Multi-Sensor Aircraft Classification

Sarah Bolton
Department of Electrical and Computer Engineering
Air Force Institute of Technology
Wright-Patterson AFB, OH, USA
sarah.bolton@afit.edu

Richard Dill
Department of Electrical and Computer Engineering
Air Force Institute of Technology
Wright-Patterson AFB, OH, USA
richard.dill@afit.edu

Michael R. Grimaila
Department of Systems Engineering and Management
Air Force Institute of Technology
Wright-Patterson AFB, OH, USA
michael.grimaila@afit.edu

Douglas D. Hodson
Department of Electrical and Computer Engineering
Air Force Institute of Technology
Wright-Patterson AFB, OH, USA
douglas.hodson@afit.edu

***Abstract*—Automatic Dependent Surveillance–Broadcast (ADS-B) is a useful tool to for air traffic controllers, military and other sources that are invested in understanding a national or global air picture. While it is highly available, it can sometimes lack integrity due to hacking, spoofing or, even, unintentional inaccuracies in the broadcast. Unlike primary radar, ADS-B's lack of trustworthiness makes it not feasible to rely on it alone. Fusing other data sources with ADS-B can help confirm the accuracy of the broadcasts or allow ADS-B to act as a surrogate for primary radar and bolster the information that primary radar can provide. This paper presents an effective method of using ADS-B data as a surrogate for primary 3D radar by combining the kinematic information that ADS-B data provides with weather and aircraft images to make predictions about aircraft characteristics.**

***Index Terms*—Multivariate Long Short-Term Memory – Fully Convolutional Network, Automatic Dependent Surveillance-Broadcast, open-source data, classification, machine learning, sensor fusion**

I. Introduction

Automatic Dependent Surveillance–Broadcast (ADS-B) has become an important research tool over the last decade. Since its inception in 1995, many countries have begun to mandate the use of ADS-B within controlled airspace. Notably, the United States mandated the use of ADS-B within controlled airspace on 1 January 2020 with Federal Regulation 14 CFR 91.225 and 14 CFR 91.227 which was followed by Europe's ADS-B mandate in June 2020 with the Commission Implementing Regulation (EU) Number 1028/2014. These mandates apply to a large number of aircraft even if they are not based in Europe or the United States since international flights that fly through either of these areas are also required to comply with the mandate. While the exact number of aircraft that must comply with these regulations is difficult to determine, ADS-B repositories, such as the ADS-B Exchange, collect flight information on over 15,000 aircraft daily [1]. In the US alone, 163,216 aircraft were ADS-B compliant as of April 2023 [2].

Funding provided by the Air Force Research Laboratory.

One of the main features of the ADS-B system is open sharing and wide availability. For ease of access, there is no encryption or authentication. This makes ADS-B extremely vulnerable to cyber attacks and other inaccuracies. Jamming, eavesdropping, spoofing and injections of fake tracks would be easy for an attacker since ground stations cannot distinguish the difference between a real broadcast or a fake one. Since much of the broadcasted data is in the form of free text, even typos are an issue. Costin et al proved the vulnerabilities to ADS-B in 2012 [3]. For this reason, it's important to have a backup source of information when using ADS-B to ensure an accurate air picture. While ADS-B provides plenty of useful information for aircraft tracking, it cannot entirely replace primary radar due to its lack of integrity.

ADS-B's vulnerability concerns make fusing the ADS-B broadcast with other related sensors an important research area. While ADS-B on its own is not entirely trustworthy, using a backup sensor, like primary radar or images of the aircraft, allows ADS-B receivers to confirm the information being provided by the broadcast. For this research, we use only the kinematic data (speed, altitude, track and location) within ADS-B so that our model could act as a surrogate for a non-ADS-B data source. The kinematic data is able to simulate primary 3D radar and potentially eliminate the need to rely on ADS-B altogether. Using the kinematic data within ADS-B, this paper presents a method to predict an aircraft's wake turbulence category (WTC), description and designator as described in the ICAO doc 8643 [4]. To improve the accuracy of predictions using only kinematic data, we present a method of combining weather with the ADS-B data. Then, to ensure the integrity of the data, we use images to verify the authenticity of the ADS-B transmission. Section II discusses relevant literature. Section III reviews how the research and experiments were setup. Section IV presents the results from the research. Finally, we conclude in section V.

DOI 10.1109/CSCE60160.2023.00136

II. Background & Previous Work

A. Aircraft Prediction with ADS-B

ADS-B provides aircraft tracking and identification information via an onboard ADS-B Out transponder. It is broadcast in the clear at 1090 MHz in many countries worldwide and at both 1090 MHz and 978 MHz within the United States. In recent years, ADS-B has become an important tool for researchers. Due to its prevalence, ease of use, wide availability, and the dense information contained within the broadcast, ADS-B has been shown to be useful for a wide variety of aircraft research topics. The areas of research most related to this paper include enhancing aircraft operations [5]–[8], predicting flight patterns [9]–[13], identifying vulnerabilities [3], improving ADS-B transmission security [14]–[17] and predicting aircraft characteristics [18]–[22].

Created by Karim et al. in 2017, the Multivariate Long Short-Term Memory - Fully Convolutional Network (MLSTM-FCN) has been shown be effective in predicting characteristics about aircraft using the kinematic data found within ADS-B [19]–[23]. [20] and [19] used a Dual-Stage Deep Engine Classifier (DSDEC) that implemented the MLSTM-FCN to make predictions about aircraft engine types. They were able to achieve an overall accuracy of 89.2% with jet engine accuracy at 98.4%, turboprop accuracy at 79.2% and piston engine accuracy at 89.9% when looking at only the take-off phase of flight. The research completed in [22] predicted more definitive aircraft characteristics, specifically, WTC, description and designator using all flight phases instead of only the take off phase. While the accuracy was reduced overall due to the use of the entire flight path and the increase in prediction classes (3 engine types vs 5 WTC classes and 9 description classes), the take-off phase was able to produce slightly improved accuracy to [19] when predicting WTC and description.

B. Image Classification

While unrelated to ADS-B aircraft classification, aircraft image classification is important to understand for the purposes of sensor fusion. Many techniques have been shown to be effective in classifying an aircraft model based on its image. In 2013, Maji et al. developed an image benchmark that was gathered mostly from https://www.airliners.net/ to facilitate a baseline for other researchers with a focus on aircraft image classification [24]. The image repository, called the Fine-Grained Visual Classification of Aircraft (FGVC-Aircraft) benchmark, has been cited nearly 1,000 times in other aircraft classification research. The benchmark includes 10,000 airplane images spanning 100 different models. Each image is organized hierarchically by model, variant, family and manufacturer.

In their initial attempt at classifying the images in the benchmark, Maji et al. achieved a variant classification accuracy of 48.69% with a non-linear support vector machine [24]. More recent research improved the classification accuracy to about a 90% variant classification accuracy. In 2021, Rong et al. achieved a 93.3% accuracy using a Separated Smooth Sampling Network (SSSNet) [25]. In 2022, Yanfeng Wang et al. achieved a 89.2% accuracy using a Bat Algorithm and Convolutional Neural Network (BA-CNN) [26]. Also in 2022, Lei Wang et al. achieved a 91.4% accuracy using a Multilayer Feature Fusion (MFF) network with Parallel Convolutional Block (PCB) mechanism [27]. A more recent effort was completed by Liu et al. in March 2023 that developed a Cross-Layer Mutual Attention Learning Network (CMAL-Net) which achieved 94.7% accuracy [28]. Their CMAL-Net uses a Residual Network (Resnet) as its backbone. The architecture develops multiple classifiers that are trained to make predictions based on a specific layer from shallow to deep. The prediction from each layer is combined to create the final output. The backbone, the Resnet, is a Convolutional Neural Network (CNN) that uses a technique called skip connections to solve the exploding and vanishing gradient problems.

C. Sensor Fusion

Information fusion is described as the blending of data acquired from multiple sources. Sensor fusion is a subset of information fusion and is the merging of data derived from only sensory sources [29]. In general, sensor/information fusion provides better reliability, improved coverage, increased confidence, reduced ambiguity, better resolution and is more robust than using a single data source [29].

Sensor fusion is a vast area of study with no common model or architecture. However, one way to split sensor fusion methods is by the two predominant areas when it can occur, pre-classification or post-classification [30]. Like the name implies, pre-classification occurs before the model is built and post-classification occurs afterwards. Some research refers to pre-classification as low or intermediate-level fusion and post-classification as high or decision-level fusion [29], [30].

The most straightforward method to fuse data is a pre-classification method called data level fusion. In data level fusion, the raw data is combined before training the model. The model is built using data from all fused sources. Another pre-classification method, feature level fusion, uses data association and grouping techniques to improve performance over just the use of the raw data. Post-classification fuses the output or prediction from models that were built using a single data source. A few common methods include voting, ranking, fuzzy logic and other statistical methods. In the case of ADS-B, fusion from other sensor sources could be invaluable since the integrity of ADS-B data cannot be guaranteed.

III. Methodology

The intent behind this research is to fuse multiple aircraft data sources to predict aircraft characteristics. We will make predictions by training a model to determine an aircraft's WTC, description and designator as listed in the ICAO doc 8643. ADS-B, weather and images are selected as the data sources to be combined. Weather and ADS-B are fused during pre-processing, while aircraft images are fused during post-processing. Due to the success in previous research, the MLSTM-FCN model trains the ADS-B and weather models while the CMAL-Net trains the image models [21], [28], [31]. The steps required to fuse the three data sources can be seen in Figure 1.

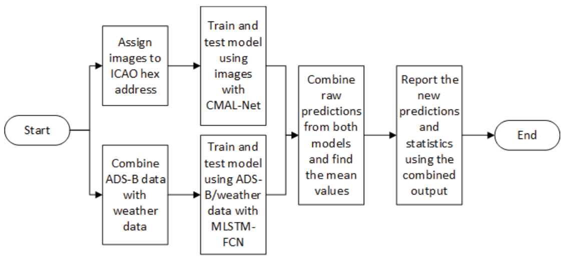

Fig. 1. Flowchart to combine ADS-B, weather and image data

A. Data Preparation

The ADS-B data in this research is obtained from the ADS-B Exchange [1]. The original dataset spans 1 Jan 2021 to 17 Jan 2022 and is 3.7 TB of ADS-B data. However, this research focuses on 1 Oct 2021 to 2 November 2021 with October being used for the training set and November being used for the test set. After obtaining the data, the files are converted from JSON to CSVs, cleaned and broken into 300-time step tensors.

The weather data, obtained from Meteostat [32], is shown in Table I. Calls are made to Meteostat for each ADS-B tensor in order to obtain weather for the timeframe and location of the aircraft during the tensor. The weather data is combined with the ADS-B data and saved as a CSV file.

TABLE I
WEATHER FEATURES USED FROM METEOSTAT

Feature	***Description***	***Type***
temp	The air temperature in °C	Float64
dwpt	The dew point in °C	Float64
rhum	The relative humidity in percent (%)	Float64
prcp	The one hour precipitation total in mm	Float64
snow	The snow depth in mm	Float64
wdir	The average wind direction in degrees (°)	Float64
wspd	The average wind speed in km/h	Float64
pres	The average sea-level air pressure in hPa	Float64

The image data is obtained from the FGVC-Aircraft benchmark [24]. Since there are aircraft included in the FGVC-Aircraft benchmark where the ADS-B data does not have enough samples to train a model, and there are aircraft present in ADS-B data that the FGVC-Aircraft benchmark does not utilize (i.e. the two sets are not equal), the intersection of aircraft classes between the ADS-B data and the FGVC-Aircraft images is found. The similarities among the two sets results in 55 remaining aircraft for the classification model. All other aircraft samples are removed from both datasets. In the training set this produces a total of 325,822,800 observations which translates to 1,086,076 tensors or approximately 452,5322 flight hours. For testing, the intersection yields 67,435 ADS-B/weather tensors which is equivalent to 20,230,500 timesteps or roughly 28,098 hours of flight data. In the FGVC-Benchmark, the remaining images include 3,736 training images and 1,864 testing images.

Since ICAO hex address is not annotated in the FGVC-Aircraft benchmark, each ICAO hex address within the ADS-B data is randomly assigned an image that matches its designator. Due to the number of ICAO hex addresses outnumbering the number of images, each image could be assigned to multiple ICAO hex addresses. This is assumed to be a negligible concern due to the random assignment.

B. Model Development

A total of nine models are created. Three models are trained on ADS-B only, three models are trained on both ADS-B and weather, and three models are trained on the image benchmark. The six models created with the ADS-B and weather data use the MLSTM-FCN algorithm to classify each of the aircraft characteristics of interest (WTC, description and designator). Each ADS-B/weather model is trained for 150 epochs with a dropout of 0.5 and a learning rate of 0.001. These parameters are selected due to success with previous research in this area [21], [22]. Using the test set, precision, recall, accuracy, loss and F1 scores are found. Raw predictions are recorded for each model. Similarly, three models are trained on the FGVC-Aircraft images using the CMAL-Net architecture for 200 epochs. Loss, precision, recall, accuracy, F1 scores and raw predictions are recorded. The list of models can be seen in table II.

TABLE II
MODELS

Model #	***Input Data***	***Algorithm***	***Prediction***
1	ADS-B	MLSTM-FCN	WTC
2	ADS-B	MLSTM-FCN	Description
3	ADS-B	MLSTM-FCN	Designator
4	ADS-B & Weather	MLSTM-FCN	WTC
5	ADS-B & Weather	MLSTM-FCN	Description
6	ADS-B & Weather	MLSTM-FCN	Designator
7	Images	CMAL-Net	WTC
8	Images	CMAL-Net	Description
9	Images	CMAL-Net	Designator

C. Post-processing

Since the results for the fusion between the weather and ADS-B models is found by running the MLSTM-FCN model on the test set, the bulk of post-processing focuses on combining the outputs from the models trained by the MLSTM-FCN and the image models trained by the CMAL-Net. To combine the outputs from the nine models, the raw predictions for each class are normalized per tensor. The predictions from the six ADS-B models and the three image models are combined and the element-wise mean is found for each class per tensor. The class with the largest mean value per tensor is selected as the predicted class for that tensor. The predicted classes are compared to the actual classes for each tensor to determine the precision, recall, accuracy and F1 score.

IV. Results and Discussion

The results from the combination of weather and ADS-B data using the MLSTM-FCN can be seen in Table III. Both predictions on ADS-B alone and predictions with ADS-B and weather are represented in the table. As can be seen, adding the weather data provided improvements with the models that predicted description and designator, but did not improve the WTC model.

TABLE III
ADS-B and Weather Results

Class	***Wx***	***Loss***	***Precision***	***Recall***	***Accuracy***	***F1***
WTC	No	0.370	0.873	0.869	0.870	0.871
	Yes	0.389	0.859	0.853	0.857	0.856
Description	No	0.223	0.953	0.941	0.947	0.947
	Yes	0.233	0.958	0.929	0.949	0.943
Designator	No	1.849	0.619	0.201	0.410	0.303
	Yes	1.748	0.638	0.253	0.448	0.362

The results for the FGVC-Aircraft benchmark image classification can be seen in Table IV. The models performed slightly better than previous research due to the reduction in the number of aircraft samples [24]–[28].

TABLE IV
FGVC-Aircraft Image Results

Class	***Loss***	***Precision***	***Recall***	***Accuracy***	***F1 Score***
WTC	0.090	0.983	0.984	0.983	0.984
Description	0.083	0.989	0.969	0.987	0.978
Designator	0.283	0.954	0.953	0.953	0.953

The fusion between all three data sources can be seen in Table V. In all cases, the image data improved the results compared to the ADS-B and weather data alone. Additionally, adding the ADS-B and weather data to the image data further improved the results for WTC and description when compared to just images alone. With designator, the accuracy was reduced by 3% compared to the image classification by itself. This is likely due to the low accuracy of predicting designator using only ADS-B and weather.

TABLE V
ADS-B with FGVC Aircraft Results

Class	***Wx***	***Precision***	***Recall***	***Accuracy***	***F1 Score***
WTC	No	0.9880	0.9750	0.9898	0.9814
	Yes	0.9847	0.9770	0.9890	0.9808
Description	No	0.9693	0.9113	0.9956	0.9334
	Yes	0.9736	0.9323	0.9960	0.9493
Designator	No	0.8972	0.9435	0.9245	0.9152
	Yes	0.8893	0.9316	0.9201	0.9049

V. Conclusion

In this work we propose an architecture for combining the kinematic data found within ADS-B with weather and aircraft images. Experiments using this method demonstrate when looking at sensor fusion that occurs during pre-processing (i.e combining weather with ADS-B kinematic data) there are negligible improvements when predicting WTC and description, but there is a slight (4%) improvement when predicting designator. When looking at the entire architecture, there is near perfect accuracy when predicting description and 99% accuracy when predicting WTC. This is a 1-2% improvement over using only images. Designator's accuracy is slightly reduced from the accuracy with only images (-5%), but is nearly doubled when making predictions from only ADS-B and weather.

The architecture has two main advantages. First, with only using the kinematic data found within ADS-B, it allows the model to utilize ADS-B as a surrogate for primary 3D radar and potentially eliminate the need for ADS-B altogether. Second, it allows the use of ADS-B in combination with other sources of aircraft data to verify the accuracy of ADS-B. This method could easily be modified to provide more weight to the image results which would further improve the accuracy. It could also allow for predictions to be made without images, but would provide a smaller confidence value on its accuracy.

Acknowledgment

The views expressed are those of the authors and do not reflect the official guidance or position of the United States Government, the Department of Defense or of the United States Air Force. The material was assigned a clearance of CLEARED on 08 JUN 2023. Originator Reference Number: AFIT-23-ENV03100. Number: 88ABW-2023-0601

References

[1] ADS-B Exchange, "Home - serving the flight tracking enthusiast - ads-b exchange," Mar 2023. [Online]. Available: https://www.adsbexchange.com/

[2] United States Department of Transportation, "Current equipage levels," April 2023. [Online]. Available: https://www.faa.gov/air_traffic/technology/equipadsb/installation/current_equipage_levels

[3] A. Costin and A. Francillon, "Ghost in the air(traffic): On insecurity of ads-b protocol and practical attacks on ads-b devices," in *BLACKHAT 2012, July 21-26, 2012, Las Vegas, NV, USA*, EURECOM, Ed., Las Vegas, 2012, © EURECOM. Personal use of this material is permitted. The definitive version of this paper was published in BLACKHAT 2012, July 21-26, 2012, Las Vegas, NV, USA and is available at :.

[4] "Icao document 8643," International Civil Aviation Organization, Quebec, Canada, Standard, 2016.

[5] N. Ruseno, C.-Y. Lin, and S.-C. Chang, "Uas traffic management communications: The legacy of ads-b, new establishment of remote id, or leverage of ads-b-like systems?" *Drones*, vol. 6, no. 3, 2022. [Online]. Available: https://www.mdpi.com/2504-446X/6/3/57

[6] A. Filippone, B. Parkes, N. Bojdo, and T. Kelly, "Prediction of aircraft engine emissions using ads-b flight data," *The Aeronautical Journal*, vol. 125, no. 1288, p. 988–1012, 2021.

[7] F. Peyrin, P. Fréville, N. Montoux, and J.-L. Baray, "Original and low-cost ads-b system to fulfill air traffic safety obligations during high power lidar operation," *Sensors*, vol. 23, no. 6, 2023. [Online]. Available: https://www.mdpi.com/1424-8220/23/6/2899

[8] M. Yue, H. Zheng, H. Cui, and Z. Wu, "Gan-lstm-based ads-b attack detection in the context of air traffic control," *IEEE Internet of Things Journal*, pp. 1–1, 2023.

[9] L. Qian, N. Woods, and M. Rahman, “(U) identifying ads-b flight patterns,” 2019, paper presented at the 2019 MSS Joint (BAMS and NSSDF) Conference, San Diego, CA, 21–24 October 2019.
[10] J. Sun, J. Ellerbroek, and J. Hoekstra, “Large-scale flight phase identification from ads-b data using machine learning methods,” in *7th International Conference on Research in Air Transportation*, 2016, pp. 1–8.
[11] V. Kumar, S. Kalam, A. K. Das, and D. Sinha, “Attack detection scheme using deep learning approach for iot,” *Advanced Computing and Systems for Security*, vol. 14, pp. 17–30, 2021.
[12] J. Sun, J. Ellerbroek, and J. Hoekstra, “Flight extraction and phase identification for large automatic dependent surveillance–broadcast datasets,” *Journal of Aerospace Information Systems*, vol. 14, no. 10, pp. 566–572, 2017.
[13] S. T. Kanneganti, P. B. Chilson, and R. Huck, “Visualization and prediction of aircraft trajectory using ads-b,” in *NAECON 2018-IEEE National Aerospace and Electronics Conference*. IEEE, 2018, pp. 529–532.
[14] F. Hasin, T. H. Munia, N. N. Zumu, and K. A. Taher, “Ads-b based air traffic management system using ethereum blockchain technology,” in *2021 International Conference on Information and Communication Technology for Sustainable Development (ICICT4SD)*, 2021, pp. 346–350.
[15] N. Pearce, K. J. Duncan, and B. Jonas, “Signal discrimination and exploitation of ads-b transmission,” in *SoutheastCon 2021*, 2021, pp. 1–4.
[16] T. Wu, S. Zhang, J. Yang, and P. Lei, “An ads-b signal poisoning method based on generative adversarial network,” *Electronics Letters*, vol. 59, no. 2, p. e12699, 2023. [Online]. Available: https://ietresearch.onlinelibrary.wiley.com/doi/abs/10.1049/ell2.12699
[17] J. Wang, Y. Zou, and J. Ding, “ADS-B spoofing attack detection method based on LSTM,” *EURASIP Journal on Wireless Communications and Networking*, August 2020.
[18] R. Ginoulhac, F. Barbaresco, J. Y. Schneider, J. M. Pannier, and S. Savary, “Target classification based on kinematic data from AIS/ADS-B, using statistical features extraction and boosting,” *Proceedings International Radar Symposium*, vol. 2019-June, no. 681, pp. 1–10, 2019.
[19] K. Basrawi, R. Dill, G. Peterson, B. Borghetti, and J. Lopez, “Aircraft identification from ads-b kinematic data,” *Journal of DoD Research and Engineering*, vol. 5, no. 2, pp. 31–43, 2022.
[20] K. Basrawi, “Aircraft classification from ads-b kinematic data,” Ph.D. dissertation, Air Force Institute of Technology, 2021.
[21] S. Bolton, R. Dill, M. R. Grimaila, and D. Hodson, “Ads-b classification using multivariate long short-term memory–fully convolutional networks and data reduction techniques,” *The Journal of Supercomputing*, vol. 79, no. 2, pp. 2281–2307, 2023.
[22] S. Bolton, R. Dill, M. Grimaila, and D. Hodson, “Aircraft classification using flight phase identification,” 2023, unpublished.
[23] F. Karim, S. Majumdar, H. Darabi, and S. Chen, “LSTM fully convolutional networks for time series classification,” *IEEE Access*, vol. 6, pp. 1662–1669, 2018.
[24] S. Maji, E. Rahtu, J. Kannala, M. B. Blaschko, and A. Vedaldi, “Fine-grained visual classification of aircraft,” *CoRR*, vol. abs/1306.5151, 2013. [Online]. Available: http://arxiv.org/abs/1306.5151
[25] S. Rong, Z. Wang, and J. Wang, “Separated smooth sampling for fine-grained image classification,” *Neurocomputing*, vol. 461, pp. 350–359, 2021.
[26] Y. Wang, Y. Chen, and R. Liu, “Aircraft image recognition network based on hybrid attention mechanism,” *Computational Intelligence and Neuroscience*, vol. 2022, 2022.
[27] L. Wang, K. He, X. Feng, and X. Ma, “Multilayer feature fusion with parallel convolutional block for fine-grained image classification,” *Applied Intelligence*, vol. 52, no. 3, pp. 2872–2883, 2022.
[28] D. Liu, L. Zhao, Y. Wang, and J. Kato, “Learn from each other to classify better: Cross-layer mutual attention learning for fine-grained visual classification,” *Pattern Recognition*, vol. 140, p. 109550, 2023. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0031320323002509
[29] W. Elmenreich, “An introduction to sensor fusion,” *Vienna University of Technology, Austria*, vol. 502, pp. 1–28, 2002.
[30] A. Vakil, J. Liu, P. Zulch, E. Blasch, R. Ewing, and J. Li, “A survey of multimodal sensor fusion for passive rf and eo information integration,” *IEEE Aerospace and Electronic Systems Magazine*, vol. 36, no. 7, pp. 44–61, 2021.
[31] F. Karim, S. Majumdar, H. Darabi, and S. Harford, “Multivariate lstm-fcns for time series classification,” *Neural Networks*, vol. 116, pp. 237–245, 2019. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0893608019301200
[32] C. S. Lamprecht, “Meteostat Python.”