Improved basis functions for dynamic calibration of semi-empirical thermospheric models

Eric K. Sutton, Samuel B. Cable, Chin S. Lin
Air Force Research Laboratory

Frank A. Marcos
Boston College

ABSTRACT

State-of-the-art satellite drag models require upgrades to meet operational Precision Orbit Determination requirements for collision avoidance, reentry predictions and catalog maintenance. Accurate model representations of the upper atmosphere are not currently possible without the use of data assimilation, or model calibration. Due to incomplete global data sampling in the thermosphere, such calibration has only been successfully demonstrated by fitting the available data to a low-dimensional model. The High Accuracy Satellite Drag Model (HASDM), used by the Space Surveillance Network to track low-earth orbiting satellites, fits recent data to a truncated set of spherical harmonics. In our study, the goal is to replace this low-order spherical harmonic expansion with a set of basis functions better suited to capture the spatial variability and response of the thermosphere. By comparing the base model of HASDM with the Thermosphere-Ionosphere-Electrodynamic General Circulation Model (TIEGCM), we create a new set of orthogonal basis functions that can be used to calibrate semi-empirical models such as HASDM with increased accuracy in the presence of sparse data. We present initial comparisons between the conventional and new approaches.

1. INTRODUCTION

Atmospheric drag is the dominant and most difficult force to determine and predict, in the orbit propagation model of low earth orbiting satellites [Marcos and Wise, 2002]. The orbital drag acceleration, $\vec{a}_D$, can be related to satellite properties and atmospheric density, ρ, by:

$$\vec{a}_D = -1/2\left(C_D A_{ref}/m\right)\rho|\vec{V}|\vec{V} \tag{1}$$

where C_D is the coefficient of drag, A_{ref} is the reference satellite area projected into the ram direction, m is the satellite mass, and $\vec{V}$ is the satellite velocity with respect to the atmosphere. Neutral density contributes the most to the total variability of drag acceleration, however, the $C_D A_{ref}$ term and thermospheric winds (entering through the $\vec{V}$ term) can also contribute significant amounts at times.

The thermosphere is a strongly driven dynamic system. Variability of neutral density in the thermosphere depends not only on location but on solar and geophysical conditions as well. Accelerations can change by more than an order of magnitude during the solar cycle with an approximate period of 11 years, and by a factor of 2-4 during moderate geomagnetic events. During such events, the spatial distribution and temporal response strongly depends on latitude and local time, both in geographic and magnetic coordinates.

Air Force Space Command requires knowledge and forecasting of thermospheric density accurate to within 5% in the thermosphere between 90 and 500 km. This requirement is focused on improving the efficiency of satellite catalog maintenance, the accuracy of reentry predictions, and the reliability of satellite conjunction analysis. In support of this goal, the current pursuit focuses on the incremental improvement of empirical model calibration techniques.

2. DESCRIPTION OF MODELS

The Jacchia 70 (J70 hereafter) [Jacchia, 1970] is a static diffusion model of the upper atmosphere. The thermospheric portion of the model begins at 105 km. Above this level, J70 specifies the total number density, *n*, for each species, *i* = {Ar, He, N_2, O, or O_2}, by vertically integrating the diffusion equation:

$$\frac{dn_i}{n_i} = -\frac{m_i g}{kT} dz - \frac{dT}{T}(1 + \alpha_i) \tag{2}$$

where *m* is the molecular weight, *g* is the acceleration due to gravity, *z* is the height, *k* is the Boltzmann constant, *T* is the temperature, and α is the diffusion coefficient (α = -0.38 for He, and α = 0 for all other species). The total mass density, ρ, is the summation of the mass densities of all species: $\rho = \sum_i n_i m_i / N_A$, where N_A is Avogadro's number. The vertical temperature profile, *T(z)*, required to carry out the integration in (1) is parameterized in the following form:

$$T(z) = T_x + A \tan^{-1}\left\{\frac{G_x}{A}(z - z_x)[1 + B(z - z_x)^j]\right\} \tag{3}$$

where $A = \frac{2}{\pi}(T_\infty - T_x)$, $B = 4.5\times10^{-6}$ for *z* in km, T_X is the temperature at a prescribed inflection point z_x = 125 km, G_x is the temperature gradient at the inflection point, T_∞ is the exospheric temperature, j = 2.5. With the application of several additional constraints [see Jacchia, 1970], the model produces a vertical density profile that is uniquely specified by T_X and T_∞.

Marcos et al. [1997] first attempted to calibrate this model by estimating a global correction to the value T_∞, designated as ΔT_C, to bring the model into better agreement with recent satellite tracking data. Storz et al. [2002] and Casali and Barker [2002] extended this technique by estimating a spherical harmonic field for ΔT_C as well as for a correction to the value of T_X, designated as ΔT_X. Due to the sparse data sets available, it was necessary to truncate the spherical harmonic fields at degree and order (2,2) for ΔT_C and (1,1) for ΔT_X. This extension has come to be known as the High Accuracy Satellite Drag Model (HASDM). The goal of our study is to replace the truncated set of spherical harmonic functions used to specify ΔT_C and ΔT_X with a global basis set that more accurately represents the true variability of the thermosphere.

The Thermosphere-Ionosphere-Electrodynamics General Circulation Model (TIE-GCM) [Richmond et al., 1992], developed by the National Center for Atmospheric Research (NCAR), is a comprehensive, first-principles, three-dimensional, non-linear representation of the coupled thermosphere and ionosphere system that includes a self-consistent solution of the middle and low-latitude dynamo field. The model solves the three-dimensional momentum, energy and continuity equations for neutral and ion species at each time step, using a semi-implicit, fourth-order, centered finite difference scheme on each pressure surface in a staggered vertical grid. It can be run in either serial or parallel mode on a variety of platforms, including Linux workstations and supercomputers [http://www.hao.ucar.edu/modeling/tgcm/tie.php, accessed 8-30-2011]. The time step used in this study is 120 seconds. The horizontal grid spacing is 5° in latitude and longitude, with a vertical spacing of a half scale height. Energy and momentum sources originating in the magnetosphere were parameterized by coupling to the Weimer-05 [Weimer, 2005] empirical convection electric field model near the poles. A realistic empirical model for the annual and semiannual variations of eddy diffusivity at the lower boundary is provided by Qian et al. [2009].

3. DEVELOPMENT OF A NEW BASIS SET

The end product of this study is an algorithm that operates in much the same way as those of Marcos et al. [1997] and Storz et al. [2002]. However, we establish an improved set of basis functions to more efficiently and accurately correct the J70 model in the presence of recent satellite data. This is done using TIEGCM output as input data to drive a ΔT_C and ΔT_X correction in J70. Through analysis of these corrections, we can capture the most significant modes of variability from the physics-based GCM.

We must first calculate the principal components of the ΔT_C and ΔT_X fields [see Björnsson and Venegas, 1997], using TIEGCM as input data. To this end, we use J70 in a similar fashion to the work of Storz et al. [2002]; however, instead of estimating spherical harmonic expansion coefficients, ΔT_C and ΔT_X are estimated directly at each latitude/local time grid point of TIEGCM using a vertical density profile from 200 km to the upper boundary of TIEGCM. Figure 1 shows this basic process at one location in latitude and local time. An iterative weighted least-squares fit brings the J70 density into agreement with the TIEGCM vertical profile by adjusting ΔT_C and ΔT_X within J70. This requires knowledge of the partials $\partial\rho/\partial T_C$ and $\partial\rho/\partial T_X$ with respect to the J70 model at each location and height, which are calculated by the J70DCA algorithm [Storz et al., 2002].

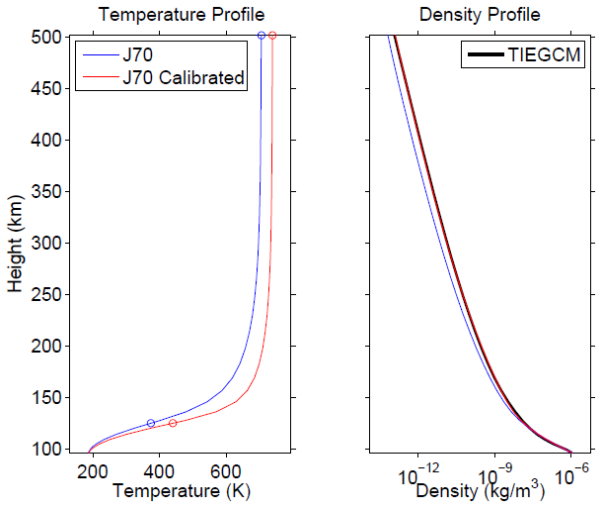

Fig. 1. Temperature (left) and density (right) vertical profiles for TIEGCM (black, density only), unadjusted-J70 (blue), and adjusted-J70 (red) at one latitude/local time grid point. Circles indicate unadjusted and adjusted T_X at 125 km and T_∞ near the upper boundary.

For this study, we estimate ΔT_C and ΔT_X at every latitude and local time grid point of TIEGCM for every hour of 2008. After arranging these fields in an m-by-n matrix called F, m being the number of times used (366 days * 24 hours) and n being the number of unique grid point locations (36 latitudes * 72 local times), we compute the covariance matrix $R = F^T F$. At this point, we solve the matrix eigenvector/eigenvalue problem: $RV = VD$, where the columns of V are the corresponding eigenvectors of R, and D is a diagonal matrix formed from the eigenvalues, λ_i, of R. The eigenvectors, V_i, are referred to as Empirical Orthogonal Functions (EOFs), while the eigenvalues relate to the amount of variability of the original field that is captured by each corresponding eigenvector. This amounts to finding the set of orthonormal basis functions that maximizes the projection of the row-vectors of F onto each basis function.

This procedure is carried out separately for ΔT_C and ΔT_X. Figures 2 and 3 show the first 9 EOFs for ΔT_C and the first 4 EOFs for ΔT_X, respectively, after rearranging into the latitude/local time coordinate frame. The different modes are a combination of the true thermospheric variability with the error in both TIEGCM and J70. Much of the variability caused by physical processes that cannot be captured by the J70 model – even when corrected by a truncated spherical harmonic expansion – can be represented by these basis functions. More importantly, the variability captured by each mode drops off much more quickly for EOFs than it does for a spherical harmonic expansion.

Although these functions are a convolution of model error – introduced by J70 and TIEGCM – with the true thermospheric variability, some information on the latter can still be gleaned from them. Modes 1 and 2 of ΔT_C are essentially a diurnal correction at low to mid latitudes. Modes 3 and 4 have a strong semidiurnal component evident at low to mid latitudes. Modes such as 6, 8, and 9 exhibit what appears to be a strong diurnal variation near either pole. This is most likely the signature of the longitude-UT effect [Hedin et al., 1979] related to the offset between

the geographic and geomagnetic poles, an effect that isn't well represented by the longitudinally-independent J70 model formulation. Additionally, some effects reminiscent of seasonal variations can be seen in mode 3, manifested as a hemispherically antisymmetric increase in amplitudes near the poles.

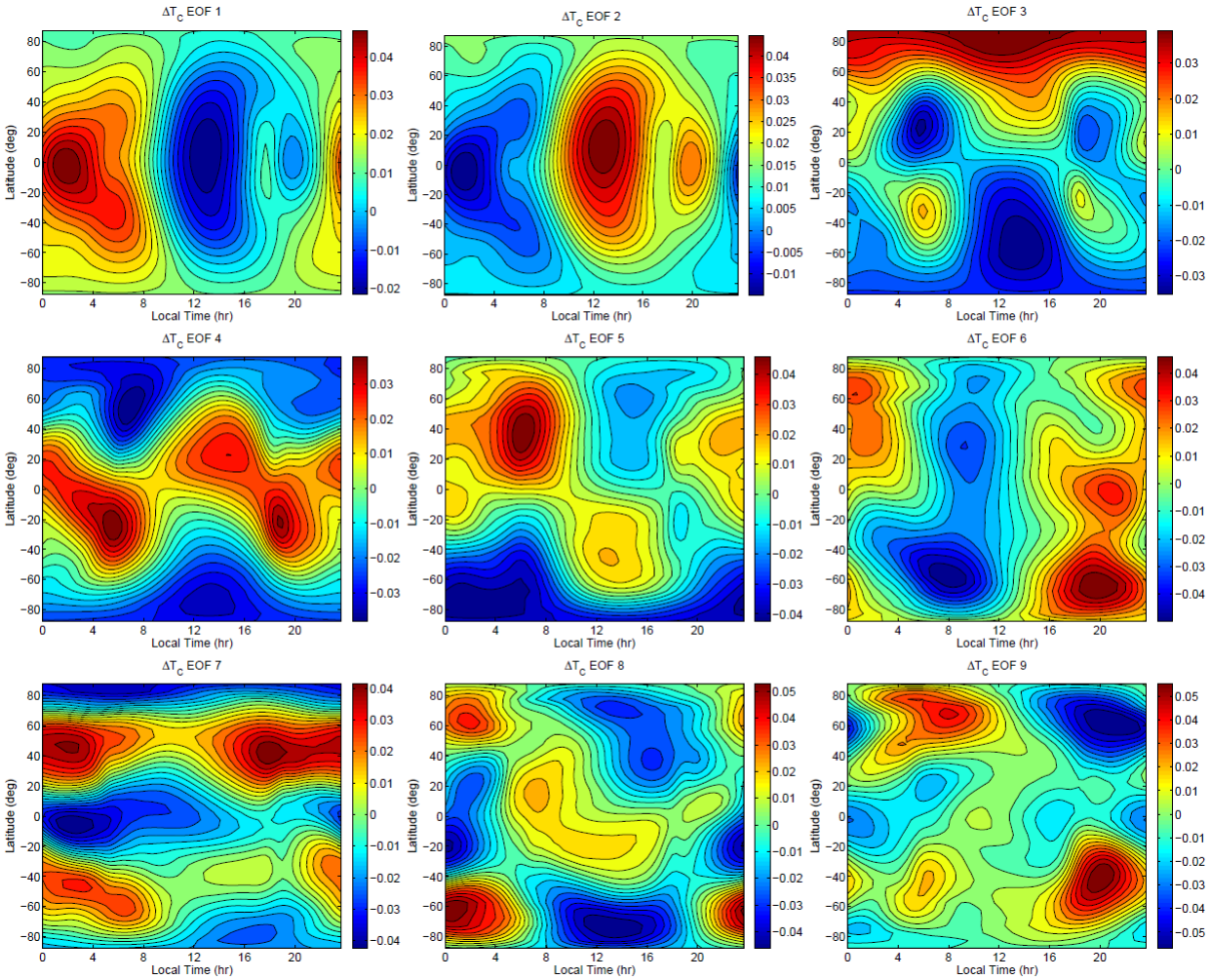

Fig. 2. EOF modes of ΔT_C estimated for TIEGCM during 2008. These 9 modes can be used in place of the 9 functions that currently comprise the [2x2] spherical harmonic expansion for ΔT_C. Each EOF is normalized such that $\sum_\varphi \sum_\theta V_i(\varphi,\theta)^2 = 1$, where the summations are performed over all latitudes, φ, and local times, θ.

For ΔT_X, mode 1 has a strong diurnal component, however the latitudinal structure is much different than modes 1 and 2 of ΔT_C. A strong antisymmetric semidiurnal component is apparent in mode 2. A strong seasonal component is seen in modes 3 and 4.

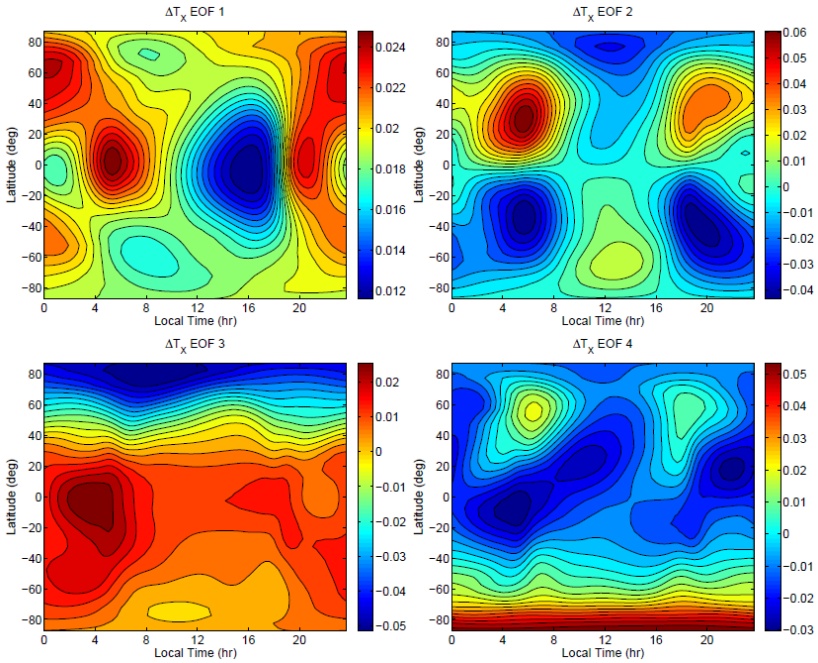

Fig 3. EOF modes of ΔT_X estimated for TIEGCM during 2008. These 4 modes can be used in place of the 4 functions that currently comprise the [1x1] spherical harmonic expansion for ΔT_X. Each EOF is normalized as in Fig. 2.

4. INITIAL COMPARISON OF BASIS FUNCTIONS

For an initial comparison of the performance between the new and currently used basis functions, we reconstruct the ΔT_C and ΔT_X fields using both methods. The [2x2] spherical harmonic representation of ΔT_C uses 9 orthogonal functions while the [1x1] spherical harmonic representation of ΔT_X uses 4 orthogonal functions. To stay consistent with the conventional approach, we also limit the EOFs to 9 functions representing ΔT_C and 4 functions representing ΔT_X. At each time step of TIEGCM within 2008, the truncated set of expansion functions (either spherical harmonics or EOFs) can be fit to the ΔT_C and ΔT_X fields using a least squares method to estimate the expansion coefficients:

$$\Theta(\varphi,\theta) = \sum_i C_i \Theta_i(\varphi,\theta) \tag{4}$$

where Θ is the reconstructed field, the C_i*'s* are the estimated expansion coefficients, the Θ_i's are the expansion functions (either spherical harmonics or EOFs), φ is latitude, and θ is local time.. Figure 4 shows the original ΔT_C field estimated using TIEGCM neutral density as input data, as well as the fits using modes 1-9 of the EOFs and of the spherical harmonics. Figure 5 shows a similar fit for ΔT_X using modes 1-4. The epoch for figures 4 and 5 is chosen such that the performance of the truncated spherical harmonic expansion is typical of 2008.

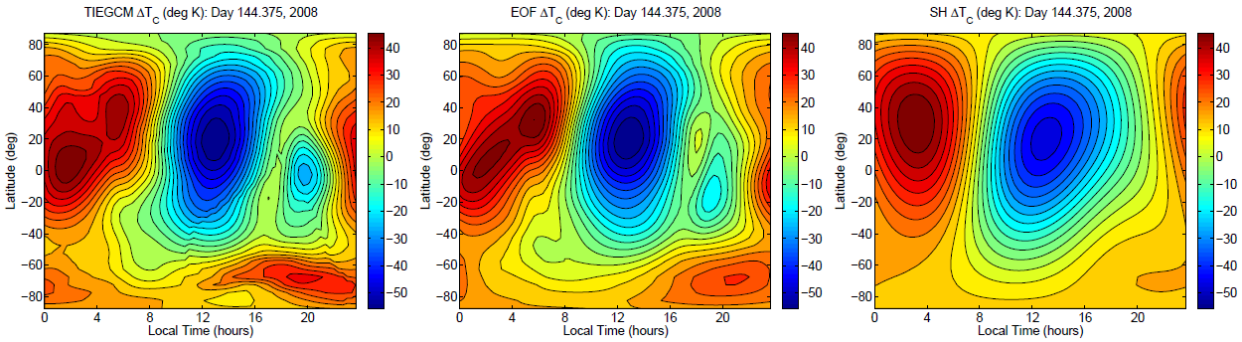

Fig 4. ΔT_C estimated for TIEGCM for a typical time epoch in 2008 (left), the reconstruction of ΔT_C using the 9 lowest order EOF expansion functions (center), and the reconstruction of ΔT_C using 9 [2x2] spherical harmonic expansion functions (right).

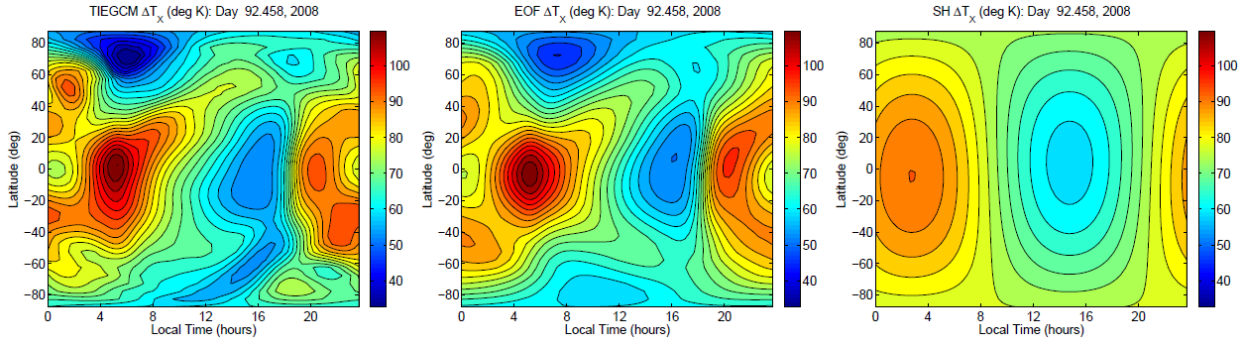

Fig 5. ΔT_X estimated for TIEGCM for a typical time epoch in 2008 (left), the reconstruction of ΔT_X using the 4 lowest order EOF expansion functions (center), and the reconstruction of ΔT_X using 4 [1x1] spherical harmonic expansion functions (right).

For the initial validation of the new basis set, we compare the reconstruction of TIEGCM-derived ΔT_C and ΔT_X using the EOF expansion functional approach with the conventional spherical harmonic expansion functional approach. The metric used is the *% RMS error*, defined as:

$$\%\ RMS\ error = 100 \times \sum_{\varphi}\sum_{\theta}\left(\Theta(\varphi,\theta) - \Delta T(\varphi,\theta)\right)^2 \Big/ \sum_{\varphi}\sum_{\theta}\Delta T(\varphi,\theta)^2 \qquad (5)$$

For the reconstructions of ΔT_C in figure 4, the *% RMS error* is 3.1% when using the EOF approach and 11.3% when using the spherical harmonic approach. For the reconstructions of ΔT_X in figure 5, the *% RMS error* is 0.37% when using the EOF approach and 2.43% when using the spherical harmonic approach.

Figures 6 shows the *% RMS error* over all of 2008 for the ΔT_C and ΔT_X fields, reconstructed using both methods. It should be noted that the EOF reconstructions of ΔT_C and ΔT_X always outperform those of the spherical harmonic expansions. During the year, spikes in error on the order of ~1 day are seen. These are well correlated with geomagnetic activity indices. While these spikes are still present in the EOF reconstructions, their amplitudes are much smaller than in the spherical harmonic reconstructions indicating that the EOFs capture the geomagnetically-induced variations of TIEGCM more efficiently.

Another salient feature is an annual/semiannual variation in the *% RMS error* of the spherical harmonic reconstruction of ΔT_X. Indicated by an increase in error around northern hemisphere summer, this is most likely related to seasonal variations as well as the annual/semiannual correction for eddy diffusivity that has been applied to TIEGCM. While the spherical harmonic expansions produce significant error, the EOFs ability to capture this feature is promising.

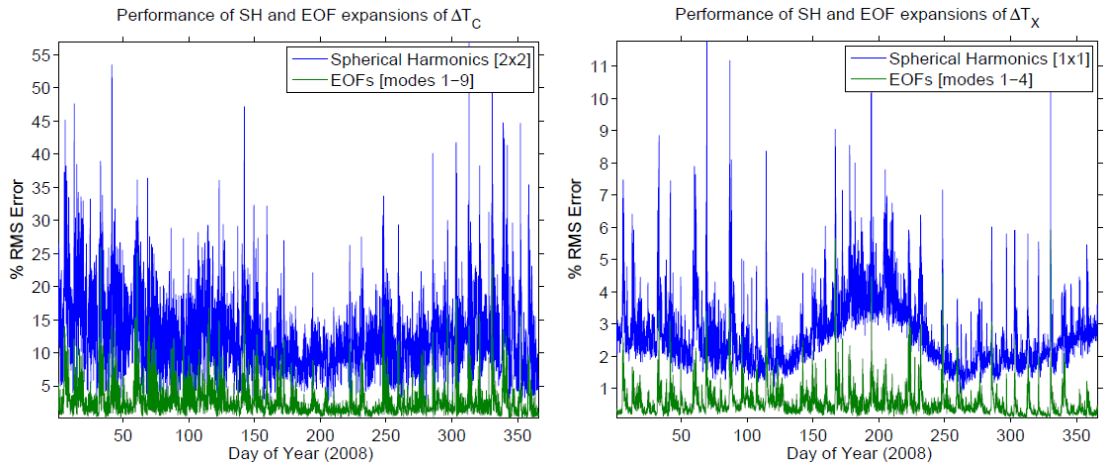

Fig 6. Percent RMS error when reconstructing the ΔT_C (left) and ΔT_X fields from EOFs (green) and spherical harmonics (blue).

5. SUMMARY AND FUTURE WORK

We have presented a new set of basis functions capable of representing the most important modes of thermospheric variability within the framework of an adjusted static diffusion model. Our approach demonstrates the potential for significant upgrades to current operational satellite drag modeling capabilities. Thermospheric variability is specified by TIEGCM, and thus, several limiting assumptions should be pointed out. First, the simplifying assumptions of hydrostatic equilibrium and constant gravity are imposed. Secondly, several physical processes are not fully accounted for, such as the influences of the lighter neutral and ion species, [H] and [He]. In addition, eddy diffusivity at the lower boundary is specified by an empirical parameterization. In spite of these simplifying assumptions and the subsequent error that they impose on TIEGCM, the modes of variability of TIEGCM are more realistic than any existing empirical model. The purpose of this study is to extract the most important of these modes and use them to efficiently calibrate empirical models, without the increased overhead that would be required to calibrate a high-dimension general circulation model.

The validation presented in this paper is only a first step. A comparison of the new and currently used techniques using actual satellite tracking data will be required before this basis set can be considered validated. As mentioned in Section 3, several of the EOF modes exhibit traits which are not included in the J70 model, e.g. the longitude-UT effect. Thus we are able to include and correct for modes of thermospheric variability not captured by J70. However, some of these traits may not be observable from the ground-based satellite tracking data set currently used to drive HASDM. The operational data validation study will resolve these issues, as well as guide any required modifications to the EOFs to provide the needed upgrade to satellite drag modeling capabilities.

6. REFERENCES

Björnsson, H., and S. A. Venegas (1997), A manual for EOF and SVD analyses of climatic data, Center for Climate and Global Change Research Report #97-1, 1-53.

Hedin, A., C. Reber, N. Spencer, H. Brinton, and D. Kayser (1979), Global model of longitude/UT variation in thermospheric composition and temperature based on mass spectrometer data, *J. Geophys. Res.*, 84(A1), 1-9.

Jacchia, L. G. (1970), New static models of the thermosphere and exosphere with empirical temperature profiles, *SAO Special Report #313*.

Marcos, F. A., M. J. Kendra, J. M. Griffin, J. N. Bass, D. R. Larson, and J. J. F. Liu (1997), Precision low earth orbit determination using atmospheric density calibration, *Adv. Astronaut. Sci.*, 97(1), 501-513.

Marcos, F. A., and J. O. Wise (2002), Towards a golden age of satellite drag, *40th AIAA Aerospace Sciences Meeting and Exhibit*, 14-17 January 2002, Reno, NV, AIAA 2002-0092.

Qian, L., S. C. Solomon, and T. J. Kane (2009), Seasonal variation of thermospheric density and composition, *J. Geophys. Res.*, 114(A01312), doi:10.1029/2008JA013643.

Richmond, A. D., E. C. Ridley, and R. G. Roble (1992), A thermosphere/ionosphere general circulation model with coupled electrodynamics, *Geophys. Res. Lett.*, 19(6), 601-604, doi:10.1029/92GL00401.

Storz, M. F., B. R. Bowman, and J. I. Branson (2002), High Accuracy Satellite Drag Model (HASDM), *AIAA/AAS Astrodynamics Specialist Conference and Exhibit*, 5-8 August 2002, Monterey, CA, AIAA 2002-4886.

Weimer, D. R. (2005), Improved ionospheric electrodynamic models and application to calculating Joule heating rates, *J. Geophys. Res.*, 110, A05306, doi:10.1029/2004JA010884.