition that for diabatic processes the change of total potential energy in the actual atmosphere may differ from that in the reference atmosphere. Examination of eq (2) shows that available potential energy is produced when the change of atmospheric total

potential energy exceeds the change of the reference atmosphere's total potential energy. A positive (negative) generation results when heating (cooling) occurs at pressures higher and cooling (heating) at pressures lower than the reference pressure.

In this study, the spatial distribution of efficiency factors is obtained from isentropic analyses of a mid-latitude cyclone, while observed surface precipitation rates are used to establish total latent heat release. Models developed in the next section, which approximate the vertical distribution of latent heat release, are then used in estimating the generation.

3. MODELS OF LATENT HEAT RELEASE

Latent heat release by the condensation process represents the sink of water vapor within the storm, while sources of water vapor are primarily from low-level advection of warm moist air from oceanic areas into the storm. The relation between the storm's water vapor sources, sinks, and storage is first developed in a water vapor budget equation before modeling the latent heat release.

THE WATER VAPOR BUDGET EQUATION

The development of a water vapor budget equation for a volume moving with the storm is based on the generalization of the Leibnitz rule for the differentiation of integrals. In general, the total water vapor mass in an arbitrary volume defined in pressure coordinates is

$$Q=\int_{V_p} qJ_p dV_p \tag{6}$$

where q is the specific humidity, J_p is the transformation Jacobian $|\partial z/\partial p|$ and V_p is the volume of integration bounded by vertical walls and lower and upper pressure surfaces p_1 and p_2. The differential dV_p is equal to $dpdxdy$. From the generalization of the Leibnitz rule to three dimensions, the time rate of change of the storm's water vapor mass is

$$\frac{dQ}{dt}=\int_{V_p}\left[\frac{\partial}{\partial t}(q\rho J_p)+\nabla_p\cdot(q\rho J_p\mathbf{W})+\frac{\partial}{\partial_p}(q\rho J_p\omega_B)\right]dV_p \tag{7}$$

where again $\mathbf{W}$ is the lateral boundary velocity and ω_B is the velocity of the vertical boundary with respect to the vertical pressure coordinate.

By substituting the continuity equation in pressure coordinates

$$\frac{d}{dt}(\rho J_p)+\rho J_p\left(\nabla_p\cdot\mathbf{U}+\frac{\partial\omega}{\partial p}\right)=0 \tag{8}$$

into eq (7), the budget equation becomes

$$\frac{dQ}{dt}=\int_{V_p}\left\{\rho J_p\frac{dq}{dt}+\nabla_p\cdot[q\rho J_p(\mathbf{W}-\mathbf{U})]+\frac{\partial}{\partial p}[q\rho J_p(\omega_B-\omega)]\right\}dV_p \tag{9}$$

where ω is the fluid vertical velocity in pressure coordinates.

From use of the divergence theorem, the budget equation for the storm's water vapor mass is

$$\frac{dQ}{dt}=\int_A\int_{p_2}^{p_1}\rho J_p\frac{dq}{dt}dpdA-\int_{p_2}^{p_1}\int_L q\rho J_p(\mathbf{U}-\mathbf{W})\cdot\mathbf{n}dLdp+\int_S\{[q\rho J_p(\omega-\omega_B)]_{p_2}-[q\rho J_p(\omega-\omega_B)]_{p_1}\}dS. \tag{10}$$

The lower vertical boundary of the volume will be the surface of the earth. This condition is maintained by equating p_1 to p_0 and $\omega_B(x,y,p_0,t)$ to $\omega(x,y,p_0,t)$. Thus the fluid always remains in contact with the earth's surface. Now, by defining the relative velocity of the fluid, $\mathbf{U}_r$, to be the difference between $\mathbf{U}$ and $\mathbf{W}$ and letting the upper surface of the volume be a fixed pressure surface [so that $\omega_B(p_2)$ vanishes], the storm's water vapor budget equation becomes

$$\frac{dQ}{dt}=\int_A\int_{p_2}^{p_0}\rho J_p\frac{dq}{dt}dpdA-\int_{p_2}^{p_0}\int_L\mathbf{n}\cdot(q\rho J_p\mathbf{U}_r)dLdp+\int_S(q\rho J_p)_{p_2}\omega_2 dS.$$

From the hydrostatic assumption and the expression

$$\frac{dq}{dt}=\omega\frac{dq}{dp}, \tag{11}$$

the final budget equation for the storm's water vapor mass is

$$\frac{dQ}{dt}=\frac{1}{g}\int_A\int_{p_2}^{p_0}\omega\frac{dq}{dp}dpdA-\frac{1}{g}\int_{p_2}^{p_0}\int_L\mathbf{n}\cdot q\mathbf{U}_r dLdp+\frac{1}{g}\int_S q_2\omega_2 dS. \tag{12}$$

The first term in eq (12) represents sources and sinks by evaporation and condensation of water vapor within the moving volume and will be denoted by $S(Q)$. The last two terms measure the flux of water vapor through the lateral boundary by the relative velocity of the fluid and by vertical advection through the volume's upper surface. If the volume is extended to the top of the atmosphere, the last term also vanishes under the condition that $\omega(0)$ is zero.

In this study, the generation of available potential energy by the release of latent heat is emphasized. Thus, only the sink of water vapor by the process of condensation

is considered. In a mature storm, the water vapor storage tends to be constant, such that the sink by condensation is primarily balanced by lateral inflow in the low troposphere. Vertical advection through the upper surface must be small for budget volumes extending into the stratosphere or high troposphere.

In all three models employed in this study, it is assumed that condensation occurs for upward vertical motion when q is equal to q_s. Thus, the change of water vapor mass in the subregion of the source integral where these conditions are satisfied is

$$S(Q_s)=\frac{1}{g}\int_{A_s}\int_{p_{2s}}^{p_{1s}}\omega\frac{dq_s}{dp}\,dp dA_s \quad \{q=q_s;\ \omega<0\} \tag{13}$$

where the subscript s denotes the condition of saturation and the pressures p_{1s} and p_{2s} denote the lower and upper boundary pressures for the precipitating portion of the volume.

Under the assumption that all the condensed water vapor precipitates immediately, the relation between $S(Q_s)$ and the precipitation rate is

$$S(Q_s)=-\int_{A_s}\rho_w\dot{P}dA_s \tag{14}$$

where $\dot{P}$ is the rate of precipitation (cm hr^{-1}), ρ_w is the density of water, and A_s is the area of the precipitating region.

By combining eq (13) and (14) and using the knowledge that the precipitation rate is a measure of the rate of latent heat release, the total heating in the region defined for $S(Q_s)$ is

$$H_L(S)=\int_{A_s}L_{ls}\rho_w\dot{P}dA_s=-\frac{1}{g}\int_{A_s}\int_{p_{2s}}^{p_{1s}}L_{ls}\omega\frac{dq_s}{dp}\,dp dA_s \tag{15}$$

where L_{ls} is the latent heat of condensation. Note the integrand of eq (15), a function of the vertical motion distribution and the individual change of specific humidity with pressure, is equal to the latent heating component of Q_m. In the next section, three models of latent heat release are developed to estimate the spatial distribution of latent heat release in the cyclone.

PHYSICAL MODELING

Models for the vertical distribution of latent heat release are obtained by assuming different profiles of the vertical motion and the individual change of specific humidity with pressure that satisfy the integral constraint of eq (15). Each model utilizes the observed surface precipitation rate to establish the total latent heat release in the atmospheric column.

In the models, heat is distributed between the lifting condensation level p_{1s} and an assumed cloud top level corresponding to the upper boundary of the precipitating volume p_{2s}. Examination of soundings at stations reporting steady precipitation showed close agreement between the lifting condensation level and the observed cloud bases. However, because cloud tops are not easily determined from soundings, several different cloud top levels are selected to study the influence of various heating profiles in the estimates for the generation of available potential energy.

First model of latent heat release. Bradbury (1957) and others have shown that the major flux of moisture into extratropical cyclones occurs below 800 mb. Because the equivalent potential temperature of air below 800 mb tends to be constant due to mixing in the boundary layer, Bradbury's result supports the assumption of a constant equivalent potential temperature for the first model. Thus, the moist adiabat through the lifting condensation level (LCL) defines the profile of the individual change of specific humidity. It is also assumed in the first model that the vertical velocity from the LCL to the cloud top is constant. The observed precipitation rate at the surface and the assumed cloud top uniquely determine the value of the vertical motion necessary to satisfy eq (15).

Typical profiles of vertical motion and latent heat release from this model are shown in figure 1A. The small heating maximum in the middle troposphere is associated with the maximum value of dq_s/dp on the pseudoadiabat.

Second model of latent heat release. The assumption of constant vertical motion made in the first model, though not unreasonable for shallow saturated layers, is not representative of large-scale vertical velocities over any substantial vertical extent of the atmosphere. In the second model, the vertical motion profile is assumed to be quadratic between the earth's surface and the cloud top, while the assumption of constant equivalent potential temperature for the ascending air is retained. A quadratic form for the vertical motion is

$$\omega=\omega_m(1-X^2) \tag{16}$$

where ω_m is the maximum upward vertical velocity at X equal to zero. The scaled pressure variable is

$$X=\frac{p-\frac{1}{2}(p_{2s}+p_0)}{\frac{1}{2}(p_{2s}-p_0)} \quad \begin{matrix} p_0\geq p\geq p_{2s} \\ -1\leq X\leq +1 \end{matrix} \tag{17}$$

where p_0 and p_{2s} are the surface and the cloud top pressures, respectively.

Figure 1B presents the profiles of vertical motion and latent heat release from the second model for the same initial data used in the first model. Note that the heating maximum in the middle troposphere for the second model is substantially greater than the first model.

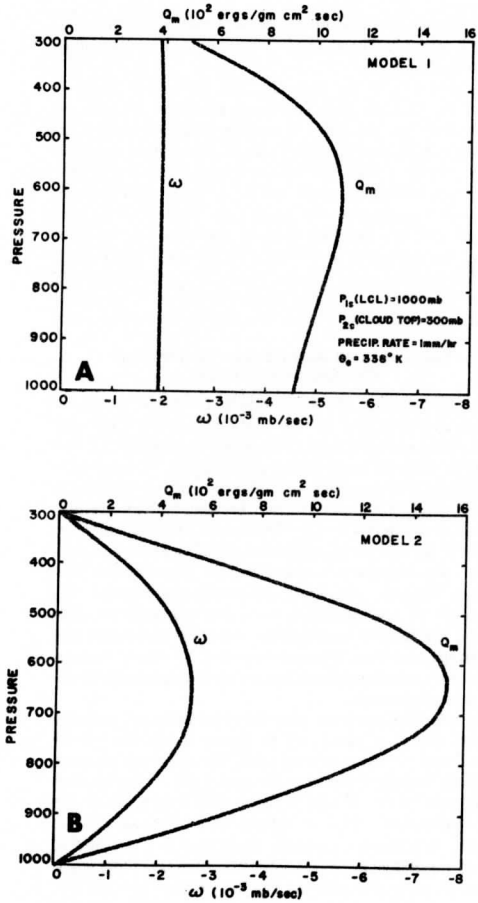


FIGURE 1.—Profiles of vertical motion and latent heat release for (A) model 1 and (B) model 2. Parameters for both models are indicated in the lower left corner of (A).

Third model of latent heat release. The assumption of constant equivalent potential temperature made in the previous two models implies that the lapse rate in the saturated portion of the column is moist adiabatic. Examination of soundings with deep saturated layers producing precipitation shows departures from this condition. In the third model, vertical variations in the equivalent potential temperature are allowed, while using the quadratic modeling of the vertical motion. The assumption of saturation on isentropic surfaces above the LCL but below the cloud top level together with pressure analyses of these surfaces determines the spatial distribution of equivalent potential temperature.

The vertical distribution of heating from the second and third models is quite similar. However, in the stable atmosphere, the third model releases more heat in the upper portion of the atmosphere than the second model because the equivalent potential temperature is assumed to increase with height.

4. COMPUTATIONAL PROCEDURES

PRECIPITATION RATE AND LATENT HEAT ADDITION

Precipitation rates used to establish the latent heating distribution were obtained from hourly amounts recorded at several hundred stations throughout the United States. At each location, an average precipitation rate for a 4-hr period was determined. The area distribution of precipitation for a 1° grid was obtained by averaging the station data lying within a 0.6° latitude radius of each grid point. The objectively analyzed field using this scheme compared favorably with a subjective analysis.

Obviously, some error is made when using rainfall amounts to estimate latent heat release, since condensation can occur without precipitation. Wexler and Atlas (1958) state, however, that cloud storage of liquid water is small, especially in stratiform clouds producing precipitation and need not be considered in the calculation of rainfall rates. This indicates that errors in latent heat estimates induced by condensation without precipitation are minimal and the vertical profile of heating may be determined by using the total latent heat release based on the precipitation rate. The finite estimate of the latent heating per unit mass corresponding to the condensation rate at the ith grid point on the kth isentropic surface is

$$Q_{m_{ik}} = -L_{iv}\omega_{ik}\left(\frac{dq_s}{dp}\right)_{ik}. \tag{18}$$

EFFICIENCY FACTOR

The efficiency factors were obtained from 10 isentropic surfaces equally spaced at 5°K intervals from 280° to 325°K. In computing the efficiency factor, the reference pressure $p_r(\theta_k)$ was obtained for a circular volume with a 1700-km radius centered about the storm; and on underground portions of isentropic surfaces (Lorenz 1955), the pressure was set equal to the surface pressure. The grid-point estimate of the efficiency factor is

$$\epsilon_{ik} = 1 - \left(\frac{p_{rk}}{p_{ik}}\right)^{\epsilon}. \tag{19}$$

GENERATION

Under the hydrostatic assumption, the finite approximation of the generation of available potential energy is

$$G=\frac{1}{g}\sum_{k=1}^{10}\sum_{i=1}^{n}G_{ik}\Delta p_{ik} \tag{20}$$

where n is the number of grid points and Δp_{ik} is one-half the difference between p_{k-1} and p_{k+1}. The generation estimate at each ikth point is

$$G_{ik}=\epsilon_{ik}Q_{m_{ik}}. \tag{21}$$

FRICTIONAL DISSIPATION

Frictional dissipation estimates for the planetary boundary layer are obtained using Lettau's (1961) formulation. The dissipation at the ith grid point is

$$D_i=\rho_i C_{g_i}^2\, U_{g_i}^3 \cos\alpha \tag{22}$$

where U_g is the geostrophic wind and α is the angular departure of the geostrophic wind vector from the surface shear stress. The geostrophic drag coefficient (Kung 1963) is

$$C_{g_i}=k\,[\ln(U_{g_i}/z_0 f_i)-1.865]^{-1} \tag{23}$$

where k is the von Kármán constant equal to 0.4, z_0 is the roughness parameter, and f is the Coriolis parameter. In this study, z_0 was assumed to be 50 cm, a representative value for the geographical region of the storm (Kung 1963), and α was assumed to be 20°.

5. THE CASE STUDY

SYNOPTIC DESCRIPTION

The cyclone selected for study developed rapidly on Mar. 25–26, 1964. The surface and 300°K isentropic analyses for three successive stages of the storm are presented in figure 2. At 1200 GMT on the 25th, the cyclone was located in northern Arkansas and southern Missouri along a frontal zone associated with strong low-level northward advection of warm, moist subtropical air (mT) from the Gulf of Mexico and southward advection of continental polar air (cP) from the northern Plains States.

The region of lowest pressure on the 300°K isentropic surface at 1200 GMT on the 25th shows that the coldest air was situated over Colorado and Wyoming. A strong baroclinic zone, indicated by the large gradient of pressure, was located to the rear of the developing depression. Above the baroclinic zone, 500-mb winds attained speeds of 80 kt in a portion of the jet core extending from northern Texas to Iowa.

By 0000 GMT on the 26th (fig. 2C), the cyclone was well developed and was moving northeastward at about 30 kt while deepening at the rate of approximately 1 mb hr^{-1}. The vertical axis of the system still sloped to the west, with the 500-mb trough extending from the Dakotas to western Texas. The baroclinic zone (fig. 2D) that was located west of the surface position of the cyclone had intensified. The 500-mb winds in excess of 90 kt extended from Oklahoma to northern Illinois.

By 1200 GMT on the 26th (fig. 2E), the system had just begun to occlude after reaching southern Ontario. A closed Low on the 500-mb surface centered over Lake Michigan indicated that the axis of the system was now nearly vertical. The 300°K isentropic analysis (fig. 2F) shows the baroclinic zone extending from directly over the surface system to the southwest. This is to be contrasted with its location well to the west of the Low during the cyclogenetic (March 25 1200 GMT) and mature (March 26 0000 GMT) stages. Again, the maximum 500-mb wind region extending from Tennessee to southern Ontario was situated above the 300°K baroclinic zone.

VERTICAL AND HORIZONTAL DISTRIBUTION OF EFFICIENCY FACTORS

The generation of the storm's available potential energy is a function of the diabatic heating field and thermodynamic structure. The latter is implicitly manifested in the efficiency factor distribution. The pressure analysis on 10 isentropic surfaces was utilized to provide the efficiency factor distribution because isobars are also isolines of efficiency factors. In figures 2B, 2D, and 2F, efficiency factors for the 1700-km radial area corresponding to the isobars are presented in parentheses. The pressure analysis, primarily indicating the vertical displacement of the isentrope from low-tropospheric subtropical air to high-tropospheric polar air illustrates the relationship between thermal structure, efficiency factors, and regions where net heating and cooling generate available potential energy.

The reference pressure, also the area-averaged pressure for the storm, is presented in figure 3. As the storm evolved and moved northeastward, the reference pressure for all the isentropic surfaces decreased due to the relative inward flux of cP air and outward flux of mT air at the lateral boundary. Note in figures 2B, 2D, and 2F the decrease in reference pressure with time for the 300°K surface resulted in increasing the positive efficiency factors at high pressures in the subtropical air but decreasing the magnitude of the negative efficiencies at low pressures in the polar air.

Cross sections of potential temperature and efficiency factors along latitude circles were also prepared to study the time and vertical variation of efficiency factors within all stages of the storm. However, only three cross sections through the mature stage are presented in figure 4. For their position with respect to the storm center, refer to figure 2C. The cross sections show that the largest positive values occur at about 750 mb southeast of the surface low-pressure center, while the largest negative values occur at the 300-mb level in the polar troposphere northwest of the surface frontal zone. The three cross sections portray the contraction of the positive efficiency area and expansion of the negative area as one moves northward.

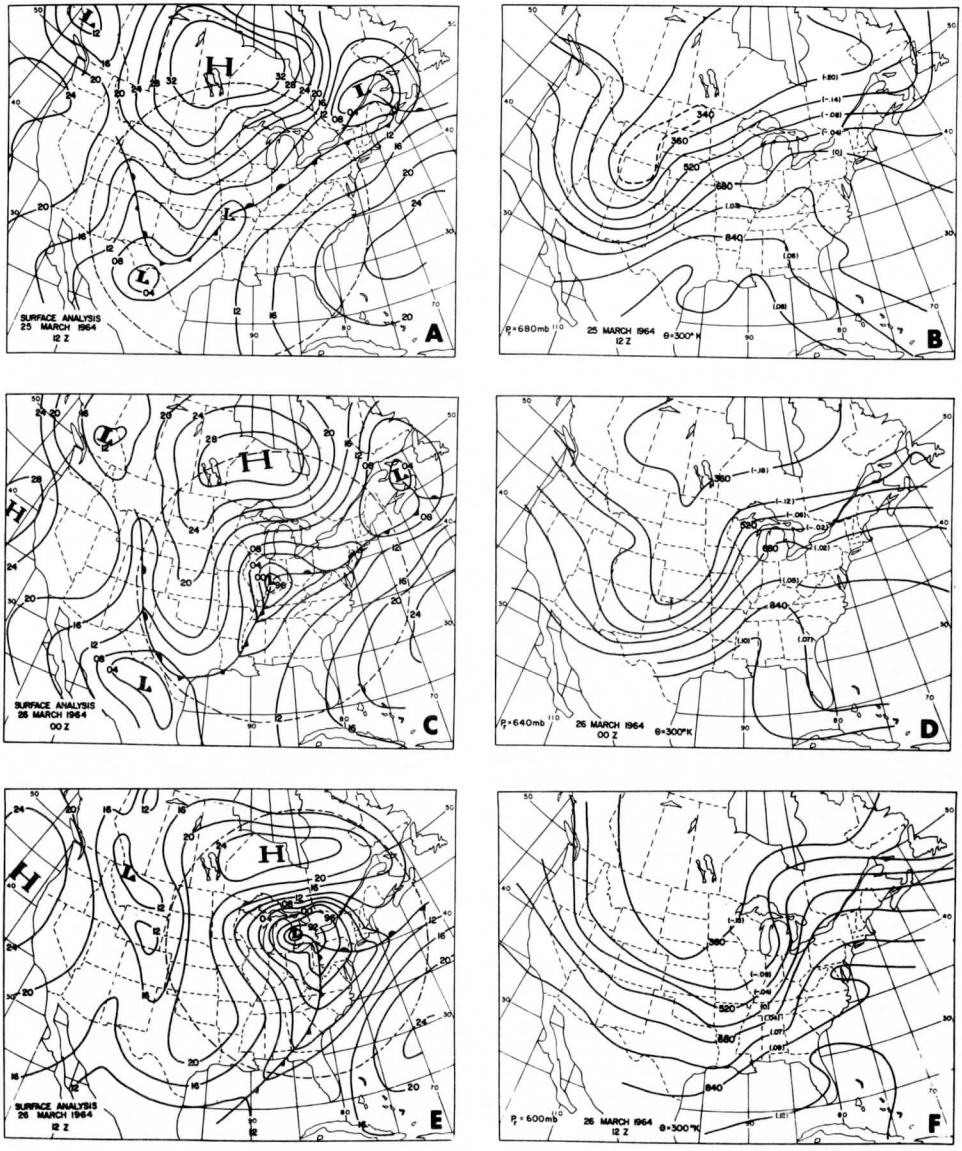


FIGURE 2.—Surface analyses for the cyclogenetic, mature, and occluded stages are presented in (A), (C), and (E). The dashed circular line outlines the region of the storm for which the generation was estimated. Pressure analyses (in mb) on the 300°K surface for the three stages are presented in (B), (D), and (F). The values of the efficiency factor isopleth for the 1700-km area that corresponds to the isobars are indicated in parentheses.

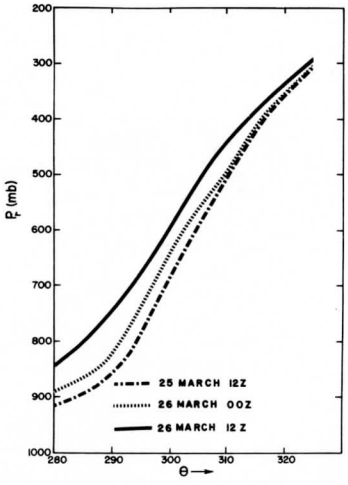


FIGURE 3.—Vertical profile of the reference pressure for the three stages of the storm.

While the same general features prevailed in the cross sections for both the developing and occluding stages, there were some differences. During the 24-hr period, the positive values in the subtropical air increased, and the absolute values of the efficiency factors in the polar air decreased. This tendency was related to the decrease in reference pressures as the storm evolved (fig. 3).

The cross sections and the 300°K isentropic analyses for all stages of the storm show that the zero efficiency factor isoline approximately coincides with the location of a major precipitation area commonly located north of the Low center. Released latent heat in these regions would therefore contribute to very small or even negative generation. In contrast, the peripheral areas of the cyclone where efficiency factors attain their largest absolute values are the regions where diabatic processes assume greater importance in the generation of available potential energy for the storm.

Another interesting feature noted in the evolution of the efficiency structure is the tendency for the zero efficiency isopleth to be more horizontal in the early stage and more vertical in the mature and occluded stage. In considering the effects of total heating, the efficiency structure indicates that low tropospheric heating and high tropospheric cooling are more efficient mechanisms in generating available potential energy in the cyclogenetic stage and that horizontal differential heating will become more efficient in the subsequent stages. A similar evolution was noted by Hahn and Horn (1969).

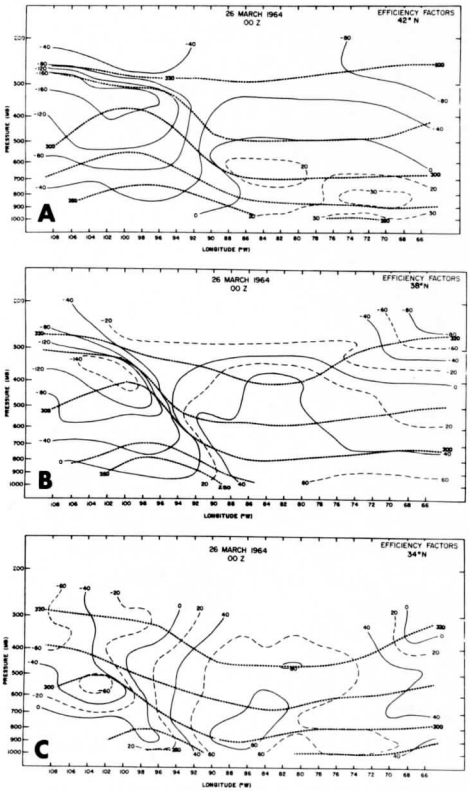


FIGURE 4.—East-west cross sections of efficiency factors (solid line, values times 10^{-3}) and potential temperature (dotted line, °K) for the mature stage.

PRECIPITATION DISTRIBUTION

Precipitation rates used to establish total latent heat release are presented in figures 5A, 5C, and 5E. During the cyclogenetic stage, precipitation occurred in both the subtropical and polar air. Snow showers in the Western States resulted from upslope motion and low-level instability in the cold air. The region of maximum warm frontal precipitation extending from northeastern Missouri to western New York just north of the warm front resulted from thunderstorm activity and convective instability in the subtropical air. In some cases, snow was observed with these convective cells. The region of heaviest precipitation in northern Georgia, Alabama, and eastern Tennessee

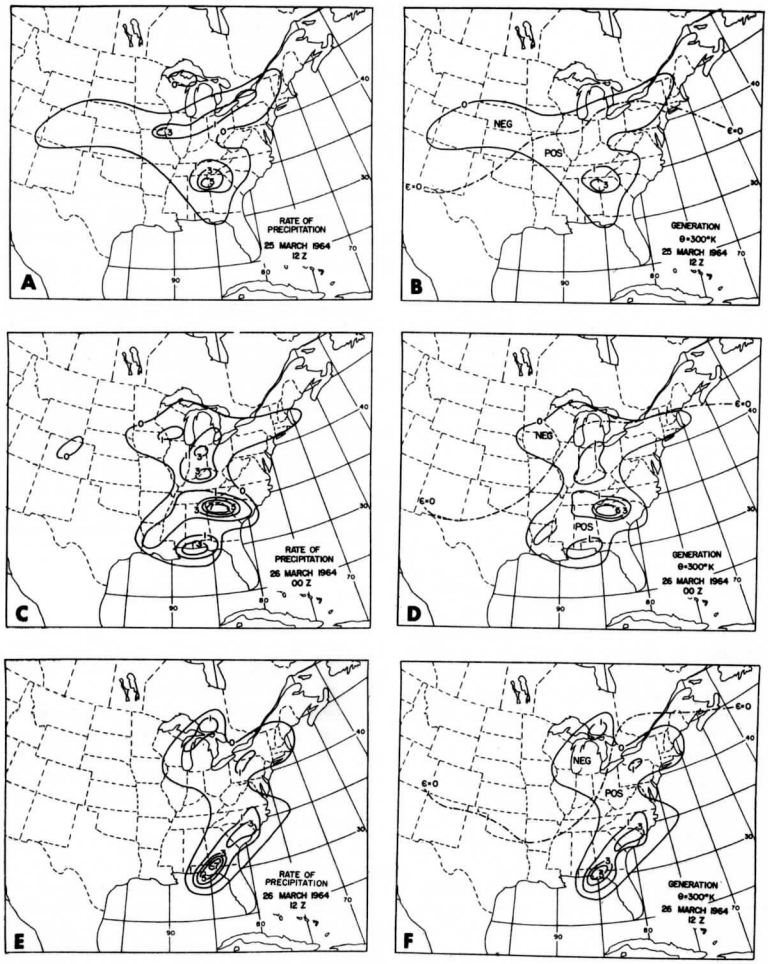


FIGURE 5.—Rate of precipitation (mm hr^{-1}) and generation estimates (10^2 ergs g^{-1} cm^{-2} s^{-1}) on the 300°K surface for the cyclogenetic, mature, and occluded stages.

resulted from thunderstorms substantially ahead of the storm center in the maritime air.

During the mature stage of the storm (March 26 at 0000 GMT, fig. 5C), precipitation in the cold air north and east of the Low resulted primarily from the large-scale vertical motion field, while the areas of maximum precipitation in the warm sector of the cyclone (fig. 2C) resulted mainly from convective activity.

By March 26 at 1200 GMT, in the beginning of the occluded stage (fig. 5E), lower precipitation rates were observed in the northern portions of the system, probably resulting from the fact that the low-level moisture source

TABLE 1.—*Estimates of the generation* [watts (W) m^{-2}] *of available potential energy by latent heat release and of the frictional dissipation* (W m^{-2}) *in the boundary layer*

Model of latent heat release	1	2	3			Frictional dissipation
Assumed cloud top (mb)	300	300	200	300	400	
Mar. 25, 1964, 12 (GMT)	1.0	1.0	0.7	0.9	1.1	3.1
Mar. 26, 1964, 00	7.9	8.0	7.8	8.0	8.3	5.1
Mar. 26, 1964, 12	6.0	6.0	5.9	6.0	6.2	9.1

was being cut off in the occlusion process. Convective activity along the cold front produced the largest precipitation amounts in the southeastern United States.

Using all three models described in section 3, the rate of release of latent heat was computed for isentropic surfaces between the lifting condensation level and the assumed cloud top, and analyses were prepared for the 300°K surface. The distribution of latent heat release on this surface closely resembled that of the surface precipitation rates, and the analyses are not presented.

THE GENERATION AND FRICTIONAL DISSIPATION

Generation estimates using the three models described in section 3 were computed for the storm. A cloud top of 300 mb was assumed for the computations with models 1 and 2, while cloud tops of 200, 300, and 400 mb were used with model 3 to study the importance of vertical variations of latent heat release in generating available potential energy. The results of the computations (table 1) indicate positive generation for each stage of the storm. However, the generation during the mature and early occluded stage is nearly an order of magnitude greater than that in the cyclogenetic stage.

The small generation values for the cyclogenetic stage resulted from the conditions that precipitation tended to be slight and evenly distributed in the warm and cold air. This is evident on a comparison of the positive and negative generation areas presented in figures 5B, 5D, and 5F that illustrate the variations of the generation contribution for the three stages. Notice that the negative generation area is largest during the cyclogenetic stage and that the zone of maximum latent heat release in the northern portion of the precipitating region of the cyclogenetic stage (fig. 5B) caused little generation by virtue of its coincidence with the zero efficiency isoline. The maximum generation during the mature stage resulted from the large precipitation amounts in the cyclone's warm sector, while the slightly smaller generation in the occluded stage was associated with reduced precipitation amounts in the same region.

Estimates of the area-averaged frictional dissipation of the planetary boundary layer computed from eq (22) are presented in the last column of table 1. In the cyclogenetic and early occluded stages, the boundary layer dissipation exceeded the generation by the latent heat component. In the mature stage, generation exceeds dissipation. The increase of the generation by the latent heat component from approximately 1.0 W m^{-2} in the cyclogenetic stage to 8.0 W m^{-2} in the mature stage and the excess of the generation over the frictional dissipation of 5.1 W m^{-2} in the mature stage strongly suggests that an in situ source of energy exists within the storm that aids in its development. All of the estimates suggest that the generation by the latent heat component at the secondary scale is a significant fraction of the rate of kinetic energy dissipation within the storm.

VERTICAL DISTRIBUTION OF HEATING AND GENERATION

The latent heat component of the generation on an isentropic surface is largely due to the covariance between heating and the efficiency factor. The efficiency cross sections showed that the largest positive factors occur at high pressure in the subtropical air and the largest negative values occur at low pressure in the polar air. The decrease of efficiency factors with height in both polar and subtropical air indicates that the vertical variation of the heating is also important in the generation process. Vertical profiles were computed from the mass-weighted integral defined by

$$F(\theta) = \int \rho J_\theta f dA. \tag{24}$$

For f equal to Q_m and ϵQ_m, eq (24) provides profiles of latent heat release and the generation contribution as a function of potential temperature.

The physical interpretation of these profiles is difficult if one attempts to associate the heating or generation on an isentropic surface with geometric height. However, the profiles explain the relationship between isentropic structure, distribution of latent heat release, and generation.

In figure 6A, profiles of latent heat release for three stages of the cyclone are presented. These profiles were computed using the third model of latent heat release and a constant cloud top of 300 mb. The most interesting feature is the time variation of the relative maximum within the 300°–310°K layer associated with the latent heat release in the subtropical air. In the cyclogenetic stage, the maximum is somewhat lower than in the latter two stages. This difference is due to the distribution of precipitation about the storm. In the cyclogenetic stage, approximately equal amounts of precipitation fell within the warm and cold sectors. In the mature and occluded stages, a larger percentage of the total rainfall occurred in the warm sector, which contributed to latent heat release at higher potential temperatures. While profiles of heating with respect to pressure or geometric height would display a similar quadratic variation, it is doubtful

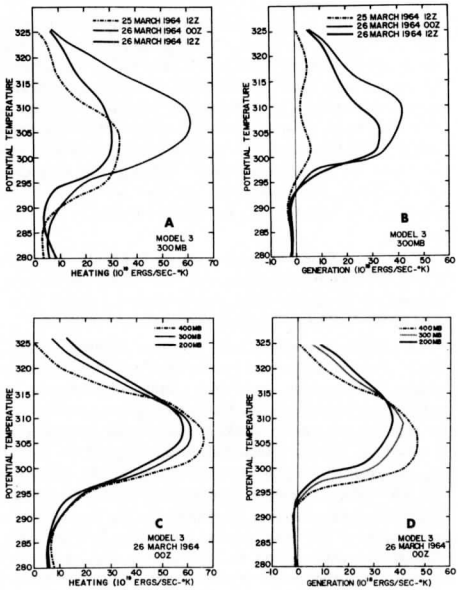


FIGURE 6.—Profiles of latent heat release and generation for the three stages from model 3 are presented in (A) and (B). Profiles of latent heat release and generation for the mature stage incorporating a variable cloud top in model 3 are presented in (C) and (D).

that the upward displacement of the relative maximum with respect to time would be detected.

The generation profiles in figure 6B confirm the importance of the vertical variation of precipitation with respect to potential temperature. While approximately the same amount of total heating occurred in the cyclogenetic and occluded stages (7.3×10^{21} ergs s^{-1} vs. 7.4×10^{21} ergs s^{-1}), the much smaller generation for the cyclogenetic stage resulted from the condition that more precipitation occurred in the negative efficiency region of the polar air. See the negative and positive generation areas on the 300°K isentropic surface for the two time periods in figures 5B and 5D. The maximum generation for the mature stage is due not only to a more favorable horizontal distribution of precipitation but also to the condition that the maximum latent heat release of 13.1×10^{21} ergs s^{-1} occurred for this stage.

In figure 6C, the effect of varying the cloud top parameter is shown for the mature stage. Lowering the cloud top increases the heating at lower levels but decreases it at higher levels. A similar effect is evident in the generation profiles in figure 6D. The results in table 1 show that the total generation increases slightly with a lower cloud top height. The assumption of a constant cloud top for a cyclone is, of course, an oversimplification. In reality, the cloud tops vary, and the latent heat release is distributed within different layers of the atmosphere in various positions of the storm. However, the rather slight dependency of the generation estimates to cloud top variations indicates that the generation is not critically dependent upon this parameter in this study.

This conclusion is reinforced by the rather close agreement in generation estimates between the three models of heat release for a 300-mb cloud top. Because the profiles of latent heat release for model 1 showed a very uniform distribution with potential temperature while those of models 2 and 3 were of a quadratic nature, the agreement in generation estimates must be due to the special nature of the vertical distribution of the efficiencies. Note in figure 1A that Q_m tends to be a constant value; and in figure 1B, Q_m tends to be quadratic. If in the precipitating subregion the product of ϵ and $\partial p/\partial\theta$ tends to be linear, the integrated product of Q_m, constant or quadratic function, with $\epsilon(\partial p/\partial\theta)$, a linear function, will tend toward the same value for all three models.

While one may conclude for the diagnostic modeling that the differences in this study between models and within models are relatively unimportant for estimates of the generation of available potential energy, the same conclusion is not valid for a prognostic model in isentropic coordinates. For an isentropic model, the heating function is the "vertical motion" with respect to that coordinate system. Thus, since heating determines the vertical exchange of mass, momentum, and energy in an isentropic model, the prognostic heating profile must converge to its corresponding profile in the atmosphere for accurate numerical weather prediction.

TOTAL GENERATION AND FRICTIONAL DISSIPATION AS A FUNCTION OF AREA

In the results presented thus far, a circular area with a 1700-km radius about the storm center was selected for computing the reference pressure. One naturally questions whether or not changes in area will cause significant variations in estimates of the storm's generation of available potential energy. To answer this question, table 2 presents a comparison of the generation by the latent heat component in the mature stage from the third model for smaller circular areas centered about the storm. A cloud top of 400 mb was arbitrarily selected for the comparison.

The results in table 2 show that the total generation estimates for the volume decreased 18 percent when the radius of the circular area was decreased from 1700 to 1100 km. In contrast, the total boundary layer frictional dissipation decreased about 38 percent for the same change in area. The 38-percent decrease in total dissipation is due to the reduced areal extent of the boundary layer. Two factors account for the decrease in total generation. First,

TABLE 2.—*Generation* ($\bar{G}$), *frictional dissipation* ($\bar{D}$), *total generation* ($\bar{G}A$), *and total boundary frictional dissipation* ($\bar{D}A$) *for various circular areas, A, about the mature stage of the storm. The cloud top is assumed to be 400 mb.*

Radius (10^3 km)	A (10^{12} m^2)	$\bar{G}$ (W m^{-2})	$\bar{D}$ (W m^{-2})	$\bar{G}A$ (W)	$\bar{D}A$ (W)
1.7	9.0	8.3	5.1	74.7	45.9
1.4	6.2	11.3	5.8	70.0	36.0
1.1	3.8	16.0	7.5	60.8	28.5

TABLE 3.—*Total heating* (10^{14} W) *and average generation* (W m^{-2}) *of available potential energy by latent heat release and infrared emission for the storm volume with a 1700-km radius. Estimates from latent heat release are by the third model with a cloud top of 200 mb. Infrared estimates are: test 1, uniform* (-1.4°C day^{-1}); *test 2, clear* (-1.6°C day^{-1}) *versus cloudy* (-1.2°C day^{-1}); *test 3 clear* (-1.8°C day^{-1}) *versus cloudy* (-1.0°C day^{-1}).

March	Time (GMT)	Total heating		Generation			
		Latent model 3	Infrared test 1	Latent model 3	Infrared Test 1	Infrared Test 2	Infrared Test 3
25	12	7.3	−9.9	0.7	0.9	1.1	1.3
26	00	13.1	−10.4	7.8	1.2	1.6	2.0
26	12	7.4	−10.9	5.9	1.1	1.7	2.3

the increased reference pressures for the smaller areas tended to increase the region of negative efficiency factors; and second, the reduced areas contained less total precipitation. Fortunately, the rather small variation in total generation estimates shows that selection of the area for computing the total generation by the storm is not critical.

One extremely interesting implication of these results is evident in the estimates of mean generation. Note in table 2 that the average generation increased from 8.3 to 16.0 W m^{-2}, a result which indicates that the energy generated is primarily available within the region where development and intensity of the storm is a maximum. Very likely at this time, the generation rate by this component exceeds the total kinetic energy dissipation within the 1100-km volume.

6. ESTIMATES OF THE GENERATION BY INFRARED EMISSION

While the diabatic process of latent heat release is of primary interest in this study, two simplified distributions for the divergence of infrared irradiance were also assumed to assess the possible importance of this process in generating the storm's available potential energy. For the first test, uniform heating throughout the volume of -1.4°C day^{-1} was assumed. The rate of specific heat addition was computed by

$$Q_m = c_p \left(\frac{\partial T}{\partial t}\right)_I \tag{25}$$

where $(\partial T/\partial t)_I$ is the instantaneous temperature change due to infrared emission. Average generation estimates for the storm volume with a 1700-km radius are presented under test 1 in table 3. Note that the average generation increased from the cyclogenetic to the mature stage and then decreased by a very small amount from the mature to the occluding stage. This trend in the estimates parallels those for latent heat release, although the magnitudes of the generation estimates by this model of infrared emission are substantially lower than those for latent heat release.

The importance of the covariant nature between a given heating distribution and the efficiency factors in generating available potential energy is evident in comparing the total heating and generation in table 3.

The order of magnitude of the total heating by the infrared and the latent heat components is the same for all stages. In the cyclogenetic stage, approximately the same generation resulted from these two processes. However, in the mature and occluding stages, the generation by latent heat release was substantially higher because the major areas of precipitation occurred in the subtropical air. The enhanced generation by latent heat release in the latter two stages is due primarily to the increase of the covariance of the efficiency and heating. For the case of uniform infrared cooling, the covariance of Q_m and $\epsilon(\partial p/\partial\theta)$ is zero, and the positive generation results because the mass-averaged efficiency on an isentropic surface is negative. If Lorenz' (1955) isobaric approximation had been utilized, the estimate of generation from the uniform cooling would be identically zero in contrast to the isentropic estimate of approximately 1 W m^{-2}.

Because extensive middle and upper level stratoform clouds substantially reduce the infrared cooling, two additional tests incorporating this effect were conducted. Within the volume about the storm, the clear and cloudy areas were determined from the synoptic data. Two different infrared cooling rates were then assumed to be representative. The heating rates for clear and cloudy regions for test 2 were -1.6°C day^{-1} versus -1.2°C day^{-1}; and for test 3, they were -1.8°C day^{-1} versus 1.0°C day^{-1}, respectively. Admittedly, this approach is crude, since the strong cooling at cloud tops was not incorporated in the model. Results of the generation computations are presented in table 3. The generation estimates from tests 2 and 3 are slightly greater than test 1. This increase is due to the covariance between Q_m and $\epsilon(\partial p/\partial\theta)$ because the average of the heating remained -1.4°C day^{-1}. However, from a comparison of the results, the generation by the latent heat component is the primary diabatic component that provides an in situ source of energy to partially offset the kinetic energy dissipation within this extratropical cyclone.

7. SUMMARY AND CONCLUSIONS

In this study, the generation of storm available poten-

tial energy (Johnson 1970) due to the process of latent heat release has been estimated for the cyclogenetic, a mature, and an early occluding stage of a mid-latitude cyclone. Preliminary results for the process of infrared cooling have also been presented. Three different models for the vertical distribution of latent heat release produced positive generation estimates for all stages of the storm.

Results from the third model, physically the most realistic, show that the generation is primarily dependent upon the distribution of precipitation about the cyclone. The largest generation by the latent heat component, 6–8 W m^{-2}, occurred in the latter two stages when the maximum latent heat release occurred in the cyclone's warm sector. In the cyclogenetic stage, the generation was about 1.0 W m^{-2}. Because efficiency factors for the storm tend to decrease with height, maximum condensation in the lower levels generates more available potential energy than when condensation is a maximum at higher levels.

Two simplified infrared cooling distributions produced generation estimates ranging from 1 to 2 W m^{-2}. While these results are tentative, they indicate that the process of infrared cooling contributes a smaller portion of the total generation for well-organized disturbances than for developing systems.

The fact that latent heat release produced positive generation values for all stages of the storm is significant. However, from the small generation in the developing stage, we conclude as Danard (1964) did that latent heat release alone cannot initiate the development of a storm but can provide an important source of energy for its maintenance after its circulation is established.

Of the estimates for the three stages, the generation by the latent heat component for the cyclogenetic stage is probably the most questionable. In the modeling, it was assumed that all condensed water vapor immediately precipitated. However, in a developing cyclone, it is possible that a significant percentage of the latent heat was released in the formation of the extensive cloud cover present in later stages of the cyclone. The diagnostic model of this study is incapable of estimating this contribution. Thus, one cannot conclusively state that the energy generated by the latent heat release in the incipient stage is not a factor in its early development.

The percentage error in the generation estimates in the mature and occluding stages of 8 and 6 W m^{-2} is likely less susceptible to errors from net condensation not precipitated. The magnitude of both estimates demonstrates the importance of latent heat release in well-organized synoptic systems. The time variations from 1 to 8 to 6 W m^{-2} for the three stages indicate that the in situ source of energy available for conversion to kinetic energy by the latent heat release is an important factor in the transient behavior of the storm. The estimates of boundary layer frictional dissipation by Lettau's (1961) model from the surface pressure gradient shows that dissipation increases from 3 to 5 to 9 W m^{-2} during the three stages. In contrast, the generation by the latent heat release and infrared emission reached its maximum intensity during the mature stage, being nearly twice the dissipation. It decreased by 2 W m^{-2} from the mature to the early occluding stage. From the condition that the storm's low-level moisture supply will continue to be reduced in the occluding process and more latent heat will be released in the cold core vortex of the occluded system, very likely the generation will continue to decrease. The results of this study indicate that the decrease in the generation precedes and likely exceeds the decrease of kinetic energy dissipation in the storm. Thus, the implication that the in situ energy generated by the latent heat release to offset kinetic energy dissipation may be an important factor in the transient behavior of the disturbance is extremely interesting.

The results of this study compare favorably with those obtained by Danard (1966). Using Lorenz' approximate equation, involving the covariance of isobaric deviations of temperature and heating, Danard obtained a generation of approximately 11 W m^{-2} for a region containing a rapidly developing extratropical cyclone. Because a large portion of the precipitation occurred in the warm sector of the cyclones used for both this and Danard's study, the agreement between results re-emphasizes the importance of the distribution of precipitation in generating available potential energy at the storm scale. The distribution of precipitation about the storm studied here is not typical of all extratropical cyclones. Many have relatively dry warm sectors with major precipitation areas north of the warm front. The generation by latent heat release within these systems is probably less. Further studies are needed to clarify the role of diabatic processes within extratropical cyclones.

The interaction energetically of the middle-latitude cyclone with the circumpolar vortex during the growth, maturation, and decay of the storm is an unsolved problem. The results of this investigation that the in situ generation of the storm is transient, at times being large and being sufficient to offset a major portion of the local frictional dissipation of the system, represents another step in the search for the energy sources of the extratropical cyclone.

ACKNOWLEDGMENTS

The research was jointly supported by the National Environmental Satellite Center, Environmental Science Services Administration, ESSA Grants WBG–52 and E–8–69(G), and the Atmospheric Sciences Section, National Science Foundation, NSF Grant GA–10997. Partial support was provided by the Wisconsin Alumni Research Foundation and the National Science Foundation in the form of a grant for computational funds and the use of computer facilities.

REFERENCES

Aubert, Eugene J., "On the Release of Latent Heat as a Factor in Large Scale Atmospheric Motions," *Journal of Meteorology*, Vol. 14, No. 6, Dec. **1957**, pp. 527–542.

Bradbury, Dorothy L., "An Experiment in Stability and Moisture Analysis," *Final Report*, Contract AF19(604)-1293, Department of Meteorology, University of Chicago, Feb. **1957**, 32 pp.

Danard, Maurice B., "On the Influence of Released Latent Heat on Cyclone Development," *Journal of Applied Meteorology*, Vol. 3, No. 1, Feb. **1964**, pp. 27–37.

Danard, Maurice B., "On the Contribution of Released Latent Heat to Changes in Available Potential Energy," *Journal of Applied Meteorology*, Vol. 5, No. 1, Feb. **1966**, pp. 81–84.

Dutton, John A., and Johnson, Donald R., "The Theory of Available Potential Energy and a Variational Approach to Atmospheric Energetics," *Advances in Geophysics*, Vol. 12, **1967**, pp. 333–436.

Eliassen, Arnt, "On the Vertical Circulation in Frontal Zones," *Geofysiske Publikasjoner*, Vol. 24, Oslo, **1962**, pp. 147–160.

Hahn, D., and Horn, Lyle, "The Generation of Availabie Potential Energy in a Mid-Latitude Cyclone," *Studies of Large Scale Atmospheric Energetics, Final Report*, ESSA Grant WBG-52, Department of Meteorology, The University of Wisconsin, Madison, June **1969**, pp. 1–56.

Johnson, Donald R., "The Available Potential Energy of Storms," *Journal of the Atmospheric Sciences*, Vol. 27, No. 5, Sept. **1970**, pp. 727–741.

Kung, Ernest C., "Climatology of Aerodynamic Roughness Parameter and Energy Dissipation in the Planetary Boundary Layer Over the Northern Hemisphere," *Annual Report*, Contract DA-36-039-AMC-00878, Department of Meteorology, The University of Wisconsin, Madison, Nov. **1963**, pp. 37–96.

Lettau, Heinz H., "Theoretical Wind Spirals in the Boundary Layer of a Barotropic Atmosphere," *Annual Report*, Contract No. DA-36-039-80282, Department of Meteorology, The University of Wisconsin, Madison, Aug. **1961**, pp. 115–170.

Lorenz, Edward N., "Available Potential Energy and the Maintenance of the General Circulation," *Tellus*, Vol. 7, No. 2, May **1955**, pp. 157–167.

Lorenz, Edward N., "The Nature and Theory of the General Circulation of the Atmosphere," *WMO Bulletin*, Vol. 16, No. 2, World Meteorological Organization, Geneva, Apr. **1967**, pp. 74–78.

Margules, Max, "Über die Energie der Stürme" (On the Energy of Storms), *Jahrbuch Zentralamt für Meteorologie und Geodynamik*, Vol. 40, Vienna, **1903**, 26 pp.

Petterssen, Sverre, "On the Influence of Heat Exchanges on Motion and Weather Systems," *Final Report*, Contract No. AF19(604)-2179, Department of Meteorology, The University of Chicago, Feb. **1960**, 18 pp.

Rao, Gandikota V., "On the Influences of Fields of Motion, Baroclinicity and Latent Heat Source on Frontogenesis," *Journal of Applied Meteorology*, Vol. 5, No. 4, Aug. **1966**, pp. 377–387.

Starr, Victor P., "What Constitutes Our New Outlook on the General Circulation?," *Journal of the Meteorological Society of Japan*, Ser. II, Vol. 36, No. 5, Oct. **1958**, pp. 167–173.

Wexler, Raymond, and Atlas, David, "Moisture Supply and Growth of Stratiform Precipitation," *Journal of Meteorology*, Vol. 15, No. 6, Dec. **1958**, pp. 531–538.

[*Received February 24, 1970; revised May 13, 1970*]

THE GENERATION OF AVAILABLE POTENTIAL ENERGY BY SENSIBLE HEATING: A CASE STUDY

Robert L. Gall
and
Donald R. Johnson

Abstract:

The importance of the sensible heat transfer from the ocean to the atmosphere within an east coast cyclone is studied using the concept of available potential energy. Both the generation of the storm's available potential energy by this diabatic component and the boundary layer frictional dissipation are estimated.

The sensible heat transfer through the interface is calculated by employing the bulk aerodynamic method. The ten meter wind speed is estimated from the surface geostrophic wind models of Rossby and Montgomery (1935) and Lettau (1961). For each model, both a variable wind dependent and a fixed roughness parameter are utilized. The sensible heat flux from the four different computations varied by 50 percent showing the sensitivity of the flux estimates to the modeling of the boundary layer winds.

From the flux calculations both upper and lower bound estimates of the storm generation by the sensible heating are determined. The upper bound estimate is made by adding all the thermal energy to the superadiabatic surface layer, a few millibars in vertical extent, while the lower bound is determined by distributing the energy uniformly within the dry adiabatic mixing layer extending from a few to several hundred millibars vertically. These bounds reflect the uncertainty in the nature of the energy transfer between the turbulent scale of the boundary layer and the quasi-horizontal scale of the storm. It is postulated that the true generation is bounded by these upper and lower estimates. For the small area of 5.4×10^2 m^2, the lower bound generation estimates

range from 0.7 to 1.9 watts m^{-2} at 00Z and 1.3 to 3.1 watts m^{-2} at 12Z while the upper bound ranges from 1.5 to 3.9 at 00Z and 2.6 to 6.1 at 12Z for the four different estimates of sensible heat transfer. The boundary layer frictional dissipation increases from 2.7 at 00Z to 7.1 watts m^{-2} at 12Z. The results indicate that sensible heat generation is significant in the creation of the cyclone's available potential energy, particularly during the early cyclogenetic stage.

List of Symbols

A	available potential energy of the storm volume
π	total potential energy of the storm volume
p	pressure
T	temperature
ρ	density
θ	potential temperature
θ_T	potential temperature at the top of the atmosphere
V_θ	volume of the region encircling the storm and extending from the earth's surface to the top of the atmosphere
J	Jacobian of the transformation from Cartesian to isentropic coordinates
ϵ	efficiency factor
Q_m	rate of heat addition per unit mass
F	sensible heat flux
F_s	upward sensible heat flux from the ocean
U_g	surface geostrophic wind speed
$U_a(z)$	wind speed at anemometer height z
$\theta_a(z)$	potential temperature of the air at anemometer height z

$O_s(0)$	sea surface potential temperature
$C_a(z)$	aerodynamic drag coefficient at anemometer height z
C_g	geostrophic drag coefficient
k	von Karman's constant
z_0	surface roughness parameter
α	angle between the geostrophic wind vector and the surface shear stress
D	dissipation of kinetic energy
p_t	pressure at top of the mixing layer
c_p	specific heat at constant pressure
c_v	specific heat at constant volume
R_d	gas constant for dry air
κ	R_d/c_p
g	gravity
f	Coriolis parameter
r	subscript to denote variable of the reference atmosphere

I. INTRODUCTION

In the wintertime, frequent and intense cyclogenesis occurs along the east coasts of the Asian and North American continents. In these regions large amounts of sensible heat are transferred to the atmosphere when very cold air, originating over the continents, moves eastward over warm oceanic currents. Although this energy transfer is thought to be a major factor in the cyclogenesis, its relative importance to the production of kinetic energy within the storm is an unsolved problem.

Studies of the last decade indicate that the inclusion of sensible heating at the earth's interface is important in predicting storm development. Reed (1958), using a graphical prediction model, found that the

addition of this diabatic component in the lower layers of the atmosphere improved forecasts of cyclogenesis in the Gulf of Alaska. Winston (1955) showed that although horizontal advection of absolute vorticity could partially account for the cyclogenesis in the Gulf of Alaska, substantial errors were present in adiabatic models. Petterssen, Bradbury, and Pedersen (1962) found that inclusion of sensible and latent heat transfer from the ocean was necessary to explain the thickness tendency fields in developing storms that formed off the east coast of the United States. Manabe (1958), Pyke (1964), Laevastu (1965), and others showed that sensible heating attained maximum values during and just preceding the rapid cyclogenesis in these regions.

These investigations, however, were not able to answer unequivocally the question of whether the energy that is supplied to the atmosphere from the ocean is an in situ energy source for the cyclogenetic process. The total potential energy of the atmosphere (gravitational plus internal) is increased by the heating, but there is no assurance that this process is important in increasing the kinetic energy in the cyclogenetic region. Some insight is provided by noting that the differential heating between the land and ocean causes a strong temperature gradient to develop along the coast. Within the framework of the theory of baroclinic instability (Charney (1947), Eady (1949), and others), the increasing horizontal temperature gradient and decreasing static stability by surface heating are important factors in the instability mechanism by which the conversion of potential to kinetic energy is insured.

An alternative approach based on energy principles is to investigate how the inclusion of sensible heating changes the total potential and kinetic energy of the cyclone. However, a diagnostic limitation of this method is that total potential energy changes calculated from data cannot be equated to kinetic energy changes. Errors in the estimates of the time derivatives of the potential energy are of the same order as the actual time derivatives of kinetic energy and equality for the changes of energy is not realized in empirical studies. In order to avoid this problem and others, Lorenz (1955) formulated the theory of available potential energy to explain and study the general circulation. Because the magnitude of available potential energy is closer to that of kinetic energy, errors in the estimates of the temporal derivatives are less than the actual kinetic energy derivatives. Thus the effects of measurement errors are reduced.

The equations that have been primarily used in available potential energy studies were only approximate. Dutton and Johnson (1967) pointed out that for some processes these approximations led to erroneous conclusions. For instance, the approximate expression for the generation of available energy depends on the covariance of heating and the

temperature deviation on an isobaric surface. With sensible heating to cold air at the earth's interface, a destruction of available energy is inferred. Dutton and Johnson (1967), however, emphasize that by using the precise expression, which depends on the product of the heating and the deviation of pressure from its average on an isentropic surface, the generation of available energy by sensible heating to cold air will usually be positive.

The equations of both Lorenz (1955) and Dutton and Johnson (1967) were developed for estimating changes of available potential energy of the entire atmosphere. Recently, Johnson (1970) derived the available potential energy of an open system and its rate of change for a limited domain encircling the storm center and extending over the entire vertical extent of the atmosphere. In this study, this framework for the open system is used to estimate the generation of available potential energy by sensible heating for an extratropical cyclone along the east coast. Because the cyclone is the central feature of the open system, the available potential energy and kinetic energy of the open system as well as their rate of change will be identified with those of the storm. The boundary layer frictional dissipation is also estimated in order to compare the energy source by sensible heating with the energy sink.

II. ENERGY EQUATIONS FOR THE AVAILABLE POTENTIAL ENERGY OF A STORM

A. The definition

The available potential energy after the definition of Margules (1903) is

$$A = \pi - \pi_r \tag{1}$$

where π is the total potential energy of a region encircling the storm and π_r is the total potential energy of the same region after an adiabatic redistribution to a hydrostatically stable horizontal stratification (called the reference state) in the absence of lateral boundary work and advection. From Johnson (1970), the time rate of change of the available potential energy of a volume surrounding the storm in symbolic form is

$$\frac{dA}{dt} = - C(A, K) + B(A) + W(A) + G(A) \tag{2}$$

where $C(A, K)$ is the conversion of the available potential energy to kinetic energy, $B(A)$ is the relative advection of energy into the storm's moving volume, $W(A)$ is the change of available potential energy from

boundary work and G(A) is the generation of available energy within the storm volume. The expression for the generation is

$$G = \int_{V_\theta} (1 - (p_r/p)^\kappa) \rho J Q_m \, dV_\theta \tag{3}$$

where the reference pressure defined by the mass distribution in the actual atmosphere is

$$p_r = g \int_{\theta}^{\theta_T} \overline{\rho J} \, d\theta \tag{4}$$

The overbar denotes a horizontal average over that portion of the potential temperature surface contained in the limited region. Without any approximations an alternate generation expression is

$$G = \int_m \epsilon \, Q_m \, dm \tag{5}$$

where m is the total mass of the open system and the efficiency factor is defined to be

$$\epsilon = (1 - (p_r/p)^\kappa) \tag{6}$$

The physical basis for the efficiency factor lies in the definition of the available potential energy of the open system. Normally, heating or cooling changes the total potential energy of both the natural and the reference atmospheres differently. The efficiency factor provides the weighting of the heating field to determine this difference, i.e., the generation of available potential energy. Heating (cooling) at pressures higher and cooling (heating) at pressures lower than the isentropic mean will result in generation (destruction) of the storm's available potential energy.

III. SENSIBLE HEATING

In theory, the generation of available potential energy expressed by (3) is valid for all scales of the continuum including those with hydrostatic imbalances and static instabilities (Dutton and Johnson, 1967). In application, however, the data density, the necessary employment of the hydrostatic assumption, and the approximate model of sensible heating preclude any claim for precision. Attention will be focused on assumptions which may invalidate conclusions, while approximations that may introduce 20 to 30 percent estimation errors are accepted as inevitable. The goal is to establish from the viewpoint of available potential energy whether sensible heating is an important process in the frequent cyclo-

genesis of these oceanic regions and not the precise magnitude of the generation which undoubtedly varies with each synoptic event.

The form of the generation shows that the diagnostics may be separated into two distinct aspects—the efficiency distribution and the sensible heat addition. The efficiency distribution, which is determined by the quasi-hydrostatic large scale structure of the storm, is presented later. The modeling of the sensible heat addition is presented in this section.

A. Sensible heat flux

For the turbulent sensible heat flux within the time and space scale of the cyclone, the ocean behaves as a strong and continuous heat source in comparison with energy added to the boundary layer over continental regions. Thus, only the sensible heat flux through the ocean's interface need be considered.

Figure 1 portrays a schematic showing three layers of the boundary region: the surface layer, the adiabatic layer, and the free troposphere. The surface layer, at most a few millibars in depth, is characterized by superadiabatic lapse rates. The adiabatic or mixing layer with its con-

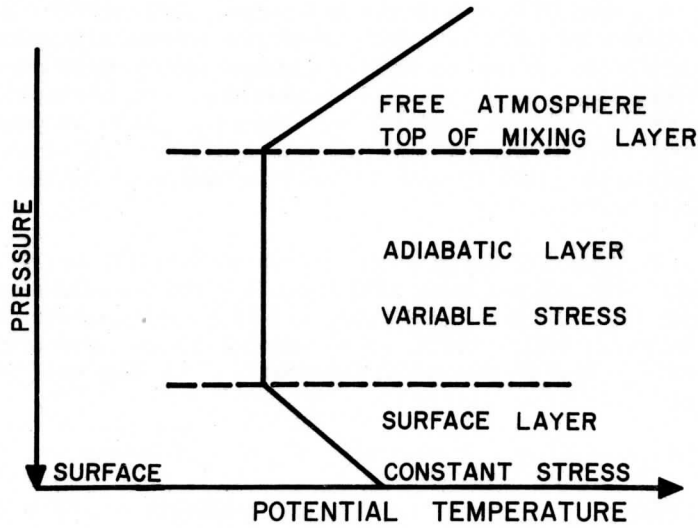


Fig. 1. Schematic distribution of potential temperature through the boundary layer.

stant potential temperature extends from a few to several hundred millibars in depth. Ideally, the top of the mixing layer is the lowest level for which the vertical sensible heat flux becomes negligible in comparison with its surface value. Deardorf (1966) emphasized that significant sensible heat flux in stable air only occurs if the lapse rate is nearly adiabatic. Consequently, the top of the mixing layer should coincide with the level where the stability increases. In general the top of the mixing layer does not coincide with the level for which the wind approaches geostrophic balance, i.e., the upper level normally defined for the planetary boundary layer.

A practical method for estimating the sensible heat flux is the bulk aerodynamic method given by

$$F_s = c_p C_a(z)[\theta_s(0) - \theta_a(z)]U_a(z) \tag{7}$$

Its utility and limitations are thoroughly discussed in several excellent summaries (e.g., Deacon and Webb, 1962; Roll, 1965). The choice of a height for the air temperature and wind speed is not extremely critical since the drag coefficient, $C_a(z)$, is adjusted for the appropriate height. In this study a commonly selected height of 10 meters is used.

The use of air temperature ordinarily observed from a ship's bridge height of less than 10 meters (from 6 to 9 meters, Roll (1965)) will produce a conservative estimate of the sea-air temperature difference for the model. Within the surface layer of constant stress where the upward sensible heat flux is large, potential temperature decreases rapidly with height. See the schematic distribution in Figure 1. An air temperature measurement taken in the region of six to nine meters produces an underestimate of the idealized 10 meter temperature difference and thus also the sensible heat flux.

In the derivation of the bulk aerodynamic method it is assumed that the sensible heat transfer coefficient is equal to the momentum transfer coefficient, i.e., the drag coefficient, in a neutral atmospheric surface layer (Deacon and Webb, 1962). Observational results for neutral conditions indicate that the functional dependence of the drag coefficient is of the form

$$C_a(10\text{ m}) = (a + b\,U_a(10\text{ m}))10^{-4};\quad 1 \leq U_a \leq 20\text{ m/sec}. \tag{8}$$

Sheppard (1958) suggests that the empirical constants a and b are 8.0 and 1.14 while Deacon and Webb (1962) suggest 11.0 and 0.7. However, Manier's (1962) empirical result from a least squares fit to the data of 9 different investigators that a is 7.9 and b is 0.96 was utilized primarily because of his interest in the heat balance for the North Atlantic

region (Manier and Möller, 1961). For the same region, Petterssen, Bradbury and Pedersen (1962) set a equal to 26.0 and b equal to zero. In this study, the maximum 10 meter wind speed was 27 m sec^{-1} while the predominate speeds were 12 m sec^{-1}. As a result, the magnitudes of the transfer coefficients were usually 10 to 20 percent less than the values suggested by Sheppard (1958) and Deacon and Webb (1962), and approximately equal to Petterssen et al. (1962). From (7), the flux estimates are similarly conservative by the same proportionality since the sensible heat transfer is directly proportional to the transfer coefficient's magnitude.

No attempt is made to incorporate the dependence of the sensible heat transfer coefficient upon bulk stability because of the limited vertical resolution of synoptic data. Garstang's (1967) analysis for low latitude sensible heat transfer within synoptic scale systems indicates that the effective transfer coefficient is double the neutral coefficient for the large negative lower limit of his bulk Richardson number equal to -0.325. Deardorff's (1968) recent theoretical analysis of the ratio of the diabatic transfer coefficient to its neutral value also provides support for an argument that the estimated transfer coefficients based on neutral theory are only 50 to 75 percent of their effective values in the regions of extreme free convection. Thus these factors indicate that the sensible heat flux estimates of this study will tend to be conservative.

In the bulk aerodynamic calculations, the ten meter wind speed was estimated from the surface geostrophic wind field by two methods, one from the model proposed by Rossby and Montgomery (1935), the other by Lettau (1961). In Rossby and Montgomery's formulation, hereafter called Model 1, the wind speed for the constant stress layer is

$$U_a(z) = \frac{3}{2}[(0.065\, U_g)k \sin\alpha \ln((z + z_0)/z_0)] \qquad (9)$$

where α is the angle between the geostrophic wind vector and the surface shear stress. In this study a constant α of 20° was used. In Lettau's (1961) formulation, hereafter denoted Model II, the wind speed for the constant stress layer is

$$U_a(z) = U_g(C_g/k) \ln((z + z_0)/z_0) \qquad (10)$$

where C_g, the geostrophic drag coefficient from Kung's (1963) empirical regression relation, is

$$C_g = .401\,[\ln(U_g/z_0 f) - 1.865]^{-1} \quad . \qquad (11)$$

At this point the only undetermined parameter is the surface roughness. Over the land, a constant value of 100 cm was assumed. Over the ocean, the state of the sea, i.e., its roughness, depends strongly on the wind

speed. Empirical data indicate that z_0 varies from 10^{-4} cm for light winds to 50 cm for winds in excess of 30 m/sec. For the oceanic regions, two calculations were made for each model, one with a constant z_0, the other with a variable z_0. For the constant z_0 calculations, a value of 1 cm was selected as it applies to moderate wind speeds of 10 - 15 m sec^{-1}. For the variable z_0 calculations, the surface roughness over the ocean was determined from the approximate relation for adiabatic theory (Manier, 1962; Roll, 1965) that

$$z_0 \approx 10^3 e^{-k/\sqrt{C_D}} \tag{12}$$

A first guess of one-half the geostrophic wind was used for the ten meter wind speed to estimate the drag coefficient by (8), then from (12) and either (10) or (11), the 10 meter wind estimate was revised using this estimate of the drag coefficient. Three such iterations were sufficient to achieve a stable estimate for the 10 meter wind speed.

B. Vertical distribution of sensible heat addition

After the sensible heat enters the atmosphere through the earth-air interface, it diffuses vertically. If this vertical transfer within the atmosphere were solely by molecular conduction and the motion were quasi-hydrostatic, this transfer would be confined to the surface layer. The generation equation (5) could be applied directly because the efficiency factor is essentially constant within this layer. However, the extreme heating of this layer produces large hydrostatic instabilities and free convection. This in combination with the mechanically forced turbulence within the boundary layer results in a continued upward vertical transfer of sensible heat throughout the boundary layer. The theoretical aspects of this scalewise heat transfer have been discussed by Priestley (1967).

From the aspect of the generation of available potential energy, distribution of all the thermal energy throughout the surface layer is justified if the frictional dissipation of the vertical kinetic energy in the boundary layer is negligible or small. A portion of the available potential energy generated within the surface layer manifests itself through hydrostatic instability as the vertical kinetic energy of free convection. If this convective kinetic energy is adiabatically converted back into potential energy by the negative buoyancy forces, an entropy conserving process, the entire amount of available potential energy generated within the surface layer must be realized at the hydrostatic synoptic scale. However, if the convective kinetic energy is frictionally dissipated into thermal energy, a local entropy increasing process that serves to diffuse the thermal energy within the mixed layer, not all the available potential energy generated can be realized at the synoptic scale.

Because the exact nature of the energy exchange between the micro-, meso- and macro-scales is not known, the generation estimates are computed for two distributions of heating. One generation estimate representing the upper bound is computed by confining the thermal energy to the surface layer; a lower bound is computed by assuming that the heat addition per unit mass is uniform within the mixing layer. For horizontal homogeneity and the hydrostatic approximation, the heat addition per unit mass is

$$Q_m = g\frac{\partial F}{\partial p} \tag{14}$$

where F is the turbulent heat flux. Thus for constant flux divergence, Q_m is independent of pressure. This distribution, possibly the most realistic, is consistent with the view that the turbulent redistribution serves to diffuse the internal energy and maintain the condition of $\theta \neq \theta(p)$ in the mixing layer. For a uniform translating mixing layer, the substantial derivative $d\theta/dt$ must remain independent of pressure, implying that Q_m is a constant except for its pressure dependency with $d\theta/dt$. However, this dependency remains slight due to the restricted vertical extent of the mixing layer with respect to pressure.

IV. THE CASE STUDY

In February, 1968, an intense upper tropospheric vortex existed over Eastern Canada for a duration of several weeks. In the lower troposphere a series of polar outbreaks occurred with cold air flowing southward from Canada over the Gulf Stream. South of Newfoundland several cyclones developed rapidly in the lower troposphere along the poleward boundary of the Gulf Stream. The strong sea surface temperature gradient in this region is shown in the North Atlantic portion of a hemispheric analysis obtained from the Naval Fleet Weather Central of Monterey (Figure 2). Of these cyclones, the storm that formed on February 14, 1968 intensified at a rate much faster than is usually observed, and since the sensible heat flux from the ocean to the atmosphere was extremely large, this storm was selected for this study.

A. The synoptic description

In Figures 3a through f the surface and 500 mb analyses portray the sequence of synoptic events for the 24 hour period from 0000Z on the 14th through 0000Z on the 15th of February. On February 13 at 0000Z (not shown) the upper tropospheric vortex was centered just northwest of Maine. A short wave with a strong vorticity maximum was embedded in the vortex just north of Lake Huron. Within the next 24 hours, the center

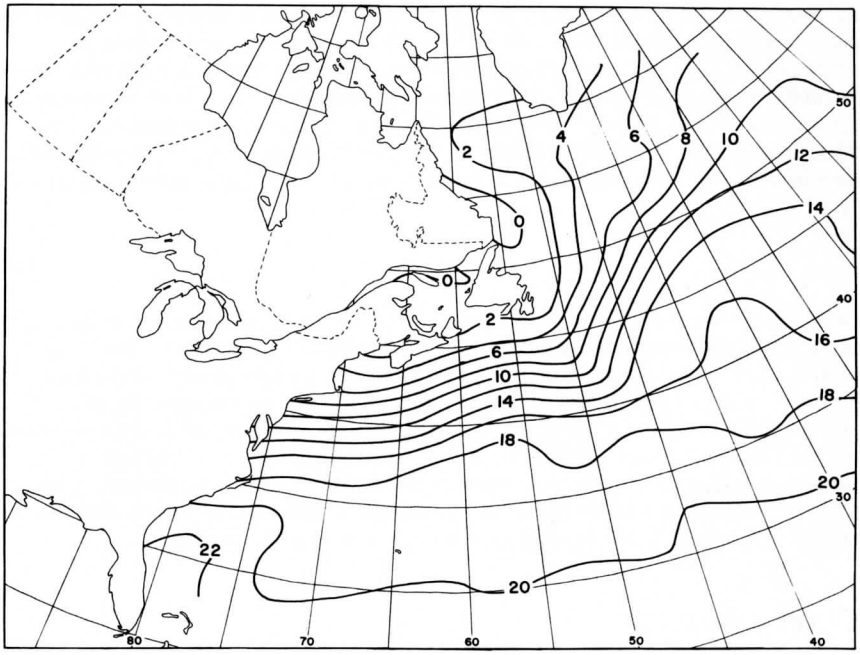


Fig. 2. Sea surface temperature (° C) for the North Atlantic Ocean, February 14, 1968.

of the vortex moved slowly southeast to the position over Maine shown in Figure 3b, while the vorticity maximum moved counterclockwise about the vortex to a position southeast of the vortex center. In the low troposphere, the short wave was accompanied by a surge of cold air whose leading edge at 1800Z, February 13 (not shown) was marked at the surface by a weak trough extending south-southwestward along the east coast of the United States. For the succeeding six hours, this trough advanced rapidly eastward and reinforced the cold front lying well offshore that was associated with the previous cold outbreak.

At 0000Z on the 14th and at a point just in advance of the 500 mb short wave trough the first evidence of cyclogenesis appeared along the old cold front about 100 nautical miles south of Sable Island. See Figure 3a. Ships in the vicinity reported pressure falls of 6 mb in three hours. The low's central pressure decreased at a rate of almost 2 mb per hour throughout the following 24 hours. By 1200Z (Figure 3c and d) the center was located 200 nautical miles south of Newfoundland, the central pres-

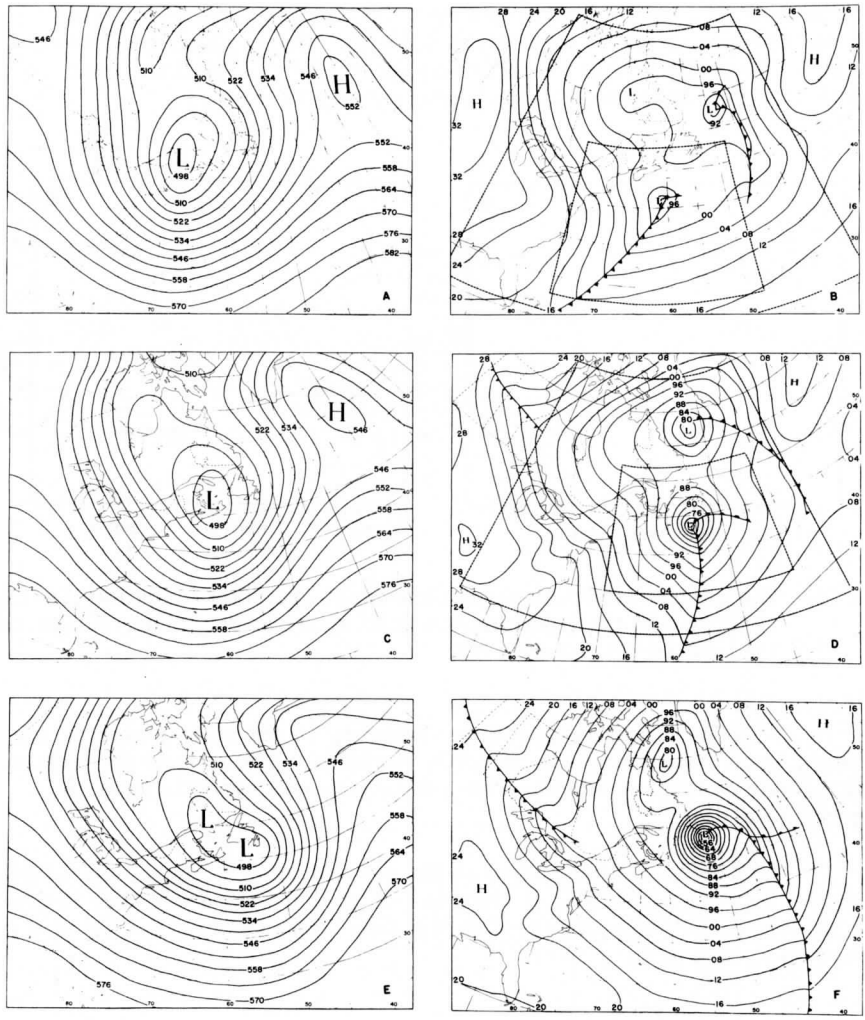


Fig. 3. The 500 mb height analysis (geopotential decameters) and surface analysis (millibars) for 0000Z, February 14 (A and B), 1200Z, February 14 (C and D), and 0000Z February 15 (E and F), 1968. The dashed rectangles in B and D delineate the large and small regions used to compute the generation of storm available potential energy.

sure had reached 967 mb and winds near the center exceeded 50 knots. The surface low attained its greatest intensity shortly after 0000Z February 15. The pressure and wind profiles for St. Johns, Newfoundland, in Figure 4b attest to the remarkable intensity of this storm.

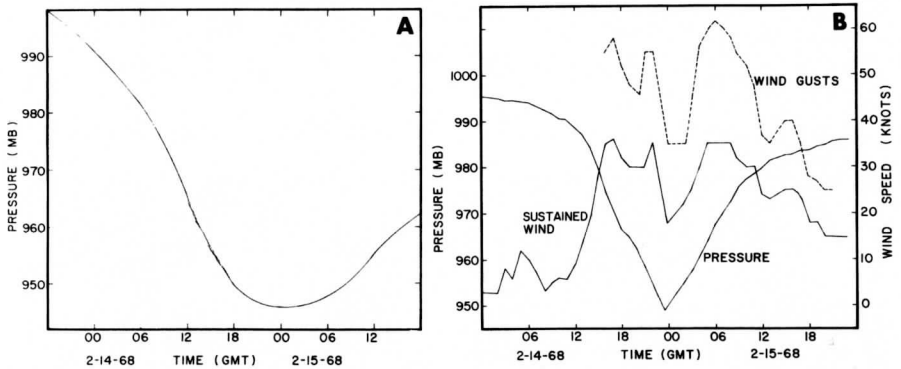


Fig. 4. The central pressure of the east coast cyclone of February 14-15, 1968 (A) and the pressure and wind speed at St. Johns, Newfoundland during its passage (B).

In the ESSA composite picture for February 14 shown in Figure 5, the cyclone already appears to be strongly occluded. The time of the picture corresponds to local noon which is midway between synoptic time of Figures 3c and e.

For the diagnostic computations the synoptic times, 0000Z and 1200Z, February 14, were selected. Figure 4a shows that the earlier time is near the beginning of cyclogenesis while the latter time is in the period of the most rapid storm development.

B. The analysis

The existence of a dry adiabatic lapse rate in the mixing layer means that potential temperature is no longer a single-valued function of height and the isentropic representation fails. Constant height or pressure analysis must be utilized in this layer. Above the mixing layer, the isentropic portrayal of atmospheric structure no longer fails. The depth of the mixing layer must be known for matching of the two representations.

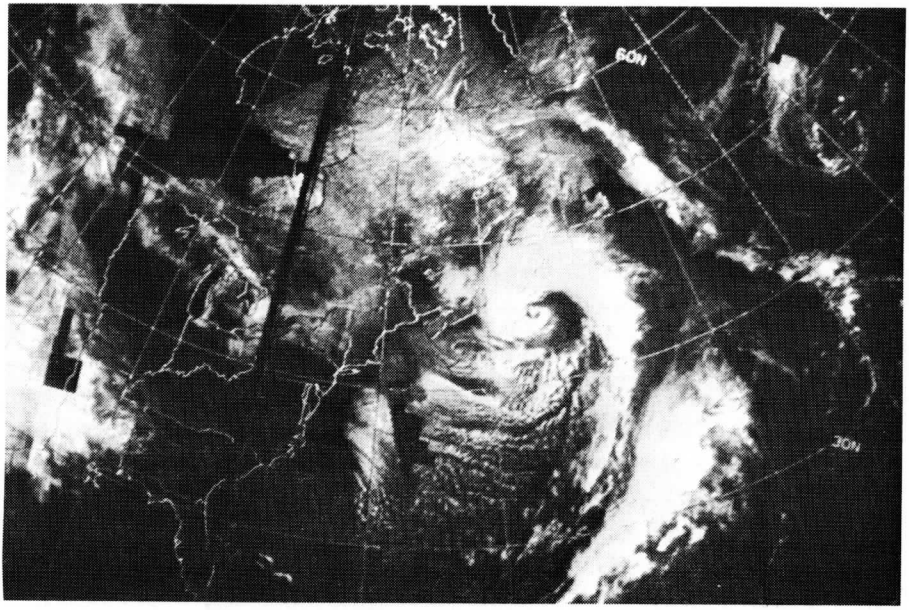


Fig. 5. Composite mosaic of satellite pictures from ESSA 3 for the North Atlantic Ocean for February 14, 1968.

Above the mixing layer isentropic pressure analyses for every five degrees from 250° to 320° were prepared. The potential temperature distribution of the mixing layer was specified by surface analyses of potential temperature. For the specification of the upper pressure of the mixing layer, approximate soundings were constructed for the oceanic regions from the isentropic analysis of the stable troposphere. The lapse rate of the lowest stable portion of these soundings was extended downward until it intersected the dry adiabat of the surface potential temperature. The intersection defined the upper pressure of the mixing layer.

This interpolation-extrapolation technique was first conducted for regions in close proximity to actual data and was extended laterally to sparse data regions. The pressure distribution of the upper boundary of the mixing layer that is presented in Figures 6b and d shows that its depth sometimes exceeds 200 mb. When these analyses are combined with the surface pressure (Figures 3a and c) and potential temperature analysis (Figure 6a and c) the mass-potential temperature distribution of the mixing layer is specified. Although this procedure is approximate, it is probably the best possible with the limited data.

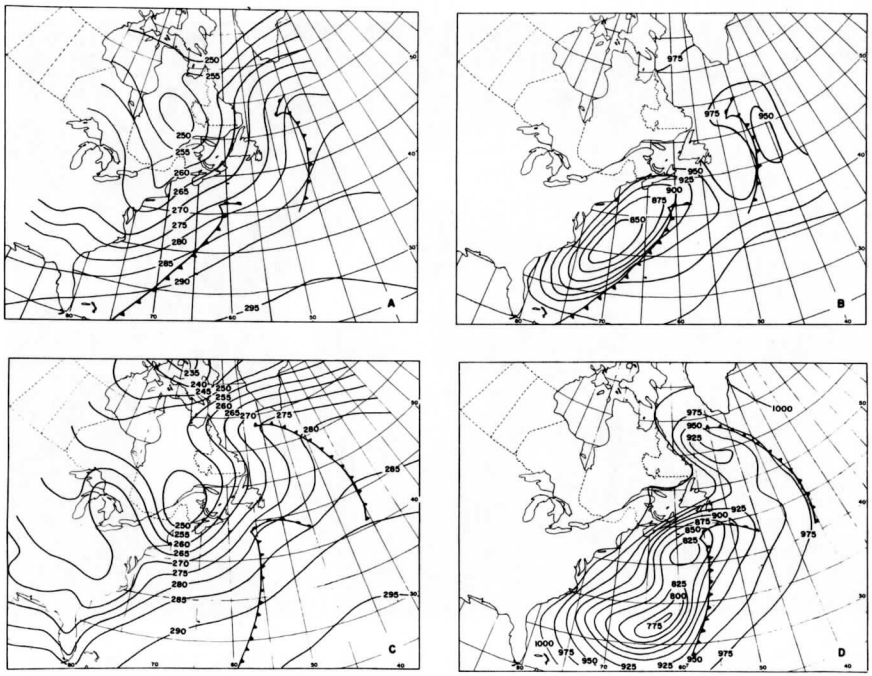


Fig. 6. Surface potential temperature (° K) distribution (A and C) and the pressure (mb) of the top of the mixing layer (B and D). Upper figures—0000Z and lower figures—1200Z, February 14, 1968.

C. The isentropic topography

Only six isentropic surfaces are presented in Figures 7a through f to portray the three-dimensional structure. The most striking feature at 1200Z on the 14th is the extremely cold air which appears as a mound from the earth's surface to the stratosphere. Near the surface, the center of the cold air is located northwest of Maine. Its axis tilts southeastward with height to a position just east of Maine on the 295° K surface. On the 275° K and 285° K surfaces, a ridge of cold air extends from the main center of cold air southeastward toward the surface pressure center located south of Newfoundland. The isentropic surfaces along this ridge slope from relatively high levels over the cold air, and descend steeply towards the ocean with the 275° K and 285° K surfaces becoming vertical in the mixing layer. The very steep gradient of pressure is evidence of the vortex's strong baroclinic zone associated with the short wave jet core that was important in initiating cyclogenesis.

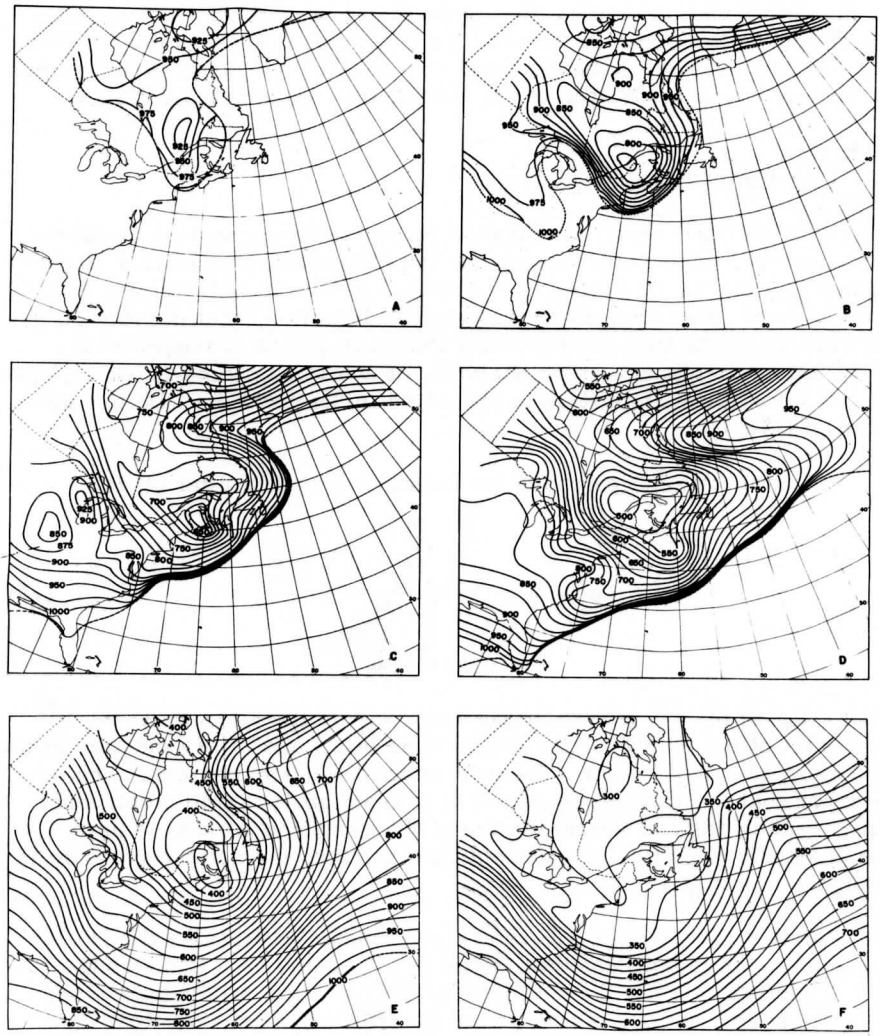


Fig. 7. Pressure analysis on the isentropic surfaces 255° K (A), 265° K (B), 275° K (C), 285° K (D), 295° K (E), and 305° K (F) for 1200 Z, February 14, 1968.

Over the Great Lakes modification due to sensible heating is evident on the 265° K surface (Figure 7b). The intersection of this surface with the ground partially circumscribes this region because the water is warmer than 265° K. To the west and south, the cold air has subsided and spread laterally into a thin sheet to cover most of the eastern United States.

D. The efficiency factor analysis

The efficiency factor distribution on an isentropic surface is determined solely by the pressure because the reference pressure is a single-valued function of the potential temperature. This is still true for adiabatic regions since even there the original definition, that the reference atmosphere is realized by preserving the mass-potential temperature distribution, is utilized. The analytics of available energy in atmospheres with local static instabilities is presented in Section 2.6 of Dutton and Johnson (1967).

For computational purposes the reference pressure after (4) is

$$p_r(\theta_k) = \frac{\sum_{j=1}^{N}\sum_{j=1}^{M} p_{ij}(\theta_k)A_{ij}}{\sum_{i=1}^{N}\sum_{j=1}^{M} A_{ij}} \qquad (15)$$

where θ_k is the k^{th} isentropic surface, p_{ij} is the pressure at the ij^{th} grid point of the $N \times M$ latitude-longitude grid, and A_{ij} is the area of the grid mesh at the ij^{th} point. The grid interval was two degrees. The restriction that the surface air potential temperature at a grid point is not equal to the potential temperature of a k^{th} isentropic surface is used. On the underground portion of the isentropic surface, the surface pressure is assigned (Lorenz, 1955).

Figure 8 presents the 12Z reference pressure profiles for the small and large regions delineated in Figure 3. The large area includes the 500 mb vortex; the small area includes the extratropical cyclone of the lower troposphere. The similarity of the profiles is evidence that the selection of the lateral extent of the volume is not a critical factor in this study.

Two cross sections of the efficiency distribution for 12Z, February 14, are presented to illustrate the regions where diabatic heating is important. The first cross section, Figure 9a, is orientated north-south along 56° W, transects the atmosphere just west of the storm and parallels the cold

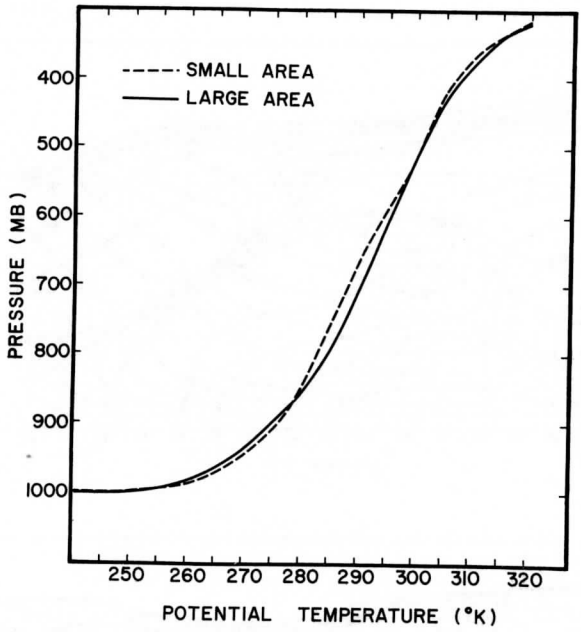
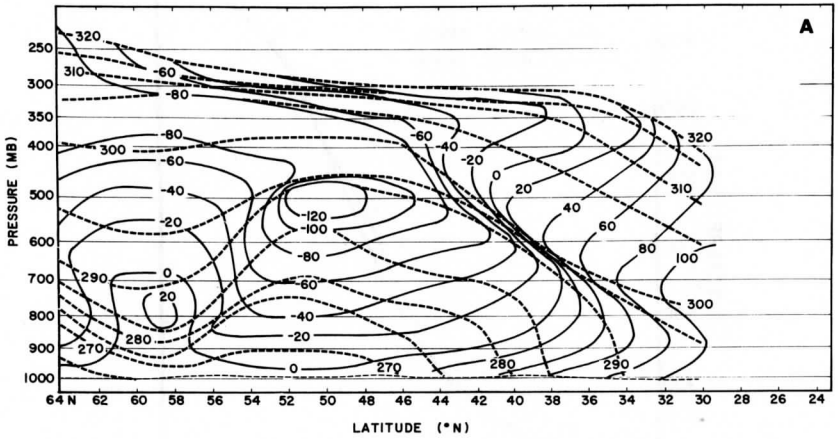


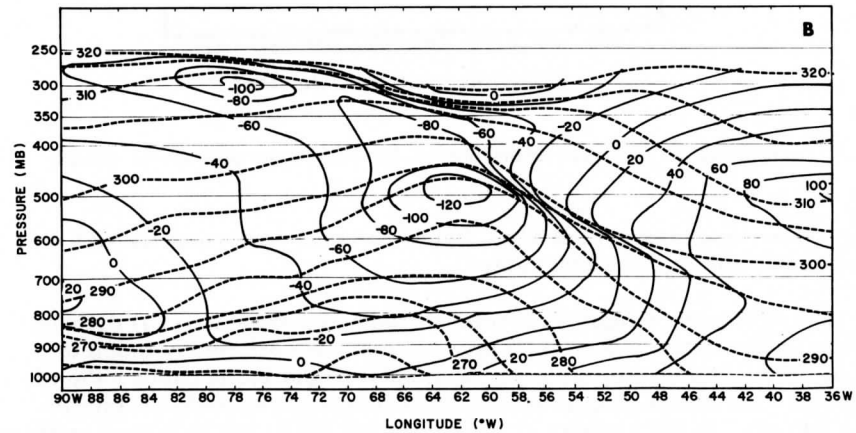
Fig. 8. Reference pressure distribution for the large and small areas.

front. The second cross section, Figure 9b, runs east-west through a point just south of the storm center.

The efficiency pattern is similar to Hahn and Horn's (1969) distribution for maturing cyclones. A nearly vertical zero efficiency isopleth exists in the middle troposphere. Southeast of the storm center, the efficiencies are positive up to relatively high altitudes, with a maximum near 700 mb. In the vicinity of the baroclinic zone, the zero line assumes an "S" shape with a positive region extending over the negative region of the colder air.

The 00Z and 12Z surface analyses presented in Figures 10a and c show that efficiencies are positive with one exception. In eastern Canada the cold air and lower-than-average surface pressures combine to give a limited region of small negative values. Along the east coast of the U. S., the efficiencies were small. The large positive efficiencies encountered in the southern Atlantic illustrate that for a given amount of sensible heating, the generation is largest when the energy is added to warmer air. Still, the maximum sensible heating occurs further north in colder air and regions of reduced efficiencies. Thus, maximum generation occurs where





and east-west along 24° N latitude (B), 1200Z, February 14, 1968. Solid and dashed isopleths are efficiency factor ($\times 10^{-3}$) and potential temperature (° K) respectively.

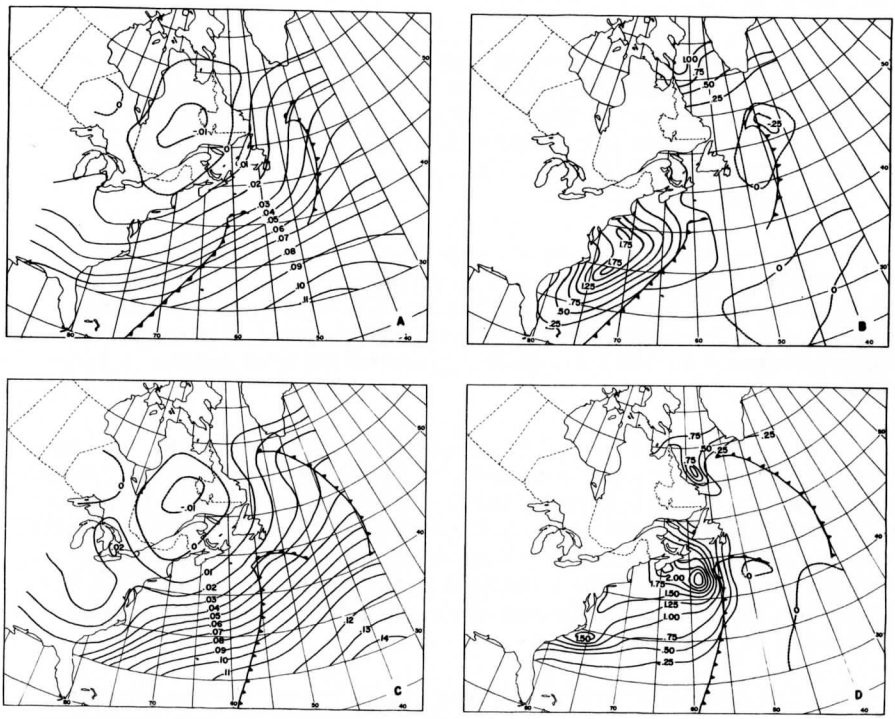


Fig. 10. Surface efficiency factors for the large area (A and C) and sensible heat flux (cal cm^{-2} sec^{-1} $\times 10^{-2}$) from the ocean (B and D). Upper figures—0000Z and lower figures—1200Z, February 14, 1968.

the turbulent transfer of energy is large and the efficiency is still significantly positive.

E. The sensible heat flux

The surface sensible heat fluxes determined from the bulk aerodynamic method and Model I with z_0 variable are presented in Figures 10b and d. Positive upward flux exists over the entire low pressure area. Even ahead of the cold front the flux to the relatively cold polar air from previous outbreaks is still positive. At 00Z, the greatest upward flux occurs along the U. S. coast. As the system develops, the area of maximum values translates with the storm center and a secondary maximum moves southward along the U. S. coast.

Exceptions to upward sensible heat flux occurs approximately 1000 kilometers in advance of the cold front (Figure 10b and d), and at 00Z in a region surrounding the northern cyclone. The modified polar air in these regions is caught in the return circulation of the developing vortex. In general, downward flux extrema are an order of magnitude less than upward flux maxima.

In the computations of the ten meters wind speed, estimates by Models I and II differed by two meters per second in areas of light geostrophic winds and by 4 meters per second for winds of 10 - 15 meters per second. In a comparison of predicted and reported ship winds Model I with z_0 variable consistently produced the best results in this study.

In the aerodynamic method with a variable drag coefficient, the dependence between sensible heat flux and wind speed is quadratic. Thus the differences between Model I and II wind speed predictions is amplified in estimating sensible heat flux. In Tables 1 and 2 the average heat transfer predicted by Model I with z_0 variable is at least twice that of Model II with z_0 fixed. The variations of the maximum flux in Table 3 are similar. The results illustrate the sensitivity of flux estimates to differences in modeling.

In a comparison of maximum fluxes (Table 4), our results compare favorably with Petterssen's et al. (1962) and Manabe's (1958) estimates. The difference between our estimate of 25.8×10^{-3} cal cm^{-2} sec^{-1} and Petterssen's et al., estimate of 16.7 cal cm^{-2} sec^{-1} is due to the quadratic dependence on wind speed in this study in contrast to the linear dependence utilized by Petterssen's et al. (1962). At a ten meter wind speed of 18 m sec^{-1} and an equal sea-air temperature difference, the flux predicted in both studies is nearly equal.

Table 1. Average sensible heat flux (cal cm^{-2} sec^{-1} $\times 10^{-3}$) for large area 14 February 1968

Hour of Observation	0000Z		1200Z	
Model I	I	II	I	II
Variable z_0	1.5	0.8	2.6	1.5
Constant z_0	0.9	0.6	1.7	1.1

Table 2. Average sensible heat flux (cal cm^{-2} sec^{-1} $\times 10^{-3}$) for small area 14 February 1968

Hour of Observation	0000Z		1200Z	
Model II				
Variable z_0	2.9	1.6	5.0	2.8
Constant z_0	1.8	1.2	3.5	2.1

Table 3. Maximum sensible heat flux (cal cm^{-2} sec^{-1} $\times 10^{-3}$) 14 February 1968

Hour of Observation	0000Z		1200Z	
Model	I	II	I	II
Variable z_0	18.5	10.2	25.8	15.1
Constant z_0	12.0	7.6	23.4	13.0

Table 4. Maximum sensible heat flux (cal cm^{-2} sec^{-1} $\times 10^{-3}$) found by other investigators

Investigator	Sensible heat flux
Petterssen, et al. (1962)	16.7
Pyke (1964)	2.5
Manabe (1958)	11.9
Gall and Johnson (1970)	25.8

Manabe's estimate of 11.9×10^{-3} cal cm^{-2} sec^{-1} is the maximum average sensible heat flux during a cold outbreak computed for a budget volume encompassing the entire Sea of Japan. Very likely the heating along the Siberian and Korean coasts just as the coldest air commences its over-water trajectory exceeded the average value by a factor of two or three. Pyke's (1964) estimate of 2.5×10^{-3} cal cm^{-2} sec^{-1} for weather ship P in the Gulf of Alaska during a period of strong northerly flow is for a region far south of the maximum heating over the Japanese current.

F. The generation

In the discussion of the vertical distribution of sensible heating it was argued that generation for the synoptic scale should be some intermediate value between upper and lower bounds. The maximum generation for the large and small areas (Tables 5 and 6) is from Model I with z_0 variable. In the large area upper bound estimates are 1.9 watts m^{-2} at 00Z and 2.8 watts m^{-2} at 12Z and the lower bound estimates are 1.0 and 1.6 watts m^{-2} respectively. The minimum lower bound estimates from Model II with z_0 constant are 0.7 watts m^{-2} at 00Z and 1.2 watts m^{-2} at 12Z. The variation between lower and upper bound estimates is only slightly larger than the variation between the various model estimates. Thus the uncertainty of the generation due to our lack of knowledge of the vertical exchange of energy is compounded by the uncertainty in the modeling of sensible heat flux.

The estimates for the small area (Table 6) are roughly double the estimates for the larger area (Table 5), and indicate the concentration of the diabatic effects within the storm circulation. At 12Z all the estimates for the small area are significantly positive. If the wind model with the most realistic predictions (Model I - z_0 variable) provides the most realistic sensible heat flux estimates, then the generation should be within the range of 1.9 and 3.9 watts m^{-2} for 00Z and 3.1 and 6.1 watts m^{-2} for 12Z.

The results of a test to determine if the depth of the mixing layer critically affects the computed generation are presented in Table 7. In this test the generation was estimated by assigning a uniform pressure of 900 mb for the top of the mixing layer. The comparison of the test and former results shows that in the large area generation estimates with a constant p_t are slightly lower. For the small area, this result is reversed. The use of a constant p_t of 900 mb in the small area effectively lowers the top of the mixing layer with the result that the energy is added at higher pressures.

Table 5. Average generation (watts m^{-2}) for the large area with a variable mixing layer depth. Upper and lower bounds are from uniform heating of the surface and mixing layers, respectively.

Time of Observation	00Z		12Z	
Model	I	II	I	II
Upper bound				
Variable z_0	1.9	1.0	2.8	1.5
Constant z_0	1.1	0.7	1.8	1.2
Lower bound				
Variable z_0	1.0	0.6	1.6	0.9
Constant z_0	0.6	0.4	1.0	0.6

Table 6. Average generation (watts m^{-2}) for the small area with a variable mixing layer depth.

Time of Observation	00Z		12Z	
Model	I	II	I	II
Upper bound				
Variable z_0	3.9	2.1	6.1	3.4
Constant z_0	2.4	1.5	4.2	2.6
Lower bound				
Variable z_0	1.9	1.0	3.1	1.7
Constant	1.1	0.7	2.0	1.3

The spatial distribution of the generation estimates from uniform heating within the surface and mixing layers are presented in Figures 11a through d. While these patterns are for the large volume, they are nearly identical to distributions based on the small volume because of their similar reference pressure distributions. An important feature when the energy is added to the surface layer is that nearly all the estimates are positive since the reference pressure is almost always less than the pressure of the top of the surface layer. In contrast, when the heat is added to the mixing layer there is a region of negative values west of the storm center (Figure 11d). Here the reference pressure is greater than the mean pressure of the mixing layer, thus, the local generation becomes negative.

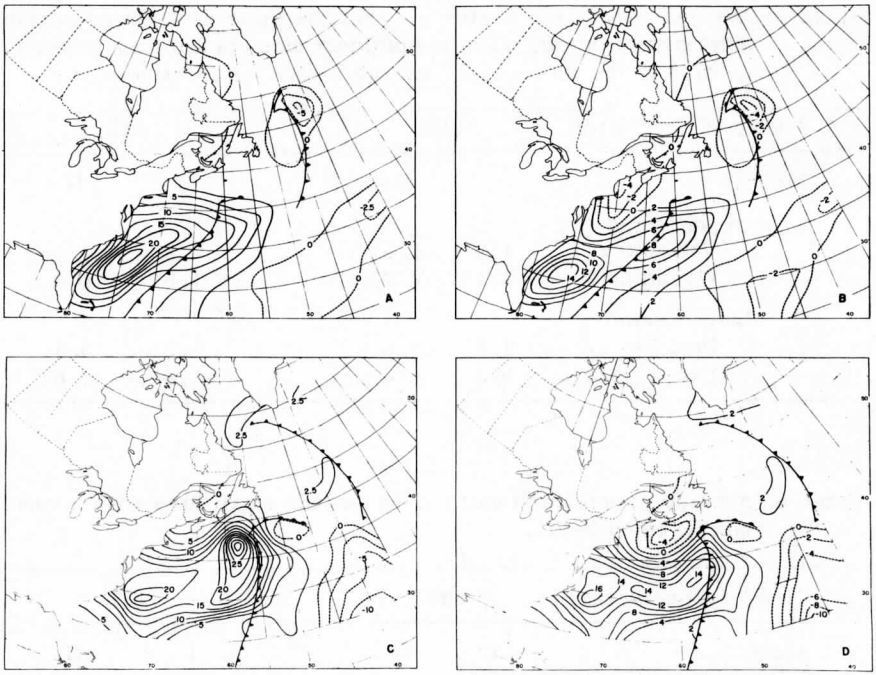


Fig. 11. Distribution of generation (watts m^{-2}) within the large area for uniform heating of the surface layer (A and C) and the mixing layer (B and D). Upper figures—0000Z and lower figures—1200Z, February 14, 1968.

Further from the dome of cold air, the reference pressure decreases until it is less than the mean pressure of the mixing layer and the generation becomes positive. This factor causes the maximum generation in Figures 11b and d to be displaced further southward than the corresponding maximum in Figures 11a and c.

If the approximate generation expression (Lorenz, 1955) which depends on the covariance of the isobaric temperature and heating deviation is utilized, the estimates, shown in Table 8, are negative. The standard atmosphere's static stability of γ equal to 6.5° C/km was used in these computations. Negative estimates such as these have led to the conclusion that sensible heating destroys eddy available potential energy by reducing the temperature contrast between air masses. Thus, the role of sensible heating in eddy generation has been erroneously interpreted by some investigators.

Table 7. Average generation (watts m^{-2}) with the top of the mixing layer assumed to be 900 mb. Variable z_0 was used in both models.

Time of Observation	00Z		12Z	
Model	I	II	I	II
Large area	0.9	0.5	1.2	0.7
Small area	1.9	1.0	3.2	

Table 8. Comparison between the approximate and exact generations estimates (watts m^2) by sensible heating for 12Z, February 14, 1968.

Area	Large	Small
Approximate	-1.1	-7.6
Exact/upper bound	2.8	6.1
Exact/lower bound	1.6	3.1

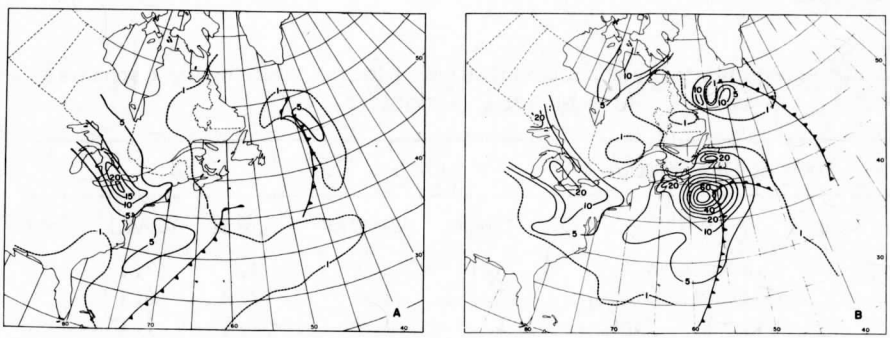


Fig. 12. Frictional dissipation of kinetic energy (watts m^{-2}) in the boundary layer for 0000Z (A) and 1200Z (B), February 14, 1968.

G. The comparison of dissipation and generation

The frictional dissipation of kinetic energy in the boundary layer was estimated using Lettau's (1961) formulation given by

$$D = \rho C_g^2 U_g^3 \cos \alpha$$

The departure angle between the geostrophic wind vector and the surface shear stress, α, was assumed to be 20°.

Table 9 presents the frictional dissipation using the land ocean distribution of the roughness parameter from Model I, variable z_0. Generation estimates are included for comparison. The results show that the storm generation by sensible heating constitutes a significant fraction of the frictional dissipation. At 00Z the average frictional dissipation for the large area exceeds the upper bound generation estimate. For the small area, the lower bound estimate is approximately equal to the dissipation while the upper bound is nearly double the dissipation. By 12Z the dissipation exceeds the upper bound estimate by the sensible heat component. The results indicate that in situ generation by sensible heating is important in maintaining the kinetic energy of the storm against frictional dissipation. In the relative sense, the generation by sensible heating is more important for the smaller intense storm scale than for the larger scale.

In the bulk aerodynamic computations the enhancement of sensible heat transfer in the super adiabatic surface layer was disregarded by not including a drag coefficient dependence on the bulk Richardson number. Thus, the upper and lower generation bounds are very likely conservative. Just how important the sensible heating is can only be determined with more precise modeling.

Table 9. Boundary layer frictional dissipation and generation estimates (watts m^{-2}) from Model I, variable z_0.

Time of Observation	00Z		12Z	
Area	Large	Small	Large	Small
Dissipation	2.1	2.2	4.7	7.1
Generation/upper	1.9	3.9	2.8	6.1
bounds /lower	1.0	1.9	1.6	3.1

Table 10. Total sensible heat flux, dissipation, and generation estimates 10^{13} watts) by Model I with variable z_0. Large and small areas are 1.67×10^{13} and 0.54×10^{13} m^2 .

Time of Observation	00Z		12Z	
Area	Large	Small	Large	Small
Sensible heat flux	124.0	79.0	184.0	113.0
Dissipation	4.0	1.4	7.6	3.8
Generation/upper	3.6	2.5	4.7	3.3
bounds /lower	2.0	1.2	2.7	1.7

The spatial distribution of the frictional dissipation is presented in Figures 12a and b. The most striking feature is the large dissipation of 70 watts/m^2 associated with the maximum winds located just to the rear of the vortex center (Figure 12b). This is higher than the dissipation reported by Petterssen et al. (1962). Their values for a composite cyclone which are effectively averages did not exceed 40 watts/m^2. However, because dissipation is proportional to the cube of the surface geostrophic wind, the wind speed associated with 70 watts/m^2 need be only slightly greater than one corresponding to 40 watts/m^2.

Elsewhere the dissipation did not exceed 10 watts/m^2 except near the Northern storm and over land. The higher dissipation over land is due to the greater roughness. Kung's (1963) average zonal values for the Northern American continent indicate that the value for z_0 of 100 cm over land is possibly too large and that the dissipation may be overestimated.

In Table 10 total sensible heat flux, generation and frictional dissipation are summarized. The generation appears to remain relatively constant during the cyclogenetic period. This behavior is accounted for by two off-setting factors. Figures 10a and c show that efficiency factors along the coast decrease with time as the center of the cold air mass moves toward the coast, a factor leading to decreased generation. During the same period sensible heating increases and tends to increase the generation. Eventually the translation of the polar air over the ocean will lead to its modification and the sensible heating and generation will decrease. Thus, the effect of sensible heating must be more important early in the storm's development than in later stages.

VI. REMARKS AND CONCLUSIONS

Although precise estimates of the generation by sensible heating cannot be stated from this study, the condition that sensible heat flux generates a significant amount of available potential energy during cyclone development is established. Evidence is provided that energy added to the atmosphere by molecular transfer across the earth's interface is actually a direct energy source to the quasi-horizontal scales of atmospheric motion. The extended horizontal domain of the molecular process primarily determined by large scale conditions was undoubtedly an important element in the forcing of the large scale cyclogenesis.

Estimation of generation is crucially dependent on the nature of the turbulent exchange of energy within the boundary layer. The upper bound estimates given by confining the thermal energy within the surface layer are nearly twice the lower bound estimates given by distributing the heating throughout the mixing layer. From the thermodynamic viewpoint and the exact energy equations, the upper bound estimate would be true if the buoyant mass element retained its identity and entropy in seeking its equilibrium level. Thus its turbulent vertical kinetic energy would be transformed back into available potential energy of the synoptic scale by negative buoyancy. However, because the mass element loses its identity through diffusion within the boundary layer the entropy must increase and the generation of available potential energy at the synoptic scale will tend toward the lower bound estimate.

From the upper and lower bounds estimated in this study, the computations indicate that the in situ generation by sensible heating is somewhere between 34 to 177 percent of the boundary layer frictional dissipation during the cyclogenic stage and from 18 to 86 percent during the mature stage.

Some importance for other diabatic processes in development may be inferred from Bullock and Johnson's (1971) results for the mature stage of a mid-latitude cyclone. For an area of 0.62×10^{13} m^2 over the continental U. S. Bullock and Johnson found that the storm generation by latent heating and infrared emission was 7.0×10^{13} watts and 0.7×10^{13} watts, respectively, a total of 7.7×10^{13} watts. If these estimates are adjusted to the small area of this study and added to the generation by sensible heating at 12Z, the total generation would be 8.5×10^{13} watts for the lower bound and 10.1×10^{13} watts for the upper bound. Both estimates exceed the boundary layer frictional dissipation by at least a factor of two. These results provide support for the hypothesis that the in situ generation within the storm scale is sufficient to offset its frictional dissipation in the formative and mature stages, and is important in the development process.

A complete theory for cyclogenesis along the east coast or for that matter anywhere, awaits a better understanding of the thermodynamic and mechanical interaction between all scales of the motion. The solution will involve mass and angular momentum as well as energy. It is undoubtedly the interaction of several processes that result in the complex sequence of events we call cyclogenesis.

ACKNOWLEDGMENTS

The authors wish to thank Professor Lyle Horn for his helpful suggestions, Mr. Daniel Dolson for his assistance in data preparation and Miss Sharon Nicholson and Mr. Si Chung Ok for drafting.

The research was jointly sponsored by the National Environmental Satellite Center, Environmental Science Services Administration, ESSA Grant E-8-69(G) and the Atmospheric Sciences Section, National Science Foundation, NSF Grant GA-10997. Partial support was provided by the National Aeronautics and Space Administration through a fellowship and the Wisconsin Alumni Research Foundation and the National Science Foundation in the form of a grant for computational funds and the use of computer facilities.

REFERENCES

Bullock, B. R., and D. R. Johnson, 1971: The generation of available potential energy by latent heat release in a mid-latitude cyclone. Mon. Wea. Rev., Vol. 99, No. 1.

Charney, J. G., 1947: The dynamics of long waves in a baroclinic current. Journ. of Met., Vol. 4, No. 5, 135-162.

Deacon, E. L. and E. K. Webb, 1962: Interchange properties between the sea and air. The Sea, Vol. 1, M. N. Hill, gen. ed., John Wiley and Sons, New York, 43-87.

Deardorf, J. W., 1966: The counter gradient heat flux in the lower atmosphere and in the laboratory. Journ. of the Atm. Sci., Vol. 23, 503-506.

Deardorf, J. W., 1968: Dependence of air-sea transfer coefficients on bulk stability. Journ. of Geophys. Res., Vol. 73, No. 8, 2549-2557.

Dutton, J. A. and D. R. Johnson, 1967: The theory of available potential energy and a variational approach to atmospheric energetics. Advances in Geophysics, Vol. 12, Academic Press, New York, 333-436.

Eady, E. T., 1949: Long waves and cyclone waves. Tellus, Vol. 1, No. 3, 33-52.

Garstang, M., 1967: Sensible and latent heat exchange in synoptic scale systems. Tellus, Vol. 19, No. 3, 492-508.

Hahn, D. G. and L. H. Horn, 1969: The generation of available potential energy in a mid-latitude cyclone. Final Rept. ESSA Grant WBG 52, Dept. of Meteorology, University of Wisconsin, Madison.

Johnson, D. R., 1970: The available potential energy of storms. Journ. of the Atm. Sci., Vol. 27, No. 5, pp. 727-741.

Kung, E. C., 1963: Climatology of aerodynamic roughness parameter in the planetary boundary layer over the northern hemisphere. Annual Report, Contract DA-36-039-AMC-00878, Department of Meteorology, University of Wisconsin.

Laevastu, L., 1965: Synoptic scale heat exchange and its relations to weather. Proceedings of the Air-Sea Interaction Conference, Tallahassee, Fla., 1965, Tech. note 9-SAIL-1.

Lettau, H. H., 1961: Theoretical wind spirals in the boundary layer of a barotropic atmosphere. Annual Report, Contract DA-36-039-80282, Department of Meteorology, University of Wisconsin, Madison.

Lorenz, E. N., 1955: Available potential energy and the maintenance of the general circulation. Tellus, Vol. 7, No. 2, 157-167.

Manier, Gerhard, 1962: Zur Berechnung der latenten und fühlbaren Wärmeströme von der Meeresoberfläche an die Luft. Geofisica Pura e Applicata, Milan, Vol. 52, pp. 189-213.

Manier, G. and F. Möller, 1961: Determination of heat balance at the boundary layer over the sea. Final Report, Contract AF 61(052)-315, Meteorologisch-Geophysikalisches Institut, Johannes Gutenberg-Universitat, Mainz, Germany.

Manabe, S., 1958: On the estimation of energy exchange between the Japan Sea and the atmosphere during the winter based upon the energy budget of both the atmosphere and the sea. Journ. Met. Soc. Japan, Vol. 36, 123-133.

Margules, M., 1903: Über die energie der stürme. Jahr. Kais. Zent. für Meteor., Vienna. In a collection of translations by Cleveland Abbe, 1910, "The Mechanics of the Earth's Atmosphere," 3rd collection, Smithsonian Misc. Coll., Vol. 51, No. 4, 1910.

Petterssen, S., D. L. Bradbury, and K. Pedersen, 1962: The Norwegian cyclone models in relation to heat and cold sources. Geophys. Publik., Vol. 24, No. 9, 243-280.

Priestley, C. H. B., 1967: Handover in scale of the fluxes of momentum, heat, etc., in the atmospheric boundary layer. In Boundary Layer and Turbulence, The Physics of Fluids Supplement 1967, 38-46.

Pyke, C. B., 1965: On the roll of air-sea interaction in the development of cyclones. Bull. Amer. Met. Soc., Vol. 46, No. 1, 4-15.

Reed, R. J., 1958: A graphical prediction model incorporating a form of non-adiabatic heating. Journ. of Met., Vol. 15, 1-8.

Roll, H. U., 1965: Physics of the Marine Atmosphere, Academic Press, New York.

Rossby, C. G., and R. B. Montgomery, 1935: The layer of frictional influence in wind and ocean currents. Papers in Physical Oceanography and Meteorology, Massachusetts Institute of Technology and Woods Hole Oceanographic Institute, Vol. 3, No. 3.

Sheppard, P. A., 1958: Transfer across the earth's surface and through the air above. Quart. Journ. Royal Met. Soc., Vol. 84, 205-224.

Winston, J. S., 1965: Physical aspects of rapid cyclogenesis in the Gulf of Alaska. Tellus, Vol. 7, 481-500.

Scanner's note:

This page is blank.

AN INVESTIGATION OF CLOUD DISTRIBUTION FROM SATELLITE INFRARED RADIATION DATA

Robert C. Lo
and
Donald R. Johnson

Abstract:

A physical-statistical model using NIMBUS II MRIR Channel 2 (10-11 microns) data is formed to estimate cloud cover. In the model, cloud cover is estimated from the distribution of DV2 values defined to be the differences between the earth surface temperatures and the corresponding Channel 2 observed equivalent black body temperatures. A preliminary test for four separate areas comparing model and corresponding estimates by four experienced meteorologists from HRIR photo-strip was conducted. The results from the model and the estimates by the meteorologists are in excellent agreement for regions in which there was a distinct temperature difference between the earth surface and the cloud top.

For cloud type estimation, Channel 1 (6.4 - 6.9 microns), and Channel 2 data are used simultaneously in a joint distribution model. Bivariate CV1 and DV2 frequency distributions portray distinct features associated with different cloud patterns from which it was possible to estimate the cloud coverage for different levels in the troposphere. No significant improvement in the estimation of cloud cover is attained by forming empirical bivariate frequency distributions from Channel 3 (14 - 16 microns) and either Channel 1 or Channel 2 data. The major limiting factor in estimating cloud coverage in this study was the marginal resolution of the MRIR radiometer. With higher resolution radiometers combined with irradiance information from other spectral intervals, the results indicate that a reliable global climatology may be found.

1. INTRODUCTION

One of the major problems in atmospheric heat budget studies is the specification of the global cloud distribution and the release of latent heat. Not only is heat realized by the condensation through convection, but the cloud distribution also strongly influences reflection and direct absorption of solar energy as well as the flux of infrared radiation. One major limitation of the energy budget studies conducted prior to the satellite era that was fully recognized by authors was the limited data available for oceanic areas, e.g., Houghton (1954), London (1957).

Beginning with the satellite era, irradiance data and cloud pictures are now available for oceanic areas and the opportunity exists for the development of a comprehensive global climatology of cloud distribution. For maximum utility, the global cloud climatology must include information concerning the vertical as well as the horizontal distribution of cloud cover and types. Such information would aid in the specification of the release of latent heat and the influence of clouds on radiative transfer (Manabe and Möller, 1961; Davis, 1963 and Katayama, 1966). Rasool (1964), Arking (1964), Winston and Taylor (1967) and others have together initiated such studies. However, several problems concerning the inference of cloud distribution from irradiance information remain. Among the problems are areal and spectral resolution of the instrument, variable atmospheric emissivity, and the reduction of the information to manageable and meaningful statistics.

The aim of this paper is to develop a model using satellite radiation data and statistical considerations to estimate the areal distribution of the cloud cover and to study techniques to infer cloud type and height. Channel 1 (6.4 - 6.9μ), 2 (10 - 11μ) and 3 (14 - 16μ) MRIR data from the NIMBUS II satellite are used in the estimation model to form bivariate and marginal frequency distributions from which the cloud distributions are inferred.

II. CONCEPTS, TECHNIQUES AND PROCEDURES

Of the infrared radiation emitted by the earth's surface, a significant portion of the energy is absorbed by the atmospheric constituents, water vapor, carbon dioxide and ozone. The absorption by these constituents is primarily confined to the region of the infrared spectrum with wavelengths less than 8 and greater than 12 microns. In the 8 to 12 micron region of the infrared spectrum, the window region, the atmosphere tends to be nearly transparent to the flux of the black-body radiation emitted by the earth's surface. Because of the selective absorption, the satellite observations of infrared radiation flux within different regions of the

infrared spectrum can be indicative of the content and distribution of the absorbing constituents.

The intensity of absorption and re-emission within the atmosphere depends on the absorptivity of the selective absorbing constituents. In spectral regions of strong absorption, the infrared energy emitted by the earth's surface is almost completely absorbed and re-emitted in the lower atmosphere. Very little energy passes directly into space. Since the absorption and re-emission process may effectively occur many times before the energy escapes to space, a troposphere with decreasing temperatures will give mean temperatures of re-emission that will be significantly less than the surface temperature. Thus the stronger the absorption, the lower will be the characteristic temperature of re-emitted energy sensed by the satellite. In spectral regions of strong absorption, the satellite observation will be indicative of the distribution of absorbing constituents within the upper tropospheric and stratospheric layers. However, in spectral regions of weak absorption, only a small fraction of the outgoing energy is absorbed and re-emitted. Most of the energy is transmitted through the atmosphere directly into space, and the satellite observations are indicative of the earth's surface temperature.

The MRIR radiometer on NIMBUS II was designed with four channels to measure the infrared flux in different spectral regions: Channel 1—the water vapor band from 6.4 - 6.9 microns, Channel 2—the window region from 10 - 11 microns, Channel 3—the carbon dioxide absorption band from 14 - 16 microns, and Channel 4—the broad band infrared region from 5 - 30 microns. Because of the strong selective absorption by water vapor in the 6.4 to 6.9 microns region, clear atmosphere Channel 1 observations are indicative of the water vapor content of the upper troposphere. In contrast, the Channel 2 observations provided good estimates of the surface temperature since a clear atmosphere is nearly transparent in the infrared region of the window channel.

In an atmosphere with liquid and solid forms of water, the atmosphere tends to be uniformly opaque to the entire infrared spectrum. Thus the energy received in the window channel is a measure of the energy emitted from cloud tops. When the clouds have extensive vertical development and are dense, the infrared radiative characteristics of the cloud top are nearly those of a black-body and the equivalent black-body temperature is an excellent estimate of the cloud top temperature, provided that the resolution of the satellite sensor is perfect. Since it is primarily only clouds that interfere with the infrared transfer through the atmosphere, Channel 2 data are either estimates of the surface temperatures under clear skies or cloud top temperatures under overcast conditions. In the following a model based on selective absorption within different spectral regions is developed to estimate cloud coverage in a finite area.

In the formation of an idealized model for a partly cloudy region, it is assumed: 1) that the clouds are of one type with a characteristic cloud top temperature, 2) that the atmosphere is perfectly transparent within the window region, 3) that the satellite sensor has perfect resolution, 4) that the tropospheric temperature decreases with height, 5) that the surface temperature is constant, and 6) that the clouds are effectively black with respect to the infrared flux. Under these assumptions, a histogram of Channel 2 equivalent black-body temperature from a partial cloud covered region would be described by two discrete lines. One line would represent the sample of observations of the irradiance emitted within the window region from the earth's surface; the other, the irradiance emitted from the upper surface of the clouds. See Figure 1 which portrays a hypothetical sample composed of an equal number of observations from a cloudy region with a cloud top temperature of 255° K and from a clear region with a surface temperature of 295° K. The existence of the temperature difference depends on the condition that the tropospheric temperature decreases with height since the intercepted irradiance from the earth or cloud top surface is proportional to the fourth power of temperature.

Under the additional assumptions that the observations are evenly distributed in the region of interest and that each observation is representative of the cloudiness in each small subregion, the ratio of the number of points for each peak to the total number of points becomes an estimate of the fractional area of either the cloudy or clear region. In the example portrayed in Figure 1, the fractional cloud cover for the area is the ratio 260 to 520, i.e., one-half.

In actuality, some of the assumptions made for the idealized cloud cover model are not satisfied. For certain conditions, the cloud thickness and its top height, thus its top temperature, will vary. Even in the window region the atmosphere is not perfectly transparent, and the outgoing infrared energy is affected, however slightly, by the variable absorbing constituents in the atmosphere. Thus in the model's application to real situations the Channel 2 histogram will not be represented by the two discrete lines when the assumptions are not true. Some insight into its expected form is given by a modified form of Central Limit Theorem (Cramer, 1945). The dominant physical process being measured is the irradiance received from the emitting surface. However, since the violation of the assumptions influence the observations of the physical process in a variety of ways, according to the Central Limit Theorem, the observations will tend to be normally distributed, and centered about a representative equivalent black-body temperature for the irradiance emitted by either the cloud top temperature or the earth's surface. The application of the Central Limit Theorem is valid, insofar as there is no single dominating effect from the violation of any assumption, combination or assumptions, or if the effects violating the assumptions are independent.

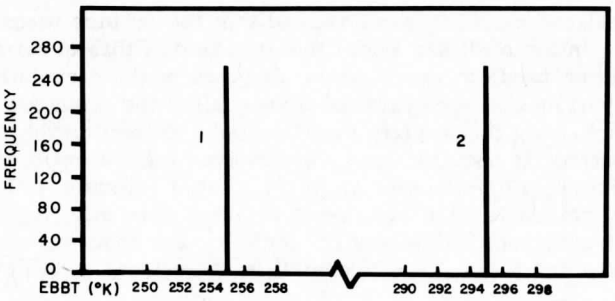


Fig. 1. Frequency histogram of equivalent black body temperatures for an idealized model.

Thus a more realistic model to estimate cloud cover in a given region is developed by noting that the assumptions stated for the idealized model need not be vigorously satisfied to obtain meaningful results. A more realistic histogram corresponding to Channel 2 observations of the same physical process portrayed in Figure 1 is constructed in Figure 2. Instead of two discrete lines, there are now two separate modes, more or less normally distributed, indicating the cloudy and the clear subregions. The fractional area of each subregion is given by the ratio of the area under each mode to the total area.

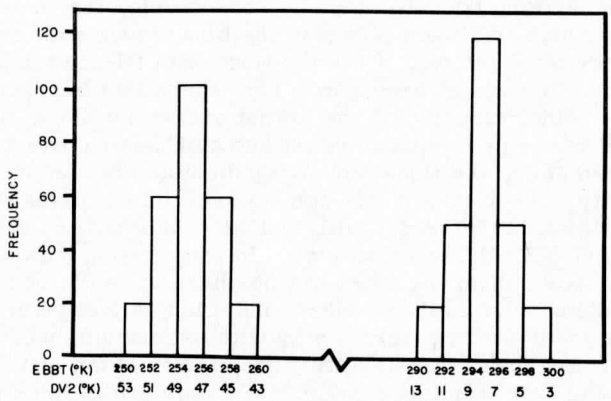


Fig. 2. Frequency histogram of equivalent black body temperatures for a realistic model.

In the idealized model it was assumed that the surface temperature was constant. In the realistic model the violation of this assumption in an area of interest tends to increase the variance of the distribution. This component of variance may be reduced by rescaling the abscissa from equivalent black-body temperature to a DV scale defined as the difference between the surface temperature and the observed equivalent black-body temperature (Shenk, 1963; Panofsky, et al., 1965). Thus a model formed in DV as the abscissa is more applicable to areas with differing surface temperature distributions. The new DV scale for the abscissa is presented in Figure 2 for which the surface temperature is assumed to be 303° K.

In situations in which more than two types of sky conditions exist in region of interest, more than two modes will be found in the actual distribution of Channel 2 observations. If the heights of the cloud tops tend to be distinct for the different layers of clouds, then separate modes for each cloud type will also tend to be distinct, i.e., if the resolution of the instrument is sufficient. In this situation the definition used in defining the fractional area covered by a single cloud layer can be extended to the fractional area covered by the i^{th} cloud layer by

$$C_i = \frac{N_i}{N} \tag{1}$$

where C_i is the cloud coverage of i^{th} layer in percent, N_i is the number of observations under the i^{th} mode and N is the total number of observations in the region of interest.

Besides the cloud coverage information, the histograms also reveal information concerning the cloud types. For example, the mode characterized by very high DV2's are observations from clouds with cold equivalent black-body temperatures, while the mode with the lowest DV2's will be observations from clear regions with high equivalent black-body temperatures. Since clouds with the lowest equivalent black-body temperatures will have large vertical extent and high water content in a liquid or frozen state, the mode probably indicates a region of cumulonimbus activity or deep large scale convection. Intermediate sky conditions will be indicated by modes with typical values ranging between the extreme DV2's for the two examples. In some cases, however, no definite information about their types or heights can be derived directly from the histograms due to the condition that cloud effective emissivities of the middle and upper troposphere vary with extremes from 0.16 to 1.0 (Kuhn, 1963). Because of the resulting uncertainty of the emissivity variation, the upward irradiance value from a region with thick warm clouds of the middle troposphere may be identical to that from a region with the thin cold clouds of the upper troposphere. In the thin cloud case, the infrared irradiance observed is partly contributed by flux from cold clouds and partly by the irradiance from the lower warmer regions. If one

assumed an effective emissivity of one, the cloud would seem to be from a region lower than its actual height in the troposphere. Thus in an area which has both low-emissivity valued clouds of the upper troposphere and thick high-emissivity valued clouds of the middle troposphere, Channel 2 histograms will not provide any discriminating information concerning either the type or the height of its upper surface. However, the histograms will still provide total cloud coverage information.

In the situation in which the cloud model using only Channel 2 observation fails to discriminate between middle and upper troposphere clouds, the most feasible alternative is to extend the modelling to single distribution histograms of Channel 1 and to joint distribution histograms of Channels 1 and 2. Channel 1, the 6 micron water vapor channel, is more sensitive to cloudiness and the water vapor in the upper troposphere (Raschke, 1966). Fritz and Rao (1967) have shown that the fractional transmission of the radiation through cirrus clouds is much larger for 10 micron radiation (Channel 2) than for 6 micron radiation (Channel 1). In one of the examples for an upper cloud temperature of 235° K, 60 percent of the upward irradiance was transmitted in the 10 micron region while in the 6 micron region the cloud was opaque. Thus, in histograms of Channel 1 observations, one would expect to find separate distinct modes associated with cirrus and middle level cloudiness so that in the situation where Channel 2 fails both cloud type and cloud top heights could be inferred. The single distribution model for Channel 1 observations is formed in the same manner as for Channel 2 model. Large DV1's indicate cirrus cloudiness while low valued DV1's indicate a dry middle and upper troposphere. Because of the strong absorption in the 6 micron region, only a negligible portion of the upward irradiance on the water vapor region that the satellite intercepts actually is emitted from the earth's surface. Again one expects that the actual distribution of Channel 1 observations will contain modes that according to the Central Limit Theorem will tend to be approximately normally distributed instead of discrete lines because of the many factors influencing the 6 micron irradiance.

Since both Channel 1 and Channel 2 histograms provide information of cloud type and Channel 1 is to be used in situations when Channel 2 fails, it is important to examine their joint distribution. In the actual construction of the two-dimensional arrays, the isopleths for local frequency maxima will delineate the models of the specific sky conditions. From statistical considerations, i.e., the Central Limit Theorem, these modes should be in the form of a bivariate joint normal distribution function. From physical considerations the center with the lowest DV1 and DV2 values should indicate clear sky conditions, while the center with the highest DV1 and DV2 values indicate regions of cumulonimbus activity or deep large scale convection. As for those centers with DV1 and DV2 values in between the two extremes, the sky conditions they indicate

may be quite complicated. However, it is reasonable to assume tentatively that in situations where two centers have approximately the same DV2 values but different DV1 values, according to the results of Fritz and Rao (1967), the mode with higher DV1 should indicate cirrus clouds while the mode with a lower DV1 should indicate lower clouds.

In the next section actual results from NIMBUS II MRIR data using both single and joint distribution models are presented.

III. RESULTS FROM CHANNELS 1 AND 2

Because the purpose of this paper is to develop a technique to estimate cloud distributions from radiation data, a preliminary study was conducted using only Channel 2 data to test the cloud cover model. For the preliminary work, four separate areas were selected. These four areas are designated in Table 1, and outlined in the HRIR photo-strip shown in Figure 4. Only the detailed results for the region B2 covering the area of the South Pacific from 105° W to 108° W and from 14° S and 23° S are presented to illustrate the preliminary tests. The MRIR Channel 2 data were taken from NIMBUS II orbit 213 of May 31, 1966. Sea surface temperature for the DV2 values were estimated from the May and June monthly mean temperature for the Southern Pacific (U. S. Navy Marine Climatic Atlas of the World, 1959).

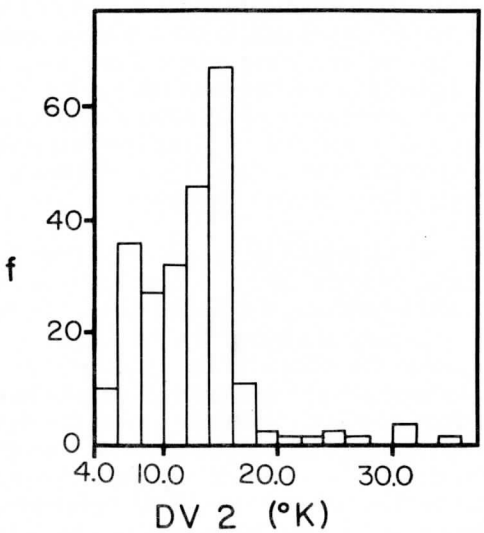


Fig. 3. Frequency histogram of DV2 data for block B2.

Table 1. Location, population, model and visual estimates of cloud cover for blocks B1 to B4.

Block No.	Latitudinal Extent	Longitudinal Extent	Channel 2 Model Estimates (%)	Individual Visual Estimates (%)				Averaged Visual Estimates (%)
B1	1S to 12S	102W to 105W	5.0	25	20	20	20	21.2
B2	14S to 23S	105W to 108W	69.5	40	80	75	75	67.4
B3	26S to 34S	109W to 113W	82.4	75	90	85	65	78.7
B4	40S to 43S	113W to 118W	67.2	55	75	80	55	66.3

Fig. 4. HRIR photo-strip for South Pacific. NIMBUS II data orbit 212, May 31, 1966.

In Figure 3 the histogram of Channel 2 radiation data is presented. Two modes are found, one is peaked at DV equal to 7 and the other is peaked at 15. For the region with DV2's greater than 20, the minor peaks that exist are not well defined due to the limited frequency of occurrence.

To estimate cloud coverage, the number of observations from the first three data intervals which corresponds to observations from clear regions were assigned to the lowest mode. Thus the fractional area of the clear region is 73/240 which is a percentage of 30.4% for clear conditions and 69.4% for cloudy conditions. In an independent check, the HRIR photostrip was given to four experienced meteorologists who were asked to estimate the cloud coverage of the area. They were not aware of the estimates from the MRIR data. Their individual estimates for the region B2 are presented in Table 1. It is interesting to note that while their estimates range from 40 to 80 percent, the mean estimate of 67.4% agrees remarkably well with the estimate of 69.5% from the model.

The cloud estimation results for the other three areas are summarized on Table 1 which presents the number of observations contained in each area as well as the results of visual and model estimates of percentage cloud cover. In comparing these results, the agreement between the cloud coverage estimates from the model and by the four meteorologists are excellent for Blocks 2, 3 and 4, while the poor results for B1 are likely due to the condition that the mode associated with the clear area dominates the mode associated with the low clouds. The two modes are not distinct, since the temperature difference between the ocean surface and the top of the low clouds is small. However, Young (1967) has noted a tendency for observers to overestimate cloud cover from satellite pictures for scattered conditions. Note in the HRIR photo shown in Figure 4 that the contrast between the clouds and ocean surface is slight. Because the HRIR contrast is a direct measure of the difference of the irradiance emitted from the two surfaces, it is clear that the clouds, regardless of their coverage, are warm low clouds. In these situations where there is little contrast in the upward irradiance from the earth's surface and warm low clouds, this technique will fail because of the low resolution of the MRIR radiometer.

The failure of this technique is due to the condition that the instrumental field of view is receiving the combined irradiance from both warm low cloud tops and the ocean surface. The field of view of the MRIR radiometer on NIMBUS II is a circle with a diameter of approximately 68 kms at the earth's surface. Thus the observed value is the average of the irradiance from a circular area that is much larger than the characteristic scale of small scale convective clouds. In the poor results for B1, the resolution of the MRIR radiometer is undoubtedly the primary reason for failure since in regions of variable low cloud cover, the observations due to averaging will completely smear the histogram of what should be two

distinct modes, i.e., if the resolution were perfect. If the cloud model were used with HRIR observations, the modes associated with various cloud features should be more distinct since the HRIR field of view, which spans a circular area with an 8 kms diameter, would provide much greater resolution. Also the reliability of the modes would be increased from the higher rate of sampling.

Although the preliminary study indicated certain limitations of the direct inference of the cloud distribution, the results indicated that large scale area cloud coverage could be inferred from the MRIR radiation data. Thus, the bivariate model using both Channel 1 and Channel 2 data was used for the same four regions of the preliminary tests. The first result to be discussed corresponds to the B2 case of the preliminary study just presented. In Figure 5 the empirical joint distribution of DV1 and DV2 data for the region B2 is portrayed. The isopleths indicate the number of observations. Also indicated on the figures are the marginal probability densities for each channel. Two relative maxima are found. One is

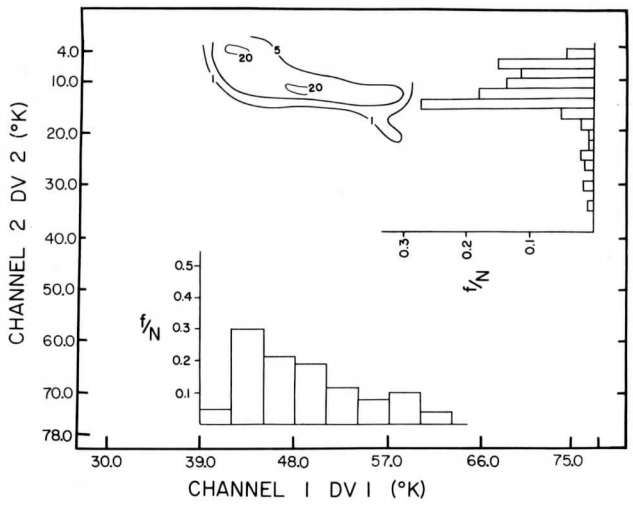


Fig. 5. Bivariate frequency and marginal probability density distributions of DV1 and DV2 data for region B2.

located at a DV1 equal to 43 and DV2 equal to 6, and the other at DV1 equal to 48 and DV2 equal to 14. The location of the upper left-hand maxima has the lowest possible DV2 and DV1 values of all centers from the scatter diagrams for the four regions. Thus it indicates not only a clear but also a dry atmosphere. It is interesting to note that the relative range of DV1 values is considerably less than for the DV2 data. This is due to the condition that even in a relatively dry troposphere, there is still sufficient water vapor for the irradiance in the 6 micron emitted at the earth's surface to be completely absorbed and re-emitted at a low temperature. Thus the black-body temperatures of Channel 1 are confined to a smaller range of tropospheric temperature than for Channel 2.

The mode in Figure 5 located at DV values of 48 and 14 indicates the distinct cloudy region that was estimated in the preliminary study to constitute 69.5% of the total B2 area. Note the DV2 marginal probability distribution along the right-hand edge of the figure portrays the two modes that, except for scaling, are identical to Figure 3. In the DV1 marginal distribution shown in the lower portion of the figure, there are no distinct modes. However, when both distributions are considered simultaneously, it is clear that the low value DV1's were primarily observed over the clear region while the DV1's greater than 49 were all observed over the cloudy region. In this study, no quantitative assessment of the amount of water vapor in the upper troposphere using Raschke's technique (1966) was made. Still, it is possible to infer from Raschke and Bandeen's (1967) statement, "that contributions (to satellite observation) from irradiance in the 6.3 micron region below 600 mb are small," and that the region of the troposphere above the low level clouds contained higher water vapor content than the clear regions. Furthermore, since the subregion about the cloudy maxima in the bivariate distribution is literally extended to the left indicating a great variation of DV1 values above the clouds, it is clear that variance of water vapor content above the clouds was greater than in the clear drier subregion.

The DV2 histogram also indicates a few higher valued observations which along with the high valued DV1 observations indicate a trace of clouds higher in the troposphere than the relative uniform layer covering 69% of the area. From the photo-strip, Figure 4, this condition appears to be true. The B2 area is partly clear, partly covered with warm low clouds, and in the middle of the region there is a trace of brighter clouds.

The bivariate frequency distribution for Block 1 showing only one mode is presented in Figure 6. Note that the one maxima in this Figure is centered at the same DV1 and DV2 values of the relative maxima indicating the clear condition in Figure 5. It also is the example in the preliminary study in which it was difficult to assess the area covered by low scattered clouds. Very likely the highest valued DV1 and DV2 data are indicative

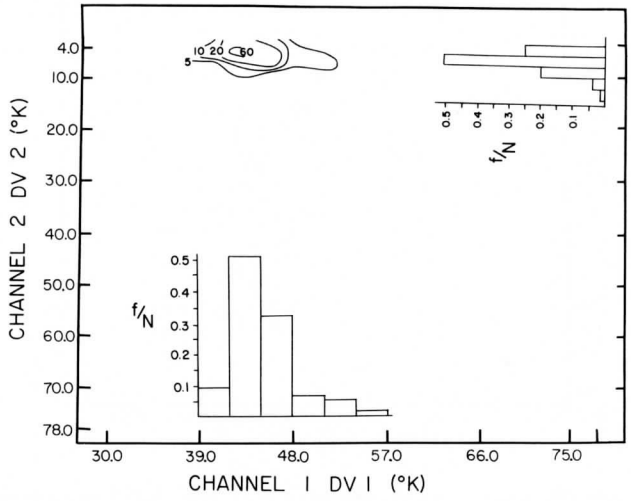


Fig. 6. Bivariate frequency and marginal probability density distribution of DV1 and DV2 data for region B1.

of the 10 to 20 percent low cloud cover indicated in the HRIR photo-strip. However, as discussed previously, there is no distinct mode associated with the clouds because of the limited resolution of the MRIR instrument, and it is difficult to estimate the cloud cover in this situation.

The bivariate frequency distribution of Channel 1 and 2 for B3 is presented in Figure 7. The elongated region containing nearly all the observations portrays a general correlation between DV1 and DV2 which is primarily due to the presence or absence of clouds. The relative modes II, III and VI deviate from the linear relation indicated in the diagram. Because the emissivity of cirrus clouds in the spectral region of Channel 1 is greater than Channel 2 (Fritz and Rao, 1967), relative modes indicative of these cirrus clouds should be displaced to the right of the general linear relation. Mode VI belongs in this category. The low DV2 values for Mode II indicates a clear region but one with higher water vapor content than for Mode I.

The results for Block 4 are presented in Figure 8 in which the three relative maxima are identified I, II and III. The DV1 density distribution indicates an upper troposphere with a uniformly higher water vapor content over the entire B4 area while the DV2 data indicates the three modes.

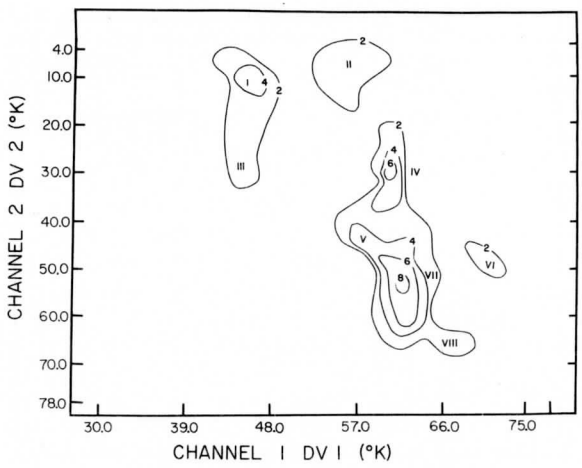


Fig. 7. Bivariate frequency distribution of DV1 and DV2 data for regions B3.

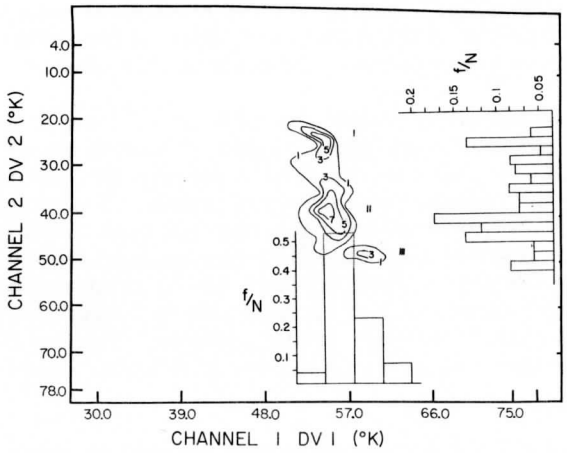


Fig. 8. Bivariate frequency and marginal probability density distribution of DV1 and DV2 data for region B4.

In the preliminary study the estimate of cloud cover was 67%. However, in a comparison of the DV2 values in this figure with the DV2 values for the clear regions of the previous three figures, one might conclude that the entire region was overcast. However, from the HRIR photo-strip there are darker areas which from a preliminary visual examination appear to be cloud-free areas.

There are at least two possibilities for this contradiction: 1) the contrast between the brightness of the highest clouds and the darker area is so great that the dark blotches appear to be cloud-free even though they are not or 2) the lack of resolution in the MRIR data does not allow for representative irradiance sampling from the dark areas. The latter possibility is indicated in the HRIR photo. Since the dark areas are of such limited horizontal extent (maximum = 1.5 degree of latitude), it is likely that the averaging of the irradiance from the limited MRIR resolution is the best explanation for the high DV values associated with Mode I. Regardless whether the dark blotches are cloud-free or not, one can conclude that the cloudiness of the highest layer is 67%.

Figure 9 is a composite bivariate frequency distribution for all four regions. Seven relative maxima are identified by roman numerals I through VII. One of the most interesting aspects of the combined figure is the definite organization of the field. Distinct maxima stand out in the two dimensional portrayal while the local maxima of the medium to high valued

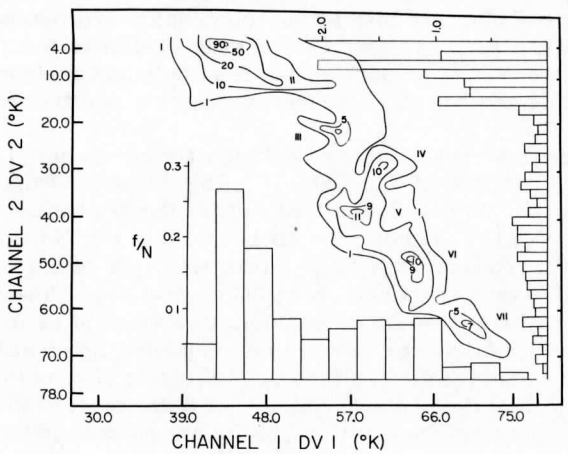


Fig. 9. Combined bivariate frequency and marginal probability density distributions of DV1 and DV2 data for regions B1, B2, B3 and B4.

portion of the marginal DV1 and DV2 distributions are much less pronounced. Thus, if one considers only the distributions separately, one loses information concerning the cloud and water vapor structure of the atmosphere. For example, consider that the observations associated with modes III, IV, V and VI were combined. Their marginal DV1 distribution would not be significantly different from the relative uniform DV1 distribution in the interval from 51 to 66 indicated in Figure 8. However, when the two dimensional array is considered, it is clear that distinct cloud layers were present. Thus, it is important to form a cloud or irradiance model which preserves this important information in any extended general circulation climatology.

The most distinct maximum, I, is associated with the clear conditions. Note that local maxima are also present in both marginal distributions. Thus, one can conclude that clear regions are associated with lower values of middle and upper tropospheric water vapor content than cloudy areas.

The highest DV valued maximum, VII, is undoubtedly associated with the most intense convection in the cloudy regions. This mode is from region B3 data and is associated with the highest clouds in the extreme upper right-hand corner (see Figure 4). Other relative maxima can be identified with cloud features in the various blocks and indicate that various synoptic systems have identifiable irradiance fields which are distinct from each other. Another interesting feature of Figure 9 is the indication of a strong positive correlation between the data of the two channels. However, for many purposes it would be misleading to attempt to infer one from the other by correlation techniques. For example, if a certain portion of the earth yielded a subsample similar to the subsample about modes IV and V, the hypothetical results would indicate a negative correlation while the complete sample indicates a positive correlation.

The percentage area coverage of distinct cloud patterns associated with each mode is presented in Table 2. The division of the distribution according to mode is not difficult except for the separation of modes I and II. Since most of the variance about strongly peaked mode I is likely associated with scattered low level cloudiness, the region of the distribution for DV1 greater than 46 and DV2 greater than 10 was combined with mode II. While it would be possible to speculate as to the type and height of cloud for all modes, there is no verifying information available. Hence, it seems advisable to attempt cloud identification in a later study when more complete data is available. The data indicate in the bivariate frequency distributions that distinct modes are present which allow quantitative estimates of percentage cloud coverage and that the possibility of the quantitative identification of cloud type and cloud top height to form a global cloud climatology deserves additional study.

Table 2. Percentage area coverage by relative mode

Relative Mode	Observational Frequency	Percentage Coverage (%)
I	377	46.7
II	183	22.7
III	33	4.1
IV	25	3.1
V	99	12.3
VI	63	7.8
VII	27	3.3
Total	807	100.0

IV. ADDITIONAL RESULTS FROM CHANNELS 3 and 4

The additional channels in the MRIR experiment provide the opportunity to construct multidimensional distributions. However, due to the preliminary nature of this study only two-dimensional histograms were studied.

Channel 4 measurements which are indicative of the total infrared irradiance emitted by the earth atmospheric system are highly correlated with Channel 2. Because Channel 2 is more sensitive to clear versus overcast conditions, it does not seem worthwhile to attempt bivariate distributions utilizing Channel 4.

Due to the nearly constant distribution of CO_2 in the atmosphere, Channel 3 measurements of the irradiance emitted in the CO_2 region are an indication of the effective mean temperature of the atmosphere from around 700 mb to one mb (Raschke, 1966). Because of the strong absorption in the 14 to 16 micron band, however, the observations primarily measure the temperature of the lower stratosphere and are not directly related to the detection of clouds.

Single mode distributions of DV3 were determined for the same areas utilized in the DV1 and DV2 histograms. No additional information concerning the cloud distributions were found. In the four histograms, the center of the modes shifted towards lower values with the increasing latitude of the region due to the condition that the decrease of stratospheric

temperatures with latitude is less than the decrease of the surface temperature. The scaled DV values did not adequately compensate for this effect because they are only adjusted for surface temperature.

Bivariate frequency distributions of DV1 and DV3 and also DV2 and DV3 were constructed. The DV1 - DV3 distribution showed that the two channels tended to be linearly correlated. In the distributions for B1, B2 and B4, only one mode was found and the diagrams are not presented. The distribution for region B3 portrayed in Figure 10 displays two modes. The left one indicates a dry upper troposphere and warm stratosphere while the right one indicates a moist upper troposphere and cold stratosphere. Very likely at this time the polar-subtropical break in the tropopause is located in B3.

The DV2 - DV3 bivariate distribution was similar to the DV1 - DV2 histogram, only the latter revealed more information about cloud types. Figure 11 shows the DV2 - DV3 distribution for B3. In a comparison with Figure 9, the first four modes along the DV1 scale are distinct but are not indicated by the DV3 scale. Thus for cloud type information, Channel 1 data seems preferable to Channel 3 information.

V. CONCLUSIONS AND SUGGESTIONS FOR FUTURE RESEARCH

The cloud cover estimation model developed and tested in this study shows considerable value in being able to quantitatively estimate cloud cover from satellite irradiance data. It will fail for situations in which the temperature difference between the earth's surface and the equivalent black-body temperature of the cloud top surface is not significantly large. However, these situations are normally only associated with low clouds, thin cirrus or for clouds over the extremely cold ice and snow covered polar regions. Over most regions in which there is considerable contrast in the upward irradiance from the earth's surface and the clouds, the technique using MRIR data shows considerable promise. With a higher resolution instrument, the relative modes in the bivariate distributions should be more pronounced with less variance, making the estimation more precise.

In future research for the development of a global cloud climatology, it is important to estimate the cloud type and heights. This should be possible with a higher resolution instrument since the variance of the relative modes would provide information on the uniformity of the irradiance from the upper cloud surface. Such information should identify whether the cloud is stratiform or cumuloform.

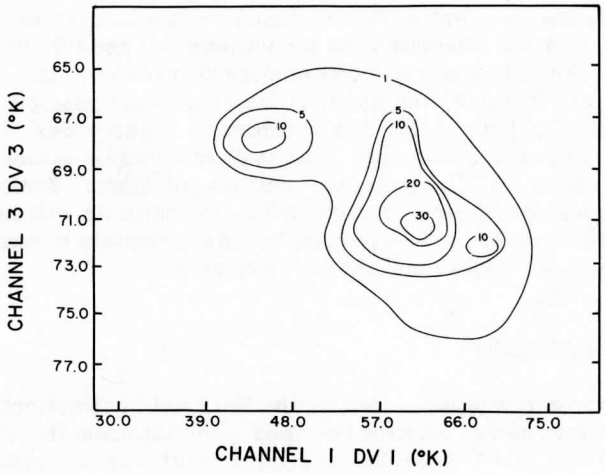


Fig. 10. Bivariate frequency distribution of DV1 and DV3 data for region B3.

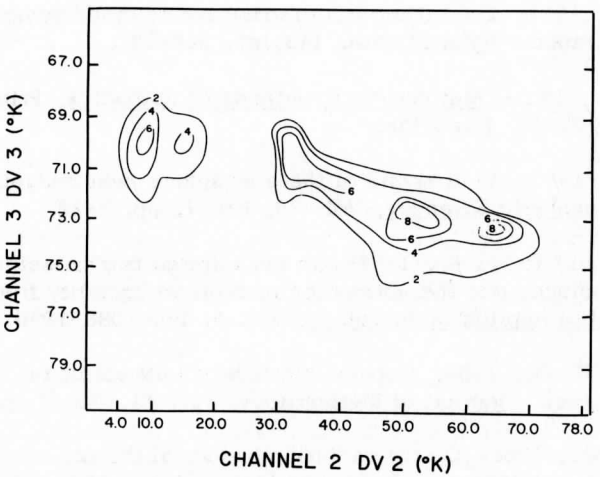


Fig. 11. Bivariate frequency distribution of DV3 and DV2 data for region B3.

Another improvement in this technique could be obtained through instrumental design. Because it is the variation of the effective emissivity of clouds within different spectral regions that leads to the multi-modal frequency distributions, the spectral interval of the channels should be selected such that the difference in effective emissivities is maximized. By emphasizing this difference, peaks of the various modes will be more pronounced and discrimination will be enhanced. Eventually as techniques and instruments are improved, refinements will lead not only to estimates of areal cloud coverage but also possibly to estimates of type, cloud type heights and vertical extent.

ACKNOWLEDGMENTS

The research was supported by the National Environmental Satellite Center, Environmental Science Services Administration through ESSA Grants WBG-52 and E-8-69(G). Partial support was provided by the Wisconsin Alumni Research Foundation and the National Science Foundation in the form of a grant for computational funds and the use of computer facilities.

BIBLIOGRAPHY

Arking, A., 1964: The latitudinal distribution of cloud cover from TIROS photographs. Science, Vol. 143, pp. 569-572.

Cramer, H., 1945: Mathematical methods of statistics, Princeton University Press, Princeton.

Davis, P., 1963: An analysis of the atmosphere heat budget. Journal of Atmospheric Sciences, Vol. 20, No. 1, pp. 5-22.

Fritz, S. and Rao, P. K., 1967: On the infrared transmission through cirrus clouds and the estimation of relative humidity from satellites. Journal of Applied Meteorology, Vol. 6, pp. 1088-1096.

Houghton, H. G., 1954: On the annual heat balance of the northern hemisphere. Journal of Meteorology, Vol. 11, No. 1, pp. 1-9.

Katayama, A., 1966: On the radiation budget of the troposphere over the northern hemisphere. Journal of Meteorological Society of Japan. Series II, Vol. 44, No. 6.

Kuhn, P. M., 1963: Measured effective long-wave emissivity of clouds. Monthly Weather Review, Vol. 91, No. 10-12, pp. 635-640.

London, J., 1951: Study of the atmospheric heat balance. New York University, N. Y. (Prog. Rept. 131. 06, Contract AF 19(122)-165), 21 pp.

Manabe, S. and Möller, F., 1961: On the radiative equilibrium and heat balance of the atmosphere. Monthly Weather Review, Vol. 89, No. 12, pp. 503-532.

Panofsky, H. A., Schwalb, A., Lethbridge, M. D., Huang, C., Frasch, C., Bradford, R., 1965: Synoptic applications of infrared satellite data. Interim Report, Grant No. WBG-48, Dept. of Meteorology, Pennsylvania State University.

Raschke, E., 1966: Tropospheric water vapor content and surface temperature from TIROS IV radiation data. NASA Contractor Report, NASA CR-595, Washington, D. C.

Rasool, S. I., 1964: Cloud heights and night-time cloud cover from TIROS radiation data. Journal of Atmospheric Sciences, Vol. 21, pp. 152-156.

Shenk, William E., 1963: TIROS II window radiation and large scale vertical motion. Journal of Applied Meteorology, Vol. 2, No. 6, pp. 770-775.

U. S. Navy Marine Climatic Atlas of the World, Vol. V, South Pacific Ocean, 1959, NAVAER 50-1C-532.

Widger, W. K. and Barnes, J. C., 1967: Comments on "variability in estimating total cloud cover from satellite pictures." Journal of Applied Meteorology, Vol. 6, No. 6, pp. 1126-1127.

Winston, J. S., 1965: Comments on "Cloud heights and night-time cloud cover from TIROS radiation data." Journal of the Atmospheric Sciences, Vol. 22, pp. 333-339.

Young, M. J., 1967: Variability in estimating total cloud cover from satellite pictures. Journal of Applied Meteorology, Vol. 6, No. 3, pp. 573-579.

Young, M. J., 1967: Reply to comments on "Variability in estimating total cloud cover from satellite pictures." Journal of Applied Meteorology, Vol. 6, No. 6, pp. 1128.

Scanner's note:

This page is blank.

THE USE OF APPROXIMATING POLYNOMIALS TO ESTIMATE PROFILES OF WIND, DIVERGENCE AND VERTICAL MOTION

Phillip J. Schmidt
and
Donald R. Johnson

Abstract:

"Least squares" approximating polynomials are used to suppress bias and random errors in estimating vertical profiles of winds, divergence and vertical motion. A quadratic polynomial is used to filter each wind profile. Profiles of divergence and vertical motion computed from a linear, a cross-product and a quadratic two-dimensional (horizontal) approximating polynomial model and from the Bellamy technique are compared. The random error variance component of the wind observations is estimated from the filtering polynomial prediction errors. In turn, the random-error variance component of the filtered wind, divergence and vertical motion is determined from the wind observational error variance for the various models.

In the presence of nonlinear variation in the horizontal wind field the Bellamy modeling assumption of linear wind variation introduces biased divergence errors. The divergence bias errors will persist through a considerable portion of the troposphere through the thermal wind relation and in the vertical integration cause large "spurious" vertical motion estimates of ω at the top of the profile. Divergence estimates from both the cross-product and the quadratic approximating polynomial models of the horizontal wind field tend to be less biased in this situation and normally produce superior vertical motion profiles.

I. INTRODUCTION

In the past two decades many attempts have been made to evaluate the large-scale fields of divergence and vertical motion. The methods most widely used have been the adiabatic method, the kinematic method and several methods employing vorticity. Previous studies (Palmén (1956), Endlich and Clark (1963), and Eddy (1964)) indicated that reliability of the kinematic technique was sufficient for it to be used in diagnostic studies if some added means were employed to reduce the influence of observational errors and small-scale effects. The intent of our study is to develop improved diagnostic techniques for atmospheric energy studies by using filtering polynomials. This is accomplished by incorporating vertically-filtered profiles of wind at each station to estimate a horizontally-filtered profile of divergence. Then vertical motion is computed from the divergence by "least squares" polynomial integration. Through these techniques the effects of random errors are suppressed and only the variations associated with the scale of the primary and secondary circulation are retained.

II. THE FILTERING POLYNOMIAL TECHNIQUE

The diagnostic model is based on application of "least squares" approximating polynomials (Hildebrand, 1956), that was initiated by Panofsky (1949) for the purpose of objective analysis and further applied by Gilchrist and Cressman (1954), D. H. Johnson (1957) and others. This approach assures that if "least squares" estimates are unbiased, they are better estimates of the true wind structure than basic observations since a portion of the random error component is filtered. In addition, if the random errors are also independent, the polynomial estimates are minimum variance estimates (Kendall and Stuart, 1958 and 1961).

In the filtering of each wind profile a second degree polynomial is fitted to seven adjacent observations of a wind component, thus leaving four degrees of freedom for the random error component. The filtered estimate is made for the mid-observation of the seven. The next filtered estimate is made by adding the next adjacent observation in height, deleting the lowest observation in height from the prior seven points and predicting for the new mid-observation. Only mid-point polynomial estimates are utilized since the random error variance of this estimate is a minimum (Johnson, 1965). Through this stepping procedure profiles of filtered wind components are obtained.

The horizontally-filtered divergence estimates are made by selecting from three to eight station wind profiles to estimate one profile of divergence near the center of the station array. The predicted eastward wind

component by the highest order approximating polynomial employed in the study is

$$\hat{u} = A_0 + A_1 x + A_2 y + A_{12} xy + A_{11} x^2 + A_{22} y^2 \qquad (1)$$

where the A_i's are the "least squares" polynomial coefficients estimated from the wind component data for each isobaric level and x and y are scaled longitudinal and latitudinal distances. The relation between x and longitude λ is

$$x = \frac{\lambda - (\lambda_{max} + \lambda_{min})}{(\lambda_{max} - \lambda_{min})} \qquad (-1 \leq x \leq +1) \qquad (2)$$

where λ_{max} and λ_{min} are the two extreme longitudes of all stations selected to estimate a profile of divergence. A similar relation exists between y and latitude ϕ. The square area circumscribed by λ_{max}, λ_{min}, ϕ_{max} and ϕ_{min} is called the local region since all stations used for the determination of the approximating polynomial are located within this region. The local region may change as new stations are added or a different combination of stations is utilized.

In the experimental results, profiles of divergence and vertical motion will also be presented for linear and cross-product models of polynomial approximation based on a truncation of (1). For the linear model only the first three coefficients are estimated, for the cross-product model the first four coefficients are estimated, while all six coefficients are estimated for the quadratic model. Since the approximations are by "least squares" the only restriction on the number of stations employed is that it must be equal to or exceed the number of coefficients in each model.

When the number of stations is equal to the number of coefficients the normal equations of the "least squares" method reduce to a set of ordinary simultaneous equations. In the situation where only three stations are employed for the linear model, estimates of divergence from the polynomial approximation will be identical to estimates by the Bellamy method (Bellamy, 1949).

The divergence is estimated by

$$\nabla \cdot \underline{U} = \frac{1}{R}\left(\sec \phi \frac{\partial u}{\partial \lambda} + \frac{\partial v}{\partial \phi} - v \tan \phi\right) \qquad (3)$$

when ∇ is the isobaric del operator, $\underline{U}$ is the velocity, R is the earth's radius, v is the northward wind component and λ and ϕ are the earth's longitude and latitude, respectively. At the midpoint of the scaled region

(0, 0), the relation between the filtered divergence estimate and the polynomial coefficients reduces to

$$\nabla \cdot \underline{U} = \left(\frac{A_1 \sec \phi}{\lambda_{max} - \lambda_{min}}\right) + \left(\frac{B_2}{\phi_{max} - \phi_{min}}\right) - B_0 \tan \phi)/R \qquad (4)$$

where B_0 and B_2 are from the set of coefficients used in representing the northward component by a polynomial expression similar to (1).

An alternative method would be to objectively analyze data from the rawinsonde network to a regular grid and then utilize conventional finite differencing operations that are valid through either second- or fourth-order to estimate horizontal divergence. Such a procedure is consistent with the idea of using higher order approximating polynomials to gain unbiased estimates of the first-order derivatives. However, this technique would produce unbiased estimates only if the rawinsonde winds were interpolated to a regular grid by an approximating function that produced unbiased wind component estimates. Due to the present irregular rawinsonde network, it is unlikely that any objective analysis technique exists that will produce unbiased estimates at fixed grid points under all possible configurations of atmospheric variables for the synoptic scale.

The vertical motion ω at the d^{th} isobaric level is

$$\omega = \omega(d - 1) + \Delta\omega(d) \qquad (5)$$

where the incremental vertical motion from the isobaric equation of continuity is

$$\Delta\omega(d) = -\int_{P_{d-1}}^{P_d} \nabla \cdot \underline{U}\, dp \quad . \qquad (6)$$

In evaluating this integral a second degree "least squares" polynomial was used to approximate the profile of filtered divergence over a vertical distance spanned by seven adjacent points. Then the segment of the polynomial between the midpoint observation and the adjacent lower point is used to represent the divergence between the limits in (6) and the integral is evaluated analytically. The entire profile of vertical motion is estimated by stepping through the divergence profile by the same manner utilized in computing the filtered wind profiles. This "ad hoc" technique of polynomial integration (after Johnson, 1965) has the advantage of combining polynomial filtering and integration. A more detailed derivation of this technique of integration and also the filtering technique for the wind and divergence profiles is presented in Appendix A.

III. EXPERIMENTAL RESULTS

The results obtained with the "least squares" technique are presented in three sections. Filtered profiles of wind are discussed in the first section. In the second section profiles of divergence determined by the Bellamy technique using vertically filtered winds and profiles of vertical motion from the polynomial integration are presented. In the third section the results from higher order approximating polynomials are compared.

The rawinsonde wind data obtained from the National Weather Records Center were listed at 50 mb intervals from the surface to 200 mb. Above this level the data interval decreased. In this study we shall be unable to analyze the validity of two assumptions utilized in the data reduction. These assumptions are that a balloon's trajectory may be used to estimate the wind and that discrete data at 50 mb intervals provide representative observations of a true, large-scale wind profile.

The first assumption should be valid since the balloon acts as an integrating factor of small-scale wind fluctuations and the relative, large-scale motion of the air past the balloon may be neglected (Perkins (1952)). The validity of the second assumption appears questionable from Reiter's example (Figure 2.122.1, p. 26 (1963)) showing systematic departure of the coded wind message from the profile determined from original rawin ascent records and the work of De Mandel and Scroggins (1967). However, for the scale for which horizontal divergence is determined in this study, the 50 mb vertical data interval should be sufficient.

Filtered Winds

In Figure 1, both the basic data—open circles—and the filtered estimates—solid line—for the eastward wind component are presented for Montgomery. The scatter of the open circles about the smooth profile is consistent with the suggestion by "least squares" theory that the random and small-scale components will be removed. Note the excellent description of the jet maximum of 65 m/sec.

An added advantage of using approximating polynomials is that estimates of random-error variance may be determined from the residual sums of squares provided that observational errors are random and independent and the approximating polynomial is unbiased.

In Figure 1a one-standard-deviation confidence interval is denoted by dotted lines. If this confidence interval were plotted about the true profile and the random errors were normally distributed, it should contain two-thirds of the wind observations. In the 1000 to 50 mb interval,

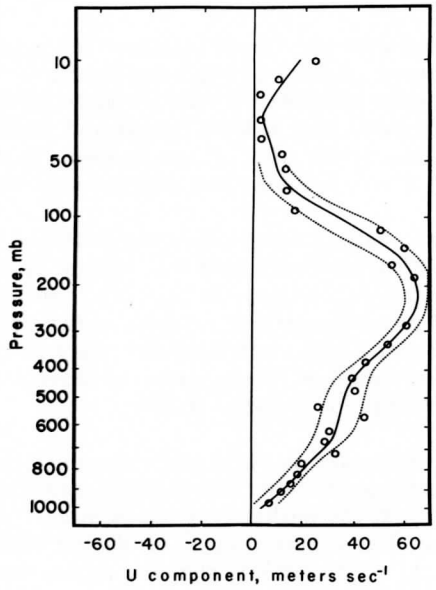


Fig. 1. Filtered eastward wind component profile for Montgomery on 0000Z, December 12, 1960. Open circles indicate wind observations. Dotted curves represent a one-standard-deviation confidence interval.

roughly two-thirds of the observations are within the limits. The variance of the filtered midpoint estimate from the seven point filtering interval for the quadratic approximation will be 0.333 times the observational variance. Thus the confidence interval for filtered wind estimates would be slightly less than 0.6 of the interval shown in Figure 1.

Divergence and Vertical Motion Estimates—Bellamy Techniques

In using the Bellamy technique for estimating divergence, the vertical motion profiles in many instances appeared reasonable. However, as a check, vertical motions were computed from overlapping triangles. In some cases the results were contradictory. Typical examples from 0000Z, December 12, 1960 are presented for the surface synoptic situation shown in Figure 2. In Figure 3 the divergence and vertical motion profiles for the triangle bounded by Columbia, Dayton and Nashville show excellent agreement with the synoptic situation. The convergence that occurs below 775 mbs and the upper level divergence display Dines' compensation.

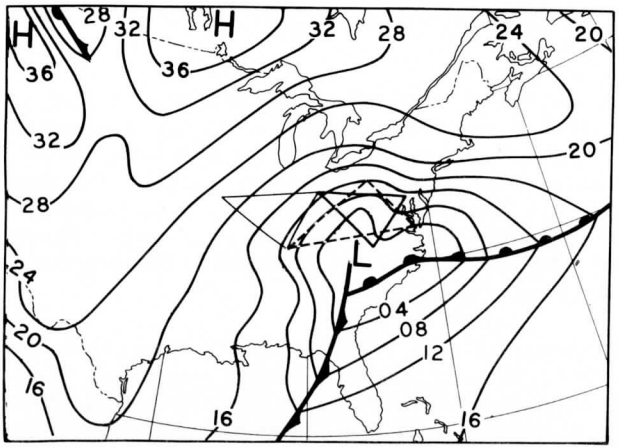


Fig. 2. Sea level pressure distribution for 0000Z, December 12, 1960. The three triangles designate areas for which divergence and vertical motion are estimated.

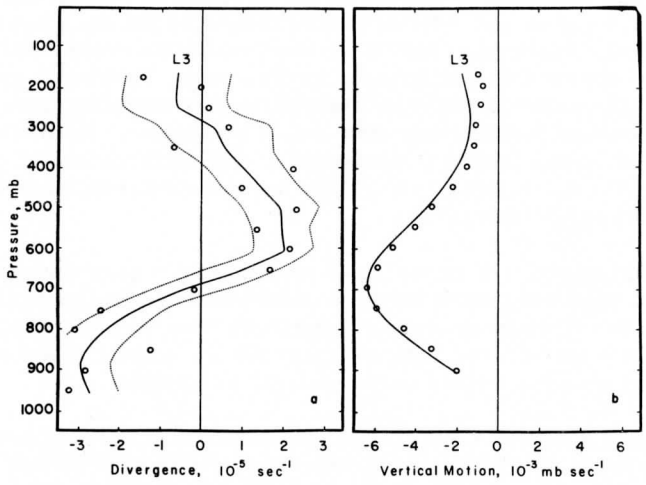


Fig. 3. Filtered divergence and vertical motion profiles computed by Bellamy technique (denoted L3) for a profile bounded by Columbia, Dayton and Nashville for 0000Z, December 12, 1960. Circles in 3a are unfiltered divergence estimates. Circles in 3b are vertical motion estimates from unfiltered divergence. Dotted lines represent a one-standard-deviation confidence interval.

The profile is similar to the profile of divergence for a developing cyclone presented by Petterssen (1956).

In Figure 3b, the filtered vertical motion profile-solid line-with an extremum of -6.4×10^{-3} mb sec^{-1} occurring at 700 mb agrees with Danielson's (1966) and Krishnamurti's (1968) results for a similar cyclonic situation. The estimate of ω also approaches a small value in the stratosphere. The circles in Figure 3b are vertical motion estimates from the polynomial integration of unfiltered divergence estimates. The differences are small and indicate that if only vertical motion estimates are desired, the preliminary filtering is unnecessary.

The dashed curves in Figure 3a bound the 68% (one standard deviation) confidence interval for the unfiltered divergence estimates. They include 12 of 17 points and indicate validity for standard deviation estimates, that are determined through a variance analysis of the residual sums of squares. A corresponding confidence interval for filtered divergence, is 0.577 times the indicated interval. This supports use of polynomial filtering since the range is considerably reduced.

The divergence and vertical motion profiles in Figure 4 show conflicting results for the overlapped triangles shown in Figure 2. The solid curves portray divergence and vertical motion estimates for the triangle bounded by Dayton, Washington and Greensboro while the dashed curves portray estimates for the triangle bounded by Nashville, Pittsburgh and Norfolk. Since the centroids of these triangles nearly coincide, the divergence and vertical motion estimates should be similar. Neither divergence profile displays Dines' compensation. The vertical motion profiles above 500 mb exhibit large, unrealistic values similar to those noted by Kurihara (1960) and Pfeffer (1962).

The standard deviation of the vertical motion random error component estimated for the uppermost value of the dashed profile in Figure 4b is 2.1×10^{-3} mb/sec. This is an order of magnitude less than the computed value of 3.5×10^{-2} mb/sec and indicates that the unrealistic result must be due to bias errors accumulated in the vertical integration.

In the example shown in Figure 3 our conclusion was that polynomial filtering is able to remove sufficient random-error variance to enable filtered profiles to closely approximate true atmospheric profiles. Now a distinction must be made between the true profile and the expected profile. If all modelling assumptions were satisfied, then the expected profile would be identical to the true profile. When assumptions are violated, the expected profile will diverge from the true profile due to presence of

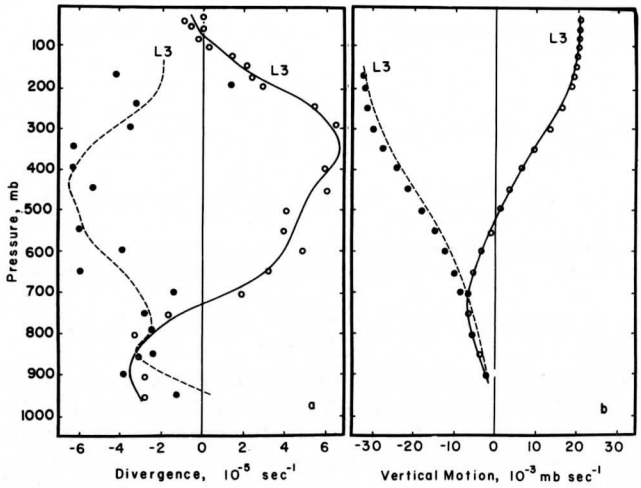


Fig. 4. Filtered divergence and vertical motion for two overlapping triangles for 0000Z, December 12, 1960. Solid curves—Dayton, Washington and Greensboro triangle; dashed curves—Nashville, Pittsburg and Norfolk triangle. Circles portray unfiltered estimates.

bias errors. Two possible sources of bias error in this Bellamy technique are from violation of the assumptions that (1) an unbiased filtering polynomial was selected in vertical filtering, or (2) the wind field is linear within the area bounded by the triangle. The likelihood that the first assumption is satisfied was illustrated in Figure 1. The possibility that the second assumption is violated in certain instances will now be considered.

The presence of nonlinearity in the wind field does not affect the ability of filtering polynomials to reduce the effects of an observational random-error component. However, if the wind field possesses nonlinear variation in the triangular region then estimates of divergence by the Bellamy method may contain a bias error.

To illustrate this condition, let the eastward component of a true, nonlinear wind field in a local region be given by the second order Taylor Series expansion:

$$u(x, y) = u_0 + \frac{\partial u}{\partial x})_0 (x - x_0) + \frac{\partial u}{\partial y})_0 (y - y_0) + \frac{1}{2} \frac{\partial^2 u}{\partial x^2})_0 (x - x_0)^2$$

$$+ \frac{\partial^2 u}{\partial x \partial y})_0 (x - x_0)(y - y_0) + \frac{1}{2} \frac{\partial^2 u}{\partial y^2})_0 (y - y_0)^2 \qquad (7)$$

For simplicity, let u_0, $\frac{\partial u}{\partial y})_0$, $\frac{\partial^2 u}{\partial x^2})_0$ and $\frac{\partial^2 u}{\partial y^2})_0$ be zero and $\frac{\partial u}{\partial x})_0$ be equal to 3.0×10^{-5} sec^{-1}. Since there usually is significant lateral shear variation in the horizontal, let $\frac{\partial^2 u}{\partial x \partial y})_0$ be equal to 3×10^{-10} sec^{-1} m^{-1}. In Figure 5, four triangles denoted by A, B, C and D are located so that their vertices lie on a circle with radius of 100 km centered at a common origin. The Bellamy divergence estimates at the single, common centroid computed using the true, discrete winds at the apex of the four triangles are 1.5×10^{-5} sec^{-1} for A; 4.5×10^{-5} sec^{-1} for B; and 3.0×10^{-5} sec^{-1} for C and D. Clearly two estimates, A and B, are biased while two, C and D, are unbiased even in the presence of the nonlinear field. The geometrical reason for the bias errors is associated with the orientation of the triangle in the nonlinear field and truncation error. Triangles C and D are laterally symmetric with respect to the x axis while triangles A and B are not. In estimates from triangle B, divergence information comes from discrete winds along a line of y equal to +50 km where $\frac{\partial u}{\partial x}$ is larger than $\frac{\partial u}{\partial x})_0$, but the estimates are assigned to the common origin. For triangle A information comes from the line y equal to -50 km where $\frac{\partial u}{\partial x}$ is small, but is again assigned to the origin. In the triangles C and D the lateral symmetric property averages the tendency to over- and under-estimate divergence from discrete wind information above and below the x axis. Thus in a nonlinear wind field, Bellamy divergence estimates for the centroid are not unique, but depend on orientation of the triangle with respect to the varying wind shear field.

Bias divergence errors must be present in either one or both of the contradictory profiles in Figure 4a. From an inspection of the 500 mb wind field, the cause of the bias error component is likely associated with the opposite orientation of the Dayton-Washington-Greensboro and Nashville-Pittsburg-Norfolk triangles in a nonlinear wind field and similar to the Taylor Series example.

Certainly higher order wind variations exist in the atmosphere. For example, $\frac{\partial^2 u}{\partial x \partial y}$, $\frac{\partial^2 v}{\partial x \partial y}$, $\frac{\partial^2 u}{\partial y^2}$ and $\frac{\partial^2 v}{\partial x^2}$ are elements in horizontal advection of vorticity that tend to be extrema in regions of significant vertical motion. In the atmosphere such nonlinear fields are maintained in the vertical through the thermal wind relation and thus the bias error accumulates in the integration to estimate vertical motion. We conclude that significant bias errors exist in Bellamy divergence estimates and that bias errors will likely be largest in regions where the gradient of vorticity and vorticity advection is a maximum.

Divergence and Vertical Motion from Cross-Product and Quadratic Models

The methods of Bellamy (1949) and Endlich and Clark (1963) are special cases of the general polynomial model. In their methods, the number of coefficients to be estimated is equal to the number of observations and an exact fit is always determined for the horizontal wind field. Two undesirable results of this condition are that no degrees of freedom are retained to study the "lack of fit" error for both bias and random components and that no horizontal filtering is accomplished. In the general model for which 5 and 8 stations are used to estimate horizontal wind fields, these undesirable features are avoided. However, in using more stations it is necessary to increase the size of the local region. The a priori problem is whether or not the expansion of the local region can be kept small enough to permit an unbiased approximating polynomial.

Our first consideration was the selection of a network arranged so that station location would constitute the basis of a desirable statistical design (Box and Wilson, 1951; Box and Hunter, 1957) while the local region of expansion would remain as small as possible. A region with five stations (Pittsburgh, New York, Washington, Greensboro and Cape Hatteras) arranged in quasi-rectangular fashion about one center station was selected. When postulating that the highest variation is quadratic in the local region, the quasi-rectangular array insures unbiased estimates in that the estimates of $\frac{\partial u}{\partial x})_0$ and $\frac{\partial u}{\partial y})_0$ are independent of $\frac{\partial^2 u}{\partial x^2})_0$ and $\frac{\partial^2 u}{\partial y^2})_0$. The center station provides one degree of freedom for filtering. If the station array formed a perfect rectangular array, the estimate of $\frac{\partial u}{\partial x})_0$ would also be independent of $\frac{\partial^2 u}{\partial x \partial y})_0$. A similar model is also valid for estimating $\frac{\partial v}{\partial y})_0$.

When quadratic variations are present, estimates of v_0 by the cross-product model will be biased. However, because the v_0 term in the divergence equation is usually an order of magnitude less than the divergence, it is likely that v_0 bias error does not contribute significantly to any vertical-motion bias component. Later three additional stations—Buffalo, Albany and Norfolk—were added to form an expanded eight station region and a full quadratic model is employed to eliminate the possibility of a bias error in v_0.

After the five station array was selected, preliminary checks were made to 1) ascertain whether the nonlinear wind field evident in the results of the overlapping triangles extended downstream through the five station region, and 2) test the assumption of the unbiased vertical filtering. In the first check Bellamy divergence and vertical motion estimates were made for eight different triangle combinations in the five station region. Station vertices for eight triangles identified by a letter and number are listed in Table I and the vertical motion profiles with corresponding identification are presented in Figure 6. The varying contradictory results indicate failure of the Bellamy model to portray reasonable synoptic similarity. Thus the nonlinear wind field in the local region provides data suited to study estimation by higher order models.

Table I

Overlapping triangles in local region used for Bellamy technique

T1:	New York, Greensboro, Pittsburgh
T2:	Cape Hatteras, Greensboro, Washington
T3:	New York, Cape Hatteras, Greensboro
T4:	New York, Pittsburgh, Washington
T5:	Greensboro, Pittsburgh, Washington
T6:	Cape Hatteras, Greensboro, Pittsburgh
T7:	New York, Cape Hatteras, Washington
T8:	New York, Cape Hatteras, Pittsburgh

The second check consisted of both an inspection of filtered profiles and a random-error variance analysis. No significant bias errors introduced by vertical filtering were present. Twelve-day mean RMS vector wind error estimates for five stations are presented and compared with Ellsaesser's (1957) results in Table II. The similarity achieved by averaging 60 profiles is excellent.

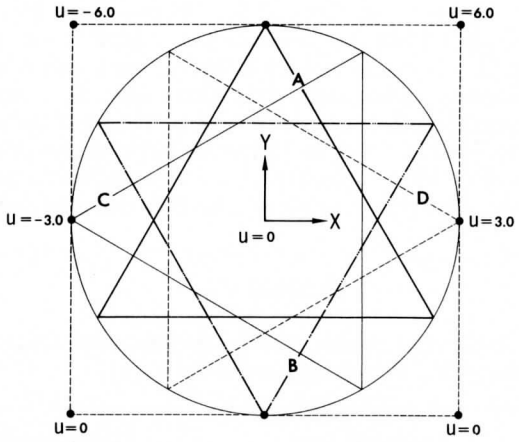


Fig. 5. Hypothetical distribution of four triangles to estimate divergence by the Bellamy method in a nonlinear wind field.

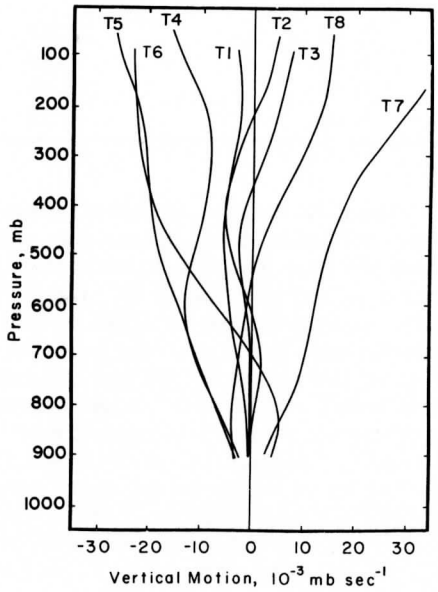


Fig. 6. Profiles of vertical motion by the Bellamy method for eight different triangles in the eastern U. S. for 0000Z, December 12, 1960. Triangle identification is given at the profile top and listed in Table I.

Profiles of divergence and vertical motion computed from four different models for 12 December 1960 are presented in Figure 7. A linear model using 5 stations (L5), a cross-product model using 5 and 8 stations (C5 and C8) and a quadratic model using 8 stations (Q8) were also used for comparison. All vertical motion profiles were computed by setting ω (950 mb) equal to zero in order to facilitate comparison of the different models. Since ω (950 mb) is usually small in comparison to mid-tropospheric values, the error introduced by this simplification is minor.

Table II

Twelve day mean RMS wind vector error estimates from Cape Hatteras, Greensboro, Washington, Pittsburgh and New York.

Level	Estimate from "lack of fit" (m/sec)	Ellsaesser's estimates (m/sec)
700	2.76	2.42
500	3.14	2.88
300	4.98	5.16
200	5.29	5.70

The four profiles in Figure 7a indicate low-level divergence, mid-tropospheric convergence, and upper-level divergence. The low-level divergence appears reasonable since it is associated with anticyclonic isobaric curvature in the local region (see Figure 2). Figure 7b shows that stratospheric estimates of ω from the higher order models tend to zero, a feature which suggests superior results. From Table III, the one standard deviation estimates of ω at 100 mb for the C5 and C 8 models are 2.14×10^{-3} and 1.59×10^{-3} mb/sec. Thus their actual departure from zero for this situation may be entirely explained by a random error component. The L5 model produced less acceptable results. A departure of 6.4×10^{-3} mb/sec from zero at the 100 mb level for the L5 model exceeds a three-standard-deviation estimate (Table III) and is almost certainly too large to be explained by the random-error component. However, the uppermost values of ω in Figure 7 are not nearly as great as those indicated for the eight Bellamy profiles in Figure 6.

In these results the decision whether or not the uppermost value of ω is biased is determined from estimates of the vertical motion random-error variance. The standard deviation estimates for divergence and vertical

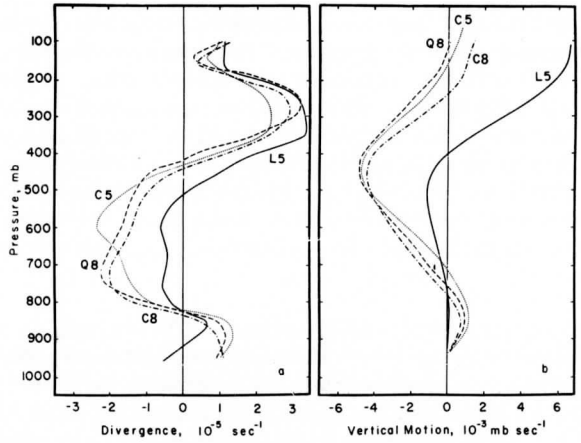


Fig. 7. Filtered divergence and vertical motion profiles for the origin of the approximating polynomial region in the eastern United States for 0000Z, December 12, 1960.

Table III

Standard deviation estimates for divergence (10^{-5} sec^{-1}) and vertical motion (10^{-3} mb/sec) for 0000Z, December 12, 1960.

	L3		L5		C5		C8	
Level	DIV	ω	DIV	ω	DIV	ω	DIV	ω
850	0.75	0.69	0.61	0.63	0.63	0.61	0.42	0.38
700	0.67	1.40	0.57	1.12	0.64	1.24	0.42	0.78
500	0.63	1.67	0.43	1.33	0.49	1.48	0.37	0.96
300	1.16	2.00	0.76	1.55	0.87	1.73	0.67	1.17
200	1.10	2.45	0.82	1.76	0.94	1.98	0.70	1.39
100	----	----	0.93	1.90	1.06	2.14	0.94	1.59
50	----	----	----	----	----	----	----	----

motion presented in Table III were computed from estimates of random-error variance from the initial vertical filtering. The tendency for the divergence standard deviation estimates to increase with height is due to the increase of the wind variance with height. The increase in standard deviation estimates of vertical motion is due to this effect and the accumulation of a random-error component through vertical integration of divergence. In comparing the models the random-error variance tends to be reduced as the local region is expanded and the degrees of freedom for horizontal filtering, i.e., the number of stations minus the polynomial coefficients, are increased.

In Figure 7 one cannot conclusively state which divergence or vertical motion profile is closest to the true value. In the midtroposphere, the systematic difference of divergence estimates between the linear and higher order models is large enough to indicate the presence of higher order variation in the wind field that was evident in the contradictory results from the Bellamy method. These results tentatively support the use of higher order polynomial models in preprocessing wind fields used for estimating kinematic divergence and vertical motion.

In computing kinematic divergence only estimates of linear derivatives of the wind field are utilized. Consequently, in selecting a model one need not be concerned with estimating the entire set of higher order polynomial coefficients, but rather with choosing a model which will produce unbiased estimates of the linear derivatives. Since the order of the selected model is truncated by necessity due to the marginal rawinsonde station density, the design of the model (Box and Wilson, 1951; Box and Hunter, 1957), i.e., the location of the reporting stations within the local region, determines to what extent the linear derivatives may be biased by the various orders of variation in the wind field.

In the next example the eastern U. S. was under the influence of a surface high pressure area and a 500 mb ridge on 0000Z, 5 December 1960. Because 500 mb winds were light and relatively uniform in the local region, one might expect that no significant nonlinear variations might be present. In Figure 8 profiles from a linear, L5, and two cross-product models, C5 and C8, are compared. In Table IV, the standard deviation estimate of ω at 50 mb for the L5 model is 1.49×10^{-3} mb/sec. The 50 mb ω estimate of -2.5×10^{-3} mb/sec is within two standard deviations of zero. It might be plausibly accepted, if it were not for the very systematic difference between the linear and cross-product estimates. From an inspection of the 500 mb wind field, the improvement in the cross-product estimates appeared to be associated with presence of significant $\frac{\partial^2 v}{\partial x \partial y}$ variation. $\frac{\partial v}{\partial y}$ was negative along a line from Pittsburgh to Greensboro, while between New York and Cape Hatteras it was positive. Thus

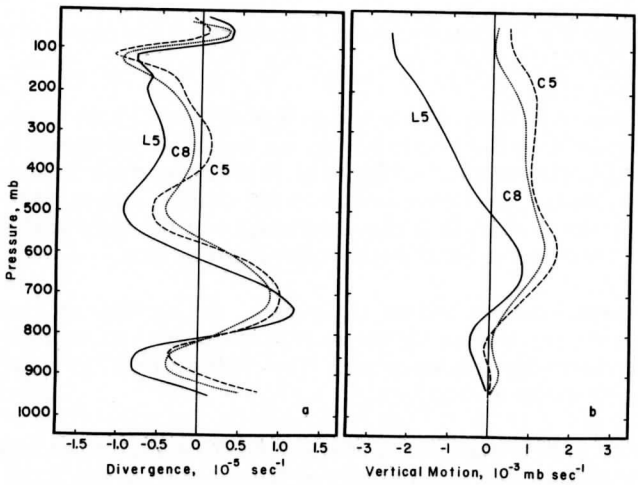


Fig. 8. Filtered divergence and vertical motion profiles for the origin of the approximating polynomial region in the eastern United States for 0000Z, December 5, 1960.

Table IV

Standard deviation estimates for divergence (10^{-5} sec^{-1}) and vertical motion (10^{-3} mb/sec) for date-times indicated in December 1960.

	05/00Z		12/00Z		05/00Z		05/12Z	
	L5		C5		C8		Q8	
Level	DIV	ω	DIV	ω	DIV	ω	DIV	ω
850	0.29	0.27	0.63	0.61	0.24	0.21	0.21	0.19
700	0.24	0.52	0.64	1.24	0.22	0.43	0.26	0.41
500	0.43	0.68	0.49	1.47	0.21	0.53	0.28	0.56
300	0.74	1.08	0.87	1.73	0.33	0.64	0.58	0.81
200	0.70	1.33	0.94	1.98	0.33	0.73	0.57	1.02
100	0.61	1.43	1.06	2.13	0.28	0.77	0.34	1.10
50	0.75	1.49	----	----	0.22	0.79	0.24	1.13

$\frac{\partial v}{\partial y}$ changed in the x direction and estimates by any model which neglects the term $\frac{\partial^2 v}{\partial x \partial y}$ will tend to be biased. In Figure 8, biased L5 divergence estimates appear to extend from the 650 mb to the 250 mb level. The marginal performance of the linear model can again be attributed to presence of a gradient of the vertical component of vorticity even in areas of light to moderate winds. In a comparison of the cross-product models the results in Figure 8 show that the oscillations of the C8 model are only slightly reduced by the expansion of the local region.

In the next example shown in Figure 9, the profiles of divergence and vertical motion are similar for all models up to the 600 mb level, above which systematic departures occur. The profiles from C5 to L5 are more biased than from C8 to Q8.

The bias in this example appears to be from neglect of higher than quadratic variation due to proximity of the jet core in the local region. If the five stations of the cross-product model were arranged in a perfect rectangle with one center station, bias errors in linear derivatives could be from the neglect of the functions x^2y, y^2x, x^3 and y^3 (Box and Wilson [3]; Box and Hunter [2]). In the 500 mb analysis for the local region x-quadratic variation appeared to vary in the y direction and is the most likely source of the bias error. In these situations selection of an unbiased model is extremely complicated. There is no method of knowing

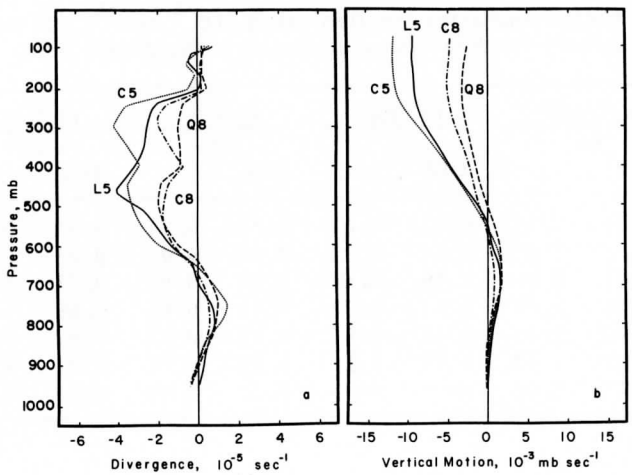


Fig. 9. Filtered divergence and vertical motion profiles for the origin of the approximating polynomial region in the eastern United States for 0000Z, December 25, 1960.

the order of variation a priori, and of insuring an optimum selection of a higher order model for a particular situation.

Another possible source of the bias error in Figure 8 is the violation of the assumption that winds are observed directly above the observing station. In reality high altitude observations are measurements made at an appreciable distance downstream. If lateral wind shears are present, the relative positions of the balloon will change with time causing biased divergence estimates. Although this bias is usually small (Kurihara, 1960), the deformation of relative balloon positions is appreciable in the vicinity of a jet core. When this systematic divergence error is integrated, vertical motion estimates must be biased.

Temporal Continuity of Vertical Motion Estimates

To portray temporal continuity, the cross-product model will be compared with the Bellamy method for a synoptic situation in which the Bellamy method produces reasonable estimates. The best situation in which winds were light and appeared to vary linearly in the local region occurred from 1200Z on December 2, 1960 to 1200Z on December 3, 1960. During this period the surface high pressure system and upper air ridge remained relatively stationary. The cross-product model was applied for the five station region and the Bellamy model for the eight triangles listed in Table III. The results from T2, which is a nearly equilateral triangle bounded by Cape Hatteras, Greensboro and Washington in the lower half of the five station region, were selected for the comparison. It was the only triangle that produced reasonable estimates for each of the three synoptic hours.

Successive profiles of vertical motion by the two models are presented in Figure 10. In a detailed comparison, all three profiles from the cross-product model satisfy upper boundary conditions in a more convincing fashion than the Bellamy profiles. Vertical motion profiles display an upward shift of the region of maximum downward motion as the upper-air, high pressure ridge moves slowly eastward from Western Pennsylvania into the local region. However, this feature is more prominent and shows better time continuity in the cross-product profiles. Because the cross-product model results were compared with the best of eight different results by the Bellamy method, we conclude that even in situations where the Bellamy model produces valid estimates, the cross-product estimates are better.

The final example is an extended, temporal comparison of vertical motion estimates for the Bellamy, the cross-product (C8) and the quadratic (Q8) models presented in Figure 11. The isopleth interval is not linear because of the large gradients at high levels. Estimates were not computed

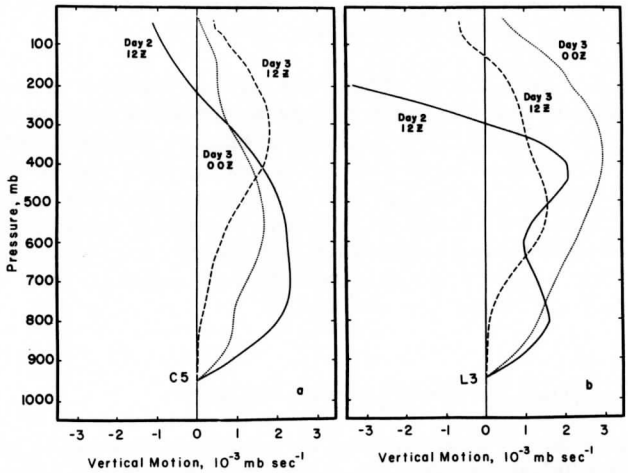


Fig. 10. Filtered vertical motion by the cross-product model (a) and the Bellamy method for triangle bounded by Cape Hatteras, Greensboro, and Washington, D. C. (b) from 1200Z, December 2 through 1200Z, December 3, 1960.

if there were insufficient data below the 100 mb level. The shaded regions indicate upward vertical motion; clear regions indicate downward motion; and speckled regions designate unrealistic magnitudes larger than 5×10^{-3} mb/sec. The frequency and vertical extent of unrealistic results in Figure 11 show that cross-product and quadratic models produce superior results. In particular regions of excessively large ω (Figure 11a) frequently extend further into the middle troposphere indicating that above 500 mb all kinematic estimates by the Bellamy method are probably suspect. Also, the 12 hour oscillations in the middle and upper troposphere from 5 through 7 December (Figure 11a) are probably unreal.

At this point we note that the method suggested by Kurihara (1960) and Pfeffer (1962) to adjust profile by the mean value of divergence is questionable when the magnitude of stratospheric ω's exceed that expected from the random-error component. An attempt should be made to locate the region of biased divergence values. Then, only after attempts to eliminate these bias errors by expanded modeling are made should the adjustment of subtracting mean value of divergence be applied so that the uppermost ω becomes equal to zero.

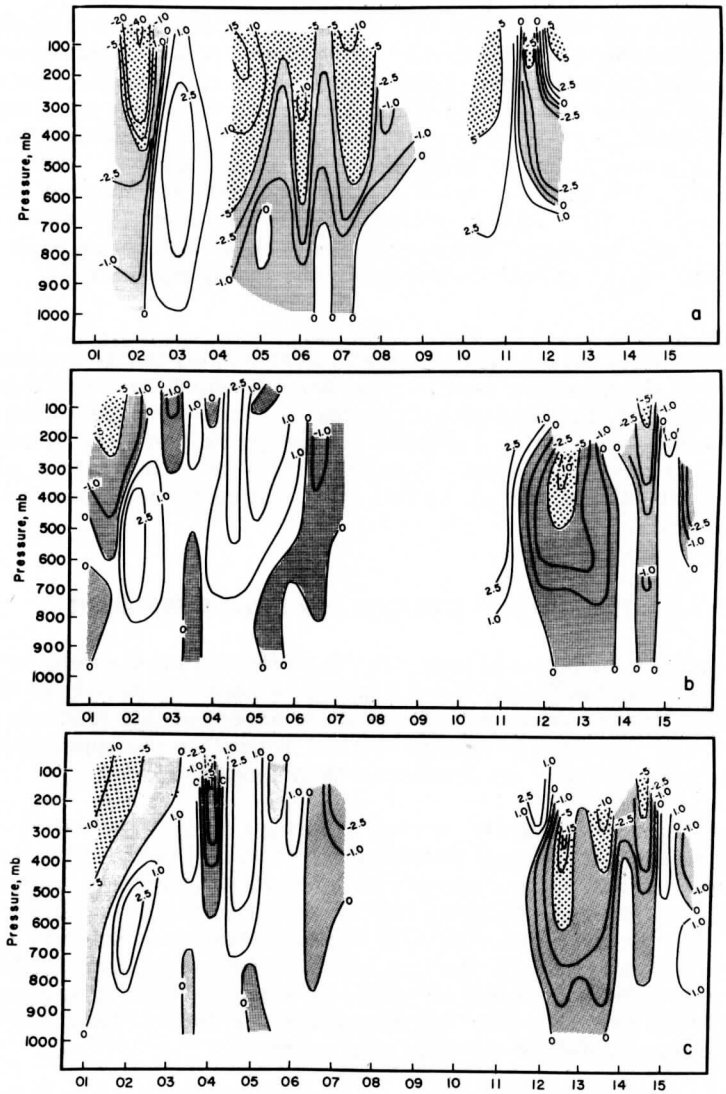


Fig. 11a, b, and c. Time cross-section of vertical motion estimates (10^{-3} mb/sec) from December 1 through December 15, 1960. In 11a the estimates are from the Bellamy method for the triangle bounded by Cape Hatteras, Greensboro, and Washington, D. C. Figures 11b and 11c portray estimates from a cross-product model, C8, and the quadratic model, Q8, respectively.

Both cross sections (Figure 11b and c) display the temporal continuity that one expects with transient weather systems. Large, upward vertical motions during the period from 1200Z December 11 to 1200Z December 13 occur during strong cyclonic activity. The downward motion occuring from 1200Z December 4 through 0000Z December 5 is associated with a relatively stationary surface anticyclone and upper air ridge. In the cross-product model, the maximum value of downward vertical motion at the 500 mb level is 2.5×10^{-3} mb/sec while the maximum upward velocities are near -5.0×10^{-3} mb/sec. The magnitudes of these values are comparable with results of both Danielsen (1966) and Krishnamurti (1968).

The patterns from the cross-product model display better continuity than those of the quadratic model. This improvement is probably due to the condition that the cross-product model, in contrast to the quadratic model, retains two additional degrees of freedom for horizontal filtering of the random-error component. Note in Table IV that model C8 standard deviation estimates of vertical motion are significantly less than the values for model Q8.

V. CONCLUSIONS AND SUGGESTIONS FOR FUTURE RESEARCH

The problem of unbiased interpolation of fields from discrete measurements is extremely difficult, if not impossible when the density of the reporting stations is marginal. The positive aspect of this research is that the problem of "spurious kinematic vertical motion" in the upper atmosphere has been isolated and alternative techniques proposed.

The results show that the vertical filtering of wind observations is desirable to suppress effects of random errors in estimating divergence and vertical motion profiles and to gain estimates of the random-error variance. An increased sampling rate to provide more vertical observations of the balloon positions is also desirable since, as noted by Rachele and Duncan (1967), the profiles would be more representative of the true wind. With an increased sampling rate, approximating polynomial estimates would be even more accurate due to the Central Limit effect in reducing the random-error variance component.

The most serious problem in kinematic, vertical motion calculations lies in estimation of the horizontal divergence which belongs to the more general problem of two-dimensional interpolation. It is emphasized that 1) utilization of diagnostic techniques which yield unbiased estimates of atmospheric variables, and 2) incorporation of a means of checking the validity of modelling assumptions are extremely important considerations. Our results indicate that estimates of horizontal divergence and

vertical motion from wind gradients computed with cross-product and quadratic polynomial models are superior to those obtained from the linear model. Bias errors present in Bellamy estimates of divergence have been traced to violation of the modelling assumption of linear wind variation. Regions of nonlinear wind variation in many instances are also regions where the gradients of vorticity, vorticity advection and vertical motion are extrema. Thus, the assumption of linearity utilized in the Bellamy technique is inconsistent with the expected velocity distributions where more active weather situations are occurring.

From considerations of typical wind distributions in quasigeostrophic balance, optimum statistical designs and our diagnostic results, the cross-product model is the lowest order model that should be used for direct kinematic vertical motion estimates using the present U. S. rawinsonde network. Even though higher order modeling tends to reduce biased estimates of vertical motion, the cross-product and quadratic results indicate that bias errors are not always eliminated. In an ideal situation the choice of a model should be based on the principal features of the wind field and the pattern of reporting stations in each local area. The danger in selecting a unique order for the interpolating polynomial is that a single model is not optimized to produce unbiased estimates in all situations. In order to reduce the polynomial order required for unbiased estimation for wind fields associated with asymmetric ridges and intense troughs or jets, translation and rotation of the independent variables should be attempted to gain symmetry of the wind field with respect to both the origin and axes of the scaled independent variables. Such techniques will improve the statistical design for the particular wind field and thereby reduce the tendency for biased estimates. In regions of strong wind shear, the use of a variable design to compensate for relative changes of the balloon's position due to wind drift should be investigated to insure that this source of bias error is unimportant.

ACKNOWLEDGMENTS

We are grateful to Dr. Christopher Hayden of the National Environmental Satellite Center for his careful review and helpful suggestions. The research was supported by the National Environmental Satellite Center, Environmental Science Services Administration through ESSA Grants WBG-52 and E-8-69(G). Partial support was provided by the Wisconsin Alumni Research Foundation and the National Science Foundation in the form of a grant for computational funds and the use of computer facilities.

APPENDIX A

The technique of filtering and integration by approximating polynomials to estimate profiles of wind, divergence and vertical motion

The General Filtering Polynomial

In general, the $(p \times 1)$ vector of discrete k^{th} degree "least squares" polynomial estimates for either the vertical or horizontal filtering operations is

$$\hat{\underline{z}}_k = \underline{X}\,\underline{B} \tag{1}$$

where the subscript k denotes the degree, $\underline{X}$ is a $(p \times q)$ design matrix and $\underline{B}$ is a $(q \times 1)$ vector of polynomial coefficients given by

$$\underline{B} = [\underline{X}^T\underline{X}]^{-1}\,\underline{X}^T\underline{z} \tag{2}$$

$\underline{z}$ is the $(p \times 1)$ vector of observations.

The general design matrix $\underline{X}$ for the k^{th} degree approximating polynomial used in estimating the horizontal wind field is

$$\underline{X} = \left[\begin{array}{c|ccc|c|ccccccc} 1 & x_1 & y_1 & x_1^2 & x_1y_1 & y_1^2 & \cdots & x_1^k & x_1^{k-1} & y_1 & \cdots & x_1 & y_1^{k-1} & y_1^k \\ 1 & x_2 & y_2 & x_2^2 & x_2y_2 & y_2^2 & \cdots & x_2^k & x_2^{k-1} & y_2 & \cdots & x_2 & y_2^{k-1} & y_2^k \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 1 & x_n & y_n & x_n^2 & x_ny_n & y_n^2 & \cdots & x_n^k & x_n^{k-1} & y_n & \cdots & x_n & y_n^{k-1} & y_n^k \end{array}\right] \tag{3}$$

where x and y are scaled longitudinal and latitudinal coordinates respectively. $\underline{X}$ has been partitioned to indicate combinations of powers of the independent variable associated with a full k^{th} degree polynomial representation. The design matrix $\underline{X}$ for k^{th} degree vertical filtering polynomials is simply

$$\underline{X} = \begin{bmatrix} 1 & x_1 & x_1^2 & \cdots & x_1^k \\ 1 & x_2 & x_2^2 & \cdots & x_2^k \\ \vdots & & & & \\ 1 & x_n & x_n^2 & \cdots & x_n^k \end{bmatrix} \tag{4}$$

where x is the scaled pressure coordinate. In general, the scaled independent variable x_i (or y_i) over the interval $[-1, 1]$ is

$$x_i = \frac{X_i - 1/2(X_{max} + X_{min})}{1/2(X_{max} - X_{min})} \tag{5}$$

where X_{max} and X_{min} are maximum and minimum values in the local region. The vector of polynomial estimates expressed as a linear transformation of the original observations is

$$\hat{\underline{z}}_k = \underline{X}[\underline{X}^T\underline{X}]^{-1}\underline{X}^T\underline{z} \tag{6}$$

The i^{th} polynomial estimate of the vector $\underline{z}$ is

$$z_i = \underline{i}\,\underline{X}[\underline{X}^T\underline{X}]^{-1}\underline{X}^T\underline{z} \tag{7}$$

where $\underline{i}$ is a $(1 \times p)$ row vector with unity in its i^{th} element and zeros elsewhere. While equation (6) provides discrete polynomial estimates at all observational points, normally only midpoint polynomial estimates from (7) are used to construct wind or divergence profiles since the random error variance of the midpoint estimate is a minimum (Johnson, 1956).

Vertical Filtering of Wind Profiles

In generating vertically filtered profiles for the eastward and northward wind components, u and v, the k^{th} degree "least squares" approximating polynomial is fitted to a subregion of the original span of the entire $(R \times 1)$ height-ordered observational vectors $\underline{U}$ and $\underline{V}$ with elements u_r and v_r $(r = 0, 1, \ldots R)$. The span of the r^{th} subregion, the filtering interval, is designed to contain an equal number of observations on either side of the midpoint prediction. Thus the span contains $(2d_1 + 1)$ observations indexed $j = 1, \ldots, (2d_1 + 1)$. The midpoint prediction $(j = d_1 + 1)$ normally made for each polynomial fit is given from (7) as

$$\hat{u}_r = \underbrace{(d_1 + 1)\underline{X}[\underline{X}^T\underline{X}]^{-1}\underline{X}^T}\, \underline{u}_r \tag{8}$$

$$\hat{v}_r = \underbrace{(d_1 + 1)\underline{X}[\underline{X}^T\underline{X}]^{-1}\underline{X}^T}\, \underline{v}_r \tag{9}$$

A subsequent $(r+1)^{th}$ midpoint estimate from a new subregion of $(2d_1 + 1)$ observations is made by deleting the $(r - (d_1+1))^{th}$ observation from the previous filtering interval and adding the $(r + d_1 + 2)^{th}$ observation. Exceptions to the midpoint prediction technique occur in the filtered polynomial estimate for the lowest d_1 and highest d_1 observations of the vectors $\underline{U}$ and $\underline{V}$. Estimates for these observations are made by selecting the appropriate d_1 predictions for these points from the fitted polynomial for the lowest and highest filtering intervals. In this manner $(R \times 1)$ vectors of filtered wind estimates $\hat{\underline{U}}$ and $\hat{\underline{V}}$ are generated. Their elements correspond one-to-one to the elements of $\underline{U}$ and $\underline{V}$.

Horizontal Filtering of Wind Profiles

The general polynomial expressions for the continuous wind field in the horizontal filtering region for the r-th pressure level (r = 0, 1, ... D_1, ... d, ... D_2, ... R) are

$$\hat{\hat{U}}_r = \underline{x}_r \underline{A}_r \tag{10}$$

$$\hat{\hat{v}}_r = \underline{x}_r \underline{B}_r \tag{11}$$

where $\underline{x}_r$ is

$$\underline{x}_r[1 | xy\ x^2\ xy\ y^2 | \ldots | x^k\ x^{k-1}\ y \ldots xy^{k-1}\ y^k]_r \tag{12}$$

The indices D_1 and D_2 denote the lowermost and uppermost levels respectively, for which constant-pressure-surface data is available for all M observing stations. The double hat indicates the condition that polynomial estimates at this step have been determined from wind estimates which have been previously filtered in the vertical. $\underline{A}$ and $\underline{B}$ are $(q \times 1)$ vectors of "least squares" coefficients given from (2) by

$$\underline{A}_r = \overbrace{[\underline{X}^T\underline{X}]^{-1}\underline{X}^T}^{r}\, \hat{u}_r \tag{13}$$

$$\underline{B}_r = \overbrace{[\underline{X}^T\underline{X}]^{-1}\underline{X}^T}^{r}\, \hat{v}_r \tag{14}$$

$\underline{u}_r$ and $\underline{v}_r$ are $(M \times 1)$ vectors of filtered wind estimates at the same pressure level from M rawinsonde stations. The wavy overbar denotes the design matrix or sequence of matrices associated with the r^{th} level, approximating polynomial. The r superscript for the horizontal approximating polynomial denotes the general condition that the horizontal

design matrix should be allowed to vary with height to allow for relative changes of the geographical position of the different, ascending rawin balloons. However, since the balloon-position data were not readily available from recorded wind data, the horizontal design matrix in our analysis was later fixed using the geographical position of rawin stations. The error induced by this restriction has been discussed by Deardorff (1957) and Kurihara (1960) and is considered in more detail in the results.

Computation of Divergence from Filtered Wind Profiles

Kinematic divergence is estimated by

$$\nabla \cdot \vec{V} = \frac{1}{R_1} (\sec \phi \frac{\partial u}{\partial \lambda} + \frac{\partial v}{\partial \phi} - v \tan \phi) \tag{15}$$

where ∇ is the isobaric del operator, $\vec{V}$ is the wind vector, R_1 is the earth's radius, and λ and ϕ are the earth's longitude and latitude, respectively.

In contrast to Panofsky's (1949) representation of divergence over an extended region by polynomial techniques, only the divergence estimate at the origin ($x = 0$, $y = 0$) of the scaled region is used in our method. Polynomial estimates near the boundary of the filtering interval are particularly susceptible to effects of bias and random errors. From a differentiation of (10) and (11) with respect to longitude and latitude, respectively, and an evaluation at the origin (0, 0) of the filtering region, the derivatives used to estimate divergence are

$$\frac{\partial \hat{\hat{u}}_r}{\partial \lambda})_0 = C_{1r} A_{1r} \tag{16}$$

$$\frac{\partial \hat{\hat{v}}_r}{\partial \phi})_0 = C_{2r} B_{2r} \tag{17}$$

The v-component estimate at the origin is simply

$$\hat{\hat{v}}_{r0} = B_{0r} \tag{18}$$

A_{1r} is the second element of $\underline{A}_r$ and B_{0r} and B_{2r} are the first and third elements of $\underline{B}_r$. The Constants C_1 and C_2 are denominators from the general scaling relation (5) for the longitudinal and latitudinal axis, respectively. Using (16), (17) and (18), the filtered divergence estimate at the r^{th} pressure level is

$$\nabla \cdot \hat{\hat{\vec{V}}}_r = K_{1r} A_{1r} + K_{2r} B_{2r} + K_{0r} B_{0r} \tag{19}$$

where

$$K_{1r} = C_{1r} \sec \phi_{0r} / R_1 \tag{20}$$

$$K_{2r} = C_{2r} / R_1 \tag{21}$$

$$K_{0r} = \tan \phi_{0r} / R_1 \tag{22}$$

The $(D_2 - D_1 + 1) \times 1$ vertical-profile vector of divergence estimates $\underset{\sim}{\nabla \cdot \hat{\hat{\vec{V}}}}$ is determined for the origin of the filtering region.

Filtered Vertical Motion Profiles

The "least squares" polynomial, vertical-motion model is based on the equation of continuity in pressure coordinates given by

$$\frac{\partial \omega}{\partial p} = - \nabla \cdot \vec{V} \tag{23}$$

ω, the vertical motion in pressure coordinates, is defined at the d^{th} isobaric level to be

$$\omega(d) = \omega(d - 1) + \Delta\omega(d) \tag{24}$$

The incremental vertical motion using (10) and (23) becomes

$$\Delta\omega(d) = - \int_{p_{d-1}}^{p_d} \nabla \cdot \vec{V} \, dp \tag{25}$$

From (24) and (25)

$$\omega(d) = \omega_0 + \Delta\omega(D_1) - \sum_{r=D_1+1}^{D_2} \int_{p_{r-1}}^{p_r} \nabla \cdot \vec{V} \, dp \tag{26}$$

where ω_0 is the boundary value at the earth's surface and $\Delta\omega(D_1)$ is the boundary-layer incremental vertical motion given by

$$\Delta\omega(D_1) = - \int_{p_0}^{p_{D_1}} \nabla \cdot \vec{V} \, dp \tag{27}$$

The necessity of defining an incremental vertical motion for the boundary

layer is due to the condition that the surface level p_0 varies for all stations.

In a determination of the r^{th} incremental vertical motion, $\Delta\omega(r)$, a "least squares" approximating polynomial is used to estimate a continuous profile of divergence from discrete, filtered estimates over a span of pressure from P_{r-d_2} to P_{r+d_2}. The number of observations in the filtering interval is $(2d_2 + 1)$. The continuous profile of estimated divergence over the r^{th} filtering interval is

$$\hat{\nabla\cdot\vec{V}}_r = \underline{x}_r \underline{D}_r \tag{28}$$

where $\underline{D}_r$ is a $((k+1) \times 1)$ vector of "least squares" coefficients and $\underline{x}_r$ is a $(1 \times (k+1))$ vector with continuous functional elements x^i given by

$$\underline{x}_r = [1, x, x^2, \ldots, x^k]_r \tag{29}$$

The vector of coefficients $\underline{D}_r$ for the integrating polynomial is related to discrete divergence estimates by

$$\underline{D}_r = \overline{[\underline{X}^T\underline{X}]^{-1}\underline{X}^T}^r \underbrace{\hat{\nabla\cdot\vec{V}}}_r \tag{30}$$

where $\underbrace{\hat{\nabla\cdot\vec{V}}}_r$ is a $(2d_2 \times 1)$ subvector of $\underbrace{\hat{\nabla\cdot\vec{V}}}$. The straight overbar indicates design matrices associated with the polynomial integration technique. Transforming to x as the variable of integration and substituting the continuous polynomial representation of divergence estimates from (28) into (25) yields the filtered incremental vertical motion given by

$$\Delta\hat{\omega}(r) = - C_r \int_{x(P_{r-1})}^{x(P_r)} \underline{x}_r \underline{D}_r dx \tag{31}$$

where C_r is the denominator from the general scaling relation (5). Since $\underline{D}_r$ is constant with respect to the region of integration

$$\Delta\hat{\omega}(r) = - C_r \underline{X}_r \underline{D}_r \tag{32}$$

where the integrated vector is

$$\underline{X}_r = \left[(x(P_r) - x(P_{r-1})), \frac{(x^2(P_r) - x^2(P_{r-1}))}{2}, \frac{(x^3(P_r) - x^3(P_{r-1}))}{3}, \ldots, \frac{(x^k(P_r) - x^k(P_{r-1}))}{k}\right] \tag{33}$$

Filtered vertical motion at the d^{th} pressure level is now

$$\hat{\omega}(d) = \omega_0 + \Delta\omega(D_1) - \sum_{r=D_1+1}^{d} \underline{C}_r \underline{X}_r \underline{D}_r \qquad (34)$$

This "ad hoc" technique of polynomial integration (after Johnson (1965)) has the advantage of combining polynomial filtering and integration.

BIBLIOGRAPHY

Bellamy, J. C., 1949: "Objective calculations of divergence vertical velocity and vorticity," Bulletin of the American Meteorological Society, Vol. 30, No. 2, pp. 45-50.

Box, G. E. P. and J. S. Hunter, 1957: "Multi-factor experimental designs for exploring response surfaces," Annals of Mathematical Statistics, Vol. 28, No. 1, pp. 195-241.

Box, G. E. P. and K. B. Wilson, 1951: "On the experimental attainment of optimum conditions," Journal of the Royal Statistical Society, Series B., Vol. 13, No. 1, pp. 1-45.

Danielsen, E. P., 1966: "Research in four-dimensional diagnosis of cyclonic storm cloud systems," Scientific Report No. 1, Contract AF 19(628)-4762, Department of Meteorology, Pennsylvania State University.

De Mandel, R. E. and J. R. Scroggins, 1967: "Mesoscale wave motions as revealed by improved wind profile measurements," Journal of Applied Meteorology, Vol. 6, No. 4, pp. 617-620.

Eddy, A., 1964: "The objective analysis of horizontal wind divergence fields," Quarterly Journal of the Royal Meteorological Society, Vol. 90, No. 4, pp. 424-440.

Ellsaesser, H. W., 1957: "Errors in upper-level wind forecasts," Bulletin of the American Meteorological Society, Vol. 38, No. 9, pp. 511-517.

Endlich, R. M. and J. R. Clark, 1963: "Objective computation of some meteorological quantities," Journal of Applied Meteorology, Vol. 2, No. 1, pp. 66-81.

Gilchrist, B. and G. P. Cressman, 1954: "An experiment in objective analysis," Tellus, Vol. 6, No. 4, pp. 309-318.

Hildebrand, F. B., 1956: Introduction to Numerical Analysis, McGraw-Hill, New York.

Johnson, D. R., 1965: "The role of terrestrial radiation in the generation of available potential energy," Ph. D. Thesis, Department of Meteorology, University of Wisconsin.

Johnson, D. H., 1957: "Preliminary research in objective analysis," Tellus, Vol. 9, No. 3, pp. 316-322.

Kendall, M. G. and A. Stuart, 1958 and 1961: "The advanced theory of statistics," Vol. 1, Distribution Theory (1958), and Vol. 2 Statistical Inference and Its Statistical Relationship (1961), C. Griffin, London.

Krishnamurti, T. N., 1968: "A study of a developing wave cyclone," Monthly Weather Review, Vol. 96, No. 4, pp. 208-217.

Kurihara, Y., 1960: "Numerical analysis of atmospheric motions," Journal of Meteorological Society of Japan, Series II, Vol. 38, No. 6, pp. 288-304.

Palmén, E., 1956: "Vertical circulation and release of kinetic energy during the development of hurricane Hazel into an extratropical storm," Tellus, Vol. 10, No. 1, pp. 1-23.

Panofsky, H. A., 1949: "Objective weather map analysis," Journal of Meteorology, Vol. 6, No. 6, pp. 386-392.

Perkins, D. T., 1952: "The response of balloons to the wind," Bulletin of the American Meteorological Society, Vol. 23, No. 4, pp. 135-139.

Petterssen, S., 1956: Weather Analyses and Forecasting, McGraw-Hill, New York.

Pfeffer, R. L., 1962: "Results of recent research in meteorology at the Lamont Geological Observatory," Proceedings of the International Symposium on Numerical Weather Prediction in Tokyo, The Meteorological Society of Japan, pp. 249-264.

Rachele, H. and L. D. Duncan, 1967: "Desirability of using a fast sampling rate for computing wind velocity from pilot-balloon data," Monthly Weather Review, Vol. 95, No. 4, pp. 198-202.

Reiter, E. R., 1963: Jet Stream Meteorology, University of Chicago Press, Chicago and London, 515 pp.