Diagnosis of Middle-Atmosphere Climate Sensitivity by the Climate Feedback–Response Analysis Method

XUN ZHU,* JENG-HWA YEE,* MING CAI,* WILLIAM H. SWARTZ,* LAWRENCE COY,# VALENTINA AQUILA,@ ROLANDO GARCIA,& AND ELSAYED R. TALAAT**

* Applied Physics Laboratory, The Johns Hopkins University, Laurel, Maryland
+ Department of Earth, Ocean and Atmospheric Science, Florida State University, Tallahassee, Florida
# Science Systems and Applications, Inc., Lanham, Maryland
@ GESTAR, NASA Goddard Space Flight Center, Greenbelt, and The John Hopkins University, Baltimore, Maryland
& Atmospheric Chemistry Division, NCAR, Boulder, Colorado
** Heliophysics Division, NASA, Washington, D.C.

(Manuscript received 12 January 2015, in final form 24 August 2015)

ABSTRACT

The authors present a new method to diagnose the middle-atmosphere climate sensitivity by extending the climate feedback–response analysis method (CFRAM) for the coupled atmosphere–surface system to the middle atmosphere. The middle-atmosphere CFRAM (MCFRAM) is built on the atmospheric energy equation per unit mass with radiative heating and cooling rates as its major thermal energy sources. MCFRAM preserves CFRAM’s unique feature of additivity, such that partial temperature changes due to variations in external forcing and feedback processes can be added to give a total temperature change for direct comparison with the observed temperature change. In addition, MCFRAM establishes a physical relationship of radiative damping between the energy perturbations associated with various feedback processes and temperature perturbations associated with thermal responses. In this study, MCFRAM is applied to both observations and model output fields to diagnose the middle-atmosphere climate sensitivity. The authors found that the largest component of the middle-atmosphere temperature response to the 11-yr solar cycle (solar maximum vs solar minimum) is the partial temperature change due to the variation of the solar flux. Increasing CO₂ cools the middle atmosphere, whereas the partial temperature change due to changes in O₃ can be either positive or negative. The application of MCFRAM to model dynamical fields reconfirms the advantage of introducing the residual circulation to characterize middle-atmosphere dynamics in terms of the partial temperature changes. The radiatively driven globally averaged partial temperature change is approximately equal to the observed temperature change, ranging from −0.5 K near 25 km to −1.0 K near 70 km between solar maximum and solar minimum.

1. Introduction

The warming of Earth’s surface and lower atmosphere due to increases in greenhouse gases (GHG) is associated with enhanced middle-atmosphere cooling and a possible strengthening of the Brewer–Dobson circulation through radiative–dynamical coupling. Because both the air density and the optical depths of major radiatively active species decrease with altitude, the physical state of the middle atmosphere as represented by various parameters such as temperature and winds is quite sensitive to climate forcing and is thus an important part of the signature of GHG-driven climate change (e.g., Akmaev et al. 2006). Hence, a more accurate quantification of the middle-atmosphere response to solar variability and anthropogenic changes in trace species is necessary to improve predictions of climate change.

Complex global circulation models are often used to study the combined climate response to specified changes in carbon dioxide, ozone, aerosols, or the solar forcing. However, the problem of identifying the individual contribution to the total predicted climate response in temperature and wind patterns by a specified change remains. A sensitivity investigation, where only one change at a time is specified in a numerical model, can give some information but neglects potentially important interactions between the different forcing...
perturbations. Moreover, examination of the climate evolution documented by reanalysis products and observations needs to include all historic forcing changes. It will be useful to be able to interpret the total climate change as a sum of perturbations with each contribution to the sum associated with specific forcing perturbations. One approach to identify forcing contributions is to use a radiation algorithm to linearly map the energy perturbations associated with forcing changes at all altitudes to small-amplitude temperature perturbations. An example of this approach is the climate feedback–response analysis method (CFRAM) that separates and estimates temperature responses due to external forcing and various climate feedbacks in the coupled troposphere–surface system based on data or model outputs that have already included all the resultant changes (Lu and Cai 2009, hereafter LC09; Cai and Lu 2009, hereafter CL09).

CFRAM is formulated based on the atmosphere–surface energy equation, and it explicitly decomposes the directly observable temperature change into partial temperature changes due to individual external forcing and feedback processes (LC09; CL09). The unique feature of CFRAM is that this decomposition into partial temperature changes is linearly additive, such that the sum of the partial temperature changes gives a total temperature change that can be compared directly to the observed temperature change at every point in space. From a modeling perspective, the so-called external forcing and its variation are akin to independent variables or parameters that would be prescribed as input values in the model. On the other hand, the feedback or internal processes of a system are similar to dependent variables or parameters that often constitute a set of model output values.

In this paper, CFRAM is extended to the middle atmosphere based on three physical features of this region: (i) the air density varies with altitude by several orders of magnitude and the energy deposition per unit mass generally slowly varies with altitude or log pressure, (ii) radiative energy exchange that can be directly evaluated from satellite observations plays a major role in the energy budget, and (iii) the energy flux associated with Earth’s surface can be considered as external forcing to the layered middle atmosphere. As a result, the middle-atmosphere climate feedback–response analysis method (MCFRAM) is formulated using an energy equation per unit mass in the commonly used units of kelvins per day for the middle atmosphere. It can be applied to both satellite observations and output fields of three-dimensional (3D) chemistry–climate models (CCM) to derive various partial temperature changes.

Because of gravity and planetary wave forcing, the middle atmosphere is often far from radiative equilibrium. Dynamical effects make important contributions to the middle-atmosphere energy budget, either through the eddy heat flux divergence or through adiabatic heating due to vertical motions. Therefore, changes to the middle-atmosphere climate due to dynamics need to be considered in addition to radiative forcing. With the radiative contributions being explicitly calculated in MCFRAM, the circulation changes can either be evaluated as residual terms in the energy budget equation or calculated directly if the winds are available. The newly developed MCFRAM method allows us for the first time to attribute temperature changes or anomalies in the middle atmosphere derived from observations and from numerical models to individual processes such as the solar cycle, anthropogenic greenhouse gas increases, changes in ozone, and changes in the Brewer–Dobson circulation. In this study, we illustrate the utility of MCFRAM by applying it to observations made by the Sounding of the Atmosphere using Broadband Emission Radiometer (SABER), a broadband radiometer on board the Thermosphere, Ionosphere, Mesosphere, Energetics and Dynamics (TIMED) satellite, and to the output of the Goddard Earth Observing System chemistry–climate model [GEOSCCM; Pawson et al. (2008) and references therein] over one solar cycle. The objective of the analysis is to ascertain the main factors that are responsible for the solar cycle variations in the middle atmosphere recorded in the SABER observations. The analysis of the GEOSCCM solar cycle climate simulations allows us to gain further insight into the role of the eddy-driven residual circulation in the middle atmosphere in response to solar cycle forcing.

In section 2, we extend CFRAM to the middle atmosphere. Then, we perform an eigenmode analysis of the generalized damping matrix derived from MCFRAM. The middle-atmosphere temperature and ozone fields needed in the analysis are derived from the TIMED/SABER instrument. Section 3 applies MCFRAM to the SABER observations while section 4 performs a set of more detailed MCFRAM analyses of the GEOSCCM. Section 5 summarizes the results.
2. Review and extension of the coupled feedback–response analysis method

This section presents the derivation of the MCFRAM equations. After a brief review of the CFRAM assumptions, the column energy balance equation for the middle atmosphere is used as a starting point for formulating the new MCFRAM equations.

a. Formulation of the middle-atmosphere CFRAM

CFRAM was originally formulated in the form of a vertical energy flux difference between two levels of a layered atmosphere for a single-column time-mean energy balance equation (LC09; CL09):

$$R^*(T, r, s, \ldots, \alpha, \beta, \ldots) = S^*(T, r, s, \ldots, \alpha, \beta, \ldots) + Q^*(T, r, s, \ldots, \alpha, \beta, \ldots),$$  

(1)

where $R^*$ and $S^*$ are column vectors of the infrared and solar flux differences corresponding to the vertical profiles of total radiative cooling and heating in each layer, respectively; $Q^*$ is the nonradiative energy flux difference in each layer; $T$ is temperature profile; $(r, s, \ldots)$ are the mixing ratios of radiatively active species such as CO$_2$, O$_3$, H$_2$O, and clouds; and $(\alpha, \beta, \ldots)$ are parameters such as the solar irradiance at the top of the atmosphere (TOA), surface albedo, and solar declination angle. The terms in Eq. (1) for CFRAM have the units of energy flux (W m$^{-2}$), which corresponds to the integrated heating or cooling rate per unit volume in a given layer. There are several advantages of adopting the flux form with units of watts per square meter in the conventional CFRAM: (i) the energy flux of the atmosphere can be seamlessly merged with the surface energy flux, (ii) the TOA version of CFRAM can be directly compared to a TOA-based climate feedback analysis such as the partial radiative perturbation method (Wetherald and Manabe 1988), and (iii) the layer thickness of the tropospheric CCMs is usually slowly varying in mass so the heating or cooling rate perturbations per unit space of different layers also vary slowly with altitude.

The air density decreases with altitude exponentially in the middle atmosphere, ranging from the tropopause ($\sim$10 km) to the turbopause ($\sim$110 km), spanning several orders of magnitude in density variation. The energy deposition or the atmospheric heating rate in models and remote sounding is often scaled by mass, and a vertical grid is used that varies as altitude or log pressure. As a result, we begin by developing our MCFRAM from an energy equation per unit mass—that is, by dividing Eq. (1) by $c_p \rho \Delta z$ with $c_p$, $\rho$, and $\Delta z$ being the specific heat at constant pressure, air density, and layer thickness, respectively,

$$R(T, r, s, \ldots, \alpha, \beta, \ldots) = S(T, r, s, \ldots, \alpha, \beta, \ldots) + Q(T, r, s, \ldots, \alpha, \beta, \ldots) + Q_{mol}(T),$$  

(2)

where $R$ and $S$ are the infrared radiative cooling and solar flux heating rates, respectively, and $Q$ is the nonradiative heating rate excluding the molecular thermal conductivity $Q_{mol}(T)$, which is only a function of temperature profile $T$ (Banks and Kockarts 1973). The units of all terms in Eq. (2) are kelvins per day. We now consider two statistical equilibrium states, 1 and 2, with two different sets of corresponding atmospheric parameters, each set satisfying the energy balance equation [Eq. (2)]. In practice, these two states can be two ensemble, time, or spatially averaged states. The difference of the energy balance equations between these two states is

$$\Delta(R - Q_{mol}) = \Delta S + \Delta Q.$$  

(3)

We now introduce the linear approximation to the responses of $R$ and $Q_{mol}$ due to the temperature variation and separate this term from the variations due to other parameters:

$$\Delta(R - Q_{mol}) \approx \frac{\partial(R - Q_{mol})}{\partial T} \Delta T + \Delta R,$$  

(4)

where $\Delta R = [R(T_1, r_2, s_2, \ldots, \alpha_2, \beta_2, \ldots) - R(T_2, r_1, s_1, \ldots, \alpha_1, \beta_1, \ldots)]$ is the change in total cooling rate due to all parameters except the temperature profile for which we set $\tilde{T} = (T_1 + T_2)/2$ to be the “mean temperature profile” between $T_1$ and $T_2$. Substituting Eq. (4) into Eq. (3), we obtain

$$\Delta T = A^{-1}(\Delta S - \tilde{\Delta} R + \Delta Q),$$  

$$A = \frac{\partial(R - Q_{mol})}{\partial T},$$  

(5)

where $A = \partial(R - Q_{mol})/\partial T$ is the generalized damping matrix (day$^{-1}$) calculated from the temperature dependence of infrared radiative cooling rate and molecular diffusion.

In the original CFRAM, where the surface and the atmosphere are in direct contact and thus strongly coupled dynamically and radiatively, the discretization of the energy equation [Eq. (1)] and the derivation of the “Planck feedback matrix” $\partial R^*/\partial T$ based on the temperature profile need to include the temperatures of both the surface level and layers of the atmosphere (LC09). The surface temperature and atmospheric temperature are treated as equally important in the setting of the problem. As a
result, Eq. (1) needs to be in the form of energy flux difference in units of energy flux \((W \cdot m^{-2})\). The middle atmosphere is not in direct contact with Earth’s surface. It can therefore be discretized solely based on a layered atmosphere in an energy equation per unit mass [Eq. (2)]. The effect of the energy flux emergent from the lower boundary on the middle atmosphere is primarily the radiative flux that can be considered as an external forcing. For example, the effect of the solar radiative flux can often be parameterized by an effective albedo of the surface and lower atmosphere \((\omega_0)\) caused by surface reflection and multiple scattering of clouds, aerosols, and air, which enhances the heating rate due to absorption of the Chappuis bands (410–750 nm) by ozone in the stratosphere (e.g., Meier et al. 1982; Nicolet et al. 1982). The calculation of the generalized damping matrix \(\mathbf{A}\) for a basic state of temperature and species distributions can be implemented by a radiation algorithm and molecular diffusive formulation. In this paper, The Johns Hopkins University Applied Physics Laboratory (JHU/APL) middle-atmosphere radiation algorithm (Zhu 1994, 2004) is adopted for radiative cooling calculations, and a temperature-dependent thermal conductivity of \(\lambda = 5.6 \times 10^{-4}T^{0.69}\) (kg m s\(^{-3}\) K\(^{-1}\)) (Banks and Kockarts 1973) is used for calculating the diffusive heat flux of \(\lambda \partial T / \partial z\). Each vertical column of \(\mathbf{A} \left( = \delta (\mathbf{R} - \mathbf{Q}_{\text{mol}}) / \delta \mathbf{T} \right)\) represents the change in vertical profile of cooling rate and diffusive heating rate difference (K day\(^{-1}\)) due to a unit change in temperature \((\delta \mathbf{T})\) with respect to \(\mathbf{T}\) at altitude \(z\) (horizontal row).

In the middle atmosphere, the effect of line overlap is negligible for the infrared radiative cooling rate calculations. As a result, the total infrared cooling rate can be evaluated as the sum of the cooling rates due to CO\(_2\), O\(_3\), and H\(_2\)O (Zhu 1994). Therefore, the term \(\Delta \mathbf{R}\) in Eq. (4) or (5) becomes

\[
\Delta \mathbf{R} = \Delta \mathbf{R}^{\text{CO2}} + \Delta \mathbf{R}^{\text{O3}} + \Delta \mathbf{R}^{\text{H2O}} .
\]

In other words, the linear separation of the partial infrared radiative cooling rate due to individual gases in the middle atmosphere is satisfied to a very good approximation, which is not the case for the troposphere. Thus, the approximations made for the feedback analysis are much better justified for the middle atmosphere than for the troposphere and surface temperatures.

For radiative heating by solar flux, we still need to invoke a linear approximation to decompose the energy perturbation into individual components; namely,

\[
\Delta S \approx \frac{\partial S}{\partial r_{\text{O3}}} \Delta r_{\text{O3}} + \frac{\partial S}{\partial r_{\text{O2}}} \Delta r_{\text{O2}} + \frac{\partial S}{\partial F_{10.7}} \Delta F_{10.7} + \frac{\partial S}{\partial \omega_0} \Delta \omega_0 ,
\]

or

\[
\Delta S \approx \Delta S^{\text{O3}} + \Delta S^{\text{O2}} + \Delta S^{\text{F10.7}} + \Delta S^{\omega_0} ,\]

where \(r_{\text{O3}}\) and \(r_{\text{O2}}\) are the O\(_3\) and O\(_2\) mixing ratios, respectively. Likewise, we also let \(r_{\text{CO2}}\) and \(r_{\text{H2O}}\) denote the mixing ratios of CO\(_2\) and H\(_2\)O, respectively. The quantity \(F_{10.7}\) is the 10.7-cm solar radio flux \((10^{-22} \text{ W m}^{-2} \text{ Hz}^{-1})\), which is a parameter representing the solar flux variations, and \(\omega_0\) is the effective albedo of surface and the lower atmosphere. The effect of \(r_{\text{O2}}\) variation on the energy perturbation \((\Delta S^{\text{O2}})\) is only important in the lower thermosphere. Unlike at the surface where the mean vertical velocity vanishes, the vertical velocity associated with the mean meridional circulation in the atmosphere plays an important role in coupling the radiation with dynamics and photochemistry. Here, invoking a linear approximation, we may explicitly extract this special “dynamical response" term from the total nonradiative energy source (Holton 2004):

\[
\Delta \mathbf{Q} = \Theta \Delta \mathbf{w}^* + (\Delta \Theta) \mathbf{w}^* + \Delta \mathbf{Q}^{\text{non-w}} ,
\]

where the diagonal matrix \(\Theta\) is the static stability parameter and column vector \(\Delta \mathbf{w}^*\) is the variation in the mean vertical velocity that yields the change in adiabatic cooling. The last term in Eq. (8) represents the contributions due to eddies and the energy transport by horizontal advection between neighboring vertical columns. Note that although the nonradiative energy source \((\Delta \mathbf{Q})\) can be evaluated from the dynamical modules during model integrations as reported in Lu and Cai (2010) and Song et al. (2014), it cannot be obtained directly from observations. It can also be evaluated as an energy residual term to balance the net radiative cooling rate and molecular thermal conduction according to Eq. (3): \(\Delta \mathbf{Q} = \Delta (\mathbf{R} - \mathbf{S} - \mathbf{Q}_{\text{mol}})\), which is an exact relation and can be evaluated accurately based on observations in the middle atmosphere. Such an approach of using better-defined thermal forcing that can be evaluated directly from observations to diagnose mechanical forcing was also proposed in Zhu et al. (2001) to diagnose the dynamical fields in the middle atmosphere. As reported in Lu and Cai (2010) and Song et al. (2014), \(\Delta \mathbf{Q}\) inferred explicitly from dynamical fields is almost identical to that inferred as an energy residual term. Given \(\Delta \mathbf{Q}\), we then use Eq. (8) to obtain \(\Delta \mathbf{Q}^{\text{non-w}}\) from the difference between \(\Delta \mathbf{Q}\) and the other two terms, which can be calculated from the available \(\mathbf{T}\) and \(\mathbf{w}^*\) profiles.

Substituting Eqs. (6)–(8) into Eq. (5), we obtain
\[
\Delta T = A^{-1} [-\Delta R^{\text{CO}2} + (\Delta S^{\text{O}3} - \Delta R^{\text{O}3}) + \Delta S^{\text{F}10.7} - \Delta R^{\text{H}2\text{O}} + \Delta S^{\text{O}2} + \Delta S^{\text{adj}} + \Theta \Delta w^* + (\Delta \Theta)w^* + \Delta Q^{\text{non-w}}] - \Delta T^{\text{err}},
\]

where \(\Delta T^{\text{err}}\) is the error due to linearization. Note that the change of \(r_{\text{O}3}\) in the middle atmosphere contributes to both the solar flux radiative heating and infrared cooling rate variations. This is similar to \(H_2O\) and clouds in the troposphere that can both radiatively heat and cool the atmosphere.

As shown in LC09, Eq. (9) is an additive formulation for the thermal response—that is, the sum of the partial component of the residual circulation, which leads to a zonal means in the meridional plane. Under such circumstances, we may choose all fields to be two-dimensional (2D) velocity. On the other hand, if Eq. (2) is zonally averaged, the well-known cancellation of the two terms on the right-hand side of Eq. (12) is expected to be dominant over the second term (mean meridional advection), and the third term makes the smallest contribution among the first three terms because of the quasigeostrophic nature of planetary-scale Rossby waves (Dunkerton 1978; Andrews and McIntyre 1976). Our MCFRAM additive relation (3) is similar to \(H_2O\) and clouds to both the solar flux radiative heating and infrared cooling rate variations. This is similar to \(H_2O\) and clouds in the troposphere that can both radiatively heat and cool the atmosphere.

The well-known cancellation of the two terms on the right-hand side of Eq. (13a) due to planetary-scale Rossby waves is the principal motivation for introducing the concept of the residual circulation. The definition of \(v^*\) by Eq. (13b) is based on the requirement that both \((\bar{v}, \bar{w})\) and \((v^*, \bar{w}^*)\) satisfy the zonal-mean continuity equation. In addition, the first term on the right-hand side of Eq. (12) is expected to be dominant over the second term (mean meridional advection), and the third term makes the smallest contribution among the first three terms because of the quasigeostrophic nature of planetary-scale Rossby waves (Dunkerton 1978; Andrews and McIntyre 1976).
that corresponds to the complete difference of the first term in Eq. (12). The third term in Eq. (8) can be decomposed into three components corresponding to the last three terms in Eq. (12) and we can rewrite it into

$$\Delta Q_{\text{non-rad}} = \Delta Q_{\text{eddy-p}} + \Delta Q_{\text{eddy-s}} + \Delta Q_{\text{dyn}}.$$  \hspace{1cm} (14)

The corresponding $\Delta T^{(o)}$ are given by

$$\Delta T^{(o)} = \Delta T^{\text{eddy-p}} + \Delta T^{\text{eddy-s}}.$$  \hspace{1cm} (15)

We can define the energy perturbation due to all the grid-resolved dynamics as

$$\Delta Q^{\text{dyn}} = \Delta Q_{\text{eddy-p}} + \Delta Q_{\text{eddy-s}}.$$  \hspace{1cm} (16)

Furthermore, when the energy perturbation by the subgrid-scale eddies is also added, the nonradiative energy perturbation can be calculated by the residual method

$$\Delta Q = \Delta Q^{\text{dyn}} + \Delta Q^{\text{eddy-s}} = \Delta R - \Delta Q_{\text{mol}} - \Delta S.$$  \hspace{1cm} (17)

The physical meanings of all $\Delta T^{(o)}$ are given in Table 1, where we have assumed that $w^* = w^s$, the vertical component of the residual circulation. The sum of the first six components forms the partial temperature change due to radiative processes $\Delta T^{\text{rad}} = \Delta T^{(1-6)}$. The nonradiative partial temperature change $\Delta T^{\text{non-rad}}$ includes changes due to both the grid-resolved ($\Delta T^{\text{dyne}}$) and subgrid ($\Delta T^{\text{eddy-s}}$) atmospheric motions. It should be noted that $\Delta T^{\text{rad}}$ has been derived from the changes in net radiative heating rate excluding the cooling rate change due to the temperature variation $\Delta T$ itself— that is, the terms $(\Delta S - \Delta R)$ in Eq. (5). At the same time, the effect of $\Delta T$, if any, is allowed and has to be included in evaluating $\Delta Q$ and $\Delta S$, leading to a well-behaved additive relation [Eq. (10)]. The components of $\Delta Q$ and $\Delta S$ induced by $\Delta T$, if of small magnitudes, could also be moved to the left-hand side of Eq. (3) to combine with the $\partial (R - Q_{\text{mol}})/\partial T$ term to form a revised generalized damping matrix. One important reason that $\Delta T$ is generically isolated out in Eq. (4) for $\Delta (R - Q_{\text{mol}})$ and has been moved to the left-hand side in Eq. (5) is that the generalized damping matrix $A$ introduced in Eq. (5) is well behaved and always invertible in the middle atmosphere owing to its dominant diagonal elements (e.g., Spiegel 1957; Fels 1982; Goody and Yung 1989; Zhu and Strobel 1991). Physically, the invertibility of $A$ also means that the radiative cooling and thermal conductivity can serve as restoring agents to stabilize the system around its equilibrium positions.

The additive relation (10) for the temperature changes is an alternative expression of the energy equation [Eq. (3)] that is also additive. A linear transformation that isolates out the temperature variation from the energy difference on the left-hand side of Eq. (3) leads to the partial temperature differences as shown in Eq. (11) and allows us to derive this alternative relationship. The principal advantage of the additive relation (10) for temperature over the additive relation (3) for energy is that $\Delta T$ on the left-hand side of Eq. (10) is a directly observable and commonly used quantity, which can serve as a natural and standard scale for comparison. It should be noted that since MCFRAM only provides a diagnostic relation between two equilibrium states it does not contain the information of how the climate system has evolved. In other words, MCFRAM cannot unravel indirect effects that interact nonlinearly as the system is evolved from state 1 to state 2.
b. Eigenmodes of the generalized damping matrix and illustration of MCFRAM

In Eq. (9), there is a common factor, matrix $\mathbf{A}^{-1}$, that multiplies all the radiative and nonradiative energy perturbation terms. As a result, both the magnitude and vertical structure of the climate feedbacks are significantly influenced by the generalized damping matrix $\mathbf{A}$ defined in Eq. (5). Its inverse ($\mathbf{A}^{-1}$) will be referred to as the generalized relaxation matrix. For a given vertical profile of $\Delta \mathbf{F}^{(\theta)}$ the spatial structure of $\Delta \mathbf{T}^{(\theta)}$ is completely determined and can be understood by the eigenvectors ($\xi_i$) and eigenvalues ($\lambda_i$) of $\mathbf{A}$ or $\mathbf{A}^{-1}$:

$$\mathbf{A} \xi_i = \lambda_i \xi_i \quad \text{or} \quad (\mathbf{A}^{-1}) \xi_i = \lambda_i^{-1} \xi_i, \quad i = 1, 2, \ldots, N, \quad (18)$$

where $N$ is the total number of vertical layers. Equation (18) indicates that the eigenvalues of $\mathbf{A}^{-1}$ are the inverse of the eigenvalues of $\mathbf{A}$ corresponding to the same eigenvectors. Here, $\lambda_i$ and $\lambda_i^{-1}$ can be called the generalized damping rate and relaxation time corresponding to the perturbation eigenvector $\xi_i$, respectively. In the absence of molecular viscosity ($\mathbf{Q}_{\text{mol}} = 0$) the generalized damping matrix is given by $\mathbf{A} = \partial \mathbf{R}/\partial \mathbf{T}$. Its eigenvalue $\lambda_i$ is the radiative damping rate of a temperature perturbation (e.g., Goody and Yung 1989; Zhu and Strobel 1991). The effect of the vertical structure of the temperature perturbation characterized by its eigenvector $\xi_i$ on the radiative damping rate has been well documented (Zhu and Strobel 1991; Zhu 1993). The occurrence of the radiative damping rate in MCFRAM is a natural consequence of the fact that the basic MCFRAM equation [Eq. (9) or (10)] is an energy perturbation equation. When the energy perturbation refers specifically to the cooling rate change associated with a temperature perturbation that has been singled out among all the other changes, it is the radiative damping rate that establishes the connection between the energy and temperature perturbations. The magnitude of $\lambda_i$ under nonvanishing $\mathbf{Q}_{\text{mol}}$ conditions increases with the magnitudes of the cooling rate and molecular viscosity. It also increases with the decreasing characteristic vertical scale of the energy perturbation—that is, the vertical scale of the cooling rate variation or the temperature variation.

Infrared radiative heat exchange by CO$_2$ and O$_3$ makes a major contribution, whereas infrared cooling by H$_2$O makes a minor contribution to the radiative cooling rate in the middle atmosphere (Zhu 1994). Here, we use the $\mathbf{T}$ and $r_{O3}$ observed from the SABER instrument on board the TIMED satellite to derive $\mathbf{A}$ and to perform the eigenmode analysis to illustrate the general characteristics of the eigenvector of $\mathbf{A}$ in the middle atmosphere. The required global-mean $r_{H2O}$ profile for the radiation algorithm is derived from the 3D GEOSCCM (Pawson et al. 2008). In Fig. 1, we show the $\mathbf{T}$ and $r_{O3}$ profiles observed by TIMED/SABER averaged over a 54°S–54°N latitudinal range and the 12-yr period 2002–13. The SABER observations ranging from 20 to 110 km in the middle atmosphere are provided on a 0.7-km grid and are merged with the U.S. Standard Atmosphere 1976 (COESA 1976) in the troposphere. The radiative heating and cooling rate calculations based on the JHU/APL radiation algorithm are performed in the entire vertical domain of 157 layers, whereas the MCFRAM is applied to the top 129 layers ($N = 129$) that corresponds to the middle atmosphere, ranging from 10 to 110 km. The matrix $\mathbf{A}$ has dimensions of $129 \times 129$ with 129 eigenmodes. Any given vertical profile of $\Delta \mathbf{F}^{(\theta)}$ can be decomposed by a complete set of the eigenvectors, with each component decaying (i.e., relaxing to 0) at a rate proportional to the inverse of their corresponding eigenvalues. Figure 2 shows a set of nine selected vertical eigenmodes of the generalized damping matrix $\mathbf{A}$ calculated from $\mathbf{T}$ and $r_{O3}$ shown in Fig. 1 based on the JHU/APL radiation algorithm (Zhu 1994, 2004). The quantity $r_{CO2}$ in the calculation is set at a 2005 level of 380 ppmv and decreases with altitude above the upper mesosphere scaled according to the U.S. Standard Atmosphere 1976 (COESA 1976). The eigenmodes describe a quantitative relationship between $\Delta \mathbf{F}^{(\theta)}$ and the corresponding $\Delta \mathbf{T}^{(\theta)}$. The eigenvalues ($\lambda_i$) of the selected eigenvectors ($\xi_i$) range from a maximum value of $\lambda_{\text{max}} = 18.97$ day$^{-1}$ (blue line marked with circles in Fig. 2a) to a minimum value of $\lambda_{\text{min}} = 0.021$ day$^{-1}$ (black dashed line in Fig. 2d). The vertical eigenmode of the largest damping rate corresponding to the smallest relaxation time ($\lambda_{\text{max}}^{-1} = 0.053$ days) exhibits large amplitude near 107 km, which varies rapidly with altitude on a very small vertical scale of ~4 km. A heating perturbation $\Delta \mathbf{T}^{(\theta)}$ with fine vertical structure near 100–110 km will project strongly onto this eigenmode and will be effectively damped in a very short time and produce a very small temperature perturbation. On the other hand, the eigenmode with the smallest damping rate has a vertical structure centered near the tropopause that does not change sign. This mode has the largest relaxation time ($\lambda_{\text{min}}^{-1} = 46.9$ days) such that heating perturbations that project strongly onto it will yield the largest response in $\Delta \mathbf{T}^{(\theta)}$. The calculated magnitude for $\lambda_{\text{min}}^{-1}$ near the tropopause is also consistent with the relaxation time estimated from the seasonal variations of upwelling in the tropical lower stratosphere (Randall et al. 2002).

There are two distinct features shown in Fig. 2. First, there exists a strong scale dependence of the eigenvectors for the generalized damping matrix $\mathbf{A}$. Eigenvectors
corresponding to large-scale vertical perturbations have small eigenvalues. Second, the magnitude of the eigenvalue decreases as the location of the characteristic perturbation shifts from the upper middle atmosphere to the lower middle atmosphere. As a result, we note that when the magnitude of eigenvalue decreases as we move consecutively from Fig. 2a to Fig. 2d the vertical scale of the eigenvector increases and the location of its main perturbation shifts to the lower altitude. This is consistent with the general nature of scale dependence for the radiative damping of temperature perturbations in the middle atmosphere (e.g., Goody and Yung 1989; Fels 1982; Zhu 1993). In addition, the effect of the molecular diffusion included in A has the effect of damping very efficiently small-scale perturbations in the lower thermosphere. To show the general nature of the scale dependence of the damping rate and its departure from a one-to-one relationship for the eigenmodes in the middle atmosphere, we perform a Fourier transform of all 129 eigenvectors and calculate their power spectral densities (PSDs) (Zhu and Strobel 1991; Zhu 1993). Figure 3 shows a scatterplot of the generalized damping rate \( \lambda_i \) versus the wavenumber of the maximum peak in the PSD for all 129 eigenvectors. Also shown in the figure are the analytic expression for the parameterized radiative damping rate proposed in Zhu (1993) and a square fit (\( \lambda_i \sim m^2 \)) to the diffusive damping. We note that there is a set of eigenmodes that nicely fits \( \lambda_i \sim m^2 \) because the thermal conductivity is only locally important in the lower thermosphere where the wavenumbers of the corresponding eigenmodes are well defined. On the other hand, the vertical inhomogeneity of the radiative heat exchange over the domain produces the eigenmodes that do not have well-defined scales for given perturbations and the relationship is no longer single valued. For example, a wavy perturbation having a large vertical scale in the mesosphere and a small vertical scale in the stratosphere as shown in Fig. 2d could decay uniformly in the entire atmosphere with the same damping rate. As a result, there is a significant dispersion in Fig. 3 when only one wavenumber can be selected to represent the structure of an eigenmode. The relationship can be improved by reducing the size of the vertical domain, which increases the atmospheric
homogeneity of the problem. A better parameterization for radiative damping in practice is to introduce a scale-dependent radiative damping rate that varies with both altitude and wavenumber (Fels 1982; Zhu 1993). The drawback for this approach is that the parameterization is a redundant expression for radiative damping of a temperature perturbation.

The effect of the vertical structure of \( D_F(n) \) on \( D_T(n) \) through \( A \) can be seen from Fig. 4 where \( D_T^{CO2}, D_T^{O3} \), and \( D_T^{F10.7} \), as defined in Table 1, are calculated based on three \( D_F(n) \) components caused by changing three atmospheric parameters: (i) \( r_{CO2} \) is doubled from 380 to 760 ppmv, (ii) \( r_{O3} \) is uniformly reduced by 50%, and (iii) the solar index \( F10.7 \) is increased from 60 to 260.

The vertical structure of the temperature differences (Fig. 4b) is smoother than and significantly different from that of the heating rate variations (Fig. 4a). This is mainly due to the scale dependence of the generalized damping rate \( \lambda_i \) where smaller-scale \( D_F(n) \) are more effectively damped; that is, \( D_T(n) \) are smoother than \( D_F(n) \). Furthermore, the lower middle atmosphere is more sensitive in \( D_T(n) \) to a smaller \( D_F(n) \) due to smaller \( \lambda_i \) than the upper middle atmosphere.

In summary, the MCFRAM equations allow for the expansion of local temperature changes in terms of changes in forcing and dynamics. The availability of more input parameters on the climate change results in more complete information on the individual contributions to the observable total change in temperature to be extracted from the MCFRAM equations. To derive the generalized damping matrix \( A \), the MCFRAM generally requires a state-of-the-art radiative transfer code (the JHU/APL radiative transfer algorithm is used here) and will be limited by the accuracy of this radiative transfer code as well as limited by the climate model, analysis, or observation quality. The linearization assumptions...
3. Application of MCFRAM to TIMED/SABER observations

Application of MCFRAM is straightforward as shown in Eqs. (5) and (9) together with Table 1 once the input fields of various parameter variations such as CO₂, O₃, winds, and solar cycle forcing are available.

While climate models (such as GEOSCCM) can provide all the needed input fields distributed globally and uniformly, satellite observations often provide only some of the needed fields to derive the balanced additive relation (10). In general, the role of observations is to provide a verification to model’s credibility in simulating the actual physical processes. The original CFRAM has only been applied to model output fields because the major energy sources of sensible and latent heating rates in the troposphere can only be derived reliably from models. On the other hand, radiative heat exchange is the major energy source in the middle atmosphere, and it can be evaluated accurately based on the observed temperature and species distributions. In this section, we show MCFRAM analyzed results by using SABER-observed T and O₃ fields (Russell et al. 1999). We use the V1.07 SABER data available to the public from the TIMED mission data center (http://www.timed.jhuapl.edu).

Figures 5a and 5b show the zonal-mean T and O₃ fields in the middle atmosphere derived from SABER observations in the low and midlatitudes averaged over the 12-yr period of 2002–13. Shown in Figs. 5c and 5d are the T and O₃ differences between two time-mean states covering the periods of 2002–03 and 2008–09, respectively. Though the temperature change in the middle atmosphere exhibits a noticeable decrease from the 2002–03 period near solar maximum to the 2008–09 period near solar minimum over most regions, there also exist regions where temperature increases between the same two periods when the solar energy input decreases. We note that the observed temperature difference represents the sum of effects contributed by...
various processes including the solar flux changes due to solar cycle and man-made variations in \( r_{\text{CO2}} \) and other chemical species.

We now apply MCFRAM to the SABER-observed \( T \) and \( O_3 \) difference between the two periods: 2002–03 and 2008–09. The corresponding mean \( r_{\text{CO2}} \) and \( F_{10.7} \) used in MCFRAM analysis for these two periods are \( r_{\text{CO2}} \approx 374.7 \) ppmv, \( F_{10.7} \approx 167.1 \) and \( r_{\text{CO2}} \approx 386.3 \) ppmv, \( F_{10.7} \approx 68.1 \), respectively. There are six yaw cycles in each year, with each yaw cycle covering about 60 days. With both ascending and descending measurements, it is possible to have complete local time coverage in one yaw cycle, such that averaging over even one yaw cycle should remove any undersampled diurnal tidal signals. The corresponding local time and latitudinal coverage in two yaw cycles separated by 6 yr are nearly identical. Care has also been taken to remove quasi-biennial signals. The temperature difference shown in Fig. 5c represents the observed temperature difference \( \Delta T \). The MCFRAM analysis is performed separately on each of the corresponding 60-day periods, with the seasonal parameters such as the solar declination angle and \( F_{10.7} \) varying with separate yaw cycles. \( \Delta T^{(n)} \) as defined in Eq. (11) or Table 1 is the mean \( \Delta T^{(n)} \) averaged over all the 60-day periods in the corresponding yaw cycles separated by 6 yr. Given the observed \( T \), \( O_3 \), and \( F_{10.7} \) variations and using the JHU/APL middle-atmosphere radiation algorithm, the first three components of \( \Delta T^{(n)} \) shown in Table 1 (i.e., \( \Delta T^{\text{CO2}}, \Delta T^{O3} \) and \( \Delta T^{F10.7} \)) can be explicitly evaluated. Since other radiatively active species including \( H_2O \) and solar flux heating by \( \text{CO}_2 \) near 4.3 and 2.7 \( \mu m \) only make minor contributions to the radiative cooling and heating rate in the middle atmosphere (López-Puertas and Taylor 2001), we expect the sum of the above three terms to be approximately the partial temperature change due to radiative transfer \( \Delta T^{\text{rad}} \) as described in Table 1. As mentioned before, we use the residual of Eq. (3) to estimate \( \Delta Q \) to calculate \( \Delta T^{\text{non-rad}} \). In Fig. 6, we show the latitude–altitude distributions of \( \Delta T^{\text{CO2}}, \Delta T^{O3}, \Delta T^{F10.7}, \) and \( \Delta T^{\text{non-rad}} \). Also shown in the figure is the MCFRAM error due to linearization \( \Delta T^m \).
We note that the middle-atmosphere cooling rate by the CO₂ 15-μm band is mainly contributed from its cool-to-space component with its escape probability slowly varying with altitude in the middle atmosphere (Zhu et al. 1992). A uniform change in \( r_{\text{CO}_2} \) leads to a near uniform change in the escape probability in the middle atmosphere. Hence, the maximum \( \Delta T_{\text{CO}_2} \) due to a uniform increase in \( r_{\text{CO}_2} \) in the middle atmosphere occurs at the equatorial stratopause (Fig. 6a), where the peak temperature as shown in Fig. 5a produces the largest cooling rate variation. On the other hand, the response \( \Delta T_{\text{O}_3} \) represents a combined effect of both the solar radiative heating and 9.6-μm-band infrared cooling. Since there are both positive and negative ozone variations between the periods 2002–03 and 2008–09 (Fig. 5d), the induced response \( \Delta T_{\text{O}_3} \) also shows a non-uniform spatial pattern (Fig. 6b). The peak variation of \( \Delta T_{\text{O}_3} \) in the upper mesosphere is mainly due to the change in localized absorption of solar ultraviolet (UV) flux heating, whereas the peak variations in the stratosphere are mainly due to the enhanced \( \text{O}_3 \) 9.6-μm-band, cool-to-space variations in a more transparent atmosphere. Here, we note that the middle-atmosphere climate responses to the cooling rate changes induced by CO₂ and O₃ variations are different. \( \Delta T_{\text{CO}_2} \) (Fig. 6a) mostly follows \( r_{\text{CO}_2} \) owing to the strong dependence of outgoing infrared radiation on the Planck blackbody emission, whereas \( \Delta T_{\text{O}_3} \) (Fig. 6b) mostly follows \( r_{\text{O}_3} \) as a result of a stronger dependence of radiative emission on more rapidly varying escape probability (Zhu et al. 1991). The \( \Delta T_{F_{10.7}} \) shown in Fig. 6c exhibits a pattern of overall monotonic increase in magnitude with altitude mainly because of the fact that solar UV fluxes vary most strongly at shorter wavelengths, which are generally absorbed at higher altitudes.

The overall spatial pattern and magnitude of \( \Delta T_{\text{non-rad}} \) shown in Fig. 6d are similar to \( \Delta T \) shown in Fig. 5c, indicating the importance of dynamical drive of the zonal-mean middle-atmospheric thermal structure. One striking feature in Fig. 6d is that \( \Delta T_{\text{non-rad}} \) is significantly greater and has richer spatial structure than any individual \( \Delta T^{(n)} \) owing to radiative processes. In other words, the major part of the temperature changes in the middle atmosphere is associated with dynamic processes over this period. We note that the three regions of significant peak in \( \Delta T_{\text{non-rad}} \) are located at equatorial mid-stratosphere and high-latitude upper mesosphere, respectively, where the effects of wave drag associated with the wave-driven equatorial quasi-biennial oscillation (QBO) and mesosphere gravity wave breaking are expected to be important. Diagnostically, the corresponding changes in thermal radiation as a whole balance the nonradiative energy source. One plausible
explanation is that the middle-atmosphere wave drag is strong and spatially and temporally inhomogeneous owing to the randomness of various wave generation and dissipation mechanisms. Furthermore, from a global perspective, the adiabatic heating and mechanical forcing are balanced in a zonally averaged meridional plane under quasi-equilibrium conditions (Fels 1987; Zhu et al. 2001). For example, in the lower stratosphere, because the tropopause is much higher (\(\sim 17\) km) in the tropics than in the extratropics (\(\sim 10\) km), an induced thermal cooling in the high-latitude lower stratosphere associated with \(\Delta T_{CO_2}\) coupled with a midtropospheric warming in the tropics would enhance a meridional temperature gradient. Such a change in thermal forcing is accompanied by an enhancement in the vertical gradient of the wave drag, which drives the strengthening of the Brewer–Dobson circulation in the lower stratosphere (Butchart et al. 2006; Garcia and Randel 2008; Shepherd and McLandress 2011). Finally, comparison between Figs. 5c and 6e suggests that the linearization from an additive energy equation [Eq.(3)] for heating–cooling rate difference to the additive MCFRAM equation [Eq.(10)] for temperature change leads to an error of \(\sim 5\%\) or less.

Though \(\Delta T_{\text{non-rad}}\) is comparable to \(\Delta T\) and significantly greater than the radiative components of \(\Delta T^{(n)}\) locally, the global average of \(\Delta T_{\text{non-rad}}\) in the middle atmosphere should be much smaller than its typical local values. This is mainly due to the fact that the globally averaged vertical velocity at a given pressure level must vanish (Olaguer et al. 1992), and the main role of propagating waves is to redistribute rather than generate momentum and heat (e.g., Zhu et al. 2008, 2010). It is only the eddy diffusion, molecular viscosity, and wave heating generated by gravity wave breaking that will be able to produce a globally averaged heating or cooling rate difference. In Fig. 7a, we plot the global mean \(\Delta T^{(n)}\) averaged over a 54°S–54°N latitudinal range as shown in Figs. 6a–c together with the sum of the three components, which gives a very good approximation of \(\Delta T_{\text{rad}}\) in the middle atmosphere. The figure shows that the magnitude of \(\Delta T_{\text{rad}}\) gradually increases from 0 K near 22 km to 1 K near 30 km. It remains about \(\sim 1\) K throughout the altitude range 30–70 km. Note that, above \(\sim 40\) km, the largest component of \(\Delta T_{\text{rad}}\) over one solar cycle is \(\Delta T_{\text{F10.7}}\). In Fig. 7b, the globally averaged \(\Delta T_{\text{rad}}\) is plotted together with the globally averaged \(\Delta T_{\text{non-rad}}\) and \(\Delta T_{\text{err}}\) components owing to the dominance of the molecular thermal conductivity that is not linear in the local temperature perturbations. Figure 7b confirms our conjecture that the globally averaged \(\Delta T_{\text{non-rad}}\) is a small difference between globally averaged \(\Delta T\) and \(\Delta T_{\text{rad}}\) in most of the middle atmosphere, although locally \(\Delta T_{\text{non-rad}}\) could be noticeably greater than either \(\Delta T\) or \(\Delta T_{\text{rad}}\). Physically, Fig. 7b also shows that the globally averaged temperature change in the middle atmosphere is radiatively driven below \(\sim 70\) km, where the vertical eddy transport due to wave breaking is expected to be small. This is also consistent with previous discoveries from observational
investigations that the seasonal temperature variations in the lower stratosphere show significant compensations between tropics and extratropics (Yulaeva et al. 1994) or between the Northern and Southern Hemispheres (Fueglistaler et al. 2011). An increase in $r_{\text{CO}_2}$ coupled with a decrease in solar radiation reduces the net radiative heating rate, which cools the atmosphere. Near and above the mesopause region, globally averaged $\Delta T^{\text{non-rad}}$ is no longer small but of the same order of magnitude as $\Delta T$ or $\Delta T^{\text{rad}}$. This is mainly because the gravity wave breaking in the upper mesosphere induces eddy diffusion that irreversibly transports and distributes tracers including the potential temperature associated with atmospheric energy.

It is worth pointing out that the results shown in Fig. 7 also verify both the SABER observations of $T$ and $O_3$ and the accuracy of the JHU/APL radiation algorithm for the middle atmosphere. One common way of verifying measurements and testing radiation algorithms is to evaluate the global radiative balance (Kiehl and Solomon 1986; Olaguer et al. 1992). A good radiation algorithm requires the globally averaged net radiative heating rate to be much smaller than the typical values of the localized net radiative heating rate. A more stringent requirement for a good algorithm is that the heating or cooling rate respond sensitively to variations in radiation parameters while still preserving the property of its globally averaged net radiative heating rate close to zero in range of altitude where the effect of eddy transport is negligible. In other words, the condition of a vanishing globally averaged net radiative heating is not artificially imposed but naturally derived. Note that $\Delta T^{\text{rad}}$ and $\Delta T^{\text{non-rad}}$ are closely related to the difference of the net radiative heating rate and the vertical velocity between two slightly different equilibrium states, respectively. The result shown in Fig. 7a suggests that the JHU/APL radiation algorithm is sensitive to variations in $\text{CO}_2$, $O_3$, and $F_{10.7}$ and yet the globally averaged $\Delta T^{\text{non-rad}}$ shown in Fig. 7b remains small, as expected for thermally driven global change, based on the premise that the SABER-observed $T$ and $O_3$ fields are accurate as well.

Using only SABER-observed $T$ and $O_3$ fields MCFRAM has the capability to identify the specific sources of middle-atmosphere climate change. The MCFRAM analysis shows that $\text{CO}_2$-induced temperature perturbations peak at the equatorial stratopause, while $O_3$-induced temperature perturbations have a more complex structure and are sensitive to the specific $O_3$ change pattern. In addition, the solar forcing response was shown to increase with increasing altitude. Large changes due to dynamics are also evident based on the MCFRAM equations.

4. Application of MCFRAM to GEOSCCM output fields

We now apply the MCFRAM to 2D zonal-mean fields derived from the GEOSCCM where several dynamical variables are also available from the model output. The 3D GEOSCCM uses the GEOS-5 atmospheric general circulation model (Rienecker et al. 2008) in its forecast model component, coupled with the stratospheric chemical solver developed as a part of the GSFC 3D chemical transport model (Douglass et al. 1996; Pawson et al. 2008). With respect to Rienecker et al. (2008) this version of GEOSCCM also includes a treatment of stratospheric aerosol (Aquila et al. 2012, 2013), a mechanism to generate the QBO using a gravity wave drag parameterization (MoIod et al. 2012), and variable solar irradiance forcing (Swartz et al. 2012).

In general, the usual output fields of GEOSCCM or any other CCMs are not specifically designed for directly performing a full MCFRAM analysis. Additional processing of some of the output fields is needed in order to produce a set of appropriate input fields for MCFRAM analysis. One potentially important input parameter as shown in Eq. (7) or (9) is the effective albedo of the surface and the lower atmosphere ($\omega_0$) that radiatively forces the middle atmosphere from the troposphere and surface. The $\omega_0$ is not saved in GEOSCCM as an output field in the simulations used but can be derived from the TOA net downward shortwave flux. In addition, because the upper boundary of the current GEOSCCM is below the mesopause, where the effect of $O_3$ variation is negligible in the energy budget, we will neglect $\Delta T^{O_2}$ in what follows.

Another issue in implementing MCFRAM analysis based on model output fields is that most CCMs such as the GEOSCCM only save separately the total solar heating and infrared cooling rates but not the individual components contributed by different absorbers and solar flux variations. Specifically, we also note that $\mathbf{T}$ needs to be used while evaluating three terms on the right-hand side of Eq. (6). Furthermore, Eq. (7) suggests a significant modification to the online radiation code in any CCM in order to derive and save heating rate contributions by different components mainly because of the nonlinear effect between solar flux and absorber. One way to get around these difficult issues is to calculate all the radiative heating and cooling perturbation terms offline and introduce two difference terms to the basic MCFRAM equations [Eqs. (10) and (11)] (Taylor et al. 2013; Sejas et al. 2014):

$$\Delta T = \sum_n \Delta T_n^{(n)} - \Delta T_{\text{diff1}} - \Delta T_{\text{diff2}} - \Delta T_{\text{err}}. \quad (19)$$
Fig. 8. (a) Changes in effective albedo of the surface and lower atmosphere \( (\Delta_a) \) scaled by the diurnally averaged solar radiation between 2002–03 and 2008–09 vs months. (b) Partial temperature change due to the changes in effective albedo \( (\Delta T_{ao}) \) by altitude–latitude. (c) Temperature change between two mean states \( (\Delta T) \) by altitude–latitude.

Here, the two partial temperature changes due to radiation differences are calculated based on GEOSCCM-saved total solar heating and infrared cooling rates together with the offline radiation algorithm:

\[
\Delta T_{\text{diff}} = (A^{-1}) \Delta S_{\text{diff}} \quad \text{and} \quad \Delta T_{\text{diff2}} = (A^{-1})(-\Delta R_{\text{diff}}),
\]

(20a)

(20b)

where \( \Delta S_{\text{diff}} \) and \( \Delta R_{\text{diff}} \) are respectively the changes in total radiative heating and cooling rates between states 1 and 2 derived from the offline and GEOSCCM online radiation algorithms

\[
\Delta S_{\text{off}} = \Delta S_{\text{ccm}} \quad \text{and} \quad \Delta R_{\text{off}} = \Delta R_{\text{ccm}}.
\]

(21a)

(21b)

It has been suggested that the difference terms are mostly due to the different averaging procedures between the offline and online calculations (Taylor et al. 2013; Sejas et al. 2014). Additional differences could also be contributed by the different radiation algorithms adopted for the offline and online radiative heating and cooling rate calculations. When \( \Delta Q \) is evaluated from the radiative forcing by the residual method, the correction terms \( \Delta S_{\text{diff}} \) and \( \Delta R_{\text{diff}} \) will also be included in \( \Delta T_{\text{non-rad}} \). The difference introduced by inferring \( \Delta Q \) from radiative forcing evaluated from the online radiation algorithm calculations can be estimated and analyzed by a comparison with that derived directly from CCM outputs saved during runtime (Sejas et al. 2014).

In this paper, we choose the same output periods of 2002–03 (near solar maximum) and 2008–09 (solar minimum) from one GEOSCCM simulation as those for SABER observations used in the last section to perform the MCFRAM analysis. In Fig. 8, we show the variation in effective albedo of the surface and lower atmosphere scaled by the diurnally averaged solar radiation as a function of month and latitude over the 24-month period. Also shown in the figure are the corresponding partial temperature change \( \Delta T_{ao} \) and the modeled temperature difference between the two mean states \( \Delta T \). The figure shows a typical variation of \( \sim 10 \text{ W m}^{-2} \) that is about 2% of the globally averaged solar flux and is about one order of magnitude greater than the variation in the solar constant over the 11-yr solar cycle (Lean 1991). Significant geographic and transient variations are present with peak values appearing near equatorial and summer polar areas, where the maximum mean solar fluxes are deposited. The corresponding \( \Delta T_{ao} \) as shown in Fig. 8b has a typical magnitude of 0.05 K peaking at low latitudes, which is more than one order of magnitude smaller than \( \Delta T \) (Fig. 8c). Here, we note that the model simulated \( \Delta T \) (Fig. 8c) is about a factor of 2 smaller than the one observed by SABER (Fig. 5c). Climate change or the system's feedback response is often associated with a radiative forcing scaled by the changes in the total radiation flux. Since the energy deposition in the atmosphere at different wavelengths varies drastically with spatial and temporal distributions of absorbers, the change in energy flux may not be a good indicator of how the system responds. On the other hand, the MCFRAM analysis based on Eqs. (10) and (11), together with Table 1, provides us with a complete view of the system response in the same variables and units scaled by an observed or modeled \( \Delta T \).

We now examine \( \Delta T^{(o)} \) associated with the dynamical processes as shown in Fig. 9. Figures 9a–c show three components of the grid-resolved partial temperature changes \( \Delta T_{\text{sw}}, \Delta T_{\text{ir}}, \text{ and } \Delta T_{\text{eddy-p}} \), respectively. Figures 9d–f show three other \( \Delta T^{(o)} \) for the purpose of comparison. Comparing \( \Delta T_{\text{sw}}^{(o)} \) with \( \Delta T_{\text{sw}}^* \) (Fig. 9d) suggests that the two are nearly identical, indicating an unimportant contribution by \( \Delta T^{(o)} \). We further note that \( \Delta T_{\text{sw}}^{(o)} \) is greater than either \( \Delta T_{\text{ir}} \) or \( \Delta T_{\text{eddy-p}} \), making a dominant contribution to \( \Delta T_{\text{dyn}}^{(o)} \) (Fig. 9e). This is consistent with the result from scale analysis that the first
term on the right-hand side of Eq. (12) is dominant (Dunkerton 1978). Another significant fact shown in Fig. 9 is that the magnitude of $\Delta T^{eddy-p}$ is the smallest among the three components of $\Delta T^{dyn}$. To further demonstrate this point, we show $\Delta T^w$ in Fig. 9f—that is, the partial temperature change due to the change in $\overline{w}$, the conventional Eulerian-mean vertical velocity. The magnitude of $\Delta T^w$ is about a factor of 2 greater than that for $\Delta T^{w,*}$. The significant difference between $\Delta T^w$ and $\Delta T^{w,*}$ with a smaller magnitude in $\Delta T^{w,*}$ confirms the rationale of introducing the TEM system based on residual circulation in the middle atmosphere. By using the TEM system to describe the advective transport of zonal-mean energy, the effect of the grid-resolved eddy heat flux by Rossby waves can be effectively included in the zonal-mean circulation. Dunkerton (1978) showed that the residual circulation defined by Eq. (13) is approximately the generalized Lagrangian-mean (GLM) meridional circulation introduced by Andrews and McIntyre (1978) that characterizes the motion of the center of mass of an ensemble of air parcels. Furthermore, eddy transport of chemical species and conservative dynamical quantities can all be parameterized by the same eddy diffusion coefficient under the TEM system (Zhu 1988).

In Fig. 10, we show the rest of the $\Delta T^{(n)}$ in Table 1 below 70 km that can be directly calculated based on GEOSCCM output fields and the offline JHU/APL radiation algorithm. Figures 10a–d show the four $\Delta T^{(n)}$ associated with the radiation processes in the middle atmosphere. Figures 10e and 10f are, respectively, the partial temperature changes due to radiative processes ($\Delta T^{rad}_{\text{online}}$) and dynamical processes ($\Delta T^{\text{non-rad}}$) including the online correction terms; that is,

\[
\Delta T^{rad}_{\text{online}} = \Delta T^{rad}_{\text{offline}} - \Delta T^{\text{diff1}} - \Delta T^{\text{diff2}} \quad \text{and} \quad (22)
\]

\[
\Delta T^{\text{non-rad}} = (A^{-1})(\Delta R - \Delta S - \Delta Q_{\text{mol}}) - \Delta T^{\text{diff1}} - \Delta T^{\text{diff2}}, \quad (23)
\]

with $\Delta T^{\text{diff1}}$ and $\Delta T^{\text{diff2}}$ shown in Figures 10g and 10h, respectively. Note that the first term in Eq. (22) is the sum of Figs. 10a–d plus Fig. 8b. Figure 10h shows the partial temperature change due to subgrid-scale eddies $\Delta T^{eddy-s}$:

\[
\Delta T^{eddy-s} = \Delta T^{\text{non-rad}} - \Delta T^{\text{dyn}}. \quad (24)
\]

We first note that the overall patterns and magnitudes of the partial temperature changes $\Delta T^{\text{CO2}}$ and $\Delta T^{\text{F10.7}}$ (Figs. 10a and 10c) that are primarily induced by the variations of the external forcing are nearly identical to those derived by SABER observations in the common domain (Figs. 6a and 6c). GEOSCCM shows a strengthening $\Delta T^{\text{CO2}}$ associated with the CO2 cooling in the high-latitude regions in the Southern Hemisphere mesosphere where the coldest temperature often occurs near the summer mesopause (Lübken 1999; Lübken et al. 1999). This is caused by heat exchange between the...
warmer stratopause and colder mesopause when the CO$_2$ 15-μm-band transmission behaves transparently and the summer mesopause receives net radiative heating from the stratopause (Zhu et al. 1992). An increase in $r_{CO2}$ increases the atmospheric opacity that leads to a reduction in summer mesopause net heating rate. Furthermore, there exists a local maximum in CO$_2$ 15-μm cooling rate near the winter polar mesopause as a result of the combination of local thermodynamic equilibrium conditions, a near-uniform temperature, and a near transparent emission to space (Zhu 1994). This too contributes to the strengthening in $\Delta T_{CO2}$ in the high-latitude and polar mesosphere regions. There exists a significant difference in $\Delta T_{O3}$ between GEOSCCM fields (Fig. 10b) and SABER observations (Fig. 6b) mainly owing to the difference in $r_{O3}$ between the two. This is not surprising because the middle-atmosphere O$_3$ and its variability are very sensitive to a strong nonlinear coupling between photochemistry and dynamics. For example, the largely off-set peaks in $\Delta T_{O3}$ in the equatorial lower stratosphere may well reflect the degree of fidelity of GEOSCCM simulation to the equatorial quasi-biennial oscillation phenomenon.

The partial temperature change $\Delta T_{H2O}$ as shown in Fig. 10d makes a very small contribution and is negative in the low latitudes but mostly positive in the middle latitudes. Middle-atmosphere H$_2$O may increase with time as a result of increasing CH$_4$ in the troposphere (Zhu et al. 1999). Its decadal change could also be well correlated with the equatorial sea surface temperature and the associated cooling of the cold point to limit the direct entry of H$_2$O into the stratosphere (Solomon et al. 2010). The existence of large regions of both positive and negative $\Delta T_{H2O}$ in the middle atmosphere is an indication that both processes play important roles in determining $r_{H2O}$ in the time period of 2002–09. Comparison between Figs. 8–9 and Figs. 10a–d gives us one example how the MCFRAM with its key additive relation provides a more direct and quantitative insight into the relative importance of different factors of

Fig. 10. As in Fig. 9, but calculated based on radiative forcing: (a) $\Delta T_{CO2}$, (b) $\Delta T_{O3}$, (c) $\Delta T_{F10}$, (d) $\Delta T_{H2O}$, (e) $\Delta T_{rad}$, (f) $\Delta T_{non-rad}$, (g) $\Delta T^{diff1}$, (h) $\Delta T^{diff2}$, and (i) $\Delta T^{eddy}$. 
climate forcing and feedback processes when they are constrained under the same scale with the same units. Figures 10e and 10f show the partial temperature changes due to radiative ($\Delta T_{\text{rad}}$) and dynamical ($\Delta T_{\text{non-rad}}$) processes, respectively, according to online radiation algorithms. We first note that the magnitudes of both changes are comparable to the observed $\Delta T$ (Fig. 8c) though significantly smaller than the individual contributions by dynamical processes (Fig. 9). This is consistent with the fact that thermal forcing and mechanical forcing are diagnostically balanced in the middle atmosphere (Zhu et al. 2001). Comparison between Figs. 6d and 10f shows certain similarities in spatial pattern between SABER observations and GEOSCCM output with peak values of $\Delta T_{\text{non-rad}}$ near the midlatitude lower mesosphere.

Figures 10g and 10h show the partial temperature changes, $\Delta T_{\text{diff1}}$ and $\Delta T_{\text{diff2}}$, due to differences in heating and cooling rates between the offline and online calculations, respectively. The figures show that the differences are small in most regions of the middle atmosphere except $\Delta T_{\text{diff1}}$ near the low-latitude upper boundary and $\Delta T_{\text{diff2}}$ near the polar region. We note that the heating rate difference associated with $\Delta T_{\text{diff1}}$ by solar radiation near model’s upper boundary is sensitive to the shielding effect of the solar flux by the absorber column above the upper boundary. Furthermore, the sensitivity decreases with increasing latitude as the slant path also increases. The high-latitude cooling rate difference associated with $\Delta T_{\text{diff2}}$ is likely sensitive to the nonlocal heat exchange when the vertical temperature gradient is large. Specifically, the heat exchange by the CO$_2$ 15-$\mu$m band becomes transparent above the stratosphere among different layers of atmosphere, whereas the O$_3$ 9.6-$\mu$m-band emission is largely transparent in the entire atmosphere (Zhu et al. 1991, 1992; Zhu 1994).

In Fig. 11, we show the global average of $\Delta T^{(e)}$ presented in Figs. 9 and 10. We first note that from Fig. 11a the globally averaged $\Delta T^{(e)}$ due to dynamical processes are negligibly small in comparison with the observed $\Delta T$. Furthermore, there exists a clearly major cancellation between $\Delta T_{\text{eddy-p}}$ and $\Delta T_{\text{w}}$ even in the globally averaged sense, again confirming the rationale of introducing the residual circulation in the middle atmosphere in terms of the partial temperature changes when planetary-scale wave activity is dominant. Several major features in Fig. 11b are consistent with those derived from the SABER observations as shown in Fig. 7: (i) $\Delta T_{\text{T10.7}}$ makes the largest contribution above $\sim$40 km, (ii) $\Delta T_{\text{CO2}}$ is negative at all altitudes whereas $\Delta T_{\text{O3}}$ is positive in some part of the altitude range, and (iii) $\Delta T$ and $\Delta T_{\text{online}}$. 
are nearly coincident. All the model fields in GEOSCCM, including the atmospheric temperature, have been integrated subject to the influence of a set of prescribed boundary conditions. Furthermore, there exists a shielding effect near the upper boundary that could lead to significant errors in heating rate calculations. We note the rapid increases with altitude of the magnitudes for several radiative components of $\Delta T(n)$ in Fig. 11b near the upper boundary. On the other hand, $\Delta T(n)$ derived from the SABER observations as shown in Fig. 7 do not systematically increase with altitude below 80 km, indicating the effect of boundary condition on the heating rate calculations for GEOSCCM fields.

Using the more complete model output fields, MCFRAM is able to make further breakdown in contributions by various climate forcing and feedback processes. The CO$_2$ and solar forcing response patterns are similar to those found in the SABER observations, whereas the patterns in the more sensitive O$_3$ response differ from the SABER results. Latitudinal changes in the H$_2$O perturbation response are also found. Additional partitioning of the dynamical term enabled us to reconfirm that the adiabatic heating induced by vertical motion makes the largest dynamical contribution.

5. Summary

In this study, we have extended the climate feedback–response analysis method (CFRAM) for the coupled troposphere–surface system to the middle atmosphere. The middle-atmosphere CFRAM (MCFRAM) is built upon the atmospheric energy equation per unit mass, with radiative heating and cooling rates as its major thermal energy sources. In addition, molecular thermal conduction is added to the energy equation when the upper boundary is extended beyond the mesopause. MCFRAM preserves the unique additive property of the original CFRAM in that, under linear approximation, the sum of all the partial temperature changes [$\Delta T(n)$] equals the total temperature change, which can be directly compared with the observed temperature change ($\Delta T$). Such an additive relation results from the generic additive relation (3) for energy differences such that all $\Delta T(n)$ are proportional to the energy perturbations [$\Delta F(n)$] associated with the variations of external forcing or feedback parameters. By introducing the generalized damping matrix ($A$) to the basic MCFRAM equation, the relationship between the partial temperature changes, $\Delta T(n)$, and their physical causes, $\Delta F(n)$, is quantitatively clarified by the well-documented theory of radiative damping of thermal disturbances in the middle atmosphere. Specifically, we show that $A$ serves as a filter that smooths the small-scale structure in $\Delta F(n)$. In addition, Figs. 2 and 4 show that, for a given $\Delta F(n)$, the maximum response in $\Delta T(n)$ occurs when $\Delta F(n)$ is peaked at the location where the cooling rate of the basic state reaches its minimum value.

The newly developed MCFRAM is applied to two sets of zonally averaged data. One is the middle-atmosphere zonal-mean fields of $T$ and O$_3$ derived from SABER observations in low and midlatitudes averaged over 60-day periods. The other comprises the zonal-mean fields of $T$ and O$_3$ plus several dynamical variables saved from GEOSCCM simulations. Results show that MCFRAM can complement both observations and climate model experiments. It is found that the spatial structures of the response, $\Delta T(n)$, to variations of CO$_2$, O$_3$, and solar flux are different. $\Delta T^{CO2}$ closely follows the temperature distribution in most of the middle atmosphere because the cool-to-space approximation is valid for an atmosphere with uniform $F_{CO2}$. On the other hand, both the solar radiative heating and 9.6-µm-band cooling by O$_3$ are strongly influenced by the O$_3$ distribution, and they contribute to $\Delta T^{O3}$ in about the same order of magnitude. $\Delta T^{H10.7}$ increases monotonically with altitude owing to the fact that the solar UV flux has greater variation at shorter wavelengths, which are generally absorbed at higher altitudes. The two periods used to derive the statistical equilibrium states are 2002–03 and 2008–09, corresponding to near solar maximum and solar minimum, respectively. The $r_{CO2}$ between these two periods increased from $\sim$374.7 to $\sim$386.3 ppmv. It is consistently found by both datasets that for a half-cycle span of the 11-yr solar cycle the largest radiative component of $\Delta T(n)$ in the middle atmosphere is the one due to the variation of the input solar flux $\Delta T^{H10.7}$. The effect of $F_{CO2}$ always cools the middle atmosphere with time ($\Delta T^{CO2} < 0$). On the other hand, depending on the relative importance of O$_3$ heating and cooling rates, $\Delta T^{O3}$ could be either positive or negative. The MCFRAM analysis of GEOSCCM fields suggests that $\Delta T^{H2O}$ and $\Delta T^{O3}$ (where the superscript is the albedo) make minor contributions to the observed $\Delta T$. Our GEOSCCM analysis of various components of $\Delta T^{H2O}$ demonstrates the major cancellation between the energy transport by the zonal mean meridional circulation ($\vec{v}, \vec{w}$) and the one by the meridional heat flux ($\nu \vec{T}$) associated with planetary-scale waves, reaffirming the rationale for introducing the residual circulation in terms of the partial temperature changes.

Because not all of the required parameters are present in the input datasets, especially those datasets derived directly from observations, the partial temperature change due to nonradiative processes ($\Delta T^{non-rad}$) often needs to be evaluated as a residual. Such an approach is well founded because of the exact additive relation for
the generic energy equation [Eq. (3)]. $\Delta T_{\text{non-rad}}$ for the SABER observations includes all dynamical effects, whereas many individual components in $\Delta T_{\text{non-rad}}$ can be evaluated separately based on the GEOSCCM model outputs. In both cases, $\Delta T_{\text{non-rad}}$ and $\Delta T_{\text{rad}}$ are comparable to the observed $\Delta T$, although the magnitudes of the individual components can be either much greater or smaller than $\Delta T$. This is consistent with the fact or an alternative interpretation based on MCFRAM that thermal forcing and mechanical forcing as a whole are approximately balanced in the middle atmosphere (Zhu et al. 2001). On the other hand, the globally averaged $\Delta T_{\text{non-rad}}$ is much smaller than either $\Delta T_{\text{rad}}$ or $\Delta T$ below $\sim 70$ km, indicating the lack of vertical transport of energy by eddies. Physically, this means that the globally averaged climate change in the middle atmosphere below $\sim 70$ km is thermally driven, which is consistent with the same conclusion in the lower stratosphere from observational studies (Yulaeva et al. 1994; Fueglistaler et al. 2011). This also means that the globally averaged partial temperature change due to all radiative processes ($\Delta T_{\text{rad}}$) is approximately equal to the observed temperature change ($\Delta T$). It ranges from $-0.5$ K near 25 km to $-1.0$ K near 70 km from the near solar maximum to the solar minimum.

Acknowledgments. This research was supported by NASA Living With a Star Program under Grant NNX13AF91G and NASA Geospace Science Program under Grant NNX13AE33G to The Johns Hopkins University Applied Physics Laboratory. Comments on the manuscripts by three anonymous reviewers and by Jae N. Lee are also appreciated.

REFERENCES


