Retrospective-Cost-Based Adaptive Input and State Estimation for the Ionosphere–Thermosphere

Asad A. Ali,* Ankit Goel,* Aaron J. Ridley,† and Dennis S. Bernstein‡
University of Michigan, Ann Arbor, Michigan 48109

DOI: 10.2514/1.I010286

The upper atmosphere is a strongly driven system in which the global state is rapidly altered by the solar drivers. One of the main drivers of the upper atmosphere is the solar irradiance in the extreme ultraviolet and x-ray bands. The solar irradiance in these bands is proxied by ground-based measurements of $F_{10.7}$, which is the solar irradiance at the wavelength of 10.7 cm. The problem of estimating $F_{10.7}$ and physical states in the upper atmosphere is considered by assimilating the neutral density measurements in the global ionosphere–thermosphere model and using retrospective-cost adaptive input and state estimation. Retrospective-cost adaptive input and state estimation is a non-Bayesian estimator that estimates the input by minimizing the difference between the estimator output and the output of the physical system. In this paper, we use retrospective-cost adaptive input and state estimation to estimate $F_{10.7}$ using simulated data as well as real satellite data.

I. Introduction

STATE-ESTIMATION techniques combine measurements with dynamical and statistical models. The external input may consist of an unknown stochastic driver and a known deterministic driver, where knowledge of the latter is used in the estimator. If the deterministic driver is unknown, it can be treated as a stochastic driver; however, the resulting estimates may be poor. State estimators that account for the unknown deterministic signal include unbiased Kalman filters, unknown input observers, and state estimators with input reconstruction [1–5].

In many applications, the response is primarily due to the deterministic driver since, as the effect of the initial state decays, the forcing alone governs the asymptotic response. In the linear case, systems with this property are asymptotically stable; in the nonlinear case, these systems are called incrementally stable or contractive, and this phenomenon is called entrainment [6–8]. Entrainment in strongly driven systems suggests that, at least in some cases, the state-estimation accuracy can be enhanced by estimating the unknown drivers.

In the near-Earth environment, the effects of space weather are primarily manifested by the properties of the ionosphere and thermosphere, which influence radio propagation and satellite drag. The sun is the primary driver in this region, strongly driving the combined ionosphere and thermosphere. In particular, the sun's extreme ultraviolet (EUV) and x-ray radiation produces photoionization that, in turn, through chemistry and heating, drives the ionosphere and thermosphere.

Unfortunately, it is not possible to measure EUV and x-ray flux from the ground due to atmospheric absorption. Consequently, measurements of solar irradiance at 10.7 cm wavelength ($F_{10.7}$) are used as a proxy for EUV and x-ray flux. This quantity, in units of 10^{-22} W-Hz^{-1} m^{-2} = 1 solar flux units (SFUs), is measured daily at local noon by the Dominion Radio Observatory in Penticton, Canada [9]. The present work is aimed at estimating $F_{10.7}$ at arbitrary temporal resolution in a manner that is consistent with features of the ionosphere and thermosphere by assimilating satellite neutral density measurements. Such model-based estimation of $F_{10.7}$ using alternative sensors can also be used to validate direct measurements made at ground stations. Furthermore, since models of the ionosphere–thermosphere are typically incomplete or inexact, the difference between estimates of $F_{10.7}$ and direct measurements provides the basis for model calibration and parameter estimation.

In this paper, we use the global ionosphere–thermosphere model (GITM) [10], which captures global temperature, density, and winds from 100 to 600 km above the Earth. Drivers include electric fields that give rise to the aurora, lower-atmosphere tides, and solar radiation. Chemistry and thermodynamics account for ions and neutrals.

We use retrospective-cost adaptive input and state estimation (RCAISE) [11] to estimate $F_{10.7}$ by assimilating satellite measurements in GITM. Assuming that the drivers (that is, inputs) of the system are unknown, RCAISE constructs an input to the estimator that minimizes the cost function based on the error between the measured and computed neutral densities. To do this, RCAISE adapts the coefficients of the driver-estimator subsystem by retrospectively optimizing the neutral density error with respect to the driver estimate: that is, the estimator input. The output of the adaptive driver-estimator subsystem is then used to asymptotically drive the neutral density error to zero. In this way, RCAISE estimates the unknown input (that is, the external driver) along with the unknown states of the system.

RCAISE differs from existing data assimilation methods in several ways. For example, unlike four-dimensional (4-D) variational data assimilation (VAR) techniques [12–17], RCAISE does not require an adjoint code; in fact, no adjoint code is available for GITM. Furthermore, like both 4-D VAR and ensemble-based methods, RCAISE does not require an explicit model of the system dynamics, and thus can be implemented based on a numerical black-box simulation of the dynamics. Moreover, like 4-D VAR, but unlike ensemble-based methods [18–20], RCAISE does not require an ensemble of models. In addition, RCAISE does not assume the existence of a stochastic model, and it does not use any prior information concerning the statistics of unknown quantities. For example, RCAISE does not use statistical knowledge of the sensor noise (instrument error), stochastic drivers, or parametric errors to stochastically optimize the state and driver estimates. Consequently, RCAISE is a non-Bayesian technique, and thus does not provide a statistical measure of the accuracy of its estimates: for example, in terms of a probability distribution or confidence intervals. Nevertheless, we consider applications in which sensor measurements are corrupted by noise; for example, the Challenging Mini Satellite Payload (CHAMP) [21] data used in this study are inherently subject to instrument error. Finally, RCAISE is

*Graduate Student, Aerospace Engineering Department, 1320 Beal Ave.
†Professor, Atmospheric, Oceanic, and Space Sciences, 2455 Hayward St.
‡Professor, Aerospace Engineering, 1320 Beal Ave.

relatively inexpensive to implement since, unlike 4-D VAR, it does not depend on nonquadratic optimization and, unlike Bayesian methods, it does not require an ensemble of models.

In accordance with the application of RCAISE to estimate $F_{10.7}$, the GITM study in this paper makes no assumption about prior statistics, which are required by Bayesian methods. In contrast, the related study in [22] is based on assumed statistical priors constructed based on ad hoc requirements. For example, the initial distribution of $F_{10.7}$ in [22] was chosen to be normal in accordance with the assumptions of the ensemble adjustment Kalman filter (EAKF) [20] used in the data assimilation research testbed (DART) code [23]. In this case, the standard deviation was chosen to create sufficient spread after an initial spinup in order to initialize the ensemble. This technique for choosing priors, along with inflation and localization, is representative of the "tuning knobs" routinely used to implement Bayesian ensemble methods.

The derivation of the RCAISE algorithm given in this paper is based on a linear dynamics model for which the modeling information needed to implement the algorithm consists of components of the impulse response. However, since RCAISE does not require an explicit model of the dynamics (which may, for example, be in the form of a computer code, as in the case of GITM), RCAISE can be applied to nonlinear systems. In this case, the required modeling information can be based on either the qualitative behavior of the system or by observing the response of the estimator. The numerical examples in this paper show that RCAISE is effective on nonlinear systems, which is due to the fact that the algorithm requires extremely limited modeling information. In analogy with ensemble methods, the tuning knobs used in RCAISE include the required modeling data and the weights used for regression. The sensitivity of the accuracy of RCAISE to variations in the model data is illustrated for both an illustrative example in Sec. III and for GITM in Sec. VI.

In [24], RCAISE was used to estimate constant $F_{10.7}$ in GITM using simulated measurement data, where the measurement was assumed to be at a fixed position in the atmosphere. The contribution of the present paper is to derive RCAISE, illustrate RCAISE using simple examples, formulate RCAISE for GITM, and apply RCAISE to space weather driver estimation using GITM in the case where the measurements are obtained from real and simulated satellites and where the driver $F_{10.7}$ is time varying.

As noted previously, RCAISE is intended for use with strongly driven systems with unknown deterministic drivers, and furthermore does not provide bounds or statistics on the states or drivers. Therefore, RCAISE is not intended to compete with Bayesian estimation techniques that use priors to obtain error density estimates. On the other hand, since RCAISE requires no ensemble or adjoint code, and it depends solely on linear regression, it is easy and inexpensive to implement. In contrast, the results of [22] required 640 processors, where one week of simulated time (which is typical of space weather applications) required about one week of computer time, for a total of about 107,500 CPU hours. This cost is partly due to the inefficient interface between the DART and GITM. Nonetheless, the computational cost of running RCAISE with GITM is close to the cost of running GITM alone, which is approximately 10–15 times faster than real time, so that one week of simulated time requires about 17 h of computer time, for a total of about 544 CPU hours. Consequently, GITM with the DART (as implemented in [22]) is about 200 times more expensive than GITM with RCAISE.

As shown in [25], RCAISE can also be used for model refinement through subsystem identification. For details, see [26].

The contents of this paper are as follows. In Sec. II, we formulate RCAISE for linear systems, followed by illustrative linear and nonlinear examples in Sec. III. Section IV describes GITM and CHAMP and Gravity Recovery and Climate Experiment (GRACE) [27] satellites that are used in the numerical experiments. In Sec. V, we formulate RCAISE for GITM and describe the implementation of RCAISE on a multiprocessor GITM simulation. Section VI presents state and driver estimation results. Finally, in Sec. VII, we give conclusions and future directions.

II. Retrospective-Cost Adaptive Input and State Estimation

In this section, we give a brief derivation of RCAISE for a linear time-invariant system. Consider

$$x(k + 1) = Ax(k) + Bu(k) \tag{1}$$

$$y(k) = Cx(k) \tag{2}$$

where $x(k) \in \mathbb{R}^n$ is the unknown state, $u(k) \in \mathbb{R}^m$ is an unknown driver, and $y(k) \in \mathbb{R}^p$ is the measured output. The matrices $A \in \mathbb{R}^{n \times n}$, $B \in \mathbb{R}^{n \times m}$, and $C \in \mathbb{R}^{p \times n}$ are assumed to be known.

To obtain an estimate $\hat{x}(k) \in \mathbb{R}^n$ of the state $x(k)$, we consider the state estimator

$$\hat{x}(k + 1) = A\hat{x}(k) + B\hat{u}(k) \tag{3}$$

$$\hat{y}(k) = C\hat{x}(k) \tag{4}$$

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

where $\hat{x}(k) \in \mathbb{R}^n$ is the estimated state, $\hat{y}(k) \in \mathbb{R}^p$ is the estimated output, $\hat{u}(k) \in \mathbb{R}^m$ is the estimated driver, and $z(k) \in \mathbb{R}^p$ are the innovations: that is, the output error. The signal $\hat{u}(k)$, which is the estimate of the unknown driver, is the output of an adaptive driver-estimator subsystem of order n_e for which the input is $z(k)$. The adaptive subsystem is formulated as

$$\hat{u}(k) = \sum_{i=1}^{n_e} M_i(k)\hat{u}(k - i) + \sum_{i=0}^{n_e} N_i(k)z(k - i) \tag{6}$$

where $M_i(k) \in \mathbb{R}^{m \times m}$, $i = 1, \ldots, n_e$, and $N_i(k) \in \mathbb{R}^{m \times p}$, $i = 0, \ldots, n_e$. The order n_e of the state estimator is chosen based on the assumed order of the unknown subsystem. RCAISE minimizes $z(k)$ by updating $M_i(k)$ and $N_i(k)$.

Figure 1 shows the structure of the adaptive input and state estimator [Eqs. (3–6)]. Note that the coefficients $M_i(k)$ and $N_i(k)$ of the adaptive driver-estimator subsystem [Eq. (6)] change based on $z(k)$. Also, Eq. (3) is intended to mimic Eq. (1), where the driver estimate $\hat{u}(k)$ replaces the true driver $u(k)$. The ability of Eq. (3) to estimate $x(k)$ depends on the implicit assumption that Eq. (1) is a strongly driven system, where the effect of the initial conditions goes to zero asymptotically, and the output of the system is solely due to the input.
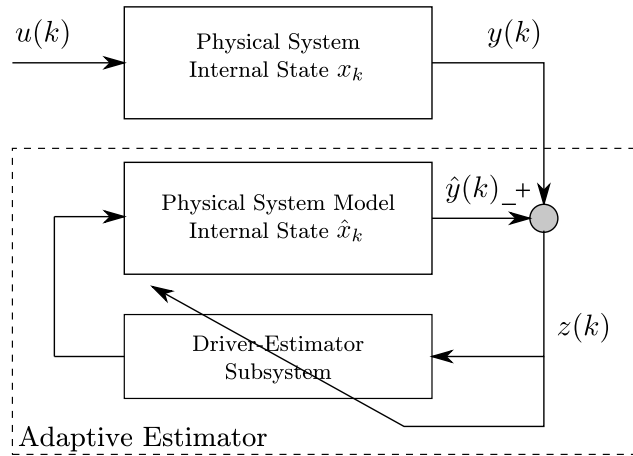
Fig. 1   Adaptive input and state estimator structure. The adaptive input and state estimator uses the error $z(k)$ between the measured output $y(k)$ and the computed output $\hat{y}(k)$ to update the adaptive driver-estimator subsystem in order to obtain an estimate $\hat{u}(k)$ of the driver $u(k)$ that minimizes the retrospective cost J defined by Eq. (22). The driver estimate $\hat{u}(k)$ asymptotically drives the error between $y(k)$ and $\hat{y}(k)$ to zero. Consequently, for a strongly driven system, the state $\hat{x}(k)$ of the physical system model converges to the state $x(k)$ of the physical system.

Note that Eq. (3) is distinct from the standard Kalman filter, which has the form

$$\hat{x}(k+1) = A\hat{x}(k) + Bu(k) + Kz(k) \tag{7}$$

where $u(k)$ represents the known deterministic driver, and the gain matrix K is updated based on the computed error covariance. Note that Eq. (7) gives only state estimates, whereas Eq. (3) gives estimates of both the state and the unknown driver, assuming that the system is asymptotically stable.

In the Kalman filter, the state estimates are improved by assimilating the output error $z(k)$. Once the state estimates converge to the true state, application of the deterministic driver $u(k)$ forces $z(k)$ to converge to zero asymptotically. However, in RCAISE, the output error $z(k)$ is used to estimate the unknown driver $u(k)$. Coupled with the assumption that the system is strongly driven, RCAISE and the estimated driver $\hat{u}(k)$ then drive $z(k)$ to zero asymptotically.

For all $i \geq 1$, define the Markov parameter H_i of (A, B, C) by $H_i = CA^{i-1}B$. Let r be a nonnegative integer. Then, substituting Eq. (3) into itself, $r - 1$ times gives, for all $k \geq r$,

$$\hat{x}(k) = A^r \hat{x}(k-r) + \sum_{i=1}^{r} A^{i-1} B\hat{u}(k-i) \tag{8}$$

It follows from Eqs. (1), (4), (5), and (8) that

$$z(k) = CA^r \hat{x}(k-r) - y(k) + \bar{H}\,\hat{\bar{U}}(k-1) \tag{9}$$

where

$$\bar{H} = [\,H_1 \quad \cdots \quad H_r\,] \in \mathbb{R}^{p \times rm}, \qquad \hat{\bar{U}}(k-1) = [\,\hat{u}^T(k-1) \quad \cdots \quad \hat{u}^T(k-r)\,]^T$$

Next, rearranging the columns of $\bar{H}$ and the components of $\hat{\bar{U}}(k-1)$, and partitioning the resulting matrix and vector, yields

$$\bar{H}\,\hat{\bar{U}}(k-1) = \mathcal{H}'\hat{U}'(k-1) + \mathcal{H}\hat{U}(k-1) \tag{10}$$

where $\mathcal{H}' \in \mathbb{R}^{p \times (rm-l)}$, $\mathcal{H} \in \mathbb{R}^{p \times l}$, $\hat{U}'(k-1) \in \mathbb{R}^{rm-l}$, and $\hat{U}(k-1) \in \mathbb{R}^l$. For example, if

$$\bar{H} = [\,H_1 \quad H_2 \quad H_3 \quad H_4 \quad H_5\,]$$

then $\bar{H}$ can be partitioned as Eq. (10) with

$$\mathcal{H}' = [\,H_1 \quad H_2 \quad H_4\,], \qquad U'(k-1) = \begin{bmatrix} \hat{u}(k-1) \\ \hat{u}(k-2) \\ \hat{u}(k-4) \end{bmatrix}$$

and

$$\mathcal{H} = [\,H_3 \quad H_5\,], \qquad \hat{U}(k-1) = \begin{bmatrix} \hat{u}(k-3) \\ \hat{u}(k-5) \end{bmatrix}$$

The partitioning of Eq. (10) provides flexibility in implementing RCAISE, in that only selected columns of $\bar{H}$ and the components of $\hat{\tilde{U}}(k-1)$ need be used.

Next, using Eq. (10), we can rewrite Eq. (9) as

$$z(k) = \mathcal{S}(k) + \mathcal{H}\hat{U}(k-1) \tag{11}$$

where

$$\mathcal{S}(k) = CA^r\hat{x}(k-r) - y(k) + \mathcal{H}'\hat{U}'(k-1) \tag{12}$$

Next, for $j = 1, \ldots, s$, we rewrite Eq. (11) with a delay of k_j time steps, where $0 \le k_1 \le k_2 \le \cdots \le k_s$, and we replace $\mathcal{H}$ by $\mathcal{H}_j \in \mathbb{R}^{p \times l}$ to obtain

$$z_j(k-k_j) = \mathcal{S}_j(k-k_j) + \mathcal{H}_j\hat{U}_j(k-k_j-1) \tag{13}$$

where Eq. (12) becomes

$$\mathcal{S}_j(k-k_j) = CA^r x(k-k_j-r) - y(k-k_j) + \mathcal{H}'_j\hat{U}'_j(k-k_j-1)$$

and Eq. (10) becomes

$$\bar{H}\hat{\tilde{U}}(k-k_j-1) = \mathcal{H}'_j\hat{U}'_j(k-k_j-1) + \mathcal{H}_j\hat{U}_j(k-k_j-1) \tag{14}$$

where $\mathcal{H}'_j \in \mathbb{R}^{p \times (rm-l_j)}$, $\hat{U}'_j(k-k_j-1) \in \mathbb{R}^{rm-l_j}$, and $\hat{U}_j(k-k_j-1) \in \mathbb{R}^{l_j}$. The subscript j in $\mathcal{H}'_j$, $\mathcal{H}_j$, $\hat{U}'_j$, and $\hat{U}_j$ indicates that the partitioning of $\bar{H}\tilde{U}(k-k_j-1)$ in Eq. (14) can be different for each delay. Now, by stacking $z(k-k_1), \ldots, z(k-k_s)$, we define the extended performance

$$Z(k) = \begin{bmatrix} z_1(k-k_1) \\ \vdots \\ z_j(k-k_s) \end{bmatrix} \in \mathbb{R}^{sp} \tag{15}$$

Therefore,

$$Z(k) = \tilde{\mathcal{S}}(k) + \tilde{\mathcal{H}}\hat{\tilde{U}}(k-1) \tag{16}$$

where

$$\tilde{\mathcal{S}}(k) = \begin{bmatrix} \mathcal{S}_1(k-k_1) \\ \vdots \\ \mathcal{S}_s(k-k_s) \end{bmatrix} \in \mathbb{R}^{sp}, \qquad \hat{\tilde{U}}(k-1) = \begin{bmatrix} \hat{u}(k-q_1) \\ \vdots \\ \hat{u}(k-q_g) \end{bmatrix} \in \mathbb{R}^{mg} \tag{17}$$

where, for $i = 1, \ldots, g, k_1 \le q_i \le k_s + r$ and $\tilde{\mathcal{H}} \in \mathbb{R}^{sp \times mg}$ is constructed according to the structure of $\hat{\tilde{U}}(k-1)$ The vector $\hat{\tilde{U}}(k-1)$ is formed by stacking $\hat{U}_1(k-k_1-1), \ldots, \hat{U}_s(k-k_s-1)$ and removing copies of repeated components.

For example, with $k_1 = 0$ and $k_2 = 1$, stacking $\hat{U}_1(k-1) = [\hat{u}(k-1) \quad \hat{u}(k-2)]^T$ and $\hat{U}_2(k-2) = \hat{u}(k-2)$ results in $\hat{\tilde{U}}(k-1) = [\hat{u}(k-1) \quad \hat{u}(k-2)]^T$. The coefficient matrix $\tilde{\mathcal{H}}$ consists of the entries of $\mathcal{H}_1, \ldots, \mathcal{H}_s$ arranged according to the structure of $\hat{\tilde{U}}(k-1)$.

Next, we define the retrospective performance

$$\hat{z}_j(k-k_j) = \mathcal{S}_j(k-k_j) + \mathcal{H}_j U^*_j(k-k_j-1) \tag{18}$$

where the past drivers $\hat{U}_j(k-k_j-1)$ in Eq. (13) are replaced by the drivers $U^*_j(k-k_j-1)$, which are obtained next using retrospective-cost optimization. In analogy with Eq. (15), the extended retrospective performance for Eq. (18) is defined as

$$\hat{Z}(k) = \begin{bmatrix} \hat{z}_1(k-k_1) \\ \vdots \\ \hat{z}_s(k-k_s) \end{bmatrix} \in \mathbb{R}^{sp}$$

and thus is given by

$$\hat{Z}(k) = \tilde{\mathcal{S}}(k) + \tilde{\mathcal{H}}\tilde{U}^*(k-1) \tag{19}$$

where the components of $\tilde{U}^*(k-1) \in \mathbb{R}^g$ are the components of $U^*_1(k-k_1-1), \ldots, U^*_s(k-k_s-1)$ ordered in the same way as the components of $\tilde{U}(k-1)$. Subtracting Eq. (16) from Eq. (19) gives

$$\hat{Z}(k) = Z(k) - \tilde{\mathcal{H}}\hat{\tilde{U}}(k-1) + \tilde{\mathcal{H}}\tilde{U}^*(k-1) \tag{20}$$

where, in analogy with Eq. (17),

$$\tilde{U}^*(k-1) = \begin{bmatrix} u^*_{k-1}(k-q_1) \\ \vdots \\ u^*_{k-1}(k-q_g) \end{bmatrix} \in \mathbb{R}^{mg}$$

Finally, we define the retrospective-cost function

$$J(\tilde{U}^*(k-1), k) = \hat{Z}^T(k)R(k)\hat{Z}(k) \tag{21}$$

where $R(k) \in \mathbb{R}^{ps \times ps}$ is a positive-definite performance weighting. The goal is to determine retrospectively optimized drivers $\tilde{U}^*(k-1)$ that would have provided better performance than the past driver estimates $\tilde{U}(k-1)$ that were applied to the adaptive state estimator.

A. Cost Function Optimization with Adaptive Regularization

To ensure that Eq. (21) has a global minimizer, we consider the regularized cost

$$\bar{J}(\tilde{U}^*(k-1), k) = \hat{Z}^T(k)R(k)\hat{Z}(k) + \eta(k)\tilde{U}^{*T}(k-1)\tilde{U}^*(k-1) \tag{22}$$

where $\eta(k) \geq 0$. Substituting Eq. (20) into Eq. (22) gives

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

where

$$\mathcal{A}(k) = \tilde{\mathcal{H}}^T R(k)\tilde{\mathcal{H}} + \eta(k)I_g,$$
$$\mathcal{B}(k) = 2\tilde{\mathcal{H}}^T R(k)[Z(k) - \tilde{\mathcal{H}}\hat{\bar{U}}(k-1)],$$
$$\mathcal{C}(k) = Z^T(k)R(k)Z(k) - 2Z^T(k)R(k)\tilde{\mathcal{H}}\hat{\bar{U}}(k-1) + \hat{\bar{U}}^T(k-1)\tilde{\mathcal{H}}^T R(k)\tilde{\mathcal{H}}\hat{\bar{U}}(k-1)$$

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

$$\tilde{U}^*(k-1) = -\frac{1}{2}\mathcal{A}^{-1}(k)\mathcal{B}(k) \tag{23}$$

The optimized drivers given by Eq. (23) are subsequently used to update the adaptive driver-estimator subsystem [Eq. (6)]. Note that $\tilde{\mathcal{H}}$ is the only modeling information required by RCAISE.

B. Adaptive Driver-Estimator Subsystem Update

The adaptive driver-estimator subsystem model [Eq. (6)] can be expressed as

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

where the driver-estimator coefficients are

$$\theta(k) = [M_1(k) \cdots M_{n_e}(k)N_0(k) \cdots N_{n_e}(k)] \in \mathbb{R}^{m \times (2n_e m + (2n_e+1)p)}$$

and the regressor vector is

$$\phi(k-1) = \begin{bmatrix} \hat{u}(k-1) \\ \vdots \\ \hat{u}(k-n_e) \\ z(k) \\ \vdots \\ z(k-n_e) \end{bmatrix} \in \mathbb{R}^{2n_e m + (2n_e+1)p}$$

Next, for all $i \in \{1, \ldots, k\}$, let d_i be a positive integer such that $\tilde{U}^*(k-1)$ contains $u^*_{k-1}(k-d_i)$. Then, we define the cumulative cost function

$$J_R(\theta(k)) = \sum_{i=1}^{k} \|\theta(k)\phi(i-d_i-1) - u^*_{k-i}(i-d_i)\|^2 \tag{24}$$

where $\| \cdot \|$ is the Euclidean norm. By minimizing Eq. (24), we update the gains $\theta(k)$ using the retrospectively optimized drivers $u^*_{k-i}(i-d_i)$. The recursive minimizer of Eq. (24) is given by the standard recursive least-squares (RLS) algorithm

$$P(k) = \lambda^{-1}P(k-1) - \lambda^{-1}P(k-1)\phi(k-d_k-1)$$
$$\times [\phi^T(k-d_k-1)P(k-1)\phi(k-d_k-1) + \lambda]^{-1} \times \phi^T(k-d_k-1)P(k-1) \tag{25}$$

$$\theta^T(k) = \theta^T(k-1) + P(k)\phi(k-d_k-1)[\phi^T(k-d_k-1)\theta^T(k-1) - u^{*T}(k-d_k)] \qquad (26)$$

with the initial conditions $\theta^T(0) \in \mathbb{R}^{[n_c(m+p)+1]\times m}$ and $P(0) = \gamma I$, where $\gamma > 0$. For details on RLS, see [28].

C. Computational Complexity

RCAISE is an ensemble-free algorithm that involves the minimization of two quadratic cost functions. For implementation, RCAISE requires the computation of Eqs. (6), (23), (25), and (26). The computational complexity of the matrix product in Eq. (6) is $\mathcal{O}(m[n_c m + (n_c + 1)p])$, and the computational complexity of the matrix products in Eqs. (25) and (26) is $\mathcal{O}((n_c m + (n_c + 1)p)^2)$. Finally, the computational complexity of the matrix products and inverse in Eq. (23) is $\max(\mathcal{O}((mg)^2 sp), \mathcal{O}((sp)^2 mg))$ and $\mathcal{O}((mg)^3)$, respectively. In this paper, we consider the case where $m = 1$, $p = 1$, $g = 1$, and $s = 1$, for which the computational complexity of RCAISE is $\mathcal{O}(n_c^2)$.

An implementation of RCAISE on a four-processor GITM simulation with low resolution (16,200 grid points) from 24 November 2002 to 28 December 2002 with $n_c = 100$ completes in 4 h, 3 min of which are devoted to RCAISE equations (6), (23), (25), and (26). Consequently, RCAISE adds about 1% to the cost of a single GITM simulation.

III.   Illustrative Examples

In the previous section, RCAISE was derived for a system with a linear dynamics model, and implementation of this algorithm requires the matrix $\tilde{\mathcal{H}}$. The matrix $\tilde{\mathcal{H}}$ can be constructed from the impulse response of the model or it can be determined by numerical testing. In this section, we demonstrate RCAISE with nonlinear models. For these models, we construct $\tilde{\mathcal{H}}$ by numerical testing.

Example III.1 (Lorenz system): The three-variable Lorenz system is known to have chaotic solutions for certain parameter values and initial conditions. This system does not have external drivers. A fictitious input, retrospectively optimized by RCAISE, is applied to the estimator in order to drive the estimator states toward the Lorenz system states.

Consider the discretized Lorenz equations

$$x_1(k+1) = x_1(k) + T_s\sigma(x_2(k) - x_1(k)),$$
$$x_2(k+1) = x_2(k) + T_s(x_1(k)(\rho - x_3(k)) - x_2(k)),$$
$$x_3(k+1) = x_3(k) + T_s(x_1(k)x_2(k) - \beta x_3(k)),$$
$$y(k) = x_1(k) + x_2(k)$$

where $T_s = 0.001$, $\sigma = 10$, $\beta = 8/3$, and $\rho = 28$, and the initial conditions $x_1(0), x_2(0), x_3(0)$ are chosen randomly from a normal distribution. For this choice of parameters, the solution of this system is extremely sensitive to the initial conditions. The only measurement is the scalar output y.

The adaptive estimator has the form

$$\hat{x}_1(k+1) = \hat{x}_1(k) + T_s\sigma(\hat{x}_2(k) - \hat{x}_1(k)) + \hat{u}(k),$$
$$\hat{x}_2(k+1) = \hat{x}_2(k) + T_s(\hat{x}_1(k)(\rho - \hat{x}_3(k)) - \hat{x}_2(k)),$$
$$\hat{x}_3(k+1) = \hat{x}_3(k) + T_s(\hat{x}_1(k)\hat{x}_2(k) - \beta\hat{x}_3(k)),$$
$$\hat{y}(k) = \hat{x}_1(k) + \hat{x}_2(k)$$

where $\hat{u}(k)$ is a fictitious input to the estimator, which is determined by RCAISE to drive the estimator states to true states. We use $\tilde{\mathcal{H}} = H_2 = 10,000$, $R(k) = 1$, $\eta = 0$, $n_c = 4$, and $\gamma = 100$, where H_2 is obtained by numerical testing, and the estimator is initialized with $\hat{x}_1(0) = \hat{x}_2(0) = \hat{x}_3(0) = 0$. Figure 2 shows that, after RCAISE is turned on at $k = 100$, the estimated states are driven to the true states of the Lorenz system, and $z(k)$ is asymptotically driven to zero.

Example III.2 (Van der Pol oscillator): We consider the Van der Pol oscillator driven by a harmonic input. Consider the nonlinear discretized Van der Pol oscillator equations in state-space form

$$x_1(k+1) = x_1(k) + T_s x_2(k),$$
$$x_2(k+1) = x_2(k) + T_s(1 - x_1(k)^2)x_2(k) - T_s x_1(k) + T_s u(k),$$
$$y(k) = x_1(k) + 0.2x_2(k)$$

where $T_s = 0.1$ s, and the unknown input is $u(k) = \sin(0.01k)$. The objective is to use the measurement y to estimate u.

The adaptive estimator has the form

$$\hat{x}_1(k+1) = \hat{x}_1(k) + T_s\hat{x}_2(k),$$
$$\hat{x}_2(k+1) = \hat{x}_2(k) + T_s(1 - \hat{x}_1(k)^2)\hat{x}_2(k) - T_s\hat{x}_1(k) + T_s\hat{u}(k),$$
$$\hat{y}(k) = \hat{x}_1(k) + 0.2\hat{x}_2(k)$$

where $\hat{u}(k)$ is the estimate of $u(k)$ obtained from RCAISE. We use $R(k) \equiv 1$, $\tilde{\mathcal{H}} = 0.01$, $\eta = 0.001$, $n_c = 2$, and $\gamma = 200$, where $\tilde{\mathcal{H}}$ is chosen by numerical testing. Figure 3 shows that, after RCAISE is turned on at $k = 80$, the retrospectively optimized input converges to the unknown driver, and the estimated states are driven to the true state of the Van der Pol oscillator.

Next, to study the sensitivity of RCAISE to $\tilde{\mathcal{H}}$, RCAISE is run with three separate values of H_1, which are chosen to span one order of magnitude. Figure 4 shows how the choice of H_1 affects the transient response and convergence time. By observing these responses, H_1 can be chosen to adjust RCAISE to obtain the desired response.
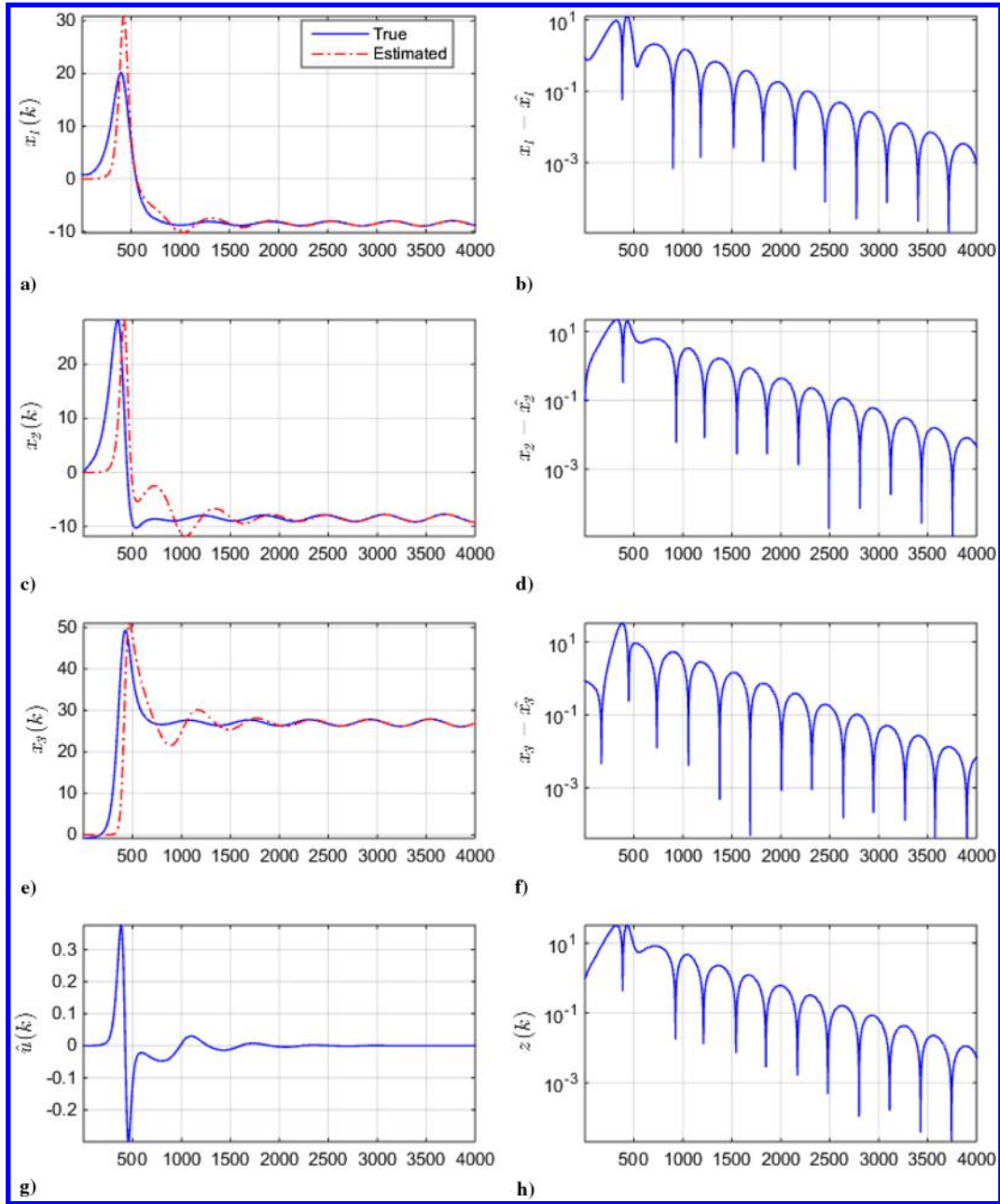
Fig. 2 True and estimated states for the three-variable Lorenz system: a, c, e) estimated states of the Lorenz system; b, d, f) error in the estimated states; g) fictitious driver; and h) estimator output error $z(k)$. The initial state of the Lorenz system is chosen randomly from a Gaussian distribution, whreeas the initial state of the estimator is set to zero. RCAISE is switched on after $k = 100$. To estimate the states, RCAISE adds a fictitious driver $\hat{u}(k)$ to the model, which converges to zero as the states converge.

Example III.3 (Van der Pol oscillator with matched, unmodeled dynamics): Consider the modified Van der Pol oscillator

$$x_1(k+1) = x_1(k) + T_s x_2(k),$$
$$x_2(k+1) = x_2(k) + T_s(1 - x_1(k)^2)x_2(k) - T_s x_1(k) + \sin(x_2(k)) + T_s u(k),$$
$$y(k) = x_1(k) + 0.2x_2(k)$$

where $\sin(x_2(k))$ is an unmodeled feature of the dynamics, $T_s = 0.1$ s, and the driver $u(k) = \sin(0.01k)$ is unknown. Note that the unmodeled term $\sin(x_2(k))$ is matched to the unknown driver in the sense that there exists a driver signal that can account for the presence of this term. In particular, replacing $u(k)$ by $u(k) - \sin(x_2(k))/T_s$ removes the unmodeled term. Since the term $\sin(x_2(k))$ is unmodeled, it is not included in the estimator, which has the form

$$\hat{x}_1(k+1) = \hat{x}_1(k) + T_s \hat{x}_2(k),$$
$$\hat{x}_2(k+1) = \hat{x}_2(k) + T_s(1 - \hat{x}_1(k)^2)\hat{x}_2(k) - T_s \hat{x}_1(k) + T_s \hat{u}(k),$$
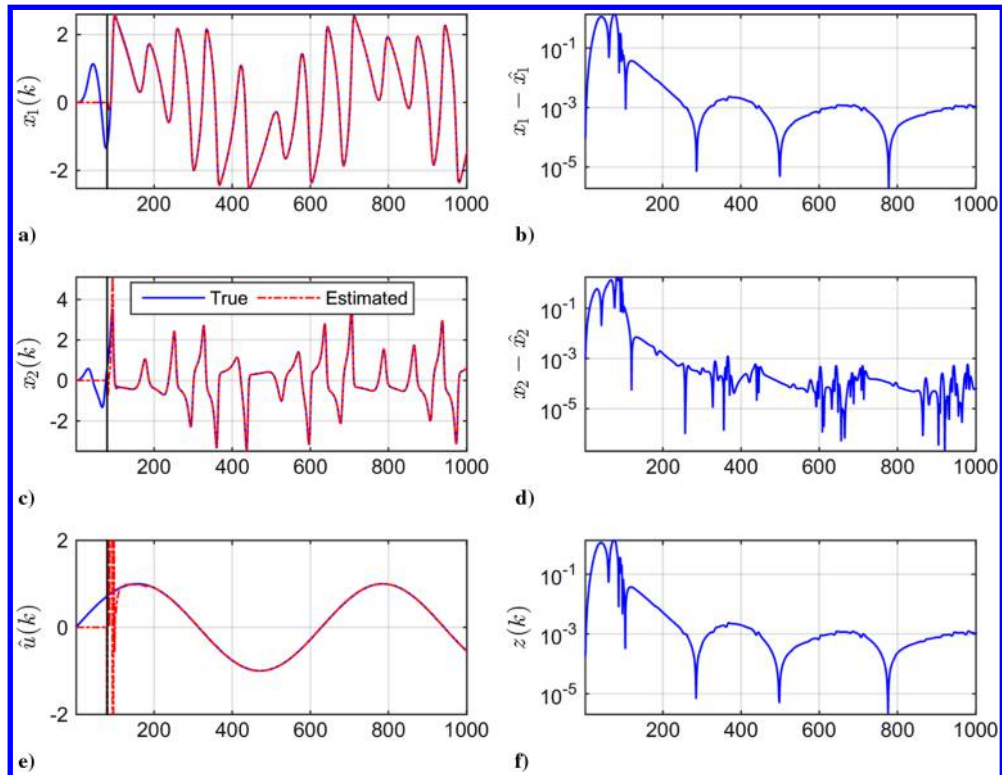$$\hat{y}(k) = \hat{x}_1(k) + 0.2\hat{x}_2(k)$$

Fig. 3 True and estimated states $x_1(k)$ and $x_2(k)$, and driver $u(k)$ for the Van der Pol oscillator: a, c) estimated states; b, d) error in the estimated states, e) true and estimated drivers; and f) estimator output error $z(k)$. The vertical line indicates the time step at which RCAISE is switched on. After $k = 80$, RCAISE reduces the error in $\hat{x}_1(k)$, $\hat{x}_2(k)$, and $\hat{u}(k)$ by approximately three orders of magnitude.

Let $R(k) \equiv 1$, $\tilde{\mathcal{H}} = 0.02$, $\eta = 0.0008$, $n_c = 4$, and $\gamma = 1000$. Figure 5 shows the true and estimated states and driver. Figure 5 shows that, after RCAISE is turned on at $k = 80$, RCAISE reduces the error in $\hat{x}_1(k)$ and $\hat{x}_2(k)$ by one order of magnitude. For this example, the estimated driver does not converge to the true driver due to the matched, unmodeled term $\sin(x_2(k))$ in the dynamics of the oscillator. In fact, the estimated driver
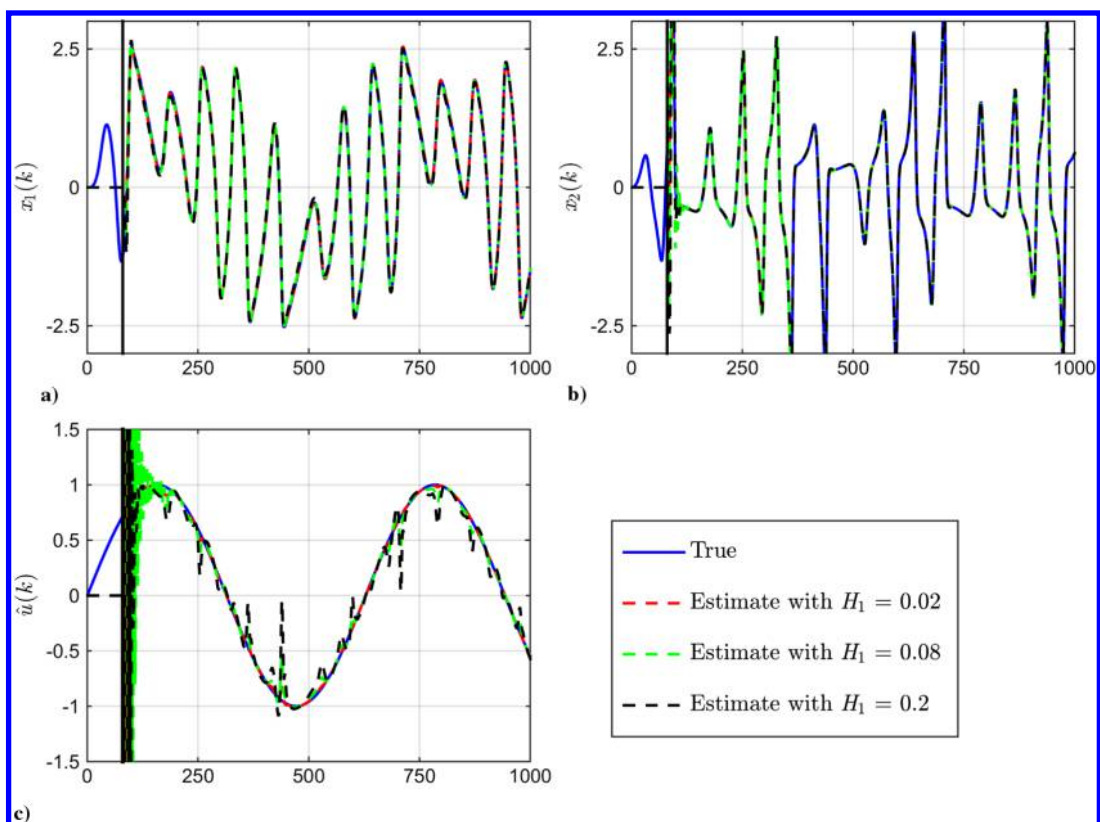

Fig. 4 True and estimated states $x_1(k)$ and $x_2(k)$, and driver $u(k)$ for the Van der Pol oscillator: a, b) estimated states; and c) true and estimated drivers. The vertical line indicates the time step at which RCAISE is switched on. RCAISE is run with three separate values of H_1, which are chosen to span one order of magnitude to illustrate the sensitivity of the estimate of the unknown input u to the choice of H_1. These plots show how the choice of H_1 affects the transient response and convergence time. By observing these responses, H_1 can be chosen to adjust RCAISE to obtain the desired response.
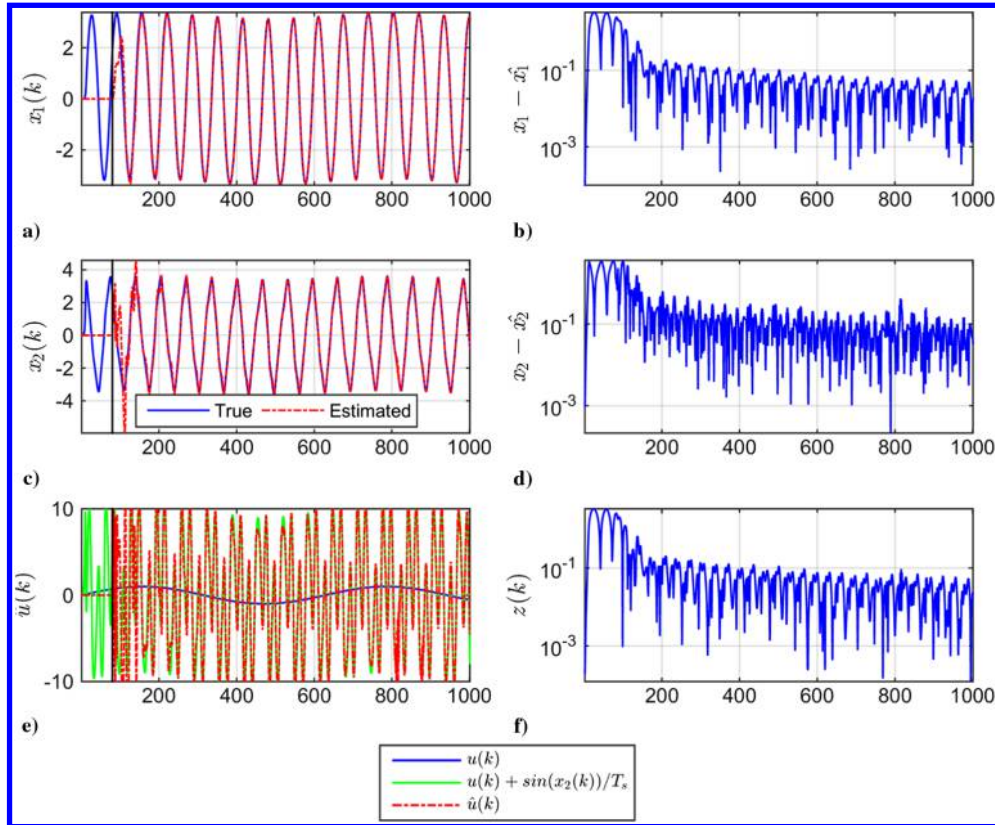
Fig. 5   True and estimated states $x_1(k)$ and $x_2(k)$, and driver $u(k)$ for the Van der Pol oscillator with matched, unmodeled dynamics: a, c) estimated states; b, d) error in the estimated states; e) true and estimated drivers; and f) estimator output error $z(k)$. The vertical line indicates the time step at which RCAISE is switched on. After $k = 80$, RCAISE reduces the errors in $\hat{x}_1(k)$ and $\hat{x}_2(k)$ by approximately one order of magnitude. For this example, the estimated driver does not converge to the true driver due to the unknown term $\sin(x_2(k))$ in the dynamics of the oscillator. However, $\hat{u}(k)$ does give an estimate of $u(k) + \sin(x_2(k))/T_s$, which shows that RCAISE reconstructs not only the unknown driver but also the matched, unmodeled feature of the dynamics.

$\hat{u}(k)$ converges to $u(k) + \sin(x_2(k))/T_s$, which shows that RCAISE reconstructs not only the unknown driver but also the matched, unmodeled feature of the dynamics.

Example III.4 (Van der Pol oscillator with unmatched, unmodeled dynamics): Consider the modified Van der Pol oscillator

$$x_1(k + 1) = x_1(k) + T_s x_2(k) + 0.1 \sin(x_2(k)),$$
$$x_2(k + 1) = x_2(k) + T_s(1 - x_1(k)^2)x_2(k) - T_s x_1(k) + T_s u(k),$$
$$y(k) = x_1(k) + 0.2x_2(k)$$

where $0.1 \sin(x_2(k))$ is an unmodeled feature of the dynamics, $T_s = 0.1$ s, and the driver $u(k) = \sin(0.01k)$ is unknown. Unlike Example III.3, however, the unmodeled term $0.1 \sin(x_2(k))$ is not matched to the unknown driver, and thus there does not exist a driver estimate that can exactly account for the presence of this term. Since the term $0.1 \sin(x_2(k))$ is unmodeled, the estimator has the form

$$\hat{x}_1(k + 1) = \hat{x}_1(k) + T_s \hat{x}_2(k),$$
$$\hat{x}_2(k + 1) = \hat{x}_2(k) + T_s(1 - \hat{x}_1(k)^2)\hat{x}_2(k) - T_s \hat{x}_1(k) + T_s \hat{u}(k),$$
$$\hat{y}(k) = \hat{x}_1(k) + 0.2\hat{x}_2(k)$$

Let $R(k) \equiv 1$, $\tilde{\mathcal{H}} = 0.005$, $\eta = 0.0008$, $n_c = 4$, and $\gamma = 1000$. Figure 6 shows that there is a persistent error in the estimated states and that RCAISE cannot estimate $u(k)$ due to the unmatched, unmodeled feature of the dynamics. This example shows that RCAISE requires knowledge of the correct model structure in order to achieve accurate input estimation.

IV.   GITM and Satellites

GITM solves the Navier–Stokes equations for the upper atmosphere in a spherical coordinate system. For the ionosphere, the continuity equation is solved, whereas the momentum and energy equations are simplified using the assumption of steady state. This approach is effective in the ionosphere, where the timescales for changes in the charged particle velocities and temperatures are extremely small. Frictional terms couple the species and capture eddy diffusion in the lower thermosphere. In the upper thermosphere, all of the species separate out and reach a roughly hydrostatic balance.

GITM is fully described in [10] and has been used to study various geophysical phenomena, such as nonhydrostatic wind structures [29], circulation in the ionosphere [29], and the influence of ion variability and grid resolution on heating in the thermosphere [30].
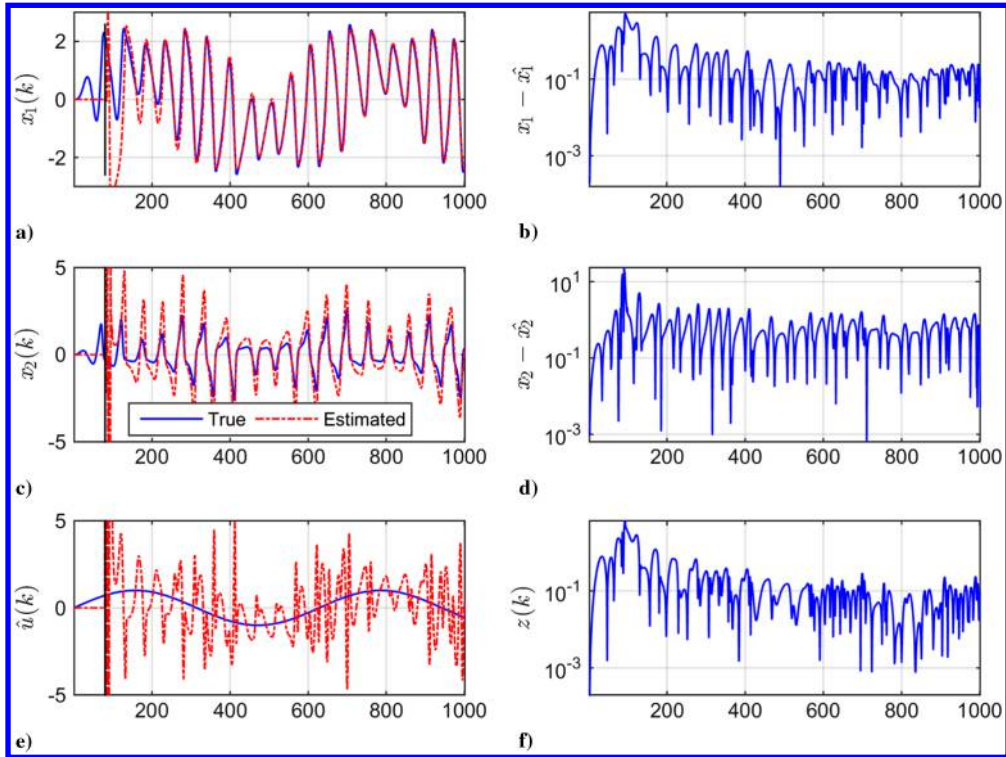
Fig. 6   True and estimated states $x_1(k)$ and $x_2(k)$, and driver $u(k)$ for the Van der Pol oscillator with unmatched, unmodeled dynamics: a, c) estimated states; b, d) error in the estimated states; e) true and estimated drivers; and f) estimator output error $z(k)$. The estimated driver does not converge to the true driver because of the unmatched, unmodeled term $0.1 \sin(x_2(k))$ in the dynamics of the oscillator, which cannot be estimated by the adaptive estimator. Consequently, errors in the estimates of $x_1(k)$ and $x_2(k)$ are persistent.

The initial state of the upper atmosphere, which includes the temperature and densities of the main species, is determined at the subsolar point from mass spectrometer and incoherent scatter radar (MSIS) [31–33] and the International Reference Ionosphere [34]. These quantities are then used to compute the scale height of the atmosphere as a function of altitude. The grid spacing is given as one-third of the initial subsolar scale height. In addition to 50 altitudes, two ghost cells are included at the lower and upper boundaries to specify the boundary conditions. The bottommost modeled cell is at a 100 km altitude, whereas the topmost cell can be as low as 450 km during solar minimum conditions or a 750 km altitude during solar maximum conditions.

In the vertical direction, the momentum is solved for all of the main species that, in Earth's upper atmosphere, are O, O_2, N_2, N, and NO. These are treated as five separate fluids, with a coupling frictional term as a source term for each of them. This frictional term forces the fluids to move with almost exactly the same speed below about 120–130 km altitude, and it is considered to be caused by eddy diffusion within the model. The remaining species, which consist of excited states of N and O, are assumed to be in chemical equilibrium.

The drivers of GITM include the solar luminosity (as described later in this study), tides at the lower boundary specified by the MSIS empirical atmosphere model [31–33], the horizontal wind model [35], the high-latitude electric field specified in [36], the aurora specified in [37], and the magnetic field based on the International Geophysical Reference Field model.

The grid structure within GITM is fully parallel and covers the entire planetary surface by using a block-based two-dimensional domain decomposition in the horizontal coordinates [29]. The number of latitude and longitude blocks can be specified at runtime 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 [10]. This flexibility can be used to validate consistency by running the estimator at various levels of resolution.

In summary, GITM can be implemented on multiple processors with a nonuniform grid, using various models of electric fields, magnetic fields, auroral particle precipitation, solar EUV and x-ray drivers, and particle energy deposition.

In this paper, we use neutral density data from CHAMP and GRACE satellites. CHAMP data are used as measurements to obtain state and driver estimates, whereas GRACE data are used as a metric for assessing the accuracy of state estimates. Figure 7 shows the trajectories of CHAMP and GRACE from 02:12:00 to 03:40:00 Coordinated Universal Time on 21 November 2002. Note that, since only CHAMP is used to obtain measurements for driver estimation, this satellite may at certain times be located where the solar zenith angle is greater than 90 deg. When this occurs, the effect of $F_{10.7}$ on the estimator output is significantly delayed [38]. This delay can be mitigated by using multiple satellites such that the solar zenith angle at the location of at least one satellite is less than 90 deg. However, in this paper, we use only one satellite.

V.   RCAISE Formulation for GITM

We now implement RCAISE with GITM. Consider the ionosphere–thermosphere system represented by

$$\hat{x}(k + 1) = \hat{f}(\hat{x}(k), v(k), \hat{u}(k)) \tag{27}$$

$$\hat{y}(k) = \hat{h}(\hat{x}(k)) \tag{28}$$

where $\hat{x}(k) \in \mathbb{R}^{31 n_g}$ is the estimated state; n_g is the number of grid points; the number of states per grid point is 31; $\hat{y}(k) \in \mathbb{R}^p$ represents the estimator outputs at the satellite locations, where p is the number of satellites; and $\hat{u}(k) \in \mathbb{R}^{37}$ is the estimated driver (that is, the EUV and x-ray
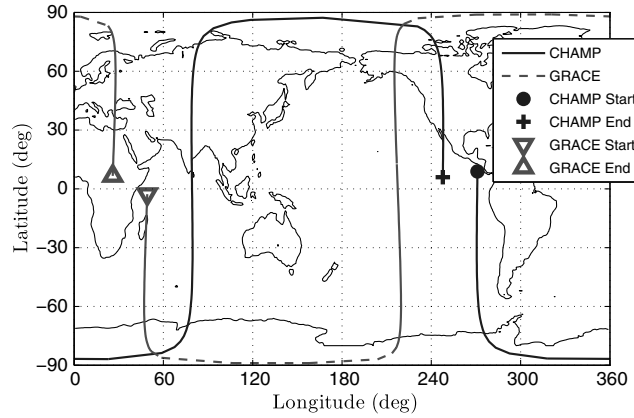
Fig. 7   CHAMP and GRACE trajectories on 21 November 2002.

flux in 37 wavelength bins). Within this formulation, GITM uses the extreme ultraviolet flux model for aeronomic calculations (EUVAC) model of the solar EUV irradiance given in [39] to provide an estimate of $u(k)$, which is given by

$$\hat{u}(k) = \hat{\bar{\mathrm{F}}}_{10.7}(k)E + Q \tag{29}$$

where

$$E \in \mathbb{R}^{37}, \qquad Q \in \mathbb{R}^{37}$$

and

$$\hat{\bar{\mathrm{F}}}_{10.7}(k) = \frac{1}{2}(\hat{F}_{10.7}(k) + \hat{F}_{10.7a}(k)) \tag{30}$$

where $\hat{F}_{10.7}(k) \in \mathbb{R}$ is the estimated solar irradiance at the wavelength of 10.7 cm, and $\hat{F}_{10.7a}(k) \in \mathbb{R}$ is an 81-day average of $\hat{F}_{10.7}(k)$. In other words, GITM uses the EUVAC model to estimate $u(k)$ using $\bar{\mathrm{F}}_{10.7}(k)$. The matrices E and Q were given in table 1 of [39]. Finally, we define the error $z(k) = y(k) - \hat{y}(k)$, where $y(k)$ is the neutral density measured by CHAMP and $\hat{y}(k)$ is the neutral density estimated by RCAISE.

Our goal is to estimate $x(k)$ by determining a driver signal $\bar{\mathrm{F}}_{10.7}(k)$ that minimizes J [22]. We implement RCAISE on GITM with $\tilde{\mathcal{H}} = H_1$. To select a suitable value of H_1, we use the following modifications of RCAISE. We use a low-pass filter that limits the rate of change of the output $\hat{\bar{\mathrm{F}}}_{10.7}(k)$ from the adaptive driver-estimator subsystem [Eq. (6)]. This prevents Eq. (6) from converging to a system that gives a highly oscillatory $\hat{\bar{\mathrm{F}}}_{10.7}(k)$. Second, we constrain the maximum and minimum values of the output from Eq. (6) in order to prevent GITM from yielding nonphysical states. We set these constraints to $70 \le \hat{\bar{\mathrm{F}}}_{10.7}(k) \le 400$.

This approach also accounts for unknown dynamics and measurement noise. In particular, if $z(k)$ is nonzero, RCAISE attempts to minimize it by changing $\hat{\bar{\mathrm{F}}}_{10.7}(k)$. However, since the effect of a change in $\hat{\bar{\mathrm{F}}}_{10.7}(k)$ on the output $\hat{y}(k)$ is slow, and since $\hat{\bar{\mathrm{F}}}_{10.7}(k)$ alone cannot counteract the effect of all other drivers in GITM, RCAISE gives an oscillatory $\hat{\bar{\mathrm{F}}}_{10.7}(k)$ in an attempt to minimize $z(k)$. For some values of H_1, this can cause $\hat{\bar{\mathrm{F}}}_{10.7}(k)$ to alternate between the saturation limits. To prevent this, we apply a low-pass filter to the signal $z(k)$. A schematic for estimating states and driver in the ionosphere–thermosphere using GITM and RCAISE is shown in Fig. 8. Note that, in Fig. 8, the input to Eq. (6) is $z_f(k)$ and the output from Eq. (6) is $\hat{\bar{\mathrm{F}}}_{10.7,f}(k)$.

The frequency of computing the RCAISE update equations and updating $\hat{\bar{\mathrm{F}}}_{10.7}(k)$ depends on the frequency of available measurements. For the purposes of this paper, GITM uses the fixed time step $T_s = 2$ s, whereas measurements are used to update $\hat{\bar{\mathrm{F}}}_{10.7}(k)$ only every 60 s.

Due to the low computational burden of RCAISE compared to GITM, we implement RCAISE on one processor and send all information needed by RCAISE from other processors to this processor. A flowchart summarizing the implementation of RCAISE with GITM on four processors is shown in Fig. 9.

VI.   State and Driver Estimation Results for GITM

We divide the RCAISE implementation into two main setups. The first setup uses simulated satellite data, whereas the second setup uses real satellite data. In the first setup, we run GITM with a specified "true" $\bar{F}_{10.7}(k)$ and store neutral mass density at CHAMP and GRACE satellite locations, which we call the "truth data." Neutral density at CHAMP's location is donated by $y_C(k)$ and at GRACE's location is donated by $y_G(k)$, whereas the estimated neutral densities at CHAMP's and GRACE's locations are denote by $\hat{y}_C(k)$ and $\hat{y}_G(k)$. Note that we use $y_G(k)$ and $\hat{y}_G(k)$ as a metric for assessing the accuracy of the state estimates, and do not use it in RCAISE. Next, we combine RCAISE with GITM and use $y_C(k)$ to estimate $\bar{F}_{10.7}(k)$ and states.

The second setup involves real satellite data. Similar to the previous setup, $y_C(k)$ denotes the neutral mass density measured by CHAMP (the truth data), whereas $y_G(k)$ denotes the neutral mass density measured by GRACE. GITM is first simulated with the ground observations of $\bar{F}_{10.7}(k)$, yielding the neutral mass density at CHAMP and GRACE locations, denoted by $y_{C,\mathrm{obs}}(k)$ and $y_{G,\mathrm{obs}}(k)$, respectively. Next, RCAISE and GITM use $y_C(k)$ to estimate both states and $\bar{F}_{10.7}(k)$. The estimated neutral mass densities from GITM with RCAISE at CHAMP and GRACE locations are denoted by $\hat{y}_C(k)$ and $\hat{y}_G(k)$. Note that GRACE data are used solely as a metric for assessing state estimate accuracy; RCAISE has no access to these data.

$u(k)$: Solar EUV and x-ray radiation
$v(k)$: Electrodynamic drivers (known)
$y(k)$: Satellite measurements
$\hat{u}(k)$: Estimated solar EUV and x-ray radiation
$\hat{y}(k)$: GITM output at satellite location
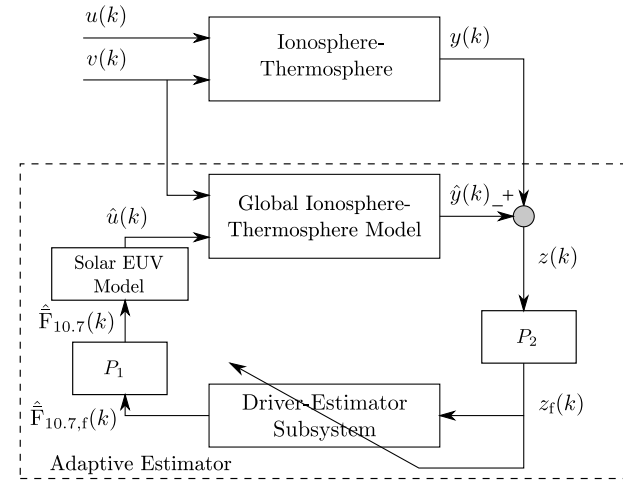P_1, P_2 : Low-pass filters

Fig. 8 RCAISE setup for GITM. The known driver $v(k)$ is used in GITM, whereas the unknown driver $u(k)$ must be estimated. The objective is to construct an estimate of $\bar{\hat{F}}_{10.7}(k)$ that, when used in the EUVAC model, minimizes J given by Eq. (22).

To implement RCAISE, $\bar{\hat{F}}_{10.7}(k)$ is set to a constant value of 100 SFUs for the first 24 h. RCAISE is switched on after 24 h. This allows the initial condition response to decay.

We implement GITM on four processors with a resolution of 5° latitude and 5° longitude. The time step for GITM is set at 2 s, and RCAISE updates $\bar{\hat{F}}_{10.7}(k)$ every 60 s. In all figures in this section, RCAISE is switched on after 24 h. All numerical experiments consider the period from 21 November 2002 to 6 December 2002, during which real data are available.
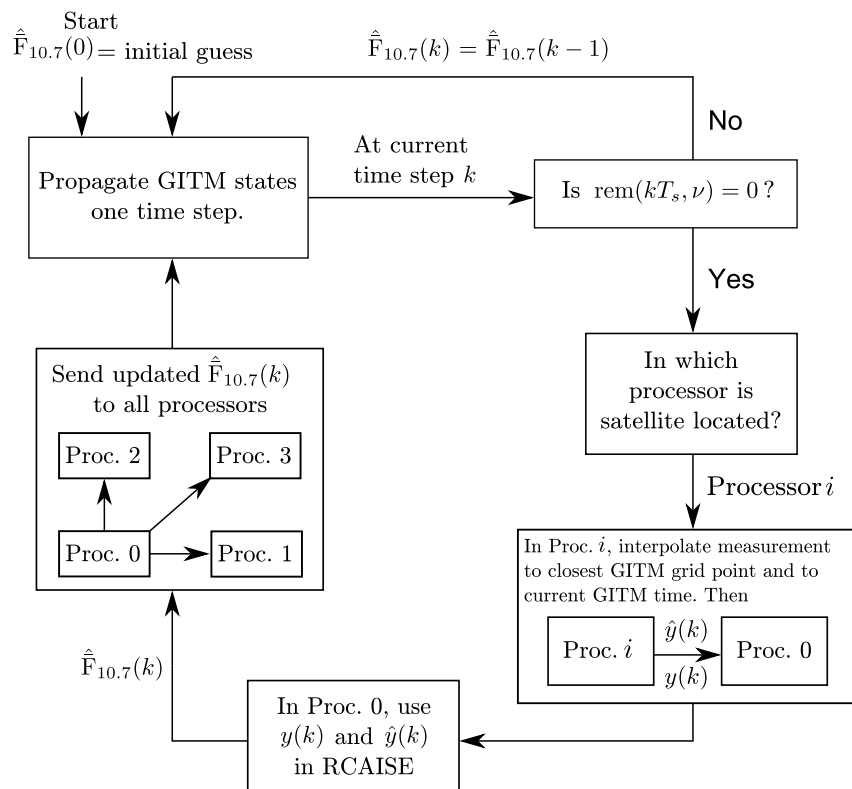

Fig. 9 Implementation of RCAISE on a four-processor GITM simulation. Each processor simulates the ionosphere–thermosphere dynamics in one quadrant of the Earth's atmosphere. If, at the current time step k, $\mathrm{rem}(kT_s, \nu) = 0$, then the processor that simulates the quadrant of the atmosphere in which the satellite is located uses satellite data to obtain $y(k)$ and uses $\hat{x}(k)$ to obtain $\hat{y}(k)$. These estimates are then sent to processor (proc.) 0, where RCAISE uses them to obtain $\bar{\hat{F}}_{10.7}(k)$, which is then sent to all processors to allow GITM to propagate states to the next time step. If $\mathrm{rem}(kT_s, \nu) \neq 0$, then RCAISE is not used and $\bar{\hat{F}}_{10.7}(k)$ remains unchanged from its previous value. Note that $T_s = 2$ s and $\nu = 60$ s; hence, RCAISE updates $\bar{\hat{F}}_{10.7}(k)$ every 60 s.
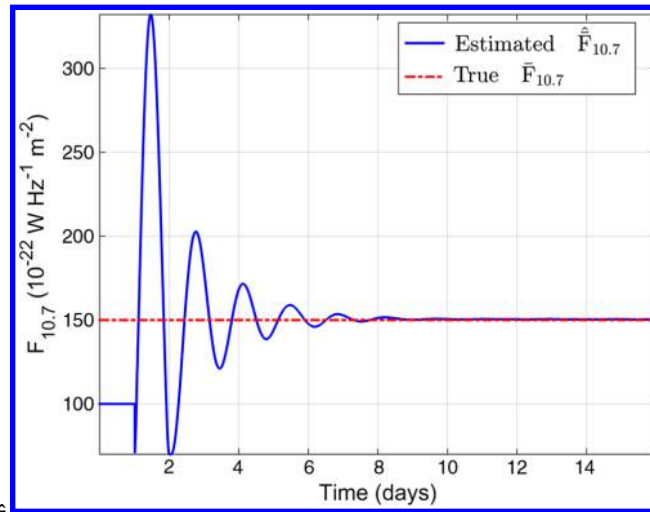
Fig. 10   True $\bar{F}_{10.7}(k)$ **and estimated** $\hat{\bar{F}}_{10.7}(k)$ **for the case of simulated satellite data with the constant true** $\bar{F}_{10.7}(k) = 150$ **SFsU. This plot shows that, after RCAISE is turned on,** $\hat{\bar{F}}_{10.7}(k)$ **converges to within 10 SFUs of** $\bar{F}_{10.7}(k)$ **within four days.**

The filters P_1 and P_2 for all of the following studies are defined as

$$\hat{\bar{F}}_{10.7}(k) = P_1(\hat{\bar{F}}_{10.7,f}(k)) = \frac{1}{90}\sum_{i=0}^{119}\hat{\bar{F}}_{10.7,f}(k-i) \quad z_f(k) = P_2(z(k)) = \frac{1}{120}\sum_{i=0}^{119}z(k-i)$$

Note that $\hat{\bar{F}}_{10.7}(k)$ is 120 min average of the driver-estimator subsystem output, and $z_f(k)$ is a 90 min average of the difference between the output of the physical system and the estimator output, as shown in Fig. 8.

The orbital period of CHAMP is 90 min, during which it traverses from the day side to the night side of Earth, causing an oscillatory behavior in $y_C(k)$ and $y_G(k)$. Therefore, for the following studies, we show the 90 min moving averages $\bar{y}_{C,90}(k)$ and $\bar{y}_{G,90}(k)$.

A.   Simulated Satellite Data (Constant $F_{10.7}$)

We consider the case where the truth data are generated by setting $\bar{F}_{10.7}(k) = 150$ SFUs for all $k \geq 0$. Neutral density is extracted from GITM output along the circular orbit in CHAMP's orbital plane at the altitude of 413.513 km and along the circular orbit in GRACE's orbital plane at the altitude of 501.171 km. RCAISE modeling data are chosen as $\tilde{\mathcal{H}} = H_1 = 0.1$, $\gamma = 100$, and $\eta = 0$.

Figure 10 shows that $\bar{F}_{10.7}$ converges to within 10 SFUs within four days after RCAISE is switched on. Figures 11 and 12 show the true and estimated 90 min averages of the neutral density along the circular orbits in CHAMP and GRACE orbital planes. These plots show that the estimated neutral densities $\hat{\bar{y}}_{C,90}(k)$ and $\hat{\bar{y}}_{G,90}(k)$ converge to the true neutral densities $\bar{y}_{C,90}(k)$ and $\bar{y}_{G,90}(k)$ within four days after RCAISE is switched on. Note that RCAISE does not use $\bar{y}_G(k)$ to estimate the driver; it is used as a metric to assess the quality of the state estimate.

As noted in Sec. IV, if one satellite measurement is available, then the solar zenith angle at the satellite location may become greater than 90 deg. When this occurs, the response of GITM to a change in $\bar{F}_{10.7}(k)$ is delayed. Figure 13 shows the time periods during which the solar zenith angle at the satellite location is greater than 90 deg, as indicated by the shaded regions. Despite this, RCAISE is able to force the output from GITM to match the true output. In addition, it may be possible to increase the rate of convergence of the estimates by using multiple satellites to reduce the time periods during which the minimum solar zenith angle at all satellites is greater than 90 deg.
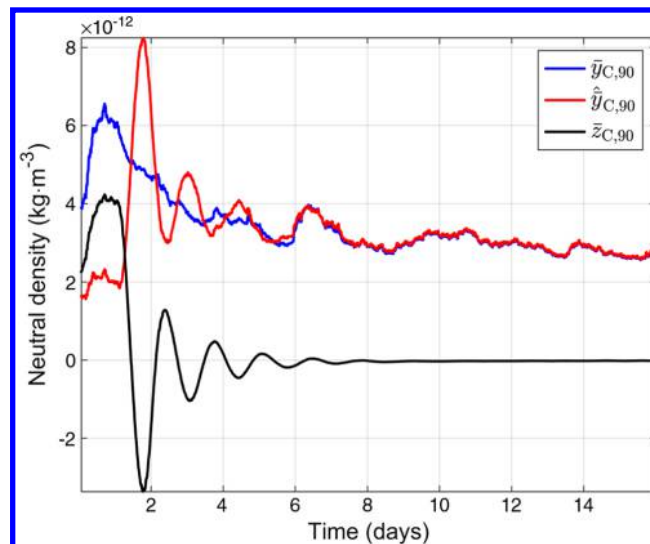

Fig. 11   Ninety-minute average of measured neutral density $\bar{y}_{C,90}(k)$ **and the estimated neutral density** $\bar{y}_{C,90}(k)$ **at CHAMP's location for the case of simulated satellite data with the constant true** $\bar{F}_{10.7}(k) = 150$ **SFUs. In this example, RCAISE is turned on after GITM simulates 24 h of atmospheric response. This plot shows that the error** $\bar{z}_{C,90}(k) = \bar{y}_{C,90}(k) - \hat{\bar{y}}_{C,90}(k)$ **converges to zero.**
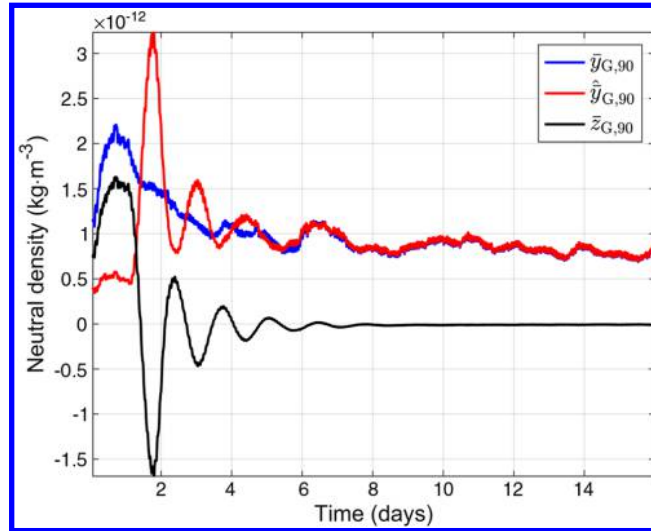
Fig. 12 **Ninety-minute average of measured neutral density $\bar{y}_{G,90}(k)$ and the estimated neutral density $\hat{\bar{y}}_{G,90}(k)$ at GRACE's location for the case of simulated satellite data with the constant true $\bar{F}_{10.7}(k) = 150$ SFUs. Note that $\bar{y}_{G,90}(k)$ is used to assess the quality of state estimation, and it is not assimilated by RCAISE.**
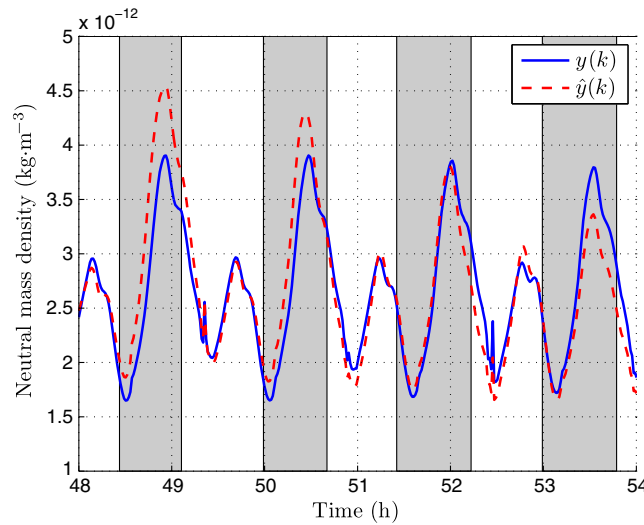

Fig. 13 **Time periods during which the solar zenith angle at the satellite location is greater than 90 deg (shaded region) for the case of simulated CHAMP satellite data with a constant of $\bar{F}_{10.7}(k) = 150$ SFUs. Note that the neutral density estimate $\hat{y}(k)$ is more accurate when the solar zenith angle at the satellite location is less than 90 deg.**
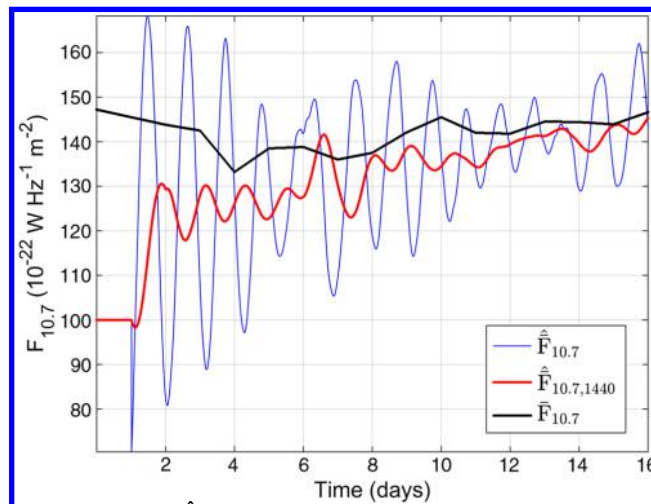

Fig. 14 **Comparison of measured and estimated $F_{10.7}(k)$. $\hat{\bar{F}}_{10.7}(k)$ is the estimate of $F_{10.7}(k)$ given by RCAISE by assimilating neutral density measured by CHAMP satellite in GITM. This plot shows that $\hat{\bar{F}}_{10.7}(k)$ averaged over one day given by $\hat{\bar{F}}_{10.7,1440}(k)$ converges to within 10 SFUs of measured $\bar{F}_{10.7}(k)$ within five days.**
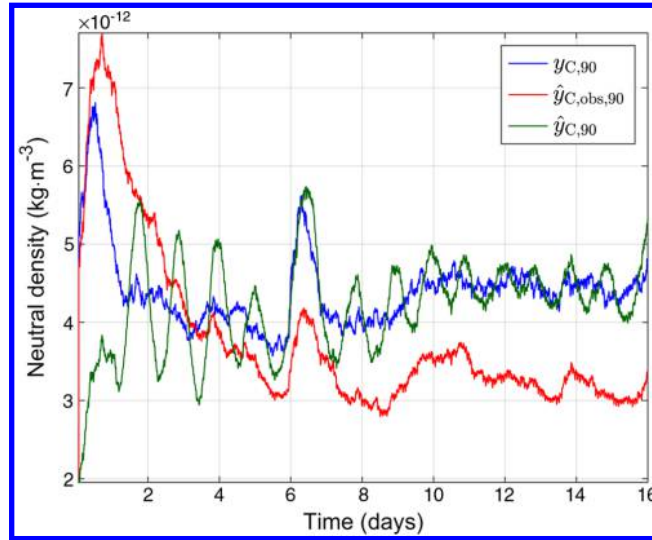
Fig. 15   Ninety-minute average of measured neutral density $y_{C,90}(k)$**, computed neutral density** $y_{C,\text{obs},90}(k)$ **by using measured** $\bar{F}_{10.7}(k)$ **in GITM, and the estimated neutral density** $\hat{y}_{C,90}(k)$ **by using** $\hat{\bar{F}}_{10.7}(k)$ **estimated by RCAISE in GITM at CHAMP's location for the case of real satellite data. In this example, RCAISE is turned on after GITM simulates 24 h of atmospheric response. Note that** $\text{RMS}(z_{C,\text{obs}} = y_{C,90} - y_{C,\text{obs},90}) = 1.0504 \times 10^{-12}$ **kg·m**$^{-3}$**, whereas** $\text{RMS}(z_{C,\text{RCAISE}} = y_{C,90} - \hat{y}_{C,90}) = 0.4603 \times 10^{-12}$ **kg·m**$^{-3}$**, and RCAISE reduces the error in the estimated neutral density by 56%.**

B.   Real Satellite Data

We now consider the case where the truth data are the neutral density measurements made by CHAMP and GRACE from 21 November 2002 to 6 December 2002. RCAISE modeling data are chosen as $\tilde{\mathcal{H}} = H_1 = 0.1$, $\gamma = 100$, and $\eta = 0$.

In addition to RCAISE, we also run GITM with $F_{10.7}$ obtained from ground measurements by the Dominion Radio Astrophysical Observatory [9]. We denote neutral density calculated by GITM using measured $F_{10.7}$ along CHAMP's and GRACE's orbits $\hat{y}_{C,\text{obs}}$ and $\hat{y}_{G,\text{obs}}$.

Figure 14 shows the estimate $\hat{\bar{F}}_{10.7}(k)$ along with one-day average $\hat{\bar{F}}_{10.7,1440}(k)$ and the measured $\bar{F}_{10.7}$. Figure 15 shows the 90 min average of the neutral density $\bar{y}_{C,90}(k)$, $\hat{\bar{y}}_{C,90}(k)$, and $\hat{\bar{y}}_{C,\text{obs},90}(k)$ along CHAMP orbit.

We calculate the root mean square of the output error z_C after 24 h in order to minimize the effect of transients generated due to the initial conditions. GITM with measured $\bar{F}_{10.7}$ gives $\text{RMS}(z_{C,\text{obs}} = \bar{y}_{C,90} - \hat{\bar{y}}_{C,\text{obs},90}) = 1.0504 \times 10^{-12}$ kg·m^{-3}, whereas GITM with RCAISE-estimated $\bar{F}_{10.7}(k)$ gives $\text{RMS}(z_{C,\text{RCAISE}} = \bar{y}_{C,90} - \hat{\bar{y}}_{C,90}) = 0.4603 \times 10^{-12}$ kg·m^{-3}, which is an improvement of 56%.

To determine the quality of the state estimate, we use the neutral density measured by GRACE. Figure 16 shows the 90 min average of the neutral density $\bar{y}_{G,90}(k)$, $\hat{\bar{y}}_{G,90}(k)$, and $\hat{\bar{y}}_{G,\text{obs},90}(k)$ along GRACE orbit. We calculate the root mean square of the error z_G after 24 h in order to minimize the effect of transients generated due to the initial conditions. GITM with measured $\bar{F}_{10.7}$ gives $\text{RMS}(z_{G,\text{obs}} = \bar{y}_{G,90} - \hat{\bar{y}}_{G,\text{obs},90}) = 0.2796 \times 10^{-12}$ kg·m^{-3}, whereas GITM with RCAISE estimated $\bar{F}_{10.7}(k)$ gives $\text{RMS}(z_{G,\text{RCAISE}} = \bar{y}_{G,90} - \hat{\bar{y}}_{G,90}) = 0.2426 \times 10^{-12}$ kg·m^{-3}, which is an improvement of 13%.
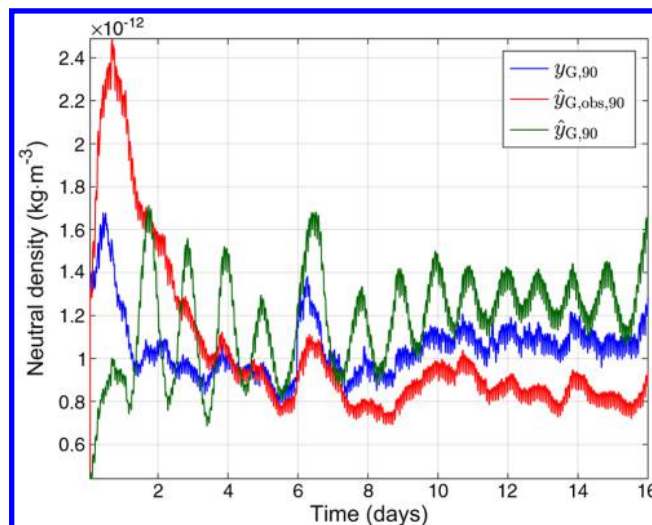

Fig. 16   Ninety-minute average of measured neutral density $y_{G,90}(k)$**, computed neutral density** $y_{G,\text{obs},90}(k)$ **by using measured** $\bar{F}_{10.7}(k)$ **in GITM, and the estimated neutral density** $\hat{y}_{G,90}(k)$ **by using** $\hat{\bar{F}}_{10.7}(k)$ **estimated by RCAISE in GITM at GRACE's location for the case of real satellite data. In this example, RCAISE is turned on after GITM simulates 24 h of atmospheric response. Note that** $\text{RMS}(z_{G,\text{obs}} = y_{G,90} - y_{G,\text{obs},90}) = 2.7963 \times 10^{-12}$ **kg·m**$^{-3}$**, whereas** $\text{RMS}(z_{G,\text{RCAISE}} = y_{C,90} - \hat{y}_{C,90}) = 2.4268 \times 10^{-12}$ **kg·m**$^{-3}$**, and RCAISE reduces the error in the estimated neutral density by 13%.**

VII. Conclusions

This paper used retrospective-cost input and state estimation for data assimilation with neutral density measurements from a satellite in the global ionosphere–thermosphere model. For the case of simulated satellite measurements, GITM was used as both the truth model and the data assimilation model, whereas for the case of real satellite measurements, GITM was used as the data assimilation model. RCAISE estimates both states and unknown drivers under the assumption that the system is strongly driven. RCAISE requires neither an ensemble nor an adjoint code, but instead involves the minimization of two quadratic cost functions at each assimilation step. RCAISE is not based on Bayesian updates; thus, unlike ensemble-based data assimilation methods, does not explicitly account for instrument errors and does not estimate probability distributions of the estimation errors.

For input and state estimation in GITM, it was assumed that the only available measurements were the neutral density data obtained from CHAMP satellite. For the case of simulated CHAMP data based on GITM truth model, RCAISE was able to estimate $\bar{F}_{10.7}$ and states within five days. For the case of real CHAMP data with GITM, the $\hat{\bar{F}}_{10.7}$ estimate provided by RCAISE reduced the rms error between the measured output and the output from GITM by 56% compared to GITM using the measured $\bar{F}_{10.7}$, and it reduced the rms error between GRACE data and the output from GITM at GRACE locations by 13% compared to GITM using the measured $\bar{F}_{10.7}$. For all of GITM state and driver estimation examples, the computational cost of the RCAISE adaptive driver-estimator subsystem update was equivalent to approximately 1% of the computational cost of GITM. In comparison, the related study in [22] required about 200 times more computational effort.

The specific application of RCAISE in the present paper, namely, the use of neutral density measurements to estimate $F_{10.7}$, is intended as an illustration of RCAISE on a highly nonlinear, physically realistic system. Although measurements of $F_{10.7}$ are available from observations, the estimates provided by RCAISE at a modest computational cost may, for practical purposes, be useful for validation. Furthermore, RCAISE can provide the estimates of $F_{10.7}$ at desired temporal spacing. In addition, this work provides a first step toward estimating drivers that cannot be measured or estimated by existing techniques.

The demonstration of RCAISE in this paper is but a first toward elucidating the properties of this method. Quantification of noise sensitivity, convergence speed, accuracy, computational cost, and ease of tuning relative to data assimilation methods such as 4-D VAR and EAKF remain to be examined beyond the results of [22]. Finally, it is noted that the ability of RCAISE to estimate drivers at low computational cost may serve as a useful aid to ensemble-based data assimilation algorithms: for example, by merging the classical state estimator [Eq. (7)] with the adaptive state estimator [Eqs. (2–6)]. In this sense, RCAISE may provide a low-cost enhancement to relatively expensive Bayesian data assimilation techniques.

Acknowledgments

References

[1] Kitanidis, P. K., "Unbiased Minimum-Variance Linear State Estimation," *Automatica*, Vol. 23, April 1987, pp. 775–778.
    doi:10.1016/0005-1098(87)90037-9

[2] Gillijns, S., and De Moor, B., "Unbiased Minimum-Variance Input and State Estimation for Linear Discrete-Time Systems," *Automatica*, Vol. 43, No. 1, 2007, pp. 111–116.
    doi:10.1016/j.automatica.2006.08.002

[3] Xiong, Y., and Saif, M., "Unknown Disturbance Inputs Estimation Based on a State Functional Observer Design," *Automatica*, Vol. 39, Aug. 2003, pp. 1389–1398.
    doi:10.1016/S0005-1098(03)00087-6

[4] Floquet, T., and Barbot, J. P., "State and Unknown Input Estimation for Linear Discrete-Time Systems," *Automatica*, Vol. 42, Nov. 2006, pp. 1883–1889.
    doi:10.1016/j.automatica.2006.05.030

[5] Sundaram, S., and Hadjicostis, C. N., "Delayed Observers for Linear Systems with Unknown Inputs," *IEEE Transactions on Automatic Control*, Vol. 52, Feb. 2007, pp. 334–339.
    doi:10.1109/TAC.2006.890371

[6] Angeli, D., "A Lyapunov Approach to Incremental Stability Properties," *IEEE Transactions on Automatic Control*, Vol. 47, March 2002, pp. 410–421.
    doi:10.1109/9.989067

[7] Russo, G., di Bernardo, M., and Sontag, E. D., "Global Entrainment of Transcriptional Systems to Periodic Inputs," *PLOS Computational Biology*, Vol. 6, April 2010, Paper e1000739.
    doi:10.1371/journal.pcbi.1000739

[8] Sontag, E., "Contractive Systems with Inputs," *Perspectives in Mathematical System Theory, Control, and Signal Processing*, edited by Willems, J., Hara, S., Ohta, Y., and Fujioka, H., Vol. 398, Springer, New York, 2010, pp. 217–228.

[9] *Dominion Radio Astrophysical Observatory* [online database], National Research Council, Ottawa, 2014, http://astro-canada.ca/_en/a2107.php [retrieved 2015].

[10] Ridley, A. J., Deng, Y., and T'oth, G., "The Global Ionosphere/Thermosphere Model," *Journal of Atmospheric and Solar-Terrestrial Physics*, Vol. 68, No. 8, 2006, pp. 839–864.
    doi:10.1016/j.jastp.2006.01.008

[11] D'Amato, A. M., Ali, A. A., Springmann, J. C., Cutler, J. W., Ridley, A. J., and Bernstein, D. S., "Adaptive State Estimation for Nonminimum-Phase Systems with Uncertain Harmonic Inputs," *Guidance, Navigation, and Control Conference*, AIAA Paper 2011-6315, 2011.

[12] Lewis, J. M., and Derber, J. C., "The Use of Adjoint Equations to Solve a Variational Adjustment Problem with Advective Constraints," *Tellus A*, Vol. 37, No. 4, 1985, pp. 309–322.
    doi:10.1111/tela.1985.37A.issue-4

[13] Thépaut, J.-N., Hoffman, R. N., and Courtier, P., "Interactions of Dynamics and Observations in a Four-Dimensional Variational Assimilation," *Monthly Weather Review*, Vol. 121, No. 12, 1993, pp. 3393–3414.
    doi:10.1175/1520-0493(1993)121<3393:IODAOI>2.0.CO;2

[14] Andersson, E., Pailleux, J., Thépaut, J.-N., Eyre, J., McNally, A., Kelly, G., and Courtier, P., "Use of Cloud-Cleared Radiances in Three/Four-Dimensional Variational Data Assimilation," *Quarterly Journal of the Royal Meteorological Society*, Vol. 120, No. 517, 1994, pp. 627–653.
    doi:10.1002/(ISSN)1477-870X

[15] Courtier, P., Thépaut, J.-N., and Hollingsworth, A., "A Strategy for Operational Implementation of 4-D-Var, Using an Incremental Approach," *Quarterly Journal of the Royal Meteorological Society*, Vol. 120, No. 519, 1994, pp. 1367–1387.
    doi:10.1002/(ISSN)1477-870X

[16] Zou, X., Navon, I., and LeDimet, F., "An Optimal Nudging Data Assimilation Scheme Using Parameter Estimation," *Quarterly Journal of the Royal Meteorological Society*, Vol. 118, No. 508, 1992, pp. 1163–1186.
doi:10.1002/(ISSN)1477-870X

[17] Stauffer, D. R., and Bao, J.-W., "Optimal Determination of Nudging Coefficients Using the Adjoint Equations," *Tellus A*, Vol. 45, No. 5, 1993, pp. 358–369.
doi:10.1034/j.1600-0870.1993.t01-4-00003.x

[18] Houtekamer, P. L., and Mitchell, H. L., "Data Assimilation Using an Ensemble Kalman Filter Technique," *Monthly Weather Review*, Vol. 126, No. 3, March 1998, pp. 796–811.
doi:10.1175/1520-0493(1998)126<0796:DAUAEK>2.0.CO;2

[19] Evensen, G., "Sequential Data Assimilation with a Nonlinear Quasi-Geostrophic Model Using Monte Carlo Methods to Forecast Error Statistics," *Journal of Geophysical Research*, Vol. 99, May 1994, pp. 10,143–10,162.
doi:10.1029/94JC00572

[20] Anderson, J. L., "An Ensemble Adjustment Kalman Filter for Data Assimilation," *Monthly Weather Review*, Vol. 129, No. 12, Dec. 2001, pp. 2884–2903.
doi:10.1175/1520-0493(2001)129<2884:AEAKFF>2.0.CO;2

[21] "GFZ German Research Centre for Geosciences: CHAMP Mission [online database]," GFZ German Research Centre, Potsdam, Germany 2012, http://op.gfz-potsdam.de/champ/ [retrieved 2015].

[22] Morozov, A. V., Ridley, A. J., Bernstein, D. S., Collins, N., Hoar, T. J., and Anderson, J. L., "Data Assimilation and Driver Estimation for the Global Ionosphere-Thermosphere Model Using the Ensemble Adjustment Kalman Filter," *Journal of Atmospheric and Solar-Terrestrial Physics*, Vol. 104, Nov. 2013, pp. 126–136.
doi:10.1016/j.jastp.2013.08.016

[23] Anderson, J., Hoar, T., Raeder, K., Liu, H., Collins, N., Torn, R., and Avellano, A., "The Data Assimilation Research Testbed: A Community Facility," *Bulletin of the American Meteorological Society*, Vol. 90, Sept. 2009, pp. 1283–1296.
doi:10.1175/2009BAMS2618.1

[24] Agarwal, K., Ali, A. A., D'Amato, A. M., Ridley, A. J., and Bernstein, D. S., "Retrospective-Cost-Based Adaptive State Estimation and Input Reconstruction for the Global Ionosphere-Thermosphere Model," *Guidance, Navigation, and Control Conference*, AIAA Paper 2012-4601, 2012.

[25] D'Amato, A. M., Ridley, A. J., and Bernstein, D. S., "Retrospective-Cost-Based Adaptive Model Refinement for the Ionosphere and Thermosphere," *Statistical Analysis and Data Mining*, Vol. 4, No. 4, 2011, pp. 446–458.
doi:10.1002/sam.v4.4

[26] D'Amato, A. M., Ali, A. A., Ridley, A., and Bernstein, D. S., "Retrospective Cost Optimization for Adaptive State Estimation, Input Estimation, and Model Refinement," *Procedia Computer Science*, Vol. 18, 2013, pp. 1919–1928.
doi:10.1016/j.procs.2013.05.361

[27] *GRACE: Gravity Recovery and Climate Experiment [online database]*, Univ. of Texas at Austin, Center for Space Research, Austin, TX, 2012, http://www.csr.utexas.edu/grace/ [retrieved 2015].

[28] Söderström, T., and Stoica, P., *System Identification*, Prentice–Hall, Upper Saddle River, NJ, 1988, pp. 101–111.

[29] Deng, Y., Richmond, A. D., Ridley, A. J., and Liu, H.-L., "Assessment of the Non-Hydrostatic Effect on the Upper Atmosphere Using a General Circulation Model (GCM)," *Geophysical Research Letters*, Vol. 35, No. 1, Jan. 2008, Paper L01104.

[30] Yiğit, E., and Ridley, A., "Role of Variability in Determining the Vertical Wind Speeds and Structure," *Journal of Geophysical Research*, Vol. 116, No. A12, 2011, Paper A12305.

[31] Hedin, A., "A Revised Thermospheric Model Based on Mass Spectrometer and Incoherent Scatter Data: MSIS-83," *Journal of Geophysical Research*, Vol. 88, No. A12, 1983, pp. 10,170–10,188.
doi:10.1029/JA088iA12p10170

[32] Hedin, A. E., "MSIS-86 Thermospheric Model," *Journal of Geophysical Research*, Vol. 92, No. A5, 1987, pp. 4649–4662.
doi:10.1029/JA092iA05p04649

[33] Hedin, A., "Extension of the MSIS Thermosphere Model into the Middle and Lower Atmosphere," *Journal of Geophysical Research*, Vol. 96, No. A2, 1991, pp. 1159–1172.
doi:10.1029/90JA02125

[34] Bilitza, D., "International Reference Ionosphere 2000," *Radio Science*, Vol. 36, No. 2, 2001, pp. 261–275.
doi:10.1029/2000RS002432

[35] Drob, D., et al., "An Empirical Model of the Earth's Horizontal Wind Fields: HWM07," *Journal of Geophysical Research*, Vol. 113, No. A12, 2008,
doi:10.1029/2008JA013668

[36] Weimer, D., "Improved Ionospheric Electrodynamic Models and Application to Calculating Joule Heating Rates," *Journal of Geophysical Research*, Vol. 110, No. A5, 2005.
doi:10.1029/2004JA010884

[37] Fuller-Rowell, T., and Evans, D., "Height-Integrated Pedersen and Hall Conductivity Patterns Inferred from the TIROS-NOAA Satellite Data," *Journal of Geophysical Research*, Vol. 92, No. A7, 1987, pp. 7606–7618.
doi:10.1029/JA092iA07p07606

[38] Pawlowski, D. J., and Ridley, A. J., "Modeling the Thermospheric Response to Solar Flares," *Journal of Geophysical Research*, Vol. 113, No. A10, Oct. 2008, Paper a10309.
doi:10.1029/2008JA013182

[39] Richards, P. G., Fennely, J. A., and Torr, D. G., "EUVAC: A Solar EUV Flux Model for Aeronomic Calculations," *Journal of Geophysical Research*, Vol. 99, No. A5, May 1994, pp. 8981–8992.
doi:10.1029/94JA00518

G. Chowdhary
Associate Editor

This article has been cited by:

1. Ahmad Ansari, Dennis S. Bernstein. Estimation of angular velocity and rate-gyro noise for sensor health monitoring 128-133. [Crossref]

2. Ahmad Ansari, Dennis S. Bernstein. Adaptive input estimation for nonminimum-phase discrete-time systems 1159-1164. [Crossref]

3. Liang Han, Zhang Ren, Dennis S. Bernstein. 2016. Maneuvering target tracking using retrospective-cost input estimation. *IEEE Transactions on Aerospace and Electronic Systems* **52**:5, 2495-2503. [Crossref]

4. Ahmad Ansari, Dennis S. Bernstein. Adaptive non-Bayesian state estimation 6977-6982. [Crossref]

5. Ankit Goel, Karthik Duraisamy, Dennis S. Bernstein. Parameter estimation in the Burgers equation using retrospective-cost model refinement 6983-6988. [Crossref]