The Forcing and Balance of Zonally Symmetric Modes in a Global Model

RONALD GELARO
 Naval Research Laboratory, Monterey, California

RONALD M. ERRICO
 National Center for Atmospheric Research, Boulder, Colorado

(Manuscript received 2 February 1992, in final form 27 May 1992)

ABSTRACT

The temporal behavior of the zonal wavenumber 0 Hadley modes in the U.S. Navy's operational numerical weather prediction model is investigated. Time series of individual gravitational normal modes are examined in forecasts using different initialization strategies to determine the existence and nature of any nonlinear balance and the impact of initial conditions. A form of diabatic initialization based on a least-squares fit to the modes' trajectories in the forecast model is developed. A straightforward interpretation of the results, particularly how they relate to the principles of nonlinear normal-mode initialization, is afforded via a simple prognostic equation for a single gravitational mode. A limited number of behavior types is observed, which varies depending on the temporal and meridional scales of the modes. Only the largest-scale zonally symmetric modes show any evidence of diabatic balance and may thus be suitable candidates for diabatic initialization. Convective heating is the primary stationary forcing for these modes. Most medium-scale modes, whose natural periods may be close to the diurnal period, behave in a forced, wave-like manner due, apparently, to near-resonant diurnal forcing. These modes with the smallest temporal and meridional scales exhibit both balanced and forced behavior. The dominant forcing for these modes appears to be diabatic.

1. Introduction

The specification of appropriate initial conditions for models used in numerical weather prediction (NWP) remains a challenging problem. The process of data assimilation, whereby a numerical forecast model is updated in a coherent and physically consistent manner from diverse, temporally inconsistent, and spatially incomplete observations, has been used successfully for this purpose, as reviewed by Bengtsson (1975). This is accomplished through the predictive capability of the numerical model, which carries ingested information forward in time so that it can be consistently fused with observations to form an analysis of the appropriate fields.

Despite the sophistication of current data-assimilation schemes, most numerical forecasts suffer from unrealistically large gravitational-wave oscillations, commonly referred to as "noise," during their early stages. For the most part, these oscillations are due to 1) inconsistencies between the analyzed fields and the internal dynamics of the forecast model, or 2) errors in the analyses themselves. Generally speaking, the first type of noise has little appreciable impact on most meteorologically significant components of the flow, while the second type may have serious and long-lasting impacts on a broad range of scales. The latter can be especially troublesome in the tropics, where analysis errors tend to be large as a result of inadequacies in both the density and quality of observations (Daley and Mayer 1986). Moreover, since many of the important atmospheric processes in the tropics (e.g., convection) are only crudely parameterized in most forecast models, the forecast fields have only limited potential as reliable substitutes for observations in the data-assimilation process.

To eliminate the noise inherent in most analyses, the analyzed fields are often modified so as to provide dynamically consistent initial conditions for the forecast model. This process is called initialization, for which several methods have been developed (e.g., Charney 1955; Phillips 1960; Nitta and Hovemale 1969; Dickinson and Williamson 1972; Sugi 1986). In recent years, nonlinear normal-mode initialization (NNMI; Baer 1977; Machenhauer 1977) has become a widely used initialization method in numerical weather prediction. This method is based on the normal-mode solutions to a linearized version of the forecast model. The analyzed data are modified according to a diagnostic relationship in which their gravitational-mode tendencies are assumed to vanish due to a balance between the inertial–gravitational and nonlinear

Corresponding author address: Dr. Ronald Gelaro, Naval Research Laboratory, Monterey, CA 93943-5006.

© 1993 American Meteorological Society
forces acting on the modes. An advantage of NNMI is that it allows great flexibility in selecting the spatial and temporal scales to be initialized. An outstanding problem, however, is the determination of an appropriate set of modes to balance, especially for those modes associated with tropically driven circulations, in which diabatic processes may be important. As with the analysis errors themselves in the tropics, this is largely a reflection of our inability to measure and define accurately the forces that govern these circulations.

The inclusion of diabatic forcing as part of the NNMI balance condition—referred to as diabatic NNMI (e.g., Wergen 1987)—has been proposed as a means of improving the initialization of tropically driven circulations. This, however, has proven to be only marginally successful, as a result of both theoretical and procedural difficulties. In particular, there appears to be little evidence that the NNMI balance condition applies to the tropical atmosphere in a general sense. For example, Errico et al. (1988) investigated the balance properties of normal modes in two forecast models and found little or no evidence of diabatic balance for modes associated with tropically driven circulations in either model. Despite these and other findings, there remains a considerable amount of uncertainty regarding the appropriateness of NNMI for these modes, and their temporal behavior in general.

In this paper, we use the U.S. Navy Operational Global Atmospheric Prediction System (NOGAPS; Hogan and Rosmond 1991) to examine the forcing and balance of large-scale zonally symmetric gravitational modes; that is, the zonal wavenumber 0 Hadley circulation. The present study differs from earlier ones on this subject in that we relate these characteristics directly to the temporal behavior of individual modes rather than only statistics derived from the complete set of modes or forecast fields. One objective of this study is to determine whether such modes exhibit the behavioral characteristics upon which the assumptions of NNMI are based. Aside from its implications for NNMI, however, an understanding of the behavior and relevant forcing of these modes is of interest in its own right given the uncertainties of modeling many of the important processes in the tropical atmosphere and their impact on the general circulation (e.g., Tiedtke 1984; Donner et al. 1982). In addition, future forms of physical initialization, where, for example, improved satellite-observing capabilities may be used to augment initial heating and moisture fields in numerical models (e.g., Krishnamurti et al. 1991), will benefit from an understanding of these processes at this level of detail.

We focus our analysis on the medium-depth zonally symmetric modes since they have a clear physical analogy with respect to the atmospheric Hadley circulation. In addition, there is some evidence that they exhibit a higher degree of balance than other medium-depth internal modes (Errico and Rasch 1988). A simple prognostic equation is used to interpret the behavior of the modes in the forecast model, with particular attention given to investigating 1) the extent to which they obey the balance constraint imposed by NNMI, 2) whether the predominant forcing is adiabatic or diabatic, and 3) the impact of imposing inappropriate initial conditions. Initial conditions are derived using conventional adiabatic NNMI as well as an independent method based on a least-squares fit to the mode trajectories in the forecast model. The latter serves as a de facto diabatic initialization scheme. Also, an extended-range model integration is used to determine the existence of steady solutions for these modes. The results of the various forecast experiments are compared to assess the nature and magnitude of the errors introduced by the initialization schemes.

A description of the NOGAPS model and its normal-mode solutions is given in section 2. In section 3, we examine the principles upon which NNMI is based in terms of a simple prognostic equation. In section 4, we examine the temporal behavior of selected gravitational modes in the forecast model and interpret their behavior in terms of this simple equation. In section 5, we examine the impact of different initialization strategies on these modes, paying particular attention to how their responses relate to the various forces acting on them. The effects of specific diabatic forcings are examined in section 6. Conclusions are presented in section 7.

2. Description of model and normal modes

The model used in this study is version 3.2 of NOGAPS, which is a complete real-time forecast system consisting of an optimal interpolation (OI) analysis scheme (Barker et al. 1989), a nonlinear normal-mode initialization (NNMI) scheme, and a global forecast model. The forecast model is spectral in both horizontal dimensions with a triangular truncation at wavenumber 79 (179 truncation), which corresponds to a transform grid with 240 equally spaced longitudinal points and 120 Gaussian latitudes (roughly, 1.5° horizontal resolution). It has 18 vertical levels defined in terms of σ coordinates. The time-integration scheme is a centered semi-implicit scheme (Kwizak and Robert 1971) with a Robert (1966) time filter. A biharmonic (\(V^4\)) spatial filter is employed to smooth extremely small-scale features.

NOGAPS contains an extensive array of parameterized diabatic processes including gravitational wave drag, vertical fluxes of moisture, heat and momentum, cumulus convection, large-scale precipitation, shallow cumulus mixing, and radiation. Of particular note is the cumulus convection scheme, which is a modified version of the Arakawa–Schubert (A–S) scheme (Arakawa and Schubert 1974; Lord et al. 1982), and the vertical flux calculations, which are based on the K-theory formulation of Louis (1979). The A–S scheme used in NOGAPS is more comprehensive than the
convection schemes used in many other large global forecast models and includes the effects of downdrafts, latent heat of fusion, and evaporation of falling precipitation. Ground temperature is predicted over land, while the SST remains fixed at its initial values. The latter are obtained from daily analyses produced by Fleet Numerical Oceanography Center (FNOC).

The normal modes are derived from the adiabatic model equations linearized about a basic state at rest with horizontally uniform surface pressure and temperature fields denoted by $\bar{p}$, and $\bar{T}(\sigma)$, respectively. The modes describe the spectrally truncated fields of vorticity, $\xi$, divergence, $D$, and a form of the geopotential $\Phi$ that combines the effects of pressure and temperature variations (Hogan et al. 1991). The horizontal and vertical structures are determined separately as eigenvectors of the linearized system. The vertical eigenvectors represent the vertical structures of the modes, while the corresponding eigenvalues are the equivalent depths $h$ in a series of shallow-water equations whose solutions, in turn, describe the horizontal structures of the modes. There are 18 vertical modes (corresponding to the number of model levels) numbered $l = 1, \ldots, 18$, in which $l = 1$ denotes the gravest, or external mode. The values of $\hat{T}(\sigma)$ (or $\hat{D}$) and $h_l$ used in this study are listed in Table 1, where $\hat{T}(l)$ are obtained from an extended integration of the model.

Associated with each mode is a natural frequency determined as an eigenvalue of the horizontal system, and a time-dependent amplitude coefficient that appears as the coefficient in a linear expansion of the model fields in terms of the normal-mode structures. The modes are separable into three distinct types, which, in the linearized model, correspond to high-frequency eastward- and westward-propagating inertial–gravitational modes, and low-frequency westward-propagating rotational modes (Matsuno 1966). For each equivalent depth (or vertical mode $l$), there are 3240 modes of each type. Individual modes of any one type are uniquely identified by the triplet of indices $(l, m, n)$, where $m$ denotes the mode's zonal wavenumber ($m = 0, \ldots, 79$), and $n$ denotes an ordering from largest ($n = 0$) to smallest ($n = 79$) meridional scale. It should be noted that those modes corresponding to $(l, 0, 0)$ are of no physical interest in the present study and are thus neglected. Further details are given in Hogan et al. (1991).

The model data examined in this study were obtained from a series of 48-h forecasts beginning with the NOGAPS analysis fields from 1200 UTC 12 October 1989. The appropriate dynamic fields were archived every other 900-s model time step, or once every half hour during the forecast. The coefficients for each mode were then determined by projection of these fields (Hogan et al. 1991), resulting in a half-hourly time series of coefficients (97 complex values including the initial values at $t = 0$) for each mode for each forecast. For comparison, coefficients were also computed for a 48-h period starting at day 20 of an extended model integration with the noninitialized initial conditions from 12 October 1989.

3. Gravitational-mode dynamics and nonlinear balance

The prognostic equation for a single gravitational mode can be written

$$\frac{dg}{dt} = -(i\lambda + \nu)g + N(t),$$

where $g$ is the mode coefficient, $t$ is time, $i = \sqrt{-1}$, $\lambda$ is the mode’s natural frequency, $\nu$ is a linear damping coefficient, and $N$ denotes all other processes acting on $g$. In particular, $N$ includes nonlinear advection and all diabatic forcing except that portion of the diffusion that is linear in $g$. As pointed out by Errico (1989), (1) is appropriate since we wish to examine only the explicit relationships between individual modes and the forces acting on them, and not the implicit feedback of a mode upon its forcing.

If we consider a discretized model with harmonic forcing of the form

$$N(t) = \sum_k \mathcal{N}_k e^{-i\mu_k t},$$

where $\mathcal{N}_k$ is a time-independent amplitude and $\mu_0 = 0$, then the general solution to (1) is that of a simple forced and damped harmonic oscillator

$$g(t) = [g(0) - \sum_k R_k \mathcal{N}_k] e^{-(\lambda + \nu)t} + \sum_k R_k \mathcal{N}_k e^{-i\mu_k t},$$

where $R_{\lambda + \nu}$ is the exponential kernel with a complex argument $\lambda + \nu$.

**Table 1.** Values of $\hat{T}_l$ calculated from a model control run, and the corresponding set of equivalent depths $h_l$. Note that $l$ is used as a model-level index for $\sigma$ and $\hat{T}$, but refers to the ordering of the vertical modes in the case of $h_l$.

<table>
<thead>
<tr>
<th>$l$</th>
<th>$\sigma$</th>
<th>$\hat{T}_l$(K)</th>
<th>$h_l$(m)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>0.008</td>
<td>229.25</td>
<td>9669.53</td>
</tr>
<tr>
<td>2</td>
<td>0.028</td>
<td>209.14</td>
<td>3224.16</td>
</tr>
<tr>
<td>3</td>
<td>0.049</td>
<td>198.15</td>
<td>801.71</td>
</tr>
<tr>
<td>4</td>
<td>0.092</td>
<td>199.12</td>
<td>308.66</td>
</tr>
<tr>
<td>5</td>
<td>0.136</td>
<td>214.05</td>
<td>130.55</td>
</tr>
<tr>
<td>6</td>
<td>0.187</td>
<td>217.82</td>
<td>63.85</td>
</tr>
<tr>
<td>7</td>
<td>0.247</td>
<td>227.33</td>
<td>40.60</td>
</tr>
<tr>
<td>8</td>
<td>0.315</td>
<td>236.96</td>
<td>21.30</td>
</tr>
<tr>
<td>9</td>
<td>0.393</td>
<td>247.70</td>
<td>11.48</td>
</tr>
<tr>
<td>10</td>
<td>0.482</td>
<td>258.23</td>
<td>6.63</td>
</tr>
<tr>
<td>11</td>
<td>0.575</td>
<td>266.92</td>
<td>3.99</td>
</tr>
<tr>
<td>12</td>
<td>0.669</td>
<td>273.94</td>
<td>2.36</td>
</tr>
<tr>
<td>13</td>
<td>0.759</td>
<td>279.79</td>
<td>1.45</td>
</tr>
<tr>
<td>14</td>
<td>0.839</td>
<td>283.94</td>
<td>0.75</td>
</tr>
<tr>
<td>15</td>
<td>0.904</td>
<td>287.66</td>
<td>0.40</td>
</tr>
<tr>
<td>16</td>
<td>0.950</td>
<td>290.40</td>
<td>0.18</td>
</tr>
<tr>
<td>17</td>
<td>0.978</td>
<td>291.83</td>
<td>0.10</td>
</tr>
<tr>
<td>18</td>
<td>0.995</td>
<td>292.08</td>
<td>0.02</td>
</tr>
</tbody>
</table>
in which $R_k = (i\lambda - i\mu_k + \nu)^{-1}$ is a response function. We may think of (3) as having two components. The first component, which includes the first two terms on the right side of (3), has the frequency of the free linear wave, $\lambda$, while the second component has the frequency $\mu_k$. The magnitude of the former depends on both the forcing and the initial conditions. Its temporal behavior is that of a damped wave with initial amplitude $A = \|g(0) - \sum_k R_k N_k\|$ and e-folding period $1/\nu$. The magnitude of the latter, or forced, component depends on the forcing and the difference between the linear and forcing frequencies. The temporal behavior of this component may be considerably more complex than that of a simple wave.

Machenhauer (1977) noted that, for some modes, over short inertial periods, $N$ tends to vary slowly compared to the linear forcing acting on $g$ so that the two components of (3) are also referred to as the “fast” and “slow” components of the solution, respectively. A mode is said to behave in a balanced manner when its evolution is dominated by the latter component with $\|\mu\| \ll \|\lambda\|$ (such scaling implies $i\lambda g \approx N$). The principle of NNMI is to eliminate the fast component by setting

$$g(0) = \sum_k R_k N_k,$$

(initially with $R_k$ approximated by $(i\lambda)^{-1}$). Clearly, the success of this method requires an accurate assessment of all the forcing acting on $g$ and the validity of the approximation for $R_k$. If the forcing used to determine $g(0)$ is inappropriate or $R_k$ is approximated inaccurately, then $\|g(0) - \sum_k R_k N_k\|$ may be large with respect to the actual model forcing. In that case, the fast component will have unrealistically large amplitude during the early stages of the forecast and the forecast will appear “noisy.”

It is easily shown that, for stationary forcing ($\mu = 0$), (4) is equivalent to the condition

$$\frac{dg}{dt} = 0 \quad \text{at} \quad t = 0. \quad (5)$$

Equation (5) is the well-known Machenhauer balance condition used in most NNMI schemes. In practice, this condition is an appropriate one provided $\|\mu\| \ll \|\lambda\|$. Thus, the temporal behavior of $N$ is a key factor in determining whether a mode is a suitable candidate for initialization.

Despite their simple forms, (1)–(3) do, in fact, describe the behavior of many gravitational modes in models (Errico and Williamson 1988), and, in particular, are useful for understanding how different initial conditions affect the temporal behavior of the modes.

4. Modeled behavior of representative modes

As a means of characterizing the balance properties of the model’s Hadley circulation, we examined the temporal behavior of selected modes during an extended (22-day) simulation beginning with the NOGAPS analysis fields on 12 October 1989. In particular, we examined the set of zonally symmetric (zonal wavenumber $m = 0$) gravitational modes corresponding to vertical modes $4 \leq l \leq 6$ and meridional indices $1 \leq n \leq 20$. These modes have structures that describe the model Hadley circulation, as demonstrated by Puri (1983) and Errico and Rasch (1988), and illustrated here in Fig. 1, which shows a meridional cross section of the mass streamfunction determined by the $l = 4$, $m = 0$, $n = 3$ gravitational mode. Hereafter, we will refer to these as Hadley modes of vertical mode $l$ and meridional index $n$. (Note, however, that, since $\psi \sim \int \bar{V} \, dp$, where $\psi$ is the streamfunction and $\bar{V}$ is the normal-mode form of the zonally averaged meridional velocity, the vertical structure in Fig. 1 is actually that of the vertically integrated structure of the mode, not the vertical structure of the mode itself.) We include modes with somewhat smaller meridional scales than that shown in Fig. 1 (e.g., $10 < n < 20$) because they may be important components of the circulation in the nonlinear model.

Since the mode coefficients are complex, their temporal behavior is presented as harmonic dials in the complex plane, as shown, for example, in Fig. 2. The distance from the origin is the modulus of the mode coefficient, while the angle $\theta$ between the positive real axis and a line connecting the origin and a point on the trajectory is its phase. Thus, a mode behaving as a linear wave will appear as a circle. If the mode has a stationary component, then the center of the circle will

![Fig. 1. Meridional streamfunction component of the $l = 4$, $m = 0$, $n = 3$, gravitational mode. The contour interval is $2 \times 10^{16}$ kg s$^{-1}$.](image-url)
be at a point determined by the phase and amplitude of that component. For zonally symmetric modes, however, the concept of eastward and westward propagation is lost, so that a wavelike trajectory in these figures actually represents a standing meridional oscillation. Thus, instead of indicating the longitudinal positions of troughs, $\theta$ represents the phase of the standing oscillation.

In general, the results of the extended simulation revealed that the balance properties of the modes varied as a function of their natural periods, $P_\nu = 2\pi/\lambda$, and to a lesser extent, their meridional indices $n$. The harmonic dials of four modes are presented in Figs. 2a–d as being representative of the different types of behavior exhibited by this subset. For each mode, we show dials corresponding to two 48-h periods during the simulation. The first period (solid curve) corresponds to the 48 h beginning with the noninitialized initial conditions on 12 October 1989. The second period (dashed curve) corresponds to the 48 h beginning at day 20 of the simulation. During the latter period, the first component of (3) will be negligible for most modes due to

Fig. 2. Harmonic dials of selected Hadley modes during a 22-day forecast. Dials in each figure correspond to the 48 h beginning with the noninitialized initial conditions (solid curve) and the 48 h beginning at day 20 of the simulation (dashed curve). Asterisks (*) denote the mode's position at 12-h intervals after the start of the forecast and the end of each day is indicated.
the effects of damping. Thus, (3) should be well approximated by

$$g(t) = \sum_{k} R_k N_k e^{-t\mu_k}$$

(6)

at this time. The scaling of the axes in Fig. 2 is mode dependent, as indicated in each plot, where the units of amplitude are such that a squared value represents the total energy per unit mass of the mode. The asterisks (*) mark the mode's position at 12-h intervals after the start of the forecast.

Figure 2a shows the harmonic dials for the $l = 4$, $n = 1$ (largest scale) Hadley mode. The mode's trajectory during the two periods differs markedly. Early in the simulation (solid curve), the trajectory has a prominent oscillatory component, especially during the first 24 h, and evolves much more rapidly than the trajectory after 20 days (dashed curve). Note that after the first 24 h, the trajectory has roughly traced an arc slightly larger than that of a semicircle. The implied period is therefore close to the mode's natural period of 48.8 h, indicating that the mode has a significant, fast, or freely oscillating, component during this time. A more slowly evolving forced component is also evident during this time. This component appears to travel more or less parallel to the imaginary axis toward zero, as indicated by the difference in the mode's position at $t = 0$ and $t = 48$ h. After 20 days, the oscillatory component is much less evident and the trajectory evolves in a slow, irregular manner indicative of a high degree of balance. The differences in the two dials may be interpreted in terms of (3). Initially, $\|g(0) - \sum_{k} R_k N_k\|$ must be nonzero so that the fast component has significant amplitude during the early stages of the simulation. After 20 days, the fast component has damped out, leaving only the slowly evolving component. The trajectory is well described by (6) during the latter period, indicating that there is no significant fast forcing acting on this mode. Most of the longest period ($P_c > 35$ h), largest meridional-scale ($n \leq 3$) modes examined in this study also exhibited high degrees of balance.

The dials for the $l = 4$, $n = 3$ Hadley mode are shown in Fig. 2b. In contrast with Fig. 2a, the trajectories in this case are remarkably similar during both time periods. Note that the mode has two distinct components. The first component is clearly oscillatory and is characterized by the series of counterclockwise loops. The period of this component is close to 24 h, which is slightly longer than the mode's natural period of 22.4 h. The second component, whose path is marked by the centers of successive loops made by the oscillating component, evolves considerably more slowly than the first and exhibits a substantial degree of stationarity. The magnitude of this component (equal to its distance from the origin) is comparable with, but somewhat larger than, the amplitude of the oscillatory component (equal to the radii of the loops). After 20 days, the oscillation shows no evidence of being damped, indicating that there is continuous, significant forcing with period close to 24 h acting on this mode. Since this is close to the mode's natural period, this may be a near-resonant response to some diurnal forcing. In terms of (6), this would imply that $\|\lambda - \mu_k\|$ is small so that the response function $R_k$ (and thus, the amplitude of the response) is large. We examine this possibility in further detail in section 6. Several of the modes examined whose natural periods were close to the diurnal period exhibited similar behavior. These modes clearly behave in an unbalanced manner.

Figures 2c and 2d show examples of two types of temporal behavior exhibited by those modes with relatively short periods and small meridional scales. Figure 2c shows the dials for $l = 4$, $n = 8$ Hadley mode. Its behavior is generally that of a damped oscillation during the first 48 h of the forecast. The period of the oscillation is close to the mode's natural period of 13.6 h, indicating that the first component in (3) is dominant during this time. Note that the damped spiral is not concentric, which indicates that there is a second, slowly varying component. After 48 h, the oscillation has damped to the point that its amplitude is comparable to that of the slowly varying component. After 20 days, the trajectory is dominated by a large stationary component indicating that the predominant forcing is stationary. Although there is also a weakly forced oscillatory response at this time (note that the period of the oscillation is substantially longer than the mode's natural period), the mode clearly exhibits a high degree of balance.

The behavior of the $l = 5$, $n = 14$ Hadley mode in Fig. 2d is considerably less regular than that shown in Fig. 2c. There is a clearly discernible free oscillation during the first 48 h—as manifested by the series of loops with periods close to the mode's natural period of 12.8 h—but the trajectory bears little resemblance to the damped spiral in Fig. 2c. After 20 days, this oscillation has damped out, but the trajectory is still irregular, indicating that the effective forcing for this mode may have many frequencies. Moreover, the mode's response to this forcing cannot be characterized as slow or balanced. The behavior types shown in Figs. 2c and 2d were both common among those modes with relatively small meridional scales and short natural periods.

5. Appropriate use of initial conditions

The results in section 4 demonstrate that the zonally symmetric Hadley modes behave in a variety of ways, ranging from free oscillations to forced motions with complex temporal characteristics. Only some of these modes, particularly those whose long-term behavior can be considered slow, may be suitable candidates for initialization. Their evolution is governed by a near balance between linear and nonlinear forcing, which
may or may not include diabatic heating. For the remaining modes, the effect of imposing any such balance condition is unclear, but is of considerable interest since most NWP models use only crude selection criteria (either by vertical mode, natural frequency, or both) for applying NNMI. In this section, we examine the impact of different initialization strategies on the behavior of the modes.

a. Initialization schemes

A comparison of two initialization schemes was made for simulations beginning with the NOGAPS analysis fields from 1200 UTC 12 October 1989. For the first scheme, we used a standard adiabatic NNMI scheme based on an iterative solution of (4), where \( N_k \) includes only adiabatic (i.e., advective) nonlinear forcing. Solutions to (4) were calculated using two full iterations of Machenhauer’s (1977) scheme after an underrelaxed initial iteration. Additional iterations yield only insignificant changes in the solutions. We initialize those gravitational modes corresponding to \( l \leq 3 \) and \( P_e \leq 24 \) h (these are the vertical mode and frequency cutoff values used operationally in NOGAPS), plus all zonally symmetric gravitational modes corresponding to \( 4 \leq l \leq 6 \). The latter group includes the modes of interest in this study. Hereafter, we refer to this as the ANMI (adiabatic nonlinear normal-mode initialization) scheme.

For the second scheme, we developed a type of diabatic initialization based on a least-squares fit to the modes’ trajectories produced from a model simulation. With this method, initial conditions were obtained from parameters determined by minimizing the quantity

\[
E = \| f(t) - \hat{g}(t) \|^2, \tag{7}
\]

in which \( \hat{g}(t) \) is the mode’s trajectory obtained from a noninitialized model run, and \( f(t) \) is a simple wave function of the form

\[
f(t) = c + r e^{i \omega t}, \tag{8}
\]

where \( c \) is the stationary point of the wave (the center of a circle on a harmonic dial), \( r \) is the wave amplitude (the radius vector of the circle), and \( \omega \) is the wave frequency. For each mode, values of \( c \) and \( r \) were computed for a range of wave values by fitting \( f(t) \) to \( \hat{g}(t) \) during the early stages of the noninitialized simulation. From these, we retained the set of parameters in (8) that minimized (7). The time interval fitted was chosen to be half the mode’s natural period (i.e., \( \pi / \lambda \)) beginning at \( t = 0 \), since this portion of the noninitialized trajectory was usually well described by the linear wave component in (3). Since the model includes diabatic forcing, the value of \( c \) thus obtained accounts for both adiabatic and diabatic contributions to the stationary component of the wave. The simulations were then rerun using the same initialization configuration as in the ANMI scheme for \( l \leq 3 \), plus initial conditions \( \dot{g}(0) = c \) as a form of diabatic initialization for all zonally symmetric modes corresponding to \( 4 \leq l \leq 6 \). Hereafter, we refer to this as the DFIT (diabatic arc fit) scheme.

It should be noted that setting \( \dot{g}(0) = c \), where \( c \) is sufficiently close to the actual stationary solution of \( \dot{g} \), is synonymous, philosophically, with applying the Machenhauer balance condition (5). That is, in both cases the model’s initial conditions are defined by the stationary component of the forcing. Accordingly, it was found that the arc-fit and NNMI methods yielded very similar initial conditions when applied to external or deep internal gravitational modes \( (l \leq 2) \), which have large, adiabatic stationary components (Errico et al. 1988). Difficulties arise when the arc-fit technique is applied to modes that behave in a highly irregular fashion since the trajectories of these modes cannot be easily fitted using the simple wave function (8). This is not a weakness of this method, but rather a somewhat tangible indication that the balance condition (5) may be inappropriate for such modes.

b. Response to initial conditions

Figures 3a–d show the impact of the ANMI and DFIT initialization schemes on the four Hadley modes discussed in section 4. The format and scaling in these figures are the same as in Figs. 2a–d, except that here the dials correspond to 48-h forecasts with initial conditions obtained from the ANMI (solid curve) and DFIT (dashed curve) initialization schemes. For reference, the point marked \( U \) in these figures indicates the noninitialized (analyzed) initial condition from Fig. 2.

Figure 3a shows the corresponding harmonic dials for the \( l = 4, n = 1 \) Hadley mode. The initial conditions and subsequent trajectories differ markedly for each scheme. The ANMI trajectory has a much stronger oscillatory component and evolves more rapidly than the DFIT trajectory, especially during the early stages of the forecast. Again, the differences in the trajectories may be interpreted in terms of (3). When adiabatic forcing alone is considered, \( \| \dot{g}(0) - \sum_k R_k N_k \| \) must be large since the fast component dominates the ANMI trajectory. Note that the ANMI trajectory is very regular, especially during the first part of the forecast in which the trajectory traces out a semicircle after 24 h, or half the mode’s natural period. In contrast, the DFIT trajectory is much slower and much less regular, indicating that \( \| \dot{g}(0) - \sum_k R_k N_k \| \) is relatively small and the slow (forced) component dominates the solution. The DFIT trajectory appears to show a substantial degree of balance, indicating that diabatic initialization may be appropriate for this mode. Also, note that the noninitialized initial condition is close to the DFIT initial condition, which indicates that the analysis of this mode is fairly well balanced in this case. Comparing
the DFIT trajectory with the noninitialized one for the corresponding period (solid curve in Fig. 2a), we see that diabatic initialization slightly reduces the amplitude of the oscillation exhibited by the noninitialized trajectory. The DFIT scheme was generally successful when applied to all the longest period, large meridional-scale modes examined in this study. Thus, these modes appear to obey balance relationships in which diabatic forcing is important. We examine the source of this forcing in more detail in section 6.

Figure 3b is similar to Fig. 3a, except for the $l = 4$, $n = 3$ Hadley mode. Again, the DFIT scheme provides a more reasonable estimate of the mode's stationary component, indicating that diabatic forcing is important. Note, for example, that the DFIT initial condition lies at the approximate center (i.e., stationary point) of the semicircle traced out by the ANMI trajectory during its first 24 h. Subsequently, the amplitude of the oscillation in the DFIT trajectory is reduced considerably during the early stages of the forecast. In this case, however, the DFIT trajectory adjusts back toward an oscillatory state fairly rapidly, and, after 48 h, its phase and amplitude are again comparable with those of the ANMI trajectory. The implication of this result
is that the oscillating component is a forced response rather than a damped oscillation produced when $\|g(0) - \sum_k R_k N_k\|$ in (3) is large. This is consistent with the results in Fig. 2b, which showed that the oscillatory component of this mode shows no evidence of damping after more than 20 days. It is thus important to note that while the stationary component of this mode appears to be predominantly diabatic, the use of a diabatic balance condition is no more appropriate than an adiabatic one since the mode's behavior is not slow in general. The evolution of the DFIT trajectory in Fig. 3b may be interpreted as a form of spinup response, in which the model forcing acts to restore the mode's amplitude after its suppression due to improper initial conditions. The balance condition (5) was found to be generally inappropriate for modes with similar natural periods.

In Figs. 3c and 3d, we show the dials for the $l = 4, n = 8$ and $l = 5, n = 14$ Hadley modes, respectively. In both cases, the ANMI and DFIT schemes yield similar initial conditions (relative to the noninitialized ones) and the resulting trajectories are nearly identical. Thus, the stationary components for these modes appear to be predominantly adiabatic. For the $l = 4, n = 8$ Hadley mode, the initialized trajectories in Fig. 3c still appear noisy, but have much less amplitude than the corresponding noninitialized trajectory (solid curve in Fig. 2c). The period of the oscillation in these trajectories is between 16 and 18 h, which is considerably longer than the mode's natural period of 13.6 h and roughly the same as that exhibited by the weak oscillatory component present in the noninitialized forecast after 20 days (dashed curve in Fig. 2c). It appears then that the initialization schemes have successfully diminished the damped oscillatory component of this mode, which was clearly evident during the first 48 h of the noninitialized forecast. This success is not surprising based on the high degree of balance exhibited by this mode after 20 days. The remaining oscillation in the initialized trajectories in Fig. 3c may be due to nonlinear interactions with other gravitational modes (recall that only the zonally symmetric modes have been initialized for this equivalent depth) that become smaller as the model approaches a balanced state. This is consistent with the fact that this component has considerable amplitude at first but then damps during the forecast. Further investigation, however, is required to confirm this.

The initialization schemes have little appreciable impact on the $l = 5, n = 14$ Hadley mode in Fig. 3d. Only during the first 12 h do the initialized trajectories differ significantly from the corresponding noninitialized one (solid curve in Fig. 3d). After this time, the phase and amplitude of the trajectories are quite similar in all three forecasts and exhibit little tendency toward balance. This result is not surprising given the mode’s highly unbalanced behavior even after 20 days (dashed curve in Fig. 2d).

The results of these experiments are summarized in Fig. 4, which shows the time-averaged total energy

$$\bar{E} = \frac{1}{T} \int_0^T \hat{g} \hat{g}^* dt,$$

(9)

for the 15 largest-scale $l = 4$ Hadley modes, where $T = 48$ h and an asterisk denotes a complex conjugate. Values of $\bar{E}$ are plotted for forecasts with the noninitialized, DFIT and ANMI initial conditions, respectively. For a given mode, the differences in $\bar{E}$ between forecasts approximately reflect the amplitude of the oscillation (noise) produced by errors in their initial conditions. (For different modes, the relative magnitudes of $\bar{E}$ reflect the energy contributions from their stationary components as well.) For those modes with the largest meridional scales ($n \leq 3$), the DFIT trajectories have significantly less noise than the ANMI ones, indicating that the diabatic initial conditions are more appropriate. The energy in these modes increases by 20%-35% (i.e., they behave more noisily) when the ANMI initial conditions are used. Note, however, that for $n = 1$ and 2, the energy of the noninitialized trajectories is comparable to that of the DFIT trajectories, again indicating that these modes are well balanced in the NOGAPS analysis. For $n > 3$, there appear to be no significant differences between the DFIT and ANMI initial conditions. Adiabatic initialization may be appropriate for some of the smaller-scale modes (as demonstrated in Fig. 3c), but they tend to be much less energetic than those modes with larger meridional structures. Similar results were obtained for vertical modes 5 and 6.

6. Diabatic-forcing responses

a. Stationary forcing

As a means of identifying specific sources of stationary heating for the largest-scale ($n \leq 4$) zonally symmetric modes, we reran the initialization experi-

![Fig. 4. Time-mean energy $\bar{E}$ for the 15 largest-scale $l = 4$ Hadley modes for 48-h forecasts beginning with the noninitialized DFIT and ANMI initial conditions. The ordinate scale is $10^{-2}$ J kg$^{-1}$. Differences in $\bar{E}$ between forecasts indicate noise introduced by initial conditions.](image-url)
ments described in section 5 with selected diabatic parameterizations excluded from the model forecasts. Figure 5 shows the harmonic dials for the \( l = 4, n = 1 \) Hadley mode for two forecasts with identical adiabatic initial conditions (the ANMI initial conditions shown in Fig. 3a), but run without radiation (solid curve), and without convective heating (dashed curve). The trajectory without radiation is very similar to that obtained in the full diabatic run (solid curve in Fig. 3a), especially during the early stages of the forecast. As in Fig. 3a, the fast component dominates this trajectory and the mode's period is close to its natural period. Since, in the absence of radiative forcing, the adiabatic initial condition remains inappropriate (i.e., \( \| \mathbf{g}(0) - \sum \mathbf{R}_k \mathbf{N}_k \| \) is apparently still large), we can conclude that radiation contributes negligibly to the quasi-stationary diabatic forcing acting on this mode. In contrast, the trajectory without convective heating differs dramatically from the other two. This trajectory progresses in a slow, irregular manner, indicating that the slow (forced) component is dominant. In this case, the mode's stationary component is clearly different from the full diabatic run and the no-radiation run, and the adiabatic initial condition is appropriate (i.e., \( \| \mathbf{g}(0) - \sum \mathbf{R}_k \mathbf{N}_k \| \) is small). These results imply that convective heating is a major source of quasi-stationary diabatic forcing for this mode. Convective heating was by far the dominant source of quasi-stationary diabatic forcing for all of the large-scale zonally symmetric modes examined in this study. Although the convective heating also has high-frequency components, they were found to be of much less importance for these modes.

**Fig. 5.** Harmonic dials for the \( l = 4, n = 1 \) Hadley mode for forecasts without radiation (solid curve) and without convection (dashed curve), run from identical ANMI initial conditions used in Fig. 3a. Scaling and format are the same as in that figure.

**Fig. 6.** Harmonic dials for the \( l = 4, n = 3 \) Hadley mode for the full physics forecast (solid curve) and a forecast without radiation (dashed curve), run from identical DFTI initial conditions used in Fig. 3b. Scaling and format are the same as in that figure.

Modes with smaller meridional scales (\( n > 8 \)) and shorter natural periods (\( P_c < 20 \) h) exhibited no tendency toward balance when diabatic processes were excluded from the model forecasts. This is consistent with the results obtained in section 5 and further demonstrates that the predominant forcing for these modes is adiabatic.

The results in Fig. 5 are consistent with those of Puri (1983), who investigated the relationship between convective adjustment and the strength of the Hadley circulation in the Australian Numerical Meteorology Research Centre’s (ANMRC) spectral model. He found that convective processes strongly influenced the lowest frequency, zonal wavenumber 0 modes corresponding to vertical mode 4 (equivalent depth of 256 m), and to a lesser degree, vertical mode 5 (equivalent depth of 105 m), and that these modes dominated the Hadley circulation in that model.

### b. Resonant responses

The effect of radiation on those modes whose periods are close to the diurnal period is illustrated in Fig. 6. In this case, we show the harmonic dials for the \( l = 4, n = 3 \) Hadley mode (\( P_c = 22.4 \) h) for two forecasts with identical diabatic initial conditions (the DFTI initial conditions shown in Fig. 3a) but run with all diabatic processes included (solid curve; this is the same as the DFTI forecast shown in Fig. 3b), and with all diabatic processes except radiation (dashed curve). As shown earlier, the DFTI scheme provides an accurate estimate of the stationary component of this mode. In
this case, however, the forecast trajectory without radiation continues to behave in a well-balanced manner throughout the forecast rather than adjust back toward a noisy state. The oscillatory component in this forecast continues to damp and has a period close to the mode's natural period. This is in contrast to the 24-h period clearly evident at all times in Fig. 2b, for example, and strongly suggests that radiation acts as a direct or indirect source of resonant forcing for this mode. Finally, it should be noted that the phase and amplitude of the quasi-stationary component of this mode are the same in both forecasts since both include convective forcing. This is consistent with the results in section 6a.

7. Summary and conclusions

The balance properties of the zonal wavenumber 0 Hadley circulation in the NOGAPS forecast model were investigated in terms of the behavior of selected inertial–gravitational normal modes. The modes were defined as the free solutions to the model linearized about a state of rest with a horizontally invariant vertical temperature profile. Time series of the 20 largest-scale, zonally symmetric modes were examined for three equivalent depths. The time series were presented as harmonic dials in the complex plane and the results interpreted in terms of a simple prognostic equation for a single gravitational mode. The data for these time series were provided by a series of 48-h forecasts using different initialization strategies, including conventional diabatic NNMI (referred to as the ANMI scheme) as well as a form of diabatic initialization based on a least-squares fit to the mode trajectories (referred to as the DFIT scheme). The latter provided a straightforward means of accounting for all forms of stationary forcing acting on the modes. Additional data were obtained from an extended model integration in order to examine the long-term behavior of the modes. These results were compared with those obtained from the initialized forecasts. In general, it was found that the balance characteristics of the modes varied as a function of their natural periods \( P_e \) and to a lesser extent, their meridional indices \( n \). Since only a limited number of basic behavior types was found, it was possible to present a fairly comprehensive overview of the results by examining several representative examples.

Only those modes with the longest temporal scales and largest meridional scales (\( P_e > 35 \text{ h}, n \lesssim 3 \)) appear to obey a diabatic balance condition to any substantial degree. Their behavior can be characterized as quasistationary, owing to an apparent near balance between inertial–gravitational forcing and convective heating. Moreover, the stationary component of this heating is, in some cases, quite large. An interesting implication of this result is that, overall, tropical convection can be a strong source of stationary forcing despite its spatial and temporal variability at any given time. From the standpoint of initialization, it was shown that diabatic NNMI is strictly inappropriate for this subset of modes. In contrast, the DFIT diabatic-initialization scheme produced well-balanced initial conditions for these modes, although most forecasts with the DFIT initial conditions were found to be only slightly less noisy than those with the noninitialized (analyzed) initial conditions. Thus, while diabatic initialization is in some sense appropriate for these modes, they appeared to be well balanced in the NOGAPS data-assimilation system.

Those modes with periods \( 20 < P_e < 30 \text{ h} \) and medium-to-large meridional scales (\( 4 < n < 10 \)) behave like regular linear oscillations to a high degree. Typically, these modes have two distinct components characterized by an oscillatory component and a more slowly varying or quasi-stationary one with comparable magnitude. Since, generally speaking, these modes have natural periods close to the diurnal period, it seems likely that their oscillatory component may be a near-resonant response to some form of diurnal forcing. This conclusion is supported by the fact that radiative heating was shown to be the primary forcing mechanism for this response, and, in many cases, the observed periods of the oscillations were remarkably close to 24 h. In addition, those modes having natural periods closest to this value exhibited particularly strong oscillatory behavior. In contrast, the slowly varying components of these modes were found to be related primarily to convective heating. Although the DFIT initialization scheme was generally successful at finding this component of the solution, the initialized forecasts quickly adjusted back to an oscillatory (noisy) state, indicating that the Machenhauer balance condition is inappropriate for these modes.

The smallest-scale modes (\( P_e < 20 \text{ h}, n > 10 \)) exhibited both balanced and forced behavior. For most of these modes, the dominant forcing appears to be diabatic. Some of these modes may have large stationary components, in which case diabatic initialization may be appropriate. This, however, was not found to be true in general. Those modes whose behavior can be characterized as forced often exhibit several oscillatory components with various periods instead of a single dominant component as often exhibited by larger-scale modes. This is most likely explained by the fact that, at shallower equivalent depths and lower frequencies, advection can become more important than inertial–gravitational forcing, so that modes defined about a resting basic state may be poor approximations to the actual modes of the nonlinear model (Daley and Williamson 1985). In that case, the defined modes may have several components whose amplitudes depend on how the data projects onto the nonlinear modes, and, in turn, on how these modes project onto the defined ones (Errico and Williamson 1988).

For the most part, these results are consistent with those of Errico (1984) and Errico et al. (1988), who
found little or no evidence that medium-depth gravitational modes obey a diabatic balance condition in general. The present findings, however, particularly those demonstrating the balanced behavior of a limited number of the largest-scale zonal wavenumber 0 Hadley modes, are clearly an extension of those authors’ results, which were based primarily on root-mean-square measures of balance applied over fairly broad frequency bands. Nonetheless, the fundamental conclusions regarding the application of NNMI remain the same with either approach. Among the most important of these are that 1) NNMI in its present form is best suited for controlling high-frequency oscillations arising from inconsistencies between the dynamic constraints of an analysis scheme and those of a forecast model, 2) the balance constraint inherent in this procedure is appropriate only for those modes that would otherwise behave in a balanced manner, and 3) medium-depth internal gravitational modes, which are important in diabatically driven circulations, are, with very few exceptions, not well suited to this balance condition regardless of whether diabatic forcing is included; that is, their tendencies are not small compared with the forces acting on them.

Finally, the results presented here demonstrate that the behavior of most inertial-gravitational modes in a sophisticated numerical model can be explained in terms of the principles that govern simple gravitational-wave dynamics. For modes whose observed behavior deviates somewhat from this simple framework, such as those at increasing shallower equivalent depths, these principles remain useful for understanding fundamental concepts regarding balance, the appropriateness of initial conditions, and the theory upon which NNMI is based.

Acknowledgments. The authors wish to thank the anonymous reviewers for their helpful comments and suggestions. We also thank Mr. S. Bishop for his efforts in preparing the figures. Dr. R. Errico contributed to this study while working at NRL (formerly the Naval Oceanographic and Atmospheric Research Laboratory) as part of its UCAR Visiting Scientist Program. Funding was provided by the Office of Naval Technology, Program Element 62435N. Contribution 432: 028:92. Approved for public release; distribution is unlimited.

REFERENCES


