Tropical Rainfall Trends and the Indirect Aerosol Effect

LEON D. ROTSTAYN
CSIRO Atmospheric Research, Aspendale, Victoria, Australia

ULRIKE LOHMANN
Department of Physics and Atmospheric Science, Dalhousie University, Halifax, Nova Scotia, Canada

(Manuscript received 30 May 2001, in final form 16 January 2002)

ABSTRACT
An atmospheric global climate model coupled to a mixed layer ocean model is used to study changes in tropical rainfall due to the indirect effects of anthropogenic sulfate aerosol. The model is run to equilibrium for present-day (PD) and preindustrial (PI) sulfur emission scenarios. As in two other recent studies, the model generally gives a southward shift of tropical rainfall in the PD run relative to the PI run. This is largely due to a hemispheric asymmetry in the reduction of sea surface temperature (SST) induced by the perturbation of cloud albedo and lifetime.

Observed precipitation trends over land for the period 1900–98 show a complex pattern in the Tropics, but when zonally averaged, a southward shift similar to (but weaker than) the modeled shift is clearly evident. The zonally averaged tropical trends are significant at the 5% level in several latitude bands. The modeled present-day hemispheric contrast in cloud droplet effective radius (which affects cloud albedo) is well supported by one long-term satellite retrieval, but not by another. A third satellite retrieval, which only covers an 8-month period, does show a marked hemispheric contrast in effective radius.

Both in the modeled changes and the observed trends, a prominent feature is the drying of the Sahel in North Africa. Modeled dynamical changes in this region are similar to observed changes that have been associated with Sahelian drought. Previous work has identified a near-global, quasi-hemispheric pattern of contrasting SST anomalies (cool in the Northern Hemisphere and warm in the Southern Hemisphere) associated with dry conditions in the Sahel. The present results, combined with this earlier finding, suggest that the indirect effects of anthropogenic sulfate may have contributed to the Sahelian drying trend. More generally, it is concluded that spatially varying aerosol-related forcing (both direct and indirect) can substantially alter low-latitude circulation and rainfall.

1. Introduction
The indirect aerosol effect has received much attention in the last few years, because of the uncertainty of its magnitude and its potential to substantially offset the positive forcing due to anthropogenic greenhouse gases (e.g., Haywood and Boucher 2000; Penner et al. 2001). Two likely effects of anthropogenic aerosols on liquid water clouds have been identified. The “first indirect effect” refers to the radiative impact of a decrease in cloud droplet effective radius that results from increases in aerosols (Twomey 1977). The “second indirect effect” refers to the radiative impact of a decrease in precipitation efficiency that results from increases in aerosols (Albrecht 1989). Observational support for the first indirect effect is substantial, and is based on in situ measurements as well as satellite retrievals (Coakley et al. 1987; Raga and Jonas 1993; Twohy et al. 1995; Hudson and Yum 1997; Kaufman and Fraser 1997; Fer- ek et al. 1998; Boers et al. 1998; Han et al. 1998; Ack- erman et al. 2000; Brenguier et al. 2000; Zhiming and Yung 2000; Heymsfield and McFarquhar 2001; Hudson and Yum 2001; Nakajima et al. 2001). In contrast, observational support for the second indirect effect is rather limited (Warner 1968; Rosenfeld 2000; Hudson and Yum 2001), and its climatic significance is still considered to be very uncertain.

Recent modeling work has suggested that the indirect aerosol effect has the potential to cause large-scale shifts in tropical circulation and rainfall patterns, because of its spatially nonuniform nature. Rotstayn et al. (2000b) used a version of the Commonwealth Scientific and Industrial Research Organisation (CSIRO) atmospheric global climate model (AGCM) coupled to a mixed layer (“q-flux”) ocean model, and ran the model to equilibrium for both present-day and preindustrial indirect aerosol forcing. In each case, the distribution of cloud droplet concentration was obtained from monthly mean
distributions of sulfate generated by a chemical transport model (Penner et al. 1994), using the empirical parameterization of Boucher and Lohmann (1995). The most prominent feature of the model’s response was a southward shift of the tropical rainfall in the present-day case relative to the preindustrial case, especially over the Pacific Ocean. Sensitivity tests showed that this shift was related to a larger sea surface temperature (SST) reduction in the Northern Hemisphere (NH) than the Southern Hemisphere (SH), due to a larger perturbation of cloud radiative properties in the NH. It was not a result of any direct impact of sulfate on the generation of convective rainfall, because no such interaction was included in the model. Williams et al. (2001) used a version of the Hadley Centre AGCM with an interactive sulfur cycle and a mixed layer ocean model, and also obtained a southward shift of tropical rainfall as a response to the indirect effect of sulfate aerosol. They showed that this was associated with a southward shift of the intertropical convergence zone (ITCZ), as measured by the change in meridional streamfunction.

In none of the above modeling studies have the authors considered the important question of whether the observed rainfall record in the Tropics bears any resemblance to the pattern of response obtained with the models. Validation of modeled responses (or attribution of observed trends) to specific radiative perturbations is a difficult task, even when the nature of the perturbation is relatively well understood (as in the case of anthropogenic greenhouse gases). The current limited understanding of the indirect aerosol effect renders this task even more difficult, because the magnitude and spatial distribution of the radiative perturbation is itself highly uncertain. In addition, any response of the real climate system to the indirect aerosol effect has occurred against a background of climate change induced by greenhouse gases and other factors, as well as natural variability. Despite these uncertainties, the striking nature of the response obtained with the above models invites at least a preliminary comparison between model and observations.

The main purpose of this study is to undertake such a preliminary comparison. As well as comparing the model with observed rainfall trends, we compare the modeled hemispheric contrasts in cloud droplet effective radius with those obtained from satellite retrievals. The motivation for the latter comparison is to see whether the large-scale modeled hemispheric contrasts in cloud radiative properties—which affect SST and hence tropical circulation and rainfall—are broadly supported by the satellite retrievals. In section 2 we briefly describe an updated version of the CSIRO AGCM that includes an interactive sulfur cycle. In section 3, we present results from the model and compare some of these with observations. An interesting result shown in section 3 is that the model simulates a drying in tropical North Africa, similar to that observed during the second half of the twentieth century. Because this drying trend has been extensively studied, we present in section 4 a minireview of previous research into this drying trend, and relate some of this research to the present study. Section 5 contains a summary and conclusions.

2. Model and experiments

We use a low-resolution (spectral R21) version of the CSIRO AGCM, with 18 vertical levels. The AGCM is an updated version of that described by Rotstayn (1997). For each of the present-day (PD) and preindustrial (PI) scenarios, it is coupled to a mixed layer ocean model with fixed monthly mean heat transports. It is run until approximate equilibrium is reached, and then for a further 10 years, for which period results are shown in the next section. The equilibration period is 20 years for the PD run, and 35 years for the PI run. In both runs, the concentration of CO2 is held at 345 ppmv, a value appropriate to the mid-1980s. The direct radiative effect of sulfate aerosol is omitted from these experiments.

The model includes a physically based treatment of stratiform clouds and precipitation, and interactive calculation of cloud radiative properties (Rotstayn 1997). The improved treatment of mixed-phase clouds described by Rotstayn et al. (2000a) is also included. The treatment of radiative transfer is based on an improved version of Lacis and Hansen (1974) in the shortwave and on Schwarzkopf and Fels (1991) in the longwave.

The first indirect aerosol effect enters the model through the effect of variations of cloud droplet number concentration (Nd) on the droplet effective radius (re) in liquid water clouds, and hence on their shortwave radiative properties. The parameterization of effective radius follows Martin et al. (1994), giving re proportional to N d 1/3. It should be noted that Martin et al. derived their scheme from measurements of stratocumulus clouds, and that we extend it to all the liquid water clouds in the model, in the absence of sufficient data for other types of cloud. When the parameterization from Martin et al. is combined with the usual relation for the cloud optical depth of liquid water clouds, namely,

$$\tau = \frac{3L}{2\rho r_e},$$  \hspace{1cm} (1)

where L is the liquid water path and \(\rho\) is the density of liquid water, we obtain the expected result that \(\tau \propto N d^{1/3}\) for fixed L (Twomey 1977). The optical depths are used to calculate the cloud reflectivities and absorptivities required by the shortwave radiation scheme, using a delta-Edington technique (Slingo 1989). For the calculation of radiative transfer, the fraction of each grid box occupied by any convective cloud is added to that occupied by any stratiform cloud, and L is calculated as a weighted mean of the liquid water path contributed by convective and stratiform cloud.

The second indirect effect enters through the parameterization of autoconversion (coalescence of cloud
droplets) in stratiform clouds. Autoconversion is suppressed until the mixing ratio of cloud liquid water reaches a critical value

\[ q_{\text{crit}} = \frac{4}{3} \pi \rho r_{\text{crit}}^3 N_d / \rho, \tag{2} \]

where \( r_{\text{crit}} = 7.5 \, \mu \text{m} \) is a prescribed critical cloud droplet radius and \( \rho \) is the air density. The autoconversion rate varies as \( N_d^{1/3} \), but it is the linear dependence of \( q_{\text{crit}} \) on \( N_d \) through (2) that effectively controls the response of the model to changes in \( N_d \). In common with other convection schemes used in AGCMs, the model’s mass flux convection scheme (Gregory and Rowntree 1990) includes only a simple treatment of the microphysics of rainfall generation, which does not depend on the distribution of aerosols. Thus, we consider the second indirect effect only for stratiform clouds.

As in several other studies, \( N_d \) is estimated empirically from the mass concentration of sulfate \((m)\), using relationship D from Boucher and Lohmann (1995), namely,

\[ N_d = \begin{cases} 10^6 \times 114.8 m^{0.48} & \text{over oceans}, \\ 10^6 \times 173.8 m^{0.26} & \text{over land}, \end{cases} \tag{3} \]

with \( N_d \) in \( \text{m}^{-3} \) and \( m \) in \( \mu \text{g} \text{ m}^{-3} \). The use of this scheme means that sulfate is used as a surrogate for all aerosols that act as cloud condensation nuclei, assuming that the fraction of sulfate in aerosols remains constant. Since this parameterization is based on measurements in regions affected by industrial aerosols, using it may underestimate the effect of aerosols from biomass burning on cloud properties, even though the model does include emissions of sulfur from biomass burning (see below). This is because the emission of sulfur from biomass burning is small compared to that from industrial sources, and sulfate generally represents a smaller fraction of the mass of biomass aerosols than industrial aerosols.

The model includes an interactive sulfur (S) cycle, based in part on that developed previously for the European Centre–Hamburg Model (ECHAM). A detailed description and evaluation of the sulfur cycle is given by Rotstayn and Lohmann (2002, hereafter ROLO). Prognostic variables are dimethyl sulfide (DMS) and sulfur dioxide \((\text{SO}_2)\) as gases and sulfate as an aerosol. Transport of these quantities is by semi-Lagrangian advection (McGregor 1993), turbulent mixing (Louis 1979; Holtslag and Boville 1993), and the convective mass flux (Gregory and Rowntree 1990). The treatment of sulfur chemistry is based on Feichter et al. (1996). DMS reacts with the hydroxyl (OH) radical during the day and with the nitrate \((\text{NO}_3)\) radical at night to form \( \text{SO}_2 \). \( \text{SO}_2 \) reacts in the gas phase with OH, and in the aqueous phase with hydrogen peroxide \((\text{H}_2\text{O}_2)\) and ozone \((\text{O}_3)\), in all cases to form sulfate. The treatment of wet scavenging is linked to the generation of stratiform and convective precipitation as described by

ROLO. The calculation of dry deposition follows Lohmann et al. (1999).

Emissions for the PD scenario are as described by ROLO, except that we have reverted to a less aggressive parameterization of the emission of DMS from the ocean surface (Liss and Merlivat 1986) than the one used in the “control” run of ROLO. (This change was made because results shown by ROLO suggested that a generally more realistic simulation of the sulfur cycle was obtained over the remote oceans when the less aggressive scheme was used.) The main anthropogenic source of sulfur is industrial (67 Tg S yr\(^{-1}\)), but there is also a small component from biomass burning (2.5 Tg S yr\(^{-1}\)). The industrial emissions are for the year 1985. The main natural sources of sulfur are emissions of DMS from the ocean surface (13 Tg S yr\(^{-1}\)) and of \( \text{SO}_2 \) from non-volcanic sources (8 Tg S yr\(^{-1}\)). The PI scenario is obtained by setting the industrial source to zero, and the biomass source to 10% of its PD value.

The column burden of sulfate is 3.42 mg m\(^{-2}\) in the PD run, and 0.95 mg m\(^{-2}\) in the PI run. The difference between these (2.47 mg m\(^{-2}\)) represents the anthropogenic component, and may be compared to the range of 1.14–4.0 mg m\(^{-2}\) from the 17 models reviewed by Ramaswamy et al. (2001). The geographical distribution of the difference in sulfate column burden between the PD and PI runs is shown in Fig. 1. The effects of the large anthropogenic sources in the NH and the smaller sources in the SH are clearly evident. In the PD run, the sulfate column burden is much larger in the NH (5.21 mg m\(^{-2}\)) than in the SH (1.64 mg m\(^{-2}\)), because of the large anthropogenic source in the NH. In the PI run, there is little difference between the hemispheres (1.00 mg m\(^{-2}\) in the NH vs 0.90 mg m\(^{-2}\) in the SH). Observational evidence compiled by Heinzenberg et al. (2000) suggests that the model may overestimate the hemispheric contrast in sulfate concentration over oceanic regions for the present climate. Based on surface observations with limited spatial and temporal coverage, they estimate a SH:NH concentration ratio of 0.41. In the PD run, the SH:NH ratio over oceans is 0.28 for the surface sulfate concentration and 0.37 for the sulfate column burden.

3. Results and comparison with observations

a. Radiative perturbation and surface temperature changes

Comparison of PD and PI simulations forced by prescribed SSTs is the usual method of evaluating the radiative perturbation due to indirect aerosol effects in an AGCM that includes the second indirect effect as well as the first (e.g., Haywood and Boucher 2000). This is because it is not possible to calculate a pure radiative forcing for the second effect, which involves changes to the hydrological cycle. (Calculation of a pure radiative forcing would require that everything other than
the initial perturbation be held fixed.) However, Rotsstyn and Penner (2001) showed that this method of calculating the change in net irradiance at the top of the atmosphere (which they termed a quasi forcing) provided a reasonable approximation to a pure forcing. We have therefore calculated this quasi forcing by rerunning the PD and PI scenarios using prescribed seasonally varying climatological SSTs, instead of using the mixed-layer ocean model. For each scenario, we allowed a spinup period of 1 year, and then ran the model for a further 8 years to obtain the net irradiance at the top of the atmosphere.

For the combined first and second indirect aerosol effect, our model gives a quasi forcing of $-1.8 \, \text{W m}^{-2}$ in the global mean. The quasi forcing equals $-2.9 \, \text{W m}^{-2}$ in the NH and $-0.7 \, \text{W m}^{-2}$ in the SH, giving a NH:SH ratio of 4.1. The global-mean value is not especially large compared to the range given by recent models that include the first and second indirect effects (Haywood and Boucher 2000), or relative to the range of 0 to $-2 \, \text{W m}^{-2}$ given for the first effect only by Ramaswamy et al. (2001). It is also substantially smaller in magnitude than the value of $-2.6 \, \text{W m}^{-2}$ obtained with the earlier version of the CSIRO model used by Rotstyn et al. (2000b).

One likely reason for the smaller indirect effect obtained with the new version of the model is the change from the use of monthly mean sulfate to an interactive sulfur cycle. Feichter et al. (1997) used the same parameterization of cloud droplet number as in the present study, and found that when they used monthly mean sulfate instead of an interactive treatment of the sulfur cycle, their estimate of the first indirect effect increased by 20% in the global mean and by more than 100% in some regions. Another likely reason is an alteration to the treatment of autoconversion. The treatment of autoconversion in Rotstyn et al. (2000b) was different from that in the present model in that it included a parameterization of the effect of subgrid moisture variations on the autoconversion process (Rotstyn 2000).

A side effect of including this modification in the model was an increase of $0.5 \, \text{W m}^{-2}$ in the globally averaged second indirect effect, for reasons that are discussed by Rotstyn (2000). This modification is not used in the present study.

The difference in annually and globally averaged near-surface air temperature at equilibrium between the PD and PI runs is $-1.34 \, \text{K}$. This may be compared with the figure of $+3.52 \, \text{K}$ obtained by Rotstyn and Penner (2001) for a doubling of CO$_2$, using a similar version of the model. Figure 2 shows the geographical distribution of the difference in near-surface air temperature between the PD and PI runs. (The difference in actual surface temperature is very similar.) The largest decreases occur at high latitudes, due to strong positive surface-albedo feedbacks. The surface cooling is generally larger in the NH than in the SH, and larger in the North Atlantic than in the North Pacific, as expected for a mechanism involving anthropogenic sulfate. The pattern of SST change does not correspond exactly to the distribution of anthropogenic sulfate shown in Fig. 1. This is in part because the change of SST in the model depends on the surface energy balance (net radiation minus the sum of latent and sensible heat fluxes) as well as the fixed oceanic heat transports. As the model’s atmospheric circulation responds to the applied radiative perturbation, changes in wind speed can greatly enhance or diminish the change in SST. An example is the tongue of relatively strong cooling that extends to the southwest from near Baja California. Examination of the modeled dynamical changes reveals that there is a marked acceleration and southward extension of the trade winds in this region, as the North Pacific anticyclone moves southward in the PD run (not shown). An increase of surface wind speed also seems to explain the region of strong cooling just west of Africa in the tropical North Atlantic (discussed further in section 4). Some small areas of slight warming are probably also due to dynamical feedbacks.
b. Comparison of modeled and observed precipitation

The annual-mean precipitation distribution from the PD run is shown in Fig. 3, together with observations averaged over the period 1980–99 from version 2 of the Global Precipitation Climatology Project (GPCP; Huffman et al. 1997). The model successfully captures the broad features of the distribution of precipitation for the present climate, including the ITCZ and the South Pacific convergence zone (SPCZ). Some deficiencies are evident, such as a poor representation of the monsoonal rainfall in northern Australia. Also, the model generally tends to overestimate precipitation; the global-mean values are 2.83 mm day$^{-1}$ in the model and 2.61 mm day$^{-1}$ in the observations.

Figure 4a shows the difference in annual-mean precipitation between the PD and PI runs. An overall southward shift of tropical precipitation is again prominent, as in Rotstayn et al. (2000b) and Williams et al. (2001), increasing our confidence that it is not just an artifact of a single model. A less prominent feature is a modest reduction in precipitation over NH midlatitudes. Rotstayn et al. (2000b) showed that this midlatitude response was caused by two factors in earlier experiments: a weakening of the hydrological cycle due to an overall cooling, and some microphysical suppression of stratiform rainfall through Eq. (2). Note that the tropical rainfall is mostly convective rather than stratiform, so microphysical suppression of rainfall (which is not included in the convection scheme) is not responsible for the tropical changes. In some regions, the pattern in the Tropics differs from that shown by Rotstayn et al. (2000b). Examples are over the Indian and South Pacific Oceans, where the current model gives a more complex pattern of response than the more uniform southward shift shown by Rotstayn et al. (2000b). In particular, the pattern over the South Pacific may be broadly described as an increase north of the SPCZ and a decrease south of the SPCZ. The model gives only small rainfall changes over tropical Australia, possibly because of the poor simulation of the summer monsoon there.

Do observed precipitation trends show evidence of a southward shift in the Tropics? Figure 4b shows observed annual-mean precipitation trends over land for the period 1900–98, based on a 5$^\circ$ gridded historical dataset, updated from Hulme (1994; see also Hulme et al. 1998). We excluded points for which at least 30 years of data were not available. The increases in the mid- to high latitudes of the NH are broadly consistent with model simulations driven by increases in greenhouse gases (e.g., Roeckner et al. 1999), an effect that is not included in our simulations. In the Tropics, the observations show a complex pattern. There is evidence of a southward shift over tropical Africa; the drying trend in tropical North Africa has been extensively studied, and is discussed further in section 4. There is a suggestion of a southward shift over the regions covering Southeast Asia and northern Australia, and Central and South America, although the pattern is more complex in these regions. In some parts of the Tropics (e.g., India), the trend in the observations is not suggestive of a southward shift. Over the Pacific Ocean, Hulme’s data do not show a very coherent pattern. However, in view of the modeled response over the South Pacific that was described above, it is interesting that Salinger et al. (1996) found there has been a decrease in rainfall south of the SPCZ, and an increase of rainfall north of the SPCZ, since the mid-1970s. As discussed by Salinger et al. (1995), there appear to be different climatic regimes on either side of the SPCZ, and these regimes differ, for example, in their response to the El Niño–Southern Oscillation (ENSO), as well as in their observed climatic trends. As we discuss in section 4, some of the observed tropical rainfall trends may be related to an El Niño–like pattern of warming in the Pacific, possibly due to increased levels of greenhouse gases.

As a reminder that the observed precipitation trends are subject to considerable uncertainty, one may com-
Fig. 3. Annual-mean precipitation in mm day$^{-1}$ (a) from the PD run, and (b) from version 2 of the GPCP for the period 1980–99. Contours are 1, 2, 4, and 7 mm day$^{-1}$.

We compare the pattern in Fig. 4b with the trends shown by Karl (1998), Easterling et al. (2000), and Folland et al. (2001). There are marked differences between these in some parts of the Tropics. In a comprehensive review, Folland et al. (2001) provide much useful additional information, such as a breakdown of the observed trends into three periods: 1910–45, 1946–75, and 1976–99. The middle period, characterized by little or no warming in the global mean, shows the most evidence of a southward shift of tropical rainfall. This was also a period during which global sulfur emissions grew rapidly (e.g., Lefohn et al. 1999). The surface temperature trends shown by Folland et al. (2001) for this period clearly show a quasi-hemispheric contrast, characterized by cooling in the NH and warming in the SH. This pattern is broadly consistent with the indirect (and direct) effects of sulfate aerosol, superimposed on a generalized warming due to greenhouse gases. However, another possible explanation is natural variability, as discussed below.

A useful way to simplify the comparison of modeled and observed precipitation changes is to zonally average the data. Figure 5 shows, in zonal-mean form, the trend in the observed precipitation data from Hulme, and the modeled change in precipitation over land points. (Note that the units differ, and the comparison is only qualitative, because we are comparing an observed trend with the difference between two equilibrium simulations.) The tropical southward shift is clearly evident in the observations, as well as in the model, although its magnitude is smaller in the observations. We tested the statistical significance of the zonally averaged observed trends using a two-tailed $t$ test; points where the trend differs from zero at the 5% level are shown as asterisks, while those that do not are shown as pluses. Evidently, the trends are significant at many of the points.

Figure 5 suggests the interesting possibility that the zonally averaged precipitation trends over the twentieth century could be understood, to a first approximation, as the result of the indirect aerosol effect (tropical southward shift) superimposed on the effect of a generalized warming due to increases in greenhouse gases. Outside the Tropics, the observed trends appear to be broadly consistent with those simulated by models forced by increases in greenhouse gases (e.g., Roeckner et al. 1999). This refers to the general increase in precipitation over the mid- to high latitudes of the NH, as well as the suggestion of a poleward shift of the storm tracks in the SH.
c. Does the model overestimate the hemispheric contrast in effective radius?

In the previous section, we found that (in response to the indirect effects of anthropogenic sulfate) the model gave a stronger southward shift of tropical rainfall than was suggested by the observed rainfall trends. In this section, we consider the possibility that this is because the model overestimates the hemispheric contrast in cloud properties for the present climate. Such an overestimate could arise from the use of a sulfate-dependent parameterization of cloud droplet number, since this is likely to underestimate the effect of biomass aerosols on cloud properties, and there are substantial sources of these in the SH (Liousse et al. 1996). It could also arise if the modeled hemispheric contrast in sulfate concentration is too large, as was suggested in section 2.

Of particular interest is whether satellite retrievals support the idea of a hemispheric contrast in droplet effective radius ($r_e$) over oceans, because the response of the model is substantially mediated through hemispheric contrasts of SST induced by changes in cloud albedo (Rotstayn et al. 2000b). The satellite-retrieved hemispheric contrasts in $r_e$ from Han et al. (1994) were used by Boucher (1995) to constrain an AGCM in order to calculate a rough estimate of the radiative forcing due to the first indirect aerosol effect. This calculation required him to make several assumptions, including that there is no anthropogenic effect on $r_e$ in the SH. Although this was a rather idealized calculation because of the assumptions that were made, it does illustrate that the hemispheric contrast in $r_e$ can be taken to represent the “signature” of the first indirect aerosol effect, at least to a first approximation.

In Table 1, we show the hemispheric contrasts in $r_e$ simulated in the PD run, and from the satellite retrievals of Han et al. (1994) and Kawamoto et al. (2001). Both retrievals are based on Advanced Very High Resolution Radiometer (AVHRR) measurements. We use data from
Han et al. for the period 1986–88, and from Kawamoto et al. for the period 1987–93. We have restricted the comparison to the region between 45°S and 45°N, which is roughly the area covered by the first retrieval. For both retrievals, data are available for January, April, July, and October; we have averaged data from these months to obtain an approximation to the annual mean.

Table 1 shows that the hemispheric contrasts in $r_e$ simulated by the model are generally well supported by the retrieval of Han et al. (1994), but not by the retrieval of Kawamoto et al. (2001). The data also suggest that the modeled $r_e$ are generally a little too large, although this is less crucial to the present discussion. Comparing the model with the data from Han et al., there is good agreement regarding the global and oceanic hemispheric contrasts in $r_e$, but the model has a smaller hemispheric contrast over land. According to Han et al. (1994), their retrieval may interpret dust particles as cloud droplets over desert regions; this could affect the hemispheric contrast in their retrieved effective radii over land. Another encouraging point is that the model and observations broadly agree that the hemispheric contrast in $r_e$ is larger over the Atlantic Ocean than over the Pacific Ocean. This is easily understood in the model, because there is more anthropogenic sulfate over the North Atlantic Ocean than over the North Pacific Ocean. A plausible explanation for why the model might overestimate the hemispheric contrast in $r_e$ over the Pacific Ocean is that the model appears to overestimate the anthropogenic enhancement of sulfate at points such as Oahu and Midway Island in the North Pacific, due to excessive transport from the North American continent (ROLO).

Disappointingly, the model shows much less agreement with the retrieval from Kawamoto et al. (2001) in terms of hemispheric contrasts. Their data show a much weaker hemispheric contrast in $r_e$ over both land and global oceans. Over the Atlantic and Pacific Oceans, they have a hemispheric contrast that is of opposite sign, though small. (A large hemispheric contrast over the Indian Ocean in their data is sufficient to give a positive hemispheric contrast over oceans in the global mean.) So, whereas the data from Han et al. (1994) increased our confidence in the model, those from Kawamoto et al. (2001) have reduced it.

Another recent retrieval of $r_e$ is from Bréon and Colzy (2000), based on measurements from the Polarization and Directionality of the Earth Reflectances (POLDER) instrument. Data from POLDER are only available for the 8-month period from November 1996 to June 1997, and spatial coverage is somewhat limited. Bearing this in mind, it is encouraging that this retrieval does broadly support the modeled hemispheric contrasts in $r_e$. Averaged over the 8-month period, Bréon and Colzy give a hemispheric contrast in $r_e$ of 1.25 μm over the region from 45°S to 45°N. This amounts to 1.10 μm over land, and 1.03 μm over oceans. (Note that it is not a contradiction that the hemispheric contrast is larger globally than over land and ocean separately. This arises because there is more ocean in the SH than in the NH, and because $r_e$ is larger over the oceans.)

Overall, there is a suggestion from the satellite retrievals that the model does overestimate the hemispheric contrast in effective radius over oceans. Such an overestimate could explain the stronger tropical southward shift in the modeled rainfall changes compared to the observed trends in Fig. 5. The severity of the overestimate is very uncertain because of the disagreement among the satellite retrievals. Further observational studies to reduce this level of disagreement would be most valuable.

4. A closer look at the tropical North African drying trend

A prominent feature of both the observations and model results shown in Fig. 4 was the drying of tropical North Africa. The dry conditions that have persisted
since the late 1960s in tropical North Africa have been well documented. For example, Ward (1998) showed that the drying trend over the period 1950–87 is statistically significant over large areas. The most affected have been the western regions. The Sahel, located between roughly 12° and 18°N, is a semi-arid region, and the drought there has resulted in much human suffering (e.g., Druyan 1989). The driest decade was the 1980s, with some recovery occurring during the 1990s (Nicholson et al. 2000). Some attempts to explain the drying trend have considered the effects of land-cover changes (e.g., Charnley et al. 1975; Xue and Shukla 1993), but most research has focused on the importance of SST anomalies. In this section, we briefly review some of this research and link it to the perspective of the present study.

Early attempts to identify changes in SST or circulation associated with variations in Sahelian rainfall focused on the tropical Atlantic Ocean (Lamb 1978a,b). Subsequently, Hastenrath (1990) documented a number of decadal-scale changes in the tropical Atlantic sector that accompanied the Sahelian drying trend during the period 1948–83. These included a surface cooling in the North Atlantic (between 15° and 30°N) and a smaller contrasting warming in the South Atlantic. The main associated dynamical changes were a pressure increase over the tropical North Atlantic (strongest between 15° and 30°N), a southward shift of the near-equatorial wind confluence, and a strengthening of the northeast trades (and possibly the southeast trades). The identified changes were present in all seasons, but were most pronounced at the height of boreal summer (July–August). From the perspective of the present study, the last point is consistent with a mechanism related to sulfate aerosol, because the incoming solar flux is largest in summer. Also, surface observations of sulfate at Barbados in the tropical North Atlantic show a maximum during the summer, due to transport from the North American continent (Savoie et al. 1989). Broadly similar Atlantic Ocean changes were found by Lamb and Peppler (1992) for 3 of the 4 driest years since 1940 (1972, 1977, 1983, 1984). The exception was 1983, possibly because this was a strong El Niño year. The link between ENSO and Sahelian rainfall is discussed below.

As reviewed by Grist and Nicholson (2001) there is some controversy in the literature regarding whether the ITCZ is generally displaced southward during dry years in the Sahel. They argue that while a southward shift can be a factor in Sahelian drought, it is not a necessary or sufficient condition. Their study compared observed data for June–September from a 4-yr “dry” composite (1982–85) and a 4-yr “wet” composite (1958–61). They confirmed two of the most consistently demonstrated dynamical differences between the boreal summers of wet and dry years. These were a stronger African easterly jet (at 600 hPa and 10°N) and a weaker tropical easterly jet (centered near 200 hPa and 5°–10°N) in dry years than in wet years (e.g., Newell and Kidson 1984). They also found a weaker southwesterly monsoonal flow in dry years than in wet years. As noted by Grist and Nicholson (2001), the specific ways in which the ITCZ, the southwesterly monsoon, the African easterly jet, and the tropical easterly jet influence Sahelian rainfall variability are not fully understood.

In a detailed review of the state of knowledge in the late 1980s, Druyan (1989) discussed the importance of these different factors. The monsoon provides the main supply of moisture required for the summer rainfall in the Sahel, although it is not clear that the moisture flux is always smaller in dry years. A strong tropical easterly jet is thought to provide the upper-level divergence necessary for the development of wave disturbances that bring much of the rainfall to the Sahel. The strength of the tropical easterly jet is sensitive to meridional gradients of temperature in the troposphere, and hence may also be affected by meridional gradients of SST (thus providing a link to the perspective of the present study). Druyan (citing Newell and Kidson 1984) considered that the strength of the lower African easterly jet in dry years in the Sahel was more likely to be a response to dry and hot conditions (via thermal wind effects) than a causal factor. Grist and Nicholson (2001) confirmed this view, and argued that the importance of the African easterly jet lies in its latitudinal location rather than its strength. In dry years, the jet is displaced southward, thereby reducing both the horizontal and vertical wind shear over the Sahel. Regarding the location of the ITCZ, it is necessary to make a distinction between the Atlantic Ocean and the African continent (Druyan 1989). While several studies (e.g., Lamb 1978a,b; Hastenrath 1990) have identified a southward shift of the Atlantic ITCZ in dry years, Grist and Nicholson found no appreciable difference in its location over Africa between their wet and dry composites.

Are the dynamical changes in our model in the Atlantic–African region similar to those identified above? We focus on the July–August period, when the bulk of the rain falls in the Sahel. Figure 6 shows the difference in sea level pressure between the PD and PI runs. The model gives an increase of sea level pressure north of about 7°N, with the largest increases north of 15°N, where Hastenrath’s analysis also gave the largest increases. Figure 7a shows the 900-hPa winds from the PD run. The northeast and southeast trades are clearly evident, as is the southwesterly monsoon, which carries moisture from the Gulf of Guinea toward the Sahel. Figure 7b shows the difference in 900-hPa winds between the PD and PI runs. There is an acceleration of the trade winds to the south of the stronger Azores anticyclone, as well as a modest weakening of the southwesterly monsoon. These features are similar to the observed changes described above. (Note that, as discussed above, the stronger easterly winds appear to enhance the surface cooling in the tropical North Atlantic, thus providing a positive feedback between the dynamical changes and the SST changes.) Figure 8 shows the
zonal wind speed, averaged over the region from 15°W to 15°E, from the PI and PD runs. The changes identified by Grist and Nicholson (2001) are also qualitatively present in the model. Compared to the PI run, the PD run has a weaker southwesterly monsoon (i.e., weaker westerlies near the surface), a stronger African easterly jet, and a weaker tropical easterly jet. Also, the center of the African easterly jet is displaced southward in the PD run. There is also a small southward shift of the near-equatorial wind confluence, more marked over the Atlantic Ocean than over Africa (not shown). In summary, there is an encouraging degree of similarity between the dynamical changes in the model, and those observed in previous studies.

It is now understood that SSTs outside the Atlantic sector can have a marked influence on Sahelian rainfall. Folland et al. (1986) identified a near-global, quasi-hemispheric pattern of contrasting SST anomalies (most pronounced in the Atlantic) that was strongly correlated with Sahelian rainfall. As reviewed by Hulme (2001), "the dominant SST anomaly configuration associated with the Sahelian dessication has been an interhemispheric temperature contrast," characterized by warm SSTs in the SH and cool SSTs in the NH. This pattern is the "SST EOF3" calculated by Folland et al. (1991) (see also Rowell et al. 1995; Ward 1998; Hulme 2001). This low-frequency relationship accounts for a large fraction of the long-term trend in Sahelian rainfall, without being able to account for interannual variations around this trend. The high-frequency variations in Sahelian rainfall seem to be related more to ENSO or to regional SST anomalies in the tropical Atlantic (e.g., Ward 1998). Experiments with AGCMs forced by observed SSTs have shown substantial skill at reproducing the interannual variability (Rowell et al. 1995) or the trend (Smith 1994) in North African rainfall, thus strengthening the argument that SST forcing is a crucial factor that modulates Sahelian rainfall. Another interesting point is that the partial recovery in tropical North African rainfall during the 1990s has coincided with a strong warming in the extratropical North Atlantic since the mid-1980s (Folland et al. 2001). This warming has reversed the trend shown by Folland et al. (2001) for the period 1946–75, when the North Atlantic cooled while the South Atlantic warmed.

What is understood about the possible cause(s) of this near-global quasi-hemispheric SST contrast? Ward (1998) has suggested that it may represent the effect of a number of low-frequency regional climatic variations that have tended to coincide during the twentieth century. There is evidence from measurements and proxy-based reconstructions of surface temperature that there exists a multidecadal oscillation affecting the North At-
lantic (Delworth and Mann 2000), and possibly other oceans as well (Minobe 1997). As discussed by Delworth and Mann (2000), this multidecadal oscillation may be a result of internal variability of the coupled ocean–atmosphere–ice system as expressed through variations of the thermohaline circulation in the North Atlantic. Thus, the quasi-hemispheric SST contrast may be, at least in part, a result of the internal natural variability of the climate system.

As discussed above, the quasi-hemispheric SST contrast is also broadly consistent with the indirect (and direct) effects of sulfate aerosol, superimposed on a generalized warming due to greenhouse gases. It is interesting that increasing expression of this SST pattern occurred between 1946 and 1975 (Folland et al. 2001), because that is the period during which global sulfur emissions grew rapidly. Since the mid-1970s, there has been a marked downward trend in sulfur emissions from Europe (Mylona 1996) and North America (U.S. Environmental Protection Agency 2000). For example, according to the U.S. Environmental Protection Agency, total U.S. sulfur emissions were 24% lower in 1990 than in 1970, and 39% lower in 1996 than in 1970. While this does not demonstrate that sulfate is the primary cause of the quasi-hemispheric SST contrast (or the drying trend in tropical North Africa), it does suggest it as a possible contributing factor. One possible interpretation is that increasing levels of sulfate during 1946–75 acted to enhance the development of an SST pattern that would have occurred anyway due to natural variability.

Could an El Niño–like SST pattern be responsible for the drying trend in tropical North Africa? There is a correlation between El Niño years and dry years in the Sahel, a correlation that has become stronger since about 1970 (Janicot et al. 1996). According to Rowell (2001), the link is principally due to a direct atmospheric teleconnection, but there is also an indirect link via the effect of ENSO on Atlantic SST. The warming trend in the Pacific Ocean since about 1970 has shown a mean El Niño–like pattern, with stronger warming in the tropical eastern Pacific than the tropical western Pacific (Meehl and Washington 1996), and some climate models have provided evidence that this may be a response to greenhouse warming (Meehl and Washington 1996; Knutson and Manabe 1998; Cai and Whetton 2000). Thus, some of the tropical precipitation trends shown in Fig. 4b (in Africa and elsewhere) may be a response to a mean El Niño–like pattern of warming due to greenhouse gases. On the other hand, it appears that ENSO does not fully explain either the interannual variability or the long-term trend in Sahelian rainfall (e.g., Druyan 1989; Ward 1998). It should also be noted that no transient GCM integrations published to date (whether forced by greenhouse gas increases or unforced) have demonstrated a Sahelian drying trend comparable to that observed, so in that sense a convincing explanation of the trend has not yet been given.

5. Summary and conclusions

A version of the CSIRO AGCM with an interactive sulfur cycle, coupled to a mixed layer ocean model, generally gave a southward shift of tropical rainfall in response to the indirect effects of sulfate aerosol. This result was broadly similar to results published previously by Rotstayn et al. (2000b) and Williams et al. (2001). The response of the model is largely due to a hemispheric asymmetry in the reduction of SST induced by the perturbation of cloud albedo and lifetime entailed in the indirect aerosol effect.

Observed rainfall trends over land for the period 1900–98 showed a complex pattern in the Tropics, suggestive of a southward shift in some, but not all regions. The southward shift was evident over Africa, and to a lesser extent over Southeast Asia and northern Australia, and Central and South America. When zonally averaged, the observed trends clearly showed a southward shift in the Tropics, similar to, but weaker than the modeled shift. The zonally averaged trends were significant at the 5% level in several latitude bands.

We also considered three satellite retrievals of cloud droplet effective radius, to see whether the hemispheric...
asymmetry in cloud properties given by the model was supported. The modeled hemispheric contrasts in effective radius were generally well supported by one long-term satellite retrieval (Han et al. 1994), but not by another (Kawamoto et al. 2001). A third retrieval, covering a period of only 8 months, and with limited spatial coverage (Bréon and Colzy 2000) did broadly support the model. Overall, there is a suggestion from the satellite retrievals that the model does overestimate the hemispheric contrast in effective radius over oceans to some extent (perhaps as a result of using a parameterization that depends only on sulfate). This is a possible explanation for the stronger southward shift of tropical rainfall in the model than in the observations.

Note that the comparison with satellite-retrieved effective radii only relates to the first indirect effect, and we have made no attempt to evaluate the more elusive second indirect effect. As shown by Rotstayn et al. (2000b), the modeled response to the first indirect effect by itself is likely to be similar to, but weaker than, the response to the combined first and second indirect effects. Thus, our conclusions do not depend heavily on the inclusion of the very uncertain second indirect effect in the model. However, a possible overestimate of the second indirect effect is another plausible explanation for the stronger southward shift of tropical rainfall in the model than in the observations.

Both in the modeled changes and the observed trends, a prominent feature is the drying of tropical North Africa. In part because of the human cost of the extended Sahelian drought during the latter part of the twentieth century, there is an extensive body of research into the variability of rainfall there. Some of this research has identified a quasi-hemispheric pattern of contrasting SST anomalies (cool in the Northern Hemisphere and warm in the Southern Hemisphere) associated with dry conditions in this region. Our results, combined with this earlier finding, suggest that the indirect aerosol effect may have contributed to the observed drying trend in the Sahel. However, we do not discount possible contributions due to land-cover changes, multidecadal SST oscillations associated with the natural variability of the coupled atmosphere–ocean–ice system, or the development of a mean El Niño-like pattern of warming in the Pacific toward the end of the twentieth century.

Although the model also gave a strong rainfall response over the Pacific Ocean, we are less confident of the validity of our simulations there. Firstly, the amount of anthropogenic sulfate over the North Pacific is not very large, and observed levels are lower than those given by the model at locations such as Oahu and Midway Island (ROLO). Even the satellite-retrieved effective radii from Han et al. (1994), which broadly supported the model over the Atlantic, showed a much smaller hemispheric contrast than the model over the Pacific. Secondly, any Pacific signal that does resemble the modeled response could easily be swamped by other factors, such as changes in the behavior of ENSO (perhaps driven by greenhouse gas increases). Indeed, a preliminary evaluation of 925-hPa winds from the National Centers for Environmental Protection (NCEP) reanalysis did not indicate that the trade winds in the eastern tropical North Pacific were stronger or extended any farther south in the 1970s than in the 1950s, as suggested by our model. On the other hand, the Atlantic sector is the region where one would expect to see most clearly an observed signal due to the indirect effects of anthropogenic sulfate. As discussed in the previous section, we found good qualitative agreement between modeled and observed dynamical changes in the low-latitude Atlantic–African region. These dynamical changes were similar to changes that have been associated with Sahelian drought in several earlier studies.

We intentionally excluded anthropogenic forcings other than the indirect aerosol effect from our simulations, in order to isolate the response to this perturbation. Even regarding the indirect aerosol effect, our model is quite idealized, in the sense that we used only a simple parameterization of the effect of sulfate on cloud droplet number. As discussed above, the indirect effects of aerosols from biomass burning have not been properly accounted for in our model, and allowing for these is likely to alter the response. We have not attempted to include the direct radiative effects of sulfate or nonsulfate aerosols. As indicated by Ramanathan et al. (2001), the radiative impact of absorbing aerosols over the northern Indian Ocean is large, and there is likely to be a marked response of the atmospheric circulation to such a radiative perturbation. The Sahara is a large source of dust aerosols, and measurements at Barbados indicate that more dust traverses the Atlantic during years that are dry in the Sahel (Prospero and Nees 1977). This could be due to changes in the land surface, or to the stronger easterly winds that seem to occur there in dry years (Newell and Kidson 1984). Since dust aerosols have a substantial impact on both shortwave and longwave radiation (Ramaswamy et al. 2001), possible feedbacks on tropical North African rainfall involving Saharan dust should not be ruled out.

The discussion in the previous paragraph suggests the more general conclusion, independent of our specific results, that spatially varying aerosol-related forcing (both direct and indirect) can substantially alter low-latitude circulation and rainfall. Indeed, while completing our final revisions to this paper, we became aware of the study by Xu (2001), who relates the recent summertime tendency toward “north drought with south flooding” in eastern China to a southward shift of the summer monsoon rainy belt induced by the direct radiative effects of sulfate aerosol. He also mentions the possible relevance of direct forcing by European sulfate to Sahelian rainfall. Further studies are desirable, with more complete treatment of aerosol–cloud interactions, and detailed treatment of direct aerosol effects (including dust and carbonaceous aerosols). Studies using fully coupled atmosphere–ocean models are also desirable, to
enable a more quantitative comparison between modeled and observed trends.

Acknowledgments. This paper was markedly improved thanks to constructive comments from Martin Dix and the anonymous reviewers. This work was supported in part by the Australian Greenhouse Office.

REFERENCES


