The Small Ice Cap Instability in Diffusive Climate Models

GERALD R. NORTH

Laboratory for Atmospheres, NASA/Goddard Space Flight Center, Greenbelt, MD 20771

(Manuscript received 21 June 1984, in final form 30 August 1984)

ABSTRACT

Simple climate models employing diffusive heat transport and ice cap albedo feedback have equilibrium solutions with no stable ice cap smaller than a certain finite size. For the usual parameters used in these models the minimum cap has a radius of about 20 degrees on a great circle. Although it is traditional to remove this peculiar feature by various *ad hoc* mechanisms, it is of interest because of its relevance to ice age theories. This paper explains why the phenomenon occurs in these models by solving them in a physically appealing way. If an ice-free solution has a thermal minimum and if the minimum temperature is just above the critical value for formation of ice, then the artificial addition of a patch of ice leads to a widespread depression of the temperature below the critical freezing temperature; therefore, a second stable solution will exist whose spatial extent is determined by the range of the influence function of a point sink of heat, due to the albedo shift in the patch. The range of influence is determined by the characteristic length in the problem which in turn is determined by the distance a heat anomaly can be displaced by random walk during the characteristic time scale for radiative relaxation; this length is typically 20–30 degrees on a great circle. Mathematical detail is provided as well as a discussion of why the various mechanisms previously introduced to eliminate the phenomenon work. Finally, a discussion of the relevance of these results to nature is presented.

1. Introduction

Since their introduction more than a decade ago by Budyko (1968, 1969) and Sellers (1969), energy balance climate models (EBMs) have been the subject of many investigations (cf. the review by North *et al.*, 1981). Even with their mathematical and physical simplicity, the models have continued to provide interesting puzzles. One curiosity which has eluded a simple interpretation over the years is the small ice cap instability (SICI). Early on, several authors found that in tuning models with discontinuous albedos at the ice cap edge and with a diffusive type of transport, ice caps smaller than a certain finite size are unstable. Most investigators tended to think that the SICI was an artifact of the idealized models and the usual approach was to dismiss it or introduce additional *ad hoc* mechanisms that would remove it. For example, Cahalan and North (1979) noted that if the ice cap albedo were smoothed by say an arctangent function SICI disappeared (cf. also Suarez and Held, 1979; Coakley, 1979; Ghil, 1976, who used the Sellers-type albedo parameterization). Lin (1978) showed that when certain nonlinear diffusive heat transport mechanisms were introduced the SICI disappeared. North (1975) noted that if the equinox solar insolation distribution were used, no SICI occurred. In short, many authors have considered the phenomenon a nuisance and while it has been shown to come and go with various assumptions little effort has been devoted to determining its physical significance. The purpose of this paper is to develop a simple interpretation for SICI which explains all of the preceding removal mechanisms and which provides enough understanding of it to ask if it is relevant to climatology.

The most popular removal mechanism has been the smoothed albedo at the ice edge because in dealing with zonally symmetric models the lack of longitudinal uniformity of the real earth's ice caps naturally suggested use of a smoothed zonal average in a simulation (e.g., Oerlemans and Van den Dool, 1978). Coakley (1979) argued similarly that the seasonal variation of the snow line justified use of a smoothed albedo at the perennial ice edge. Recently, a two dimensional EBM with longitudinal and seasonal resolution has been developed, however, whose solutions exhibit zonally asymmetrical ice caps (North *et al.*, 1983). The model employs a discontinuous albedo at the local ice borders (the published version unfortunately does not have a seasonal snow line). As parameters are varied in the model, a phenomenon reminiscent of SICI occurs in the solutions, namely, abrupt transitions from no ice cap to finite unsymmetrical ones. Moreover, transitions from finite unsymmetrical caps to larger finite unsymmetrical caps can occur as parameters are varied further. These model results suggest that the smoothing argument used in the earlier zonally averaged models was inapplicable.

In the two-dimensional model paper it was speculated that the Laurentide ice sheet may have been the result of the transition from a kind of no ice situation (actually Greenland only) to a solution having a finite stable ice cap extending across the pole and well into North America. The transition could be triggered by tiny orbital element changes. The two-dimensional model also suggests that the formation of the Greenland ice sheet might have occurred by way of an abrupt transition induced by the opening of the Norwegian Sea about 50 million years ago (North and Crowley, 1985). Another application of the two-dimensional model is to the inception of glaciation of Antarctica (North and Crowley, 1985). In that case the model solutions imply that as the carbon dioxide loading of the air decreased from several times higher than at present during the mid-tertiary, there was an abrupt transition to an ice covered land mass. In the results just cited the influence of the seasonal cycle is crucial since it can lead to strong longitudinal variation in summer temperature distribution due to continentality effects. In fact, the thermal minimum in summer may be displaced from the pole and there may even be more than one relative minimum.

In addition to these model results which imply a role for SICI in the theory of the inception of land based ice sheets, various evidence implies that perennial sea ice may exhibit SICI behavior. Budyko (1974) has argued from surface heat balance considerations that if Arctic sea ice is removed it may not return; Hunt (1984), on the other hand, has argued from GCM calculations that an ice-free Arctic sea state would not exist under present conditions. The question of an ice-free arctic ocean has interesting paleoclimatic implications regarding the availability of moisture and its potential impact upon glacier growth rates; this and other related sea ice implications are reviewed by Clark (1982). These potentially important model results involving SICI type phenomena and their likely applicability to a variety of paleoclimatic situations suggest that the phenomenon of SICI deserves further investigation.

The intention of this study is to conduct a kind of thought experiment in which a solution is found to a simple EBM in a way that is intuitively appealing. The resulting picture of the SICI is clear. Before entering the mathematical detail we find it helpful to preview the argument in words. Imagine an ice free earth with a smooth temperature field falling from about 31°C near the equator to about −9.0°C at the poles. Actually this solution corresponds to the one studied by North (1975) and Drazin and Griffel (1977) for a solar constant/4 of 333.0 W m^{-2} and with other model parameters as in those papers. Now imagine placing a small patch of ice at the pole. The resulting albedo change serves as a point sink of heat. Diffusive transport and radiative damping (infrared to space) lead to a distortion of the temperature field over a spatial scale proportional to the square root of the ratio of the diffusion to radiative parameters (this length scale was noted by Lindzen and Farrell, 1977). The finite spatial extent of the thermal depression leads to a finite sized patch of horizontal area which falls below the critical temperature for the formation of ice (−10°C) and therefore a finite sized ice cap results. The size of the ice patch determines the strength of the heat sink and this can be adjusted self-consistently until the patch is in equilibrium, i.e., until its edge is at −10°C. Roughly speaking, the finite ice cap has a size approximately that of the length scale of the influence function.

Before passing to the body of the paper, we acknowledge the intuitive arguments of Brooks (1949) that small ice caps would not exist. His arguments are similar to those presented here. Of course, they predated the development of the simple EBM framework, which makes the argument much easier to understand. The present mathematical formulation also gives the reader a much clearer basis for accepting or rejecting the argument as it applies to nature.

2. Model

The class of models under consideration are the diffusive zonally symmetric mean annual models which may be defined by the energy balance equation:

$$-D\frac{d}{dx}(1-x^2)\frac{dT}{dx} + A + BT = QS(x)a(T), \quad (1)$$

where

x	sin(latitude)
T	sea level temperature
A, B	empirical radiation coefficients (201.4 W m^{-2}, 1.45 W m^{-2} C^{-1})
D	thermal diffusion coefficient ($\sqrt{D/B} = 0.30$)
$a(T)$	coalbedo = 0.38 for $T < -10$°C, 0.68 for $T > -10$°C
Q	solar constant/4
$S(x)$	normalized mean annual sunlight distribution ($=1 - 0.482P_2(x)$; $P_2(x) = (3x^2 - 1)/2$).

All lengths are in units of earth radius.

The preceding parameter values are taken from North (1975) and are the same as those used by Drazin and Griffel (1977). While these constants are not the most up-to-date, they can be used for easy comparisons with previous exact solutions and certainly serve our present purpose.

Equation (1), together with the boundary conditions that no heat flux cross the equator or go into the pole, can be solved by a variety of techniques. The usual form of displaying the solution is to plot x_s, the sine of the latitude of the ice edge as a function of solar constant. Figure 1 shows a portion of the

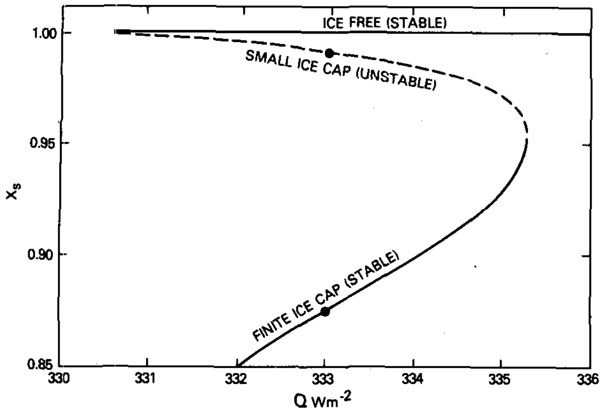

FIG. 1. A plot of the north–south symmetric equilibrium ice cap sine of latitude x_s vs solar constant/4, Q, adapted from Fig. 2 of Drazin and Griffel (1977); cf. also Fig. 1 of North (1975).

equilibrium curve as adapted from Drazin and Griffel for the north-south symmetric solutions. Note that the segments of the curve with negative slope (dashed) can be demonstrated to represent unstable solutions (Cahalan and North, 1979). According to this figure, values of the ice edge x_s greater than about 0.955 are unstable. Hence, for the solar parameter values in the range shown in the graph only ice caps extending more equatorward than about 73 degrees latitude are stable. This is what is meant by the small ice cap instability.

3. Perturbation of an exact ice-free solution

Now we wish to study the perturbation of the ice-free solution corresponding to a value of $Q = 333.0$ W m^2 due to a source or sink of heat uniformly distributed around the pole. This particular ice-free solution is the solution to a linear model since T is everywhere greater than $-10°$C. The solution is given by

$$T_F(x) = 17.3 - 26.3P_2(x), \tag{2}$$

which can be demonstrated by insertion into (1). We can find the perturbed solution due to adding a heat source whose density is $q(x)$. The shift in temperature, denoted by $T'(x)$, is the solution of a linear inhomogeneous ordinary differential equation with no-flux boundary conditions at the poles or at one pole and the equator if $q(x) = q(-x)$:

$$-D\frac{d}{dx}\left[(1-x^2)\frac{dT'}{dx}\right] + BT' = q(x). \tag{3}$$

The function T' will now be shown to have a length scale of order $\sqrt{D/B}$ if the length scale of q is small. First it is useful to digress to a short discussion of Green's functions for operators such as that in (3). To simplify the mathematics by using more familiar special functions we choose to work in a plane tangent to the pole, although the relevant spherical solution is available (see Appendix of North *et al.*, 1981, or Salmun *et al.*, 1980). In these plane polar coordinates let r be the distance from the pole. A response function $H(r)$ for the source distribution $q(r)$ is the solution of

$$\frac{1}{r}\frac{d}{dr}\left[r\frac{d}{dr}H(r)\right] - \frac{H(r)}{L} = -\frac{g(r)}{D},$$

where the length scale is defined as $L = \sqrt{D/B}$. Consider the Green's function $G(r, r')$ defined by

$$\frac{1}{r}\frac{d}{dr}\left[r\frac{d}{dr}G(r, r')\right] - \frac{G(r, r')}{L} = g\delta(r - r')/r'. \tag{5}$$

Here $G(r, r')$ may be interpreted as the response to a ring of heat source at radius r' and with strength g; $H(r)$ may be recovered for any source density $q(r')$ by the quadrature

$$H(r) = -\int_0^\infty r'q(r')G(r, r')dr'/D. \tag{6}$$

A solution for $G(r, r')$ is given by Jackson (1962):

$$G(r, r') = \begin{cases} gI_0(r'/L)K_0(r/L), & r > r' \\ gI_0(r/L)K_0(r'/L), & r < r', \end{cases} \tag{7}$$

where $I_0(z)$ and $K_0(z)$ are modified Bessel functions. For values of $r > r'$ it is clear that the dependence of the Green's function has the form of the familiar K_0 Bessel function, sketched for reference in Fig. 2. It might be recalled that asymptotically $K(z) \sim (\pi/2z)^{1/2}$ $\exp(-z)$ for large z. It follows that the physically relevant length scale of the thermal response to a point source of heat is indeed $\sqrt{D/B}$ which we defined as L. By approximating I_0 by unity for z small (cf.

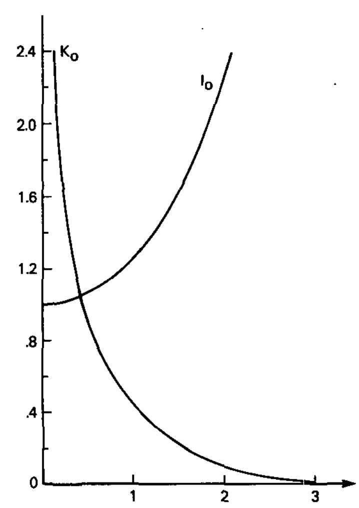

FIG. 2. The Modified Bessel functions $K_0(z)$ and $I_0(z)$.

Fig. 2), and cutting off the integral (6) at R the ice cap edge, we can obtain $H(r)$ from (6):

$$H(r) \approx gR^2K_0(r/L)/2. \tag{8}$$

The strength g of the anomalous heat source is determined by the product of the discontinuity of albedo $a = 0.30$ and the solar heating at the pole $QS(1) = 333.0 \times 0.518$ divided by the dimensional diffusion coefficient 1.45×0.30; therefore, $g = 396.5$. Note also that the thermal depression is proportional to the area of the ice cap, $R^2/2$ (a factor of 2π has been divided out of the energy balance equation).

Possible ice cap edge solutions can be found by adding the ice free solution $T_F(r)$ to (6) and equating the sum to the critical temperature $T_s = -10°C$:

$$T_F(r) + H(r, R) = T_s. \tag{9}$$

The roots of (9) for which $r = R$ are the possible ice edge solutions. By setting $r = R$ in (9), we may write the condition

$$T_F(R) + H(R, R) = T_s. \tag{10}$$

The solid curve in Fig. 3 contains a graph of the two sides of (10) with roots R determined by the intersections of the two curves. In converting from x to r (sphere to tangent plane) the approximate relation $r^2 = 1 - x^2$ was used. There are two roots, one very small ($R = 0.15$) and a larger one ($R = 0.28$). The smaller root corresponds to the small unstable solution at about $x_s = 0.985$ in Fig. 1, while the larger corresponds to the stable root at $x_s = 0.875$ of that figure. The vertical arrows in Fig. 3 indicate the roots that should be expected from Fig. 1. The dashed curve shows the improved approximation when the integrand containing I_0 and $S(x)$ is allowed to have r dependence up to terms of order r^2. In this case the resulting roots lie much closer to the exact solution. The residual error is the result of the distortion of the spherical versus plane geometry as well as the quality of this first approximation. As a final check, the spherical case was integrated numerically and both roots were essentially indistinguishable from the exact ones; the reason for the close agreement in this last case is that this solution is actually exact, and the procedure actually does represent yet another method of solving EBMs with precisely this form. The technique is not especially useful in other cases and hence we shall not dwell upon it here. We can conclude that this perturbation algorithm is valid as a device for interpreting SICI, since even the tangent plane approximation leads to roots rather close to the exact solutions for the ice cap edge.

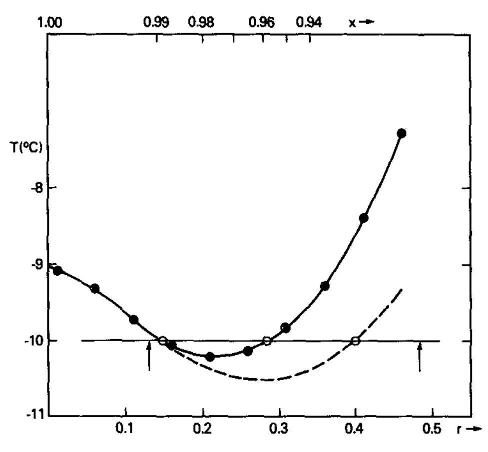

FIG. 3. Right and left sides of Eq. (10) vs argument r. The two smooth curves represent two different approximations to $H(r, r)$: the solid curve takes the ice-cap strength proportional to its area, whereas the dashed curve uses a higher order approximation. Intersections with the $-10°C$ line (open circles) give the planar model roots for ice caps. The vertical arrows show the exact roots from the spherical solution.

4. Remarks

After this simple derivation (summarized at the end of the Introduction) we can explain why the removal mechanisms work. Let us consider them one by one:

1) The smoothing mechanism introduces a new length scale, the smoothing length, which process diminishes the strength of the point sink; the net result is that the curve of Fig. 3 never goes below the critical temperature. While smoothing is effective in removing the cusp by diminishing the strength of a small patch of ice, it seems to persist under some rather drastic applications. For example, Coakley's (1979) seasonal smoothing mechanism does not always remove the cusp. This point is also made by Saltzman and Vernekar (1983).

2) If the diffusion coefficient is proportional to the gradient of temperature, then the local diffusion coefficient is vanishingly small near the pole because it is forced to be so by the no flux boundary condition at the pole. Since the local diffusion coefficient is forced to be so small near the pole an anomalous heat sink (ice patch) will only be able to spread (negative) heat a very short distance before it is radiated away to space. The length scale L is essentially zero. Lin's (1978) removal of the cusp by this mechanism is now clearly understandable. Its validity as regards nature is an open question centered around whether the atmosphere can support discontinuities in surface temperature. It is interesting that Saltzman and Vernekar (1983) find the cusp to be a very robust feature of their fairly detailed EBM which included the effects of mean meridional motions and latent heat transport.

3) Equinox heating distributions have $S(1)$ equal to zero and a small hypothetical patch of ice can have no heat sink value, since there is no incident

radiation to be affected. Therefore, we expect no cusp as was found by North (1975).

4) The climate model having no horizontal heat transport has no SICI as indicated by the figures in North *et al.* (1981); this is easily explained by the fact that the response function to a point source in such a model is a spike (delta function), and therefore no finite sized deformation is to be expected. Again the associated length scale of the influence function is zero.

5) The Budyko transport model (horizontal heat flux divergence proportional to the difference between local and global average temperatures) has a response (Green's) function which has strong spike and a weak long range part (cf., the Appendix of Cahalan and North, 1979, for its analytical form); in most cases the Budyko model does not exhibit the SICI because of the dominance of the spike in the influence function. It is important to realize that in these last two nondiffusive cases, where discontinuous temperature fields are permitted by the transport mechanism, the situation is more complicated because of the possibility of a continuous multiplicity of solutions, as has been studied by Held and Suarez (1974).

It is interesting that the characteristic length scale $\sqrt{D/B}$ is a number of the order of 0.40 to 0.60 earth radii (20 to 40 degrees on a great circle; usually the smaller value applies near the poles) depending upon the parameter choices (cf. North *et al.*, 1983, for more up-to-date values of D and B). This makes the length scale large for some applications such as the one discussed here in connection with the possibility of stable tiny ice caps. However, in some other applications we see that this same value for the length scale may be considered rather small. For instance, if we ask about whether the albedo of an ice cap at one pole can have much influence at the other, we see that the answer is no. For example, the effects of a cooling anomaly at one pole are diminished by a factor of the order of 10^{-4} of the values near the source. It is no wonder that Drazin and Griffel (1977) were able to find climate solutions with an ice cap at only one pole, quite independently of whether there was one at the other pole.

Now we come to the relevance of the inferred instability of small ice caps to the real world. First we must question the validity of the diffusive approximation in this context. Since the horizontal heat transport near the poles is dominated by transient eddy fluxes (Oort, 1974) it may be that at least in ensemble average, diffusion is a reasonable first approximation. The arguments presented here, of course, do not depend precisely upon diffusion but rather upon the random walk nature of eddy transport and its square root of time propagation characteristic. Recall that the cusp phenomenon persists even when some poloidal motion is present (Saltzman and Vernekar, 1983).

Some general circulation model results, that I am aware of, suggest that in the polar regions the model solution fields (in ensemble average) react to localized heating anomalies only over a distance of roughly the length scale introduced here. For example, Herman and Johnson (1978) in their investigation of the effects of sea ice anomalies found that the thermal effects were considerably diminished beyond a distance of about 20 degrees on a great circle. Hunt (1984) also suggests that the effects of local albedo changes near the poles are confined to latitudes poleward of 70 degrees. Similarly, Phillips and Semtner (1984) studied a simplified GCM's response to sea surface temperature anomalies and also found a similar limited but finite range of influence. Of course, as is apparent in the recent literature (e.g., Simmons *et al.*, 1983, or Phillips and Semtner, 1984) heat sources in the tropics can excite long stationary waves extending well into the midlatitudes, suggesting that diffusion would be a poor approximation in that case.

As alluded to in the Introduction, the SICI becomes much more interesting in the context of seasonal models including geography. In this case the morphology of the unperturbed field, as well as the influence function, will determine the precise shape of the smallest possible stable ice cap. If the health of nascent or thin ice caps is governed by summer temperatures, the placement of land masses can relocate the position of the thermal minimum away from the pole during seasonal extremes and therefore lead to asymmetrical SICI in some cases; presumably this is precisely what happened in the two-dimensional model of North *et al.* (1983). Small ice cap instability may then play some role in the formation of such large ice sheets as the Antarctic, Greenland, and the Laurentide. It would be premature to take such a theory too literally at present, however, since our model is highly schematized and the effects presenting themselves here as dramatic bifurcations may well be smooth but steep segments of the curves in a more realistic formulation. It is likely that the discontinuous albedo, diffusive transport and pinning the existence of perennial ice to a late summer isotherm all conspire to exaggerate the behavior. On the other hand, such simplified views can provide useful working hypotheses for experimentation with the more comprehensive models of the future.

The plausibility of SICI as inferred from diffusive models suggests that its existence in more advanced models be investigated. For general circulation model experiments this order is a tall one since these systems are not ordinarily at their best near the poles for a variety of reasons, both numerical and physical. In addition, if the small ice cap instability is indeed real in these more complicated models, it will be necessary to make extremely long runs to obtain equilibrium statistics because the characteristic time (equivalent to autocorrelation time usually) is proportional to the slope of the curve in Fig. 1, and at the bifurcation this slope is, of course, infinite. However, recent developments especially in spectral models and a continuing improvement in our collection and understanding of polar data suggest that this kind of detailed modeling of the polar regions is not far away.

Acknowledgments. I am indebted to Stephen Warren and Warren Wiscombe for pointing out the reference to Brooks. In addition, several friends at GLAS have patiently and politely listened to preliminary versions of this work; their comments have been helpful.

REFERENCES

Brooks, C. E. P., 1949: *Climate Through the Ages*. Dover, 307 pp.

Budyko, M. I., 1968: On the origin of glacial epochs. *Meteor. Gidrol.*, **2**, 3–8.

——, 1969: The effect of solar radiation variations on the climate of the earth. *Tellus*, **21**, 611–619.

——, 1974: *Climate and Life*, Academic Press, 573 pp.

Cahalan, R., and G. R. North, 1979: A stability theorem for energy-balance climate models. *J. Atmos. Sci.*, **36**, 1205–1216.

Clark, D. L., 1982: Origin, nature, and world climate effect of Arctic Ocean ice-cover. *Nature*, **300**, 321–325.

Coakley, J. A., 1979: A study of climate sensitivity using a simple energy balance model. *J. Atmos. Sci.*, **36**, 260–269.

Drazin, P. G., and D. H. Griffel, 1977: On the branching structure of diffusive climatological models. *J. Atmos. Sci.*, **34**, 1858–1867.

Ghil, M., 1976: Climate stability for a Sellers-type model. *J. Atmos. Sci.*, **33**, 3–20.

Held, I. M., and M. Suarez, 1974: Simple albedo feedback models of the icecaps. *Tellus*, **36**, 613–629.

Herman, G. F., and W. T. Johnson, 1978: The sensitivity of the general circulation to arctic sea ice boundaries: A numerical experiment. *Mon. Wea. Rev.*, **106**, 1649–1664.

Hunt, B. G., 1984: Polar glaciation and the genesis of the ice ages. *Nature*, **308**, 48–51.

Jackson, J. D., 1962: *Classical Electrodynamics*, Wiley & Sons, 641 pp.

Lin, C. A., 1978: The effect of nonlinear diffusive transport in a simple climate model. *J. Atmos. Sci.*, **35**, 337–340.

Lindzen, R. S., and B. Farrell, 1977: Some realistic modifications of simple climate model. *J. Atmos. Sci.*, **34**, 1487–1501.

North, G. R., 1975: Analytical solution to a simple climate model with diffusive heat transport. *J. Atmos. Sci.*, **32**, 1301–1307.

——, and T. J. Crowley, 1985: Application of a seasonal climate model to Cenozoic glaciation. *J. Geol. Soc. (London)*, issue 3 (in press).

——, R. F. Cahalan and J. A. Coakley, 1981: Energy balance climate models. *Rev. Geophys. Space Phys.*, **19**, 91–121.

——, J. G. Mengel and D. A. Short, 1983: Simple energy balance model resolving the seasons and the continents: Application to the astronomical theory of the ice ages. *J. Geophys. Res.*, **88**, 6576–6586.

Oerlemans, J., and H. M. van den Dool, 1978: Energy balance climate models: Stability experiments with a refined albedo and updated coefficients for infrared emission. *J. Atmos. Sci.*, **35**, 371–381.

Oort, A., 1974: Year-to-year variations in the energy balance of the arctic atmosphere. *J. Geophys. Res.*, **79**, 1253–1260.

Phillips, T. J., and A. J. Semtner, 1984: The seasonal response of the Held-Suarez climate model to prescribed ocean temperature anomalies. Part I: Results of decadal integrations. *J. Atmos. Sci.*, **41**, 2605–2619.

Salmun, H., R. F. Cahalan and G. R. North, 1980: Latitude dependent sensitivity to stationary perturbations in simple climate models. *J. Atmos. Sci.*, **37**, 1874–1879.

Saltzman, B., and A. D. Vernekar, 1983: The influence of poloidal motions and latent heat release on the equilibrium ice extent in a simple climate model. *J. Atmos. Sci.*, **40**, 1426–1434.

Sellers, W. D., 1969: A climate model based on the energy balance of the earth-atmosphere system. *J. Appl. Meteor.*, **8**, 392–400.

Simmons, A. J., J. M. Wallace and G. W. Branstator, 1983: Barotropic wave propagation and instability, and atmospheric teleconnection patterns. *J. Atmos. Sci.*, **40**, 1363–1392.

Suarez, M. J., and I. M. Held, 1979: The sensitivity of an energy balance model to variations in orbital parameters. *J. Geophys. Res.*, **84**, 4825–4836.