

# NeuralODEs for VLEO simulations: Introducing thermoNET for Thermosphere Modeling

Dario Izzo<sup>1</sup>, Giacomo Acciarini<sup>1,2</sup>, and Francesco Biscani<sup>3</sup>

<sup>1</sup>*Advanced Concepts Team, European Space Research and Technology Centre (ESTEC), Noordwijk, The Netherlands.*

<sup>2</sup>*Surrey Space Centre, University of Surrey, Guildford, United Kingdom*

<sup>3</sup>*ESA/ESOC, Darmstadt, Germany*

**Abstract** - We introduce a novel neural architecture termed thermoNET, designed to represent thermospheric density in satellite orbital propagation using a reduced amount of differentiable computations. Due to the appearance of a neural network on the right-hand side of the equations of motion, the resulting satellite dynamics is governed by a NeuralODE, a neural Ordinary Differential Equation, characterized by its fully differentiable nature, allowing the derivation of variational equations (hence of the state transition matrix) and facilitating its use in connection to advanced numerical techniques such as Taylor-based numerical propagation and differential algebraic techniques. Efficient training of the network parameters occurs through two distinct approaches. In the first approach, the network undergoes training independently of spacecraft dynamics, engaging in a pure regression task against ground truth models, including JB-08 and NRLMSISE-00. In the second paradigm, network parameters are learned based on observed dynamics, adapting through ODE sensitivities. In both cases, the outcome is a flexible, compact model of the thermosphere density greatly enhancing numerical propagation efficiency while maintaining accuracy in the orbital predictions.

## I. INTRODUCTION

The analysis of satellite trajectories, particularly when influenced by the upper atmosphere, is affected by critical uncertainties. While precise models of our thermosphere such as the widely used NRLMSIS from the United States Naval Research Laboratory [1], the Jacchia-Bowman (JB) [2], the Drag Temperature Models (DTM) [3], and the High Accuracy Satellite Drag Model (HASDM) [4], have become standard in orbital simulations, their significant discrepancies underline the challenge of accurately estimating drag effects. In a recent work by Yin et al. [5] quantitative measures of the

relative errors introduced by empirical atmospheric models are presented using as ground truth the density inferred in the thermosphere from Swarm-C [6] orbital data and show relative deviations as large as 300% and overall mostly larger than 20%. Inherent uncertainties in established thermosphere models are thus significant, suggesting that artificial neural modelling may be particularly appropriate in this context. It is thus not surprising that the idea of using an artificial neural network to represent the density in the thermosphere has been introduced and studied in the recent past in several different works [7, 8, 9, 10], producing a plethora of neural models able to directly predict the density in the thermosphere or to correct existing models.

In this work, we introduce a novel neural architecture, an implicit neural representation [11] we call thermoNET which is revealed to be able to achieve high accuracy in reproducing thermospheric empirical models while only using a limited number of learnable parameters. The compact size of the resulting networks is essential for enabling their utilization in space flight mechanics applications, including orbital decay studies, re-entry analysis, and long-term debris population analysis. The guaranteed asymptotic behaviour of thermoNETs (as the altitude grows to infinity) is also an important property that avoids numerical issues when using the model outside its training bounds. Moreover, the inherent differentiability of our neural representation allows the automated derivation of variational equations, hence of the state transition matrix, and facilitates its integration with modern numerical techniques, such as Taylor numerical integration [12], differential algebra tools [13], and the neuralODEs approach [14], thus offering additional advantages in terms of computational efficiency and overall capabilities of orbital propagation.

## II. RESULTS

### A. thermoNETs: a differentiable atmospheric model

A thermoNET takes geodetic coordinates of a point in the thermosphere and selected space weather indices as

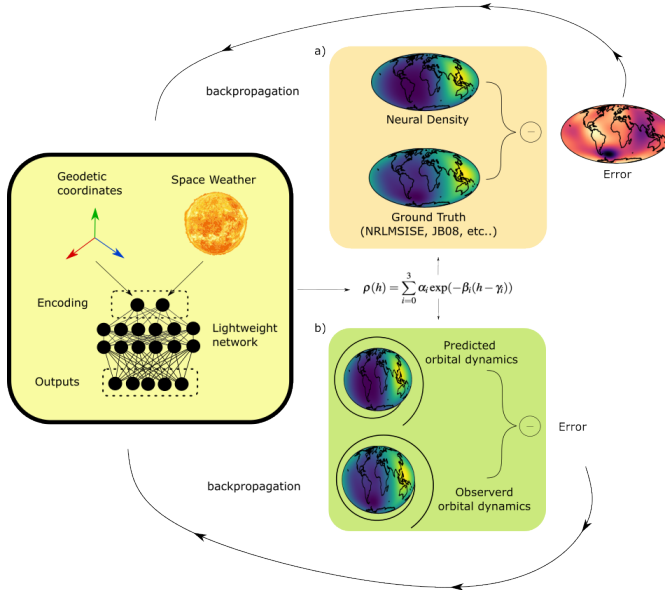


Fig. 1: Thermonets at a glance. A lightweight artificial neural network transforms geodetic coordinates and space weather data into the predefined air density model. The network parameters are learned via backpropagating the error via two distinct learning pipelines: a) to match some ground truth model (regression learning) or b) to match orbital observations (neural ODE learning).

inputs, yielding the corresponding air density at that location. Our proposed architecture, illustrated in Fig. 1, employs a straightforward linear combination of exponential terms reminiscent of isothermal altitude trends. The adjustment of the model parameters to account for latitude, longitude, and solar weather effects is delegated to a lightweight feed-forward neural network. Recent advancements in neural modeling, such as Neural Implicit Representations [11], Physical Informed Neural Networks [15], and DeepONets [16], have demonstrated the impressive representation power of neural models in real-world applications. These approaches extend beyond universal approximation theorems, providing insights into the necessary network size and complexity. Building on these ideas, our thermoNETs incorporate certain aspects of thermosphere physics directly, but they achieve this without relying on integro-differential operators. Instead, our approach involves forcing a functional form, related to underlying physical models, where the important parameters are then represented neurally. The incorporation of physics into the resultant model distinguishes thermoNETs from some of the prior works exploring neural models in analogous contexts [7, 8, 9, 10].

Most of these prior efforts have yielded more intricate architectures, applicable within narrow altitude ranges, and lacking guarantees on the asymptotic behavior of the represented density. These limitations restrict substantially their broader applicability. In contrast, our thermoNETs favor simplicity, utilizing fewer than a few thousand parameters and guaranteeing vanishing values for the density  $\rho$  as the altitude approaches infinity. They demonstrate efficacy across a broad altitude spectrum, spanning from 180 to well beyond 1000 km, and exhibit relative errors limited to a few percentage points when compared to the ground truths they seek to reproduce.

Some aspects of the thermosphere physics are included explicitly by assuming the following functional form for the density as a function of the altitude  $h$ :

$$\rho(h) = \sum_{i=0}^3 \alpha_i \exp(-\beta_i(h - \gamma_i)) \quad (1)$$

where the sum of exponential terms resembles isothermal contributions. The chosen functional form allows it to fit well, at fixed longitudes, latitudes, and space weather parameters, the variation of the air density with the altitude in the thermosphere as highlighted in Fig. 2. In the figure we have generated data using the NRLMSISE-00 model and obtained, at several values of geodetic longitude, latitude, and space weather parameters, the values  $\rho_N$  of the Earth’s atmospheric density from 160 to 1600 km. In each case, we fitted the functional form in Eq. (1) and plotted it against the ground truth together with the relative error introduced by the fit defined as  $\text{err} = \frac{|\rho - \rho_N|}{\rho_N}$ . Starting from altitudes larger than 180km the discrepancy with the NRLMSISE-00 model is well below 1% most of the time.

A thermoNET models the atmospheric density using Eq. (1), where the coefficients  $\alpha_i, \beta_i$  and  $\gamma_i$  are constructed from the output of a small feed forward neural network  $\mathcal{N}_\theta(\lambda, \varphi, \text{DOY}, \text{SID}, \mathcal{S}_w)$  receiving the geodetic coordinates  $\lambda, \varphi$  as input (see Appendix A for a definition of a differentiable geodetic coordinate transformation), as well as the day of year (DOY), the seconds in day (SID) and space weather parameters, here indicated generically with  $\mathcal{S}_w$ .

### B. Datasets

While in this study we primarily focus on the widely used empirical models NRLMSISE-00 (NRL00) [1] from the United States Naval Research Laboratory and the Jacchia-Bowman [2] (JB08) model as target densities, our approach can equivalently be used with any other data source. Both models, to compute a density  $\rho$  in a point of the thermosphere, require the geodetic coordinates  $h, \lambda, \varphi$ , the epoch and space-weather information

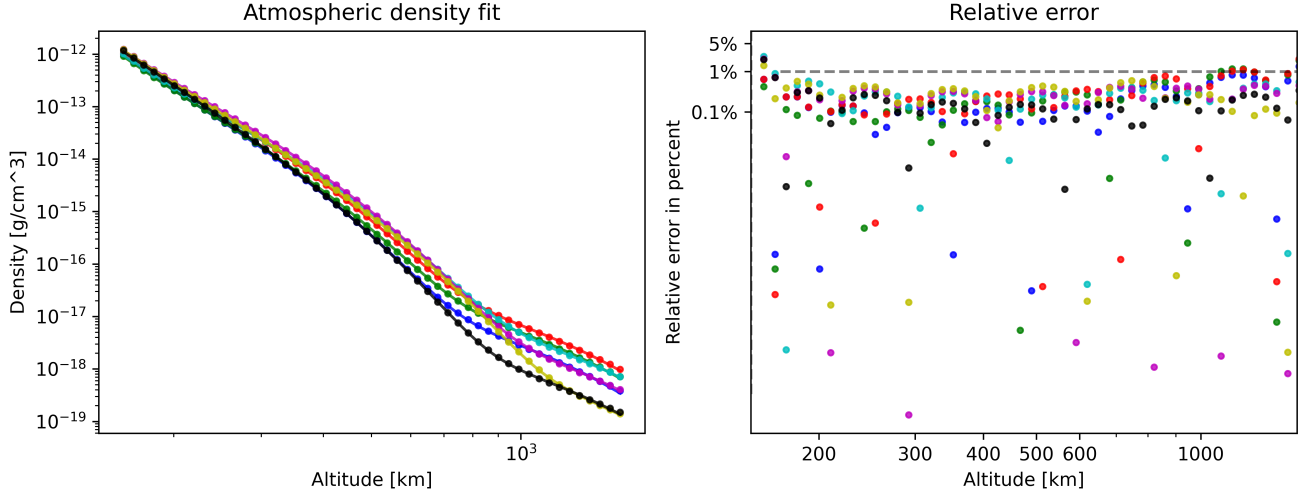


Fig. 2: Left: Ground truth (NRLMSISE-00), and model predictions fitted assuming a fixed longitude, latitude, and solar weather. Each case corresponds to an independent fit of the coefficients in Eq. (1) and the continuous line represents the ground truth, while the overlaid markers represent the fits. Right: relative error introduced by the fits. All interesting cases (above 180km) are mostly well below a 1% error as highlighted by the dotted line.

denoted by the symbol  $\mathcal{S}_w$ . We generate two comprehensive databases of air densities  $\mathcal{D}_{NRL00}$  and  $\mathcal{D}_{JB08}$  using both these empirical models. To construct the databases, we sample the geodetic longitude  $\lambda$  and geodetic latitude  $\varphi$  on a uniform 100x100 grid. For each obtained point, we randomly select an epoch within the years 2009-2022 and extract the corresponding space weather data as provided by the US National Oceanic and Atmospheric Administration (NOAA). Subsequently, utilizing the values of  $\lambda, \varphi$ , and  $\mathcal{S}_w$ , we compute the density  $\rho$  at 100 altitudes within the range of 180 to 1000 km, sampled at 100 points on a logarithmic scale as to favor low altitudes. The resulting databases contain 1 million points, representing a diverse range of epochs, and space weather conditions, and covering the selected altitude range globally. To provide an indication of the accuracy expected when utilizing empirical models to predict density values within the range considered, we report the average relative difference between JB08 and NRLMSISE-00 on the constructed dataset to be approximately 72%.

### C. Learned models

As previously noted, the dependence of the density on altitude is well captured by Eq. (1), which is then enforced in the thermoNET modelling. We thus train two feed-forward neural networks  $\mathcal{N}_{NRL00}$  and  $\mathcal{N}_{JB08}$  on each of the two datasets produced to predict the values of  $\alpha_i, \beta_i$  and  $\gamma_i$ , which are then fed into Eq. (1) to compute a predicted density  $\hat{\rho}$  and hence the loss

$\mathcal{L} = \frac{100}{N} \sum_{i=1}^N |\rho_i - \hat{\rho}_i| / \rho_i$ . Details on the exact definitions of the attributes used and the network outputs, as well as details on the training procedure, can be found in the Appendix B. Both networks, as outlined in Tab. 1, share a similar structure, differing only in the dimension of the input layer which is aligned with the specific space weather indices considered by the corresponding empirical model. With fewer than a few thousand parameters (roughly two orders of magnitude less than comparable models in [8]), they adeptly approximate the target density with errors limited to a few percentage points.

As a first comparison, we utilized our datasets to train two additional neural architectures, denoted as  $\mathcal{K}_s$  and  $\mathcal{K}_b$ , directly predicting the density through a feed-forward neural network (FFNN), akin to the approach employed in the thermosphere neural models implemented in the open-source Karman software<sup>1</sup>, developed by Trillium Technologies in collaboration with the Heliophysics Division of NASA. We trained a small network, matching the number of parameters of thermoNETs experiments, as well as the training setup: we used in both cases 2,000 epochs, decreasing the learning rate from  $10^{-3}$  to  $10^{-4}$  after 1,000 epochs. Moreover, we also trained a much larger network, mirroring the dimensions of the bigger ones presented in [10]. In the latter case, as expected, we were able to surpass for example, the performance levels of  $\mathcal{N}_{NRL00}$  as it achieved final rel-

<sup>1</sup><https://github.com/spaceml-org/karman>

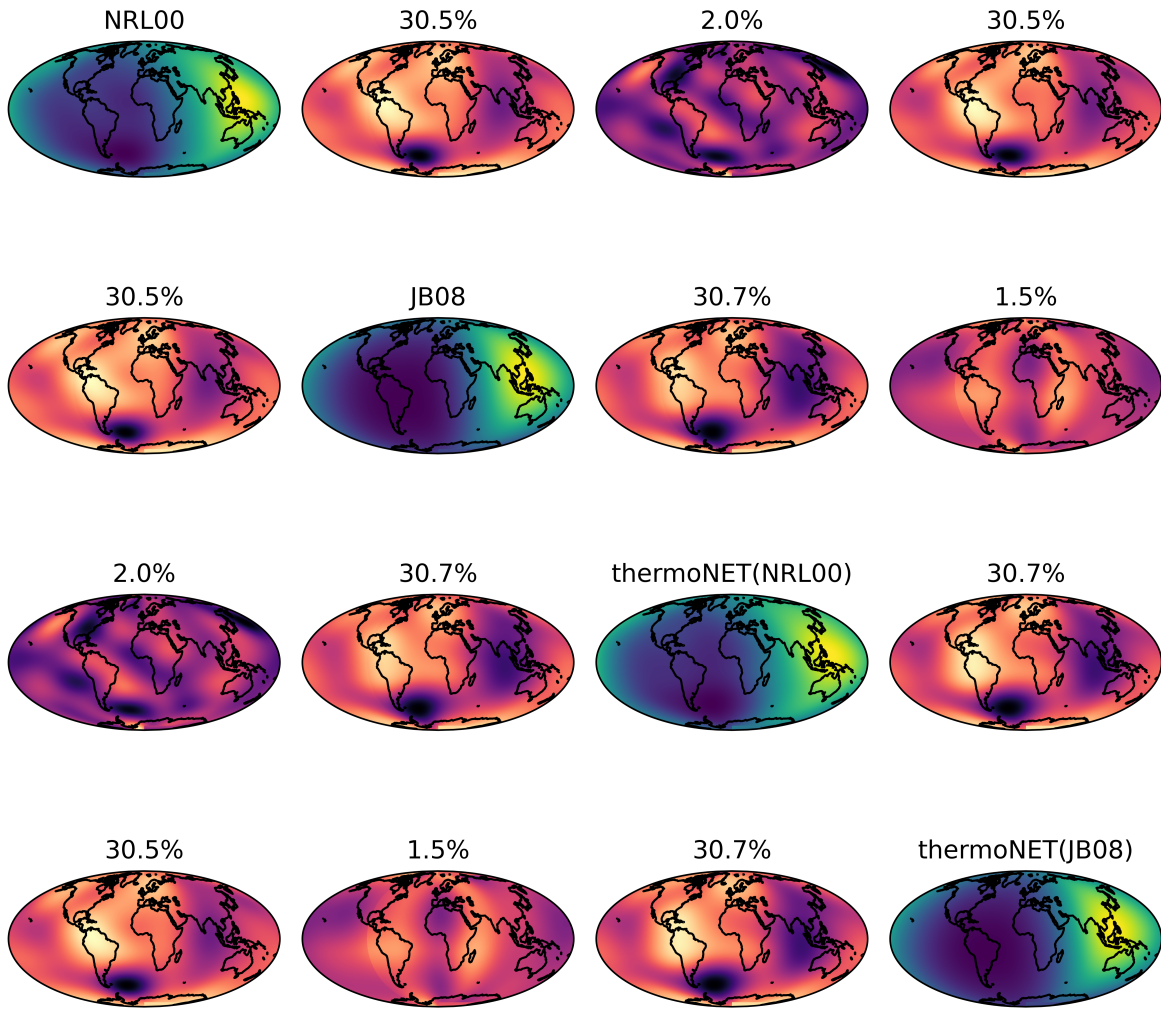


Fig. 3: Prediction (diagonal) and distribution of the absolute percentage difference between models (off-diagonal) at an altitude of 400 km and on the 22nd of April 2018 at 5:13:35 (GMT). The mean is also indicated. The ability of the neural models to approximate well atmospheric trend is evident as the error associated is one order of magnitude smaller than the difference between models.

	n. parameters	average rel. err ( $\mathcal{L}$ )	max rel. err	layers dim	dataset
$\mathcal{N}_{NRL00}$	1804	<b>2.17%</b>	32.93%	10x32x32x12	$\mathcal{D}_{NRL00}$
$\mathcal{N}_{JB08}$	1996	<b>1.43%</b>	21%	16x32x32x12	$\mathcal{D}_{JB08}$
$\mathcal{K}_S$	1977	3.88%	78.18%	11x38x38x1	$\mathcal{D}_{NRL00}$
$\mathcal{K}_B$	257,001	0.98%	15.45%	11x500x500x1	$\mathcal{D}_{NRL00}$

Tab. 1: Details on the thermoNETs developed ( $\mathcal{N}_{NRL00}$ ,  $\mathcal{N}_{JB08}$ ) and networks directly predicting the density as output ( $\mathcal{K}_S$ ,  $\mathcal{K}_B$ ), thus not enforcing Eq. (1).

ative percentage errors of about 0.98%. However, and most importantly, in the other case a significant deterioration of the mean average percentage error was observed (from roughly 2.17% of the physics-informed structure to 3.88%), confirming the intuition that informing the network with the physics model in Eq. (1) helps with the model accuracy, allowing the training and network parameters to ignore the altitude and focus solely on reproducing the other dependencies.

A further test was made comparing the networks inference time to the computational time of the original empirical atmospheric models. Without going into much details, we report that in all tests performed, we observed how the neural models consistently offered a significant speed increase (largely exceeding a factor two). The improvement in computational speed is to be attributed to the network’s small size resulting in a significant compression of the original models.

Finally, we validate our models on test data generated independently from the datasets  $\mathcal{D}$ . To do so we tested them at specific altitudes, epochs, and space weather conditions computing the quality of the reconstructed thermospheric density over the whole range of latitudes and longitudes. In Fig. 3, we show the results for the case of an altitude of 400 km and the epoch 22nd of April 2018, 5:13:35 (GMT). The high quality of the reconstructed densities is confirmed, as well as large differences between empirical models.

### III. EXPLOITING DIFFERENTIABILITY

A thermoNET possesses distinctive properties that set it apart from prior attempts at neural modeling of the thermosphere. These include its compact size, its physically informed nature (as evidenced by Eq. (1)), and the guarantees on its asymptotic behavior. These properties, coupled with the inherent differentiability of the model, open up avenues for applications that would otherwise be unattainable or less effective. In the subsequent two sections, we provide brief introductions to two such applications and present preliminary demonstrations of the anticipated advantages.

#### A. Taylor integration

The Taylor integration technique [17, 18] stands out as one of the most efficient numerical schemes for orbital propagation. Recent advancements in computer implementations [12, 19] have largely overcome the challenges associated with its use, automating the manipulations required to develop such integrators. However, the technique still encounters limitations in cases prevalent in space flight mechanics where non-differentiable terms appear on the right-hand side of the differential equations<sup>2</sup>. In instances where non-differentiable terms are present, a solution to allow the use of the Taylor method involves creating a neural representation of the non-differentiable terms and leveraging it to construct the Taylor integration scheme. This approach was introduced in [20], where a Taylor integrator with event detection was developed for dynamics incorporating complex geometrical eclipse conditions. Similarly, the thermoNETs developed in this study facilitate the incorporation of the non-differentiable NRLMSISE-00 and JB08 models. To illustrate the advantages, let us consider a satellite undergoing Keplerian motion perturbed by thermospheric drag. Its dynamics is described by:

$$\dot{\mathbf{v}} = -\frac{\mu}{r^3}\mathbf{r} - \frac{1}{2} \frac{\rho}{C_b} v_r \mathbf{v}_r \quad (2)$$

where the thermospheric density is indicated by  $\rho$ , the ballistic coefficient with  $C_b$  and the spacecraft relative velocity to the thermosphere with  $v_r$ . For our benchmark, we will ignore complicated wind profiles and thus set:

$$\mathbf{v}_r = \frac{d\mathbf{r}}{dt} - \boldsymbol{\omega} \times \mathbf{r} = \left[ \frac{dx}{dt} + \omega_z y, \frac{dy}{dt} - \omega_z x, \frac{dz}{dt} \right]^T$$

where  $\boldsymbol{\omega}$  is the Earth’s angular velocity around its spin axis. Using these equations of motion, we perform

<sup>2</sup>It is worth noting that from a strictly mathematical point of view the notion of differentiability, in the systems here considered, is generally satisfied almost everywhere. We here consider differentiability also from a computational standpoint. This pertains to the feasibility of natively performing automated differentiation on space flight mechanics code.

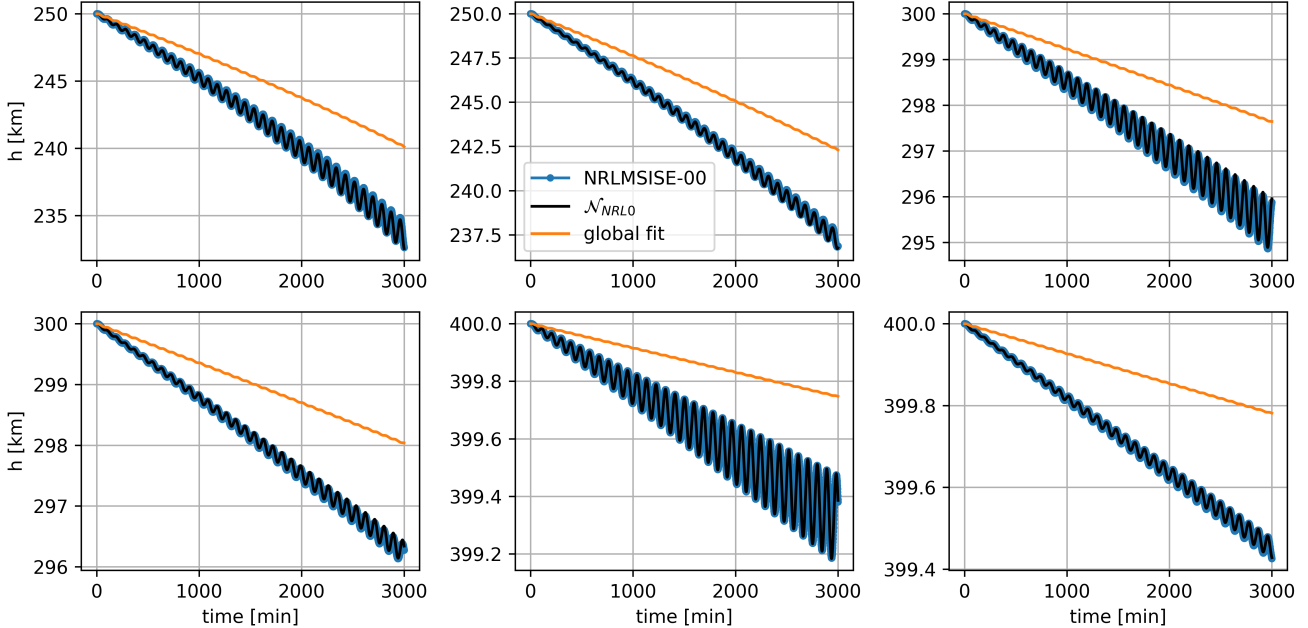


Fig. 4: Altitude trends along various numerical propagations of VLEO satellites where the thermospheric density is represented by a) the NRLMSISE-00 empirical model, b) the thermoNET  $\mathcal{N}_{NRL00}$ , or c) a fit accounting only for altitude dependence. For all practical purposes the thermoNET yields equivalent numerical results.

several numerical orbital propagations of satellites in VLEO computing the thermospheric density  $\rho$  using the NRLMSISE-00 empirical model as well as  $\mathcal{N}_{NRL00}$ . The VLEO are assumed circular and have starting altitudes of 250, 300 and 400 km. For each altitude, we consider both a highly inclined (85 deg.) and an equatorial starting condition. In the case of NRLMSISE-00, where the model lacks differentiability, we resort to the Dormand-Prince numerical scheme for numerical integration. However, for the fully differentiable model  $\mathcal{N}_{NRL00}$ , we opt for a Taylor integration method, implemented in the Heyoka open-source software [20]. The observed speedup (larger than one hundred, see Tab. 2) can be attributed to the combined advantages of the rapid inference time of thermoNETs and the utilization of the Taylor scheme [20]. To distinguish between these two effects, we report in the table one additional experiment measuring the propagation time using the Dormand-Prince scheme while evaluating the thermosphere density with the thermoNETs.

In Fig. 4 we report the altitude trends during the propagation as computed for all cases considered. The vanishingly small error introduced when using the thermoNET is a further validation of the thermospheric models produced.

### B. NeuralODEs training

We here use the term NeuralODE to indicate a set of Ordinary Differential Equations (ODEs) where the right hand side contains an artificial neural network. The term was coined in 2018 [14] and has since been widely adopted in various contexts. While there is not any formal difference between the theory of parametric ODEs and NeuralODEs, the presence of a neural network allows to use the universal approximation theorem [21] to expect a great practical flexibility of the resulting system. In our case we consider the dynamics in Eq. (2) where the expression of the density  $\rho = \rho\theta$  is defined through a neural model parameterized by  $\theta$  (e.g. weights and biases). Thanks to the differentiability of the neural model we derive, in an automated fashion, the variational equations and thus obtain the augmented system:

$$\begin{cases} \dot{\mathbf{r}} = \mathbf{v} \\ \dot{\mathbf{v}} = \mathbf{f}(\mathbf{r}, \mathbf{v}, t) = -\frac{\mu}{r^3}\mathbf{r} - \frac{1}{2}\frac{\rho\theta}{C_b}v_r\mathbf{v}_r \\ \dot{\boldsymbol{\varphi}} = \left[\frac{\partial \mathbf{f}}{\partial \mathbf{r}}, \frac{\partial \mathbf{f}}{\partial \mathbf{v}}\right]^T \boldsymbol{\varphi} + \frac{\partial \mathbf{f}}{\partial \theta} \end{cases} \quad (3)$$

where  $\boldsymbol{\varphi} = \left[\frac{\partial \mathbf{r}}{\partial \theta}, \frac{\partial \mathbf{v}}{\partial \theta}\right]^T$  is the gradient of the trajectory with respect to the ODE parameters and can thus be used to adjust them, via some form of gradient descent, and minimize a loss representing the deviation of the



Numerical Technique	Thermosphere model	CPU time	speedup
Dormand-Prince	NRLMSISE-00	7.18s	x1
Dormand-Prince	$\mathcal{N}_{NRL00}$	1.3s	x5.5
Taylor [20]	$\mathcal{N}_{NRL00}$	0.55ms	x130

Tab. 2: Indicative CPU timings (Intel(R) Xeon(R) CPU E5-2687W v3 @ 3.10GHz) for a single orbital propagation of a VLEO satellite ( $t = 50$  hours,  $h = 200$  km,  $i = 0$  deg.) undergoing a perfectly Keplerian motion perturbed by a drag term.

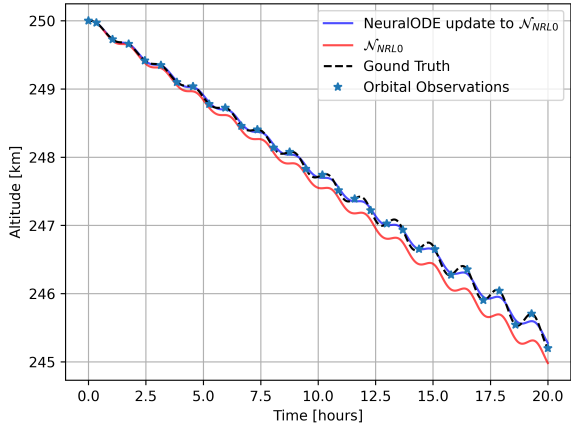


Fig. 5: The updated thermospheric model learns to match orbital observations using the NeuralODE approach and captures the orbital decay rate accurately.

trajectories from observational data. To show the use of this technique (enabled practically by the small dimension of our thermoNETs and their differentiability), we simulate noiseless observations (on the full satellite state) of a ground truth VLEO orbit using a to-be-considered-unknown thermospheric model<sup>3</sup>. We then update our  $\mathcal{N}_{NRL0}$  thermoNET via a few steps of gradient descent aiming to minimise the loss between the synthetic observations and the results of the orbital propagation. The result, shown in Fig. 5, shows how the updated thermospheric neural model captures accurately the ground truth orbital decay rate.

#### IV. CONCLUSION

ThermoNETs are introduced as an innovative implicit neural representation of the density in the thermosphere, characterized by their reduced size, increased accuracy, and guaranteed asymptotic behaviour. We find they

<sup>3</sup>For simplicity we use the JB08 empirical model when simulating a ground truth.

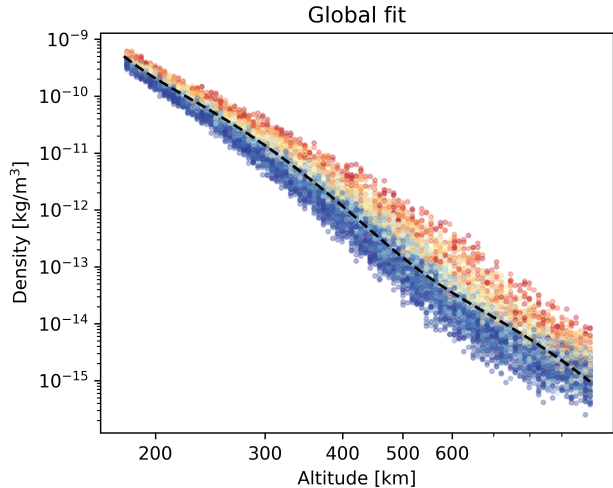


Fig. 6: Global fit (dashed line) and the fitted NRLMSISE-00 data, differing in longitude, latitude and solar weather parameters. Color is indicating the  $f_{10.7}$  coefficient with red being higher (higher solar activity).

have the capability to approximate empirical models such as NRLMSISE-00 and JB-08 within a few percentage points of relative error across a wide altitude range and to offer a computationally efficient and easily manageable differentiable model. Through their combined use with techniques such as Taylor integration and NeuralODEs, ThermoNETs allow for substantial improvements in both propagation times and prediction accuracy for decay rates, particularly when trained with observational data.

#### V. CODE AVAILABILITY

All of the figures and quantitative results in this paper can be reproduced running the python based code distributed freely under a permissive GPLv3 license here: <https://github.com/esa/thermonets>. Refer to this resource to have all the numerical details on our simulations.

## REFERENCES

- [1] Picone, J., Hedin, A., Drob, D. P., and Aikin, A. “NRLMSISE-00 empirical model of the atmosphere: Statistical comparisons and scientific issues.” *Journal of Geophysical Research: Space Physics*, Vol. 107, No. A12, pp. SIA–15, 2002.
- [2] Bowman, B., Tobiska, W. K., Marcos, F., Huang, C., Lin, C., and Burke, W. “A new empirical thermospheric density model JB2008 using new solar and geomagnetic indices.” “AIAA/AAS astrodynamics specialist conference and exhibit,” p. 6438. 2008.
- [3] Bruinsma, S. “The DTM-2013 thermosphere model.” *Journal of Space Weather and Space Climate*, Vol. 5, p. A1, 2015.
- [4] Storz, M. F., Bowman, B. R., Branson, M. J. I., Casali, S. J., and Tobiska, W. K. “High accuracy satellite drag model (HASDM).” *Advances in Space Research*, Vol. 36, No. 12, pp. 2497–2505, 2005.
- [5] Yin, L., Wang, L., Zheng, W., Ge, L., Tian, J., Liu, Y., Yang, B., and Liu, S. “Evaluation of empirical atmospheric models using Swarm-C satellite data.” *Atmosphere*, Vol. 13, No. 2, p. 294, 2022.
- [6] Friis-Christensen, E., Lühr, H., Knudsen, D., and Haagmans, R. “Swarm—an Earth observation mission investigating geospace.” *Advances in Space Research*, Vol. 41, No. 1, pp. 210–216, 2008.
- [7] Pérez, D. and Bevilacqua, R. “Neural Network based calibration of atmospheric density models.” *Acta Astronautica*, Vol. 110, pp. 58–76, 2015.
- [8] Licata, R. J., Mehta, P. M., Weimer, D. R., and Tobiska, W. K. “Improved neutral density predictions through machine learning enabled exospheric temperature model.” *Space Weather*, Vol. 19, No. 12, p. e2021SW002918, 2021.
- [9] Cui, W., Qu, W., Jiang, M., and Yao, G. “The atmospheric model of neural networks based on the improved Levenberg-Marquardt algorithm.” *Open Astronomy*, Vol. 30, No. 1, pp. 24–35, 2021.
- [10] Acciarini, G., Brown, E., Berger, T., Guhathakurta, M., Parr, J., Bridges, C., and Baydin, A. G. “Improving Thermospheric Density Predictions in Low-Earth Orbit With Machine Learning.” *Space Weather*, Vol. 22, No. 2, p. e2023SW003652, 2024.
- [11] Mildenhall, B., Srinivasan, P. P., Tancik, M., Barron, J. T., Ramamoorthi, R., and Ng, R. “NERF: Representing scenes as neural radiance fields for view synthesis.” *Communications of the ACM*, Vol. 65, No. 1, pp. 99–106, 2021.
- [12] Biscani, F. and Izzo, D. “Revisiting high-order Taylor methods for astrodynamics and celestial mechanics.” *Monthly Notices of the Royal Astronomical Society*, Vol. 504, No. 2, pp. 2614–2628, 2021.
- [13] Armellin, R., Di Lizia, P., Bernelli-Zazzera, F., and Berz, M. “Asteroid close encounters characterization using differential algebra: the case of Apophis.” *Celestial Mechanics and Dynamical Astronomy*, Vol. 107, pp. 451–470, 2010.
- [14] Chen, R. T., Rubanova, Y., Bettencourt, J., and Duvenaud, D. K. “Neural ordinary differential equations.” *Advances in neural information processing systems*, Vol. 31, 2018.
- [15] Cai, S., Mao, Z., Wang, Z., Yin, M., and Karniadakis, G. E. “Physics-informed neural networks (PINNs) for fluid mechanics: A review.” *Acta Mechanica Sinica*, Vol. 37, No. 12, pp. 1727–1738, 2021.
- [16] Lu, L., Jin, P., Pang, G., Zhang, Z., and Karniadakis, G. E. “Learning nonlinear operators via DeepONet based on the universal approximation theorem of operators.” *Nature machine intelligence*, Vol. 3, No. 3, pp. 218–229, 2021.
- [17] Jorba, À. and Zou, M. “A software package for the numerical integration of ODEs by means of high-order Taylor methods.” *Experimental Mathematics*, Vol. 14, No. 1, pp. 99–117, 2005.
- [18] Scott, J. R. and Martini, M. C. “High-speed solution of spacecraft trajectory problems using Taylor series integration.” *Journal of Spacecraft and Rockets*, Vol. 47, No. 1, pp. 199–202, 2010.
- [19] Pérez-Hernández, J. A. and Benet, L. “PerezHz/TaylorIntegration.jl: TaylorIntegration v0.4.1.”, Feb. 2019. doi:10.5281/zenodo.2562353.
- [20] Biscani, F. and Izzo, D. “Reliable event detection for Taylor methods in astrodynamics.” *Monthly Notices of the Royal Astronomical Society*, Vol. 513, No. 4, pp. 4833–4844, 2022.
- [21] Tikk, D., Kóczy, L. T., and Gedeon, T. D. “A survey on universal approximation and its limits in soft computing techniques.” *International Journal*



of Approximate Reasoning, Vol. 33, No. 2, pp. 185–202, 2003.

- [22] Featherstone, W. and Claessens, S. “Closed-form transformation between geodetic and ellipsoidal coordinates.” *Studia geophysica et geodaetica*, Vol. 52, pp. 1–18, 2008.
- [23] Slater, J. A. and Malys, S. “WGS 84—past, present and future.” “Advances in Positioning and Reference Frames: IAG Scientific Assembly Rio de Janeiro, Brazil, September 3–9, 1997,” pp. 1–7. Springer, 1998.

## APPENDIX

### A. THE GEODETIC COORDINATES

In order to compute the required thermoNETs inputs from satellite positions, we must invert the relations:

$$\begin{aligned} x &= (N + h) \cos \varphi \cos \lambda \\ y &= (N + h) \cos \varphi \sin \lambda \\ z &= ((1 - e^2)N + h) \sin \varphi \end{aligned}$$

where  $e^2 = 1 - \frac{b^2}{a^2}$ ,  $N = \frac{a}{\sqrt{1 - e^2 \sin^2 \varphi}}$ .  $a$  and  $b$  are the semi-major and semi-minor axis of the oblate ellipsoid the Geodetic coordinates refer to. The inverse relation, instead, can be written as:

$$\begin{aligned} h &= \frac{p}{\cos \varphi} - N \\ \varphi &= \arctan \left( \frac{z}{p(1 - e^2 \frac{N}{N+h})} \right) \\ \lambda &= \arctan 2(y, x) \end{aligned}$$

where  $p = \sqrt{x^2 + y^2}$ . These last relations require an iterative procedure in order to be solved as  $\varphi$  appears on the right hand side explicitly via the definition of  $N$ . The discussion on how to compute geodetic coordinates from Cartesian coordinates has been profusely discussed in the literature over the past 40 years and can be considered as a solved problem for our purposes (see [22] for a review of various approaches). We employ Algorithm 1 to perform the transformation. Note that we use a fixed number of iterations, namely four, which allow us to use the iterations to establish analytical and differentiable relations between the two sets of coordinates:  $h(x, y, z)$ ,  $\varphi(x, y, z)$  and  $\lambda(x, y, z)$ . In the case of the Earth ( $e = 0.08181919$  from the WGS-84 ellipsoid [23]), we find errors smaller than 70m when using 2 iterations, 40 cm when using 3 iterations, and 4mm when using 4 iterations.

### B. TRAINING DETAILS

For both networks,  $\mathcal{N}_{NRL00}$  and  $\mathcal{N}_{JB08}$ , we use as inputs the two geodetic coordinates  $\lambda, \varphi$  (the altitude is taken care of by Eq. (1)), the day of year (DOY) the seconds in day (SID) (a proxy for the Local Solar Time) and the relevant solar weather parameters  $\mathcal{S}_w$ . Similarly to what is done in [8] we avoid angular discontinuities by augmenting the attributes as follows:

$$\begin{aligned} t_1 &= \sin(\lambda), t_2 = \cos(\lambda) \\ t_3 &= \sin\left(2\pi \frac{DOY}{365.25}\right), t_4 = \cos\left(2\pi \frac{DOY}{365.25}\right) \\ t_5 &= \sin\left(2\pi \frac{SID}{86400}\right), t_6 = \cos\left(2\pi \frac{SID}{86400}\right) \end{aligned}$$

In the case of  $\mathcal{N}_{NRL00}$  the space weather indices used as attributes are  $\mathcal{S}_w = [f_{10.7}, f_{a10.7}, A_p]$ , representing the

---

#### Algorithm 1 Transforming Cartesian to Geodetic coordinates

---

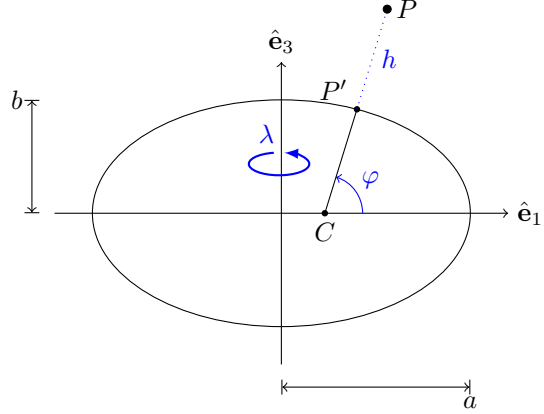
**Require:**  $n \geq 0$

```

 $\lambda \leftarrow \arctan 2(y, x)$ 
 $p \leftarrow \sqrt{x^2 + y^2}$ 
 $e2 \leftarrow 1 - \frac{b^2}{a^2}$ 
 $\varphi \leftarrow \arctan \frac{z}{p(1 - e^2)}$ 
for  $i = 1..4$  do
   $N \leftarrow \frac{a}{\sqrt{1 - e^2 \sin^2 \varphi}}$ 
   $h \leftarrow \frac{p}{\cos \varphi} - N$ 
   $\varphi \leftarrow \arctan \frac{z}{p(1 - e^2 \frac{N}{N+h})}$ 
end for

```

---



solar radio flux at 10.7 cm ( $f_{10.7}$ ), its 81 days average ( $f_{a10.7}$ ) and an average level for the geomagnetic activity ( $A_p$ ). For  $\mathcal{N}_{JB08}$  the space weather indices considered are instead those used in the Jacchia-Bowman empirical model and hence include more solar flux descriptors [ $f_{10.7}, s_{10.7}, m_{10.7}, y_{10.7}$ ], their 81 day averages as well as the geomagnetic indicators  $A_p$  and  $Dst$ . The network architecture uses tanh non-linearities and include, after the input layer, two hidden layers of 32 neurons each. The thermoNETs outputs 12 quantities  $\delta\alpha_i, \delta\beta_i, \delta\gamma_i, i = 1..4$  which reconstruct a density value using Eq.(1), where:

$$\begin{aligned} \alpha_i &= \bar{\alpha}_i(1 + \delta\alpha_i) \\ \beta_i &= \bar{\beta}_i(1 + \delta\beta_i) \\ \gamma_i &= \bar{\gamma}_i(1 + \delta\gamma_i) \end{aligned} \quad (4)$$

The reference values  $\bar{\alpha}_i, \bar{\beta}_i, \bar{\gamma}_i$  are obtained fitting Eq. (1) the whole dataset. In Fig. 6 we report the trend of such a global fit (only dependent on the altitude) used to predict the density values in the  $\mathcal{D}_{NRL00}$  dataset. The networks inputs are min-max scaled and the parameters are trained using the Adam optimizer with a learning rate scheduler.