

Retrospective-Cost-Based Adaptive Model Refinement for the Ionosphere and Thermosphere

Anthony M. D'Amato^{1*}, Aaron J. Ridley² and Dennis S. Bernstein¹

¹Department of Aerospace Engineering, University of Michigan, Ann Arbor, MI 48109, USA

²Department of Atmospheric Oceanic and Space Science, University of Michigan, Mi, 48109, USA

Received 16 December 2010; revised 7 May 2011; accepted 14 May 2011

DOI:10.1002/sam.10127

Published online 24 June 2011 in Wiley Online Library (wileyonlinelibrary.com).

Abstract: Mathematical models of physical phenomena are of critical importance in virtually all applications of science and technology. This paper addresses the problem of how to use data to improve the fidelity of a given model. We approach this problem using retrospective cost optimization, which uses data to recursively update an unknown subsystem interconnected to a known system. Applications of this technique are relevant to applications that depend on large-scale models based on first-principles physics, such as the global ionosphere–thermosphere model (GITM). Using GITM as the truth model, we demonstrate that measurements can be used to identify unknown physics. Specifically, we estimate static thermal conductivity parameters, as well as a dynamic cooling process. © 2011 Wiley Periodicals, Inc. *Statistical Analysis and Data Mining* 4: 446–458, 2011

Keywords: model refinement; adaptive; ionosphere; thermosphere; system identification; estimation

1. INTRODUCTION

Models serve a variety of purposes by capturing different phenomena at varying levels of resolution. High-resolution models are desirable when the goal is to understand scientific phenomena or assimilate data, whereas a coarser model may be preferable when the goal is to capture critical details in an efficient manner, for example, for fast prediction or control. Consequently, the fidelity of a model must be gauged against its intended usage.

Most models are constructed from collections of interconnected subsystem models, which in turn are based on a combination of physical laws and empirical observations. For example, the core of a model might be the Navier-Stokes or magnetohydrodynamic equations, while source terms, such as chemistry, heating, and friction, may be modeled using either first-principles submodels or empirical relations that have different levels of accuracy, complexity. Physical laws embody first-principles knowledge, whereas empirical observations may include relations that are based on the statistical analysis of data, for example, regression. Physics can provide the backbone of a model, while empirical relations can flesh out details such as sub-grid-scale phenomena that are beyond the ability of analytical modeling.

When input–output data are available, an empirical model can be constructed by means of system identification methods. In particular, techniques for constructing linear dynamic models that relate measured inputs to measured outputs are well developed [1–3]. A challenging extension is to develop methods for nonlinear system identification. Because nonlinear models can have a vast range of structures, the problem of nonlinear system identification requires the choice of a suitable model structure as well as an algorithm that uses data to tune the parameters of the model. Candidate model structures range from unstructured black-box models, such as neural networks, to gray-box and white-box models, where some or all of the structure of the model is specified [5–8]. The chosen model structure is assumed to be identifiable from the available measurements, which means that its independent parameters can be unambiguously estimated from sufficiently persistent data.

The ability to identify a component or subsystem of a system depends on *accessibility*, which refers to the availability of the inputs and outputs of the subsystem. The highest degree of accessibility arises when both the input and output of the unknown subsystem are measured. In the case of Hammerstein and Wiener gray-box model structures, a static nonlinear mapping is cascaded with a dynamic linear subsystem, but the intermediate signal is assumed to be unavailable for identification [8]. If a static or dynamic subsystem is completely inaccessible in the sense that neither

Correspondence to: Anthony M. D'Amato (amdamoto@umich.edu)



Fig. 1 The goal of this work is to use data to improve the accuracy of an initial model. In other words, initial model + data = improved model.

its input nor its output is measured, then the identification problem becomes significantly more challenging.

The uncertain physics of a subsystem may range from the simplest case of an unknown parameter (such as a diffusion constant), to a multivariable spatially dependent static mapping (such as a conductivity tensor or boundary conditions), to a fully dynamic relationship among multiple variables (such as reaction kinetics). The difficulty of identifying these phenomena from empirical data depends on the accessibility of the subsystem, while the ability to use data to update a model despite limited accessibility is the goal of model refinement.

Model refinement begins with an initial model, which may incorporate both physical laws and empirical observations. The components of the initial model may have varying degrees of fidelity, reflecting knowledge, or ignorance of the relevant physics as well as the availability of data. With this initial model as a starting point, the goal is to use additional measurements to refine the model. Components of the model that are poorly modeled can be updated, thereby resulting in a higher fidelity model, as shown in Fig. 1. This problem is variously known as model correction, empirical correction, model refinement, model calibration, or model updating, and relevant literature includes [9–12] on finite-element modeling, [13–15] on meteorology, [16] on feedback control, as well as algorithms [17–19] with application to health monitoring [20,21]. Model refinement is thus a specialized version of identification, which is typically concerned with the construction of a model of the entire system.

When cast in the form of a block diagram, the model refinement problem has the form of an adaptive control system [17–20,22]. This resemblance suggests that adaptive control methods may be effective for tackling the model refinement problem. To do this, we require techniques for adaptive control that are sufficiently general and computationally tractable to address the features of large-scale physically meaningful applications. We thus apply the retrospective-cost adaptive control (RCAC) technique [23–25], which differs from standard adaptive control approaches in several ways. Specifically, RCAC requires minimal modeling information concerning the known portion of the system and is applicable to a wide range of adaptive control problems, including stabilization, command following, disturbance rejection, and model

following. RCAC utilizes a surrogate cost function that entails a closed-form quadratic (and thus convex) optimization step. The controller update requires information about only the zeros of the system; no information about the poles is needed. Furthermore, the control update requires knowledge of only the nonminimum-phase zeros of the system. For model refinement, the relevant adaptive control problem is adaptive disturbance rejection, where the ‘disturbance’ to be rejected is the unknown external excitation signal. Model refinement based on RCAC is called *adaptive model refinement*.

In this paper, we formulate adaptive model refinement for linear systems. We then demonstrate the method on a linear numerical example as well as on an experimental setup. Next, we apply adaptive model refinement to a first-principles model of the ionosphere and thermosphere. Specifically, we use the global ionosphere thermosphere model (GITM) [26] to provide a known initial model. We then use data from a ‘truth model’ version of GITM in order to refine the initial model. Although the techniques developed in refs [23,24] apply to linear systems, this paper shows that model refinement based on RCAC can be effective for large-scale nonlinear systems such as GITM. Additional relevant literature on retrospective cost optimization includes refs [27–34].

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 [35–37] 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 [38].

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 [39]. The number of latitude 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 [26]. This flexibility can be used to validate consistency by running model refinement at various levels of resolution.

First principle models of the atmosphere are strongly influenced by unknowns such as thermal conductivity coefficients and cooling processes. These effects cannot be directly measured at each altitude, and thus they are inaccessible. We identify these subsystems, which are

assumed to be unknown or uncertain, using data from simulated satellites on orbit. We then correct the uncertain model to demonstrate the feasibility of implementing the adaptive model refinement technique. A preliminary version of some of the results in this paper have appeared in the conference papers [40,41].

In Section 2, we describe the model refinement problem for subsystem identification. In Section 3, a linear problem formulation is cast using transfer functions to represent the initial model and the unknown subsystem. In Section 4, we present retrospective cost optimization as a method for obtaining an estimate of the unknown subsystem. In Section 5, the technique is demonstrated on linear numerical examples, as well as an experimental example. In Section 6, we apply the technique to a nonlinear example, specifically, parameter estimation and dynamic subsystem identification in the ionosphere and thermosphere.

2. ADAPTIVE MODEL REFINEMENT FOR SUBSYSTEM IDENTIFICATION

Figure 2 shows a block diagram of the model refinement problem. Each block is labeled to denote its uncertainty status. The blocks labeled ‘Known Subsystem’ and ‘Unknown Subsystem’ represent the physical system, whose inputs include known and unknown inputs, known as

drivers. These subsystems are connected through feedback, which captures the fact that each subsystem impacts the other. The majority of the dynamics of the system are assumed to be included in the ‘Known Subsystem’, while the ‘Unknown Subsystem’ includes static or dynamic maps that are poorly known. Both the input y_0 and the output u of the ‘Unknown Subsystem’ are assumed to be unavailable, and thus this subsystem is not accessible. The objective is to use data to better understand the ‘Unknown Subsystem’. The unknown drivers v , which are unmeasured excitations to the system, may corrupt the estimated model of the unknown subsystem, despite the model-error signal z tending to zero.

The lower part of the diagram in Fig. 2 constitutes the ‘Simulated System’. The ‘Physics Model’, which is implemented in computation, captures the dynamics of the ‘Known Subsystem’ and serves as the initial model. The ‘Physics Model’ is interconnected by feedback with the block labeled ‘Identified Physics’, which is refined by the ‘Physics Update’ procedure, which is denoted by the diagonal arrow. The ‘Physics Update’ is a tuning procedure that recursively identifies the unknown physics as data become available to provide a model of the ‘Unknown Subsystem’. This tuning procedure is driven by the model-error signal z , which is the difference between the data y from the ‘Physical System’ and the computed output \hat{y} of the ‘Simulated System’.

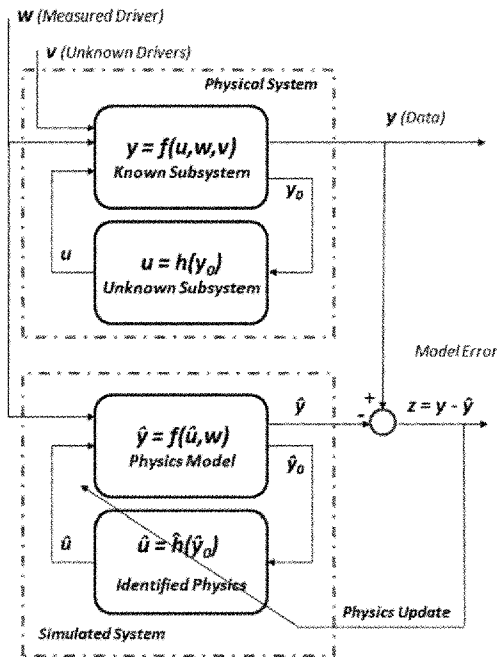


Fig. 2 This block diagram illustrates the model refinement problem, where the goal is to identify the ‘Unknown Subsystem’ of the ‘Physical System’. By depicting this problem as a block diagram, it becomes evident that the model refinement problem is equivalent to a problem of adaptive command following.

3. LINEAR PROBLEM FORMULATION

From Fig. 2, we consider a transfer function representation of the known subsystem $y = f(u, w)$, which is modeled by

$$\begin{bmatrix} y \\ y_0 \end{bmatrix} = \begin{bmatrix} G_{wy} & G_{uy} \\ G_{wy_0} & G_{uy_0} \end{bmatrix} \begin{bmatrix} w \\ u \end{bmatrix} = G \begin{bmatrix} w \\ u \end{bmatrix}, \quad (1)$$

where G is the known initial model, y the output data, w the measured input signal, y_0 the input to the unknown subsystem h , and u the output of h . Furthermore, $u = h(y_0)$ is represented by the transfer function

$$\begin{aligned} u &= G_{\Delta} \begin{bmatrix} y_0 \\ w \end{bmatrix} \\ &= [G_{\Delta, y_0} \ G_{\Delta, w}] \begin{bmatrix} y_0 \\ w \end{bmatrix} \\ &= G_{\Delta, y_0} y_0 + G_{\Delta, w} w. \end{aligned} \quad (2)$$

We stress that h is not accessible, that is, measurements of the signals u and y_0 are not available, and thus G_{Δ} cannot be identified using standard techniques. From Eqs. (1)

and (2), we obtain the closed-loop transfer function from w to y given by

$$y = [G_{wy} + G_{uy}(G_{\Delta,y_0}[I - G_{uy_0}G_{\Delta,y_0}]^{-1} \times [G_{wy_0} + G_{uy_0}G_{\Delta,w}] + G_{\Delta,w})]w. \quad (3)$$

The goal is to estimate the unknown subsystem \hat{G}_Δ such that the simulated system

$$\hat{y} = [G_{wy} + G_{uy}(\hat{G}_{\Delta,y_0}[I - G_{uy_0}\hat{G}_{\Delta,y_0}]^{-1} \times [G_{wy_0} + G_{uy_0}\hat{G}_{\Delta,w}] + \hat{G}_{\Delta,w})]w \quad (4)$$

matches the physical system, that is, the model-error signal

$$z = y - \hat{y} \quad (5)$$

is small.

To identify the feedback term G_Δ using the given initial model G , we use an adaptive feedback model structure to identify $\hat{G}_\Delta = [\hat{G}_{\Delta,y_0} \hat{G}_{\Delta,w}]$. To enforce model matching, we minimize the model-error signal z in the presence of the measured signal w . In particular, we use RCAC in a disturbance-rejection architecture. The only signals available to RCAC are the measurement y , the simulated system inputs and outputs u , \hat{y} , \hat{y}_0 , and the model-error signal z .

4. RETROSPECTIVE COST OPTIMIZATION

To model the ‘Identified Physics’, consider a strictly proper time-series model of order n_c , such that (2) is given by

$$u(k) = \sum_{i=1}^{n_c} M_i(k)u(k-i) + \sum_{i=0}^{n_c} N_i(k)\hat{y}_0(k-i) + \sum_{i=0}^{n_c} L_i(k)w(k-i), \quad (6)$$

where, for all $i = 1, \dots, n_c$, $M_i : \mathbb{N} \rightarrow \mathbb{R}^{l_u \times l_u}$, $N_i : \mathbb{N} \rightarrow \mathbb{R}^{l_u \times l_y}$ and $L_i : \mathbb{N} \rightarrow \mathbb{R}^{l_u \times l_w}$ are determined by the adaptive law presented below. Equation (6) can be expressed as

$$u(k) = \theta(k)\phi(k), \quad (7)$$

where

$$\theta(k) \triangleq [N_1(k) \cdots N_{n_c}(k) \ L_1(k) \cdots L_{n_c}(k) \ M_1(k) \cdots M_{n_c}(k)]$$

and

$$\phi(k) \triangleq [\hat{y}_0^T(k-1) \cdots \hat{y}_0^T(k-n_c) \ w^T(k-1) \cdots w^T(k-n_c) \ u^T(k-1) \cdots u^T(k-n_c)]^T \in \mathbb{R}^{n_c(l_u+l_y+l_w)}.$$

Next, we represent Eq. (5) as the time-series model with inputs u and w and output z given by

$$z(k) = y(k) - \left[\sum_{i=1}^n -\alpha_i y(k-i) + \sum_{i=d}^n \beta_i u(k-i) + \sum_{i=0}^n \gamma_i w(k-i) \right], \quad (8)$$

where $\alpha_1, \dots, \alpha_n \in \mathbb{R}$, $\beta_d, \dots, \beta_n \in \mathbb{R}^{l_z \times l_u}$, $\gamma_0, \dots, \gamma_n \in \mathbb{R}^{l_z \times l_w}$, and the relative degree d is the smallest nonnegative integer i such that the i th Markov parameter of G_{yu} is nonzero, where the Markov parameters are the components of the system’s impulse response [1,2].

Next, we define the *retrospective performance*

$$\hat{z}(\hat{\theta}, k) \triangleq y(k) - \left[\sum_{i=1}^n -\alpha_i y(k-i) + \sum_{i=d}^n \beta_i \theta(k-i)\phi(k-i) + \sum_{i=0}^n \gamma_i w(k-i) + \sum_{i=d}^v \bar{\beta}_i [\hat{\theta} - \theta(k-i)]\phi(k-i) \right], \quad (9)$$

where $v \geq d$, $\hat{\theta} \in \mathbb{R}^{l_u \times (n_c(l_y+l_u))}$ is an optimization variable used to derive the adaptive law, and $\bar{\beta}_d, \dots, \bar{\beta}_v \in \mathbb{R}^{l_z \times l_u}$. RCAC uses a retrospective performance measure, in which the performance measurement is modified based on the difference between the actual past control inputs and the recomputed past control inputs. The parameters v and $\bar{\beta}_d, \dots, \bar{\beta}_v$ must capture the information included in the first nonzero Markov parameter and the nonminimum-phase zeros from u to z [42]. In this paper, we let $\bar{\beta}_d, \dots, \bar{\beta}_v$ denote Markov parameters of the transfer function from u to z . Alternative choices of the parameters v and $\bar{\beta}_d, \dots, \bar{\beta}_v$ are discussed in ref. 42.

Next, subtracting Eq. (8) from Eq. (9) yields

$$\hat{z}(\hat{\theta}, k) = z(k) + \sum_{i=d}^n \beta_i [\hat{\theta} - \theta(k-i)]\phi(k-i). \quad (10)$$

Defining $\hat{\Theta} \triangleq \text{vec } \hat{\theta} \in \mathbb{R}^{n_c l_u (l_y+l_u+l_w)}$ and $\Theta(k) \triangleq \text{vec } \theta(k) \in \mathbb{R}^{n_c l_u (l_y+l_u+l_w)}$, it follows that

$$\hat{z}(\hat{\Theta}, k) = z(k) + \sum_{i=d}^n \Phi_i^T(k) [\hat{\Theta} - \Theta(k-i)] = z(k) - \sum_{i=d}^n \Phi_i^T(k)\Theta(k-i) + \Psi^T(k)\hat{\Theta}, \quad (11)$$

where, for $i = d, \dots, n$,

$$\Phi_i(k) \triangleq \phi(k-i) \otimes \beta_i^T \in \mathbb{R}^{(n_c l_u (l_y+l_u+l_w)) \times l_z},$$

where ‘vec’ denotes the column-stacking operator, \otimes represents the Kronecker product [43], and

$$\Psi(k) \triangleq \sum_{i=d}^n \Phi_i(k).$$

We now consider the *retrospective cost function*

$$J(\hat{\Theta}, k) \triangleq \hat{z}^T(\hat{\theta}, k) R_1(k) \hat{z}(\hat{\theta}, k) + \text{tr}[R_2(k)(\hat{\theta} - \theta(k))^T R_3(k)(\hat{\theta} - \theta(k))], \quad (12)$$

where $R_1(k) \triangleq I_{l_z}$, $R_2(k) \triangleq \alpha(k) I_{n_c(l_y+l_u+l_w)}$, and $R_3(k) \triangleq I_{l_u \times l_u}$. Using Kronecker algebra, (12) can be written as the quadratic form

$$J(\hat{\Theta}, k) = c(k) + b^T(k)\hat{\Theta} + \hat{\Theta}^T A(k)\hat{\Theta},$$

where

$$\begin{aligned} A(k) &\triangleq \Phi(k)\Phi^T(k) + \alpha(k)I, \\ b(k) &\triangleq 2\Phi(k) \left[\sum_{i=d}^n \Theta^T(k-i)\Phi_i(k) + z^T(k) \right] - 2\alpha(k)\Theta(k), \\ c(k) &\triangleq \left[\sum_{i=d}^n \Phi_i^T(k)\Theta(k-i) \right] R_1(k) \left[\sum_{i=d}^n \Theta^T(k-i)\Phi_i(k) \right] \\ &\quad + \text{tr} \left[R_2(k)\theta^T(k)R_3(k)\theta(k) \right]. \end{aligned}$$

Because $A(k)$ is positive definite, $J(\hat{\Theta}, k)$ has the strict global minimizer

$$\hat{\theta} = \frac{1}{2} \text{vec}^{-1}(A(k)^{-1}b(k)). \quad (13)$$

The gain update law is to set $\theta(k+1)$ to the global minimizer (13), that is,

$$\theta(k+1) = \hat{\theta}. \quad (14)$$

The coefficients of the time series (6) given by Eq. (14) contain information about the unknown subsystem, such as its poles, zeros, time constants, and frequency response.

RCAC requires the selection of several parameters. Specifically, n_c is the estimated order of the unknown subsystem, while ν is the number of Markov parameters obtained from the known model. The adaptive update law (14) is based on the quadratic cost function (12), which involves the time-varying weighting parameter $\alpha(k) > 0$, referred to as the *learning rate* because it affects the convergence speed of convergence of the adaptive model refinement algorithm.

The methodology for choosing these parameters is as follows. For dynamic subsystem identification, the subsystem order n_c is typically unknown. In this case, it is convenient to overestimate the subsystem order. For parameter estimation, choosing $n_c = 0$ is a natural choice in Eq. (6), since the resulting G_Δ is static. The number ν of Markov parameters is usually chosen to be 1; however, a larger value is typically needed if nonminimum-phase zeros are present in the initial model [25].

5. LINEAR EXAMPLES

5.1. Dynamic Subsystem Estimation

Consider the mass-spring-damper structure shown in Fig. 3 modeled by

$$m_1\ddot{q} + c_1\dot{q} + k_1q = w, \quad (15)$$

where m_1 , c_1 , k_1 are the known mass, damping, and stiffness, respectively, and w is a force input. As shown in Fig. 3, the mass is also connected to an unknown impedance G_Δ , which applies force to the mass in response to the velocity of the mass. We obtain the state-space representation of the known subsystem

$$\begin{bmatrix} \dot{q} \\ \ddot{q} \end{bmatrix} = A_c \begin{bmatrix} q \\ \dot{q} \end{bmatrix} + B_c u + D_{1,c} w, \quad (16)$$

$$y = C \begin{bmatrix} q \\ \dot{q} \end{bmatrix}, \quad (17)$$

where q and \dot{q} are the position and velocity, respectively, of the mass, and

$$A_c = \begin{bmatrix} 0 & 1 \\ -\frac{k_1}{m_1} & -\frac{c_1}{m_1} \end{bmatrix}, \quad B_c = D_{1,c} = \begin{bmatrix} 0 \\ \frac{1}{m_1} \end{bmatrix}, \quad C = [0 \ 1]. \quad (18)$$

Finally, we write the system in transfer function form $G(s) = C(sI - A_c)^{-1}B_c$, where s is the Laplace transform variable and the closed-loop transfer function from w to

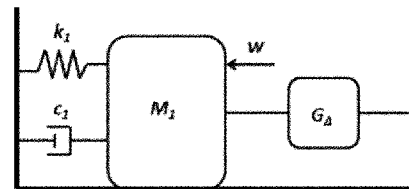


Fig. 3 A single-degree-of-freedom mass-spring-damper system connected to an unknown impedance.

y with the unknown impedance $u = G_{\Delta}y$ is $G_{cl}(s) = G(s)/(1 - G(s)G_{\Delta}(s))$.

To demonstrate adaptive model refinement, we choose $m_1 = 1 \times 10^{-4}$, $k_1 = 1$, $c_1 = 5.0275 \times 10^{-4}$, and $G_{\Delta}(s) = ((s + 30)(s + 60))/((s + 20)(s + 50)(s + 10))$. Next, the continuous-time system (17) is converted to discrete time using $A = e^{A_c T_s}$ and $B = A_c^{-1}[A - I]B$, where $T_s = 0.1$ sec is the sample time. The resulting discrete-time transfer function is $\overline{G}(z) = C(zI - A)^{-1}B$, where z is the Z-transform variable. Furthermore, $\overline{G}_{\Delta}(z)$ denotes the discretized transfer function of $G_{\Delta}(s)$.

Next, we choose $n_c = 5$, which is an overestimate of the order of \overline{G}_{Δ} , $\alpha = 1$, and $\nu = 10$, that is, we use ten Markov parameters of $\overline{G}(z)$. Figure 4(a) compares the frequency responses of the initial model and the closed-loop model consisting of the initial model and the subsystem

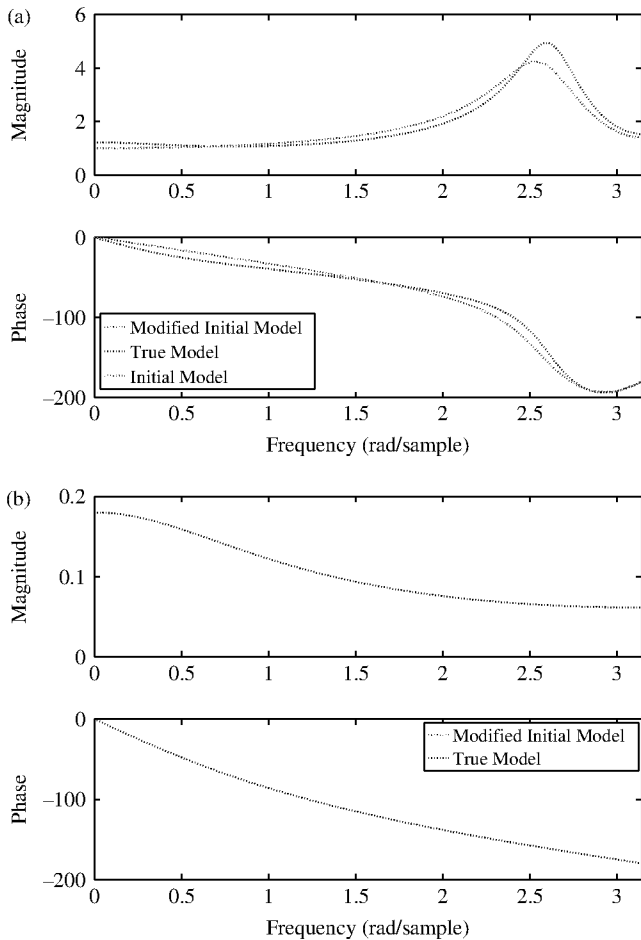


Fig. 4 (a) Comparison of the frequency response of the initial model $\overline{G}(z)$, the closed-loop $\overline{G}_{cl}(z)$, and the estimated closed loop using the identified unknown feedback $\hat{\overline{G}}_{cl}(z)$. (b) Comparison of the frequency response of the unknown feedback and the identified feedback. [Color figure can be viewed in the online issue, which is available at wileyonlinelibrary.com.]

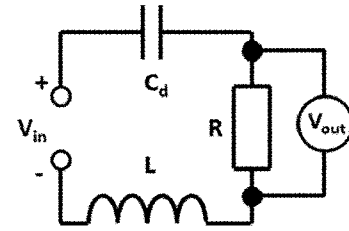


Fig. 5 A series resistor-inductor-capacitor (RLC) circuit, where voltage is measured across the resistor. The inductance L and the capacitance C_d are assumed to be uncertain.

estimate $\hat{\overline{G}}_{\Delta}$ of \overline{G}_{Δ} . The difference between the initial model and the closed-loop model is reduced by including the estimate $\hat{\overline{G}}_{\Delta}$ of \overline{G}_{Δ} ; in fact, the estimated closed-loop-model frequency response is almost identical to the frequency response of the ‘Physical System’. Figure 4(b) compares the frequency responses of \overline{G}_{Δ} and $\hat{\overline{G}}_{\Delta}$.

5.2. Static Parameter Estimation

To demonstrate adaptive model refinement for parameter estimation, we consider the series resistor-inductor-capacitor (RLC) circuit shown in Fig. 5 modeled by

$$L\ddot{x} + R\dot{x} + \frac{1}{C_d}x = u, \tag{19}$$

where L , C_d , and R are the inductor, capacitor, and resistor values, respectively, and w is the input voltage. A state-space representation of the circuit is given by

$$\begin{bmatrix} \dot{q} \\ \dot{\dot{q}} \end{bmatrix} = \begin{bmatrix} 0 & 1 \\ -\frac{1}{LC_d} & -\frac{R}{L} \end{bmatrix} \begin{bmatrix} q \\ \dot{q} \end{bmatrix} + \begin{bmatrix} 0 \\ \frac{1}{L} \end{bmatrix} u, \tag{20}$$

$$y = \begin{bmatrix} 0 & R \end{bmatrix} \begin{bmatrix} q \\ \dot{q} \end{bmatrix}, \tag{21}$$

where q and \dot{q} are the charge and current, respectively, of the circuit. Next, we write the state-space equations for the circuit with an uncertainty ΔC_d in the capacitance and ΔL in the inductance as

$$\begin{bmatrix} \dot{q} \\ \dot{\dot{q}} \end{bmatrix} = \begin{bmatrix} 0 & 1 \\ -\frac{1}{(L+\Delta L)(C_d+\Delta C_d)} & -\frac{R}{L+\Delta L} \end{bmatrix} \begin{bmatrix} q \\ \dot{q} \end{bmatrix} + \begin{bmatrix} 0 \\ \frac{1}{L+\Delta L} \end{bmatrix} w, \tag{22}$$

$$y = \begin{bmatrix} 0 & R \end{bmatrix} \begin{bmatrix} q \\ \dot{q} \end{bmatrix}, \tag{23}$$

where $u = G_{\Delta_q}q + G_{\Delta_{\dot{q}}}\dot{q} + G_{\Delta_w}w$. Estimates $\hat{\Delta C}_d$ of ΔC_d and $\hat{\Delta L}$ of ΔL can be obtained from the adaptive

model refinement estimates $\hat{G}_{\Delta,q}$ of $G_{\Delta,q}$, $\hat{G}_{\Delta,\dot{q}}$ of $G_{\Delta,\dot{q}}$, and $\hat{G}_{\Delta,w}$ of $G_{\Delta,w}$ by means of

$$\Delta \hat{L} = \frac{L}{\hat{G}_{\Delta,w}} - L = \frac{-RL}{-R + \hat{G}_{\Delta,x}} - L, \quad (24)$$

$$\Delta \hat{C}_d = -L \left(\frac{-1}{C_d} + \hat{G}_{\Delta,x} \right)^{-1} (L + \Delta L)^{-1} - C_d. \quad (25)$$

Next, we assemble a circuit with $R = 250 \Omega$, $L + \Delta L = 55 \text{ mH}$, and $C_d + \Delta C_d = 23.5 \mu\text{F}$. We assume that we do not have knowledge of either ΔC_d or ΔL , but only the initial estimates $C_d = 1 \text{ F}$ and $L = 2 \mu\text{H}$. The model (21) is similarly discretized. We drive the circuit using zero mean, Gaussian white noise, and we measure the voltage across the resistor.

We implement RCAC to obtain estimates of the transfer functions $\hat{G}_{\Delta,q}$, $\hat{G}_{\Delta,\dot{q}}$, and $\hat{G}_{\Delta,w}$. Figure 6(a) shows the history of the model-error signal z . Figure 7(a) compares the frequency responses of the initial model, the actual system, and the refined model, in discrete time.

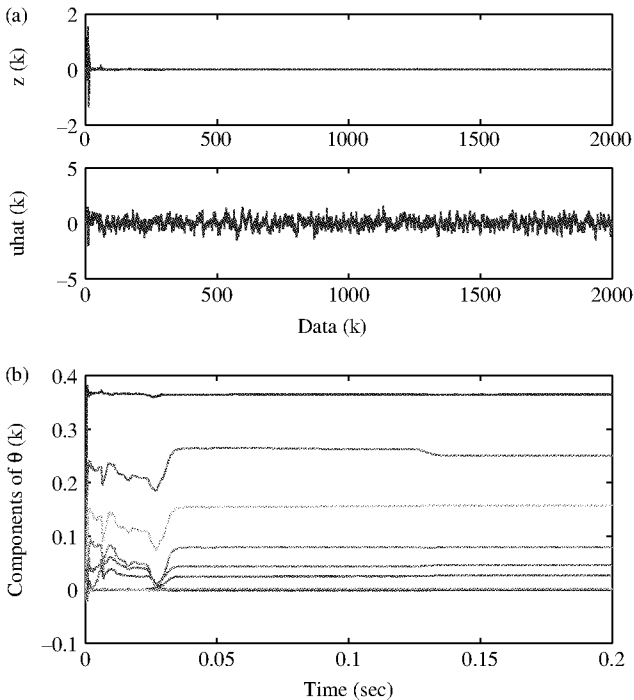


Fig. 6 (a) The history of the model-error signal $z = \hat{y} - y$ and output \hat{u} of the estimated subsystem for the simulated RLC circuit. (b) The components of the subsystem model $\theta(k)$ as functions of time. Note that z tends to zero as k becomes large, which indicates that the output of the simulated model approaches the output of the experimental circuit. [Color figure can be viewed in the online issue, which is available at wileyonlinelibrary.com.]

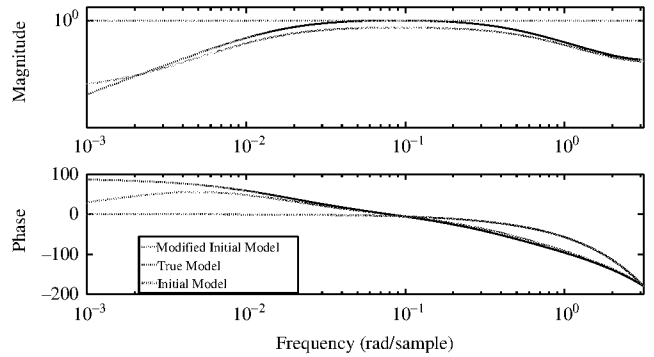


Fig. 7 This plot compares the frequency responses of the initial model \bar{G} (blue dotted line), the closed-loop \bar{G}_{cl} (black dotted line), and the estimated closed-loop $\hat{\bar{G}}_{cl}$ (red dotted line) for the simulated RLC circuit. [Color figure can be viewed in the online issue, which is available at wileyonlinelibrary.com.]

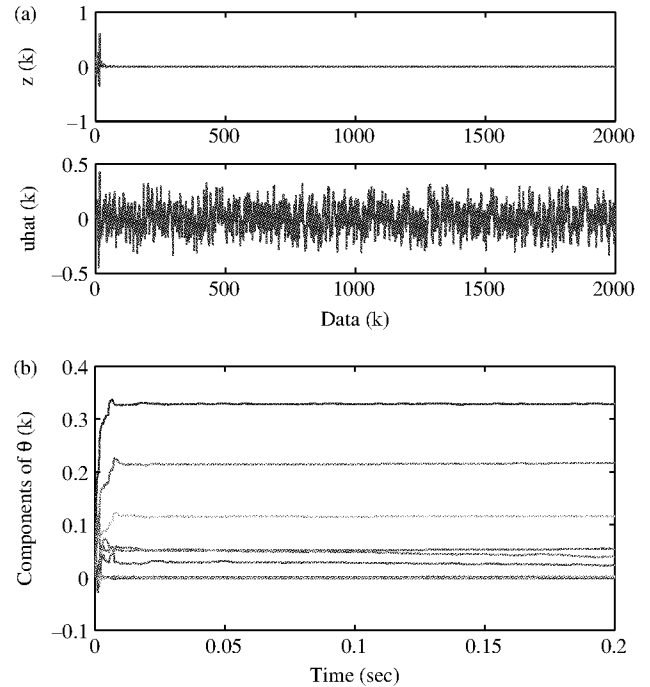


Fig. 8 (a) The history of the model-error signal $z = \hat{y} - y$ and the output \hat{u} of the estimated subsystem for the experimental RLC circuit. (b) The components of the subsystem model $\theta(k)$ as functions of time. Note that z tends to zero as k becomes large, which indicates that the output of the simulated model approaches the output of the experimental circuit. [Color figure can be viewed in the online issue, which is available at wileyonlinelibrary.com.]

Next, we generate and record the driving signal and system output. Figure 8 (a) shows the history of the model-error signal z for the experimental setup. Figure 9(a) compares the discrete-time frequency responses of the initial model, the actual system, and the refined model, for the experimental setup.

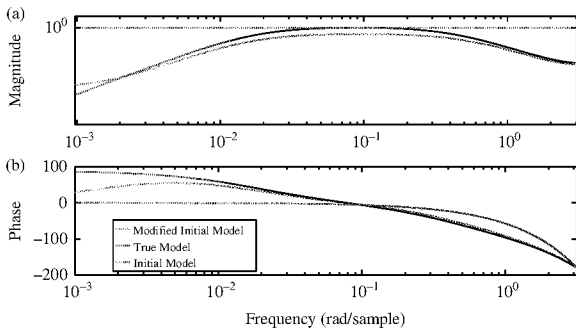


Fig. 9 This plot compares the frequency responses of the initial model \bar{G} (blue dotted line), the closed-loop \bar{G}_{cl} (black dotted line), and the estimated closed-loop $\hat{\bar{G}}_{cl}$ (red dotted line) for the experimental RLC circuit. [Color figure can be viewed in the online issue, which is available at wileyonlinelibrary.com.]

6. APPLICATION OF ADAPTIVE MODEL REFINEMENT TO IONOSPHERIC PARAMETER ESTIMATION

We now apply adaptive model refinement to a non-linear example. We consider the problem of using

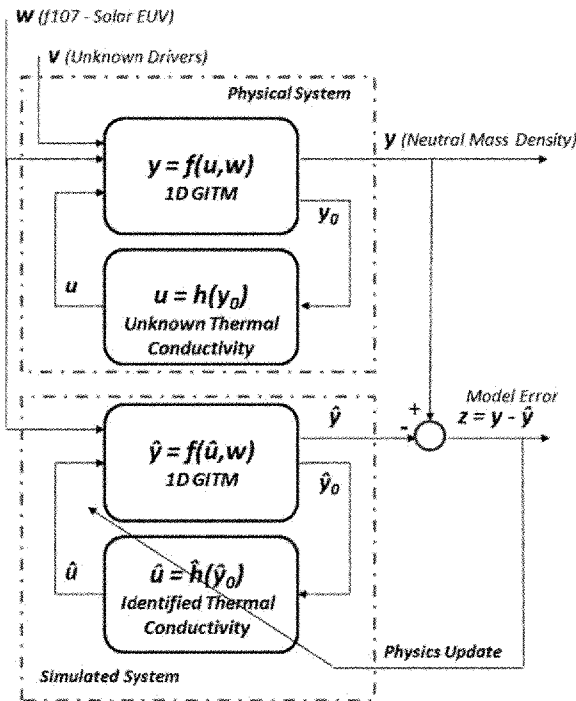


Fig. 10 This block diagram for adaptive model refinement specializes Figure 2 to a model of the ionosphere–thermosphere. Simulated data are generated by using GITM, where the thermal conductivity is assumed to be unknown. The goal is to estimate the thermal conductivity by using measurements of the neutral mass density. This problem is challenging due to the low accessibility of the unknown physics relative to the available measurements w and y .

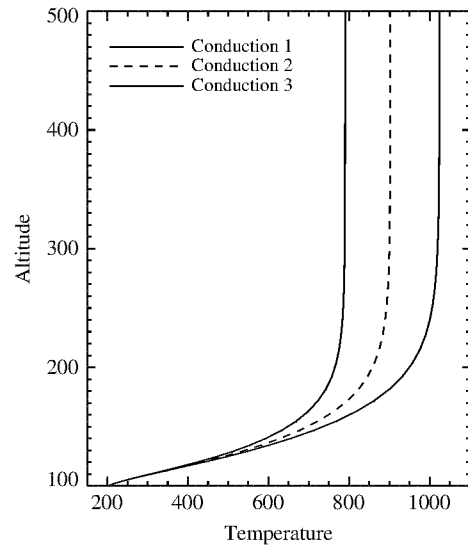


Fig. 11 Steady-state globally averaged temperature structure using three published conductivity values [45].

upper atmospheric mass-density measurements, as can be obtained from a satellite, to estimate the thermal conductivity of the thermosphere. This problem is challenging due to the fact that we do not assume the availability of measurements that can serve as inputs or outputs to the ‘Unknown

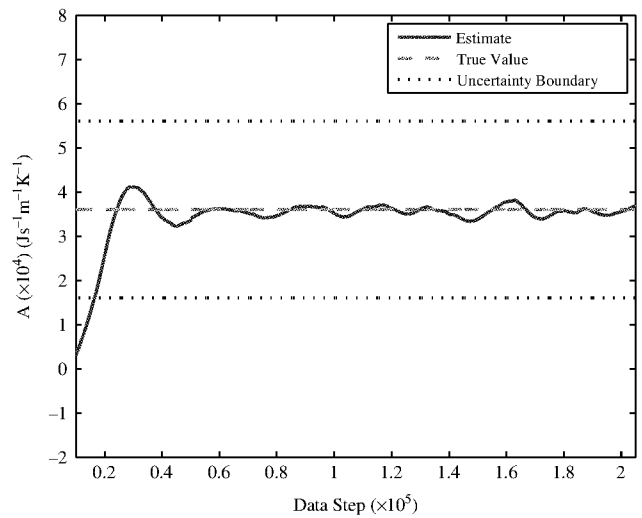


Fig. 12 This plot shows the true and estimated thermal conductivity coefficient. The initial guess for the thermal conductivity is zero, while the actual thermal conductivity is set to be the mean of the range of uncertainty. The estimate \hat{A} of A converges to a neighborhood of the true value of A within about 0.6×10^5 data points. The lack of final convergence is due to nonlinearities in the dynamics of the system. However, the oscillations are well within the uncertainty bounds, which reflect the range of published values for this coefficient. [Color figure can be viewed in the online issue, which is available at wileyonlinelibrary.com.]

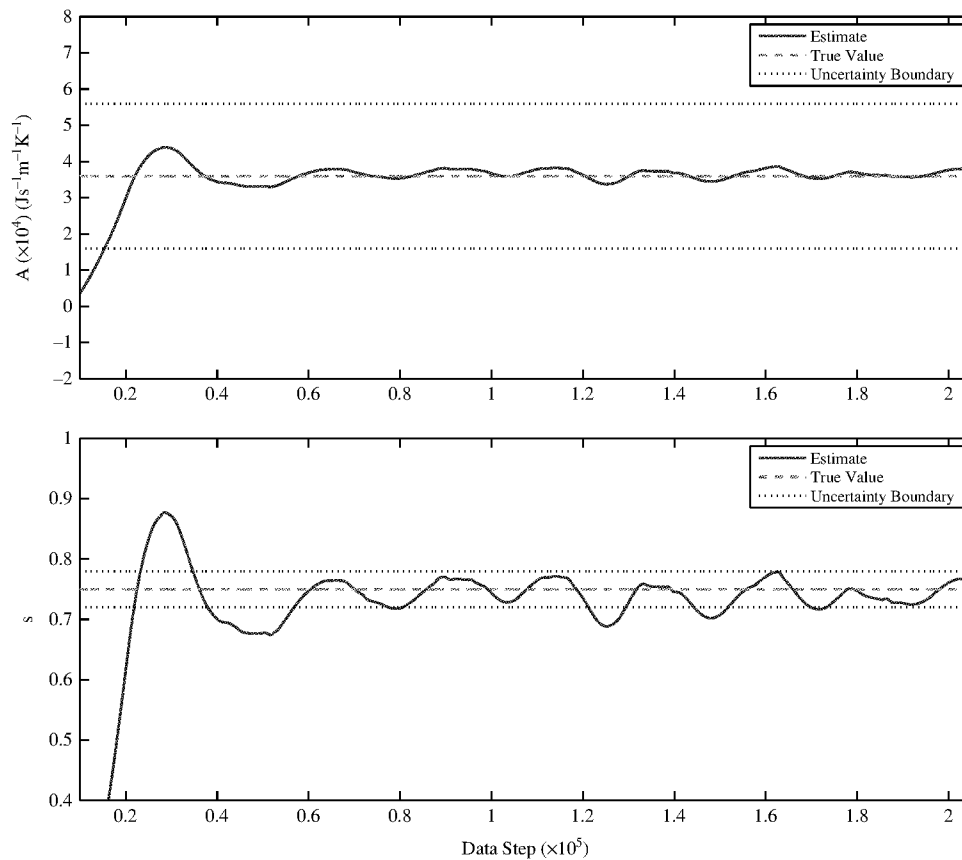


Fig. 13 These plots show the true and estimated thermal conductivity coefficient as well as the true and estimated rate coefficient. The initial guesses for both coefficients are zero. The estimates converge to a neighborhood of the true value within about 0.6×10^5 data points. The estimates are also within the uncertainty limits. [Color figure can be viewed in the online issue, which is available at wileyonlinelibrary.com.]

Subsystem', which models thermal conductivity. In other words, the objective of the identification in this application is a subsystem whose physics are inaccessible relative to the available measurements.

We use GITM to simulate the chemistry and fluid dynamics in a one-dimensional (1D) column in the ionosphere–thermosphere Fig. 10. The temperature structure of the thermosphere depends on various factors, such as the sun's intensity in extreme ultraviolet (EUV) wavelengths, eddy diffusion in the lower thermosphere, radiative cooling of the O_2 and NO , frictional heating, and the thermal conductivity.

The structure of the thermal conductivity is $\lambda = AT^s$, where A and s are the thermal conductivity and rate coefficients, respectively. The thermal conductivity may depend on chemical constituents (e.g., N_2 , O_2 , and O). Uncertainty concerning the values of A and s [44] can strongly influence the temperature structure. The need to estimate these coefficients is motivated by Fig. 11 from [45], where published values of these coefficients are shown to vary depending on the reference source. For illustration,

we assume that the true value of A is the mean of the range of values, and we seek estimates of A that are within this range.

To estimate the unknown thermal conductivity coefficient A , we apply adaptive model refinement to simulated measurements of neutral mass density provided by 1D GITM. We do this by running a 'truth model', from which we extract mass-density data at 400-km altitude, which is a typical altitude for satellites. The thermal conductivity coefficient is initialized to be zero, and its value is updated recursively. Figure 12 shows the evolution of the estimate \hat{A} of the thermal conductivity A as more data become available. The estimate \hat{A} is seen to converge to a neighborhood of the true value within about 0.6×10^4 data points.

To further illustrate the adaptive model refinement, we now assume that both the thermal conductivity A and the rate coefficient s are unknown. The parameters A and s are initialized as zero, and are updated simultaneously and recursively. Figure 13 shows the update of the estimates.

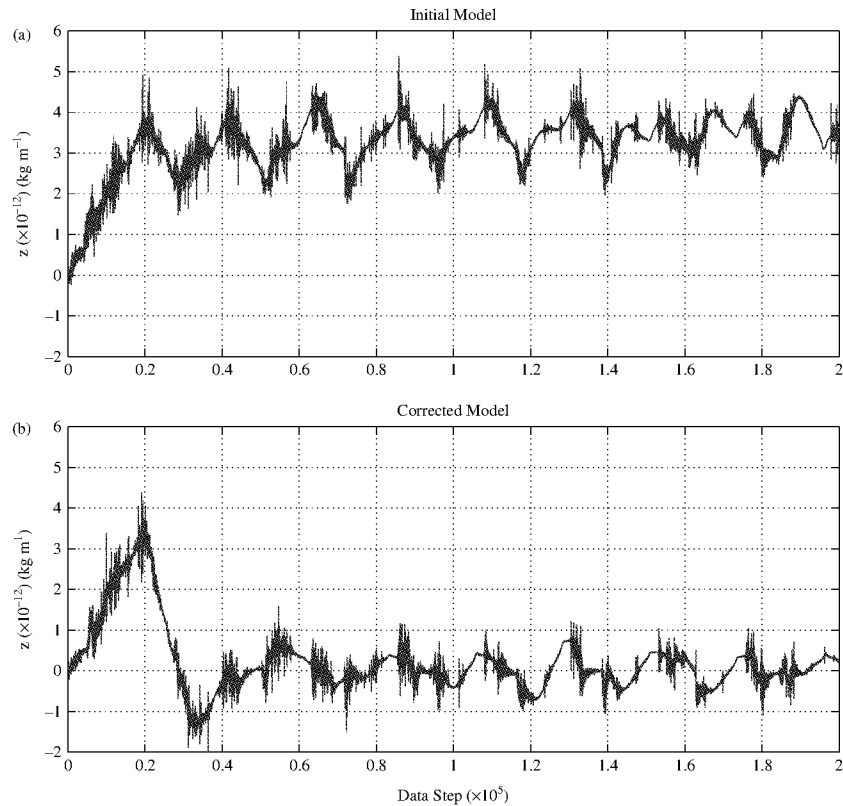


Fig. 14 (a) The model-error signal z for the difference in neutral mass-density output between the GITM truth model and the GITM initial model. (b) The difference in neutral mass-density output between the GITM truth model and the refined GITM model. By utilizing empirically refined estimates of the thermal conductivity and rate coefficient, the model error is reduced. [Color figure can be viewed in the online issue, which is available at wileyonlinelibrary.com.]

Both estimates converge to within a neighborhood of the true values within 0.6×10^5 data points.

The improvement in model accuracy attributed to the refined parameters are shown in Fig. 14. Panel (a) of Fig. 14 shows the model-error signal z for the GITM ‘truth’ model and an initial GITM model whose thermal conductivity coefficient is set to zero. Within the simulated model, this value prevents energy deposited in one layer of the atmosphere from remaining in that layer. Panel (b) of Fig. 14 illustrates the reduction in model error obtained by including the identified coefficients, thereby accounting for the thermal conductivity of this species. The benefits of refining the GITM model are evident by the improvement in model accuracy as determined by z .

7. APPLICATION OF ADAPTIVE MODEL REFINEMENT TO IONOSPHERIC DYNAMICS ESTIMATION

To illustrate adaptive model refinement in the case of an unknown dynamic subsystem, the NO radiative cooling

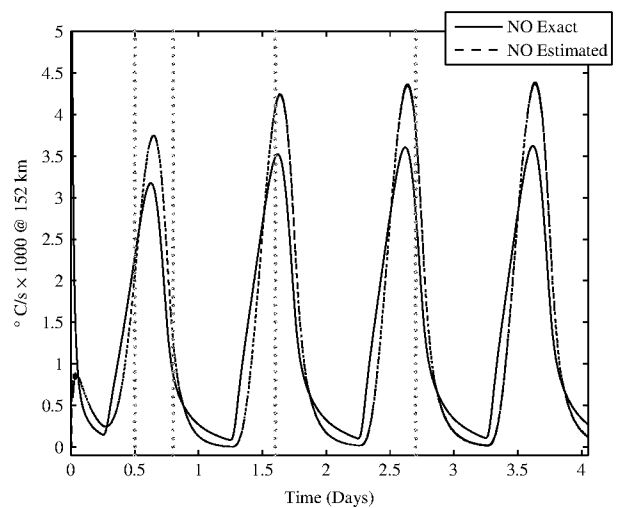


Fig. 15 This plot shows the difference between the actual NO cooling included in the GITM truth model and the cooling in the refined GITM model as a function of time at a specific altitude (152 km). The vertical dashed lines are the time instances at which the altitude versus NO cooling plots in Fig. 16 are taken. [Color figure can be viewed in the online issue, which is available at wileyonlinelibrary.com.]

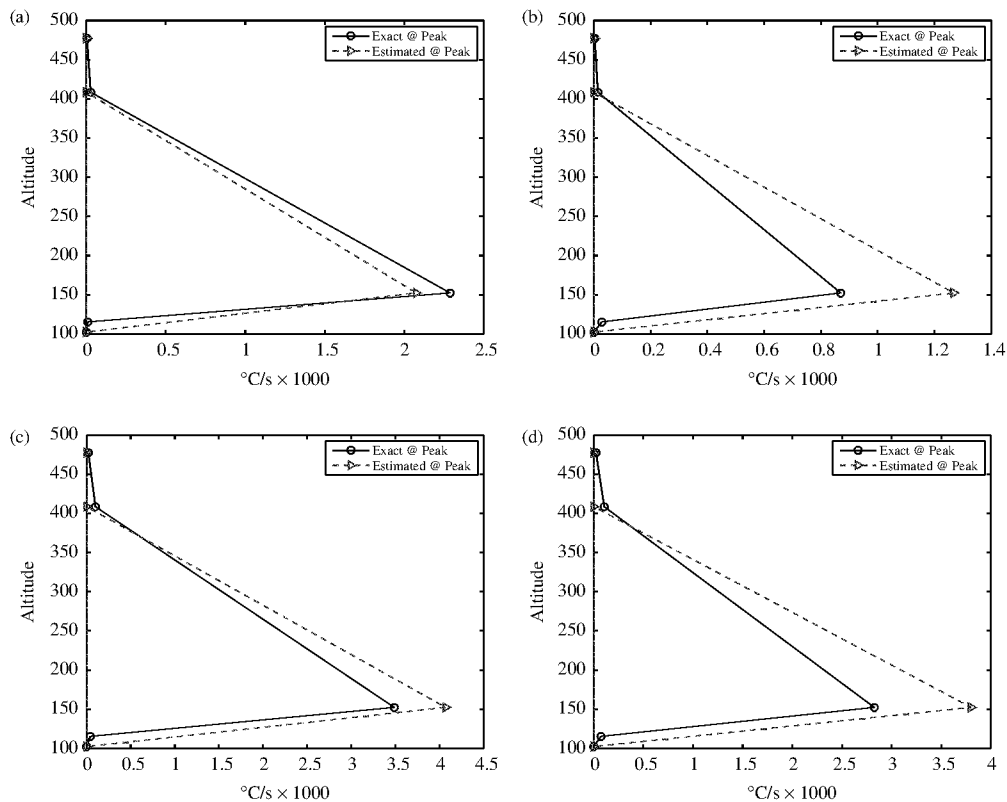


Fig. 16 These plots show the difference between the actual NO cooling included in the truth model and the cooling estimated by adaptive model refinement as a function of altitude at a given time. Cooling is along the horizontal axis, while altitude is along the vertical axis. The blue dashed line is the estimated value. The measured data are taken at an altitude of 407 km. The vertical dashed lines in Fig. 15 are the time instances at which the altitude versus NO cooling plots (a)–(d) are taken. NO cooling as function of altitude at (a) 0.5 days, (b) 0.8 days, (c) 1.6 days, (d) 2.7 days. [Color figure can be viewed in the online issue, which is available at wileyonlinelibrary.com.]

is removed from GITM to provide an initial model, but is retained in GITM for the truth model. The goal is to reproduce the missing process. This is nontrivial since the functional form of the cooling is assumed to be unknown as are the dynamics. We assume only that something is missing from the energy equation, and that this is most likely a function of temperature. The dynamics of the cooling are estimated at three different altitudes, connecting the other altitudes through linear interpolation, which is an approximation, but illustrates the technique. Nothing else about the energy sink is assumed. The thermospheric density is utilized as data at 407 km altitude from a simulated GITM truth model, which includes NO cooling. Applying adaptive model refinement with temperature as the input to the ‘Unknown Physics’, Figures 15 and 16 demonstrate that this technique captures the actual dynamics in the system. The height profile of the cooling matches the actual cooling. Furthermore, the temporal variation of the maximum cooling matches the cooling simulated by the model.

To reproduce the dynamics of the cooling, three linear dynamic equations are derived, one for each of the three chosen altitudes. This yields a profile that resembles the

natural logarithm of the NO density [46,47], indicating that this may be the source of the cooling, which it actually is. Figure 17 compares of the model without correction versus the model with correction, both of which are baselined against the truth model. Without data-based model refinement, the estimated density measurements degrade as time increases.

8. CONCLUSIONS

In this paper, we presented an adaptive model refinement technique for improving the fidelity of models using empirical data. Model refinement presents challenges relative to standard input–output system identification, specifically, a lack of accessibility to the signals that are used by standard system identification to identify the unknown subsystem. For model refinement we use retrospective cost optimization to identify the unknown subsystem. We presented a problem formulation for the linear case, and demonstrated the method on a numerical example and an experimental setup. We then demonstrated the feasibility of the method

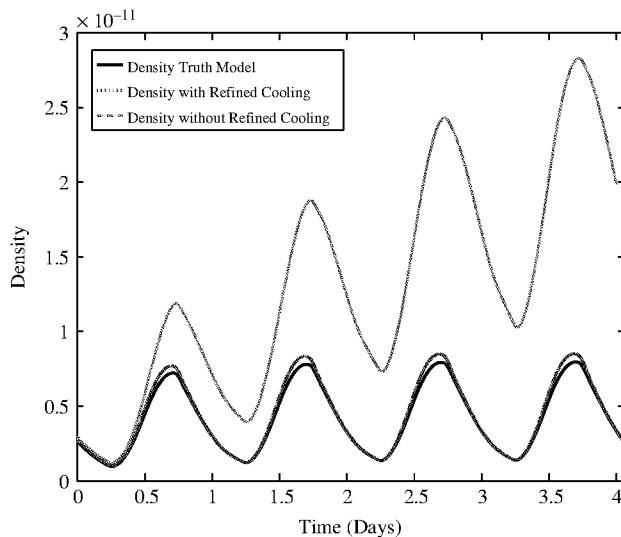


Fig. 17 This plot shows the difference between the density measurements for the initial model, where no correction is made, and the model with the refined subsystem versus the truth model. With adaptive model refinement, the refined model is able to track the truth model, whereas, if no correction is made, the density measurements degrade as time increases. [Color figure can be viewed in the online issue, which is available at wileyonlinelibrary.com.]

in refining a nonlinear model of the ionosphere and thermosphere using GITM. We demonstrated how uncertain parameters can be identified when the structure of the uncertain model is known. Furthermore, we demonstrated how unknown dynamics can be identified from data when the internal structure of the unknown subsystem is unknown. This technique can thus be used to refine and improve an initial model (or models, if several are hypothesized) that is either uncertain or erroneous. In turn, the improved model can provide a more accurate foundation for data assimilation aimed at wind and density estimates in the presence of solar storm disturbances.

ACKNOWLEDGMENTS

This research was supported in part by NASA GSRP fellowship grant NNX09AK76H, NASA grant NNX08BA57A, NSF grant CSN-1035236, NASA grant NNX09AJ59G, and NASA grant NNX11DS11B.

REFERENCES

[1] J.-N. Juang, *Applied System Identification*, New Jersey, Prentice Hall, 1993.
 [2] L. Ljung, *System Identification: Theory for the User* (2nd ed.), New Jersey, Prentice Hall, 1999.

[3] P. Van Overschee and B. De Moor, *Subspace Identification for Linear Systems: Theory, Implementation, Applications*, Berlin, Kluwer, 1996.
 [4] H. J. Palanthandalam-Madapusi, S. Gillijns, B. De Moor, and D. S. Bernstein, Subsystem Identification for Nonlinear Model Updating, *Proc Amer Contr Conf* (2006), 3056–3061.
 [5] J. S. Bendat, *Nonlinear Systems Techniques and Applications*, New York, Wiley, 1989.
 [6] J. Sjöberg, Q. Zhang, L. Ljung, A. Benveniste, B. Deylon, P. Gloennec, H. Hjalmarsson, A. Juditsky, Nonlinear black-box modeling in system identification: a unified overview, *Automatica* 31(12) (1995), 1691–1724.
 [7] R. Haber and L. Keviczky, *Nonlinear System Identification—Input-Output Modeling Approach Vol. 1: Nonlinear System Parameter Identification*, Norwell, MA, USA, Kluwer Academic Publishers, 1999.
 [8] O. Nelles, *Nonlinear System Identification*, New Jersey, Springer, 2001.
 [9] H. Palanthandalam-Madapusi, D. S. Bernstein, and A. J. Ridley, Space weather forecasting: identifying periodically switching block-structured models to predict magnetic-field fluctuations, *IEEE Control Sys Mag* 27 (2007), 109–123.
 [10] C. Minas and D. Inman, Matching finite element models to modal data, *J Vibration Acoust* 112 (1990), 84–92.
 [11] M. I. Friswell and J. E. Mottershead, *Finite Element Model Updating in Structural Dynamics*, Dordrecht, Kluwer, 1995.
 [12] S. O. R. Moheimani, Model correction for sampled-data models of structures, *J Guid Control Dyn* 24(3) (2001), 634–637.
 [13] J. B. Carvalho, B. N. Datta, W. Lin, and C. Wang, Symmetry preserving eigenvalue embedding in finite-element model updating of vibrating structures, *J Sound Vibration* 290 (2006), 839–864.
 [14] F. D'Andrea and R. Vautard, Reducing systematic errors by empirically correcting model errors, *Tellus*, 52A (2000), 21–41.
 [15] T. DelSole and A. Y. Hou, Empirical correction of a dynamical model. Part I: Fundamental issues, *Mon Weather Rev* 127(11) (2001), 2533–2545.
 [16] C. M. Danforth, E. Kalnay, and T. Miyoshi, Estimating and correcting global weather model error, *Mon Weather Rev* 135(2) (2007), 281–299.
 [17] S. Mijanovic, G. E. Stewart, G. A. Dumont, and M. S. Davies, A controller perturbation technique for transferring closed-loop stability between systems, *Automatica* 39 (2003), 1783–1791.
 [18] H. Palanthandalam-Madapusi, E. L. Renk, and D. S. Bernstein. Data-based model refinement for linear and Hammerstein systems using subspace identification and adaptive disturbance rejection. In *Proceedings of Conference on Control Applications*, Toronto, Canada, (2005), 1630–1635.
 [19] M. A. Santillo, A. M. D'Amato, and D. S. Bernstein, System identification using a retrospective correction filter for adaptive feedback model updating, In *Proceedings of the American Control Conference*, St. Louis, MO, (2009), 4392–4397.
 [20] A. M. D'Amato and D. S. Bernstein, Linear fractional transformation identification using retrospective cost optimization, *Proceedings of SYSID*, Saint-Malo, France, (2009), 450–455.
 [21] A. M. D'Amato, B. J. Arritt, J. A. Banik, E. V. Ardelean, and D. S. Bernstein, Structural health determination and model

- refinement for a deployable composite boom, AIAA SDM Conference, Palm Springs, CA, (2009), AIAA-2009-2373.
- [22] A. M. D'Amato, A. R. Wu, K. S. Mitchell, S. L. Kukreja, and D. S. Bernstein, Damage localization for structural health monitoring using retrospective cost model refinement, AIAA SDM Conference, Orlando, FL, (2010), AIAA-2010-2628.
- [23] J. B. Hoagg and D. S. Bernstein, Cumulative retrospective cost adaptive control with rls-based optimization, Proceedings of the American Control Conference, Baltimore, MD, (2010), 4016–4021.
- [24] R. Venugopal and D. S. Bernstein, Adaptive disturbance rejection using ARMARKOV system representations, IEEE Trans Control Sys Tech 8 (2000), 257–269.
- [25] J. B. Hoagg, M. A. Santillo, and D. S. Bernstein, Discrete-time adaptive command following and disturbance rejection for minimum phase systems with unknown exogenous dynamics, IEEE Trans Autom Control 53 (2008), 912–928.
- [26] M. A. Santillo and D. S. Bernstein, Adaptive control based on retrospective cost optimization, AIAA J Guid Control Dyn 33 (2010), 289–304.
- [27] A. J. Ridley, Y. Deng, and G. Tóth, The global ionosphere-thermosphere model, J Atmos Sol-Terr Phys 68 (2006), 839.
- [28] M. S. Holzel, M. A. Santillo, J. B. Hoagg, and D. S. Bernstein, Adaptive control of the NASA generic transport model using retrospective cost optimization, Proceedings of the AIAA Guidance, Navigation and Control Conference, Chicago, IL, (2009), AIAA-2009-5616.
- [29] M. S. Fledderjohn, Y.-C. Cho, J. B. Hoagg, M. A. Santillo, W. Shyy, and D. S. Bernstein, Retrospective cost adaptive flow control using a dielectric barrier discharge actuator, In Proceedings of the AIAA Guidance, Navigation and Control Conference, Chicago, IL, (2009), AIAA-2009-5857.
- [30] M. A. Santillo, M. S. Holzel, J. B. Hoagg, and D. S. Bernstein, Adaptive control using retrospective cost optimization with rls-based estimation for concurrent markov-parameter updating, Proceedings of the Conference on Decision Control, Shanghai, China, (2009), 3466–3471.
- [31] A. M. D'Amato, B. O. S. Teixeira, and D. S. Bernstein, Semiparametric identification of Wiener systems using a single harmonic input and retrospective cost optimization, IET Contr Theory Appl, (to appear).
- [32] A. M. D'Amato, K. S. Mitchell, B. O. S. Teixeira, and D. S. Bernstein, Semiparametric identification of hammerstein systems using input reconstruction and a single harmonic input, In Proceedings of the Conference on Decision Control, Atlanta, GA, (2010), 6365–6370.
- [33] H. Sane and D. S. Bernstein, Active noise control using an acoustic servovalve, In Proceedings of American Control Conference, Philadelphia, PA, (1998), 2621–2625.
- [34] S. L. Lacy, R. Venugopal, and D. S. Bernstein, ARMARKOV adaptive control of self-excited oscillations of a ducted flame, In Proceedings of the Conference on Decision Control, Tampa, FL, (1998), 4527–4528.
- [35] H. S. Sane, R. Venugopal, and D. S. Bernstein, Disturbance rejection using self-tuning ARMARKOV adaptive control with simultaneous identification, IEEE Transactions on Control Systems Technology 9(1) (2001), 101–106.
- [36] E. Yigit and A. J. Ridley, Effects of high-latitude thermosphere heating at various scale sizes simulated by a non-hydrostatic global thermosphere-ionosphere model, J Atmos Sol-Terr Phys 73 (2010), 592–600.
- [37] R. G. Roble, E. C. Ridley, A. D. Richmond, and R. E. Dickinson, A coupled thermosphere/ionosphere general circulation model, Geophys Res Lett 15 (1988), 1325.
- [38] T.J. Fuller-Rowell and D. Rees, A three-dimensional, time-dependent, global model of the thermosphere, J Atmos Sci 37 (1980), 2545.
- [39] Y. 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 35 (2008), L01104.
- [40] C. Jablonowski, M. Herzog, J. E. Penner, R. C. Oehmke, Q. F. Stout, B. van Leer, and K. G. Powell, Block-structured adaptive grids on the sphere: advection experiments, Mon Weather Rev 134 (2006), 3691–3713.
- [41] A. M. D'Amato, A. J. Ridley, and D. S. Bernstein, Adaptive model refinement for the ionosphere and thermosphere, In NASA CIDU, Mountain View, CA, 2010. (<https://c3.ndc.nasa.gov/dashlink/resources/240>) (accessed 2010).
- [42] A. M. D'Amato, S. L. Kukreja, and D. S. Bernstein, “Data-based” model refinement using retrospective cost optimization, In Proceedings of the AIAA Guidance, Navigation and Control Conference, Toronto, ON, (2010), AIAA-2010-812786.
- [43] J. B. Hoagg and D. S. Bernstein, Retrospective cost adaptive control for nonminimum-phase discrete-time systems part 1: the ideal controller and error system, part 2: the adaptive controller and stability analysis, In Proceedings of the Conference on Decision Control, Atlanta, GA, (2010), 893–904.
- [44] D. S. Bernstein, Matrix Mathematics (2nd ed.), New Jersey, Princeton University Press, 2009.
- [45] R. W. Schunk and A. F. Nagy, Ionospheres, Cambridge, UK, Cambridge Press, 2000.
- [46] D. J. Pawlowski, A. J. Ridley, Quantifying the effect of thermospheric parameterization in a global model, J Atmos Sol-Terr Phys 71 (2009), 2017–2026.
- [47] C. A. Barth, K. D. Mankoff, S. M. Bailey, and S. C. Solomon, Global observations of nitric oxide in the thermosphere, J Geophys Res 108(A1) (2003), 1027.
- [48] D. R. Marsh, S. C. Solomon, and A. E. Reynolds, Empirical model of nitric oxide in the lower thermosphere, J Geophys Res 109 (2004), A07301.