TUBAF
Die Ressourcenuniversität.
Seit 1765.

Performance of transformer-based machine learning methods for UAV time series prediction

H. Steinhoff[1], P.O. Kotowski[2], S. Weit[1], T. Günther[1]

Motivation

Problem: The acquisition of data in drone-based (UAV) geophysical electromagnetic (EM) surveys can be compromised by issues such as memory errors or equipment failure, leading to incomplete datasets.

Objective: The master's thesis aims to evaluate the effectiveness of transformer-based machine learning architectures, leveraging their attention mechanism, for full times series prediction. A trained network can be used to replace faulty time series and render the datasets viable for subsequent analysis and interpretation.

Training Data

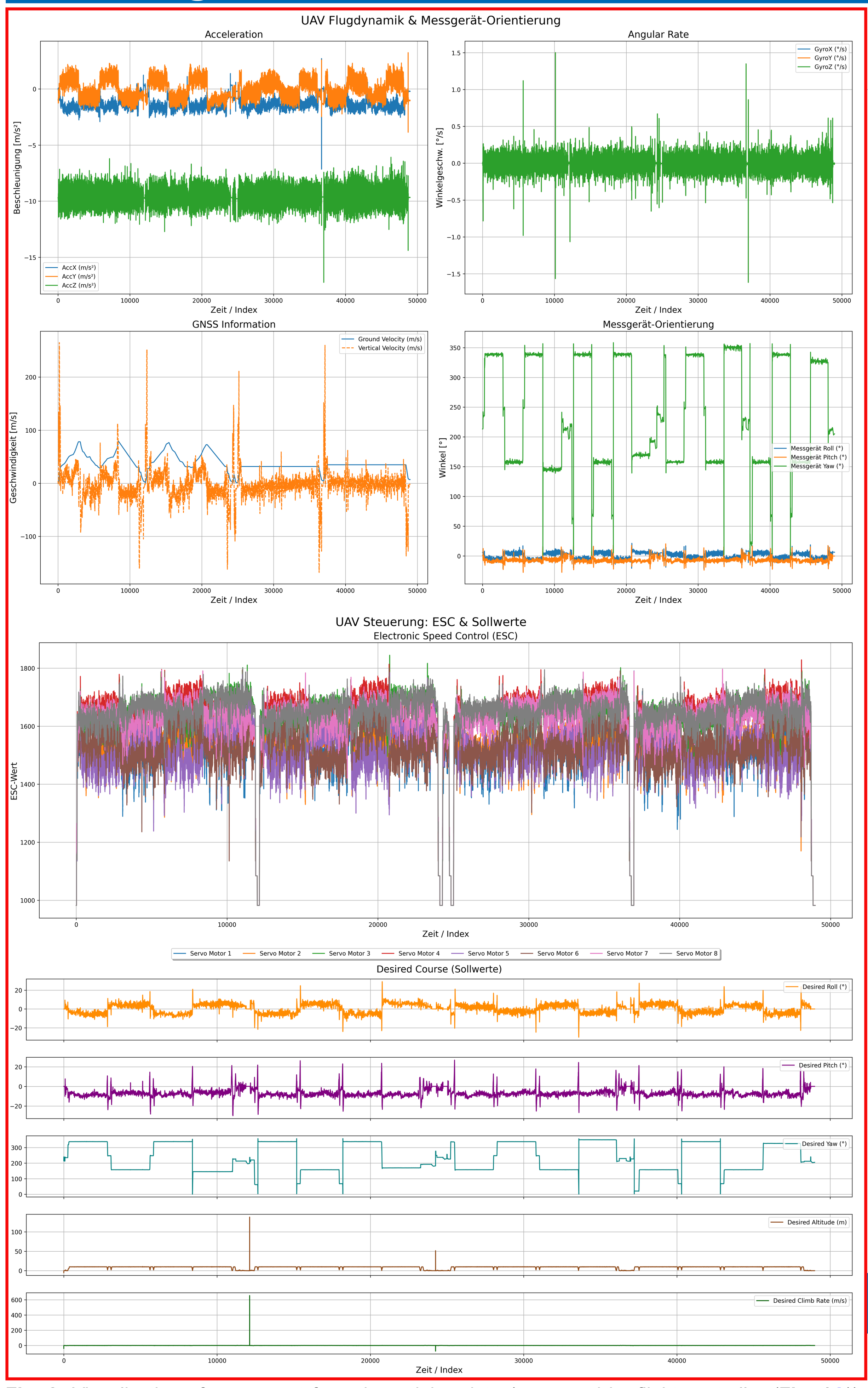

Fig. 2: Visualization of a segment from the training data (measured by flight controller (**Fig. 1A**)). The top section displays the UAV's flight dynamics and orientation. The bottom section shows the corresponding control signals.

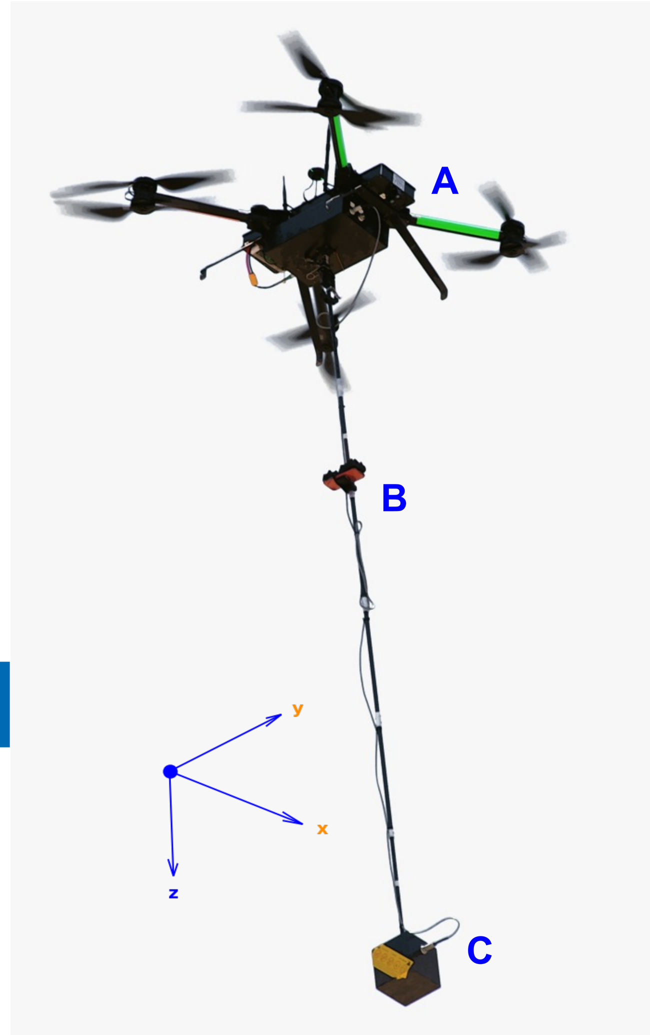

Fig. 1: UAV (**A**) with inertial sensor (**B**) and magnetometer (**C**)

Transformer Overview

The Transformer architecture is an Encoder-Decoder model that processes entire data sequences at once.

Input Processing
- Input features are converted into vectors (Embeddings).
- Positional Encodings are added to give the model information about the temporal order.

Encoder (Left Stack)
- Processes the entire historical input sequence.
- Uses Multi-Head Self-Attention to identify how influential each time step is to all other time steps in the input.

Decoder (Right Stack)
- Generates the future forecast step-by-step.
- Uses Masked Self-Attention to prevent it from seeing future forecasted values.
- Uses Encoder-Decoder Attention to focus on the most relevant parts of the input history when making a prediction.

Final Output
- A Linear Layer maps the decoder's output to a final numerical forecast.

Transformer Prediction

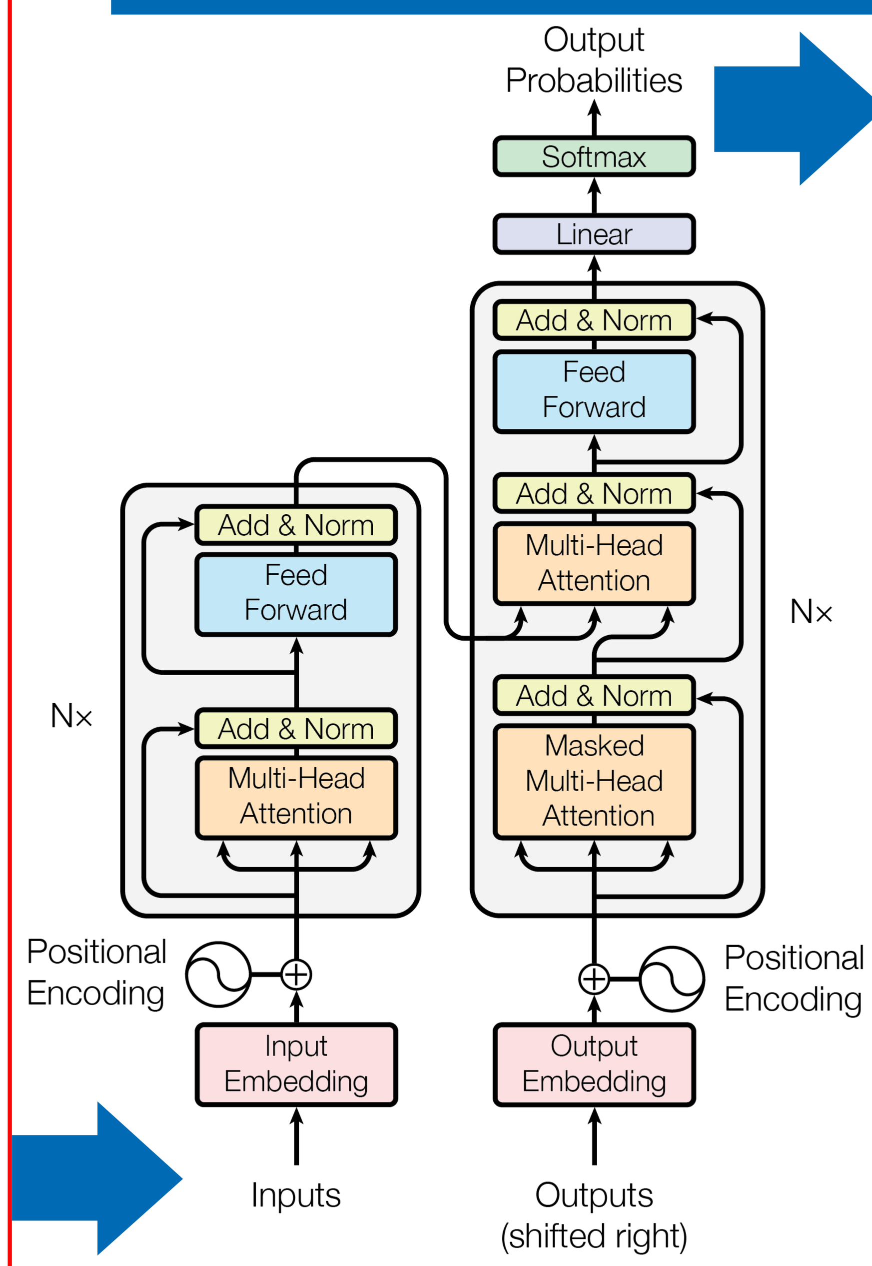

Fig. 3: The Transformer architecture (Vaswani et al., 2017), featuring an encoder stack (left) and a decoder stack (right).

Fig. 4: Performance of the Transformer model on a test data segment. The plots compare the model's predictions (red, dashed line) against the true values (blue, solid line) for the Roll, Pitch, and Yaw orientation angles (measured by **Fig. 1B**).

Residuals

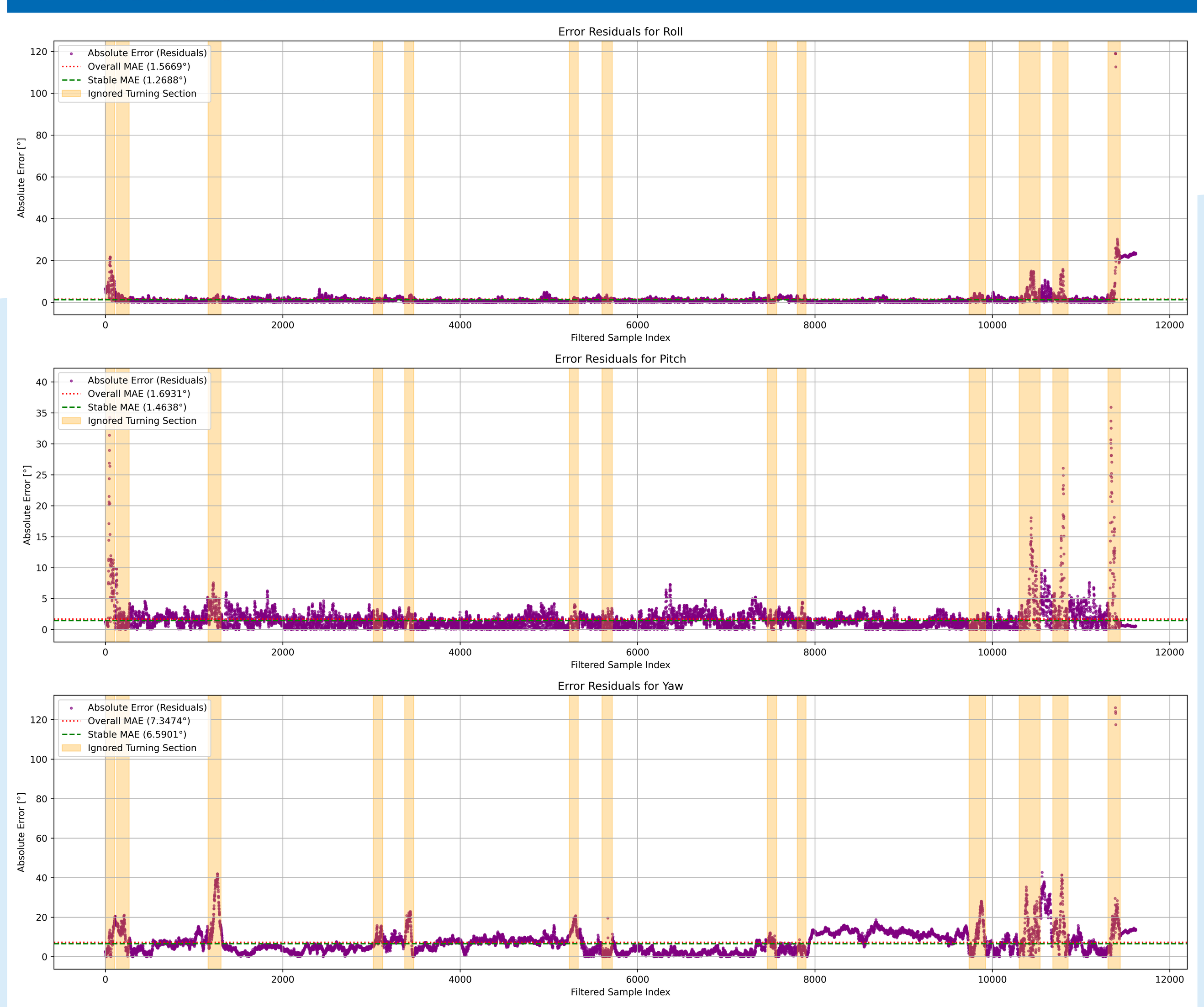

Fig. 6: Distribution of the model's absolute prediction errors for Roll, Pitch, and Yaw.

Conclusion

Conclusion: The Transformer architecture proves to be well-suited for predicting angle-based time series. Compared to the FFN, a significant improvement is observable, especially for the roll and pitch angles. For the yaw angle, however, predictions from both architectures exhibit significant deviations, a finding confirmed by the residual plots. These plots also reveal high absolute errors at the beginning and end of the time series, which can be attributed to the lack of preceding values that the Transformer requires for a context-aware prediction.

Outlook: Future work must focus on improving prediction accuracy, particularly for the yaw angle. To achieve this, further optimizations in both the preprocessing and training phases are planned. Furthermore, exploring alternative architectures (e.g., Mamba) or hybrid models (e.g., Mamba-Transformer) presents a promising option.

FFN Prediction

Fig. 5: Prediction results for the FFN model, showing its performance on the same test segment as the Transformer.

References

[1] Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A. N., Kaiser, Ł., & Polosukhin, I. (2017). Attention is all you need. arXiv. https://doi.org/10.48550/arXiv.1706.03762
[2] Kotowski, P. O., Becken, M., 2024. Forecasting Drone-towed Sensor Platform Orientation with Neural Networks. 84. Jahrestagung der Deutschen Geophysikalischen Gesellschaft (DGG), Jena.

[1] Technische Universität Bergakademie Freiberg | Institut für Geophysik und Geoinformatik | Freiberg, Deutschland
[2] LIAG-Institut für Angewandte Geophysik | Hannover, Deutschland