Sudden Stratospheric Warmings as Noise-Induced Transitions

THOMAS BIRNER

Department of Physics, University of Toronto, Toronto, Ontario, Canada

PAUL D. WILLIAMS

Department of Meteorology, University of Reading, Reading, Berkshire, United Kingdom

ABSTRACT

Sudden stratospheric warmings (SSWs) are usually considered to be initiated by planetary wave activity. Here it is asked whether small-scale variability (e.g., related to gravity waves) can lead to SSWs given a certain amount of planetary wave activity that is by itself not sufficient to cause a SSW. A highly vertically truncated version of the Holton–Mass model of stratospheric wave–mean flow interaction, recently proposed by Ruzmaikin et al., is extended to include stochastic forcing. In the deterministic setting, this low-order model exhibits multiple stable equilibria corresponding to the undisturbed vortex and SSW state, respectively. Momentum forcing due to quasi-random gravity wave activity is introduced as an additive noise term in the zonal momentum equation. Two distinct approaches are pursued to study the stochastic system. First, the system, initialized at the undisturbed state, is numerically integrated many times to derive statistics of first passage times of the system undergoing a transition to the SSW state. Second, the Fokker–Planck equation corresponding to the stochastic system is solved numerically to derive the stationary probability density function of the system. Both approaches show that even small to moderate strengths of the stochastic gravity wave forcing can be sufficient to cause a SSW for cases for which the deterministic system would not have predicted a SSW.

---

1. Introduction

The stratospheric circulation in winter is dominated by a strong cyclonic vortex over the pole. In principle, this polar vortex is driven by radiative cooling due to polar night. However, planetary waves frequently perturb the vortex in terms of both position and strength. For sufficiently strong planetary wave forcing, these disturbances can grow enough to destroy the polar vortex as a well-organized entity. These abrupt transitions are usually associated with a reversal of the zonal mean flow and a strong warming of the polar stratosphere and are thus called sudden stratospheric warmings (SSWs). Planetary wave activity is much stronger in the Northern Hemisphere compared to the Southern Hemisphere, mainly because of the difference in topography and land–sea contrast between the hemispheres, and SSWs are therefore much more frequently observed in the Northern Hemisphere. In fact, only one SSW has ever been observed in the Southern Hemisphere (see special issue of *J. Atmos. Sci.*, Vol. 62, No. 5, 2005).

Conventionally, SSWs are considered to be caused by planetary wave–mean flow interaction, more specifically by interactions of the zonal mean flow with zonal wavenumbers 1–3 (Andrews et al. 1987). Small-scale variability (e.g., due to breaking gravity waves), although well known to be ubiquitous in the winter stratosphere, is usually not considered to be sufficient to cause SSWs (an exception is the study by Duck et al. (2001), in which it is speculated that gravity waves might significantly impact the polar vortex evolution). In the present study, we ask whether such small-scale variability can modify the nature of the planetary wave–mean flow interaction to cause SSWs at planetary

wave strengths that by themselves are not sufficient to cause SSWs. This question is partly motivated by laboratory studies in which gravity waves are observed to cause transitions between different large-scale flow regimes (Williams et al. 2003). Furthermore, such transitions may be captured in a numerical model if the gravity waves are represented stochastically (Williams et al. 2004).

Holton and Mass (1976) have shown that a phenomenon resembling SSWs exists in a truncated version of a quasigeostrophic β-plane channel model of stratospheric wave–mean flow interaction (hereafter referred to as the Holton–Mass model). For a single zonal wavenumber forced at the lower boundary and with Newtonian relaxation of the zonal mean wind toward radiative equilibrium, they show that two distinct regimes exist for the zonal mean circulation: (i) a steady response for small wave forcing with the zonal mean wind close to radiative equilibrium and (ii) a vacillating response for the wave forcing above a threshold value with quasiperiodic zonal mean wind oscillations. Yoden (1987) studied the bifurcation properties of the Holton–Mass model and showed that the two circulation regimes correspond to two stable solutions of the system.

Recently, Ruzmaikin et al. (2003) introduced a highly vertically truncated version of the Holton–Mass model consisting of only three ordinary differential equations for the real and imaginary parts of the wave streamfunction and the zonal mean wind, respectively. For a given set of external parameters and very small amplitude of the wave forcing, this low-order system has only one stable solution close to radiative equilibrium. However, for higher amplitudes of the wave forcing, a second stable solution emerges that substantially deviates from radiative equilibrium and for which the wave forcing of the mean flow is of primary importance. For very high amplitude of the wave forcing, only the latter solution, far from radiative equilibrium, is stable. Thus, when initialized close to radiative equilibrium, the system undergoes an abrupt transition accompanying a sudden reduction of the zonal mean wind when the amplitude of the wave forcing exceeds a critical threshold. This abrupt transition in a low-order wave–mean flow model is likely the simplest prototype of a SSW.

In the present study, the low-order model introduced by Ruzmaikin et al. (2003) is extended to include stochastic forcing. Specifically, momentum forcing due to quasirandom gravity wave activity is introduced as an additive noise term in the zonal mean zonal momentum equation. It is thereby asked whether small-scale variability related to gravity waves can lead to SSWs given a certain amount of planetary wave activity that is by itself not sufficient to cause a SSW. Subsequent sections present the low-order model and our strategy to study SSWs in this model (section 2), followed by results and conclusions (section 3).

2. Low-order model and strategy

The stochastic low-order model used here represents a simple extension of the following deterministic low-order model as derived by Ruzmaikin et al. (2003):

$$\dot{X} = -\alpha_1 X - rY + sUY - \xi h + \delta_h \dot{h},$$
$$\dot{Y} = -\alpha_1 Y + rX - sUX + \zeta hU, \quad \text{and}$$
$$\dot{U} = -\alpha_2(U - U_R) - \eta hY. \quad (1)$$

This low-order system represents a coarse-grained version of the Holton–Mass model (Holton and Mass 1976; Yoden 1987) with only three vertical levels kept: the lower and upper boundary and a midlevel. Equations (1) represent ordinary differential evolution equations for the real and imaginary part of the wave streamfunction $\Psi = X + iY$ and the zonal mean zonal wind U at the midlevel [for details, see Ruzmaikin et al. (2003)]. The Holton–Mass model can be derived through straightforward truncation of the underlying quasigeostrophic potential vorticity equation coupled to an equation for the zonal mean zonal wind with Newtonian damping. In Eqs. (1), U_R corresponds to the zonal-mean zonal wind in thermal wind balance with the temperature of the prescribed radiative equilibrium; it is determined through the shear Λ of the radiative equilibrium and the midlevel height $z_T/2$. Thus, $U_R = \Lambda z_T/2$, where z_T is the height of the upper boundary. Throughout this study, $\Lambda = 1$ m s^{-1} km^{-1} (i.e., $U_R \approx 35$ m s^{-1}) and is fixed [note that accordingly the transient contribution proportional to $\dot{\Lambda}$ is omitted in Eqs. (1)]. The control parameter of the low-order system h parameterizes the (tropospheric) wave forcing and enters Eqs. (1) through the lower boundary condition [$\Psi(0, t) = gh/f$]; except for the experiments shown in Fig. 1, h is not considered to be time-dependent (i.e., for almost all experiments shown in this study, $\dot{h} = 0$ and the last term in the evolution equation for X is not present). All other parameters in Eqs. (1) are a function of prescribed external parameters and are given in the appendix of Ruzmaikin et al. (2003). The inverse time scales of the system $\alpha_{1,2}$ result in $\alpha_1 \approx (123 \text{ days})^{-1}$ and $\alpha_2 \approx (30 \text{ days})^{-1}$.

Stationary solutions of the system (1) can be obtained by zeroing the left-hand sides and are shown for U as a function of h in Fig. 1 [gray lines; this part of the figure is identical to Fig. 3 of Ruzmaikin et al. (2003)]. For 15 m $\lesssim h \lesssim$ 160 m, three stationary solutions exist, of which two are stable (solid lines in Fig. 1): the upper branch, the system close to radiative equilib-
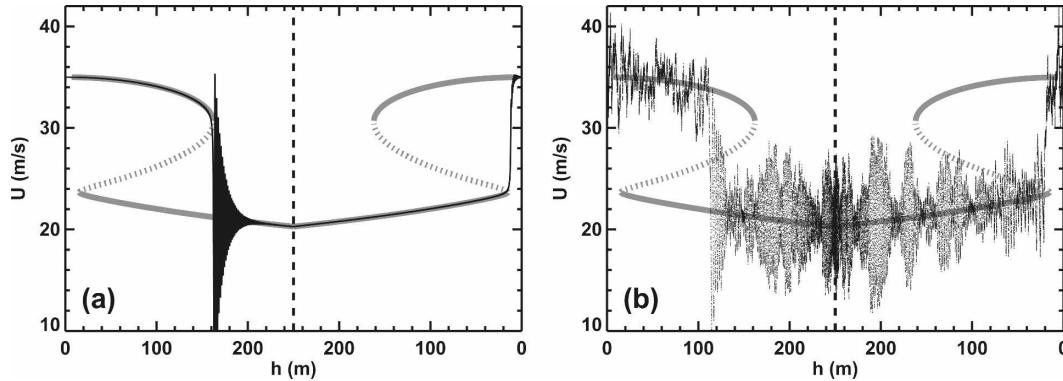
FIG. 1. Illustration of an SSW in the (a) deterministic and (b) stochastic system; U is plotted as a function of h, where h was prescribed as a slowly varying function of time given by Eq. (2) and U is obtained by forward integration of the deterministic or stochastic system. The planetary wave forcing h is initially zero and increases monotonically to 250 m (marked by vertical dashed lines) during the first 2000 days, with a reverse but otherwise identical decrease in magnitude back to zero from 2000 to 4000 days. For the integration shown in (b), $\sigma_a = 0.5$ m s^{-1} day^{-1}. Note the reversed abscissa values to the right of the dashed lines in both panels. Thick gray lines mark stationary solutions of the deterministic system, with full lines denoting stable solutions (radiative solution = upper branch; SSW solution = lower branch) and the dotted lines denoting the unstable solution. See text for details.

rium, hereafter the radiative solution (strong U); and the lower branch, the system far from radiative equilibrium, hereafter the SSW solution (weak U). The middle branch (dotted line) represents an unstable solution. For $h \lesssim 15$ m (weak tropospheric wave forcing), only the radiative solution exists; likewise, for $h \gtrsim 160$ m (strong tropospheric wave forcing), only the SSW solution exists. That is, initialized at the radiative solution, the deterministic system (1) allows for SSWs only if the tropospheric wave forcing h exceeds a certain threshold value (here 160 m). This is illustrated in Fig. 1a by overplotting a weakly time-dependent solution $U(t)$ of Eqs. (1) as obtained by prescribing h as a slowly varying function of time,

$$h(t) = 250 \text{ m} \left(\frac{1}{2} - \frac{1}{2} \cos \frac{2\pi t}{4000 \text{ days}} \right), \qquad (2)$$

such that the transience of h does not impact the behavior of the system (i.e., the solution is quasistationary). Numerically, system (1) is integrated using a simple first-order accurate forward-in-time finite difference scheme with a time step of 0.1 days. As prescribed by Eq. (2), $h(t)$ smoothly and monotonically increases from 0 to 250 m within the first 2000 days and smoothly and monotonically decreases in the same way from 250 to 0 m during the next 2000 days, after which the integration is stopped. The plotting convention in Fig. 1 is such that time is continuously running from left to right; that is, the abscissa values h reverse at the dashed lines. Initially, the quasistationary solution $U(t)$ (thin black line in Fig. 1a) corresponds to the radiative solution (upper branch), remaining on the upper branch until the radiative solution ceases to exist (at $h \approx 160$ m), at which point a SSW occurs [i.e., the system makes a

transition to the SSW solution (lower branch)]. Likewise, for decreasing h the system remains on the lower branch until the SSW solution ceases to exist (at $h \approx 15$ m), thereby closing a hysteresis loop. Of course, in the bistable regime in which both the radiative and the SSW–solution exist and are stable, it is the initial conditions that determine whether the system will eventually end up on the lower or upper branch.

In this study, the deterministic system (1) is extended to include the effect of small-scale variability (e.g., due to quasirandom gravity wave activity, which in the full equations would be represented by the vertical convergence of vertical flux of zonal momentum). This small-scale variability is represented as a simple additive noise term incorporated into the evolution equation for U (denoted by χ_a):

$$\dot{U} = -\alpha_2(U - U_R) - \eta hY + \chi_a, \qquad (3)$$

where χ_a is modeled as Gaussian white noise:

$$\overline{\chi_a(t)\chi_a(t')} = \sigma_a^2 \delta(t - t'). \qquad (4)$$

Here, σ_a represents the strength of this additive noise and $\delta(\cdot)$ denotes the Dirac delta distribution. In general, the stochastic forcing χ_a represents the effect of any small-scale variability, not necessarily limited to gravity waves, that can be parameterized decoupled from the large-scale flow (i.e., χ_a represents an additive forcing of U and χ_a does not enter the evolution equations for X and Y). However, in the present study we assume that gravity wave activity is the main cause of small-scale forcing of the mean flow in the polar stratosphere.

According to the forward-in-time numerical scheme for the deterministic system (1), the stochastic term in

Eq. (3) is appropriately integrated using the square root of the time step (Kloeden et al. 1994). The units of σ_a as defined through Eq. (4) therefore result in nonphysical units of meters-seconds$^{-1.5}$ which would complicate the interpretation of results. However, χ_a should be interpreted as parameterizing a physical random process of finite autocorrelation time T, say an Ornstein–Uhlenbeck process [see e.g., Horsthemke and Lefever (1983)]. In this case, the corresponding variance would result in $\sigma_a^2/2T$. If we measure this variance in (meters per second per day)^2 and assume $T \sim 0.5$ days (which is not unrealistic for gravity waves), then values of σ_a provide the correct corresponding physical values of units meters per second per day. We will therefore simply refer to values of σ_a having units meters per second per day throughout the paper. For reference we note that observed values of gravity wave drag in the polar stratosphere are on the order of 1 m s^{-1} day^{-1} (Hamilton 1997). In the present study, values for σ_a on the order of 0.1–10 m s^{-1} day^{-1} are considered. It should be noted that our model represents a strong idealization of the real atmosphere and that a quantitative comparison of values for σ_a to observed gravity wave drags has only limited relevance. Nevertheless, the range of values for σ_a used in the present study was chosen to broadly represent the observed parameter regime.

SSWs in the stochastic system are represented as transitions from the radiative solution (upper branch) to the SSW solution (lower branch). Given that the system is initialized at the (deterministic) radiative solution in the bistable regime, transitions can only be caused by the imposed small-scale variability parameterized as additive noise. We thus refer to these transitions as *noise-induced transitions*. One such noise-induced transition is illustrated in Fig. 1b, which shows a repetition of the above described deterministic weakly transient experiment (Fig. 1a) with prescribed $h(t)$ as in Eq. (2). This time, the system contains an additive noise term according to Eq. (3) with $\sigma_a = 0.5$ m s^{-1} day^{-1} (i.e., rather small noise strength). Evidently, a transition (SSW) from the upper to the lower branch occurs long before the threshold value of h at which the radiative solution ceases to exist is reached (transition at $h \approx 115$ m). Once in the basin of attraction of the SSW solution, the system remains there until h gets close to the threshold at which the SSW solution ceases to exist (reverse transition at $h \approx 20$ m in Fig. 1b). This indicates that in the bistable regime the SSW solution is more stable (deeper basin of attraction) than the radiative solution (cf. section 3). Note that the particular realization of the stochastic system shown in Fig. 1b was chosen to show representative transition points.

One way to study these noise-induced transitions is

to obtain statistics of first passage times (FPTs) of the system undergoing such transitions. For a given control parameter h the system is initialized at the upper branch ($U \approx U_R$) and integrated forward in time until U crosses the stationary solution of the lower branch for the first time that gives an FPT. This experiment is repeated many times with different noise realizations to derive statistics of these FPTs. Transitions occurring after 1000 days are not considered because this is much longer than the length of a given winter (i.e., integrations are stopped at $t = 1000$ days if no transition occurs by then).

Another way to study the behavior of the stochastic system is to solve the corresponding Fokker–Planck equation, which is a deterministic partial differential equation for the evolution of the probability density function in $[X, Y, U]$ space. The major advantage of this approach is that the Fokker–Planck equation describes the outcome of a hypothetical infinite-member ensemble, which is clearly unachievable by direct integration of the stochastic differential equations. The Fokker–Planck equation governing the evolution of the probability density function $p(\mathbf{x}, t)$ for this system is

$$\frac{\partial}{\partial t} p(\mathbf{x}, t) = -\sum_{i=1}^{3} \frac{\partial}{\partial x_i}[A_i p(\mathbf{x}, t)] + \frac{1}{2}\sigma_a^2 \frac{\partial^2}{\partial x_3^2} p(\mathbf{x}, t), \quad (5)$$

where $\mathbf{x}(t)$ describes the state vector with components $\mathbf{x} \equiv (x_1, x_2, x_3) = (X, Y, U)$ and $\mathbf{A}(\mathbf{x})$ is the vector describing the deterministic dynamics with its components corresponding to the right-hand sides of Eqs. (1); that is, $\mathbf{A} \equiv (A_1, A_2, A_3) = (\dot{X}, \dot{Y}, \dot{U})$.

The summation term on the right-hand side of Eq. (5) describes the dynamics of the deterministic system and is called the *deterministic drift*. The remaining term on the right-hand side is associated with diffusion of probability by the additive noise (only present in the evolution equation for U). Note that because our system contains no state-dependent (multiplicative) noise, the *noise-induced drift* term is absent from the Fokker–Planck equation, and so there is no distinction between the Stratonovich and Itô interpretations of stochastic calculus [see, e.g., Horsthemke and Lefever (1983) for more details].

Equation (5) is numerically integrated using a standard centered-in-time and centered-in-space (CTCS) finite difference scheme, which is second-order accurate. An integration domain defined by $X \in (-0.06, 0.04)$, $Y \in (-0.05, 0.05)$, and $U \in (0, 0.8)$ in nondimensional units is used (for nondimensionalization, the earth's radius is used as the length scale and 1 day as the time scale). The domain is divided into a regular numerical grid of $20 \times 20 \times 40$ boxes in the X, Y, and U dimen-
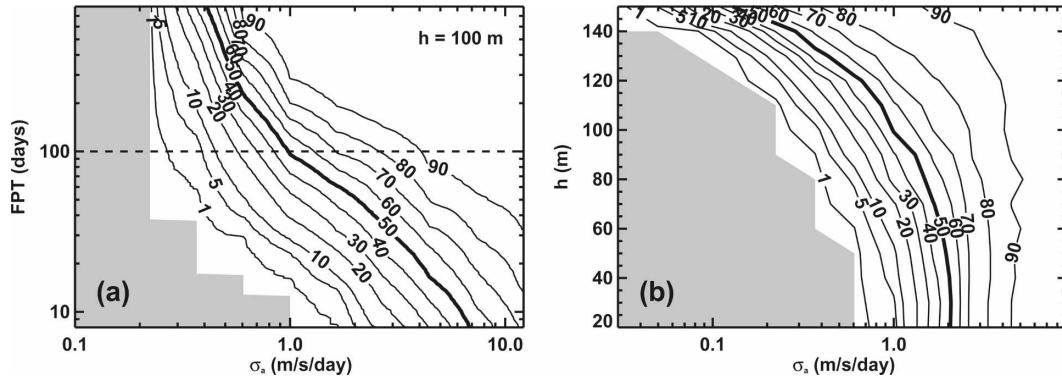
FIG. 2. (a) Cumulative frequency distribution of first passage times for $h = 100$ m as a function of noise strength σ_a. An FPT is defined as the first transition from the radiative solution to the wave SSW solution. The horizontal dashed line marks approximate length of winter (100 days). (b) Relative frequency of FPT < 100 days (i.e., the probability that a transition from a radiative to an SSW solution takes place within a given winter as a function of σ_a and h). Gray shading marks areas of zero frequency in both (a) and (b).

sions, respectively. The nondimensional time step is 10^{-4}. The leapfrog time-stepping scheme uses a Robert filter (with filter parameter 0.01) to suppress computational mode splitting. All the integrations of the Fokker–Planck Eq. (5) described in this study use $h = 100$ m and $\Lambda = 1$ m s^{-1} km^{-1}.

3. Results and conclusions

For a given control parameter $h = 100$ m, Fig. 2a shows the cumulative frequency distribution of FPTs as a function of noise strength σ_a obtained by computing 1000 different stochastic trajectories for each σ_a. As mentioned in the previous section, FPTs smaller than 100 days (i.e., with at least one transition within a given winter) are of primary interest and FPTs greater than 1000 days are not considered (trajectories were stopped after 1000 days if no transition had occurred by then). For very small noise strength ($\sigma_a \lesssim 0.2$ m s^{-1} day^{-1}), there are no transitions (gray shading); that is, the stochastic system qualitatively behaves like the deterministic system. However, for larger noise strengths, transitions become more and more likely. The probability of transitions with a certain FPT sharply increases once the minimum noise strength for transitions to occur is exceeded. That is, once σ_a exceeds this minimum value, even a small to moderate extra enhancement in noise strength makes transitions very likely to occur. At $\sigma_a \approx 1$ m s^{-1} day^{-1}, a transition in a given winter (within the first 100 days) has 50% probability (i.e., an SSW would be expected statistically to occur every other winter). The median FPT (50% contour in Fig. 2a) roughly scales with noise strength σ_a (roughly linear behavior on the log–log scale in Fig. 2a).

Figure 2b shows the cumulative frequency distribution of a transition to occur at least once over the

course of a given winter (FPT < 100 days), that is, in our terminology, the probability of a SSW as a function of σ_a and h. The minimum noise strength needed for transitions to occur (at least once out of the 1000 trajectories that were run) is not a strong function of h for $h \lesssim 100$ m. The median (SSWs every other winter) for $h \lesssim 100$ m is located at $\sigma_a \sim 1$ m s^{-1} day^{-1}, even for small h near the threshold where the SSW solution ceases to exist. That is, noise-induced transitions can still occur even when the large-scale forcing (parameterized by h) is far from causing a (deterministic) transition to the SSW solution. Applied to the real atmosphere, this means that even when planetary wave activity is not sufficient to cause a SSW by itself, small-scale variability (e.g., due to gravity waves) could trigger a SSW.

For $h \gtrsim 100$ m, transitions become more and more likely and transitions can occur even for very small noise strength ($\sigma_a \sim 0.1$ m s^{-1} day^{-1} and smaller). This is not surprising because the system for $h > 100$ m gets closer to the threshold at which the radiative solution ceases to exist.

The FPT analysis above can only determine whether a transition occurs in principle (i.e., at least once, from the radiative to the SSW solution). In general, transition back from the SSW to the radiative solution might also occur, as well as repeated transitions between the two stable solutions. In the latter case, one would expect a bimodal occupation statistic of the two stable solutions. Such a behavior can be tested more conveniently by solving the Fokker–Planck Eq. (5). Numerical solutions of Eq. (5) are only obtained for $h = 100$ m (and $\Lambda = 1$ m s^{-1} km^{-1} as before) in the present study.

Figure 3 shows the evolution of the probability density function when $\sigma_a = 3$ m s^{-1} day^{-1}. Following the FPT analysis above, an initial probability density func-
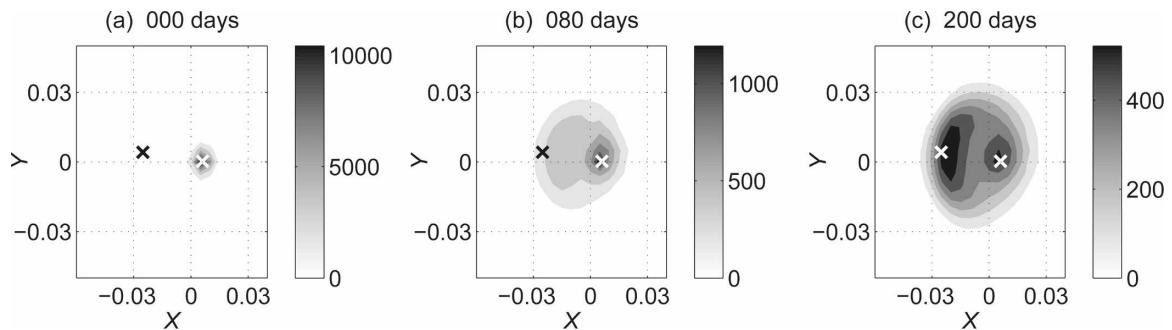
FIG. 3. Numerical solution of the Fokker–Planck equation for the case $h = 100$ m and $\sigma_a = 3$ m s^{-1} day^{-1}. The quantity shown is the probability density function integrated over all values of U, plotted as a function of X and Y (nondimensionalized), after (a) 0, (b) 80, and (c) 200 days. The initial state, shown in (a), is a narrow three-dimensional Gaussian centered on the radiative stable equilibrium point. The crosses indicate the locations of the two stable equilibrium points of the deterministic system: the radiative solution ($X = 0.006$, $Y = 0.0003$, closer to the origin) and the SSW solution ($X = -0.025$, $Y = 0.0041$).

tion that is Gaussian and centered on the radiative solution is used, with widths equal to one-twentieth of the domain extent in each dimension (Fig. 3a). The diffusion term in Eq. (5) broadens the probability density function preferentially in the direction of the SSW solution over the first 80 days (Fig. 3b). The maximum probability density remains at the radiative solution during this time interval, but a secondary maximum appears near the SSW solution after around 120 days. The secondary maximum gradually intensifies and eventually becomes the global maximum after around 200 days (Fig. 3c). There is little subsequent evolution; thus, the state after 200 days is very close to the equilibrated state, that is, the solution of Eq. (5) with $\partial p / \partial t = 0$. The equilibrated probability density function is clearly bimodal and the system is evidently more likely to be found near the SSW solution than the radiative solution. It is interesting to note that bimodality was also obtained from an EOF analysis of observed variability of the winter stratosphere (Monahan et al. 2003). However, bimodality in the case of Monahan et al. (2003) results from variability in initial conditions (dif-

ferent winters) and in the large-scale forcing (which would translate into variability of h and is not considered in the present study). On the other hand, bimodality as in Fig. 3c can only result from the additive noise term, that is, from small-scale variability (to the extent that for initial conditions as in Fig. 3a the deterministic system would always approach the radiative solution).

Figure 4 shows how the equilibrated probability density function varies with noise strength when initiated near the radiative solution, as described in the previous paragraph. For small values of the noise ($\sigma_a = 1.5$ m s^{-1} day^{-1}; Fig. 4a), the equilibrated probability density function is centered on the SSW solution, with no evidence of a secondary maximum at the radiative solution. Therefore, for this noise strength, the system initialized near the radiative solution will *always* undergo a transition to the wave solution eventually. This finding is consistent with the FPT analysis above. Furthermore, the figure implies that, once the transition to the SSW solution has occurred, the system will remain there indefinitely (i.e., that the reverse transition does
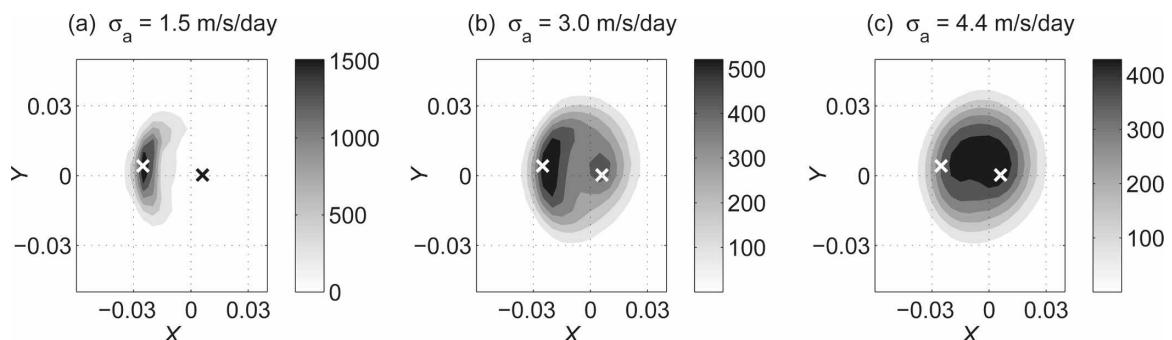

FIG. 4. Equilibrated solutions of the Fokker–Planck equation for the case $h = 100$ m. The quantity shown is the equilibrated probability density function as in Fig. 3. The noise levels are $\sigma_a =$ (a) 1.5, (b) 3.0, and (c) 4.4 m s^{-1} day^{-1}. The initial state is that shown in Fig. 3a in each case. The crosses indicate the locations of the two stable equilibrium points of the deterministic system.

not occur). As the noise strength is increased, bimodality in the equilibrated probability density function occurs around $\sigma_a = 3$ m s^{-1} day^{-1} (Fig. 4b), implying that the reverse transition from the SSW to the radiative solution becomes possible at these intermediate noise levels. For higher noise strengths ($\sigma_a = 4.4$ m s^{-1} day^{-1}; Fig. 4c), the equilibrated probability density function becomes unimodal once again, but it is so wide that the concept of two distinct stable states (which is derived from the deterministic system) is no longer meaningful. Further experiments for a range of values of σ_a between 0.5 and 7 m s^{-1} day^{-1} revealed that the system is more likely to be found near the SSW solution for $\sigma <$ 3 m s^{-1} day^{-1}, whereas the system is more likely to be found near the radiative solution for $\sigma > 3$ m s^{-1} day^{-1} (not shown).

In summary, the low-order system (1), as introduced by Ruzmaikin et al. (2003) and extended by an additive noise term, Eq. (3), exhibits noise-induced transitions from the radiative solution to the SSW solution even for small noise strength σ_a. That is, SSWs can occur for values of h (mimicking planetary wave activity) that are too small to cause a transition to the SSW solution in the deterministic case. Even for small values of h far from the critical value at which transitions would occur in the deterministic system, noise-induced transitions do occur, albeit very infrequently for moderate gravity wave noise levels. Furthermore, the system with noise preferentially resides around the SSW solution for small to intermediate noise strengths (cf. Figures 4a and 1b).

A shortcoming of the present approach is the neglect of transience in planetary wave activity—for example, even for small background planetary wave forcing (i.e., small h), a burst of additional forcing that exceeds the threshold above which only the SSW solution exists ($h \gtrsim 160$ m) might trigger a transition to the SSW solution. The importance of such transience in planetary wave activity has long been recognized since the original work of Matsuno (1971). Here we restricted ourselves to constant planetary wave forcing as in Holton and Mass (1976); we focused instead in this note on noiselike small-scale forcing because this represents a largely unexplored mechanism to cause SSWs.

Our numerical experiments represent a proof-of-concept study using a very simple conceptual model. It remains to be seen whether the mechanism here described can be found in the real atmosphere or even in more comprehensive models. A recent numerical study employing the Holton–Mass model found that perturbations applied to the zonal mean zonal wind in the upper stratosphere can lead to significant downward propagation, especially in the bistable regime, that is, when both the radiative and the SSW solutions are stable (Hardiman and Haynes 2008, manuscript submitted to *J. Geophys. Res.*). Our findings suggest that the role of small-scale variability in triggering SSWs may be important and should be studied further.

REFERENCES

Andrews, D. G., J. R. Holton, and C. B. Leovy, 1987: *Middle Atmosphere Dynamics.* Academic Press, 489 pp.

Duck, T. J., J. A. Whiteway, and A. J. Carswell, 2001: The gravity wave–Arctic stratospheric vortex interaction. *J. Atmos. Sci.,* **58,** 3581–3596.

Hamilton, K., Ed., 1997: *Gravity Wave Processes: Their Parameterization in Global Climate Models.* Springer, 414 pp.

Holton, J. R., and C. Mass, 1976: Stratospheric vacillation cycles. *J. Atmos. Sci.,* **33,** 2218–2225.

Horsthemke, W., and R. Lefever, 1983: *Noise-Induced Transitions.* Springer, 318 pp.

Kloeden, P. E., E. Platen, and H. Schurz, 1994: *Numerical Solution of SDE through Computer Experiments.* Springer, 292 pp.

Matsuno, T., 1971: A dynamical model of the sudden stratospheric warming. *J. Atmos. Sci.,* **28,** 1479–1494.

Monahan, A. H., J. C. Fyfe, and L. Pandolfo, 2003: The vertical structure of the wintertime climate regimes of the Northern Hemisphere extratropical atmosphere. *J. Climate,* **16,** 2005–2021.

Ruzmaikin, A., J. Lawrence, and C. Cadavid, 2003: A simple model of stratospheric dynamics including solar variability. *J. Climate,* **16,** 1593–1600.

Williams, P. D., P. L. Read, and T. W. N. Haine, 2003: Spontaneous generation and impact of inertia–gravity waves in a stratified, two-layer shear flow. *Geophys. Res. Lett.,* **30,** 2255, doi:10.1029/2003GL018498.

——, T. W. N. Haine, and P. L. Read, 2004: Stochastic resonance in a nonlinear model of a rotating, stratified shear flow, with a simple stochastic inertia–gravity wave parameterization. *Nonlinear Processes Geophys.,* **11,** 127–135.

Yoden, S., 1987: Bifurcation properties of a stratospheric vacillation model. *J. Atmos. Sci.,* **44,** 1723–1733.