A GCM Study of Volcanic Eruptions as a Cause of Increased Stratospheric Water Vapor

MANOJ M. JOSHI AND KEITH P. SHINE

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

(Manuscript received 25 November 2002, in final form 21 May 2003)

ABSTRACT

Recent general circulation model (GCM) experiments have shown that idealized climatic perturbations that increase the temperature of the tropical tropopause region can cause larger than expected surface temperature increases. This is because the extra water vapor that is transported into the stratosphere acts as a positive radiative forcing agent. Since major volcanic eruptions in the Tropics also perturb temperatures in the region of the tropical tropopause, evidence is sought for the existence of such a mechanism by comparing model simulations and observations of the period following the eruption of Mount Pinatubo in 1991. Lower-stratosphere temperature perturbations are simulated using a GCM of intermediate complexity, and increases in water vapor concentration are found in the lower stratosphere that decay slowly in the years following the peak temperature increase. Such increases are consistent with the limited observations that exist over this period, as well as earlier work done with simpler 2D models. Since the amount of extra water transported into the stratosphere is sensitive to the absolute temperature change in the tropical tropopause region, this mechanism will be modulated by the phases of ENSO and the quasi-biennial oscillation, as well as being dependent on the vertical profile of the heating due to the volcanic aerosol. The global-mean radiative forcing associated with the perturbation in stratospheric water vapor is approximately $0.1 \text{Wm}^{-2}$, and thus slightly counteracts the negative radiative forcing effect of the volcanic aerosol itself. It is suggested that repeated volcanic eruptions might at least partially account for increases in stratospheric water vapor over the last 40 years.

1. Water vapor and the stratosphere

Much attention has been focused on recent increases in water vapor concentrations in the stratosphere (e.g., Kley et al. 2000; Rosenlof et al. 2001). The increases, thought to be due to both increased oxidation of CH$_4$ (Considine et al. 2001) and changes in the amount of transport of water vapor into the stratosphere from below, display a significant amount of interannual variability (Rosenlof et al. 2001). In this paper, we present general circulation model results that indicate that some transient increases in lower-stratospheric (hereafter LS) water vapor might be associated with the presence of volcanic aerosol from the 1991 eruption of Mount Pinatubo. The mechanism would also apply to other tropical eruptions that result in enhanced aerosol in the tropical lower stratosphere, and indeed any other mechanism that modulates the temperature of the tropical tropopause region.

Observations of stratospheric water vapor are very limited. The longest coherent dataset is balloon-borne frost-point hygrometer measurements at Boulder, Colorado, at 40°N. These show increases in water vapor at the 24–26 km ($\approx$30 hPa) level during 1992, a year after the eruption of Mount Pinatubo (Oltmans et al. 2000, their Fig. 1). Intriguingly, the same measurements also show smaller spikes of water vapor in 1982–83, a few months after the eruption of the El Chichón volcano.

It is well known that the stratospheric aerosol associated with Pinatubo itself made remotely sensed retrievals of constituents difficult, especially in the LS region below about 40–50 hPa (Harries et al. 1996; Elson et al. 1996). Evans et al. (1998) only considered variations in humidity above an altitude of 30 km, or $\approx 15 \text{hPa}$. However, some satellite observations of water vapor do display elevated amounts in the 5–20-hPa level in 1994 (Rosenlof et al. 2001). They are consistent with increased water vapor in the LS region in early 1992, which is then transported to the 5–20-hPa level 2 yr later via the tropical tape recorder. Observations show that water vapor pulses are transported from the 20-km level to 40 km on a 2-yr timescale [see Fig. 1 in Geller et al. (2002)]. Circulation anomalies associated with aerosol heating might contribute to the explanation of these water vapor changes (Rosenlof 2002).

Evans et al. (1998) suggested that if the eruption of Pinatubo had caused an increase in LS water vapor in...
1992, then transport of water vapor into the middle stratosphere would result in an increase in water vapor trend with height over the period 1992–96, after the trend in methane was taken into account. Since the \( 2\text{CH}_4 + \text{H}_2\text{O} \) trend did not increase with height, no external injection of water vapor was evident. However, examination of their Fig. 4 does show some evidence for a small trend of the order of 20 ppbv yr\(^{-1}\) in the region 30–45 km. We comment on this in section 4.

The most immediate effect of volcanic eruptions is to dry the stratosphere, even though volcanoes themselves can directly inject large amounts of water vapor into the stratosphere (Glaze et al. 1997), because \( \text{SO}_2 \) emitted from the volcano reacts with stratospheric water vapor to produce sulfate aerosol over a period of some weeks (Pinto et al. 1989). The resulting aerosol then persists over a much longer period; the Pinatubo aerosol cloud lasted in the stratosphere for 1–2 yr (Stenchikov et al. 1998), and it is the heating effect of the aerosol cloud on the tropical tropopause that is the subject of our paper.

Considine et al. (2001) have modeled the response of stratospheric water vapor to the Pinatubo eruption using an interactive 2D transport-chemistry model with a simplified four-layer troposphere. They found that transport of water across a warmer post-Pinatubo tropopause, as well as methane oxidation could help to explain observed trends in water vapor in the 1990s (see their Fig. 16).

Another motivation for this study is that Stuber et al. (2001) have recently shown that increases in LS water vapor can be brought about by increasing the temperature of the tropopause region. The reason is that upward water transport through this region is modified by dehydration in it (e.g., Brewer 1949; Dessler 1998). The Stuber et al. (2001) mechanism involved increasing \( \text{O}_3 \) between 20 and 90 hPa, which caused an increase in temperature of \( 10\text{K} \) in the LS region. Rosenfield et al. (1998) also simulated increases in LS water vapor associated with subvisible cirrus clouds heating the tropopause region. National Centers for Environmental Prediction (NCEP) reanalyses of LS temperature show increases associated with the presence of Pinatubo sulfate aerosol peaking in early 1992 (see, e.g., Stenchikov et al. 1998; Ramachandran et al. 2000; Considine et al. 2001), so we propose that such aerosol heating allows more water vapor into the LS region in a similar fashion.

The particular importance of the Stuber et al. (2001) results is that the enhanced concentration of water vapor leads to a significant climate feedback, which enhances the climate sensitivity of their GCM by a factor of 1.8 compared to its sensitivity to \( \text{CO}_2 \) changes. Joshi et al. (2003) have shown that a similar mechanism exists in other GCMs, albeit with smaller enhancements in climate sensitivity of 1.2–1.4. If we can find evidence for increased stratospheric water vapor in the period following the Mount Pinatubo eruption, this would add support to the existence of the mechanism in Stuber et al. (2001).

Rather than trying to model explicitly the aerosol heating resulting from Pinatubo, which has already been done (Kirchner et al. 1999; Considine et al. 2001), the present work simply applies a heating function that simulates the effects of perturbations to globally averaged NCEP reanalysis temperature (data available online at http://www.cdc.noaa.gov/cdc/reanalysis/reanalysis.shtml), as it is the actual temperature changes in the region of the tropical tropopause that allow extra water vapor into the LS region. We then see if any changes in LS water vapor occur in the model, and compare these with observations. Although there is some discrepancy between the NCEP analysis and radiosonde soundings (e.g., Randel et al. 2000; Zhou et al. 2001), the NCEP analysis does give a more complete global picture. The processing of the NCEP data is described below.

The reason for using globally averaged reanalysis temperature at a certain level is that such an averaging process will average out dynamical heating and cooling effects to zero, thus ensuring that any temperature perturbations are due to diabatic perturbations, such as the presence of aerosol.

The surface albedo is changed as a proxy for the effect of the aerosol on the surface radiation budget. This is done in a time-varying fashion in order to reproduce the tropospheric cooling reported by Soden et al. (2002). Soden et al. (2002) have presented observational evidence for the model-simulated tropospheric drying that accompanied the tropospheric cooling.

2. The Intermediate General Circulation Model (IGCM)

The IGCM is a three-dimensional global atmospheric circulation model. It has relatively simple physical parameterizations, which reduce its computational demand, enabling a wider range of parameter space to be explored. It has been used extensively in studies of the earth’s climate and climate change (e.g., Forster et al. 2000). The climateology of the model in its present incarnation is described by Forster et al. (2000). The evaporation, transport, and precipitation of water are accounted for. The IGCM should allow a more realistic representation of near-tropopause processes than the 2D model of Considine et al. (2001).

The IGCM has spectral resolution T21, corresponding to a gridpoint resolution of about 5.5\(^\circ\). There are 22 layers between the surface at 1000 hPa and the model top at the 1-hPa level. Oceanic heat transport is represented by previously determined monthly averaged fluxes, which are added to the heat balance in the model’s mixed layer ocean, forcing it toward climatologically determined values (Forster et al. 2000).

The seasonal cycle of water vapor at the IGCM tropical tropopause varies between 3 ppmv in January and
4.5 ppmv in July, which is consistent with observations (Kley et al. 2000). In reality, the transport of water into the stratosphere is strongly modified by interannually varying phenomena such as the El Niño–Southern Oscillation (ENSO) and the quasi-biennial oscillation (QBO) (Geller et al. 2002). The IGCM does not represent the QBO, but we would actually cite this as an advantage for the present study. The reason is that it is desirable that the component of the LS interannual variability of water vapor that is a result of the aerosol heating should be as large as possible. The absence of a QBO in the IGCM should therefore make isolation of such a signal easier to detect. The lack of a QBO is shared with most other GCMs, and does not lessen the validity of the IGCM in addressing this particular problem.

The choice of heating function is such that it should primarily match globally averaged NCEP temperature reanalyses, while also approximating the shape of the heating function shown in Ramachandran et al. (2000) in their Fig. 3. NCEP perturbation temperatures are produced by first removing the mean and linear trend from a 20-yr time series from January 1980 to December 1999. An annual composite of the 20 yr of processed data is produced and removed from the processed data, to remove the effects of the annual cycle, and the data averaged over longitude and latitude.

The heating function inserted into the IGCM is given by

\[ Q = Q_t \times q_p \times q_\phi, \]  

where \( Q \) is in K day\(^{-1}\), \( t \) is time, \( p \) is pressure, and \( \phi \) is latitude. The component that varies in time is given by

\[ Q_t = 2.75t^{1.5} \exp(-t^{1.5}), \]  

where \( t \) is measured in years after the start of the integration, and \( Q_t \) reaches a maximum value of 1 K day\(^{-1}\) at \( t = 1 \text{ yr} \) and decays to zero after that. Here, \( q_p \) represents how the heating scales as a function of pressure, and is given by

\[ q_p = A \exp\left\{\left[ \frac{\ln p - B}{C} \right]^2 \right\} - 0.05, \]  

where \( p \) is pressure in hPa. The meridional scaling of the diabatic heating is given by

\[ q_\phi = (\cos 2.7 \phi + 0.33)/1.33, \]  

where \( \phi \) is latitude in radians. The surface albedo change follows the same time and latitude variation as Eq. (1). The peak amplitude at time \( t = 1 \text{ yr} \) is +0.02.

The IGCM is integrated for 6 yr after the start of the simulated heating perturbation. Globally averaged temperatures are then compared with the NCEP reanalysis temperatures. After some experimentation, the following values for the constants in the heating field are chosen as a good match to the observations: \( A = 0.3, B = 3.55, C = 1.1, D = 4 \). The spatial distribution of peak heating in this integration (denoted run 1), that is, the heating when

\[ Q = 1 \text{ K day}^{-1}, \]  

is shown in Fig. 1. The heating function bears more resemblance to Ramachandran et al. (2000) than Stenchikov et al. (1998) in that it includes the slight midlatitude cooling in the former study.

At 50 hPa, the maximum heating amplitude needed to simulate the NCEP reanalysis temperatures is slightly less than Ramachandran et al. (2000). Kirchner et al. (1999) and Yang and Schlesinger (2002) speculated that the presence of the QBO, as well as ozone loss following the eruption, could account for some of the discrepancy.

The cold point tropical tropopause in the IGCM lies between the 80 and 100 hPa levels, and has a temperature of 190–195 K (see Fig. 4 of Forster et al. 2000). There are model layers in this area at \( \sigma = 0.05, 0.081, \) and 0.118 (where \( \sigma \) is pressure divided by surface pressure), so the cold point tropopause in the IGCM is significantly heated by the perturbation shown in Fig. 1.

Three model integrations are carried out under the conditions used for run 1, except two are slightly perturbed in temperature at the start of the integration. The result is a three-run “ensemble,” whose output is analyzed. The output is compared with a 50-yr control IGCM integration.

3. Results

The evolution of globally averaged temperature with time is shown in Fig. 2. The NCEP values in Fig. 2 have been produced by the method shown above. A similar processing has been performed on the IGCM results, except that no linear trend has had to be removed.

In each panel, the modeled year is placed above the top axis for reference. The eruption occurred in June 1991, which corresponds to year 0.5 on the bottom of each of the axes. Peak aerosol loading occurred in late 1991 to early 1992 (Stenchikov et al. 1998). As shown in Fig. 2, the maximum globally averaged temperature perturbation also occurs during 1992, which corresponds to years 1.0–2.0 in the IGCM simulation.
The first panel of Fig. 2 shows that the IGCM represents the tropospheric cooling reported and modeled by Soden et al. (2002) fairly well. In fact the IGCM cools slightly more than the NCEP reanalysis shown, although effects such as ENSO would tend to keep the observations cooler than they might otherwise have been (Soden et al. 2002).

In the tropical tropopause region, the observations
show a warming of over 1 K, peaking in early 1992, and then decaying over the next two years (e.g., Angell 1997). The perturbation is a factor of 2 higher than that achieved by Considine et al. (2001) in their 2D modeling study. The peak amplitudes and decay timescales of the data are captured quite well by the IGCM, indicating that the representation of aerosol heating in the model is successfully reproducing the resulting impact on temperature.

A heating of about 1 K at the tropical lapse-rate tropopause after the eruption of El Chichón is evident in both Randel et al. (2000) and Zhou et al. (2001), although the interannual variability, probably associated with ENSO and the QBO, precludes a clear identification of a similar peak associated with the eruption of Mount Pinatubo. However, Angell (1997) shows a warming of the equatorial region between 50 and 100 hPa peaking at 2 K following Pinatubo after accounting for the impact of the QBO.

One difference between model and observations at the 100- and 70-hPa levels is that the data show slight evidence of a double peak structure, which might be related to the QBO (e.g., Angell 1997): one peak occurs in late 1991 and the other in late 1992. The IGCM, with its relatively simple heating function, does not reproduce the two peaks. The 2-K peak warming at 50 hPa is similar to that of Considine et al. (2001).

The effect of the temperature perturbation on specific humidity is shown in Fig. 3, which shows that when averaged globally, LS water concentration is up to 10% higher than normal for 18 months after the model eruption. These changes are larger than those found in similar analysis on globally averaged LS water vapor values in the 50-yr control run, variations in the 100–50-hPa level are never more than 4%, and always last less than 6 months.

Most of the water appears to be transported up through the equatorial region, as shown in Fig. 3 (bottom). In this area water amounts are 20%–30% above their climatological values. The effect of lowering tropospheric temperature is mirrored in tropospheric drying, and is also shown in Fig. 3. The humidity is lowered by 5%–10% for almost two years, consistent with Soden et al. (2002). The values shown in Fig. 3 (bottom) are similar to those reported by Considine et al. (2001) in their Fig. 16.

A zonal-mean picture of the perturbation is shown in Fig. 4, which shows the humidity perturbation averaged between years 0.5 and 2.5. The midlatitudes of both hemispheres in the stratosphere during this time have peak perturbations of about 10%, which is less than the amplitude measured by Oltmans et al. (2000). This could represent a failing in the IGCM’s representation of transport water vapor away from the LS Tropics over a timescale of a year or more, and is discussed below.

We hypothesize that transport of the extra water vapor in the LS region will be dependent on the actual temperature perturbation that the tropopause feels. To test the hypothesis, we repeat run 1, but with changed values for $A$, $B$, $C$, and $D$ (denoted run 2). The effects of slightly changing these constants are shown in Fig. 5, which shows the heating at the equator when $Q_t = 1$. In this
However, we show here that the rate of change of trend with height associated with water injection would not be very large. Figure 8 shows the linear trend associated with two Gaussian-shaped pulses of water vapor. One is centered in mid-1993 and representing a water vapor perturbation at the 30-km level; the other represents that perturbation a year later at the 45–50-km level. The location of the center of each curve is gleaned from an analysis of Fig. 1 of Geller et al. (2002). The peak amplitude of the pulse is 300 ppbv (when spread out globally), which represents a 10% increase to the IGCM mean stratospheric value of 3 ppmv, and is consistent with Fig. 3.

The difference between the two trends at each level is \( O(30 \text{ ppmv yr}^{-1}) \), which is entirely consistent with what was calculated by Evans et al. (1998); in other words, their trend analysis does not preclude transport of a water vapor pulse having a globally averaged size of \( O(10\%) \) of the background stratospheric level through the tropopause.

The Boulder balloon data actually show an increase in specific humidity of \( O(1) \text{ ppmv} \) in early 1992 at the 24–26- and 20–22-km levels (Oltmans et al. 2000). The two levels correspond approximately to pressures of 30 and 50 hPa, respectively. These humidity perturbations are much greater than those seen in the IGCM. The difference is certainly partially due to the oxidation of \( \text{CH}_4 \) producing more water throughout the stratosphere. Other phenomena such as ENSO, or model deficiencies such as vertical transport of water vapor through the hydrodropause, would also play a role.

Hopefully in the future, improvements in satellite technology since the Pinatubo eruption will ease identification of changes in LS water vapor following a large tropical eruption. Hopefully in this way our hypothesis can be tested directly against observations.

5. Discussion

The results generated by the IGCM, especially the differences in the results between runs 1 and 2, do show that the hypothesis presented by Stuber et al. (2001) can be demonstrated using a realistic tropopause temperature perturbation from a real forcing mechanism. However, we do not believe that the currently available observations allow us to categorically confirm the Stuber et al. (2001) hypothesis yet. The extra water vapor transported into the LS region will act as a positive radiative forcing agent in the same manner as in Stuber et al. (2001). Radiative forcing calculations performed using the extra LS water vapor shown in Fig. 4 give a globally averaged value of \( +0.1 \text{ W m}^{-2} \). The peak annually averaged direct aerosol forcing of Pinatubo was \( \approx -3.5 \text{ W m}^{-2} \) (e.g., Myhre et al. 2001), so the forcing due to increased LS water vapor acts modestly in the opposite sense to volcanic forcing of climate.

It should be stressed that the forcing value obtained above is extremely dependent on the actual location of

---

4. Comparison of the model with observations

As stated earlier, Evans et al. (1998) suggested that no water vapor was injected into the stratosphere by the eruption of Mount Pinatubo, using an analysis of trends in Halogen Occultation Experiment (HALOE) data.
the water perturbation. Similar perturbations at slightly different latitudes can yield much higher values of $>0.2$ W m$^{-2}$, especially when the water vapor perturbation spreads into the extratropics, as it is here that the greatest forcing in the longwave region occurs (Forster and Shine 2002).

The principal aim of our study is to investigate the transport of water into the LS region, not the transport of water into the middle and upper stratosphere. This is because water entering the LS region is not transported up to the 1–10-hPa level in the IGCM at the same speed that observations suggest (Geller et al. 2002). In addition, the timescale of water movement from tropopause to 1 hPa is over 2 yr: on these long
timescales, water concentration is strongly modulated by stratospheric chemistry, which has not been included in the IGCM as yet. In the future we suggest that the phenomenon might be examined in more complex models that have better water transport schemes, and include stratospheric chemistry.

Our results are intriguing when considered in the context of other volcanoes. Frost-point hygrometer data do suggest increases in water vapor concentration after the 1982 eruption of El Chichón (Oltmans et al. 2000). The saturation vapor pressure of water is a strongly nonlinear function of temperature. Such nonlinearity is suggested by comparing our results (tropopause warming of 1 K gives 10% extra water vapor) to those of Stuber et al. (2001) (tropopause warming of 5 K gives 100% extra water vapor), although we note that this is a comparison between two different climatic perturbations in two different models. Changes in the radiative properties of aerosol ejected by the volcano would also change the amplitude of the temperature anomaly produced at the tropopause, and hence the amount of anomalous stratospheric water produced.

Figure 2 of Geller et al. (2002) shows that the entry value of stratospheric water vapor is modified by the combined effects of ENSO and the QBO. Such changes in entry value, when combined with the effect of aerosol heating, will significantly change the long-term perturbation in LS water associated with any given volcanic eruption. Moreover, because of the nonlinearity in the Clausius–Clapeyron equation, the different influences on the entry value cannot simply be summed up, but must be studied together.

Finally we speculate on the long-term climatic effects
of repeated volcanic eruptions. Aerosol clouds resulting from volcanic eruptions dissipate after 1–2 yr (Robock 2000). However, age of air arguments suggest that the time it would take for stratospheric water vapor anomalies to be dissipated by transport is >5 yr (Hall and Waugh 1997). If the frequency of volcanic activity was high enough, then a water vapor anomaly would be introduced into the lower stratosphere before the anomaly due to the previous eruption had disappeared. The result would be threefold in the long term: stratospheric cooling, stratospheric humidification, and surface warming due to the positive radiative forcing associated with the water vapor. Oxidation of CH$_4$ into water vapor would also contribute to such an effect (Considine et al. 2001). Volcanically induced stratospheric humidification might also explain the long-term trend in stratospheric water vapor in spite of the long-term negative trend in cold-point tropopause temperature (Zhou et al. 2001).

In the period 1960–2000, five volcanoes have erupted with sufficient strength to cause a globally averaged stratospheric aerosol radiative forcing more negative than −1 W m$^{-2}$, compared with none in the period 1920–60 (Myhre et al. 2001). It is therefore possible that part of the long-term trend in stratospheric water vapor of 1% yr$^{-1}$ since 1955 (Rosenlof et al. 2001) may be due to volcanic activity since the eruption of Agung in 1963.

We illustrate the potential effect of repeated eruptions using the following simple model: we use the radiative forcing associated with volcanic eruptions (Myhre et al. 2001) as a zeroth-order proxy for the globally averaged stratospheric water vapor perturbation induced, and assume that the perturbation exponentially falls off with a timescale $\tau$, which is longer than the residence timescale of the aerosol itself. The effects of the five eruptions between 1963 and 1991 are then simply summed up. The globally averaged stratospheric water vapor perturbation due to Mount Pinatubo is assumed to be 15%, consistent with Fig. 3 and Considine et al. (2001), although in reality the perturbation could have been higher (Oltmans et al. 2000).

The results are shown in Fig. 9, which shows the evolution of water vapor in the simple model for two values of $\tau$: 5 and 10 yr. We believe that these two values constrain the actual destruction timescale of stratospheric water vapor given age of air values at the 60-km level, and the fast photolysis rate of water vapor at that height. Figure 9 displays water vapor trends of 0.25%–0.5% yr$^{-1}$, which account for a large fraction of the actual trend of 1% yr$^{-1}$ reported by Rosenlof et al. (2001). Interestingly, Mastenbrook and Oltmans (1983) found a decrease in water vapor over Boulder, Colorado, at the 60-hPa level in the middle to late 1970s, which followed an increase in the late 1960s. Such a pattern is hinted at by Fig. 9.

The simple model does illustrate that the repeated volcanic eruptions of the second half of the twentieth century could potentially explain some of the long-term increase in stratospheric water vapor during the same time. We aim to study this phenomenon in more detail in future.

**Acknowledgments.** This research was conducted while Manoj Joshi was funded by the European Commission through the Fifth Framework Programme (project EV2K-CT-1999-000021). We would like to acknowledge helpful comments made by Hannah Clark, Ellie Highwood, John Thuburn, Brian Topp, and the reviewers of the original manuscript.

**REFERENCES**


