AAS 20-571

USING ARTIFICIAL NEURAL NETWORKS FOR OFFLINE GRAVIMETRY

J. R. Martin[*] and H. Schaub[†]

Resolving a reliable, high-fidelity mapping between position and gravitational acceleration is paramount for accurate trajectory design and mission planning. These mappings are typically constructed through complex formulations that include large spherical harmonic expansions or polyhedral models that require thousands of facets. Such models are expensive to compute, difficult to extract, and demand particular assumptions of the gravitational environment or operational conditions of the effected spacecraft. Recent literature suggests that artificial neural networks may be capable of circumventing such limitations – however it is unclear if such networks can model the dominant, surface level perturbations. This paper attempts to fill that hole, looking specifically at how well these networks can model perturbations beyond J_2.

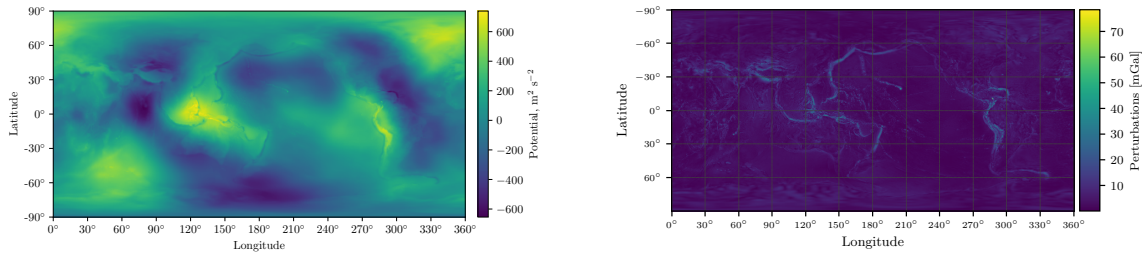
INTRODUCTION

The need for high-fidelity models of the gravity field for a given celestial body is critical for astrodynamicists, as these estimates are responsible for generating precise trajectories on the ground and ensuring spacecraft remain on those trajectories during flight. The means to computing these fields hinge on the availability of an accurate representation of the potential which is typically provided as a series expansion of spherical harmonics. Pending that dynamicists have access to sufficiently many spherical harmonic coefficients, $C_{l,m}$ and $S_{l,m}$, the potential is expressed to some finite degree l and order m from which the gradient is taken to produce the corresponding gravitational acceleration.[1,2]

The utility of the spherical harmonic model lies in the fact that major gravitational perturbations like those caused by a planet's oblateness can be captured with only a few terms. Moreover, the convergence and accuracy of the model can be well-parameterized by the highest harmonic in the series and distance from the body.[3,4] As such, significant energy and resources have been spent resolving spherical harmonic coefficients for bodies like Earth and the Moon. Missions like GRACE, GRACE-FO, GRAIL, and CHAMP each sought to resolve increasingly high-order spherical harmonic models for these masses.[5–9] The highest-resolution model generated by such missions (in conjunction with supplemental measurements from the ground) is the time-invariant EGM2008 model which maps Earth's potential to degree and order 2190.[10] These potential models are valuable data products not just for astrodynamicsts but also planetary scientists – providing insights into

[*]NSF Graduate Research Fellow, Ann and H.J. Smead Department of Aerospace Engineering Sciences, University of Colorado, Boulder, 431 UCB, Colorado Center for Astrodynamics Research, Boulder, CO, 80309.

[†]Glenn L. Murphy Chair of Engineering, Ann and H.J. Smead Department of Aerospace Engineering Sciences, University of Colorado, Boulder, 431 UCB, Colorado Center for Astrodynamics Research, Boulder, CO, 80309. AAS Fellow, AIAA Fellow.



(a) Potential, U

(b) Acceleration Norm, $||\nabla U||$

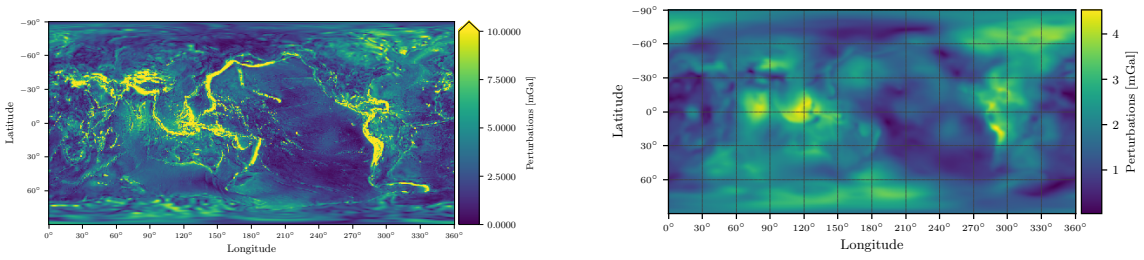
Figure 1: The EGM2008 spherical harmonic model expressed to degree and order 1000 at the reference sphere $R_0 = 6378136.3$m having removed contributions below $C_{2,2}$ and $S_{2,2}$ of the series expansion.

other planetary mechanics like mantle flow, atmospheric circulation, and total landmass distributution.[11–13]

Despite the benefits of the spherical harmonic basis, this representation of the potential does come with unique challenges. Foremost, spherical harmonics are costly to evaluate on-the-fly. Popular algorithms that compute the gradient of these models scale as $\mathcal{O}(l^2)$ making them increasingly complex and burdensome to calculate as models grow more accurate.[14] In addition, these models are challenging to regress, requiring large volumes of precise data for high-order models. Moreover, one of the primary conveniences of the model – the assumption that the body's dominant perturbations have smooth, spherical symmetry – is often only true at the coarsest of scales. As modern missions aim to explore more irregularly-shaped bodies like asteroids and comets, where perturbations from surface obstacles or discontinuous features can skew dynamics, the convenience of such a model becomes less clear.

To characterize the performance of the spherical harmonic model in these regimes, one must first identify the unique and dominant surface-level perturbations of a given celestial body – henceforth referred to as gravitational features. For large bodies like planets and moons, this demands excluding the point mass and oblateness terms within the expansion – corresponding to all coefficients up to and including $C_{2,2}$ and $S_{2,2}$ – as they are features that are unanimous for nearly all large-scale celestial bodies, and are not representative of unique perturbations on a given surface. Removing these low-order terms from the EGM2008 model and projecting the potential and corresponding acceleration magnitude onto Earth's Brillouin sphere, the bounding sphere enclosing all mass elements, produces Figure 1.[15] Figure 1 shows that unique gravitational features are those of finer, discontinuous regions on the surface. For Earth these regions include the Andes, Himalayas, and edges of the Pacific, Caribbean, and Indian tectonic plates. Such observations are consistent with the physics, as dominant perturbations come from steep gradients in the potential – generated by sudden changes in mass density or mass displacement.

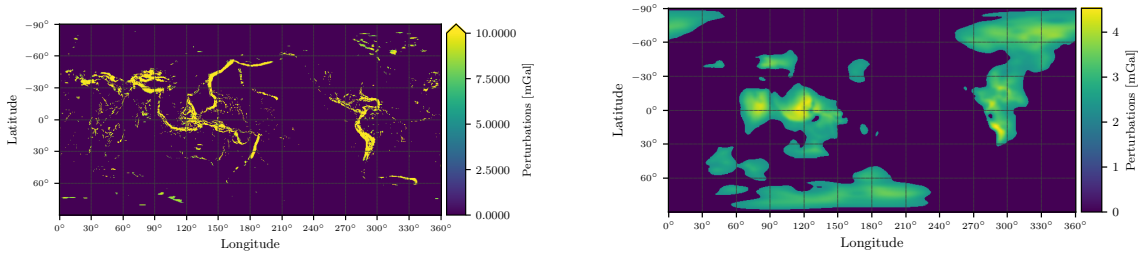
For the astrodynamicist, it is important to model these features accurately as failure to account for such perturbations can lead to erroneous dynamics over time. Using a spherical harmonic model, however, this becomes difficult. Just as a Fourier transform struggles to converge near discontinuities like the edges of a square-wave, spherical harmonics also struggle to converge in regions of discontinuity. To accurately capture these features in a spherical harmonics representation, many weakly signaled, interfering terms in the series expansion are needed to suppress the three dimen-



(a) $||\nabla U||$ at Brillouin Sphere, R_0

(b) $||\nabla U||$ at LEO Sphere, R_{ISS}

Figure 2: Gravitational features as seen at the Brillouin sphere and at LEO.



(a) R_0 Feature Mask

(b) R_{ISS} Feature Mask

Figure 3: Gravitational features masks at the Brillouin sphere and at LEO.

sional analog of Gibbs phenomenon.[16,17] As such, despite the the analytic convenience of the spherical harmonic representation, it can fail the astrodynamicist when they need it most – in regions of dominant gravitational perturbations.

The natural question becomes, is there a better way to model these gravitational features? If one could instead design a basis around dominant features rather than common geometries, then the need to resolve high-order models could be avoided in the first place. More boldly, with such a basis there may not even be a need to extract an explicit representation of potential, and researchers could map directly from position to acceleration without an intermediate model.

The solution may exist through the use of artificial neural networks (ANN). ANNs have grown increasingly popular among the regression community as a means to resolve a non-linear mapping from an input space to a desired output space through a collection of hidden layers.[18] The convenience of such a method is that the basis is not prescribed by the user, but is instead learned from training data. Provided enough flexibility in the network (sufficiently many layers and nodes per layer) the user can pass position data as input and acceleration measurements as outputs and the ANN can regress an optimal basis that spans the directions of maximal variance.[19] For gravimetry, this means rather than requiring large spherical harmonic expansions to produce an accurate gravitational model of a tectonic shelf or a mountain range and then computing the acceleration they impart on a spacecraft orbiting overhead – ANNs may ofter a more compact representation of those features, bypassing the representative and computational inefficiencies of the spherical harmonic basis entirely.

Modeling gravity fields using machine learning techniques for astrodynamics purposes is a rel-

atively new technique and the literature thus far has placed a heavy emphasis on its computational efficiency over analytic methods. Gao and Liao show that Gaussian processes can model the irregular gravity field of an asteroid in only 10^{-1} seconds compared to the 10^4 seconds needed to model the equivalent polyhedral model.[20] Furfaro et. al. uses extreme learning machines to model the gravity field of an asteroid both globally, but also in localized regions for use in proximity operations.[21] Cheng et. al. uses ANN to represent gravity field of an asteroid and uses that to facilitate the design of optimal control landing algorithms.[22] Each effort also discusses the accuracy of their machine learning models – some citing errors as low as 1.27% – however the accuracy was always measured relative to the global gravity scale. That is, the errors were defined relative to the total gravitational acceleration which included the contributions from μ/r^2 and celestial scale features like J_2. In most cases, including just the point mass contribution in the relative error calculation biases the errors to be quite low. For example, excluding J_2 in Earth's field causes $\ll 1\%$ error in total gravitational magnitude, so for an ANN or Gaussian process to resolve the gravitational field to $< 1\%$ accuracy is not necessarily proving a more compact or powerful representation than that of traditional spherical harmonics.

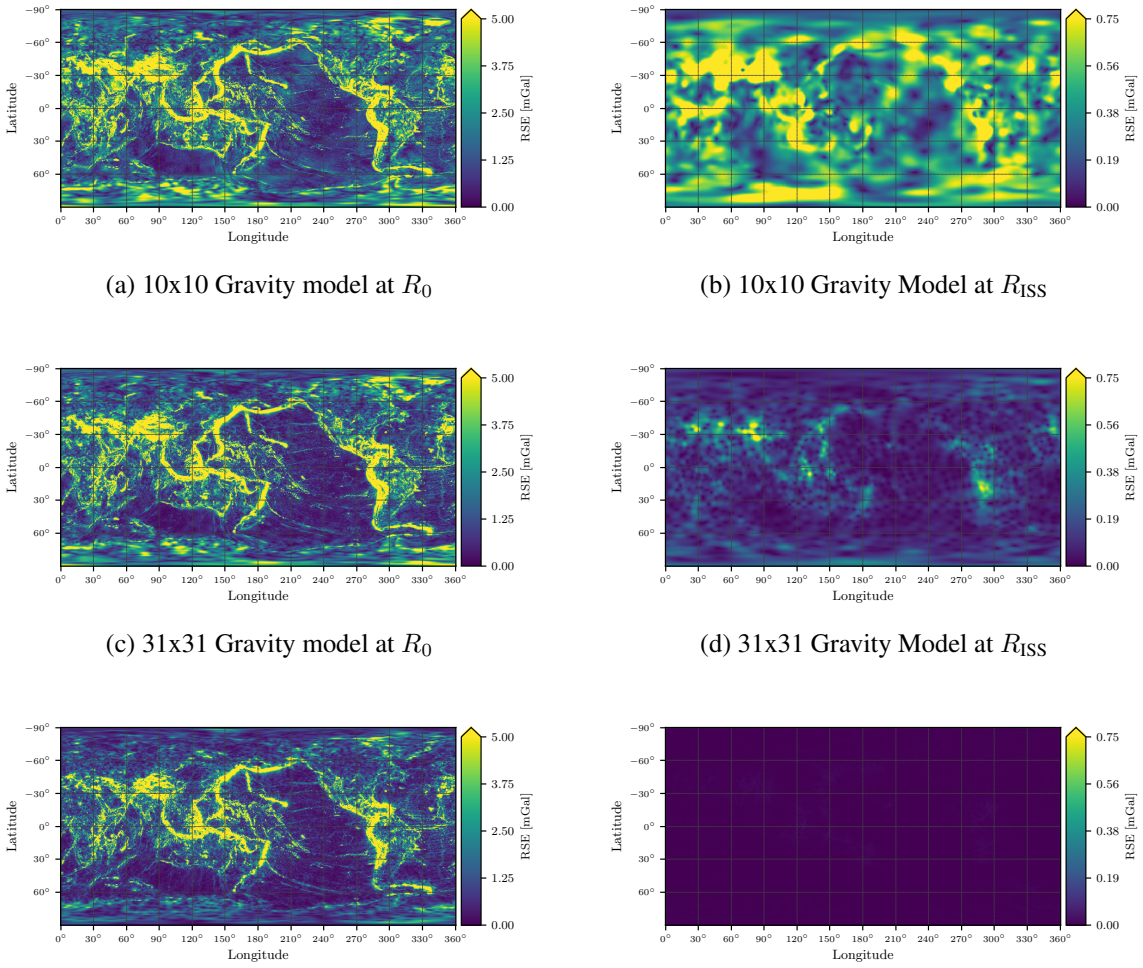
This study attempts to investigate and clarify such claims by looking specifically at how well ANN model gravitational features. By looking beyond the the point mass and oblateness terms, it becomes easier to quantify how well ANN learn the sub-1% features and if such machine learning models should be given credence for this application.

SPHERICAL HARMONIC GRAVITY FIELD

To determine how well spherical harmonics and ANNs can express gravitational features, the features must first be quantitatively defined. Using a uniformly distributed grid of 175x350 points, a map of the gravitational accelerations produced by the EGM2008 model at the Brillouin sphere ($R_0 = 6378136.3$m) is produced in Figure 1b. From the data, the gravitational features are selected by forming a mask around all accelerations in that map which exceed 3σ above the average. The mask of the selected features at the Brillouin sphere is shown in Figure 3a and best highlights the sources of these features as there is no attenuation from the the $\frac{R_0}{r}^l$ term in the potential. It is worth noting that some features within this mask reach as high as 70 mGal, but are capped at 10 mGal in the image to show greater structure in the features. Because there is attenuation at higher altitudes, a second acceleration map and corresponding feature mask is generated at radius $R_{\text{ISS}} = 6708136.1$m as shown in Figures 2b and Figure 3b respectively to show how the features smooth and decay at higher altitudes.

To demonstrate the efficiency with which the spherical harmonic representation captures these gravitational features, three spherical harmonic model fidelities are tested – degree and order 10, 31, and 100. These fidelities are chosen to show how an order-of-magnitude increase in total coefficients (110, 992, and 10100 coefficients respectively) improves the modeling of the gravitational features at each altitude. The accuracy of these models are gauged by computing the root square error (RSE) at each point in the original map (see Figure 4), then averaging the total error across the map and also separately within just the feature mask as plotted. The results of these averages are presented in Figure 5.

Figures 4, 5 show that the gravitational features consistently struggle to be expressed in a spherical harmonic basis despite order-of-magnitude increases in coefficients. Features like the Andes, Himalayas, and shelves off the Pacific coast remain visibly erroneous even at LEO altitudes – where their signal is smoothed, a condition which is otherwise advantageous for spherical harmonics. Near



(a) 10x10 Gravity model at R_0

(b) 10x10 Gravity Model at R_{ISS}

(c) 31x31 Gravity model at R_0

(d) 31x31 Gravity Model at R_{ISS}

(e) 100x100 Gravity model at R_0

(f) 100x100 Gravity Model at R_{ISS}

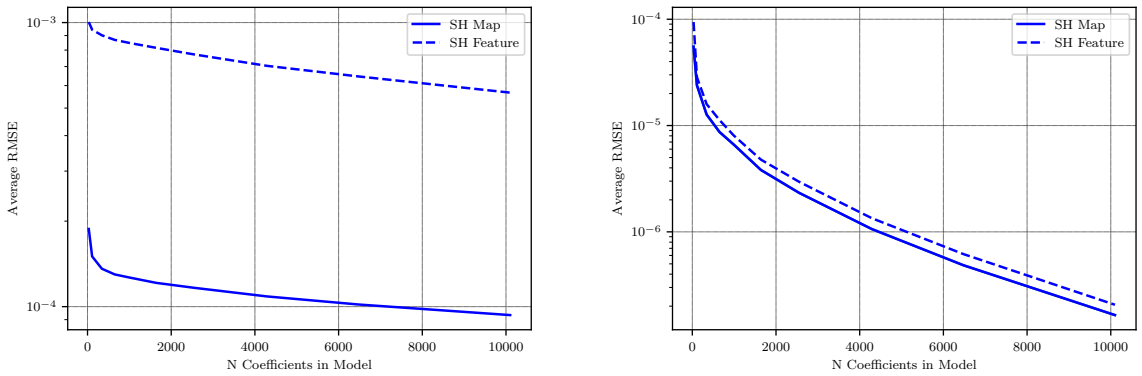
Figure 4: Root squared error of gravitational features produced by lower fidelity spherical harmonic models with respect to the EGM2008 1000x1000 field.

the Brillouin sphere, the error is even more exaggerated despite the increased coefficient counts.

These findings suggest that spherical harmonics are not a natural basis to represent surface level features. After the first few terms in the expansion, the error gradient grows increasingly flat and the case for using higher order models is not especially compelling. This highlights the need to explore alternative bases in efforts to more efficiently represent these perturbations for astrodynamics applications.

ARTIFICIAL NEURAL NETWORKS

Artificial neural networks (ANN) provide a unique alternative to the spherical harmonic basis for representing these features. Rather than prescribing coefficients based on the total correlation of the potential against some geometries like spherical harmonics, ANNs do not assume similarity or correlation with any particular geometry. Instead, ANNs learn a basis that naturally minimizes a chosen cost function. In this experiment, the cost function is the mean squared error (MSE),



(a) RSE R_0Gravity Model

(b) RSE LEO Gravity Model

Figure 5: Average root squared error of the gravitational feature within the feature mask and across the entire map as a function of number of coefficients within the spherical harmonic (SH) model.

and such a choice forces a sensitivity to outliers.[23] Using the MSE cost function, the ANN must update its weights to accommodate the large gravitational features first to minimize its total cost. As such the ANN generates a basis that prioritize features based on their magnitude rather than their geometry.

To test this hypothesis in a way that can be compared to spherical harmonics, the ANNs investigated are designed to have approximately the same number of regression parameters (weights) as there are total coefficients in the spherical harmonic models of Figure 5. If the ANN representation can prioritize the features well, the RSE produced by the ANN should be of similar or lesser magnitude than that of the spherical harmonic models, particularly within the feature mask.

Given that ANN can be designed in multiple different configurations, this study limits its investigation to six specific ANNs which were individually optimized and tested. The first three ANN are shallow, single-layer, feedforward networks (SLFNs) and the second three are three-layer deep neural networks (DNNs). These classes of ANN architectures are chosen both for their popularity for regression tasks as well as their contrasting advantages. SLFNs are popular for their simple design, fast execution and learning speeds, whereas deep neural networks have demonstrated impressive accuracy in variety of applications at the cost of interpretability and long training times.[24–27]

Each of these ANN begin with a three-node input layer which corresponds to the position vector and end with a three-node output layer corresponding to the acceleration vector. The number of nodes per hidden layer, N, in both architectures are calculated assuming one bias parameter per node beyond the input layer, plus the additional weights connecting the current node with nodes from the prior and subsequent layer. As such, the width of the SLFN can be computed by equating

$$3N + 3N + N + 3 = M \tag{1}$$

where M is the number of coefficients of an equivalent spherical harmonic model (100, 1000, 10000 coefficients). Likewise, the number of nodes per hidden layer in the three-layer deep neural network can be found using

$$3N + (3-1)N^2 + 3N + 3N + 3 = M \tag{2}$$

The corresponding layer sizes, total trainable parameters, and the number of spherical harmonics

coefficients being equated to are presented in Table 1.

Table 1: ANN Architectures

ANN Type	# Hidden Layers	# Nodes per Hidden Layer	Total Params	# SH Coef Equivalent
SLFN	1	13	94	100
SLFN	1	141	990	1000
SLFN	1	1428	9999	10000
DNN	3	5	98	100
DNN	3	20	983	1000
DNN	3	68	9863	10000

Training Data

The study only aims to investigate the representative ability of the ANN approach, therefore no bounds were placed on the amount of training data that could be provided to the network. Moreover, the study prioritizes capturing features at the Brillouin sphere R_0 where the gravitational features are of highest magnitude and greatest discontinuity.

The first data set used to train the ANNs consisted of 259,200 position vectors as inputs and their corresponding 259,200 acceleration vectors as output. The acceleration vectors were calculated using the EGM2008 gravity model expressed to degree and order 1000 using the non-singular Pines formulation.[1] While the EGM2008 model includes higher degrees, the model was truncated at degree 1000 for its computational tractability and because its Nyquist rate is significantly higher than that of the sample grid to avoid aliasing. The position vectors are randomly sampled in latitude and longitude, and constrained to a 0-5km altitude shell above the Brillouin sphere. Given that the experiment seeks to evaluate how well ANNs can resolve gravitational features, the contributions to the acceleration from terms $C_{2,2}$, $S_{2,2}$ and below were also removed from the outputs vectors prior to training.

The data set is divided into a training set (70% of all data) and a validation set (30%). The data was preprocessed by first converting to spherical coordinates, and then performing a min-max normalization for each input component to fit the bounds of $[0, 1]$. The min-max transformation is chosen due to the finite domain for the training inputs. Specifically, θ and ϕ coordinates can only scale between $[0°, 360°]$ and $[0°, 180°]$ respectively, therefore is reasonable to normalize those to bounds $[0, 1]$. Similarly, a standard normalization was applied to each output component such that the mean of each component becomes 0 and its standard deviation becomes 1. A standard transform was used for the output data as the accelerations can span a continuous space. Moreover, the standard transform extenuates features as outliers which will increase their contribution to the ANN cost thereby forcing greater accommodation from the ANN. Such data normalizations techniques are common to standardized features and accommodate most ANN activation functions which typically span the domain of [0,1] or [-1, 1]. Normalization has been shown to affect convergence rate and final performance of ANNs.[28] The validation data was subsequently scaled, based on the results of the training data transformation.

Hyperparameter Optimization

Each ANN is optimized individually using Talos*. Talos is a hyperparameter optimization python package that allows users to sample across a defined hyperparameter space and report the sensitiv-

*Autonomio Talos [Computer software]. (2019) Retrieved from http://github.com/autonomio/talos.

ities of those hyperparameters with respect to performance metrics like the validation loss. Using that sensitivity information, the user can then refine the parameter space to find a heuristic optimum to be used in production.

This study used Talos' random search hyperparameter optimization algorithm which is considered more efficient than a traditional grid search methods for problems with low intrinsic dimensionality.[29] Such a method recommends sampling only a fraction of the total hyperparameter space to understand coarse sensitivities before attempting finer tuning. The initial hyperparameter search space used for each configuration is presented in Table 2. The two parameters worth noting are the optimizers and activation functions. Only the Adadelta and Nadam optimizers are tested for their adaptive learning rates, and only the exponential linear unit and rectified linear units activation functions are used to avoid vanishing gradients during backpropogation.

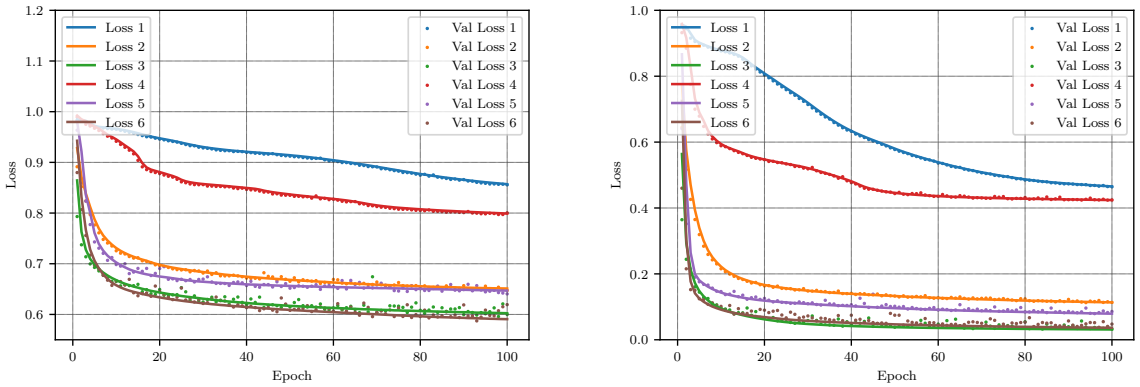
After randomly sampling 25% of this hyperparameter space, the top performing ANNs are selected to compete against their equivalent spherical harmonic model. This optimization strategy should not be considered extensive, and it is reasonable to assume that another ANN with the same number of parameters may be capable of better performance. The final parameters selected for each model can be found in Table 3. With the optimal parameters selected, each ANN is retrained for 100 epochs and with batch size of 1 on the original data set spanning the 0-5km shell. Longer training times allow for better convergence and lower batch sizes can improve generalizability of the model, more efficiently exposing idiosyncratic data/features to the network. In addition, a separate ANN of the same hyperparameter configuration, design, and training length / batch size is trained on a shell of data ranging from $R_{\text{ISS}} - 2.5\text{km}, R_{\text{ISS}} + 2.5\text{km}$ to investigate performance on attenuated features. The loss curves generated by each ANN are provided in Figure 6. Such figures show that the ANNs have not overfit to their training data as the validation loss closely follows the training loss. Moreover, each ANN has converged to some local minimum as demonstrated by the plateaued loss.

Table 2: Initial hyperparameter search space

Hyperparameter	Search Space
Batch Size (BS)	[10, 30, 50]
Epochs	[100]
Dropout	[0.0, 0.3]
Learning Rate (LR)	[0.1, 0.2, 0.3]
Kernel Initializer (KI)	[glorot normal, glorot uniform]
Kernel Regularizer (KR)	[l2, l1, None]
Optimizer	[Nadam, Adadelta]
Shapes	[brick]
Losses	[mean squared error]
Activation	[relu, elu]

Table 3: Final hyperparameters for each ANN configuration

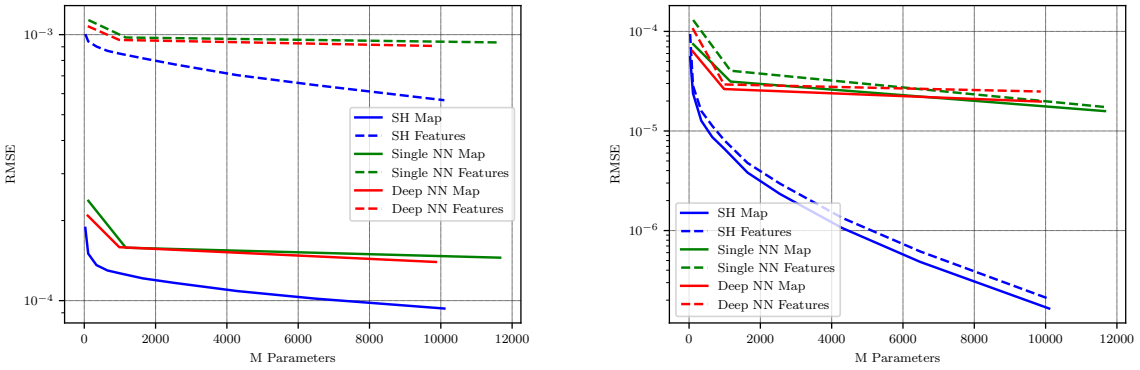
Case	ANN Type	Optimizer	Activation	KI	KR	Dropout	LR	BS
1	SLFN	Nadam	elu	glorot normal	None	0.0	0.30	10
2	SLFN	Adadelta	relu	glorot uniform	None	0.0	0.30	10
3	SLFN	Nadam	relu	glorot uniform	None	0.0	0.20	10
4	DNN	Adadelta	relu	glorot uniform	None	0.0	0.30	10
5	DNN	Adadelta	relu	glorot normal	None	0.0	0.25	10
6	DNN	Nadam	relu	glorot normal	None	0.0	0.25	10



(a) Loss for R_0 distribution

(b) Loss for R_{ISS} distribution

Figure 6: Loss functions for each ANN trained from distributions drawn from R_0 + [0, 5] km and R_{ISS} + [-2.5, 2.5] km



(a) RSE R_0 Gravity Model

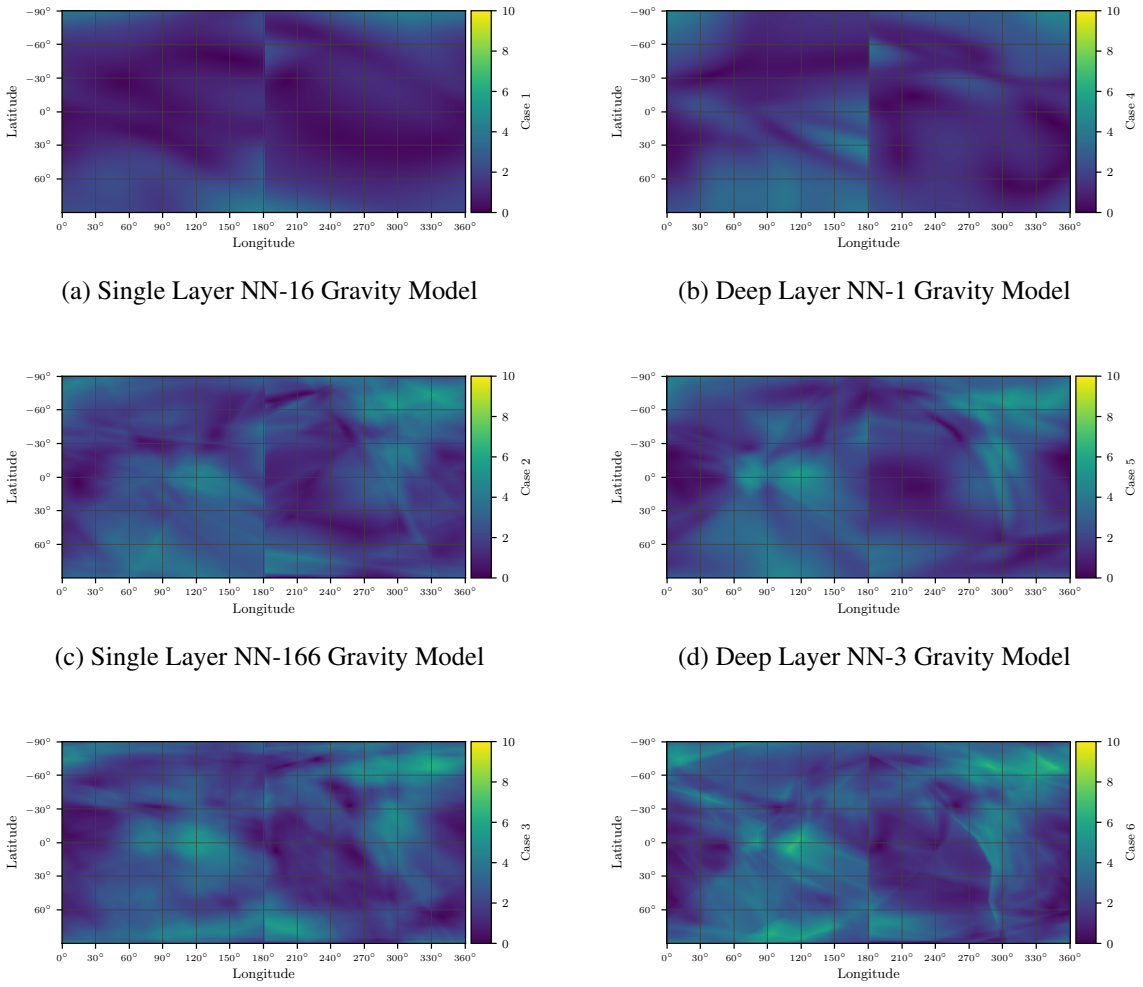
(b) RSE R_{ISS} Gravity Model

Figure 7: Average root squared error of gravitational feature mask and entire grid as a function of number of coefficients for both the spherical harmonic (SH) and ANN models

RESULTS

To evaluate the performance of the ANN model against the spherical harmonic equivalent representation, the trained networks were tested on the same data set used to generate Figure 3. These are 175x350 map grids sampled uniformly in latitude and longitude at R_0 and R_{ISS}. Just as the spherical harmonic RSE was averaged across the entire map and within the feature mask, the ANN are evaluated in the same way and their results are provided in Figure 7.

The errors are unsatisfactory. Each ANN, regardless of configuration, resulted in higher error for the same number of parameters as the equivalent spherical harmonic model – both in their average across the map and within the feature mask. While the optimization of the ANNs was not exhaustive, these early results suggest that ANN are not adequately suited to model gravitational features even as they attenuate at higher altitudes. Intuitively this is reasonable, the ANNs are



(a) Single Layer NN-16 Gravity Model

(b) Deep Layer NN-1 Gravity Model

(c) Single Layer NN-166 Gravity Model

(d) Deep Layer NN-3 Gravity Model

(e) Single Layer NN-1666 Gravity Model

(f) Deep Layer NN-4 Gravity Model

Figure 8: Generated output by each of the NN at the Brillouin sphere

attempting to accommodate the infrequent, yet extremely high-magnitude perturbing features while also trying to find a model that accommodates the more uniform magnitude in the background. This results in averaging the acceleration signal of the features and map background, amplifying the magnitude experienced near the surfaces around the features and lowering the magnitude of the features themselves. This can be seen explicitly by projecting the results of the network at R_0 in Figure 8. These maps show extremely coarse recognition of grouped features like the Andes and Himalayas in the SLFN and DNN models with approximately 1000 and 10000 free parameters, but do not accurately capture the highly discontinuous ridges of the tectonic plates.

It is technically possible that a more elaborate ANN configuration could accurately regress the gravitational features, but such success likely comes at the cost of additional parameters. If this is true, one must then decide if the ANN model is in anyway advantageous over more traditional approaches. ANN have the benefit of being quickly parallelized resulting in faster computation than the recursive formulations common to spherical harmonic models, however if exorbitant numbers of

nodes are needed to capture the equivalent accuracy of only a handful of spherical harmonic terms, it could be argued that such advantage is negligible.

CONCLUSION

This paper investigates the use of artificial neural networks to generate a more compact basis capable of regressing surface level gravitational perturbations. Despite different ANN configurations, training optimization strategies, and diverse data sets, the results suggest that such a model is not more efficient at expressing gravitational features. It is possible with further optimization effort, more specialized architectures, and additional feature engineering that a compact representation may be found – but until such ability has been demonstrated, it is encouraged to be wary of using such machine learning models for modeling any gravitational perturbations beyond those generated by the point mass contribution and J_2.

ACKNOWLEDGEMENTS

This material is based upon work supported by the National Science Foundation Graduate Research Fellowship under Grant No. 2040434.

REFERENCES

[1] S. Pines, "Uniform Representation of the Gravitational Potential and its Derivatives," *AIAA Journal*, Vol. 11, No. 11, 1973, pp. 1508–1511.

[2] J. B. Lundberg and B. E. Schutz, "Recursion formulas of Legendre functions for use with nonsingular geopotential models," *Journal of Guidance, Control, and Dynamics*, Vol. 11, No. 1, 1988, pp. 31–38, 10.2514/3.20266.

[3] B. Hofmann-Wellenhof and H. Moritz, *Physical geodesy*. Springer Science & Business Media, 2006.

[4] C. Kholshevnikov, "On Convergence of an Asymmetrical Body Potential Expansion in Spherical Harmonics," Vol. 16, 1977, pp. 45–60.

[5] C. Reigber, R. Schmidt, F. Flechtner, R. König, U. Meyer, K. H. Neumayer, P. Schwintzer, and S. Y. Zhu, "An Earth gravity field model complete to degree and order 150 from GRACE: EIGEN-GRACE02S," *Journal of Geodynamics*, Vol. 39, No. 1, 2005, pp. 1–10, 10.1016/j.jog.2004.07.001.

[6] B. D. Tapley, "Gravity model determination from the GRACE mission," *Journal of the Astronautical Sciences*, Vol. 56, No. 3, 2008, pp. 273–285, 10.1007/bf03256553.

[7] E. S. Davis, C. E. Dunn K, R. H. Stanton, and J. B. Thoman, "The Grace Mission: Meeting the Technical Challenges," *50th International Astronautical Congress*, 1999.

[8] F. G. Lemoine, S. Goossens, T. J. Sabaka, J. B. Nicholas, E. Mazarico, D. D. Rowlands, B. D. Loomis, D. S. Chinn, G. A. Neumann, D. E. Smith, and M. T. Zuber, "GRGM900C: A degree 900 lunar gravity model from GRAIL primary and extended mission data," *Geophysical Research Letters*, Vol. 41, No. 10, 2014, pp. 3382–3389, 10.1002/2014GL060027.

[9] S.-C. Han, C. Jekeli, and C. K. Shum, "Efficient gravity field recovery using in situ disturbing potential observables from CHAMP," *Geophysical Research Letters*, Vol. 29, No. 16, 2002, pp. 36–1–36–4, 10.1029/2002gl015180.

[10] N. K. Pavlis, S. A. Holmes, S. C. Kenyon, D. Schmidt, and R. Trimmer, "A preliminary gravitational model to degree 2160," *International Association of Geodesy Symposia*, Vol. 129, 2005, pp. 18–23, 10.1007/3-540-26932-0-4.

[11] S. K. Runcorn, "Flow in the Mantle Inferred from the Low Degree Harmonics of the Geopotential," *Geophysical Journal of the Royal Astronomical Society*, Vol. 14, No. 1-4, 1967, pp. 375–384, 10.1111/j.1365-246X.1967.tb06253.x.

[12] S. Manabe and T. B. Terpstra, "The effect of moutains on the general circulation of the Atmosphere.pdf," 1974.

[13] J. Kalvoda, J. Klokočník, J. Kostelecký, and A. Bezděk, "Mass distribution of earth landforms determined by aspects of the geopotential as computed from the global gravity field model EGM 2008," *Acta Universitatis Carolinae, Geographica*, Vol. 48, No. 2, 2013, pp. 17–25, 10.14712/23361980.2015.1.

[14] B. A. Jones, *Efficient Models for the Evaluation and Estimation of the Gravity Field*. PhD thesis, CU Boulder, 2010.

[15] M. Brillouin, "Équations aux dérivées partielles du 2e ordre. Domaines à connexion multiple. Fonctions sphériques non antipodes," Vol. 4, 1933, pp. 173–206.

[16] D. Gottlieb and C. W. Shu, "On the Gibbs phenomenon and its resolution," *SIAM Review*, Vol. 39, No. 4, 1997, pp. 644–668, 10.1137/S0036144596301390.

[17] E. Hewitt and R. E. Hewitt, "The Gibbs-Wilbraham phenomenon: An episode in fourier analysis," *Archive for History of Exact Sciences*, Vol. 21, No. 2, 1979, pp. 129–160, 10.1007/BF00330404.

[18] I. Goodfellow, Y. Bengio, and A. Courville, *Deep learning*. MIT press, 2016.

[19] T. Sanger, "Optimal unsupervised learning in a single-layer network," *Neural Networks*, Vol. 2, 1989, pp. 459–473.

[20] A. Gao and W. Liao, "Efficient gravity field modeling method for small bodies based on Gaussian process regression," *Acta Astronautica*, Vol. 157, No. December 2018, 2019, pp. 73–91, 10.1016/j.actaastro.2018.12.020.

[21] R. Furfaro, R. Linares, V. Reddy, J. Simo, and L. Le Corre, "Modelling irregular small bodies gravity field via extreme learning machines," *Advances in the Astronautical Sciences*, Vol. 160, 2017, pp. 471–486.

[22] L. Cheng, Z. Wang, Y. Song, and F. Jiang, "Real-time optimal control for irregular asteroid landings using deep neural networks," *Acta Astronautica*, Vol. 170, No. January 2019, 2020, pp. 66–79, 10.1016/j.actaastro.2019.11.039.

[23] K. Liano, "Robust error measure for supervised neural network learning with outliers," *IEEE Transactions on Neural Networks*, Vol. 7, No. 1, 1996, pp. 246–250, 10.1109/72.478411.

[24] G. B. Huang, H. Zhou, X. Ding, and R. Zhang, "Extreme learning machine for regression and multiclass classification," *IEEE Transactions on Systems, Man, and Cybernetics, Part B: Cybernetics*, Vol. 42, No. 2, 2012, pp. 513–529, 10.1109/TSMCB.2011.2168604.

[25] J. M. Martínez-Martínez, P. Escandell-Montero, E. Soria-Olivas, J. D. Martín-Guerrero, R. Magdalena-Benedito, and J. Gómez-Sanchis, "Regularized extreme learning machine for regression problems," *Neurocomputing*, Vol. 74, No. 17, 2011, pp. 3716–3721, 10.1016/j.neucom.2011.06.013.

[26] Y. Xu, J. Du, L. R. Dai, and C. H. Lee, "A regression approach to speech enhancement based on deep neural networks," *IEEE/ACM Transactions on Audio Speech and Language Processing*, Vol. 23, No. 1, 2015, pp. 7–19, 10.1109/TASLP.2014.2364452.

[27] J. Du and Y. Xu, "Hierarchical deep neural network for multivariate regression," *Pattern Recognition*, Vol. 63, No. April 2016, 2017, pp. 149–157, 10.1016/j.patcog.2016.10.003.

[28] T. Jayalakshmi and A. Santhakumaran, "Statistical Normalization and Back Propagationfor Classification," *International Journal of Computer Theory and Engineering*, Vol. 3, No. 1, 2011, pp. 89–93, 10.7763/ijcte.2011.v3.288.

[29] J. Bergstra and Y. Bengio, "Random search for hyper-parameter optimization," *Journal of Machine Learning Research*, Vol. 13, 2012, pp. 281–305.