STATISTICAL PERFORMANCE MEASURES OF THE HWM-93 AND MSISE-90 EMPIRICAL ATMOSPHERIC MODELS AND THE RELATION TO INFRASONIC CTBT MONITORING

Douglas P. Drob and J. M. Picone, Upper Atmospheric Physics Branch,
E.O. Hulburt Center for Space Research, Naval Research Laboratory

Sponsored by the U.S. Department of Energy
Office of Nonproliferation Research and Engineering
Office of Defense Nuclear Nonproliferation
National Nuclear Security Administration
Contract No. DE-AT04-99AL66330

ABSTRACT

An enforceable Comprehensive Nuclear-Test-Ban Treaty (CTBT) will require accurate detection and location of low-yield nuclear detonations. Thorough knowledge of the upper atmosphere and advanced modeling techniques are required for reliable infrasonic detection and location of clandestine events. The purpose of this paper is to document a statistical performance measures study of the Naval Research Laboratory's (NRL) empirical upper atmospheric models. These models known as the MSISE-90 and HWM-93 models were originally developed at the NASA Goddard Space Flight Center and made freely available to the public and scientific research communities. The Upper Atmospheric Physics branch at NRL is working toward the production of improved versions of these empirical upper atmospheric models for use in verification and compliance with the CTBT. This study identifies weak areas in the current models for planning and implementation of future upgrades. In this report, we detail model performance in two atmospheric regions; 0 to 50 km and 50 to 120 km. This report primarily focuses on systematic biases in the wind model (HWM).

Key Words: infrasound, atmospheric models, statistical performance

OBJECTIVE

The Comprehensive Nuclear-Test-Ban Treaty (CTBT) global infrasonic network will permit the detection, identification, and location of natural and man-made impulsive events. The ability to characterize and locate sources with infrasound data is related to the accuracy of atmospheric state estimates used in the propagation calculations. Conversely, because infrasound propagation depends on the atmospheric conditions, signals from naturally occurring (e.g. mirco-baroms, volcanic activity) and man-made infrasonic sources (e.g. mining blasts) can be used to improve and validate local atmospheric conditions later used to investigate a clandestine event.

The atmosphere is variable on all length and time scales making it challenging to measure and model. Of the available atmospheric models, the Horizontal Wind Model (HWM-93) and Mass Spectrometer and Incoherent Scatter Radar Extended Model (MSISE-90) provide the best and most readily available time-dependent global specification of the atmosphere from the ground to the thermosphere for infrasound propagation modeling. The HWM-93 model provides vector winds and the MSISE-90 model provides density and temperature. This report highlights the statistical performance of HWM-93 and MSISE-90 empirical atmospheric models and explores how atmospheric variability and modeling biases can influence infrasound-modeling calculations.

To provide global specifications of the atmospheric state variables over the entire infrasonic propagation domain the HWM-93 and MSISE-90 models draw upon a forty-year, multi-instrument database of satellite- and ground-based atmospheric measurements. The models internal mathematical formulations, sets of spherical and vector spherical harmonic basis functions and corresponding empirical coefficients are used to provide statistical estimates of the atmospheric state variables at any given location, day-of-year, time-of-day, and level of solar and geomagnetic activity. The two models, FORTRAN source codes, are available for download at http://uap-www.nrl.navy.mil/models_web.

We performed this study to determine if the current NRL empirical models provide sufficiently accurate atmospheric specifications for CTBT verification purposes. The limitations of these two atmospheric models stem from two kinds of uncertainty – one is statistical and one systematic. The statistical uncertainties are caused by unresolved random fluctuations in the upper atmosphere such as gravity waves, planetary waves, and turbulence. The systematic errors are caused by observation and modeling biases.

RESEARCH ACCOMPLISHED

Atmospheric model performance evaluation can be divided into two categories; objective statistical measures and consistency checks against physical laws. For this investigation, a database of standard deviations and mean biases was created to identify systematic model errors and tendencies. In addition, this database can be used in a number of ways to supplement the empirical models.

Model bias $<\varepsilon>$ describes whether a model is over- or under-predicting atmospheric behavior. Bias is defined as the average of the difference between a subset of N observations (d_i) and the corresponding model predictions (m_i). The subset of observations $\{d_0 \ldots d_N\}$ are chosen from a larger set over some arbitrary, yet statistically meaningful sample space (a bin). Although bias is an easy measure to compute and understand, it is often necessary to make an objective statement as to whether a given bias is significant. Additional measures must, therefore be considered. RMSE is defined as the positive square root of the mean square error (σ^2_ε). This is also called the standard error of estimate and written as:

$$\sigma^2{}_\varepsilon = <(d_i - m_i)^2> = \frac{1}{N-1}\sum_{i=1}^{N}(d_i - m_i)^2 \, . \qquad (1)$$

Here $< >$ denotes the ensemble average of N observations and model predictions. RMSE is straightforward to calculate, but is sometimes difficult to interpret. This is especially true in the upper atmosphere where spatial and temporal variability is significant and tends to dominate the measure. In other words, RMSE is sensitive to the natural variability of the observed field. The natural variability of the upper atmosphere is also highly dependent on geographic region and season. Standard deviation is a statistical measure of the variability about the mean in a series of observations. It is the positive square root of the variance σ^2_d, which is defined as:

$$\sigma^2{}_d = <(d_i - <d>_i)^2> = \frac{1}{N-1}\sum_{i=1}^{N}(d_i - <d>_i)^2 = <d_i^2> - <d>_i{}^2 \, . \qquad (2)$$

Other statistical analysis techniques such as anomaly correlation, principal component analysis, and multidimensional spectral analysis are useful in evaluating and improving the atmospheric models. As part of this project, we developed software tools to calculate, investigate, and utilize these statistical performance measures. These analysis tools will be utilized in our ongoing research activities.

STRATOSPHERIC STATISTICS (0 – 50 km)

The first set of statistical performance measures for the HWM/MSIS models was calculated in the 0 to 50 km region of the atmosphere using eight years of near-real-time global meteorological specifications produced by the United Kingdom Meteorology Office. These specifications include winds (U, V), temperature (T), and geopotential height (z) at 12:00 UT on a 73 x 96 degree horizontal grid on 18 pressure levels from 1000 to 0.4 millibars (mb). The 0.4 mb upper pressure level occurs at an approximate altitude of 50 km in the winter hemisphere and 55 km in the summer hemisphere. Our first step was to regrid the data from a pressure grid to an altitude grid that ranged from 0 to 50 km at 2.5 km intervals (21 levels). Then from these results model bias, RMSE, and standard deviation were calculated at 15-day intervals. Measures of RMSE were not calculated for the entire data set because the synoptic and sub-synoptic variability contained within the UKMO analysis were found to dominate RMSE.

It is known that a quasi-biennial oscillation (QBO) exists in the zonal (east-west) wind component of the equatorial stratosphere (*Naujokat*, 1986). The role that this oscillation plays in the general circulation and random variability of the atmosphere is an active area of upper atmospheric research. This feature is currently not included within the HWM wind model. As a result, a coherent model bias in the zonal wind field, with average magnitudes around 20 m/s, exists over a broad altitude region roughly ± 15° around the equator. This is shown in Figure 1.

In addition to biases caused by the QBO, other biases were discovered in the 0 to 55 km portion of the HWM model. Figure 2 shows two slices of the zonal mean wind bias over the same time span for two latitudes (45° S and 45° N). These results show that HWM does a reasonable job of predicting 15-day averages of the stratospheric winds in the northern hemisphere. However, localized stratospheric jet enhancements and planetary wave modulation can introduce significant model biases, especially in the southern hemisphere.

Atmospheric wave variance is a function of latitude, longitude, altitude, and season. In the stratosphere, the majority of wave energy resides in disturbances with horizontal wave numbers from 1 to 5 that migrating eastward or westward with periods ranging from 2 to 45 days. These waves have vertical wavelengths of 4 to 25 km and tend to grow exponentially with altitude. These waves are more frequent and intense in the northern hemisphere. Figure 3 shows the average 15-day standard deviations (1993-1995) of the zonal wind in m/s for four representative months. The temperature, pressure, and meridional wind fields also exhibit similar characteristics. The causes of these features are the mid-latitude jet stream fluctuations, the wintertime stratospheric zonal jet fluctuations, and wave growth with altitude.

In addition to the latitudinal and altitudinal variations of planetary waves, a fair amount of longitudinal variation also exists due to topology. Figure 5 shows representative months of a five-year average of 15-day standard deviations for the 50 km zonal wind field. The other geophysical fields exhibit behaviors

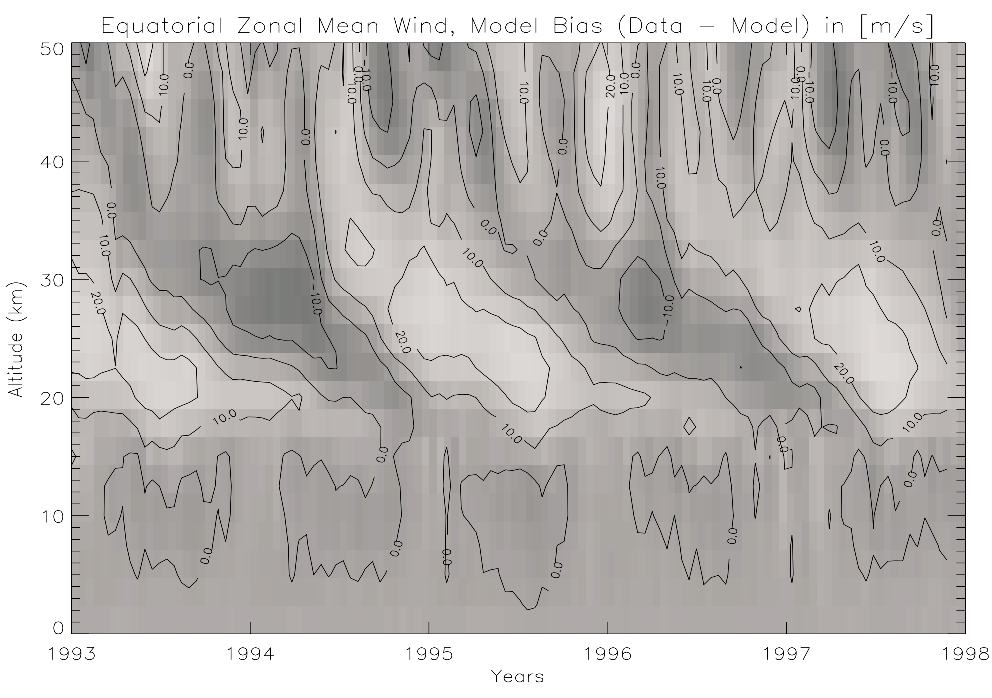

Fig. 1 – The equatorial zonal mean wind bias in m/s showing the effect of the missing stratospheric quasi-biennial oscillation (QBO).

similar to those in figure 4. The NRL-HWM/MSIS statistical performance database contains these fields at 15-day interval over the years 1993 to 1997, for the wind, temperature, and pressure fields from 0 to 50 km at 2.5 km intervals.

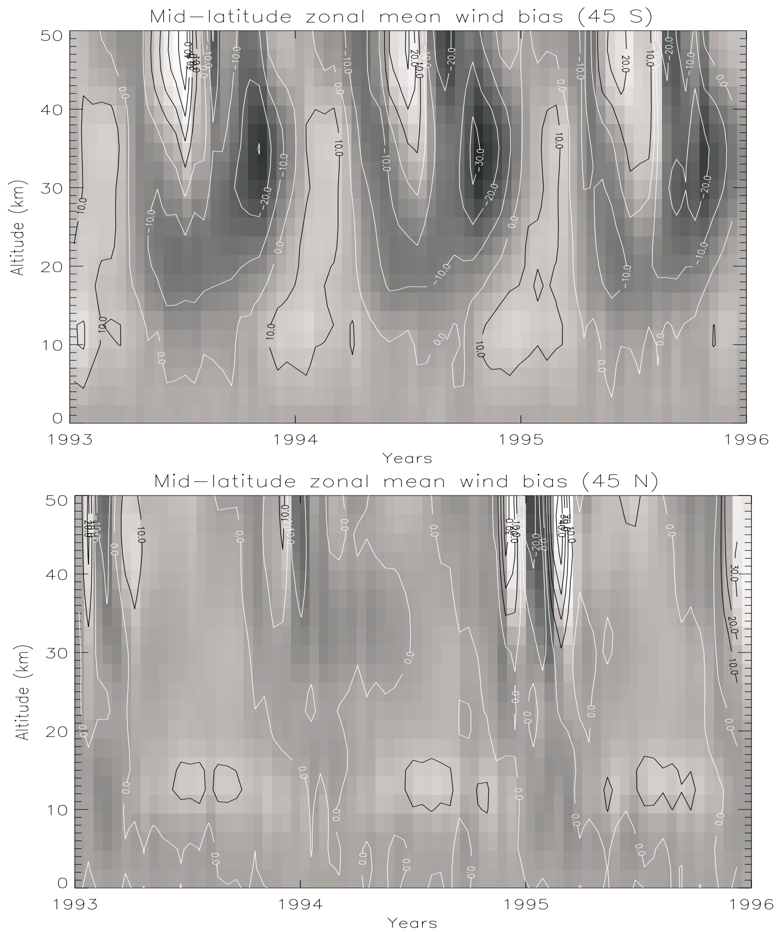

Fig. 2. – Zonal mean winds biases at 45° S (upper panel) and 45° N (lower panel).

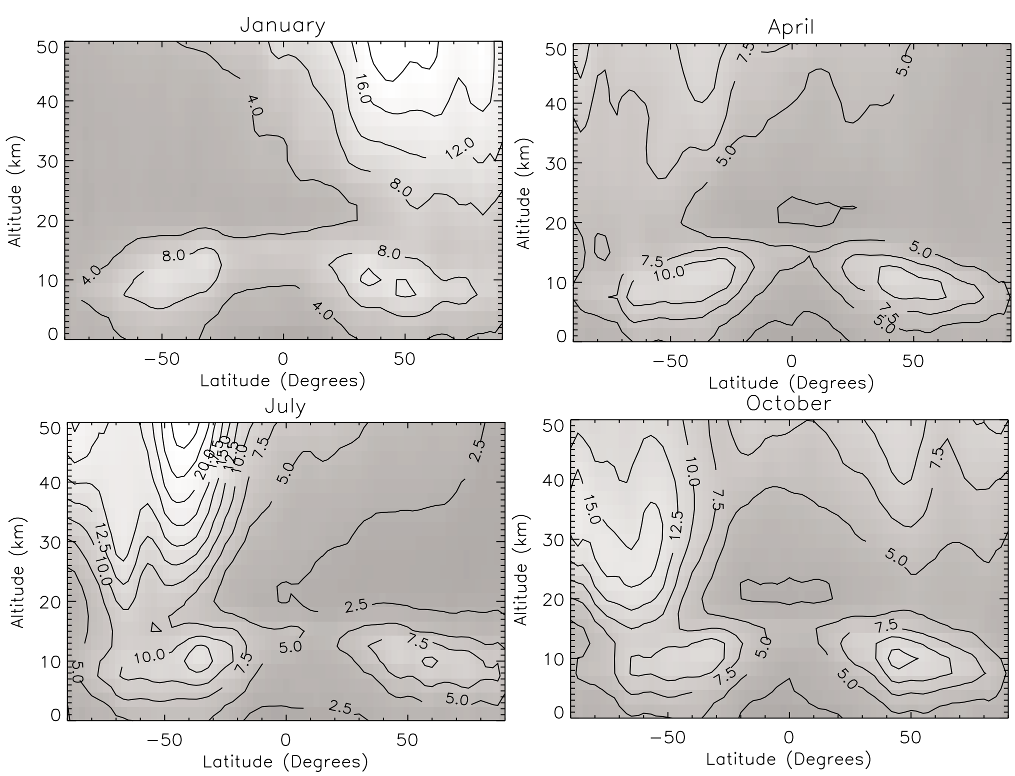

Fig. 3 – Representative months for 5-year climatological averages of the 15-day zonal wind standard deviations in units of m/s.

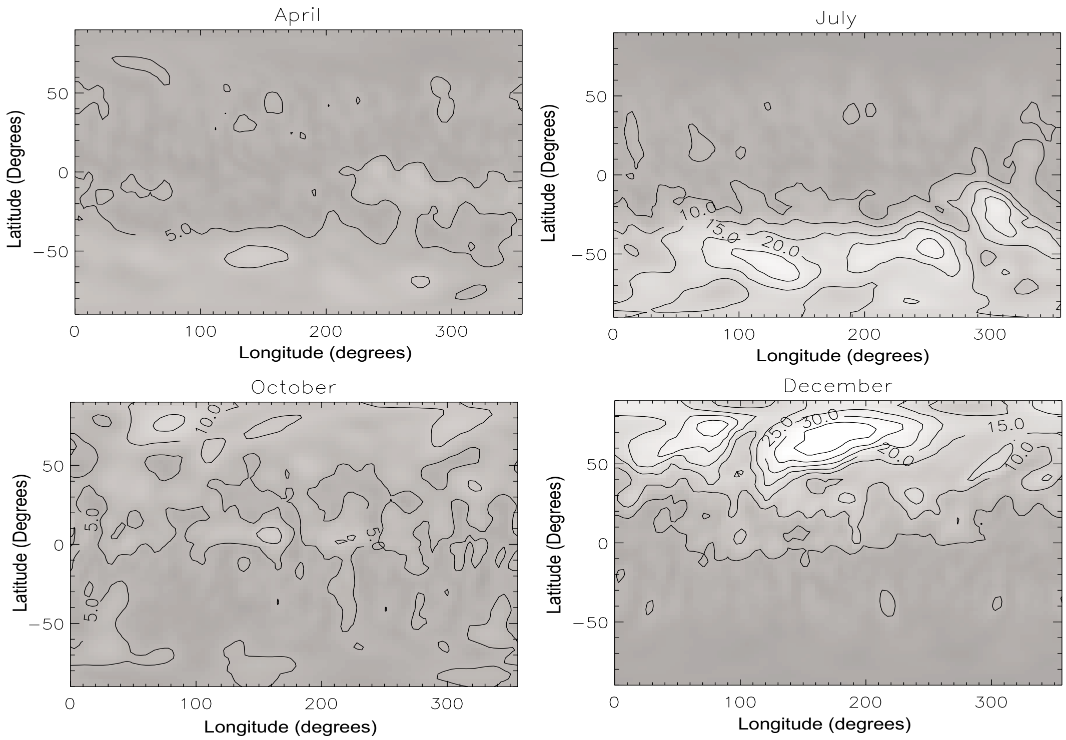

Fig. 4 – Climatological averages of the 15-day merdional wind, standard deviations in m/s at an altitude of 50 km. Four representative months are shown.

MESOSPHERE AND LOWER THERMOSPHERE STATISTICS (60 – 120 km)

The predominant variations in the Mesosphere and Lower Thermosphere (henceforth, MLT) are caused by direct diurnal solar heating of atmospheric ozone and water vapor. The resulting oscillations have harmonics of 24 (diurnal), 12 (semidiurnal), and 8 hours (terdirunal). The amplitudes and phases of these oscillations can vary significantly with latitude and season. On seasonal time scales, changes in the global general circulation pattern are also important. In addition, global scale planetary- and sub-synoptic gravity-waves are superimposed on the atmospheric general circulation and tidal oscillations. Reviews of upper atmospheric dynamics may be found in *Andrews et al.* (1987) and *Rees et al.* (1989).

Compared to the amount of data obtained daily for the atmosphere below 50 km, only a small amount of MLT data is available on a routine basis. Furthermore, the statistical uncertainties of these data are large, the sampling sparse, and the inherent natural variability large. No 'operational' weather monitoring systems currently exist for specifying the state of the MLT region.

Since the last update of the HWM and MSIS models, however, the NASA Upper Atmospheric Research Satellite (UARS) has significantly advanced our understanding of the dynamics in the MLT region. The UARS – High Resolution Doppler Interferometer (HRDI) was the first research instrument to measure global wind profiles throughout the MLT region (*Hays et al.*, 1994). The WINDII instrument onboard the UARS satellite also measured atmospheric wind profiles in the 85 to 180 km region (*Shepard et al.*, 1993). The UARS data provides a good opportunity to evaluate and improve the performance of the HWM model. In addition to obtaining the HRDI data set, we also obtained over one year of ground-based wind measurements from the Bribe Island, MF-radar (28º S, 153º E).

HWM-93 Performance

For the HWM model, the HRDI level 2-B wind measurements (version 11) from 1993 to 1996 were used to calculate a large set of monthly statistical performance measures. Point-for-point HWM wind estimates were first generated. Then zonally averaged, monthly-mean model biases and RMSE estimates where calculated on a 5° latitude by 2.5 km altitude grid that ranged from ± 65° and covered an altitude range from 52.5 to 117 km.

Meridional (north-south) wind biases between the HWM model and HRDI data for 1993 are shown in Figure 5. The model does a good job of reproducing the merdional wind measurements on a monthly-

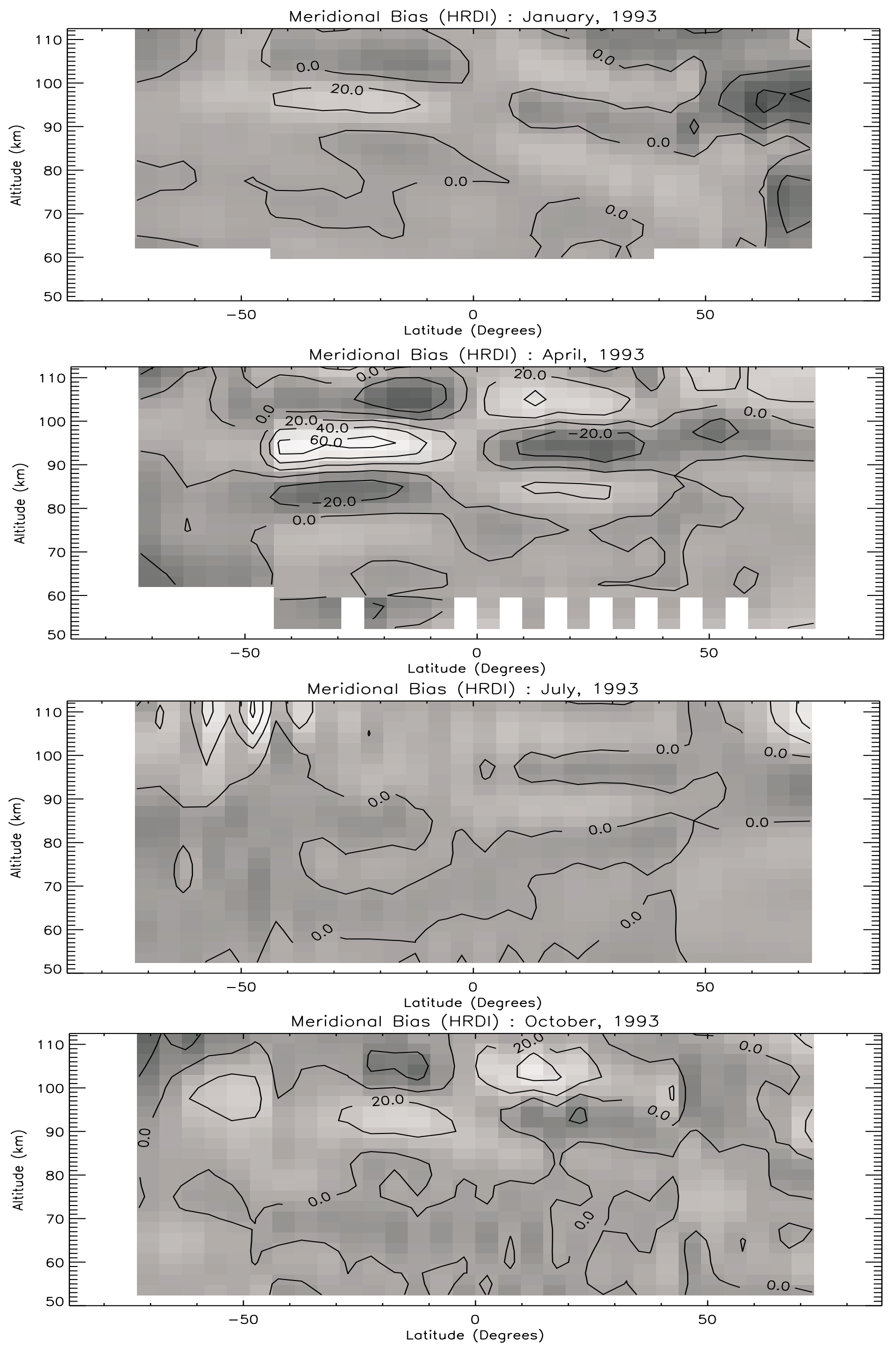

Fig 5. – Average meridional wind biases between the HRDI measurements and HWM model for representative months in 1993.

average basis, but there are notable exceptions. Systematic discrepancies appear in the sub-tropics during the equinox periods with magnitudes approaching 60 m/s near 90 km. These biases are caused by the misrepresentation of the diurnal (1-1) migrating solar tide within the model. Further analysis of the Bribe Island MF data tends to support this conclusion. While the HWM-93 mathematical formulation is capable of representing this component, the early data sets used to estimate diurnal tidal amplitudes coefficients in the MLT region of the model was somewhat limited.

Calculations of RMSE provide additional information on the performance and functionality of the HWM model. In general, this measure reflects atmospheric wave variances not represented in the empirical models. Between the altitudes of 50 and 90 km, typical RMSE magnitudes range from 20 to 30 m/s. Above 90 km, RMSE grows rapidly to 60 m/s. Regions of high RMSE at these altitudes also appear to be clustered around the equator and subtropics. In addition, there are several episodic RMSE enhancements at mid-latitudes near 70 km in the zonal wind field. These are observed at random times throughout the analysis period.

MSISE-90 Performance

It is now accepted that the MSISE-90 model, as well as the CIRA-86 international reference atmosphere (*Fleming et al.*, 1988), both underestimate the mesopause region (80 – 100) temperatures by 15 K at mid-latitudes and overestimate high-latitude mesopause region temperatures by about the same amount (*She and von Zahn*, 1998, *Lubken et al.*, 1999). We used two ground-based instruments and the HRDI temperatures measurements to perform an evaluation of the MSISE-90 model near the mesopause region (80 to 110 km). Because our findings are consistent with those published in the literature, we will not discuss the MSISE-90 mesopause region temperature biases in any detail here.

More importantly, we have recently completed the generation of NRLMSIS-00. In this new version of the MSIS model, we have added a significant amount of new data to the 120 to 600 km region. Changes to the model formulation were made at 120 km and several new physical constraints were used during the model generation process. These adjustments had several positive consequences on model accuracy as far down as 70 km. Using the data collected for this study, initial analysis indicates that NRLMSIS-00 is significantly better than MSISE-90 for CTBT verification and compliance applications. We anticipate a public release of the new model by the end of this year.

CONCLUSIONS AND RECOMMENDATIONS

The objective of this study was to: 1) define weaknesses in the current models for planning and implementation of future upgrades, 2) provide users with information about the models, and 3) develop a set of statistical uncertainties to supplement the models.

The dominant fluctuations of the upper atmosphere are primarily driven by direct solar heating. Because the resulting periodic wave processes are statistically repeatable from day to day and year to year, they can be parameterized with empirical models. The atmosphere also contains significant random and pseudo-random components. In the absence of direct observations, these components cannot be deterministically resolved or predicted with an empirical model. Their occurrence frequency and net influence, however can be parameterized with a statistical measure like standard deviation. This work represents the first step toward a global parameterization of atmospheric random variability that will extend from the ground to 120 km.

The systematic errors we discovered in the two models can exist for a number of reasons. For example, there are regions of the atmosphere were only a handful of data are available. In these regions, the model output represents an interpolation or extrapolation of data from other regions in space and time. While the use of physical constraints such as hydrostatic equilibrium and geostrophic balance can help reduce uncertainties, errors in the model predictions can still be significant. Additionally, the spectral fidelity of the models is determined by the density of the available data. Using our new database this can be subsequent improved in future model versions.

The set of finite basis functions used in the models to represent the atmosphere provides another source of error. While a given feature may be parameterized mathematically and resolved with the data, the feature may be unknown or simply ignored for computational reasons. Adjustment of the model formulation and assimilation of available data can rectify these errors.

A final source of systematic errors in our models comes from the data themselves. Almost 99.99% of all upper atmospheric measurements are made by indirect means, i.e. remote sensing techniques. Errors associated with instrumentation and poor data inversion procedures all contribute to the possibility of systematic, nonphysical errors in the observations. Owing to the fact that the upper atmospheric measurements are difficult to make, there are often no other coincident measurements available to provide consistency checks, especially where they are most needed.

Based on our findings there is a need for HWM model improvements. When compared to the recently available data sets, the historical database used to determine the present HWM empirical model coefficients was limited. For this reason, and minor limitations in the present model formulation, tidal amplitudes in the HWM-93 model have been underestimated in certain regions and seasons. Due to the lack of a parameterization of the equatorial quasi-biennial oscillation, systematic errors in the zonal wind field from 35 to 120 km ranging from 20 to 50 m/s are also present. Unlike random errors due to transient wave phenomena, these two errors have large temporal and spatial coherence. Under a significantly broad range of conditions, these errors will impact science and engineering calculations that use current versions of the HWM model to provide wind estimates.

This study illustrates that from 0 to 55 km, the models are inferior to operational data products available from a number of national and international Numerical Weather Prediction (NWP) centers. However, for analysis of early historical data, or in situations where NWP estimates and/or direct meteorological observations are not available, these models are extremely useful. Furthermore, these models are superior to other public domain atmospheric climatologies owing to their extension to higher altitudes and the ability to be directly embedded into engineering and scientific codes to produce estimates with minimal computation requirements.

The effect of the above mentioned limitations on infrasound propagation calculations for CTBT monitoring was recently illustrated in a study by *Garces et al*. (1999). In this study, we produced atmospheric estimates from the ground to 150 km for the entire year of 1996 over the Alaska and Hawaii CTBT infrasound sites by carefully merging the UKMO analysis data with the HWM and MSIS models. To do this we developed a methodology called the NRL-CAMPFIRE procedure (Combined Atmosphere Measurement Profiles For Infrasound Range Estimation). This method works well under most conditions; however, where large differences between the empirical models and the stratospheric analysis products occur, spurious ringing at the model boundaries can result. This is particularly true at high latitudes, where HWM often underestimates the localized velocity of the winter polar stratospheric jets by as much as 60 m/s. As expected, this is less of a problem over Hawaii, where the upper-stratosphere wind velocities are smaller.

Given the lack of accuracy of the HWM model in the stratosphere and the lack of operational meteorological specifications into the thermosphere, the above effects could be viewed as a minor inconvenience. However, this problem occurs close to the region where stratospheric infrasound reflections take place. Because infrasound rays are sensitive to errors and perturbations near their turning points, errors in the procedure will translate in to errors in infrasound propagation calculations. The rectification of the systematic biases in the current models would suppress these effects in the CAMPFIRE procedure by greatly reducing the differences between the current HWM model and NWP stratospheric analyses. We are now in the process of using our modeling tools to address this problem.

In this document, we have alerted current and potential HWM/MSIS users to weaknesses in the models. For those users that require additional information we have created a performance measures database. Having now identified several problematic areas, this information can be used to develop and add missing parameterizations to the models. Additionally, the new database of standard deviations can be used to weight the various datasets during the model building process. Furthermore, these new measures will be useful as a-priori statistics in making improvements to the CAMPFIRE procedure.

REFERENCES

Andrews, D. G., J. R. Holton, and C. B. Holton, “Middle Atmospheric Dynamics”, Academic Press, Orlando, Florida, 1987.

Burrage, M. D., M. E. Hagan, W. Skinner, D. L. Wu, P. B. Hays, Long-term variability in the solar diurnal tide observed by HRDI and simulated by GSWM, *Geophys. Res. Lett.*, **22**, 2641-2644, 1995.

Burrage, M. D., R. A. Vincent, H. G. Mayr, W. R. Skinner, N. F. Arnold, and P. B. Hays, Long-term variability in the equatorial middle atmosphere zonal wind, *J. Geophys. Res.*, **101**, 12847-12854, 1996.

Burrage, M. D. *et al.*, Validation of mesosphere and lower thermosphere winds from the high resolution Doppler imager on UARS**,** *J. Geophys. Res.*, **101**, 10365-10392, 1996.

Fleming, E.S., S. Chandra M.R. Schoeberl, and J.J. Barnett, Monthly mean global climatology of temperature, wind, geopotential height and pressure for 0-120 km, NASA technical memorandum, 100697, 1988.

Forbes, J. M., M. Kilpatrick, D. C. Fritts, A. H. Manson, and R. A. Vincent, Zonal mean and tidal dynamics from space: an empirical examination of aliasing and sampling issues, *Ann. Geophys.*, **15**, 1158-1164, 1997.

Garces, M, R. Hansen, K. Lindquist, D. P. Drob, and J. M. Picone, Variability of the lower atmosphere and its effects on infrasonic waves, Proc. 21 Annual Seismic Symposium on Monitoring the CTBT, 1999.

Garces, H. D. P. Drob, J. M. Picone, A theoretical study of the effects of geomagnetic fluctuations and and solar tides on the propagation of infrasounic waves in the atmosphere, *Geophys. J. Int.*, submitted, 1999.

Hays, P. B., D. L. Wu, Observations of the Diurnal Tide from Space, J. Atmos. Sci., **51**, 3077-3093, 1994.

Hedin, A. E., N. W. Spencer, and T. L. Killeen, Empirical global model of the upper thermopshere winds based on Atmopshere and Dynamics Explorer satellite data, *J. Geophys. Res.*, **93**, 9959-9978, 1988.

Hedin, A. E., M. A. Biondi, R. G. Burnside, G. Hernandez, R. M. Johnson, T. L. Killeen, C. Mazaudier, J. W. Meriwether, J. E. Salah, R. J. Sica, R. W. Smith, N. W. Spencer, V. B. Wickwar, and T. S. Virdi, Reviesed global model of upper thermospheric winds using satellite and ground-based observations, *J. Geophys. Res.*, **96**, 7657-7688, 1991.

Hedin, A. E., E. L. Fleming, A. H. manson, F. J. Schmidln, S. K. Avery, R. R. Clark, S. J. Franke, G. L. Fraser, T. Tsuda, F. Vial, and R. A. Vincent, Empirical wind model for the upper, middle and lower atmosphere, *J. Atmos. Terr. Phys.*, **58**, 1421-1447, 1996.

Hedin, A. E., M. J. Buonsanto, M. Cordrescu, M.-L. Duboin, C. G. Fesen, M. E. Hagan, K. L. Miller, and D. P. Sipler, Solar activity variations in mid-latitude thermospheric meridional winds, *J. Geophys. Res.*, **99**, 17601-17608, 1994.

Lubken, F. J., Thermal Structure of the Artic Summer Mesopause, J. Geophys. Res., **104**, 9135-9149, 1999.

Naujokat, B., An update of the observed quasi-biennial oscillation of the stratospheric winds over the tropics, *J. Atmos. Sci.*, **43**, 1873-1877, 1986.

Rees, M. H., “Physics and chemistry of the upper atmosphere”, Cambridge University Press, New York, NY, 1989.

She, C.Y., U. von Zahn, Concept of a two-level mesopause: Support through new lidar observations, *J. Geophys. Res.*, **103**, 5855, 1998.

Shepard, G. et al., WINDII, the Wind Imaging Interferometer on the Upper-Atmosphere Research Satellite, *J. Geophys. Res.*, **98**, 10725-10750, 1993.