
AI translation · View original & related papers at
chinaxiv.org/items/chinaxiv-202504.00070

A precise magnetic modeling method for scientific satellites based on a self-attention mechanism and Kolmogorov-Arnold Networks post-print

Authors: Ye Liu, Xingjian Shi, Wenzhe Yang, Zhiming Cai, Huawang Li, Huawang Li

Date: 2025-03-14T00:00:00+00:00

Abstract

As the complexity of scientific satellite missions increases, requirements for their magnetic fields, magnetic field fluctuations, and even magnetic field gradients and variations have become increasingly stringent. Additionally, there is a growing need to address the alternating magnetic fields produced by the spacecraft itself. This paper introduces a novel modeling method for spacecraft magnetic dipoles using an integrated self-attention mechanism and a transformer combined with Kolmogorov-Arnold Networks. The self-attention mechanism captures correlations among globally sparse data, establishing dependencies between sparse magnetometer readings. Concurrently, the Kolmogorov-Arnold Network, proficient in modeling implicit numerical relationships between data features, enhances the ability to learn subtle patterns. Comparative experiments validate the capability of the proposed method to precisely model magnetic dipoles, achieving maximum Root Mean Square Errors of $24.06 \text{ mA} \cdot \text{m}^2$ and 0.32 cm for size and location modeling, respectively. The spacecraft magnetic model established using this method accurately computes magnetic fields and alternating magnetic fields at designated surfaces or points. This approach facilitates the rapid and precise construction of individual and complete spacecraft magnetic models, enabling the verification of magnetic specifications from the spacecraft design phase.

Full Text

Preamble

Astronomical Techniques and Instruments, Vol. 2, January 2025, 1-9

Article

Open Access

A Precise Magnetic Modeling Method for Scientific Satellites Based on a Self-Attention Mechanism and Kolmogorov-Arnold NetworksYe Liu^{1, 2}, Xingjian Shi¹, Wenzhe Yang¹, Zhiming Cai¹, Huawang Li^{1*}¹Innovation Academy for Microsatellites of Chinese Academy of Sciences, Shanghai 201304, China²University of Chinese Academy of Sciences, Beijing 100049, China

*Correspondence: lihwa@microsat.com

Received: August 30, 2024; Accepted: September 19, 2024; Published Online: October 16, 2024

<https://doi.org/10.61977/ati2024049>; <https://cstr.cn/>© 2025 Editorial Office of Astronomical Techniques and Instruments, Yunnan Observatories, Chinese Academy of Sciences. This is an open access article under the CC BY 4.0 license (<http://creativecommons.org/licenses/by/4.0/>)**Citation:** Liu, Y., Shi, X. J., Yang, W. Z., et al. 2025. A precise magnetic modeling method for scientific satellites based on a self-attention mechanism and Kolmogorov-Arnold Networks. *Astronomical Techniques and Instruments*, 2(1): 1–9. <https://doi.org/10.61977/ati2024049>.

Abstract: As the complexity of scientific satellite missions increases, the requirements for their magnetic fields, magnetic field fluctuations, and even magnetic field gradients and variations become increasingly stringent. Additionally, there is a growing need to address the alternating magnetic fields produced by the spacecraft itself. This paper introduces a novel modeling method for spacecraft magnetic dipoles using an integrated self-attention mechanism and a transformer combined with Kolmogorov-Arnold Networks. The self-attention mechanism captures correlations among globally sparse data, establishing dependencies between sparse magnetometer readings. Concurrently, the Kolmogorov-Arnold Network, proficient in modeling implicit numerical relationships between data features, enhances the ability to learn subtle patterns. Comparative experiments validate the capability of the proposed method to precisely model magnetic dipoles, achieving maximum Root Mean Square Errors of $24.06 \text{ mA} \cdot \text{m}^2$ and 0.32 cm for size and location modeling, respectively. The spacecraft magnetic model established using this method accurately computes magnetic fields and alternating magnetic fields at designated surfaces or points. This approach facilitates the rapid and precise construction of individual and complete spacecraft magnetic models, enabling the verification of magnetic specifications from the spacecraft design phase.

Keywords: Magnetic dipole model; Self-attention mechanism; Kolmogorov-Arnold networks; Alternating current magnetic fields

1. INTRODUCTION

As the complexity of spacecraft missions increases, the requirements for the magnetic field, magnetic field fluctuations, magnetic field gradients, and magnetic gradient fluctuations of the spacecraft themselves become more stringent. For instance, the China Seismo-Electromagnetic Satellite (CSES), primarily tasked with spatial magnetic field measurements, mandates that the residual direct current (DC) magnetic field and stray magnetic field interferences, generated at the top of a 4.5 m boom, be less than 0.5 nT [1]. The Laser Interferometer Space Antenna (LISA), serving space gravitational wave detection missions, requires the magnetic field at the test mass to be less than 10 T, with a magnetic gradient less than 10 m, and magnetic field fluctuations and magnetic gradient fluctuations in the range of 0.1 mHz to 1 Hz to be less than 650 nT/Hz^{1/2} and 25 nT/Hz^{1/2}, respectively [2].

Similarly, satellites such as the Mars Global Surveyor (MGS) [3] and the Solar wind Magnetosphere Ionosphere Link Explorer (SMILE) satellite [4,5], which aims to capture panoramic images of the interactions between the solar wind and Earth's magnetosphere, also stipulate specific requirements for the magnetic fields generated at locations on the spacecraft where magnetometers are mounted.

Satellites with magnetic field specifications typically undergo rigorous magnetic field testing before launch. The prevalent method is in-situ magnetic measurement, which involves directly measuring the magnetic field produced by the spacecraft at the location of the onboard magnetometer. However, in-situ measurements are limited to validating magnetic field specifications and cannot adjust or compensate for the magnetic field of a spacecraft. This method is also unsuitable for satellites used in space-based gravitational wave detection, and similar spacecraft, where direct measurements in magnetically sensitive regions are not feasible. Consequently, an increasing number of satellites are now verifying the impact of spacecraft components on residual magnetism and striving to construct more accurate overall magnetic models. The MGS and CSES have conducted magnetic tests on their solar arrays, verifying that the stray magnetic fields generated at the positions of the magnetic detectors are less than 0.6 nT for MGS and less than 0.011 nT for CSES [6].

Researchers such as Carrubba [7] and Spantideas [8] have employed particle swarm algorithms for magnetic dipole modeling of individual spacecraft components, achieving notable results. Jin et al. [9] have characterized the magnetic source properties of individual spacecraft units within a magnetically shielded room, but due to limitations in the magnetic measurement workflow, it was impossible to model the spacecraft's alternating magnetic sources. This lim-

itation precludes validating related specifications like alternating current (AC) magnetic fields in magnetically sensitive areas of the spacecraft during the design phase. Preliminary attempts at modeling spacecraft magnetic sources using neural networks by Spantideas [10] and others did not address the modeling and validation of alternating magnetic sources.

Given the increasingly demanding nature of scientific satellite missions for astronomical and space observations [11], which require high standards for both DC and AC magnetic fields generated by the spacecraft, this paper proposes a rapid and accurate magnetic modeling method, referred to as a Transformer with Kolmogorov-Arnold Networks (TKAN), for scientific satellites that integrates self-attention mechanisms with Kolmogorov-Arnold Networks (KANs). Using the ability of self-attention mechanisms to capture global data interdependencies and the proficiency of the KAN in handling symbolic formulae, this method enhances the network's capability to learn subtle patterns from data features. It achieves rapid and precise construction of individual spacecraft magnetic models, ultimately facilitating the creation of an accurate overall magnetic model for the spacecraft.

2.1. Data Source

This study employs Multiple Magnetometer Facilities (MMFs) [12,13], as shown in Fig. 1 [Figure 1: see original paper]. This figure illustrates the placement of the MMF in a zero-magnetic space to avoid interference from the Earth's magnetic field on the magnetic field measurements of the equipment under test. Composed of several magnetometers, the MMF measures the magnetic field at the center of the spacecraft equipment. The magnified area in the image shows the process of measuring the equipment's magnetic field using the MMF. An MMF is typically positioned within a zero-magnetic space to eliminate interference from the Earth's geomagnetic field, employing P magnetic sensors to capture the magnetic fields generated by the equipment under test, where P represents the p-th measurement point (where $p = 1, 2, \dots, P$). The measurements are denoted by B .

Due to the extensive data requirements for neural network training, this paper also employs a formula for calculating magnetic fields using magnetic moments to generate training data. Considering that a single machine can be approximated as a magnetic dipole, the parameters of an equivalent magnetic dipole represent the magnetic characteristics of that single machine, serving as its actual magnetic moment. The magnetic field can be calculated using the relationship:

$$B'_p = \frac{\mu_0}{4\pi} \left[\frac{3(m \cdot r)r}{r^5} - \frac{m}{r^3} \right] \quad (1)$$

where B'_p is the magnetic field vector caused by the magnetic dipole at the p -th measurement point, r is the relative 3D position between the p -th measurement point and the magnetic dipole, m is the magnetic moment of the dipole, and μ_0 is the magnetic permeability of free space, with a value of $4\pi \times 10^{-7}$ H/m.

2.2. TKAN Model

2.2.1. Overall Structure

The structure of the TKAN model is shown in Fig. 2 [Figure 2: see original paper], and mainly consists of two components: the TKAN encoder and the KAN. The TKAN encoder incorporates the KAN [14] in place of the Multilayer Perceptron (MLP) layer [15] found in traditional transformers [16], enhancing the efficiency and interpretability of the conventional transformer encoding layers. The TKAN combines the self-attention mechanism of the transformer with the KAN. This integration not only preserves the ability of the transformer to establish long-distance semantic relationships, but also addresses the issue of poor interpretability associated with the transformer.

Although there has been extensive research attempting to decode the intrinsic mechanisms of the transformer through visualization [17,18], attention maps [19], transformer interactions among patches [20], or self-attention attribution methods [21], most studies focus only on understanding specific target problems rather than providing a comprehensive explanation from the perspective of the transformer, such as the complexity of self-attention mechanisms and the vast number of parameters. Moreover, the flexibility of spline basis functions, known as learnable activation functions, in the KAN allows for adaptive modeling of complex relationships within the data by adjusting their shapes, thereby minimizing approximation errors and enhancing the capacity of the network to learn subtle patterns from high-dimensional datasets. In traditional transformers and other classical neural networks (such as Visual Geometry Group [22] or ResNet [23]), an MLP serves as the output layer, integrating outputs from encoding layers and/or hidden layers. To enhance the network's nonlinear expression capabilities and interpretability, this study employs a KAN instead of the MLP layer for the output layer. At the edge of the network, the KAN uses learnable activation functions instead of the fixed activation functions used by an MLP, parameterizing the activation functions in the form of spline functions. This approach offers high flexibility and allows for the modeling of complex functions with fewer parameters, facilitating the resolution of magnetic dipole moment modeling issues.

2.2.2. TKAN Encoding Layer

As shown in Fig. 2, the TKAN encoding layer comprises a multi-head attention (MHA) mechanism and a feed-forward neural network. The MHA mechanism in the transformer encoding layer is pivotal for establishing long-distance

dependencies between data, which is crucial for modeling magnetic dipole moments. In the multiple magnetometer facility, different measurement points can capture the magnetic fields produced by the equipment under test at various locations, and the magnetic fields measured at different points exhibit varying degrees of correlation. The MHA mechanism effectively captures the relationships and dependencies between magnetic fields across different measurement points and channels. Additionally, the classic transformer uses a feed-forward neural network to integrate the outputs from the MHA mechanism, enhancing the nonlinear expressive capacity of the model.

(1) MHA Mechanism

The MHA mechanism consists of several sub-attention modules. Each self-attention module learns the interactions between input vectors to calculate distinct weights for different entities, thereby identifying more significant areas. The self-attention mechanism is formulated as:

$$\text{Self-attention}(Q; K; V) = \text{softmax}\left(\frac{QK^T}{\sqrt{d_k}}\right)V \quad (2)$$

where Q is the query information, K is the key information (used for matching Q), V is the value vector, and d_k is the dimensionality of K .

Self-attention is termed as such because the matrices Q , K , and V are all derived from a sequence of input vectors through linear mappings, expressed as:

$$\begin{cases} Q = XW_Q \\ K = XW_K \\ V = XW_V \end{cases} \quad (3)$$

where W_Q , W_K , and W_V represent trainable weight matrices and X is the input of the network. Since each self-attention mechanism has distinct weight matrices, the output self-attention matrices are also different. The MHA mechanism contains several different self-attention modules to create multiple subspaces, each focusing on different information. Ultimately, the outputs from each sub-attention module are merged and integrated using a fully connected layer to consolidate the outputs from each self-attention mechanism.

(2) Fully Connected Feed-Forward Neural Networks

MLP layers, also known as fully connected feed-forward neural networks, are fundamental components in deep learning architectures, and are typically employed to approximate nonlinear functions within machine learning frameworks. An MLP consisting of L layers operates through a series of transformation matrices W and activation functions σ , mathematically expressed as:

$$\text{MLP}(X) = (W_{L-1} \circ \sigma \circ W_{L-2} \circ \sigma \circ \dots \circ W_1 \circ \sigma \circ W_0)X \quad (4)$$

MLPs strive to emulate complex functional mappings via a sequence of non-linear transformations across multiple layers. Despite their ubiquity, MLPs encounter notable limitations. This is particularly the case for predicting magnetic dipoles, in which measurement points are sparse. The disparity between sparse magnetic fields and the limited expressive capability of MLPs exacerbates the challenging nature of prediction. Drawbacks include high computational complexity, substantial memory demand, and limited interpretability when handling high-dimensional data, potentially leading to inadequate feature learning across varying magnetic fields.

2.2.3. KANs

Recently, Liu et al. [14] introduced KANs as a promising alternative to traditional MLPs. While MLPs draw inspiration from the generalized approximation theorem, a KAN is grounded in the Kolmogorov-Arnold representation theorem. KANs maintain a fully connected structure akin to MLPs; however, unlike MLPs that deploy fixed activation functions on nodes (neurons), KANs implement learnable activation functions on edges (weights). Analogous to an MLP, an L -layer KAN is defined as a nesting of multiple KAN layers, as:

$$\text{KAN}(X) = (\Phi_{L-1} \circ \Phi_{L-2} \circ \dots \circ \Phi_1 \circ \Phi_0)X \quad (5)$$

where Φ_i is the i -th layer in the KAN. In each KAN layer with n_{in} -dimensional input and n_{out} -dimensional output, Φ encompasses $n_{in} \times n_{out}$ learnable activation functions $\phi_{q,p}$, with the relationship:

$$\phi_{q,p}(x) = w_b b(x) + w_s \text{spline}(x) \quad (7)$$

The learnable activation function is the sum of the basis function and the spline function, where, in most cases, $\text{spline}(x)$ is parameterized as a linear combination of B-splines [14] such that:

$$\text{spline}(x) = \sum_i c_i B_i(x) \quad (8)$$

where w_b , w_s , and c_i are the trainable parameters.

The computation from layer k to layer $k+1$ in the KAN can be represented in matrix form as:

$$X_{k+1} = \Phi_k X_k \quad (9)$$

Expanding upon this, we propose the TKAN encoder architecture, which substitutes the MLP layers in the transformer encoder with KAN layers from the original transformer model to address the issues noted. The TKAN encoder consists of improved MHA layers and KANs. Layer normalization precedes each block, and residual concatenation follows each block. Differing from the traditional transformer, the magnetic fields detected by sensors are directly input into the TKAN encoder. Composed of serially stacked TKAN encoder blocks, the features are subsequently fed into the KAN for prediction. This design effectively mitigates the limitations of traditional MLPs and boosts the performance and interpretability of magnetic dipole models.

3.1. Experiment Setup

Given the extensive data requirements for training neural networks, this study uses modeling and validation strategies for data generation. Specifically, the experiment initially uses an MMF equipped with 18 magnetometers (as shown in Fig. 1) to collect ten sets of real magnetic field data created by the magnets. The positions of the magnetometers and the measurement coordinate system within the facility are modeled. Subsequently, the simulation model uses the relationship shown in Equation (1) to calculate deviations between the magnetic fields produced by the modeled facility at each magnetometer, and the actual magnetic fields. When the deviation in each dimension at every magnetometer is less than 1 nT, the simulation model of the multiple magnetometer facility is validated as capable of replicating the magnetic fields generated by actual magnetic moments.

The validated multiple magnetometer facility model is then used to generate a dataset for this study, comprising 10,000 data points. Each data point includes magnetic field values across 54 dimensions measured by 18 magnetometers, along with the corresponding 6-dimensional parameters of the magnetic dipole. The dataset is divided into a training set and a test set with a ratio of 9:1. The experiment employs K-fold cross-validation on the training set to assess the model, dividing the training data into 9 parts. Each training iteration selects one part as the validation set, using the remaining parts as the training set, and the final model's accuracy is determined by the average of results from K iterations (with K set to 9 in this experiment). The test set serves as an independent dataset for the final evaluation of the model's performance. The experiment uses mean absolute error as the loss function, with the relationship:

$$\text{loss} = \frac{1}{n} \sum_{i=1}^n |y_i - \hat{y}_i| \quad (10)$$

where n is the number of data, y_i represents the true value, and \hat{y}_i denotes the model prediction. For optimization, the study uses the Adam optimizer, with

epochs set to 5,000, an initial learning rate of 10^{-5} , and a momentum of 0.9.

3.3. Experimental Results

In the process of modeling, several metrics are employed to reflect the discrepancies between model predictions and actual values, and to assess the stability of the model fit. These metrics include Root Mean Square Error (RMSE), Mean Absolute Error (MAE), and Coefficient of Determination (R^2). Additionally, the standard deviation (SD) is used to assess the stability of the model fit. These metrics together provide a comprehensive evaluation of model performance, and are defined as:

$$\text{RMSE} = \sqrt{\frac{1}{p} \sum_{i=1}^p (y_i - \hat{y}_i)^2} \quad (11) \quad (1)$$

$$\text{MAE} = \frac{1}{p} \sum_{i=1}^p |y_i - \hat{y}_i| \quad (12) \quad (2)$$

$$R^2 = 1 - \frac{\sum_{i=1}^p (y_i - \hat{y}_i)^2}{\sum_{i=1}^p (y_i - \bar{y}_i)^2} \quad (13) \quad (3)$$

$$\text{SD} = \sqrt{\frac{1}{p} \sum_{i=1}^p (y_i - \bar{y}_i)^2} \quad (14) \quad (4)$$

where y_i represents the actual value, \hat{y}_i denotes the predicted value, and \bar{y}_i is the mean of the actual values. A model is considered to perform better as RMSE and MAE approach 0 and R^2 approaches 1, indicating a good fit to the data. A smaller standard deviation indicates greater stability in model performance.

3.3.1. Comparative Experiments

To verify the advantages of the TKAN model in magnetic dipole modeling, this study compares it with MLP, KAN, and transformer models. In the experiments conducted for this study, the network architectures for MLP and KAN are well-defined and consist of an input layer, five hidden layers, and an output layer. The number of nodes in each hidden layer is correlated with N_{in} , the number of nodes in the input layer. The numbers of nodes in the hidden layers are $2N_{in}$, $4N_{in}$, $8N_{in}$, $4N_{in}$, and $2N_{in}$. The network structure of the transformer is shown in Fig. 2.

Quantitative metrics and comparative results obtained on the test dataset for different networks are displayed in Table 1. This table presents the results for modeling the size, position, and modulus of magnetic dipoles using MLP, KAN, transformer, and TKAN. Results are quantified by RMSE, MAE, R^2 ,

and SD. The best performance of each metric for each column is highlighted in bold. M represents the magnitude and direction of the magnetic dipole, and P denotes the position of the magnetic dipole. Both are three-dimensional vectors. $|M|$ and $|P|$ respectively represent their magnitudes. As the most widely used network model, the fully connected layers of the MLP have been adopted and utilized by many networks.

In modeling the size of magnetic dipoles, the MLP shows poorer performance than the KAN across all four metrics. However, in modeling the position of magnetic dipoles, the reverse is true. Notably, the KAN achieves the best results in modeling the magnitude of the magnetic dipole size, demonstrating its ability to accurately capture the relationship between sparse input measurements and the target magnetic dipole size.

The TKAN proposed in this paper outperforms the MLP, KAN, and transformer in both size and position modeling of magnetic dipoles, with an average RMSE reduction of 20.31, 5.01, and 3.08, average MAE reduction of 19.19, 3.34, and 2.50, average R^2 improvement of 0.081, 0.15, and 0.0060, and average SD reduction of 30.66, 24.91, and 3.30.

The larger improvements in RMSE and MAE compared with R^2 can be attributed to the nature of these metrics. RMSE and MAE are absolute error metrics that provide direct measures of average model error magnitudes, which are more sensitive to reductions in error outliers and variability in predictions. The exceptional reduction in these metrics indicates that the TKAN is particularly effective at handling outliers and fluctuations in magnetic dipole modeling. This capability stems from the integration of the KAN module, enhancing the ability of the network to model fine details and manage sparse measurement data more effectively through improved data encoding by the self-attention.

In addition, the improvement in the R^2 metric demonstrates that the TKAN not only excels in enhancing prediction accuracy but also significantly boosts its explanatory power over the total variability in the data. With an R^2 value approaching 1, TKAN outperforms the MLP, KAN, and transformer, showcasing its superior ability to account for variation in magnetic dipole modeling.

The exceptional performance of TKAN across these four metrics not only demonstrates that it inherits the advantages of the transformer by globally modeling sparse measurement data using the self-attention mechanism, but also shows that the integration of the KAN module allows for more detailed encoding of data by the self-attention mechanism, enhancing the network's performance in modeling magnetic dipoles.

3.3.2. Multiple Magnetic Dipole Modeling for Spacecraft

In this study, 25 magnetic dipoles are randomly generated as different subsystems or devices and various networks are then used to estimate these magnetic dipoles to construct a complete spacecraft magnetic model. The magnetic field

and its errors at the plane of the onboard magnetometers are calculated, with results shown in Fig. 3 [Figure 3: see original paper]. The 25 simulated magnetic sources in a simulated spacecraft are shown in Fig. 4 [Figure 4: see original paper].

Consistent with the results in Table 1, the spacecraft magnetic models constructed by MLP and KAN show significant deviations from actual magnetic fields at the target plane. However, for transformer and TKAN, the magnetic fields generated at the target plane are very close to the actual fields, with TKAN showing notably smaller errors than the transformer, indicating that TKAN provides a more accurate model of the spacecraft' s magnetic field.

3.3.3. AC Magnetic Field Modeling for Spacecraft

Traditional methods of magnetic dipole modeling fail to meet the requirements for rapid modeling, making neural networks one of the most effective solutions for modeling alternating magnetic sources. This study simulates the internal layout and magnetic source characteristics of a satellite based on the actual structure of the “Taiji” space-based gravitational wave detector (as shown in Fig. 5 [Figure 5: see original paper]), including battery arrays, gas cylinders, and moving optical subassemblies. The computed AC magnetic field power spectra density (ASD) at the test mass location, established by different networks, are illustrated in Fig. 6 [Figure 6: see original paper], and are consistent with the corresponding performance metrics presented in Table 2 .

Although MLP performs more poorly than the transformer in establishing magnetic dipoles, it achieves better MAE scores and is the fastest, taking only 9.94 s to establish the AC magnetic model of the complete spacecraft. The TKAN proposed in this paper achieves excellent results in both RMSE and MAE, with RMSE scores second only to the transformer, and the best MAE scores, demonstrating TKAN' s capability to more accurately construct the complete alternating magnetic model of the spacecraft and support calculations of alternating magnetic fields at target points, thereby verifying whether the spacecraft' s specifications meet design requirements. Furthermore, although TKAN takes the longest time to construct the complete model, the time taken is only 95.76 s, which is within an acceptable range.

3.4. Ablation Experiments

To analyze the contributions of different modules to the TKAN method, this study designs three sets of ablation experiments: Replacing the KAN layer in the TKAN encoder with an MLP (TKAN-MHA), and substituting the KAN layers after the TKAN encoder layer with MLP layers (TKAN-MLP), alongside a standard transformer configuration. The performance of each method, relative to the full TKAN approach, is assessed using metrics such as RMSE, MAE, R^2 , and SD, as detailed in Table 3 .

The data presented in Table 3 reveal that replacing the KAN layer in the TKAN encoder with the original MLP results in a slight decline in the network's ability to model the magnitude of magnetic dipoles. This suggests that the self-attention mechanism can precisely capture and model the relationships among sparse measurements, while the inclusion of the KAN enhances the encoding capabilities of the MHA mechanism. Similarly, substituting the TKAN's KAN layers with MLP layers leads to a decrease in network performance, which is marginally worse than that observed with TKAN-MHA. This indicates that the KAN is better suited as the output layer of TKAN, improving the network's ability to integrate outputs from the encoding layers.

4. CONCLUSION

As the magnetic field requirements for scientific satellites become increasingly stringent, there is a heightened demand to fully understand not only precise DC magnetic fields but also AC magnetic fields generated by the spacecraft itself. This paper introduces a novel modeling approach that integrates a self-attention mechanism with a KAN for magnetic dipole modeling. The self-attention mechanism captures correlations among globally sparse data, establishing dependencies between measurements of magnetometers. Concurrently, the KAN is proficient in handling symbolic formulas to process the implicit numerical relationships among data features, thereby enhancing the ability of the network to discern subtle patterns within the data features.

Experimental results validate the capability of TKAN in modeling magnetic dipoles, with dipole magnitudes having maximum RMSE and MAE of $24.06 \text{ mA} \cdot \text{m}^2$ and $15.78 \text{ mA} \cdot \text{m}^2$, respectively. The maximum RMSE and MAE for dipole position are 0.32 cm and 0.18 cm, respectively. The experiments also demonstrate that the magnetic model of the entire spacecraft, established using the TKAN, accurately calculates the magnetic and AC magnetic fields at specific target surfaces or points. This achievement fulfills the objective of swiftly and precisely constructing both individual component and whole spacecraft magnetic models, thereby facilitating the verification of spacecraft magnetic standards from the design phase of spacecraft development.

ACKNOWLEDGEMENTS

This work was supported by the National Research Development Program of China (2020YFC2200901).

AUTHOR CONTRIBUTIONS

Ye Liu wrote most of the manuscript, conceived the idea of the algorithm, and conducted the experiments. Xingjian Shi conceived the idea of the project

and revised the paper. Wenzhe Yang conducted the practical tests and provided experimental data. Zhiming Cai revised the paper. Huawang Li initiated the project and revised the paper. All authors read and approved the final manuscript.

DECLARATION OF INTERESTS

Zhiming Cai is the editorial board member for *Astronomical Techniques and Instruments* and was not involved in the editorial review or the decision to publish this article. The authors declare no competing interests.

REFERENCES

1. Meng, L. F., Chen, J. G., Xiao, Q., et al. 2024. A method of magnetic modelling and simulation for high magnetic cleanliness satellite. *Spacecraft Environment Engineering*, 41(3): 296–300. (in Chinese)
2. Amaro-Seoane, P., Audley, H., Babak, S., et al. 2017. Laser interferometer space antenna. arXiv:1702.00786.
3. Zhou, L., Zheng, Y. J., Ouyang, M. G., et al. 2017. A study on parameter variation effects on battery packs for electric vehicles. *Journal of Power Sources*, 364: 242–252.
4. Branduardi-Raymont, G., Wang, C., Escoubet, C. P. et al. 2018. SMILE definition study report. European Space Agency.
5. Dai, L., Zhu, M. H., Ren, Y., et al. 2024. Global-scale magnetosphere convection driven by dayside magnetic reconnection. *Nature Communications*, 15: 639.
6. Zhang, Y., Zhao, Y., Liu, Y. M. 2020. Magnetic cleanliness design and test verification of solar array. *Chinese Journal of Power Sources*, 44(2): 223–226. (in Chinese)
7. Carrubba, E., Junge, A., Marliani, F., et al. 2014. Particle swarm optimization for multiple dipole modeling of space equipment. *IEEE Transactions on Magnetics*, 50(12): 1–10.
8. Spantideas, S. T., Kapsalis, N. C., Kakarakis, S. D. J., et al. 2014. A method of predicting composite magnetic sources employing particle swarm optimization. *Progress In Electromagnetics Research M*, 39: 161–170.
9. Jin, Y. X., Yao, C. Z., Li, L. Y., et al. 2024. Magnetic characterization of spacecraft equipment in a magnetic shielded room. *IEEE Transactions on Instrumentation and Measurement*, 73: 1–9.
10. Spantideas, S. T., Giannopoulos, A. E., Kapsalis, N. C., et al. 2021. A deep learning method for modeling the magnetic signature of spacecraft

- equipment using multiple magnetic dipoles. *IEEE Magnetism Letters*, 12: 1–5.
11. Dai, L., Han, Y. M., Wang, C., et al. 2023. Geoeffectiveness of interplanetary Alfvén waves. I. Magnetopause magnetic reconnection substorms. *The Astrophysical Journal*, 945(1): 47.
 12. Dai, L., Wang, C., Cai, Z. M., et al. 2020. AME: a cross-scale constellation of CubeSats to explore magnetic reconnection in the solar-terrestrial relation. *Frontiers in Physics*, 8: 89.
 13. Tsatalas, S., Vergos, D., Spantideas, S. T., et al. 2019. A multi-magnetometer driven facility for on-ground novel characterization of spacecraft equipment. *Measurement*, 146: 948–960.
 14. Liu, Z. M., Wang, Y. X., Vaidya, S., et al. 2024. Kan: Kolmogorov-arnold networks. arXiv:2404.19756.
 15. LeCun, Y., Bengio, Y., Hinton, G. 2015. Deep learning. *Nature*, 521(7553): 436–444.
 16. Vaswani, A., Shazeer, N., Parmar, N., et al. 2017. Attention is all you need. arXiv:1706.03762.
 17. Li, Y., Wang, J. P., Dai, X., et al. 2023. How does attention work in vision transformers? A visual analytics attempt. *IEEE Transactions on Visualization and Computer Graphics*, 29(6): 2888–2900.
 18. Ghiasi, A., Kazemi, H., Borgnia, E., et al. 2022. What do vision transformers learn? a visual exploration. arXiv:2206.10838.
 19. Chefer, H. 2021. Transformer interpretability beyond attention visualization. *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*.
 20. Ma, J., Bai, Y. L., Zhong, B. N., et al. 2023. Visualizing and understanding patch interactions in vision transformer. *IEEE Transactions on Neural Networks and Learning Systems*, 35(10): 13671–13680.
 21. Hao, Y. R., Dong, L., Wei, F. R., et al. 2021. Self-attention attribution: Interpreting information interactions inside transformer. In *Proceedings of the AAAI Conference on Artificial Intelligence*.
 22. Simonyan, K., Zisserman, A. 2014. Very deep convolutional networks for large-scale image recognition. arXiv:1409.1556.
 23. He, K. M., Zhang, X. Y., Ren, S. Q., et al. 2016. Deep residual learning for image recognition. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*.

Note: Figure translations are in progress. See original paper for figures.

Source: ChinaXiv – Machine translation. Verify with original.