Orbit-centered atmospheric density prediction using artificial neural networks

CrossMark

David Pérez [a], Brendt Wohlberg [b], Thomas Alan Lovell [c],
Michael Shoemaker [d], Riccardo Bevilacqua [e,*]

[a] Mechanical, Aerospace and Nuclear Engineering Department, Rensselaer Polytechnic Institute, JEC 1034 110 8th Street, Troy,
NY 12180, USA
[b] T-5, Los Alamos National Laboratory, MS-B284, Los Alamos, NM 87545, USA
[c] Air Force Research Laboratory, Space Vehicles Directorate, Kirtland AFB, NM, USA
[d] ISR-1, Los Alamos National Laboratory, P.O. Box 1663, Mail Stop D466, Los Alamos, NM 87545, USA
[e] Mechanical, Aerospace and Nuclear Engineering Department, Rensselaer Polytechnic Institute, JEC 5048 110 8th Street, Troy,
NY 12180, USA

A R T I C L E   I N F O

A B S T R A C T

At low Earth orbits, drag force is a significant source of error for propagating the motion of
a spacecraft. The main factor driving the changes on the drag force is neutral density.
Global atmospheric models provide estimates for the density which are significantly
affected by bias due to misrepresentations of the underlying physics and limitations on
the statistical models. In this work a localized predictor based on artificial neural
networks is presented. Localized refers to the focus being on a specific orbit, rather than
a global prediction. The predictor uses density measurements or estimates on a given
orbit and a set of proxies for solar and geomagnetic activities to predict the value of the
density along the future orbit of the spacecraft. The performance of the localized predictor
is studied for different neural network structures, testing periods of high and low solar
and geomagnetic activities and different prediction windows. Comparison with previously
developed methods show substantial benefits in using artificial neural networks, both in
prediction accuracy and in the potential for spacecraft onboard implementation. In fact,
the proposed neural networks are computationally efficient and would be straightforward
to integrate into onboard software.

1. Introduction

Due to their ease of accessibility, low Earth orbits (LEO) contain the majority of artificial satellites currently in operation. At LEO below 700 km, atmospheric drag is the most significant force acting on spacecraft after gravity. Given that atmospheric drag is not easy to estimate, it constitutes the largest source of error force models. The drag force is a function of several time varying factors, such as atmospheric winds, drag coefficient, and density. However, the largest variations in the drag force are caused by changes in the atmospheric density, as the spacecraft flies through different regions of the thermosphere with different densities, and also as those densities fluctuate in response to solar and geomagnetic activity. Consequently, precise models for the density are necessary for accurately estimating the drag force, which in turn is necessary for precise onboard orbit determination. Reliable onboard orbit determination will be a key factor in the development of better methods for maneuver planning and coverage calculations. Furthermore, in the past 30 years starting with the work of Leonard et al. in Ref. [1] there has been an increasing body of work focusing on using the drag

* Corresponding author.
E-mail addresses: perezd4@rpi.edu (D. Pérez),
brendt@lanl.gov (B. Wohlberg),
thomas.lovell@kirtland.af.mil (T.A. Lovell),
shoemaker@lanl.gov (M. Shoemaker), bevilr@rpi.edu (R. Bevilacqua).

force for maneuvering spacecraft in LEO [2–5]. Accurate onboard estimation of the density can be used to improve some of the methods proposed for maneuvering with the drag force, since it will provide the controllers with an accurate estimate of the control force.

Over the past 50 years several different global atmospheric models have been developed for calculating the main characteristics of the thermosphere including density (a summary of the different models available is presented by Vallado in [6] chapter 8.6.2.). Global models can be classified into empirical and physics-based models. The seminal work for empirical global atmospheric models is Jacchia's 1960 [7] model, which uses an empirical formula that estimates the density as a function of the geometric height, the 20-cm solar flux (F_{20}) and the angular distance to the center of the diurnal solar bulge. Further improvements of this approach include Jacchia models from 1971[8], 1977 [9], and up to Jacchia–Bowman 2006 (JB2006) [10] and 2008 (JB2008) [11]. The High Accuracy Satellite Drag Model (HASDM) uses calibration data from up to 75 inactive satellites and a Dynamic Calibration of the Atmosphere (DCA) method to correct older models such as the Jacchia models [12]. DCA methods use available current measurements to correct the current density estimate; an example of one of these methods developed by two of the authors can be seen in ref [13]. Another highly used empirical global model is the Mass Spectrometer and Incoherent Scatter Radar model (MSIS-77) [14]. MSIS-77 uses data from satellites and also from ground-based measurements from incoherent scatter radars to estimate density. Several improvements to the original MSIS from 1977 have been made, including MSIS-86 [15], MSISE-90 ([16]), and NRLMSISE-00 developed by the U.S. Naval Research Laboratory [17]. An additional empirical model is the Drag Temperature Model (DTM) [18] developed in terms of spherical harmonics, using data covering nearly two solar cycles. This model has been further developed as DTM-94 [19] and DTM-2000 [20].

Global circulation models (physics-based models) are an alternative to the global empirical models for predicting the density. Among these is the thermosphere–ionosphere–mesosphere–electrodynamics General Circulation Model (TIME-GCM) [21]. This model calculates the global circulation, temperature and compositional structure with coupled electrodynamics. An additional global circulation model is the Coupled Thermosphere–Ionosphere–Model (CTIM) [22]. CTIM is a time dependent, nonlinear model that consists of the union of two elements: a neutral thermospheric model and a mid and high latitude ionospheric convection model. CTIM was further developed by including a model of the plasmasphere and low latitude ionosphere, thus producing the Coupled Thermosphere–Ionosphere–Plasmasphere model (CTIP) [23]. Later on the Coupled Thermosphere–Ionosphere–Plasmasphere Electrodynamic model (CTIPE), presented in Refs. [24] and [25], was created by combining an electrodynamic model with CTIP. The Coupled Middle Atmosphere and Thermosphere model (CMAT) [26] and its updated version (CMAT-2), first applied in [27], are extensions of CTIP developed at the University College London. Another global circulation model is the Global Ionosphere–Thermosphere Model (GITM) [28], developed at the University of Michigan.

GITM consists of a three dimensional spherical code that solves the energy, momentum and continuity equations.

Global atmospheric models are often designed to estimate much more than just the density, which unfortunately results in longer computation times and less accurate results for a specific quantity such as density. Furthermore, the physics can be misrepresented in the case of the physical models, while the data used for generating the empirical ones can be limited. These three factors result in errors in the prediction of the local density. Furthermore, the physics based models are computationally expensive and require several real-time inputs, which hampers onboard calculations. For these reasons it is desirable to use a different approach for designing a density predictor capable of running onboard a satellite.

An alternative originally proposed by Stastny et al. in Ref. [29] is a localized density model. Such an approach consists of limiting the model to estimate only the density along the orbit of a single spacecraft. By introducing these restrictions, the ability of the model to accurately estimate the density is greatly enhanced. Provided that measurements or estimates of the density of the medium around the spacecraft are available on-board, time series forecasting techniques can be used to predict the future density along the orbit of the spacecraft. In their work, Stastny et al. [29] used a linear model as the predictor and showed that such a model provided accurate results, with less bias than two of the latest empirical models (HASDM and JB2006) for predicting one orbit into the future.

A similar approach to that of Stastny et al. is used in this work. However, instead of using a linear model as the predictor, artificial neural networks (ANNs) are used. A neural network is capable of forecasting nonlinear behaviors since it contains nonlinearities in its neurons, and therefore it has the potential to accurately model the nonlinear behavior of the density along the orbit of the spacecraft. To train, validate, and test the neural networks, density data from the CHAllenging Minisatellite Payload (CHAMP) [30], mission was used.

The foremost contributions of this work are

1) Development of neural network-based localized models for the density that are capable of forecasting the density to be encountered by a spacecraft along its orbit for prediction windows of one, eight and 32 orbits into the future (i.e. approximately 90 min, 12 h and two days respectively).
2) Appropriate design of the neural network structure using different parameters such as the sampling rate of the data, the number of neurons in the hidden layer and the number of delays of the input.
3) Tests of the neural network predictors over periods of high and low solar and geomagnetic activities.
4) Comparison of the results of the neural network predictors with a simple persistence model, a linear model, JB2006, and HASDM (the latter three obtained from Ref. [29]) for the one-orbit forecast.

The paper is organized as follows: Section 2 presents the concept of atmospheric drag and density. Section 3

provides an introduction to neural networks and explains the neural network architecture used. Section 4 is dedicated to the data used for training, validation, and testing the neural network predictors. Section 5 presents the results of the different tests performed using the neural network predictors and evaluates their performances. Section 6 draws the conclusions.

In addition, the neural networks designed for this work are available as MATLAB files at the Mathworks File Exchange webpage.[1]

2. Drag acceleration

The drag acceleration experienced by spacecraft in LEO orbits is a function of the atmospheric density, the atmospheric winds, the velocity of the spacecraft relative to the medium, the geometry, altitude, drag coefficient and mass of the spacecraft. The interdependence of these parameters (e.g. the drag coefficient is affected by the temperature of the medium which also determines the density of the medium) and the lack of knowledge in some of their dynamics make the modeling of the drag force a challenging problem.

The magnitude of the aerodynamic drag force is usually expressed as (Ref. [31])

$$F_d = a_d m = \frac{1}{2} m \rho B v_s^2 \qquad (1)$$

where F_d is the magnitude of the drag force, a_d is the magnitude of the drag acceleration, m is the spacecraft mass, ρ is the density of the local atmosphere, v_s is speed of the spacecraft relative to the medium and the ballistic coefficient B is given by

$$B = \frac{C_D A}{m} \qquad (2)$$

where C_D is the drag coefficient and A is the cross-sectional area perpendicular to the atmosphere's relative velocity. One must properly estimate the density, drag coefficient, and atmospheric wind to forecast the drag acceleration. The mass and the cross-sectional area are known by design, provided that the altitude of the spacecraft is known. Once all the parameters are known, Eq. (1) can be used to obtain the drag acceleration. This can result in errors when calculating the drag since there will be errors for each of the individually estimated parameters and they will accumulate. As pointed out in [31], Eq. (1) has limited application at LEO, but still it is useful for showing the main parameters that determine the drag force.

Another alternative would be to directly forecast the drag acceleration. However, as can be observed from Fig. 1, once normalized the different elements in the drag force (ρ, ρC_D and F_d) show a very high degree of agreement. This is confirmed by the Pearson correlation coefficient (which measures the linear dependence between two values) which are 0.9948, 0.9998 and 0.9941 for the pairs (ρ, ρC_D), (ρ, F_d) and (F_d, ρC_D) respectively. Since these values are very close to one, it can be said that there is a

significant linear correlation between the pairs. This indicates that, if properly designed, a predictor could be used for modeling any of these three values and then a scaling factor could be used to obtain the other two values with a relatively low error. Since values for density are widely available to the scientific community from the CHAMP and the Gravity Recovery and Climate Experiment (GRACE) [32], missions, density was selected for the developments of this work.

2.1. Density in the thermosphere

At the thermosphere (80–640 km as defined in Ref. [33]), the solar activity creates large variations in temperature, and therefore in density, through two related mechanisms illustrated in Fig. 2. The first of them is the Extreme Ultra Violet (EUV) radiation emitted by the sun. This radiation fluctuates depending on the solar activity and deposits energy in the thermosphere through excitation, dissociation or ionization of the molecules in the thermosphere. The second mechanism is the solar wind, which consists of electrically charged particles that are funneled into the auroral zones in the thermosphere by the terrestrial magnetic field. These charged particles contribute energy to the thermosphere increasing the temperature and the density.

2.2. Indices of solar and geomagnetic activities

The EUV radiation and geomagnetic activity play a key role in the temperature and density variations. Solar activity is driven by the 11-year solar cycle and the 27-day solar rotation cycle. A summary of the indices used to represent solar and geomagnetic activities is presented in Table 1. These indices have different sampling rates which can be processed to obtain better temporal resolutions of the solar and geomagnetic activities. For onboard applications such as the predictor presented here, the use of these indices requires the ability to regularly send data to the satellite.

3. Neural networks

A biological brain contains a large network of neurons (computing cells), capable of modeling nonlinear processes in a parallel structure. During infancy human brains already have a plastic structure capable of performing complex computations, but their most remarkable feature is that as experience is accumulated, brains are capable of developing their own rules of behavior based on their evolving environment [42]. During this learning process, the brain takes advantage of its structural elasticity to further develop its structure and functionality to accommodate for performing new tasks required by the changing environment. Artificial neural networks are mathematical systems specially designed to resemble biological brains. In a similar fashion, neural networks are composed of artificial neurons as the one illustrated in Fig. 3. These artificial neurons are nonlinear processing units that are characterized by sets of inputs, outputs, biases, weights and a nonlinear transfer function.

---

[1] Link http://www.mathworks.com/matlabcentral/fileexchange/45193-neural-networks-for-localized-atmospheric-density-forecasting
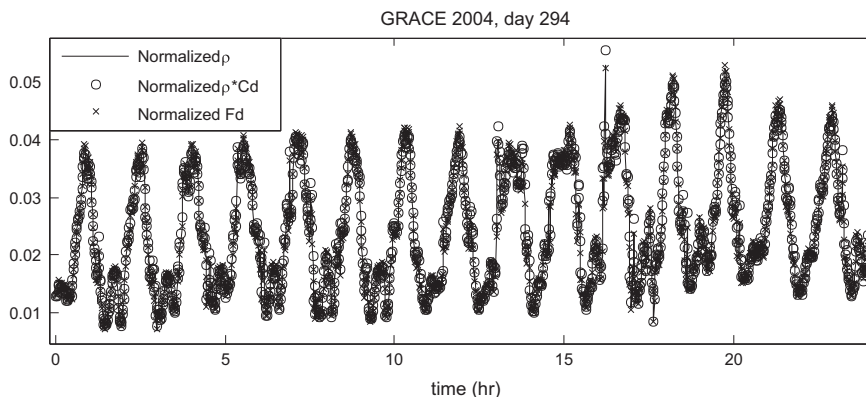
GRACE 2004, day 294

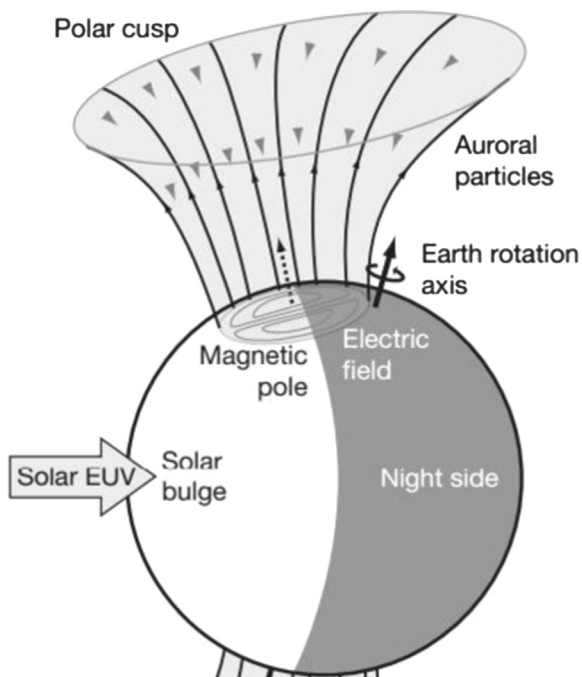Fig. 1. Normalized values for ρ, ρC_D and F_d.

Fig. 2. Mechanisms driving changes in temperature in the thermosphere (obtained from Ref [34]).

The following equation is the mathematical expression for the neuron shown in Fig. 3

$$y = f\left(\sum_{j=1}^{N} W_j x_j + B\right) \qquad (3)$$

where y is the output of the neuron, x_j is the jth input, N is the number of inputs, W_j is the jth weight, f is the nonlinear transfer function and B is the bias.

The artificial neurons are organized in layers, which are structures of artificial neurons arranged in parallel (as shown in Fig. 4). Each layer may have a different numbers of neurons, which allows for different structures that may work better for different applications, as discussed below. Furthermore, the neural networks can include delays (computational devices that store information flowing through the network) and feedbacks. A higher number of neurons in the layers allow for modeling more complex relations; however, the use of too many neurons can result in overfitting. Overfitting (or overtraining) is the condition in which a neural network has retained too much information about the training set (training is explained in the next paragraphs), and consequently it has lost its generalization ability and cannot represent the relations of new inputs and outputs. Conversely, if there are too few neurons, the neural network will not be able to learn the relation between the inputs and outputs, a condition which is known as underfitting.

The possibility of having different structures including delays and or feedback grants neural networks the flexibility to solve problems from many different fields and that computationally have completely different natures. In general, neural networks are capable of modeling nonlinear input–output relations and thus can be applied to classification, data processing, nonlinear systems modeling and time series forecasting among other things.

Through the biases and weights, in their artificial neurons, neural networks are capable of storing information that relates the inputs to the outputs. This information is embedded into the artificial neurons through a learning process (learning algorithm) analogous to the acquisition of experience and the resulting development of their rules of behavior in biological brains. The learning algorithm consists of the adaptation of the weights and biases to minimize some performance function (for example the mean squared error) that measures how well the input–output relation of the specific problem is being represented by the neural network. This process is usually called training, and a data set of inputs and outputs (training set) is selected specifically for this purpose. The minimization of the performance function must be stopped at some point in order to prevent overtraining, i.e., memorization of old data and lack of generalization capabilities. The training process is stopped when one of the following conditions is satisfied: the performance function converges, the maximum number of epochs is reached, or the performance function of the neural network for a different data set (validation set) reaches a minimum. That way, the ability of the neural network to generalize to new data is preserved. The selection of the

Table 1
Summary of solar and geomagnetic indices.

| Nomenclature	Parameter measured	Measured from	Description	Source	Periodicity
Dst (from refs [35] and [36])	Magnetic field at 4 points of Earth's equator	Four magnetic observatories – Hermanus, Kakioka, Honolulu, and San Juan are used	Indicates disturbances on the magnetic field at the Earth's dipole	World Data Center	Hourly
$F_{10.7}$ (from refs [37], [38] and [39])	Integrated emission from the solar disc at 2800 MHz (10.7 cm wavelength)	Dominion Radio Astrophysical Observatory, near Penticton, British Columbia	It measures a combination of EUV emissions (chromospheric, transition region, and coronal solar) modulated by bright solar active regions whose energy is deposited in the thermosphere.	Canadian National Research Council's Herzberg Institute of Astrophysics	3 per day
Kp (from refs [40] and [41])	Mean value of the disturbance levels in the two horizontal field component	Observed at 13 selected, subauroral stations	K variations are all irregular disturbances of the geomagnetic field caused by solar particle radiation within the 3-h interval concerned. All other regular and irregular disturbances are non K variations. Geomagnetic activity is the occurrence of K variations.	International Association for Geomagnetism and Aeronomy (IAGA)	Every 3 h
ap (from refs [40] and [41])	Kp is tranformed into a linear scale to obtain ap	Derived from Kp	ap may be regarded as the range of the most disturbed of the two horizontal field components,	International Association for Geomagnetism and Aeronomy (IAGA)	Every 3 h
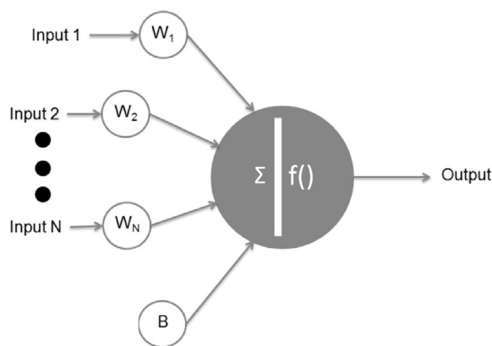
Fig. 3. Artificial neuron diagram.

training and validation data sets is very important; the better they represent the behaviors to be modeled by the neural network, the better the neural network will perform. After the neural network has been trained and validated, it is usually tested on a third group of data points (testing set) to assess the quality of the neural network model. It should be noted that the training and validation sets are used for creating the neural network, while the testing set is used for evaluating its performance.

3.1. Neural network used

A time-delay feed-forward neural network structure was chosen for the development of the density predictor. This neural network architecture does not include any feedback loops, hence the feed-forward part of its name. This neural network architecture contains a set of delays at the input layer that allow for retention of the evolution of the inputs in time, and enhances the ability of the network for forecasting applications. Furthermore, the neural network predictor contains two layers (hidden or input layer, and output layer). The output layer contains one single linear neuron. The number of neurons and delays in the

hidden layer were determined by testing different configurations. The results of these experiments are included in the results shown in Section 5. Fig. 4 shows a Simulink block diagram for a time-delay feed-forward neural network with two delays and three neurons in the hidden layer.

In the context of density prediction, the input to the neural network is the present value of the density and the output is the predicted value over a predefined prediction window. Additional inputs such as the current values of the solar and geomagnetic indices (Dst and $F_{10.7}$) can also be included. For purposes of inputting past density values, the inputs are delayed a defined number of times inside the neural network in order to capture some of their evolution in time. Such formulation is shown in the following expression:

$$\hat{y}(t+W_p) = g \begin{pmatrix} x(t-t_s), \dots x(t-Dt_s), \\ \mathrm{Dst}(t-t_s), \dots \mathrm{Dst}(t-Dt_s), \\ F_{10.7}(t-t_s), \dots F_{10.7}(t-Dt_s) \end{pmatrix}, \quad (4)$$

$$x = \ln(\rho), \quad (5)$$

$$\hat{y} = \ln(\hat{\rho}) \quad (6)$$

where g is the overall nonlinear function of the neural network, ρ is the measured density, $\hat{\rho}$ is the predicted density value (neural network output), W_p is the prediction window, t_s is the sampling period of the data, t is the time, and D is the number of delays in the hidden layer.

The Levenberg–Marquardt algorithm, developed in Refs. [43] and [44], was used to train the neural networks. This algorithm, which is included in MATLAB's Neural Network Toolbox [45] was chosen since it often has higher rates of convergence than the other algorithms provided in the Toolbox. This method solves the least squares problem (in this case it finds the weights that minimize the performance function) using regularization, by interpolating between Newton's and gradient descent methods, thus
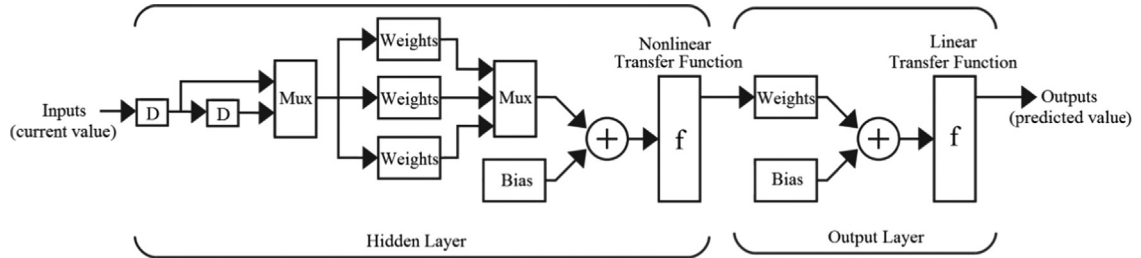
Fig. 4. Diagram for a time-delay feed-forward neural network with two layers, two delays, three nonlinear neurons in the hidden layer and one linear neuron in the output layer.

providing some of the robustness of gradient descent and some of the speed of convergence of Newton's method. The mean squared error (MSE), as explained in Eq. (7), was selected as the performance function.

$$MSE = \frac{1}{n} \sum_{i=1}^{n} (\hat{y}_i - x_i)^2 \qquad (7)$$

where n is the number of samples.

4. Data used

4.1. Density

The use of neural networks requires data sets for training, validation, and testing the model's performance. The training and validation sets must contain data covering the different behaviors to be modeled by the neural network. Over the past 10 years there have been several satellite missions equipped with reliable accelerometers from which density values can be estimated with high temporal resolutions. These data cover periods of high and low solar and geomagnetic activities. Among these missions is CHAMP, which was equipped with high precision accelerometers. CHAMP was launched in 2000 into a circular, almost polar orbit with an initial altitude of 460 km. Sutton et al. [46] presented a method for estimating density and winds from accelerometers such as the ones included in CHAMP. The estimated density from CHAMP, which was used for this work, is available online. The density data used were obtained from Ref. [47]. For the purposes of this work, it is assumed that the publicly available, pre-processed CHAMP data represents true density measurements.

For each neural network the training density data was divided into two segments: one segment of past values, assumed to be available for the training and validation of the neural network; and one segment of future values, which are values of density that would not be available during training and validation, but instead are used exclusively for testing the neural networks. The past values were sampled randomly and 70% were used for training, and the remaining 30% for validation. Furthermore, the available density data were not evenly distributed in time, therefore, for implementing the neural network, a linear interpolation was applied to make sure that there was a constant difference in time between consecutive samples in the data. This method consists on linearly interpolating the data between the available data points at the times
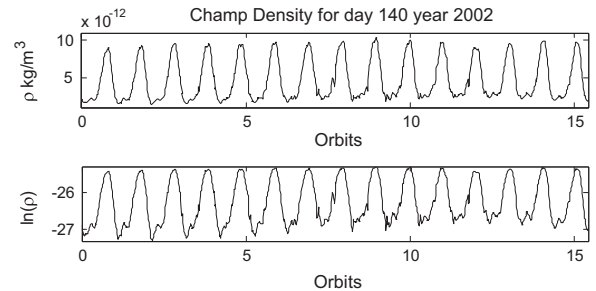

Fig. 5. Density (top) and natural log of density (bottom) from CHAMP day 140 of 2002.

defined by the desired sampling rate. The values of the density are in the order of magnitude of 10^{-12} kg/m^3 for day 140 of 2002 (see Fig. 5). This results in numerical problems during the training of the neural networks. To address this issue, the natural logarithm of the density values was used for the neural networks instead of the density values themselves. Another advantage of using the natural logarithm shown in Ref. [48] is that it often stabilizes the variance of the series, which allows for better modeling of the time series.

Several different periods of interest for training, validation, and testing the neural networks were identified. Stastny et al. [29] chose two representative days for low and high geomagnetic activities for testing his linear model and for comparing it to JB2006 and HASDM. The first of these days was day 141 of 2002; during this day there was very low geomagnetic activity (Dst$=-16$, ap$=10$ and $F_{10.7}=190.4$). The second day used by Stastny et al. was day 276 of 2001; during this day a moderate geomagnetic storm occurred, so there was a higher geomagnetic activity (Dst$=-107$, ap$=69$ and $F_{10.7}=191.8$) For obtaining the linear model, Stastny et al. used the data from day 140 (Dst $=-12$, ap$=10$ and $F_{10.7}=175.4$). This same data set was used to train, validate and test the neural networks. These data sets included $n=1080$ data points for each day with a sampling rate of 80 s.

To study the long-term performance of the neural networks, it was decided to test them over one-year intervals. Out of the years that CHAMP was collecting data, years 2003 and 2007 were certainly very interesting from the point of view of space weather and therefore were selected for testing. During 2003 (Dst $=-22$, ap$=21.8$ and $F_{10.7}=128$) the geomagnetic activity was the highest of that solar cycle [49]. In contrast, during 2007 (Dst $=-8$,
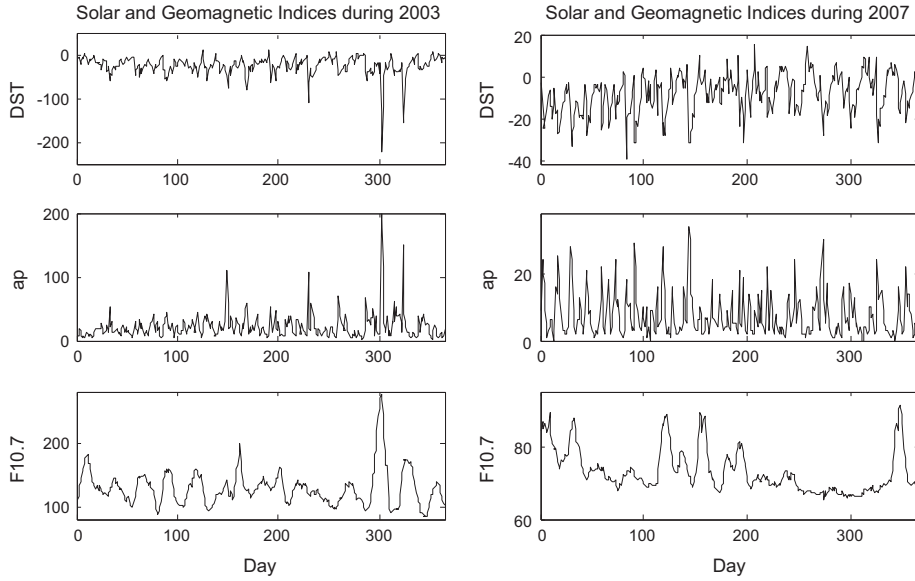
Fig. 6. Dst, ap and $F_{10.7}$ indices averaged daily for years 2003 (left) and 2007 (right).

ap=7.5 and $F_{10.7}$=73) the solar cycle went through a period of very low activity (solar minimum, see Ref. [50]) and therefore the solar and geomagnetic activities were very low. The training and validation sets consisted of data from years previous to the testing years. Specifically, data from 2002 was used to train the neural network for 2003 prediction, and data from 2006 was used to train the neural network for 2007 prediction. For these long-term experiments, the density data sets included n=3,152,812 data points for each year with a sampling rate of 120 s. Fig. 6 shows the daily averaged values for Dst, ap, and $F_{10.7}$ during 2003 and 2007. When only using the density values, the input vector includes the density at each sample time, while, when using the external inputs, it is made of the density vector, the Dst vector, and the $F_{10.7}$ vector.

4.2. Solar and geomagnetic indices

By including additional inputs other than the known present values (in this case the density values), the performance of a neural network as a predictor improves, provided that the output of the neural network is a function of these inputs. Because the density is driven by the solar and geomagnetic activities, one proxy for each of these was selected as additional inputs. Out of the many possible choices (e.g. S_{10}, M_{10}, Mg II), $F_{10.7}$ was assumed to be a suitable proxy for the solar activity affecting the density. This is a valid assumption since, as pointed out by Ref. [31], one index has not yet been proven to be clearly superior to any other for satellite operations. Dst was assumed to be a valid proxy for geomagnetic activity. The ap index was not used since, as Fig. 6 shows, Dst and ap are closely related. Furthermore as pointed out in Ref. [11], replacing ap for Dst reduced density errors especially during geomagnetic storms for Jacchia 70, NRLMSIS and JB 2008.

The indices were averaged hourly and were included in the corresponding training, validation, and testing sets (of 1080 and 3,152,812 data points for the one day and one year data sets respectively). No interpolation was necessary for the indices since the values of the indices assumed for each density data point were the latest hourly-averaged values of the indices as of the time of that particular data point. For the one orbit prediction case at a sampling rate of 80 s, 68 samples per window are used; for the eight orbits case at a sampling rate of 120 s, 360 samples per window are used; and for the 32 orbits case at a sampling rate of 120 s, 1440 samples per window are used. As with the density values, during operation the neural networks only have access to present values of the indices. The values for the Dst and $F_{10.7}$ indices used in this work were obtained from Ref. [51].

5. Test results

The training, validation, and testing of the neural networks was done in MATLAB using the Neural Network Toolbox [52]. As a benchmark for all the tests, a model using the persistence method was used. The persistence method is a very simple technique for forecasting in which the prediction is equal to the input. In other words, a predictor forecasting the density using the persistence method will predict the density in the future to be the same as it is in the present, so if the prediction window is one orbit, then each predicted orbit is equal to the previous measured orbit.

To assess the performance of the different models, different metrics were used: the MSE (shown in Eq. (7)), the mean of the ratio between the target and the outputs, its standard deviation (shown in Eqs. (8) and (9)) and the Pearson correlation coefficient of the targets to the model outputs (shown in Eq. (10)). The former two were the metrics used by Stastny et al. [29], and therefore allow
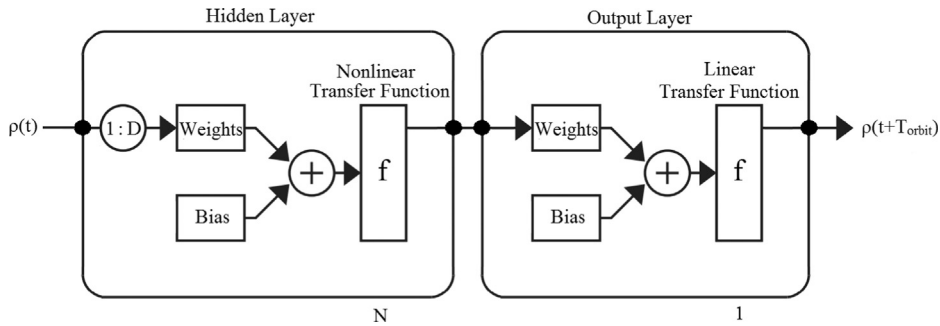
Fig. 7. Neural network diagram where N and D are the numbers of neurons and delays in the hidden layer.

Table 2
Neural network performance for different number of neurons in the hidden layer (bold for best results, italics for worst).

Testing data set	Model configuration	MSE	R	Mean target/output $\bar{r}$	St. dev. target/output s_r
Day 141 2002 CHAMP	Persistence model	0.023	0.969	1.0003	0.0058
	1 Neuron	**0.016**	0.977	1.0008	**0.0046**
	5 Neurons	0.016	**0.980**	*1.0011*	**0.0046**
	10 Neurons	0.016	0.979	1.0007	0.0047
	50 Neurons	0.016	0.979	1.0006	0.0047
	100 Neurons	0.025	0.968	1.0003	0.0059
	150 Neurons	*0.026*	*0.965*	**1.0001**	*0.0059*
Day 276 2001 CHAMP	Persistence model	0.033	0.872	**1.0001**	0.0070
	1 Neuron	**0.023**	**0.911**	0.9995	**0.0058**
	5 Neurons	0.025	0.907	0.9999	0.0058
	10 Neurons	0.037	0.891	0.9967	0.0066
	50 Neurons	0.045	0.879	0.9963	0.0072
	100 Neurons	0.171	0.649	0.9958	0.0151
	150 Neurons	*0.267*	*0.743*	*0.9885*	*0.0154*

comparing the results of this work to his.

$$\bar{r} = \frac{1}{n} \sum_{i=1}^{n} r_i, \quad r_i = \frac{x}{\hat{y}_i} \qquad (8)$$

$$s_r = \left(\frac{1}{n-1} \sum_{i=1}^{n} (r_i - \bar{r})^2 \right)^{1/2} \qquad (9)$$

$$R = \frac{1/(n-1) \sum_{i=1}^{n} (x_i - \bar{y})(\hat{x}_i - \bar{\hat{y}})}{s_{\hat{\rho}} s_{\rho}} \qquad (10)$$

where $s_{\hat{\rho}}$ and s_{ρ} are the standard deviations of the neural network outputs and targets, respectively.

5.1. Different number of neurons, delays and sampling rates

Given the selected structure of the neural network predictors, the appropriate number of neurons, delays in the hidden layer, and the data sampling rate for the localized density forecasting problem were found. This was accomplished empirically by trying different combinations. All the tests performed for this purpose were run on days 141 of 2002 and 276 of 2001, with the training and validation sets being day 140 of 2002. Again, these days cover high and low geomagnetic activity and were used also by Stastny et al. [29] to test his linear model. To find the appropriate number of neurons and delays in the hidden layer the neural network shown in Fig. 7 was used.

To find the appropriate number of neurons in the hidden layer several tests were performed in which all the other parameters were fixed and the number of neurons was varied. The prediction window was set to one orbit into the future; 17 delays (1/4 of the prediction windows) and a sampling rate of 80 s were used. The results for these tests are summarized in Table 2.

As shown in Table 2 as the number of neurons is increased, the performance of the neural networks worsens. This occurs because increasing the number of neurons causes the neural network to become overtrained. For the problem of predicting density, the results included in Table 2 indicate that for the given neural network structure, having one neuron in the hidden layer gives the best results.

To find the suitable number of delays in the hidden layer all other parameters were fixed and the number of delays was varied. The prediction window of one orbit into the future was used, the neural networks had one neuron in the hidden layer (the best result from the previous test) and again a sampling rate of 80 s was used. The results for these tests are shown in Table 3.

As can be observed from Table 3, the number of delays in the hidden layer significantly determines the neural network performance. Having too few delays results in underfitting, while having too many results in overfitting. The best results are obtained when the number of delays is between 17 and 68 (between 1/4 and one full prediction window). It is important to note that there is not a

Table 3
Neural network performance for different number of delays in the hidden layer (bold for best results, italics for worst).

Testing data set	Model configuration	MSE	R	Mean target/output $\bar{r}$	St. dev. target/output s_r
Day 141 2002 CHAMP	Persistence model	*0.023*	*0.969*	**1.0003**	*0.0058*
	1 delay	0.016	0.978	1.0005	0.0048
	17 delays(1/4 orbit)	0.016	0.977	1.0008	0.0046
	34 delays(1/2 orbit)	**0.014**	**0.982**	1.0009	**0.0044**
	51 delays (3/4 orbit)	**0.014**	**0.982**	1.0010	**0.0044**
	68 delays(1 orbit)	0.015	0.981	1.0011	0.0045
	102 delays (3/2 orbit)	0.017	0.978	*1.0013*	0.0048
	204 delays (3 orbits)	0.018	0.978	*1.0013*	0.0048
Day 276 2001 CHAMP	Persistence model	0.033	*0.872*	1.0001	*0.007*
	1 delay	0.028	0.891	**1.0001**	0.0064
	17 delays(1/4 orbit)	**0.023**	0.911	0.9995	0.0058
	34 delays(1/2 orbit)	0.027	**0.913**	0.9973	**0.0056**
	51 delays (3/4 orbit)	0.027	**0.908**	0.9973	**0.0057**
	68 delays(1 orbit)	0.028	0.907	0.9972	0.0058
	102 delays (3/2 orbit)	0.033	0.899	0.9967	0.0061
	204 delays (3 orbits)	*0.037*	0.877	*0.9965*	0.0065

significant difference in the performance of the neural networks when the number of delays is in this range.

The last parameter determined was the data sampling rate. In this case one neuron was selected for the hidden layer; the prediction window was set to one orbit into the future and 17 delays (1/4 of the prediction window) were selected. The results for these tests are summarized in Table 4 and show that there is no significant difference in the performance of the neural network when the sampling rate is below 180 s. Beyond this limit the performance decays.

One interesting result, that can be observed from Tables 2–4, is that the ratio of the targets to the outputs ($\bar{r}$) does not seem to change significantly for all the parameters tested (number of delays and neurons, and sampling rate) even for the neural networks with the worst performance (both underfitted and overfitted). This is explained by the fact that this metric measures the bias in the model. The bias in models that use actual values as inputs, such as the neural networks, will be very small provided that the mean of the time series does not vary significantly.

5.2. Predicting one orbit into the future on days 141 of 2002 and 276 of 2001

Once the appropriate structure of the neural network was found (one neuron and enough delays to store from 1/4 to one prediction window) several different neural networks were tested again on days 141 of 2002 and 276 of 2001. This was done to evaluate the improvements in performance by increasing the size of the training and validation sets from one day to a year and also by using the solar and geomagnetic indices (Dst and $F_{10.7}$) as additional inputs. The one year training and validation data set used for testing the networks on day 141 of 2002 contained the data from the 365 preceding days (day 140 2001 to day 14 2002). The one year training and validation data set used for testing the networks on day 276 of 2001 contained the data from year 2002 (day 1 2002 to day 365 2002), since the CHAMP data did not go back a year before day 276

2001. Even though this means that the neural network used was trained on data corresponding to the future of day 276 2001, the training data and validation data set is still different to the testing set which makes the test valid (of course for practical implementation of the neural networks the training and validation set would always be past and therefore available values). A sampling rate of 80 s was used since it is the same used by Stastny et al. [29]. The results of the tests for days 141 of 2002 and 276 of 2001 are summarized in Table 5.

For the 2001 scenario, training data from 2002 is used, as Stastny et al. also did in their work [29]. Training with future values and "predicting" past values is valid from the point of view of neural network, since the training/validation and testing data sets are still different. From the point of view of spacecraft onboard implementation this of course would be invalid, but here the purpose is just showing the neural network's performances and comparing with existing literature. The results in Table 5 indicate that the global models (HASDM and JB 2006 results obtained from Stastny et al. [29]) suffer from large biases in their results. This causes their performance to be much worse than the performance of all the other models including the persistence model. The neural network predictors give significantly better results than the linear model from Ref. [29], the global models, and the persistence model. For day 141 of 2002, by increasing the size of the training and validation sets, the performance of the neural networks increases; however, for day 276 of 2001 there is not a significant improvement by increasing the size of the training and validation sets nor by including the solar and geomagnetic indices. The addition of the indices does not benefit the neural networks because the number of delays (17 which corresponds to 1/4 of the prediction window) cannot capture more than one value in time of the indices since, the indices are averaged hourly. This might be solved by increasing the number of delays; however, they cannot be increased beyond 68 (one prediction window) since as shown in Table 3, this results in overfitting the neural network. An alternative solution would be retaining the same number of delays, but space them non-uniformly in time.

Table 4
Neural network performance for different sampling rates (bold for best results, italics for worst).

Testing data set	Model configuration	MSE	R	Mean target/output $\bar{r}$	St. dev. target/output s_r
Day 141 2002 CHAMP	10 s sampling rate	0.0157	**0.979**	1.0008	0.0047
	30 s sampling rate	0.0157	0.979	1.0008	0.0047
	60 s sampling rate	0.0157	0.979	1.0008	0.0047
	80 s sampling rate	0.0156	0.977	1.0008	**0.0046**
	120 s sampling rate	**0.0155**	0.980	1.0008	**0.0046**
	180 s sampling rate	0.0159	0.980	**1.0005**	0.0047
	300 s sampling rate	*0.0220*	*0.971*	*1.0009*	*0.0056*
Day 276 2001 CHAMP	10 s sampling rate	**0.0227**	**0.912**	0.9996	**0.0058**
	30 s sampling rate	0.0229	0.911	0.9996	**0.0058**
	60 s sampling rate	0.0232	0.910	0.9995	**0.0058**
	80 s sampling rate	0.0229	0.911	0.9995	**0.0058**
	120 s sampling rate	0.0232	0.910	0.9996	**0.0058**
	180 s sampling rate	0.0235	0.910	0.9995	0.0059
	300 s sampling rate	*0.0285*	*0.899*	*0.9986*	*0.0063*

Table 5
Results for predicting one orbit into the future (bold for best results, italics for worst).

Testing Data Set	Model configuration	MSE	R	Mean target/output $\bar{r}$	St. dev. target/output s_r
Day 141 2002 CHAMP	NN, preceding 365 days for training	**0.0108**	**0.9843**	**0.9998**	**0.0039**
	NN, preceding 365 days for training, Dst and $F_{10.7}$	**0.0108**	0.9842	**0.9998**	**0.0039**
	NN, day 140 of data for Training	0.0156	0.9774	1.0008	0.0046
	Persistence model	*0.0234*	*0.9685*	1.0003	0.0058
	Linear model[a]	N/A	N/A	1.0058	0.0822
	HASDM[a]	N/A	N/A	0.8662	*0.1204*
	JB2006[a]	N/A	N/A	*0.8564*	0.095
Day 276 2001 CHAMP	NN, year 2002 of data for training	0.0229	0.9086	**0.9999**	**0.0058**
	NN, Year 2002 of data for training, Dst and $F_{10.7}$	**0.0225**	**0.9099**	**0.9999**	**0.0058**
	NN, day 140 Year 2002 of data for training	0.0229	0.9106	0.9995	**0.0058**
	Persistence model	*0.0328*	*0.8718*	**1.0001**	0.007
	Linear model[a]	N/A	N/A	1.0094	0.0822
	HASDM[a]	N/A	N/A	0.8415	0.1344
	JB2006[a]	N/A	N/A	*0.6471*	*0.1355*

[a] Obtained from Ref. [29].

For day 141 of 2002, utilizing one years' worth of data to train and validate the neural network provided the best results. The actual output of this neural network and the targets are shown in Fig. 8 along with the prediction error. For day 276 of 2001, the neural network that uses the additional inputs (Dst and $F_{10.7}$) and that was trained and validated using the data from one year yielded the best results. The actual output of this neural network, the targets, and the prediction error are shown in Fig. 9.

5.3. Predicting eight and 32 orbits into the future

For most applications of the neural network density predictors, longer prediction windows are desired. For this reason additional neural network predictors were trained, validated, and tested for predicting eight and 32 orbits into the future (roughly half a day and two days respectively). For these results, the neural networks were tested on years 2003 and 2007, in order to evaluate their performance over much wider data sets including periods of low and high solar and geomagnetic activities. Again, the use of additional inputs (Dst and $F_{10.7}$) was studied along with the use of different numbers of delays. Since having
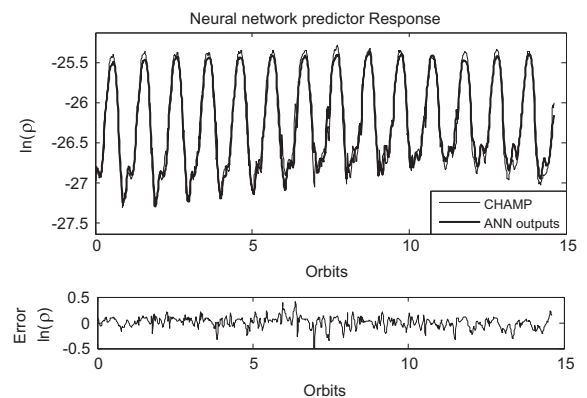

Fig. 8. Neural network response for best case with a prediction window of one orbit over day 141 of 2002.

different sampling rates, as long as they are below 180 s, does not affect significantly the neural network performance (as shown in Table 4), a sampling rate of 120 s was used in order to reduce time for training and validation the neural networks. The results are summarized in Tables 6

and 7 for the prediction windows of eight and 32 orbits respectively.

The best case included in Table 6 for both years 2003 and 2007 were those obtained with the neural network that included the additional inputs and that had 360 delays (one prediction window). Figs. 10 and 11 show the MSE over the entire years 2003 and 2007 for the best cases along with the Dst and $F_{10.7}$ averaged daily.

As indicated in Table 7, for the 32 orbit predictions the neural networks yielded a better performance over the high activity period (2003) than the low activity one (2007). This was not observed in any of the other tests performed (one orbit and 16 orbits) where the performance for both periods was almost the same or better for periods of low activity (see Tables 5 and 6). This indicates that for longer prediction windows over periods of low activity other unknown factors affect the density behavior, that are not well represented by the data used by the neural networks (current value of the density, Dst and $F_{10.7}$ indices). This is here considered a topic for further investigation and beyond the scope of this work. Further investigation may lead to the discovery of unknown effects during periods of low activity. Waves, tides, and thunderstorm phenomena in the lower atmosphere have received an increased amount of attention lately, as potential

factors affecting upper atmosphere behavior during low geomagnetic and solar activity [53].

The best case shown in Table 7 for both years 2003 and 2007 were those obtained with the neural network that used the Dst and $F_{10.7}$ indices. The MSE over the entire years 2003 and 2007 for the best cases included in Table 7 (32 orbits prediction), along with the Dst and $F_{10.7}$ averaged daily are shown in Figs. 12 and 13.

One very interesting feature can be observed in Figs. 10–13. The peaks in the MSE correspond to peaks in the Dst index. This indicates that the neural network predictors will have larger errors during geomagnetic storms. This is further confirmed by the results from the test performed for predicting one orbit into the future, in which the performance for the predictions on day 141 of 2002 were always better than the performance on day 276 of 2001, during which there was a geomagnetic storm.

From a computational point of view the training and validation of the neural network predictors can be costly especially when dealing with large prediction windows (32 orbits into the future). For this reason, for future implementation of this work the training and validation processes are not recommended to be done onboard. Rather, it is proposed that density values obtained onboard via accelerometers be sent periodically to the mission team on the ground for training and validation. Once the weights and biases of the trained neural network are obtained, they can be uplinked to the onboard computers and then the predictor can be used for onboard orbit propagation. The neural network could then be re-trained on the ground as necessary. Given that even neural networks trained with small data sets (e.g. one day) give good results (see Table 5), this prediction method could be employed even for early stages of a mission, and as more data is available, more accurate neural network predictors can be trained and sent to the spacecraft. The predictor developed here is to be included in the Propellant-less Atmospheric Differential Drag Low Earth Orbit Spacecraft (PADDLES) mission currently in development at the ADvanced Autonomous MUltiple Spacecraft (ADAMUS) laboratory at the Rensselaer Polytechnic Institute [54]. PADDLES is a 3-unit CubeSat [55] that will host a dedicated ram sensor capable of measuring atmospheric neutral density. The neural network approach proposed herein is
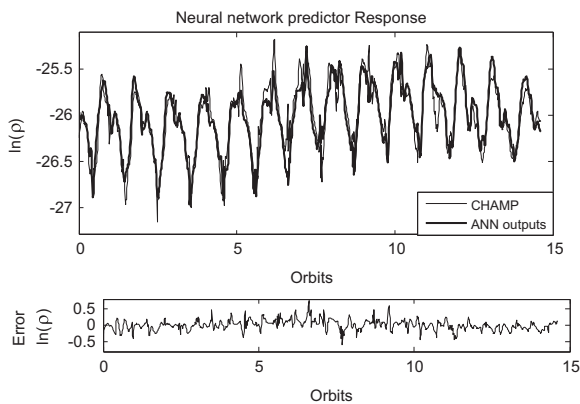

**Fig. 9.** Neural network response for best case with a prediction window of one orbit over day 276 of 2001.

Table 6
Results for predicting eight orbits into the future (bold for best results, italics for worst).

Testing data set	Model configuration	MSE	R	Mean target/output $\bar{r}$	St. dev. target/output s_r
CHAMP 2003	ANN, 90 delays (2 orbits), 1 year 2002 of data for training	0.0433	0.8971	*1.0007*	0.0078
	ANN, 90 delays (2 orbits), 1 year 2002 of data and Dst and $F_{10.7}$ for training	0.0429	0.8976	**1.0000**	0.0078
	ANN, 360 delays (8 orbits), 1 year 2002 of data and Dst and $F_{10.7}$ for training	**0.0401**	**0.9044**	0.9999	**0.0075**
	Persistence model	*0.2614*	*0.4037*	1.0002	*0.0192*
CHAMP 2007	ANN, 90 delays (2 orbits), 1 year 2006 of data for training	0.0417	0.9093	*1.0002*	0.0075
	ANN, 90 delays (2 orbits), 1 year 2006 of data and Dst and $F_{10.7}$ for training	0.0407	0.9114	1.0000	0.0074
	ANN, 360 delays (8 orbits), 1 year 2006 of data and Dst and $F_{10.7}$ for training	**0.0403**	**0.9122**	**1.0000**	**0.0074**
	Persistence model	*0.1902*	*0.6031*	1.0001	*0.0160*

Table 7
Results for predicting 32 orbits into the future (bold for best results, italics for worst).

Testing data set	Model configuration	MSE	R	Mean target/output $\bar{r}$	St. dev. target/output s_r
CHAMP 2003	ANN, 360 delays (8 orbits), 1 year 2002 of data for training	0.0917	0.7702	*1.0013*	0.0113
	ANN, 360 delays (8 orbits), 1 year 2002 of data and Dst and $F_{10.7}$ for training	**0.0895**	**0.7740**	**1.0000**	**0.0112**
	Persistence model	*0.1813*	*0.5874*	1.0001	*0.0160*
CHAMP 2007	ANN, 360 delays (8 orbits), 1 year 2006 of data for training	0.1564	0.6058	*1.0009*	0.0145
	ANN, 360 delays (8 orbits), 1 year 2006 OF data and Dst and $F_{10.7}$ for training	**0.1515**	**0.6215**	**1.0002**	**0.0143**
	Persistence model	*0.603*	*−0.2582*	1.0005	*0.0286*
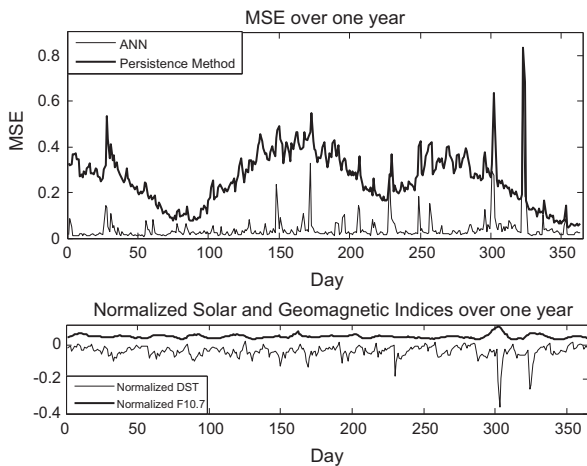
Fig. 10. MSE for best case with a prediction window of eight orbits and normalized indices over year 2003.

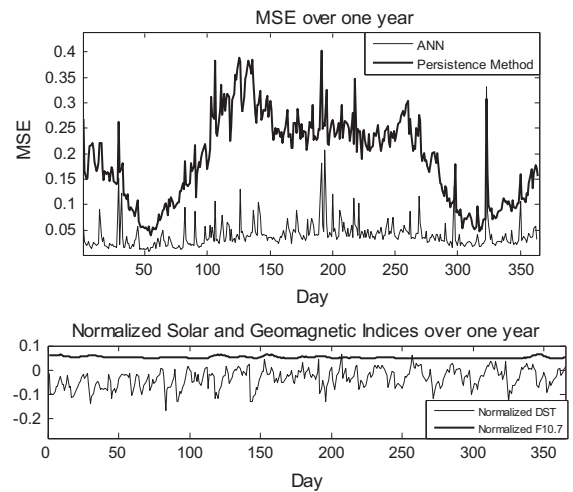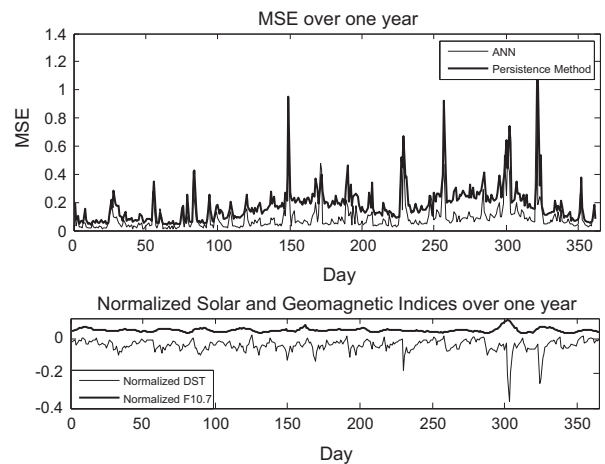**Fig. 11.** MSE for best case with a prediction window of eight orbits and normalized indices over year 2007.

Fig. 12. MSE for best case with a prediction window of 32 orbits and normalized indices over year 2003.

envisioned as part of the onboard algorithms for PADDLES, expected to be delivered by end of 2015.

6. Conclusions

The Low Earth Orbit atmospheric density neural network predictors presented in this work provide significantly better results than a linear model, and the global models HASDM and JB2006 for predicting the value of the density one orbit into the future, for periods of both high and low geomagnetic activity. Moreover, the large bias in the density predictions, present in the density predicted by the global models (which is even more severe during high geomagnetic activity), is not present in the results from the neural networks, from the linear model and from the persistence method, both during the periods of low and high geomagnetic activity. This suggests that one way for reducing the bias on the global models may be the use of past measured values of the density whenever they are available.

An interesting finding is the unexpected behavior for the 32 orbit prediction cases, as the neural network seems to work less accurately for long term predictions during low activity. This may indicate unknown dynamics driving the behavior of the density during phases when the Sun is not the main player. One possibility is the consideration of waves, tides, and thunderstorms in the lower atmosphere

as potential factors affecting upper atmosphere. Further investigation may lead to the definition of new indices, similar to the ones used for solar and geomagnetic activity, carrying information about lower atmosphere.

The neural network predictors can also use the current value of the Dst (geomagnetic activity) and $F_{10.7}$ (solar activity)
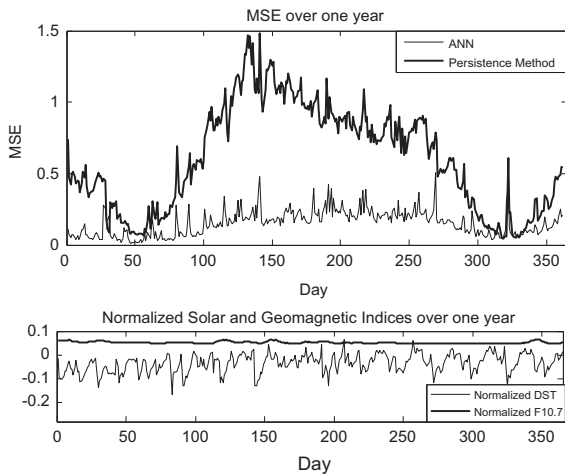
Fig. 13. MSE for best case with a prediction window of 32 orbits and normalized indices over year 2007.

indices averaged hourly as additional inputs which resulted in an improvement of the performance of the neural networks, provided that enough delays were included in the hidden layer to store some of the behavior in time of the indices. However, the number of delays cannot be increased beyond those required to store one prediction window or the neural networks will suffer from overfitting in terms of the density values. For this reason, it is expected that the performance of the neural network predictors can be improved further by having different number of delays for the density values and the solar and geomagnetic indices. An alternative solution would be retaining the same number of delays, but space them non-uniformly in time. This will allow for storing more information of the indices in the neural network predictors, which may improve the performance during periods of high geomagnetic activity.

The neural network predictors are computationally simple and can be implemented onboard a spacecraft and therefore allow for precise onboard orbit propagation at low computational cost. An example of immediate application of the algorithms presented herein is a CubeSat satellite currently developed at Rensselaer Polytechnic Institute, called Propellant-less Atmospheric Differential Drag Low Earth Orbit Satellite.

Acknowledgments

References

[1] C.L. Leonard, W.M. Hollister, E.V. Bergmann, Orbital formation keeping with differential drag, AIAA J Guidance Control Dyn. 12 (1) (1989) 108–113, http://dx.doi.org/10.2514/3.20374.

[2] T.D. Maclay, C. Tuttle, Satellite station-keeping of the ORBCOMM Constellation via active control of atmospheric drag: operations, constraints, and performance, Adv. Astronaut. Sci. 120 (Part I) (2005) 763–773. (AAS 05-152).

[3] Kumar, B.S., and Ng, A., A bang-bang control approach to maneuver spacecraft in a formation with differential drag, in: Proceedings of the AIAA Guidance, Navigation and Control Conference and Exhibit, Honolulu, Hawaii, August 2008, http://dx.doi.org/10.2514/6.2008-6469.

[4] D. Pérez, R. Bevilacqua, Differential drag spacecraft rendezvous using an adaptive Lyapunov Control Strategy, Acta Astronaut. 83 (2013) 196–207, http://dx.doi.org/10.1016/j.actaastro.2012.09.005.

[5] Pérez, D., Bevilacqua, R., Spacecraft maneuvering via atmospheric differential drag using an adaptive Lyapunov Controller, in: Paper AAS 13-440 at the 23rd AAS/AIAA Spaceflight Mechanics Meeting, Kauai, Hawaii, February 2013.

[6] D.A. Vallado, Fundamentals of Astrodynamics and Applications, 3rd ed. Springer, New York, 2001 (Chapter 8).

[7] L.G. Jacchia, A variable atmospheric-density model from satellite accelerations, J Geophys. Res. 65 (1960) 2775–2782, http://dx.doi.org/10.1029/JZ065i009p02775.

[8] L.G. Jacchia, Revised static models of the thermosphere and exosphere with empirical temperature profiles, Smithsonian Astrophysical Observatory Special Report no. 332, 1971.

[9] L.G. Jacchia, Thermospheric temperature, density, and composition: new models, Smithsonian Astrophysical Observatory Special Report no. 375, 1977.

[10] B.R. Bowman, et al., The JB2006 empirical thermospheric density model, J. Atmos. Sol. Terr. Phys. 70 (2008) 774–793, http://dx.doi.org/10.1016/j.jastp.2007.10.002.

[11] Bowman, B.R., W.K. Tobiska, F.A. Marcos, C.Y. Huang, C.S. Lin, W.J. Burke, A new Empirical Thermospheric Density Model JB2008 using new Solar And geomagnetic indices, in: Proceedings of the AIAA/AAS Astrodynamics Specialist Conference, AIAA2008-6438, 2008a, http://dx.doi.org/10.2514/6.2008-6438.

[12] Storz, M.F., Bowman, B.R., Branson, J.I., High accuracy satellite drag model (HASDM), in: Proceedings of the AIAA/AAS Astrodynamics Specialist Conference, Monterey, CA, August 2002, http://dx.doi.org/10.1016/j.asr.2004.02.020.

[13] Shoemaker, M.A., Wohlberg, B., Koller, J., Atmospheric Density reconstruction using satellite orbit tomography, in: Proceedings of the 23rd AAS/AIAA Spaceflight Mechanics Meeting, Lihue, HI, USA, February 2013, pp. 1235–1250.

[14] A.E. Hedin, J.E. Salah, J.V. Evans, C.A. Reber, G.P. Newton, N.W. Spencer, D.C. Kayser, D. Alcaydé, P. Bauer, L. Cogger, J.P. McClure, A global thermospheric model based on mass spectrometer and incoherent scatter data MSIS, 1. N_2 density and temperature, J. Geophys. Res. 82 (16) (1977) 2139–2147.

[15] A.E. Hedin, MSIS-86 thermospheric model, J. Geophys. Res. 92 (A5) (1987) 4649–4662, http://dx.doi.org/10.1029/JA082i016p02139.

[16] A.E. Hedin, Extension of the MSIS thermospheric model into the middle and lower atmosphere, J. Geophys. Res. 96 (1991) 1159, http://dx.doi.org/10.1029/JA092iA05p04649.

[17] J.M. Picone, A.E. Hedin, D.P. Drob, A.C. Aikin, NRLMSISE-00 empirical model of the atmosphere: statistical comparisons and scientific issues, J. Geophys. Res. 107 (A12) (2002) 1468, http://dx.doi.org/10.1029/2002JA009430.

[18] F. Barlier, C. Berger, J.L. Falin, G. Kockarts, G. Thuillier, A thermospheric model based on satellite drag data, Ann. Geophys. 34 (1978) 9–24. (doi: 10.1016/S1364-6826(03)00137-8).

[19] C. Berger, R. Biancale, M. Ill, F. Barlier, Improvement of the empirical thermospheric model DTM: DTM94 a comparative review of various temporal variations and prospects in space geodesy applications, J. Geod. 72 (3) (1998) 161–178, http://dx.doi.org/10.1007/s001900050158.

[20] S. Bruinsma, G. Thuillier, F. Barlier, The DTM-2000 empirical thermosphere model with new data assimilation and constraints at lower boundary: Accuracy and properties, J. Atmos. Sol.-Terr. Phys. 65 (9) (2003) 1053–1070, http://dx.doi.org/10.1016/S1364-6826(03)00137-8.

[21] R.G. Roble, E.C. Ridley, A thermosphere–ionosphere–mesosphere–electrodynamics general circulation model (TIME-GCM): equinox solar cycle minimum simulations (30–500 km), Geophys. Res. Lett. 21 (6) (1994) 417–420, http://dx.doi.org/10.1029/93GL03391.

[22] T.J. Fuller-Rowell, D. Rees, S. Quegan, R.J. Moffett, M.V. Codrescu, G.H. Millward, A coupled thermosphere–ionosphere model (CTIM) STEP Handbook of Ionospheric Models, Utah State Univ., Logan, Utah, 1996, 217–238, http://dx.doi.org/10.1029/GL015i012p01325.

[23] G.H. Millward, R.J. Moffett, S. Quegan, T.J. Fuller-Rowell, A coupled thermosphere–ionosphere–plasmasphere model (CTIP) STEP

Handbook on Ionospheric Models, Utah State Univ., Logan, Utah, 1996, 239–279.

[24] G.H. Millward, I.C.F. Müller-Wodarg, A.D. Aylward, T.J. Fuller-Rowell, A.D. Richmond, R.J. Moffett, An investigation into the influence of tidal forcing on *F* region equatorial vertical ion drift using a global ionosphere–thermosphere model with coupled electrodynamics, J. Geophys. Res: Sp. Phys. 106 (A11) (2001) 24733–24744, http://dx.doi.org/10.1029/2000JA000342. (1978–2012).

[25] T.J. Fuller-Rowell, G.H. Millward, A.D. Richmond, M.V. Codrescu, Storm-time changes in the upper atmosphere at low latitudes, J. Atmos. Sol.-Terr. Phys. 64 (12) (2002) 1383–1391, http://dx.doi.org/10.1016/S1364-6826(02)00101-3.

[26] M.J. Harris, N.F. Arnold, A.D. Aylward, A study into the effect of the diurnal tide on the structure of the background mesosphere and thermosphere using the new coupled middle atmosphere and thermosphere (CMAT) general circulation model, Ann. Geophys. 20 (2) (2002) 225–235.

[27] I. Cnossen, M.J. Harris, N.F. Arnold, E. Yiğit, Modelled effect of changes in the CO_2 concentration on the middle and upper atmosphere: sensitivity to gravity wave parameterization, J. Atmos. Sol.-Terr. Phys. 71 (13) (2009) 1484–1496, http://dx.doi.org/10.1016/j.jastp.2008.09.014.

[28] A.J. Ridley, Y. Deng, G. Toth, The Global Ionosphere–Thermosphere Model (GITM), J. Atmos. Sol.-Terr. Phys. 68 (8) (2006) 839–864, http://dx.doi.org/10.1016/j.jastp.2006.01.008.

[29] Stastny, N.B., Chavez, F.R., Lin, C., Lovell, A.T., Bettinger, R.A., and Luck, J., Localized density/drag prediction for improved onboard orbit propagation, in: Proceedings of the Advanced Maui Optical and Space Surveillance Technologies Conference, Maui, Hawaii, September 1–4, 2009, pp. 51–58.

[30] C. Reigber, H. Lühr, P. Schwintzer, CHAMP mission status, Adv. Sp. Res. 30 (2) (2002) 129–134, http://dx.doi.org/10.1016/S0273-1177(02)00276-4.

[31] Vallado, D.A., & Finkleman, D., A critical assessment of satellite drag and atmospheric density modeling, Acta Astronautica, ISSN 0094-5765, Available online http://dx.doi.org/10.1016/j.actaastro.2013.10.005 (accessed 25.10.13).

[32] B.D. Tapley, S. Bettadpur, M. Watkins, C. Reigber, The gravity recovery and climate experiment: mission overview and early results, Geophys. Res. Lett. 30 (L09607) (2004) 1–4, http://dx.doi.org/10.1029/2004GL019920.

[33] Ray, S., The layers of Earth′s atmosphere, Jet Propulsion Laboratory [government website], URL: ⟨http://airs.jpl.nasa.gov/maps/satellite_feed/atmosphere_layers/⟩ (accessed 14.12.13).

[34] E. Doornbos, M. Forster, B. Fritsche, T. van Helleputte, J. van den IJssel, G. Koppenwallner, H. Luhr, D. Rees, P. Visser, ESTEC contract 21022/07/NL/HE air density models derived from multi-satellite drag observations—final report, Technical Report DEOS/TU Delft Scientific Report 01/2009, TU Delft, 2009, 7–39.

[35] Sugiura, M., Hourly values of equatorial Dst for the IGY, Annals of the International Geophysical Year, 35, 9, Pergamon, New York, 1964.

[36] Sugiura, M., Kamei, T., Equatorial Dst index 1957–1986, in: International Association of Geomagnetism and Aeronomy Bulletin No. 40, 1991, pp. 1–14.

[37] A.E. Covington, Microwave solar noise observations during the partial eclipse of November 23, 1946, Nature 159 (1947) 405–406.

[38] A.E. Covington, Solar noise observations on 10.7 centimeters, Proc. IRE 36 (4) (1948) 454–457.

[39] W.K. Tobiska, S.D. Bouwer, B.R. Bowman, The development of new solar indices for use in thermospheric density modeling, J. Atmos. Sol.-Terr. Phys. 70 (5) (2008) 803–819, http://dx.doi.org/10.1016/j.jastp.2007.11.001.

[40] M. Menvielle, A. Berthelier, The *K*-derived planetary indices: description and availability, Rev. Geophys. 29 (3) (1991) 415–432.

[41] Linthe, H.J., Indices of Global Geomagnetic Activity, Helmholtz Centre Potsdam [government website], URL: ⟨http://www-app3.gfz-potsdam.de/kp_index/description.html⟩ (accessed 14.12.13).

[42] S. Haykin, Neural Networks and Learning Machines, 3rd ed. Prentice Hall, New York, 2009 (Chapter 1).

[43] K.A. Levenberg, Method for the solution of certain problems in least squares, Q. Appl. Math. 2 (1944) 164–168.

[44] D. Marquardt, An algorithm for least-squares estimation of non-linear parameters, SIAM J. Appl. Math. 11 (1963) 431–441, http://dx.doi.org/10.1137/0111030.

[45] The MathWorks, Levenberg-Marquardt backpropagation, Documentation Center [commercial website], URL: ⟨http://www.mathworks.com/help/nnet/ref/trainlm.html⟩ (accessed 14.12.13).

[46] E.K. Sutton, R.S. Nerem, J.M. Forbes, Density and winds in the thermosphere deduced from accelerometer data, J. Spacecr. Rocket. 44 (6) (2007) 1210–1219, http://dx.doi.org/10.2514/1.28641.

[47] Forbes, J.M., Nerem, R.S., Sutton, E.K., Zhang, X., and Bruinsma, S., Thermosphere studies using accelerometer measurements from the CHAMP and GRACE Satellites, Department of Aerospace Engineering Sciences University of Colorado [online database], URL: ⟨http://sisko.colorado.edu/sutton/⟩ (accessed 14.12.13).

[48] H. Lütkepohl, F. Xu, The role of the log transformation in forecasting economic variables, Empirical Economics, vol. 42, Springer, Berlin, 2012, 619–638, http://dx.doi.org/10.1007/s00181-010-0440-1.

[49] A.V. Belov, S.P. Gaidash, K.G. Ivanov, Kh.D. Kanonidi, Unusually High Geomagnetic Activity in 2003, Cosm. Res. 42 (6) (2004) 1–10, http://dx.doi.org/10.1007/s10604-005-0001-0.

[50] A.V. Belov, S.P. Gaidash, Anomalously low solar and geomagnetic activities in 2007, Geomagn. Aeron. 49 (5) (2009) 566–573, http://dx.doi.org/10.1134/S001679320905003X.

[51] Papitashvili, N., OMNIWeb, Goddard Space Flight Center Space Physics Data Facility [online database], URL: ⟨http://omniweb.gsfc.nasa.gov/ow.html⟩ (accessed 14.12.13).

[52] The MathWorks. Neural Network Toolbox, Documentation Center [commercial website], URL: ⟨http://www.mathworks.com/help/nnet/index.html⟩ (accessed 14.12.13).

[53] Firefly: An NSF CubeSat Project, URL:⟨http://firefly.siena-space.org/⟩ (accessed 14.12.13).

[54] Measuring Spatio-Temporal Variations in Upper Atmosphere via Drag Controlled Nano Satellites and WINCS (Winds-Ion-Neutral-Composition-Suite), URL: ⟨http://onr-yip.riccardobevilacqua.com/project.html⟩ (accessed 14.12.13).

[55] Cubesat, URL: ⟨http://www.cubesat.org/⟩ (accessed 14.12.13).

Dr. David Pérez is a Postdoctoral Research Associate in the ADvanced Autonomous MUltiple Spacecraft (ADAMUS) Laboratory at the Mechanical Aerospace and Nuclear Engineering Department at Rensselaer Polytechnic Institute, Troy, New York. David′s main research interests are guidance, navigation and control of autonomous spacecraft and robotics. He received his B.Sc. (2008) in Mechanical Engineering from St. Cloud State University, St. Cloud, Minnesota and his M.Eng (2009) and Ph.D. (2013) in Mechanical Engineering and Aeronautical Engineering respectively from the Rensselaer Polytechnic Institute, Troy, New York.

Dr. Brendt Wohlberg received the BSc (Hons) degree in applied mathematics, and the MSc (Applied Science) and PhD degrees in electrical engineering from the University of Cape Town (UCT), South Africa in 1990, 1993 and 1996 respectively. After graduating, he held postdoctoral research positions at UCT and thereafter at Los Alamos National Laboratory (LANL). He has been a technical staff member in Theoretical Division since 2002. His research interests include image restoration and related inverse problems, sparse representations, exemplar-based methods, and machine learning.

Dr. Thomas Alan Lovell is a Research Aerospace Engineer in the Space Vehicles Directorate of the Air Force Research Laboratory. He received his Ph.D. from Auburn University in Aerospace Engineering in 2001. He has authored or co-authored over 60 conference papers and a dozen journal articles. He is a Senior Member of AAS, and has been nominated as an Associate Fellow of AIAA. He serves on the AIAA Astrodynamics Technical Committee. He is also an Associate Editor for the AIAA Journal of Guidance, Control, and Dynamics. His research interests include astrodynamics, orbit determination, trajectory optimization, and feedback control design.

Dr. Michael Shoemaker earned a Bachelor of Science degree in aerospace engineering from Virginia Tech in 2004. He worked for The Boeing Company in Leesburg, Virginia, doing mission operations and orbit determination for the Iridium satellite constellation until 2008. During that time, he earned a Master of Science degree in aerospace engineering from the University of Maryland. He then went to Kyushu University in Fukuoka, Japan, where he earned a Ph.D. in aeronautics and astronautics in 2012. He has worked on several international space missions, such as the European Rosetta mission and the Japanese Hayabusa mission.

Dr. Riccardo Bevilacqua is an Assistant Professor of the Mechanical, Aerospace, and Nuclear Engineering Department, at the Rensselaer Polytechnic Institute. He holds an M.Sc. in Aerospace Engineering (2002), and a Ph. D. in Applied Mathematics (2007), both earned at the University of Rome, "Sapienza", Italy. He was a US National Research Council Post-Doctoral Fellow from 2007 to 2010, before joining RPI. He also worked as project engineer in Mission Analysis at Grupo Mecanica del Vuelo, in Madrid, Spain, during 2003. Dr. Bevilacqua's research interests focus on Guidance, Navigation, and Control of multiple spacecraft systems and multiple robot systems.