In: 2023 AIAA Aviation and Aeronautics Forum and Exposition (AIAA AVIATION Forum),
San Diego, 12-16 June 2023.
DOI:10.2514/6.2023-3439. Paper number AIAA 2023-3439

Data-driven Synthetic Air Data Estimation System Development for a Fighter Aircraft

Hasan Karali*, Mevlut Uzun †, Burak Yuksek‡, Gokhan Inalhan§
Centre for Cyperphysical and Autonomous Systems,
Cranfield University, Bedford, United Kingdom, MK43AL

In this paper, we propose an AI-based methodology for estimating angle-of-attack and angle-of-sideslip without the need for traditional vanes and pitot-static systems. Our approach involves developing a custom neural-network model to represent the input-output relationship between air data and measurements from various sensors such as inertial measurement units. To generate the training data required for the neural network, we use a 6-degrees-of-freedom F-16 simulator, which is further modified to simulate more realistic flight data. The training data covers the full flight envelope, allowing the neural network to generate accurate predictions in all feasible flight conditions. Our methodology achieves high-accuracy estimations of angle-of-attack and angle-of-sideslip, with mean absolute errors of 0.534 deg and 0.247 deg, respectively, during the test phase. The results demonstrate the potential of the proposed methodology to accurately estimate important flight parameters without the need for complex and costly instrumentation systems. The proposed methodology could have significant practical applications in the aviation industry, particularly in next-generation aircraft instrumentation and control. Future research could focus on further refining the neural-network model and exploring its application in other aircraft systems to improve safety and reduce costs.

I. Introduction

The performance and safety of an aircraft are directly related to total airspeed and relative orientation of the airframe according to the surrounding airflow. The pilot or real-time operator should monitor these physical elements to assess the flight performance during an operation. These measurements are critical for flight safety as stall warning systems utilize this feedback to warn the pilot or operators, and flight control algorithms could also use them to improve flying characteristics. In conventional systems, air data booms are mounted on the aircraft's nose section to measure dynamic pressure, angle-of-attack (α), and angle-of-sideslip (β). However, these systems are unsuitable, especially for new-generation fighter aircraft, because of critical requirements such as stealthiness and aerodynamic efficiency. From a stealthiness point of view, designers aim to reduce the aircraft's radar cross-section (RCS), but adding an air data boom could increase it and affect the detectability in a warfare environment. Aerodynamic efficiency is another critical issue; adding an air-data boom would not contribute positively.

To address the limitations of conventional air data booms, there has been growing interest in developing alternative methods for measuring and estimating airspeed and orientation parameters in aircraft. One promising approach is using AI-based techniques to estimate angle-of-attack and angle-of-sideslip without the need for traditional instrumentation systems. Recent advances in machine learning have made it possible to develop customized neural network models that can accurately estimate angle-of-attack and angle-of-sideslip using data from various sensors, such as inertial measurement units. Unlike traditional instrumentation systems, these AI-based approaches do not require air data booms, making them more suitable for new-generation fighter aircraft with stringent stealthiness and aerodynamic efficiency requirements. Furthermore, the use of AI-based techniques can potentially improve the accuracy and reliability of airspeed and orientation parameter estimation, thus contributing to increased safety and performance in aircraft operations. However, developing accurate and reliable AI-based models for airspeed and orientation parameter estimation is a challenging task that requires extensive data collection, processing, and analysis. The quality and quantity of the training data, the selection of appropriate features, and the choice of neural network architecture are all critical factors affecting the model's performance. This paper presents an AI-based methodology for estimating angle-of-attack

*PhD Student, School of Aerospace, Transport and Manufacturing, hasan.karali@cranfield.ac.uk, AIAA Member
†Postdoctoral Research Fellow, School of Aerospace, Transport and Manufacturing, mevlut.uzun@cranfield.ac.uk, AIAA Professional Member.
‡Postdoctoral Research Fellow, School of Aerospace, Transport and Manufacturing, burak.yuksek@cranfield.ac.uk, AIAA Professional Member
§BAE Systems Chair, Professor of Autonomous Systems and Artificial Intelligence, inalhan@cranfield.ac.uk, AIAA Associate Fellow

and angle-of-sideslip that overcomes these challenges and achieves high-accuracy estimations without the need for air data booms. Our approach involves developing a customized neural network architecture that incorporates additional information about the wind parameters to improve the accuracy of the estimation. We validate the performance of our methodology using data from a 6-degrees-of-freedom F-16 simulator and demonstrate its potential for improving the safety and performance of aircraft operations. After considering these fundamental requirements for performance, safety, and stealthiness, the paper's primary motivation is to design a neural network-based synthetic air data system to estimate aerodynamic angles α and β. We utilize an open-source F-16 aircraft model to generate the training dataset from the JSBSim simulation environment [1]. We applied several modifications to measurement models to obtain realistic measurements. In addition, we developed an automatic maneuver generation logic to perform maneuvers that cover the whole flight envelope. Additional attitude, altitude, and flight speed control systems are developed to recover the aircraft from adverse conditions for the continuity of the data collection process in different mass and flight conditions. Collected data is utilized for training a data-driven synthetic air data system.

To find the nonlinear mapping for the air data, we utilize deep neural networks, which have been proven to capture complex input-output relationships through gradient-based optimization and can be used to model any continuous function [2, 3]. Neural networks are not a recent concept in air data estimation. Rohloff et al. [4] and Samy [5] proposed introducing neural networks to construct virtual sensors from static pressure measurements on fuselage without using inertial data. Similarly, Borup et al. [6] got measurements from pressure sensors installed over the UAV structure combined with wind tunnel tests and used neural networks to model air data. Lerro has studied the problem from both theoretical and hardware perspectives and applied the technique for UAVs, and small category aircraft [7–10]. The main advantage of neural networks over filtering algorithms is that they are model-free. Benefiting the advances in deep learning, one can seamlessly capture highly-nonlinear dynamics. Neural networks can also integrate differential equations or model-based constraints to introduce further system physics [11], [12]. An alternative method to estimate the air data is filtering. Lie [13], [14] proposed a cascaded Extended-Kalman filtering architecture to calculate the angle of attack, sideslip, and aircraft velocity. The drawback of the algorithm is the need for an aircraft model, which is challenging to achieve in high fidelity. Sun [15] carried this work further to overcome this issue by proposing a model-free filtering system. Even though the algorithm yields promising results, it relies on the wind estimation module, which is highly sensitive to wind uncertainties.

This paper presents two significant contributions that aim to enhance the capabilities of existing air data estimation methods by leveraging synthetic data generation and advanced machine learning techniques. The first contribution of this paper involves modifying the baseline flight simulator to generate more realistic flight data that covers the complete flight envelope. This is achieved by incorporating all feasible maneuvers and considering a comprehensive set of flight conditions. To accomplish this, we have meticulously incorporated critical aspects of aircraft dynamics, such as aerodynamic forces, propulsion systems, control surface deflections, and environmental factors, into the simulator. Including these factors ensures that the generated data closely resembles real-world flight data and adequately captures the complex interactions between the aircraft and its environment. This synthetic data generation approach overcomes the limitations of traditional data collection methods, which often provide an incomplete representation of the flight envelope, and significantly reduces the operational costs associated with acquiring extensive flight data. The second contribution of our paper is the proposal of a customized neural network architecture designed specifically for predicting the angle of attack and sideslip angle given a proper set of inputs. To maximize the neural network's performance, we have employed extensive feature engineering techniques to identify the most relevant input parameters for the model. The customized neural network architecture is designed with multiple layers and a large number of hidden neurons to accommodate the high dimensionality and complexity of the input data. This design allows the model to learn intricate relationships between inputs and outputs, ultimately leading to accurate predictions of the angle of attack and sideslip angle.

In the following sections of this paper, we present a comprehensive discussion of our proposed air data estimator framework and its components. First, we outline the air data estimator framework, detailing its design and the rationale behind its development. Next, we delve into the simulation environment, describing the high-fidelity aircraft model and its role in generating realistic flight data. The data generation for the full flight envelope section elaborates on our approach to creating a comprehensive dataset by incorporating all feasible maneuvers and flight conditions. Subsequently, we introduce the artificial neural networks model, detailing the architecture, feature engineering, and training process for predicting critical flight parameters. In the estimation results section, we demonstrate the effectiveness of our proposed framework by comparing the predicted values with actual flight data, highlighting the accuracy and robustness of our model. Finally, we conclude the paper by summarizing our contributions, their implications for air data estimation, and potential avenues for future research.

II. Air Data Estimator Framework

The air data estimator (ADE) framework depicted in Fig. 1 consists of data collection/generation and model training. One conventional way of collecting data is to get historical data from previous flight records for supervised learning. This approach's limitation is that the data do not cover the full flight envelope as it only represents what the aircraft had flown. This will cause the machine learning model not to learn the complete flight dynamics, as it will be fed with a subspace of the full envelope. A solution to this issue is collecting lots of flight data, which is operationally expensive. This paper proposes a synthetic environment that includes a high-fidelity aircraft model and can generate flight data for all feasible flight conditions. This data is then fed to the machine learning process to learn a predictive model that calculates the angle of attack and sideslip angle. The actual flight data is only utilized to calibrate the aircraft model parameters within the synthetic environment.

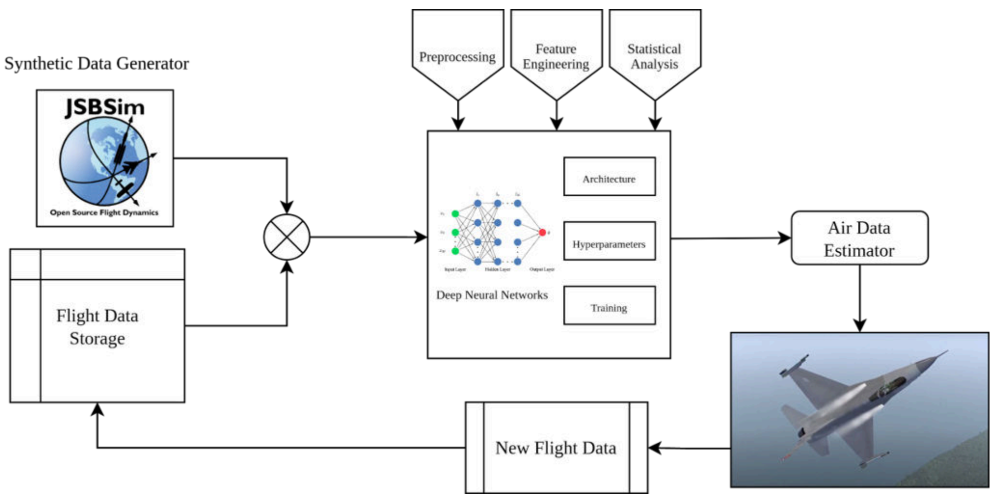


Fig. 1 Air data estimator: data collection and model training cycle

The synthetic environment developed for this research is a comprehensive flight simulator capable of generating flight data across the entire flight envelope. This environment is built upon a high-fidelity aircraft model that accurately captures the complex interactions of aerodynamic forces, propulsion systems, and flight control mechanisms. The simulator uses a six-degrees-of-freedom (6-DOF) rigid body model, accounting for all the necessary forces and moments acting on the aircraft. Additionally, it integrates realistic atmospheric conditions, including wind, temperature, and pressure variations, to ensure that the generated data closely resembles real-world flight conditions.

To create the high-fidelity aircraft model, we first establish a baseline model using available aerodynamic data, control surface characteristics, and propulsion system performance. Then, we apply an iterative process of fine-tuning and calibration, utilizing actual flight data to refine the model parameters. The primary objective of this step is to ensure that the synthetic environment accurately reflects the aircraft's response to control inputs and environmental changes. Furthermore, the refined model can be used to generate flight data for various flight conditions, including off-nominal scenarios that may not have been encountered during actual flights. The data generation process begins by defining a broad set of flight conditions, encompassing the full flight envelope. This includes varying altitudes, airspeeds, aircraft configurations, and control surface deflections. For each flight condition, the simulator generates a corresponding set of data points, including the angle of attack, sideslip angle, and other relevant parameters. The resulting dataset, comprising millions of data points, is then pre-processed to remove any outliers or inconsistencies, ensuring the integrity of the data for subsequent machine learning processes. With the dataset in hand, we proceed to train a predictive model using supervised learning techniques. In this study, we employ a deep neural network (DNN) architecture chosen for its ability to learn complex relationships between inputs and outputs. The network is designed with multiple layers and a large number of hidden neurons to accommodate the high dimensionality of the input data. The output of the trained model is the predicted angle of attack and sideslip angle for any given flight condition.

To validate the performance of the proposed ADE framework, we compare the predicted angle of attack and sideslip angle from the DNN model with simulated flight data. This comparison demonstrates the effectiveness of our synthetic environment and high-fidelity aircraft model in generating representative flight data for training the machine learning model. Moreover, the results show that the ADE framework can successfully learn the full flight dynamics, overcoming the limitations of relying solely on historical flight data. The rest of the section explains our synthetic environment, the modifications applied for a high-fidelity model, the data generation process, and the supervised learning algorithm.

III. Simulation Environment

This study uses JSBSim as our simulation environment, an open-source flight dynamics model (FDM) widely recognized in the literature [1, 16, 17]. This model incorporates 6-degrees-of-freedom equations of motion, considering the rotating Earth and quaternion representation for orientation [18]. Additionally, the simulation environment features the 1976 Standard Atmosphere Model and an appropriate gravity model to ensure realistic environmental conditions. The aerodynamic and propulsion forces and moments are modeled using look-up tables, which effectively capture the complex relationships between various flight parameters and the aircraft's response.

The wind model structure employed in our simulation environment encompasses both 1-cosine gust and turbulence effects to represent realistic atmospheric conditions during flight better [19–21]. The turbulence model parameters are derived using a discrete implementation of the Dryden Model, a widely adopted approach for simulating atmospheric turbulence in flight dynamics simulations. This implementation captures the statistical properties of turbulence and ensures that the generated flight data reflect the aircraft's response to such disturbances. By MIL-STD-1797A, the 1-cosine gust model is defined as follows:

$$
\begin{aligned}
&v = 0 && x < 0 \\
&v = \frac{V_m}{2}\left(1 - \cos\frac{\pi x}{dm}\right), && 0 \leq x \leq d_m \\
&v = v_m && x > d_m
\end{aligned}
\tag{1}
$$

The 1-cosine gust model simulates a deterministic gust profile, providing a means to assess the aircraft's performance and stability in the presence of gust disturbances. By incorporating both turbulence and gust effects into our simulation environment, we ensure that the generated flight data captures a broad range of atmospheric conditions, enhancing the robustness and accuracy of the subsequent machine-learning model. This comprehensive approach to modeling wind conditions also contributes to a more accurate representation of the aircraft's performance and stability characteristics under various flight scenarios.

The F-16A Block-32 jet fighter was chosen as the aircraft model for the simulation due to its detailed model structure and the availability of reliable references in the literature [22–24]. This aircraft has been extensively studied, and its performance characteristics are well-documented, making it an ideal candidate for our simulation environment. Moreover, the F-16A model has been subject to various modifications and developments over the years, further enhancing its suitability for our study. By using the F-16A Block-32 jet fighter as the basis for our high-fidelity aircraft model, we ensure that our simulation environment accurately captures the complex interactions between aerodynamic forces, propulsion systems, and flight control mechanisms. This choice also allows us to leverage the wealth of existing knowledge and data on the aircraft, facilitating the fine-tuning and calibration of our model to generate realistic flight data.

In addition to the aforementioned features of the simulation environment, we have also developed and integrated an inertial sensor measurement model that accounts for various sources of error commonly encountered in real-world applications. These sources of error include bias, G-dependent bias, random noise, quantization error, and measurement lag, all of which can significantly impact the accuracy and reliability of sensor measurements. To ensure that our measurement error model accurately reflects the performance of actual inertial sensors, we have adopted the parameters from the LN-200 Tactical grade Inertial Measurement Unit (IMU) specifications, which are designed to perform effectively across the entire military operating environment [25]. By incorporating these parameters into our simulation environment, we can generate realistic sensor data that takes into account the inherent limitations and uncertainties associated with inertial sensing systems.

$$
y_{m_{gyro}} = y_{gyro} + b_{gyro} + n_{gyro} + \cdots + k_{g_{gyro}} a_{body} \tag{2}
$$

Besides the core components of the simulation environment, we have also implemented additional control systems within the aircraft model, including attitude, altitude, and velocity controllers. These control systems play a crucial role during the automatic maneuver generation phase, as they enable the recovery of the aircraft from adverse flight conditions, allowing for the collection of long-duration flight data. By incorporating these additional controllers into the aircraft model, we can effectively capture the dynamic responses of the aircraft under various flight and weight conditions. One example of the integrated control system is the attitude controller, as illustrated in Fig. 2. This controller is specifically utilized to regulate the orientation of the aircraft between maneuver legs, ensuring stability and smooth transitions during complex flight maneuvers. The inclusion of such control systems not only enhances the realism of our simulation environment but also provides valuable insights into the performance and stability characteristics of the aircraft under a diverse array of operating conditions.

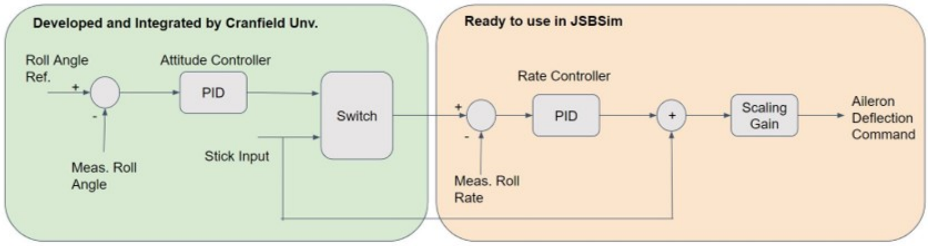


Fig. 2 Pitch attitude controller

IV. Data Generation for Full Flight Envelope

As described in the previous section, the synthetic data generator requires three input files to execute the algorithm effectively: the aircraft model, initial conditions, and maneuver files. The maneuver file is comprised of time-dependent event lists that dictate the aircraft's behavior during the simulation. However, due to the high dimensionality of the data, it is infeasible to manually generate these event lists for the purpose of creating comprehensive flight datasets. To address this challenge, we have developed an automatic maneuver script that efficiently generates maneuver files for the synthetic data generator. This script directly produces JSBSim-compatible maneuver files with an "*xml*" extension, streamlining the process of generating flight data across a wide range of flight conditions and scenarios. By utilizing the automatic maneuver script, we can significantly reduce the time and effort required to generate diverse and realistic flight data, ultimately enhancing the effectiveness of our synthetic data generation approach. Furthermore, this automated process facilitates the exploration of various flight scenarios and conditions, enabling the machine learning model to learn from a more comprehensive dataset, and ultimately improving its predictive capabilities.

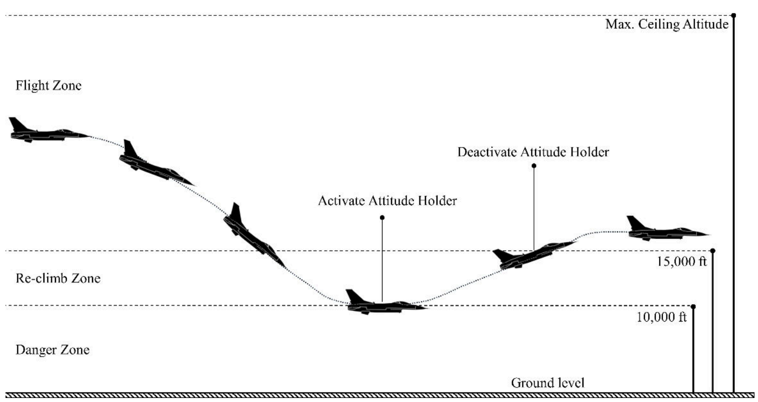


Fig. 3 Automatic maneuver generation algorithm

Fig. 3 illustrates the main working principle of the automatic maneuver generation algorithm. Within the "*Flight Zone*", input values such as body rates and thrust lever position are randomly selected from a specified range with a certain probability. This randomness ensures that a diverse set of flight conditions and scenarios are covered during the simulation. If the aircraft approaches the "*Danger Zone*", which represents a region close to the ground or other hazardous conditions, the script activates the Attitude control system to prevent the aircraft from descending further and potentially crashing. This region is referred to as the "*Re-Climb Zone*". Once the aircraft successfully climbs back to the "*Flight Zone*", it resumes its flight with random maneuvers. The altitude boundaries for the flight zones can be easily adjusted from the input section of the automatic maneuver generation algorithm. This flexibility allows for the exploration of various flight scenarios and conditions, ultimately contributing to the comprehensive nature of the dataset used for training the machine learning model. By incorporating this automated maneuver generation process into the synthetic data generator, we ensure that the generated flight data effectively spans the entire flight envelope, enhancing the robustness and accuracy of the resulting air data estimator.

Furthermore, the algorithm allows for the adjustment of the input parameters' range, such as normalized body rates and thrust level, which are included in the events. This flexibility enables the generation of time-dependent random maneuvers, utilizing variable probabilistic distributions in the created maneuver sets. It also allows users to fine-tune the algorithm by modifying the input range and probabilistic distribution parameters according to the desired flight characteristics, facilitating the exploration of various flight scenarios. Thanks to the efficiency and adaptability of the automatic maneuver generation algorithm, a maneuver file containing a diverse set of flight conditions, including aggressive agile maneuvers, can be generated rapidly. An example of such maneuvers is depicted in Fig. 4.

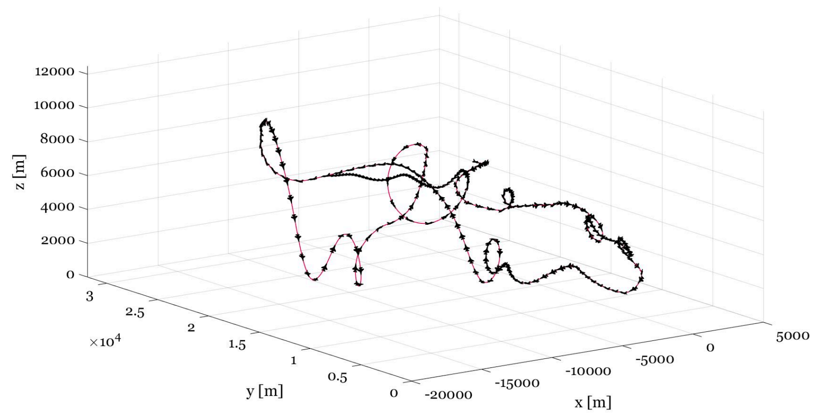
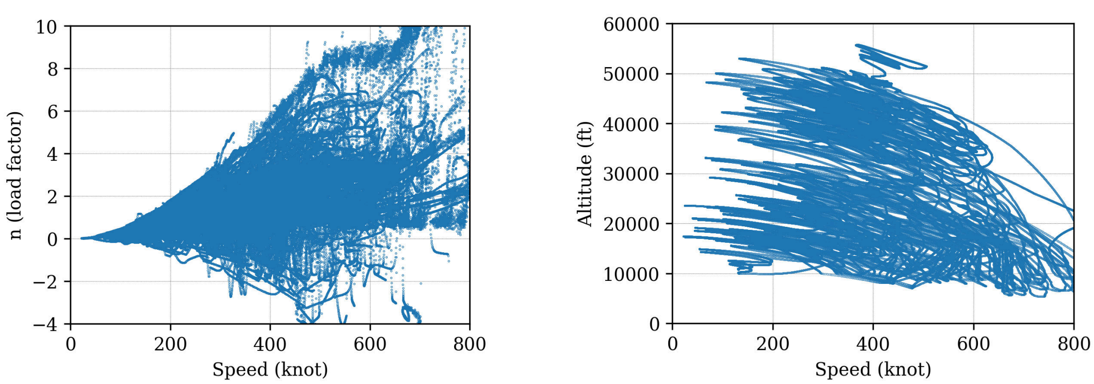


Fig. 4 Flight path of a sample maneuver set

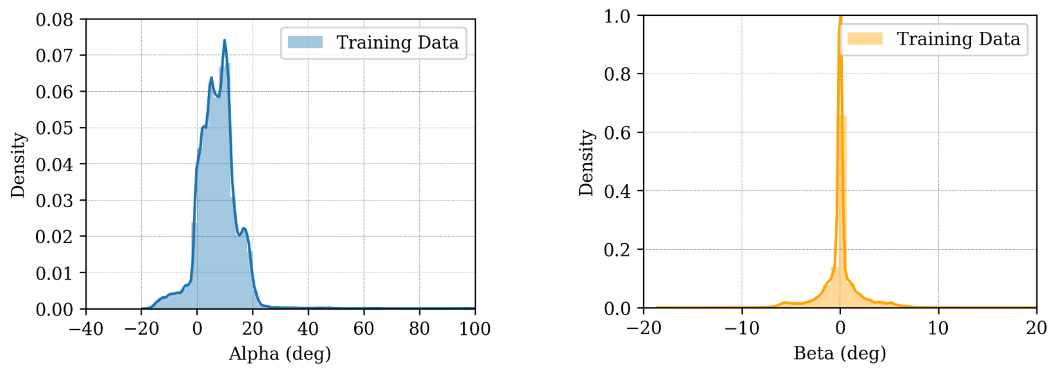
Leveraging the automatic maneuver generation algorithm, we produced a large number of training datasets to ensure comprehensive coverage of various flight conditions. The final dataset comprises 288,000 data points, equivalent to a 4-hour flight, and includes 16 distinct flight sets lasting between 15 and 20 minutes each. A sampling time of 0.05 seconds was employed during data collection from the simulation. In creating the datasets, the probability of random control surface inputs and the range of control surface deflections were adjusted to ensure that each flight set exhibited unique characteristics. For instance, 4 of the 16 flights in the final dataset consisted of non-aggressive flights with lower angles of attack. Another 4 flights feature maneuvers with high angles of attack and medium-low thrust levels at stall limits. Four additional flights in the dataset incorporate gradual climbing and descending maneuvers at varying elevator deflections and throttle levels. The remaining 4 flights consist of aggressive agile maneuvers with high kinetic energy.



(a) V-n diagram **(b) V-h diagram**

Fig. 5 Flight envelope and altitude envelope of training data

Utilizing the synthetic environment outputs and basic flight mechanics equations, we have determined the flight and altitude envelopes of the training dataset. As depicted in Fig. 5, the flight envelope adheres to the real aircraft's g-limits, ensuring a realistic representation of the aircraft's performance capabilities. This compliance is achieved by incorporating the background Stability Augmentation System (SAS) into the simulation, which provides artificial stability to maintain the proper flight characteristics of the aircraft. The incorporation of the SAS not only enhances the fidelity of the synthetic environment but also ensures that the generated flight data closely mirrors the behavior of an actual aircraft under various flight conditions. This alignment between the synthetic data and real-world performance allows for the development of a more accurate and reliable air data estimator, ultimately improving its applicability and effectiveness in practical scenarios. By simulating a realistic flight envelope and incorporating essential systems like the SAS, our approach ensures that the machine learning model is trained on a comprehensive and representative dataset, thus maximizing its predictive capabilities.



(a) Alpha distribution **(b) Beta distribution**

Fig. 6 Distribution of training data

While the generated dataset encompasses a broad range of angle of attack data, a closer examination of the distributions within the final dataset reveals a noteworthy concentration of data points, particularly in cruise trim angles. This observation suggests that the dataset may be more heavily weighted toward stable flight conditions, which could potentially influence the performance of the air data estimator in less common or more extreme scenarios. Additionally, although the incorporation of wind and turbulence models enables the generation of data with high sideslip angle values, it is observed that the sideslip angle predominantly remains around zero degrees for the majority of the flights (as illustrated in Fig. 6). This finding indicates that the dataset may not fully capture the diverse range of sideslip angle conditions that could be encountered in real-world flight scenarios. Despite these observations, the dataset still provides a valuable foundation for training the machine learning model. However, future refinements to the synthetic environment and data generation process could potentially address these limitations and further enhance the comprehensiveness and representativeness of the dataset, ultimately contributing to the development of a more robust and accurate air data estimator.

V. Artificial Neural Networks Model

We adopt a fundamental multi-layer perceptron architecture to regress the air data, Y, using X. For a fully-connected network with L hidden layers, this amounts to the following modelling equations relating the input features x, to its target prediction y. A generic structure of a deep neural network consisting of a multilayer perceptron with M input features and N layers. It is composed of sequentially connected layers, which comprise sets of neurons that are combinations of mathematical operations followed by nonlinear activation functions. The model parameters ξ are defined as $\xi = \{W, b\}$, where $W = \{w_i\}_{i=1}^{N}$, and $b = \{b_i\}_{i=1}^{N}$. The output of the lth layer is:

$$f(x, \xi_l) = f_{w_l, b_l}(x) = z_l\left(\sum_{j=1}^{N_l} w_{lj} x_j + b_l\right) \quad (3)$$

$$= Z_l\left(w_l^T x_l + b_l\right) \quad (l = 1, \dots, N) \quad (4)$$

where N_t is the neuron number and Z_l is a nonlinear activation function of a specific lth layer, and x_l is the input of the lth layer and also the output of the $(l-1)$th layer. In addition to this, w_l and b_l are learnable parameters and called weight and bias terms of that layer respectively. Finally, the output of the last layer is a result of the composite and complex mapping defined as:

$$\hat{y}(x) := \left(f_{w_n,b_n} \circ \cdots \circ f_{w_1,b_1}\right)(x) \tag{5}$$

This work also proposes a cascaded neural network architecture for estimating air data. The architecture depicted in Fig. 7 is a single neural network consisting of two parts: The first part estimates the intermediate features:

$$\hat{y}_{int} := \left(f^{[1]}_{w_n,b_n} \circ \cdots \circ f^{[1]}_{w_1,b_1}\right)(x) \tag{6}$$

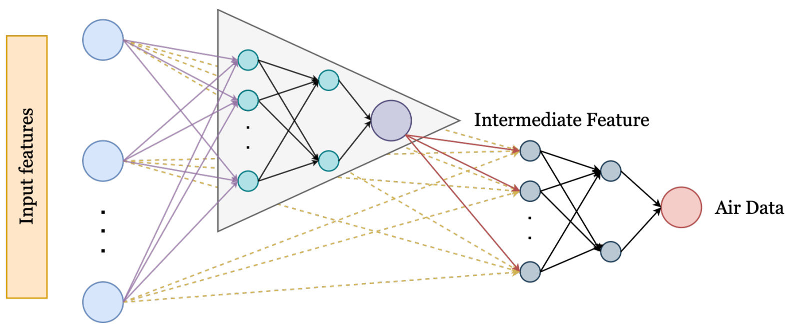


Fig. 7 Cascaded neural-network architecture for air data estimation

where $\hat{y}_{int}$ represents $V_{Z-body_{aero}}$ and $V_{Y-body_{aero}}$ for the angle of attack and angle of sideslip, respectively. Then, we concatenate the intermediate feature to the main feature set and provide this new input to the second part of the neural network, where we estimate the output variable. Note that these two parts of the neural network are not trained separately. All trainable parameters in both parts are updated simultaneously at each mini-batch.

$$x_{int} = \{x, \hat{y}_{int}\} \tag{7}$$

$$\hat{y} := \left(f^{[2]}_{w_n,b_n} \circ \cdots \circ f^{[2]}_{w_1,b_1}\right)(x_{int}) \tag{8}$$

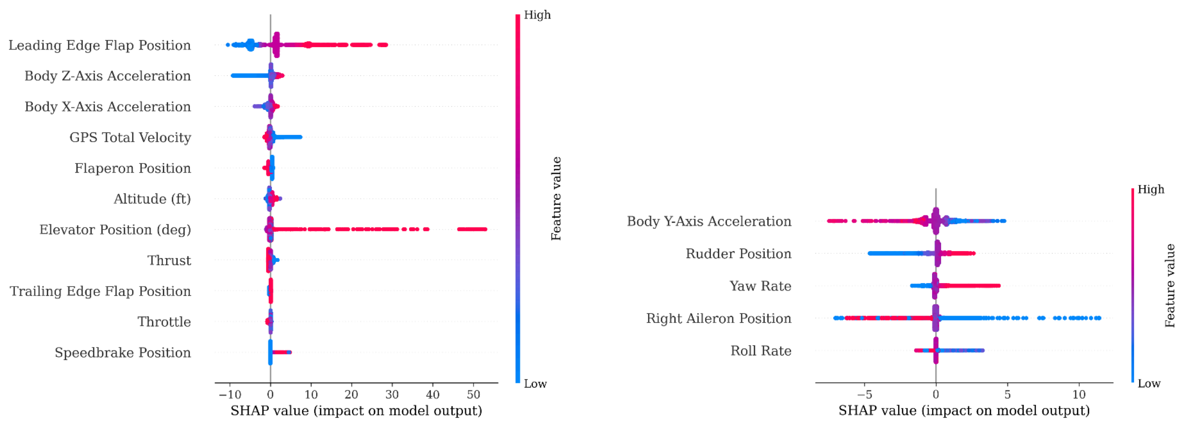
The customized neural network architecture aims to increase the accuracy of angle-of-attack and angle-of-sideslip estimation by incorporating additional information about $V_{Z-body_{aero}}$ and $V_{Y-body_{aero}}$, which are not precisely available in actual flights due to wind uncertainty. To address this issue, we introduce an intermediate estimator to provide the neural network with additional information about these variables. As the loss in the intermediate estimator decreases, the accuracy of the main output is expected to increase further. The inclusion of this additional information enables the neural network to better capture the underlying relationships between air data and measurements from various sensors, resulting in a more accurate estimation of angle-of-attack and angle-of-sideslip. Furthermore, the use of a customized neural network architecture allows us to take advantage of the specific characteristics of our data and tailor the model to our specific problem domain.

VI. Estimation Results

In this section, we present the estimation results obtained from the trained machine learning model based on the comprehensive dataset generated using the synthetic environment and automatic maneuver generation algorithm. These results provide insights into the effectiveness and accuracy of our air data estimator in predicting the angle of attack and sideslip angle under various flight conditions. By analyzing the performance metrics and examining the model's ability to handle diverse flight scenarios, we can assess the overall success of our proposed approach in developing a robust and reliable air data estimator. Furthermore, we will discuss any potential limitations or areas for improvement, offering a foundation for future refinements and enhancements to our methodology.

The input features for the models, as detailed in Table 1, are derived from flight equations and a Spearman correlation analysis applied to the data described in Section IV. We have carefully chosen features that are accessible through IMU units and other digital sensors available on the aircraft. Notably, the airspeeds u, v, and w are excluded from the feature set to prevent data leakage, as JSBSim employs these values to calculate the angle of attack and sideslip kinematically.

In addition to the feature selection process through Spearman correlation, we conducted a feature importance analysis using SHAP (SHapley Additive exPlanations) values [26–28] to gain further insights into the contributions of each input feature to the model's predictions. SHAP values are a powerful and interpretable method for explaining individual predictions of machine learning models, particularly in the context of regression problems. For regression problems, SHAP values quantify the contribution of each feature to the predicted output for a specific data point. Positive SHAP values indicate that a particular feature increases the predicted output, while negative SHAP values suggest that the feature decreases the predicted output. By computing the average absolute SHAP values for each feature across all data points, we can rank the features according to their overall importance in the model. This analysis helps us better understand the relationships between input features and the model's predictions and offers valuable insights for potential model improvements or feature engineering efforts. Figure 8 depicts the feature importance analysis for both angle of attack and sideslip, highlighting the selected features that exhibit meaningful physical relationships. For example, the leading edge flap demonstrates a considerable impact on the angle of attack, while the body Y-axis acceleration and aileron position influence the angle of sideslip significantly.



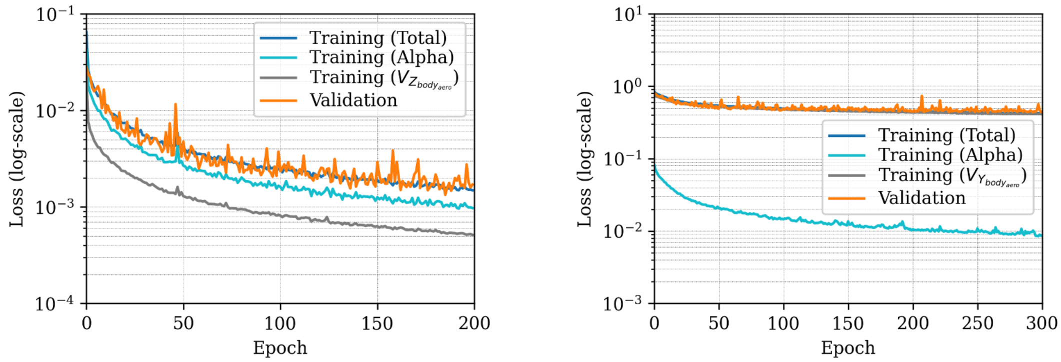
(a) **(b)**

Fig. 8 SHAP values for (a) the angle of attack and (b) the sideslip angle

Table 1 Summary of features for training phases

Angle of attack (α)	**Sideslip angle** (β)
Elevator Position	Rudder Position
Leading Edge Flap Position	Aileron Position
Trailing Edge Flap Position	Yaw Rate
Flaperon Position	Roll Rate
Speedbrake Position	Body Y-Axis Acceleration
Body X-Axis Acceleration	
Body Z-Axis Acceleration	
GPS Total Velocity	
Throttle	
Altitude	

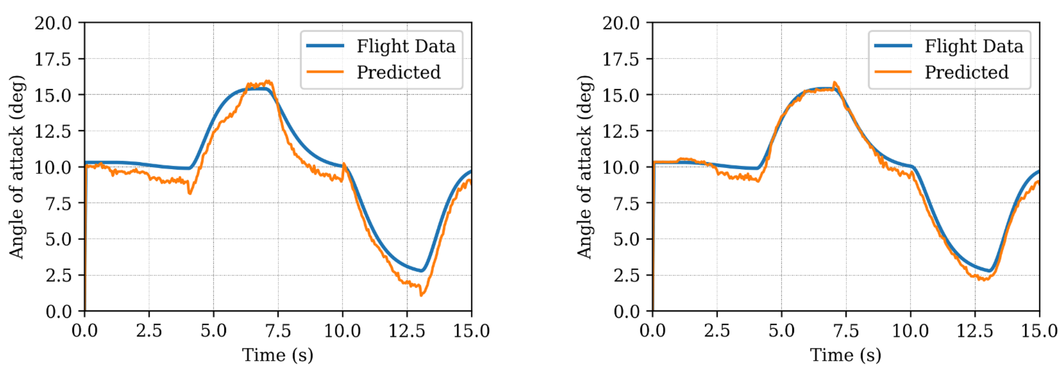
By employing the selected features, we train distinct models for the angle of attack and sideslip estimation. For the angle of attack estimation, we use a cascaded deep neural network architecture consisting of five layers, containing 512, 256, 128, 64, and 32 neurons. In contrast, the sideslip angle estimation model comprises five layers with 32, 16, 8, 4, and 2 neurons, respectively. The training data for these models encompass 288,000 flight data points, equivalent to four hours of flight time. Meanwhile, the test set consists of separate flights, totaling 20 minutes of flight time. Both the training and test datasets have a sampling rate of 20 Hz. This approach ensures a comprehensive representation of various flight scenarios, which ultimately contributes to the development of robust and accurate angle of attack and sideslip estimators.



(a) Angle of attack (α) training **(b) Sideslip angle (β) training**

Fig. 9 Learning curves of developed model

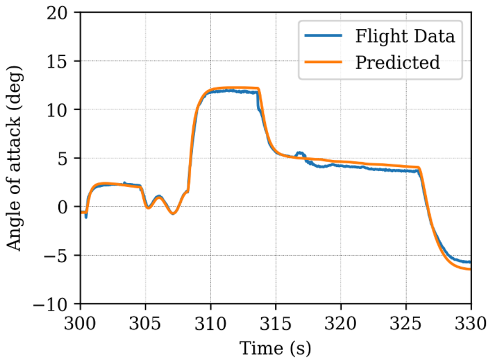
We develop separate models for the angle of attack and sideslip by employing the custom neural network architecture outlined in Section V. For the intermediate parameters, we choose $V_{Z-body_{aero}}$ and $V_{Y-body_{aero}}$ as intermediate variables for the angle of attack and angle of sideslip, respectively. The evolution of loss functions is depicted in Figure 9, where both intermediate and output values can be observed to converge toward the ground truth. Notably, the validation losses are higher than the training losses, indicating that the models do not suffer from overfitting to the training data. In Figure 10, a comparison is made between the proposed custom neural network architecture and a generic feedforward one. The results demonstrate that the customized architecture outperforms the generic feedforward network, showcasing the effectiveness of the tailored approach in capturing the underlying relationships between the input features and the target variables for the angle of attack and sideslip estimation.



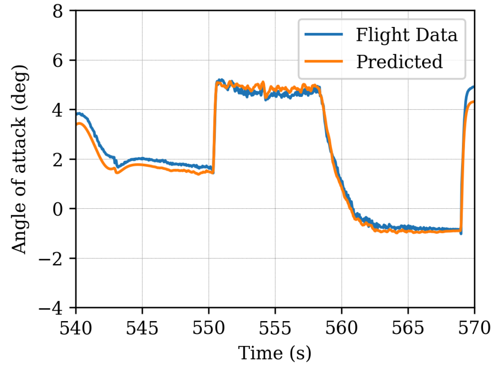
(a) Default NN Model (MAE: 0.839) **(b) Custom NN Model (MAE: 0.338)**

Fig. 10 Comparison of Neural Networks Architectures on a test case

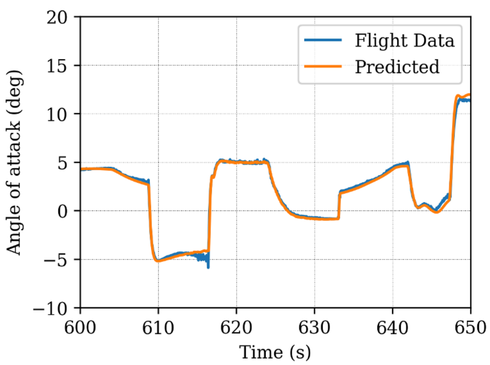
The analysis conducted within the simulation environment indicates that the proposed data-driven synthetic air data system demonstrates promising results. In Fig. 11, the predicted angle-of-attack values are illustrated as blue lines, while the actual values are represented by orange lines. During the test phase, the mean absolute error (MAE) is found to be 0.534 degrees. Of particular importance in this simulation is the observation that the proposed system can estimate the angle-of-attack with a reasonable error bound, even under high angle-of-attack conditions. This feature holds significant relevance from a flight safety perspective, as it demonstrates the system's capability to alert the pilot or operator to recover the aircraft in near-stall situations. Such a capability underscores the potential of the proposed synthetic air data system in enhancing flight safety and reliability.



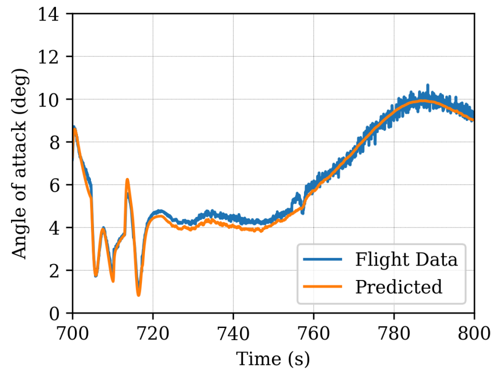
(a)



(b)



(c)



(d)

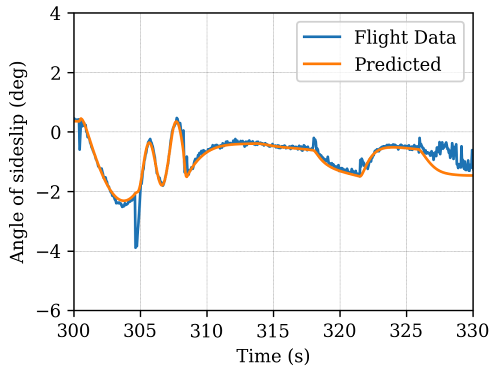
Fig. 11 Angle of attack (α) estimation results for the test scenarios (MAE: 0.534 deg)

In a similar fashion, the proposed system is capable of estimating the angle-of-sideslip with a small error bound, as demonstrated by the results presented in Fig. 12. The mean absolute error (MAE) during the test phase for the angle-of-sideslip estimation is found to be 0.247 degrees. This outcome further emphasizes the effectiveness and robustness of the proposed data-driven synthetic air data system in accurately estimating critical flight parameters, such as the angle of attack and angle of sideslip, which are essential for maintaining safe and stable flight conditions.

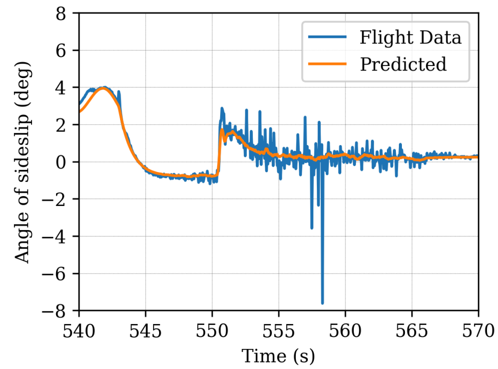
Table 2 Comparison of mean absolute errors for different models

Method	Angle of Attack Error (deg)	Angle of Sideslip Error (deg)
Linear Regression	2.538	0.486
Gradient Boosting	1.178	0.448
Regular Neural Network	0.897	0.396
Customized Neural Network	0.534	0.247

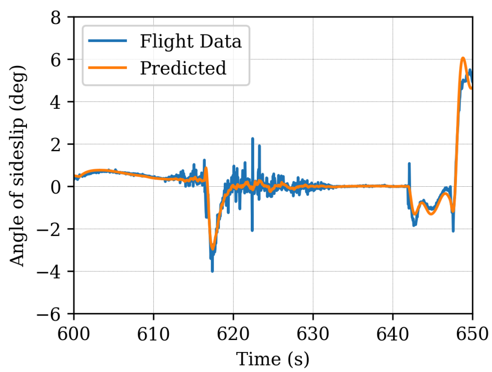
In Table 2, we compare the estimation results for the angle of attack and angle of sideslip using four different models: linear regression, gradient boosting, regular feedforward neural network, and our proposed customized neural network architecture. Our analysis shows that the linear regression model results in a significantly high error when estimating the angle of attack, which indicates its inadequacy in capturing the complex relationship between the input features and the angle of attack. On the other hand, our proposed customized neural network architecture yields the best results in estimating the angle of attack, demonstrating its effectiveness in modeling the flight dynamics and learning the complex relationships present in the data. As for the angle of sideslip estimation, the results across all four models do not vary significantly. This can be attributed to the fact that most of the data points for sideslip angle are centered around zero degrees, which simplifies the learning task for all models. Despite this, it is worth noting that our customized neural network architecture still performs well in capturing the underlying relationships in the data, making it a suitable choice for both angle of attack and the angle of sideslip estimation.



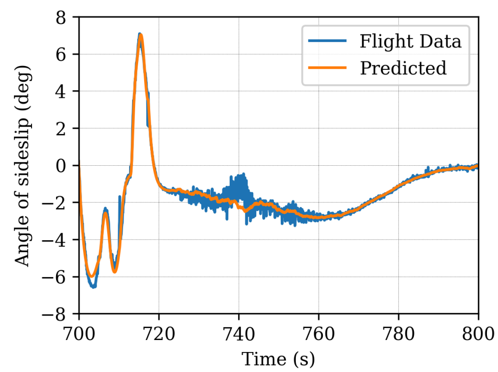
(a)



(b)



(c)



(d)

Fig. 12 Sideslip angle (β) estimation results for the test scenarios (MAE: 0.247 deg)

VII. Conclusions and Future Work

In conclusion, this study has successfully demonstrated a data-driven synthetic air data system capable of estimating the angle-of-attack and angle-of-sideslip of a fighter aircraft with high accuracy. Utilizing a high-fidelity simulation environment, we generated training data and tested the proposed method, achieving low test phase mean absolute errors of 0.534 deg for the angle of attack and 0.247 deg for the sideslip. Notably, the proposed system can estimate the angle of attack in near-stall conditions, highlighting its importance from a flight safety perspective. As part of our future work, we aim to conduct a benchmark study in which the proposed synthetic air data system will be compared to state-of-the-art filtering techniques, such as the Extended Kalman Filter (EKF) or particle filter. Additionally, we plan to explore the potential of utilizing physics-informed neural networks to enhance the estimation performance of our method further. This comparative analysis will enable us to evaluate the estimation performance of our proposed method and help identify potential areas for improvement or refinement. Overall, our work has laid the groundwork for a robust and accurate air data estimation system that can be leveraged to enhance the safety and efficiency of fighter aircraft operations. By incorporating advanced techniques such as physics-informed neural networks, we aim to develop a more sophisticated and reliable air data estimation framework for future applications.

References

[1] Berndt, J., Peden, T., Culp, D., Megginson, D., Hofman, E., Frölich, M., Marco, A., Luff, D., Duke, L., Galbrait, B., et al., "JSBSim: An Open Source, Platform-Independent," *Flight Dynamics Model in C++. Jon S. Berndt & the JSBSim Development Team*, 2011.

[2] Winkler, D. A., and Le, T. C., "Performance of deep and shallow neural networks, the universal approximation theorem, activity cliffs, and QSAR," *Molecular informatics*, Vol. 36, No. 1-2, 2017, p. 1600118.

[3] Chen, T., and Chen, H., "Universal approximation to nonlinear operators by neural networks with arbitrary activation functions and its application to dynamical systems," *IEEE Transactions on Neural Networks*, Vol. 6, No. 4, 1995, pp. 911–917.

[4] Rohloff, T. J., Whitmore, S. A., and Catton, I., "Fault-tolerant neural network algorithm for flush air data sensing," *Journal of aircraft*, Vol. 36, No. 3, 1999, pp. 541–549.

[5] Samy, I., Postlethwaite, I., Gu, D.-W., and Green, J., "Neural-network-based flush air data sensing system demonstrated on a mini air vehicle," *Journal of aircraft*, Vol. 47, No. 1, 2010, pp. 18–31.

[6] Borup, K. T., Fossen, T. I., and Johansen, T. A., "A machine learning approach for estimating air data parameters of small fixed-wing UAVs using distributed pressure sensors," *IEEE Transactions on Aerospace and Electronic Systems*, Vol. 56, No. 3, 2019, pp. 2157–2173.

[7] Lerro, A., Gili, P., and Caselle, M. S., "Development and evaluation of neural network-based virtual air data sensor for estimation of aerodynamic angles," *Politecnico di Torino*, 2012.

[8] Lerro, A., Battipede, M., Gili, P., and Brandl, A., "Advantages of neural network based air data estimation for unmanned aerial vehicles," *International Journal of Aerospace and Mechanical Engineering*, Vol. 11, No. 5, 2017, pp. 1090–1099.

[9] Lerro, A., Battipede, M., Brandl, A., Gili, P., Rolando, A. L. M., and Trainelli, L., "Test in operative environment of an artificial neural network for aerodynamic angles estimation," *28th Society of Flight Test Engineers European Chapter Symposium (SFTE-EC 2017)*, 2017, pp. 1–12.

[10] Lerro, A., Brandl, A., Battipede, M., and Gili, P., "Preliminary design of a model-free synthetic sensor for aerodynamic angle estimation for commercial aviation," *Sensors*, Vol. 19, No. 23, 2019, p. 5133.

[11] 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*, 2022, pp. 1–12.

[12] Uzun, M., Demirezen, M. U., and Inalhan, G., "Physics Guided Deep Learning for Data-Driven Aircraft Fuel Consumption Modeling," *Aerospace*, Vol. 8, No. 2, 2021, p. 44.

[13] Lie, F. A. P., and Gebre-Egziabher, D., "Synthetic air data system," *Journal of Aircraft*, Vol. 50, No. 4, 2013, pp. 1234–1249.

[14] Lie, F. A. P., and Gebre-Egziabher, D., "Sensitivity analysis of model-based synthetic air data estimators," *AIAA Guidance, Navigation, and Control Conference*, 2015, p. 0081.

[15] Sun, K., Regan, C. D., and Gebre-Egziabher, D., "Observability and performance analysis of a model-free synthetic air data estimator," *Journal of Aircraft*, Vol. 56, No. 4, 2019, pp. 1471–1486.

[16] Berndt, J., "JSBSim: An open source flight dynamics model in C++," *AIAA Modeling and Simulation Technologies Conference and Exhibit*, 2004, p. 4923.

[17] Berndt, J., and De Marco, A., "Progress on and usage of the open source flight dynamics model software library, JSBSim," *AIAA modeling and simulation technologies conference*, 2009, p. 5699.

[18] Stevens, B. L., Lewis, F. L., and Johnson, E. N., *Aircraft control and simulation: dynamics, controls design, and autonomous systems*, John Wiley & Sons, 2015.

[19] Yeager, J. C., "Implementation and testing of turbulence models for the f18-harv simulation," Tech. rep., 1998.

[20] Specification", M., "Flying qualities of piloted airplanes," Tech. rep., MIL-F-8785C, 1980.

[21] Roger, H., Mitchell, D., Ashksnas, I., et al., "MIL–STD–1797A Flying Qualities of Piloted Aircraft," , 1990.

[22] Curry, R., and Curry, R., "Dynamic ground effect for a cranked arrow wing airplane," *22nd Atmospheric Flight Mechanics Conference*, 1997, p. 3649.

[23] Corda, S., *Dynamic ground effects flight test of an F-15 aircraft*, Vol. 4604, NASA, 1994.

[24] Nguyen, L. T., *Simulator study of stall/post-stall characteristics of a fighter airplane with relaxed longitudinal static stability*, Vol. 12854, National Aeronautics and Space Administration, 1979.

[25] Groves, P. D., "Principles of GNSS, inertial, and multisensor integrated navigation systems, [Book review]," *IEEE Aerospace and Electronic Systems Magazine*, Vol. 30, No. 2, 2015, pp. 26–27.

[26] Ribeiro, M. T., Singh, S., and Guestrin, C., "" Why should i trust you?" Explaining the predictions of any classifier," *Proceedings of the 22nd ACM SIGKDD international conference on knowledge discovery and data mining*, 2016, pp. 1135–1144.

[27] Štrumbelj, E., and Kononenko, I., "Explaining prediction models and individual predictions with feature contributions," *Knowledge and information systems*, Vol. 41, 2014, pp. 647–665.

[28] Shrikumar, A., Greenside, P., and Kundaje, A., "Learning important features through propagating activation differences," *International conference on machine learning*, PMLR, 2017, pp. 3145–3153.

2023-06-08

Data-driven synthetic air data estimation system development for a fighter aircraft

Karali, Hasan

AIAA