

Retrospective-Cost Subsystem Identification for the Global Ionosphere-Thermosphere Model

Asad A. Ali^{*}, Kshitij Agarwal[†], Anthony M. D'Amato[‡],
Aaron J. Ridley[§] and Dennis S. Bernstein[¶]

University of Michigan, 1320 Beal Ave., Ann Arbor, MI 48109

We consider the problem of identifying unknown physical processes and estimating inaccessible parameters in the global ionosphere and thermosphere model (GITM). In previous work, we took advantage of developments in the retrospective cost method to estimate unknown thermal conductivity and rate coefficients in a one dimensional column in the atmosphere. In the present work, we extend the application of this method to the identification of unknown inaccessible parameters in 3D GITM, as well as the identification of the unknown subsystem that governs the radiative nitrous oxide cooling process in 1D GITM.

I. Introduction

Large-scale computational models are typically based on spatially discretized physics. These models may capture phenomena that occur over a wide range of spatial scales, and can be used for simulation, data assimilation, or control. The accuracy of these models depends, however, on the accuracy of the physics modeling as well as the accuracy of the initial state, parameters, boundary conditions, and inputs.

In recent work [1], we used retrospective cost subsystem identification (RCSI) to improve the fidelity of a given initial model by using available data to identify a subsystem whose inputs and outputs are not directly measurable. In this case, standard system identification methods are not effective. The application considered in [1] involves the dynamics of the ionosphere and thermosphere, otherwise known as space weather. This system is driven by the variable output of the Sun, and affects the drag of satellites and the propagation of radio waves. The Global Ionosphere-Thermosphere Model is the basis of these results.

GITM is a three-dimensional spherical (global Earth) code that solves the Navier-Stokes equations for the thermosphere. GITM is different from other models of the atmosphere [2–4] in that it solves the full vertical momentum equation instead of assuming that the atmosphere is in hydrostatic equilibrium, where the pressure gradient is balanced by gravity. While this assumption is valid for the majority of the atmosphere, in the auroral zone, where significant energy is dumped into the thermosphere on short time scales, vertical accelerations often occur. This heating causes strong vertical winds that can significantly lift the atmosphere [4].

The grid structure within GITM is fully parallel and covers the entire surface of the Earth by using a block-based two-dimensional domain decomposition in the horizontal coordinates [5]. The number of latitude

^{*}Graduate Student, Aerospace Engineering Department

[†]Graduate Student, Aerospace Engineering Department

[‡]NASA GSRP Fellow, Aerospace Engineering Department

[§]Professor, Department of Atmospheric, Oceanic, and Space Sciences

[¶]Professor, Aerospace Engineering Department

and longitude blocks can be specified at run time in order to modify the horizontal resolution. GITM has been run on up to 256 processors with a resolution as fine as 0.31° latitude by 2.5° longitude over the entire globe with 50 vertical levels, covering a vertical domain from 100 km to roughly 600 km [6]. This flexibility can be used to validate consistency by running model refinement at various levels of resolution.

An unknown subsystem in GITM can range from an unknown parameter to an unknown dynamical system. The challenging aspect of estimating these parameters and dynamics is the fact GITM is a large-scale, nonlinear computational fluid dynamics code, which cannot be represented analytically and thus cannot be linearized. RCSI, however, does not require an analytical representation of the dynamics of the system and thus is able to use GITM as the basis of subsystem identification.

The goal of this paper is to apply RCSI to GITM in a way that goes beyond earlier results. In particular, [1] demonstrated estimation of two parameters and estimation of nitrous oxide (NO) cooling dynamics at a single altitude for a one-dimensional version of GITM. In the present paper, we use recent developments in the retrospective cost method to estimate unknown parameters in global three-dimensional GITM as well as estimation of NO cooling dynamics at multiple altitudes in one-dimensional GITM. These extensions are based on the improved retrospective cost algorithm given in [8], which requires less modeling information than prior versions of this algorithm. This distinction is crucial for GITM, which is a highly nonlinear simulation code, whereas RCSI was originally developed in [5] for linear models.

The contents of the paper are as follows. In section 2, we briefly outline the RCSI method. In section 3, we give three numerical examples, one for parameter ID in 3D GITM, and two for subsystem ID in 1D GITM. Finally, in section 4, we give conclusions.

II. Review of Retrospective Cost System Identification for Subsystems with Inaccessible Outputs

We use the RCSI identification algorithm described in [5]. A brief outline of the method is as follows. Consider the MIMO discrete-time main system

$$x(k+1) = f[x(k)] + g_u[u(k)] + g_w[w(k)], \quad (1)$$

$$y(k) = h_y[x(k)], \quad (2)$$

$$y_0(k) = h_{y_0}[x(k), v(k)],$$

where $x(k) \in \mathbb{R}^n$, $y(k) \in \mathbb{R}^{l_y}$, $y_0(k) \in \mathbb{R}^{l_{y_0}}$, $u(k) \in \mathbb{R}^{l_u}$, $w(k) \in \mathbb{R}^{l_w}$, and $k \geq 0$. The main system (1)–(2) is interconnected with the unknown subsystem modeled by

$$u(k) = G_s(q)y(k), \quad (3)$$

where q is the forward shift operator. The system (1)–(3) represents the true system. We assume that the excitation signal $w(k)$ is known. $v(k)$ denotes measurement noise.

Next, we assume a model of the main system of the form

$$\hat{x}(k+1) = \hat{f}[\hat{x}(k)] + \hat{g}_u[\hat{u}(k)] + \hat{g}_w[w(k)], \quad (4)$$

$$\hat{y}(k) = \hat{h}_y[\hat{x}(k)],$$

$$\hat{y}_0(k) = \hat{h}_{y_0}[\hat{x}(k)], \quad (5)$$

where $\hat{x}(k) \in \mathbb{R}^{\hat{n}}$, $\hat{y}(k) \in \mathbb{R}^{\hat{l}_y}$, $\hat{y}_0(k) \in \mathbb{R}^{\hat{l}_{y_0}}$, $\hat{u}(k) \in \mathbb{R}^{\hat{l}_u}$. The model of the main system is interconnected with the subsystem model

$$\hat{u}(k) = \hat{G}_s(q)\hat{y}(k).$$

The goal is to estimate a subsystem model $\hat{G}_s(q)$ that minimizes a cost function based on the performance variable

$$z(k) \triangleq \hat{y}_0(k) - y_0(k) \in \mathbb{R}^{l_z}$$

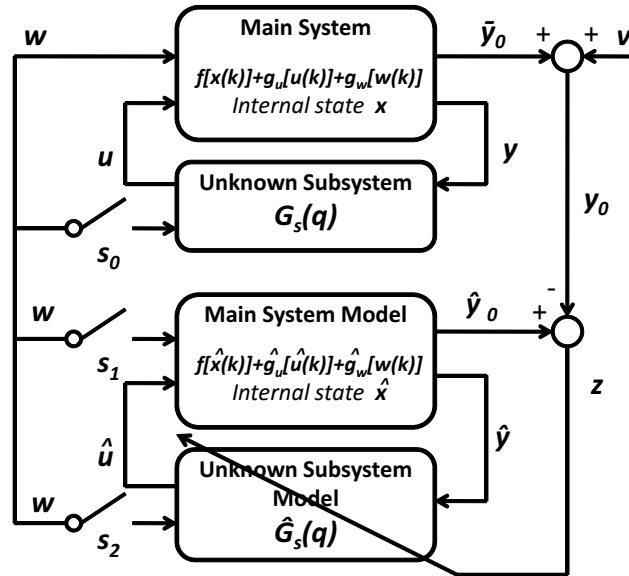


Figure 1. Model-refinement architectures. The switches s_0 , s_1 and s_2 are used to define various architectures. For GITM, only the switch s_1 is in the "on" position.

We estimate the subsystem by estimating the inaccessible subsystem output $\hat{u}(k)$ by minimizing the cost function

$$J(\tilde{U}^*(k-1), k) \triangleq \hat{Z}^T(k)R(k)\hat{Z}(k) + \eta(k)R_2\tilde{U}^{*T}(k-1)\tilde{U}^*(k-1), \quad (6)$$

where $\eta(k) \geq 0$, $R_2 \in \mathbb{R}^{\hat{U}} \geq 0$ and

$$\hat{Z}(k) = Z(k) - \tilde{\mathcal{H}}\tilde{U}(k-1) + \tilde{\mathcal{H}}\tilde{U}^*(k-1). \quad (7)$$

Furthermore,

$$\hat{Z}(k) \triangleq \begin{bmatrix} \hat{z}(k-k_1) \\ \vdots \\ \hat{z}(k-k_s) \end{bmatrix} \in \mathbb{R}^{sl_z}, \quad (8)$$

$$Z(k) \triangleq \begin{bmatrix} z(k-k_1) \\ \vdots \\ z(k-k_s) \end{bmatrix} \in \mathbb{R}^{sl_z}, \quad (9)$$

$$\tilde{U} \triangleq \begin{bmatrix} \hat{u}(k-k_1) \\ \vdots \\ \hat{u}(k-k_s) \end{bmatrix} \in \mathbb{R}^{sl_u}, \quad (10)$$

$$\tilde{U}^* \triangleq \begin{bmatrix} u^*(k-k_1) \\ \vdots \\ u^*(k-k_s) \end{bmatrix} \in \mathbb{R}^{sl_u}, \quad (11)$$

and $\tilde{\mathcal{H}}$ contains the appropriate Markov parameter estimates dependent on the choice of entries in \tilde{U}^* . Next, substituting (7) into (6) yields

$$J(\tilde{U}^*(k-1), k) = \tilde{U}^{*\text{T}}(k-1)\mathcal{A}(k)\tilde{U}^*(k-1) + \tilde{U}^{*\text{T}}(k-1)\mathcal{B}^{\text{T}}(k) + \mathcal{C}(k), \quad (12)$$

where

$$\mathcal{A}(k) \triangleq \tilde{\mathcal{H}}^{\text{T}}R(k)\tilde{\mathcal{H}} + \eta(k)R_2I_{l_U}, \quad (13)$$

$$\mathcal{B}(k) \triangleq 2\tilde{\mathcal{H}}^{\text{T}}R(k)[Z(k) - \tilde{\mathcal{H}}\tilde{U}(k-1)], \quad (14)$$

$$\begin{aligned} \mathcal{C}(k) \triangleq & Z^{\text{T}}(k)R(k)Z(k) - 2Z^{\text{T}}(k)R(k)\tilde{\mathcal{H}}\tilde{U}(k-1) \\ & + \tilde{U}^{\text{T}}(k-1)\tilde{\mathcal{H}}^{\text{T}}R(k)\tilde{\mathcal{H}}\tilde{U}(k-1). \end{aligned} \quad (15)$$

If either $\tilde{\mathcal{H}}$ has full column rank or $\eta(k) > 0$ and $R_2 > 0$, then $\mathcal{A}(k)$ is positive definite. In this case, $\bar{J}(\tilde{U}^*(k-1), k)$ has the unique global minimizer

$$\tilde{U}^*(k-1) = -\frac{1}{2}\mathcal{A}^{-1}(k)\mathcal{B}(k). \quad (16)$$

II.A. Subsystem Model Construction

The subsystem output $u(k)$ is given by the strictly proper time-series model of order n_c given by

$$\hat{u}(k) = \sum_{i=1}^{n_c} M_i(k)\hat{u}(k-i) + \sum_{i=1}^{n_c} N_i(k)y(k-i), \quad (17)$$

where, for all $i = 1, \dots, n_c$, $M_i(k) \in \mathbb{R}^{l_u \times l_u}$ and $N_i(k) \in \mathbb{R}^{l_u \times l_y}$. The subsystem output (17) can be expressed as

$$\hat{u}(k) = \theta(k)\phi(k-1), \quad (18)$$

where

$$\theta(k) \triangleq [M_1(k) \ \dots \ M_{n_c}(k) \ N_1(k) \ \dots \ N_{n_c}(k)] \in \mathbb{R}^{l_u \times n_c(l_u + l_y)} \quad (19)$$

and

$$\phi(k-1) \triangleq \begin{bmatrix} \hat{u}(k-1) \\ \vdots \\ \hat{u}(k-n_c) \\ y(k-1) \\ \vdots \\ y(k-n_c) \end{bmatrix} \in \mathbb{R}^{n_c(l_u + l_y)}. \quad (20)$$

II.B. Recursive Least Squares Update of $\theta(k)$

Let d be a positive integer such that $\tilde{U}(k-1)$ contains $u(k-d)$. Next, we define the cumulative cost function

$$J_{\text{R}}(\theta(k)) \triangleq \sum_{i=d+1}^k \lambda^{k-i} \|\phi^{\text{T}}(i-d-1)\theta^{\text{T}}(k) - u^{*\text{T}}(i-d)\|^2 + \lambda^k (\theta(k) - \theta(0))P^{-1}(0)(\theta(k) - \theta(0))^{\text{T}}, \quad (21)$$

where $\|\cdot\|$ is the Euclidean norm, and $\lambda(k) \in (0, 1]$ is the forgetting factor. Minimizing (21) yields

$$\begin{aligned} \theta^T(k) \triangleq & \theta^T(k-1) + \gamma(k)P(k-1)\phi(k-d-1)[\phi^T(k-d-1)P(k-1)\phi(k-d-1) + \lambda(k)]^{-1} \\ & \cdot [\phi^T(k-d-1)\theta^T(k-1) - u^{*T}(k-d)], \end{aligned} \quad (22)$$

where $\gamma(k)$ is either 0 or 1. When $\beta(k)$ is 1, the subsystem model is allowed to adapt, when $\gamma(k)$ is 0, the subsystem model adaption is off. The covariance matrix is updated by

$$\begin{aligned} P(k) \triangleq & (1 - \gamma(k))P(k-1) + \gamma(k)\lambda^{-1}(k)P(k-1) - \beta(k)\lambda^{-1}(k)P(k-1)\phi(k-d-1) \\ & \cdot [\phi^T(k-d-1)P(k-1)\phi(k-d-1) + \lambda(k)]^{-1}\phi^T(k-d-1)P(k-1). \end{aligned} \quad (23)$$

We initialize the covariance matrix as $P(0) = \beta I$, where $\beta > 0$.

III. Numerical results

In this section, we consider the problem of using upper atmosphere mass-density measurements to identify parameters in the thermal conductivity model in 3D GITM, as well as the problem of identifying the subsystem that governs the NO radiative cooling in 1D GITM. The NO cooling subsystem identification is further divided into two cases. For the first case, we simplify GITM by setting the NO cooling profile to be a piecewise linear function of altitude. For the second case, we do not make this simplification, and allow the physics in GITM to determine the profile of NO cooling. In all simulations, the time step is fixed at 3 seconds.

III.A. Parameter identification in 3D GITM

The thermal conductivity model in GITM is

$$\kappa(h, k) = (F_{O_2}(h, k) + F_{N_2}(h, k)) AT^{0.69}(h, k) + F_O(h, k)5.6 \times 10^{-4}T^{0.69}(h, k), \quad (24)$$

where A is the unknown thermal conductivity coefficient, h is the altitude, k is the time step, and F_{O_2} , F_{N_2} , F_O are fractional compositions of O_2 , N_2 , and O , respectively. The values of these fractional compositions are provided by GITM and are known. The true value of the coefficient is $A = 3.6 \times 10^{-4}$ and is identical for all altitudes, latitudes, and longitudes. We define the ‘‘truth model’’ as GITM with the parameter of the thermal conductivity model fixed to the true value. We use this truth model to simulate the chemistry and fluid dynamics in the global three-dimensional ionosphere-thermosphere of Earth and generate mass-density measurements at an altitude of 400 km above the Earth at a fixed latitude and longitude. This mass-density measurement is the ‘‘truth data’’, which we label as $y_0(k)$.

Next, we assume that the parameter A of the thermal conductivity model is unknown. We combine GITM with RCSI, and use RCSI to estimate A at each time step, and apply it to the thermal conductivity model in GITM at all altitudes, latitudes, and longitudes. The mass-density data at the altitude of 400 km generated from the combination of GITM and RCSI is labeled as $\hat{y}_0(k)$.

Note that since we are identifying a parameter directly, the output $\hat{u}(k)$ from the unknown subsystem model (shown in Figure 1) is the estimated value of the parameter. Furthermore, for parameter identification, we set $\hat{y}(k) = 1$ for all $k \geq 0$.

The setup of RCSI for this case uses $n_c = 1$, $\tilde{H} = -1000$, $\bar{\eta} = 0$, and $\beta = 10$. The initial values of the coefficients of the subsystem model are set to zero. Furthermore, the initial value of the parameter estimate is set to zero. Moreover, in RCSI, we set bounds on the value of the reconstructed signal $\tilde{U}^*(k)$, which is used to update the model of the subsystem. The bounds on $\tilde{U}^*(k)$ for this case are $0 \leq \tilde{U}^*(k) \leq 1 \times 10^{-2}$,

for all $k \geq 0$. These bounds are specified so that the nonlinear GITM model does not yield nonphysical states, such as negative temperature values, and become unstable. Finally, RCSI is turned on after 350 time steps. Figure 2 shows the estimate of the parameter A . The estimated value converges to the true value within 2×10^4 time steps. Figure 3 shows the output from the truth model $y_0(k)$ and the output from the combined GITM and RCSI model $\hat{y}_0(k)$, and Figure 4 shows the performance $z(k)$.

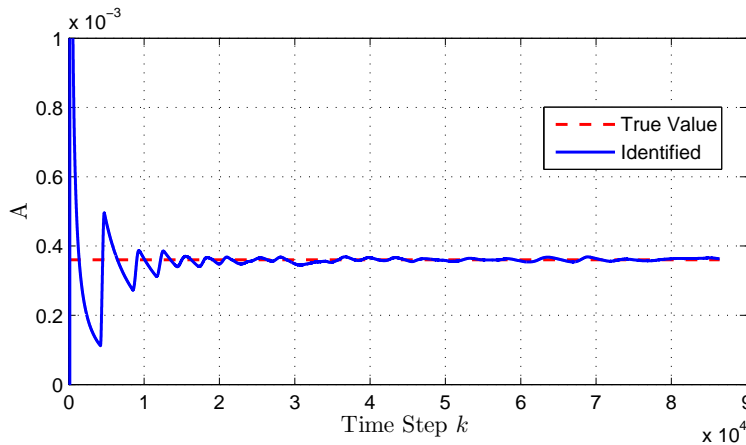


Figure 2. True and estimated coefficient A for Example III.A.

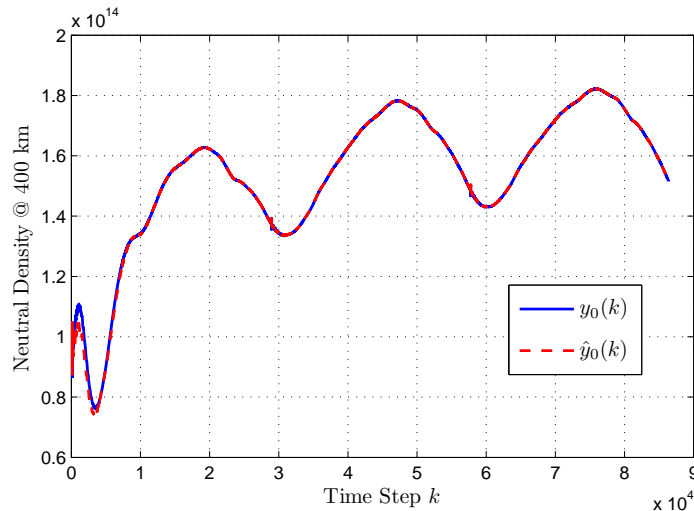


Figure 3. Comparison of the output $y_0(k)$ from the truth model and the output $\hat{y}_0(k)$ from the combined GITM and RCSI model for Example III.A.

III.B. NO Cooling subsystem identification for GITM with linear NO Cooling profile

In this section, we consider the identification of the subsystem that governs NO radiative cooling in GITM. For this example, we force NO cooling in GITM to be piecewise linear in altitude, zero at 0 km, and zero at 600 km (maximum height). First, we assume that the NO cooling process at each altitude is known, and simulate the chemistry and fluid dynamics in a one-dimensional column in the ionosphere-thermosphere and generate mass-density measurements at an altitude of 236 km and 388 km above the Earth. These mass-density measurements are the “truth data” and are labeled $y_0(k)$.

Next, we assume that the NO cooling processes are unknown, and combine GITM with RCSI to

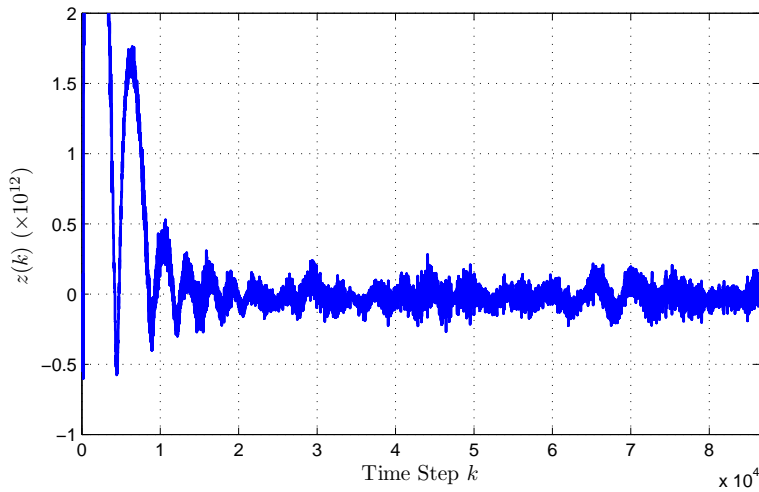


Figure 4. Performance $z(k)$ for Example III.A.

identify the NO cooling subsystems at altitudes of 236 km and 388 km at each time step k . Once we have reconstructed the NO cooling at 236 km and 388 km at k , we fit a piecewise linear function to this data such that NO cooling is zero at 0 km and 600 km. Using this function, we obtain the NO cooling at all other altitudes. The inputs to the NO cooling subsystems at 236 km and 388 km are the respective temperatures at these altitudes.

The setup of RCSI for this case uses $n_c = 8$, $\tilde{H} = -I_{2 \times 2}$, $\bar{\eta} = 1 \times 10^{-5}$, and $\beta = 0.001$. The initial values of the coefficients of the subsystem model are set to zero. Furthermore, the initial value of output of the subsystems is set to zero. The bounds on $\tilde{U}^*(k)$ for this case are $0 \leq \tilde{U}^*(k) \leq 8.0$, for all $k \geq 0$. Finally, RCSI is turned on after 400 time steps. Figure 5 shows the output from the truth model $y_0(k)$ and the output from the combined GITM and RCSI model $\hat{y}_0(k)$. Figure 6 shows the performance $z(k)$. Figure 7 compares the NO cooling at 236 km and 388 km to the output from the identified NO cooling process. Finally, Figure 8 compares the true and identified NO cooling at all altitudes for four time slices.

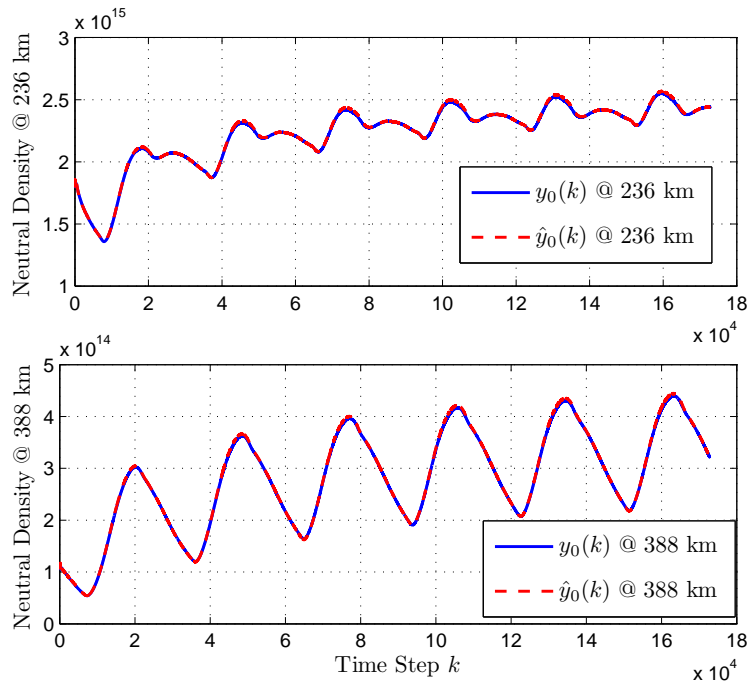


Figure 5. Comparison of the output $y_0(k)$ from the truth model and the output $\hat{y}_0(k)$ from the combined GITM and RCSI model for Example III.B.

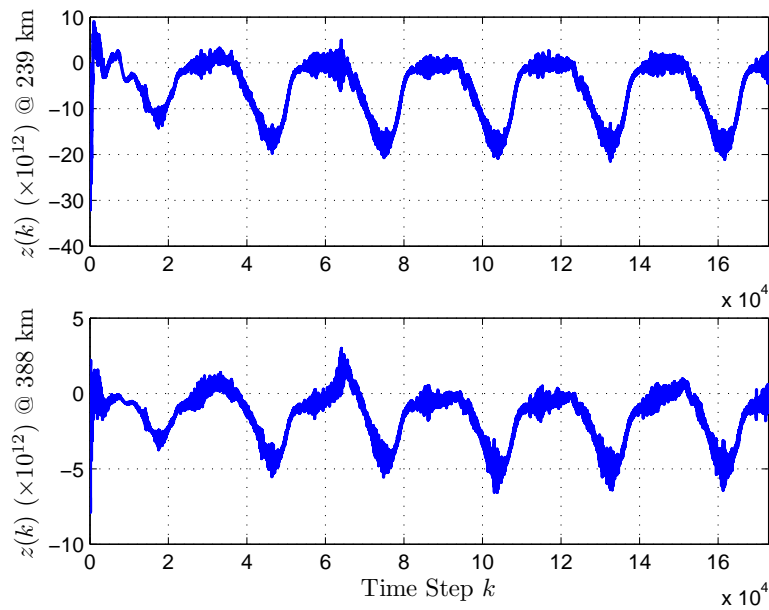


Figure 6. Performance $z(k)$ for Example III.B.

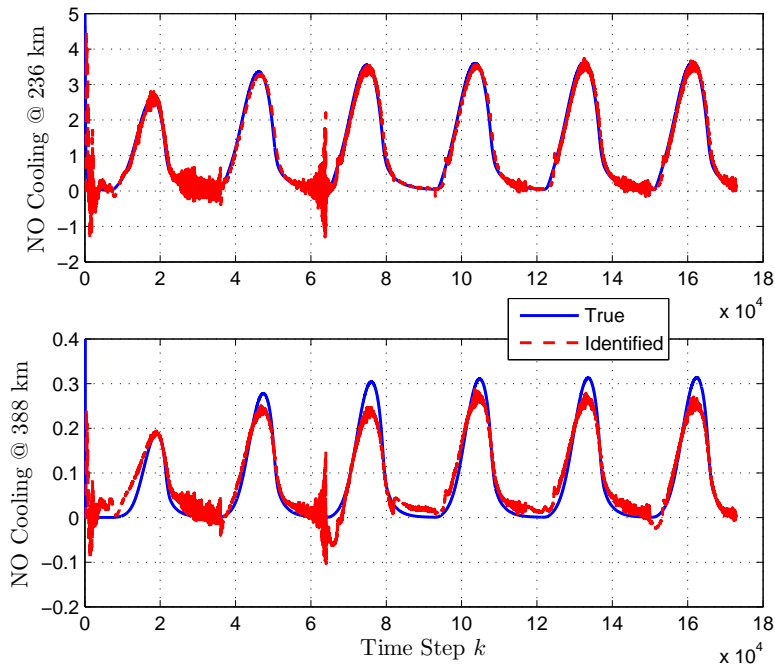


Figure 7. Comparison of the true NO cooling $u(k)$ and the output of the identified NO cooling process $\hat{u}(k)$ for Example III.B.

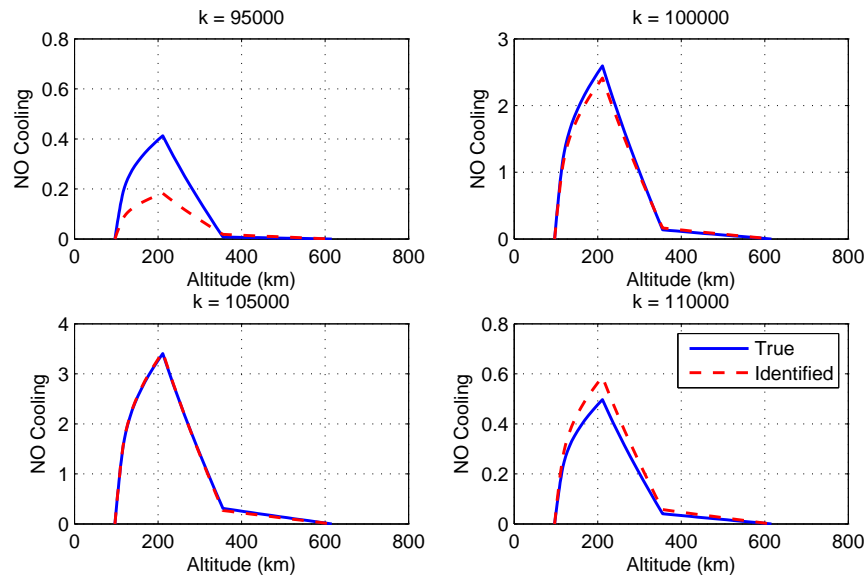


Figure 8. Comparison of the true and identified NO cooling at all altitudes for four time slices for Example III.B.

III.C. NO Cooling subsystem identification for GITM with true NO Cooling profile

In this section, we consider the identification of the subsystem that governs NO radiative cooling in GITM, but do not force the NO cooling profile in GITM to be piecewise linear in altitude, as was done in the second example. In this example, we allow the physics in GITM to determine the profile of NO cooling at

each time step. However, similar to the second example, when the NO cooling subsystems are unknown, we attempt to reconstruct the NO cooling at two altitudes and fit a piecewise linear profile to the NO cooling processes.

First, we assume that the NO cooling processes are known, and simulate the chemistry and fluid dynamics in a one-dimensional column in the ionosphere-thermosphere and generate mass-density measurements at an altitude of 138 km and 160 km above the Earth. These mass-density measurements are the “truth data”

Next, we assume that the NO cooling processes are unknown, and combine GITM with RCSI to identify the NO cooling subsystems at altitudes of 138 km and 160 km. Once we have reconstructed the NO cooling at 138 km and 160 km at time step k , we fit a piecewise linear function to this data such that NO cooling is zero at altitudes less than 110 km and at altitudes greater than 350 km. Using this function, we obtain the NO cooling at all other altitudes. The inputs to the NO cooling subsystems at 138 km and 160 km are the respective temperatures at these altitudes.

The setup of RCSI for this case uses $n_c = 6$, $\tilde{H} = -0.3I_{2 \times 2}$, $\bar{\eta} = 1 \times 10^{-5}$, and $\beta = 0.01$. The initial values of the coefficients of the subsystem model are set to zero. Furthermore, the initial output of the unknown subsystems is set to zero. The bounds on $\tilde{U}^*(k)$ for this case are $0 \leq \tilde{U}^*(k) \leq 8.0$, for all $k \geq 0$. Finally, RCSI is turned on after 400 time steps. Figure 9 shows the output from the truth model $y_0(k)$ and the output from the combined GITM and RCSI model $\hat{y}_0(k)$. Figure 10 shows the performance $z(k)$. Figure 11 compares the NO cooling at 138 km and 160 km to the output from the identified NO cooling process. Finally, Figure 12 compares the true and identified NO cooling at all altitudes for four time slices. Since the true NO cooling profile is nonlinear and we are using a linear approximation, the error $z(k)$ for this example is larger than in Example III.B. This also introduces a large amount of noise into the reconstructed subsystem outputs $\tilde{U}^*(k)$ as can be seen in Figure 11.

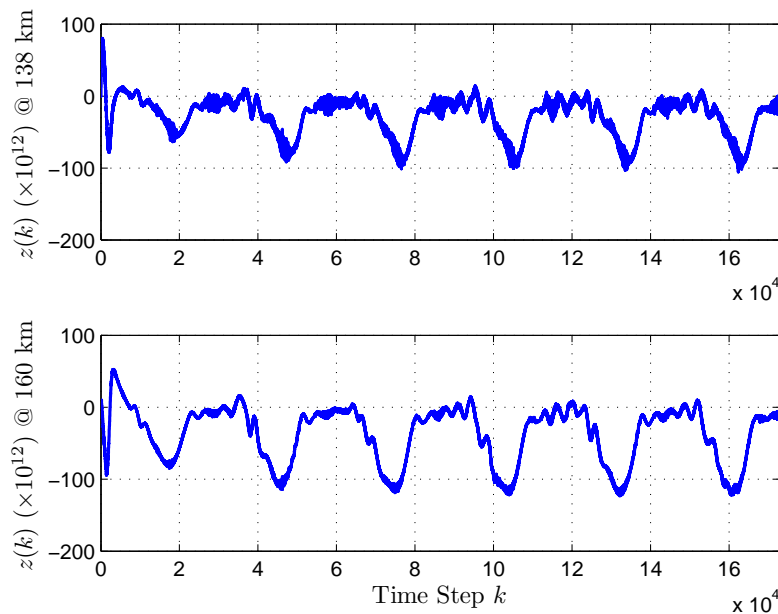


Figure 9. Comparison of the output $y_0(k)$ from the truth model and the output $\hat{y}_0(k)$ from the combined GITM and RCSI model for Example III.C.

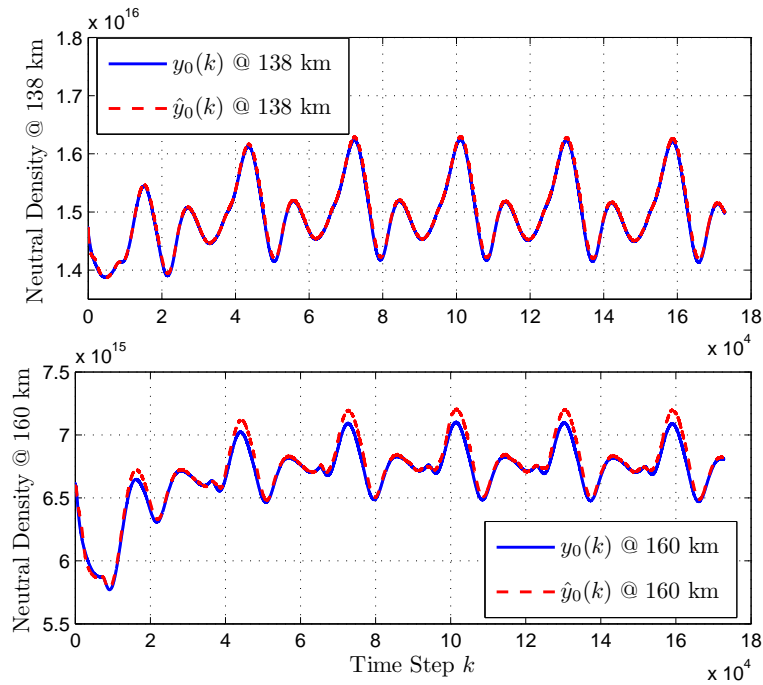


Figure 10. Performance $z(k)$ for Example III.C.

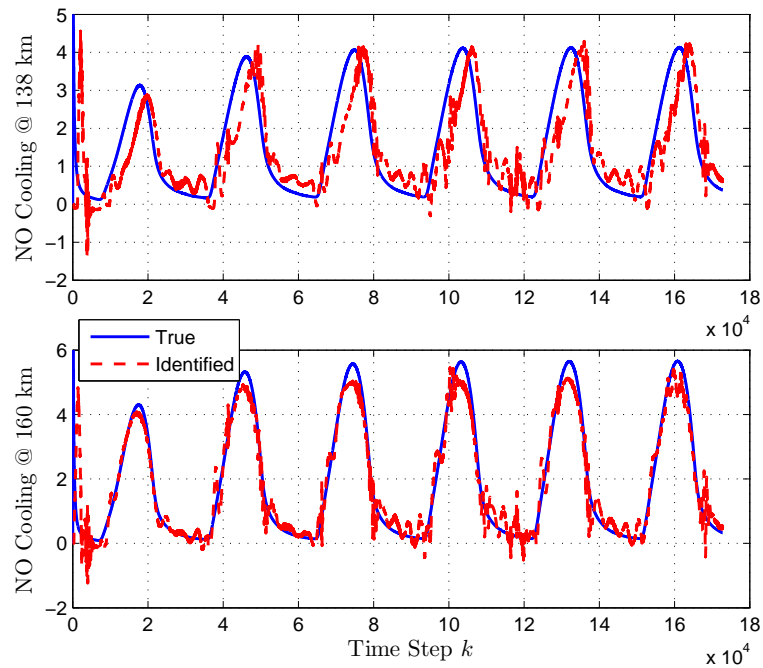


Figure 11. Comparison of the true NO cooling $u(k)$ and the output of the identified NO cooling process $\hat{u}(k)$ for Example III.C.

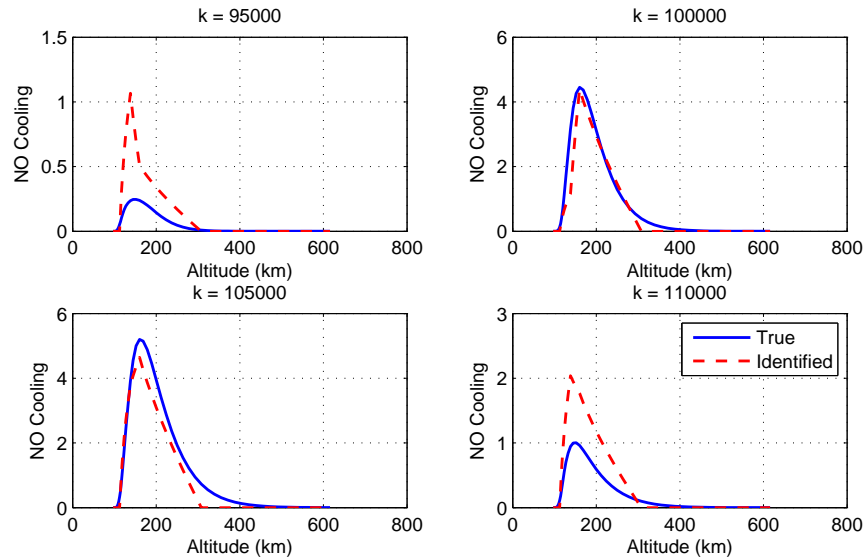


Figure 12. Comparison of the true and identified NO cooling at all altitudes for four time slices for Example III.C.

IV. Conclusions

In this paper, we used RCSI to estimate unknown parameters and identify inaccessible subsystems in GITM. As in previous work, we use one simulation of GITM as a truth model, and another copy of GITM with the unknown parameter or unknown process set to zero as the initial model.

We then apply RCSI to estimate the unknown parameter or identify the unknown subsystem using the residual between the truth model and the initial model. In the present work, we demonstrated this method by estimating the unknown thermal conductivity coefficient in 3D GITM, and identifying the NO cooling subsystem at two altitudes in 1D GITM. We note that RCSI does not require an analytical representation of GITM.

References

- ¹A. M. D'Amato, A. J. Ridley, and D. S. Bernstein, "Retrospective-cost-based adaptive model refinement for the ionosphere and thermosphere," *Statistical Analysis and Data Mining*, vol. 4, pp. 446-458, 2011.
- ²E. Yigit and A. J. Ridley, "Effects of high-latitude thermosphere heating at various scale sizes simulated by a nonhydrostatic global thermosphere-ionosphere model", *J. Atmos. Sol-Terr. Phys.*, Vol. 73, pp. 592-600, 2010.
- ³R. G. Roble, E. C. Ridley, A. D. Richmond, and R. E. Dickinson, "A coupled thermosphere/ionosphere general circulation model", *Geophys. Res. Lett.* Vol. 15, No. 12, pp. 1325-1328, 1988.
- ⁴T.J. Fuller-Rowell and D. Rees, "A three-dimensional, time-dependent, global model of the thermosphere", *J. Atmos. Sci.*, Vol. 37, pp. 2545-2567, 1980.
- ⁵Y. Deng, A. D. Richmond, A. J. Ridley, and H.-L. Liu, "Assessment of the non-hydrostatic effect on the upper atmosphere using a general circulation model (gcm)", *Geophys. Res. Lett.*, Vol. 35, L01104, 2008.
- ⁶A. J. Ridley, Y. Deng, and G. T'oth, "The global ionosphere/thermosphere model", *J. Atmos. Sol-Terr. Phys.*, Volume 68, Issue 8, pp. 839-864, 2006.
- ⁷A. M. Morozov, A. A. Ali, M. D'Amato, A. J. Ridley, S. L. Kukreja, and D. S. Bernstein, "Retrospective-cost-based model refinement for system emulation and subsystem identification," *Proc. Conf. Dec. Contr.*, pp. 2142-2147, Orlando, FL, December 2011.
- ⁸A. M. D'Amato, E. D. Sumer, and D. S. Bernstein, "Retrospective-cost adaptive control for systems with unknown nonminimum-phase zeros," *AIAA Guid. Nav. Contr. Conf.*, Portland, OR, August 2011, AIAA-2011-6203