

Evaluation of DME Network for Ensuring RNAV Capability in the European Region

Doctoral Thesis

Study Programme:	Technology in Transportation and Telecommunication
Study Field:	Air Traffic Control and Management
Thesis Supervisor:	doc. Ing. Jakub Hospodka, Ph.D.
Supervisor Specialist:	Ing. Stanislav Pleninger, Ph.D.

In Cooperation with EUROCONTROL

Ing. Tereza Topková

2022



Abstract

The Global Navigation Satellite System has become the primary source of position information for airspace users. Consequently, the role of conventional navigation aids infrastructure is evolving. On one hand, some navaids are being rationalized, and are no longer maintained in operation. On the other hand, Distance Measuring Equipment evolves from a complementary service of VHF navaids into the short-term solution for ensuring the RNAV capability for air traffic when GNSS unavailable. Therefore, the thesis focuses on evaluation of the DME infrastructure and its possibility to rationalize while providing a sufficient RNAV capability for the European air traffic. The proposed approach for the network evaluation is represented by means of a software model of the air traffic environment with relation to the current DME network. The rule-based model aims at constructing an approximation of on-board DME interrogators interacting with DME ground transponders while implemented machine learning algorithm predicts the load of the ground stations based on real measured data. The model is used for evaluating the DME network capability simulating a specific distribution of DME interrogators in the air traffic.

Keywords

Distance measuring equipment, Infrastructure Optimisation, Navaids Rationalization, Area Navigation, GNSS Back-up, Rule-based Model, Gradient Boosting Regression



Declaration

I hereby declare, that I did elaborate and write presented thesis by myself and all used information and scientific resources are in line with methodological instruction for ethical preparation of university theses.

I have no serious motives against using this educational work according to the § 60 of Act of the Czech Republic No. 121/2000, on Copyright and Rights related to Copyright and on Amendment to Certain Acts (the Copyright Act).

Tereza Topková

In Brussels, 5 November 2022



Acknowledgement

I would like to thank my supervisors, doc. Ing. Jakub Hospodka, Ph.D. and Ing. Stanislav Pleninger, Ph.D., for their support and important contributions to the thesis discussions. Special thanks to NAV Team of EUROCONTROL, namely Gerhard Berz and Valeriu Vitan, for directing to this topic, providing data and consultations during the whole process of the thesis compilation. I would also like to express my thanks to doc. Ing. Ivan Nagy, CSc. for his advice on methods for generating increased traffic. Last but not least, I would also like to thank my colleagues from the Department of Air Transport, my family and friends for their support during my PhD studies.



List of Contents

Introd	luction	1
1 Re	esearch Background/Current State of the Art	4
1.1	Navigation Systems	4
1.1.1	Performance-based Navigation	6
1.1.2	Future Evolution of Navigation	9
1.2	Distance Measuring Equipment	
1.2.1	DME Interrogator	
1.2.2	DME Transponder	
1.3	TACAN	
1.4	Alternative Position, Navigation and Timing	19
1.5	DME/DME Positioning	
1.6	RNAV Procedure Service Volume and DME Coverage Criteria	22
1.7	Navaids Infrastructure Rationalisation	
1.8	DME Network	
2 Tł	nesis Objectives and Hypotheses	
2.1	Thesis Objectives	
2.2	Hypotheses	
2.3	Methodology	
3 Re	esearch Methods	
3.1	Object Oriented Programming	
3.2	Regression	
3.2.1	Gradient Boosting Regression	
3.2.2	Regression Validation Metrics	
3.3	Data Classification	
3.3.1	Support Vector Machines	
3.4	DBSCAN Clustering	40
3.5	Generating Samples	41
3.5.1	Mixture Clustering	41
3.5.2	Inverse Transform Method	43
3.6	Monte Carlo Simulation	
4 D	ata Collection	45
4.1	Air Traffic Data	45
4.2	DME Ground Stations	
4.3	DME Interrogator	



5 1	Model Construction	
5.1	Initial Data Processing	53
5.1.1	Initial Processing of Air Traffic Data	54
5.1.2	Initial Processing of DME Ground Stations Data	58
5.1.3	Initial Processing of DME Interrogators Data	61
5.2	Classes Definition	61
5.2.1	Aircraft Class	62
5.2.2	GroundStation Class	63
5.2.3	FMS Class	63
5.2.4	Environment Class	64
5.3	DME Interrogator Assignment	65
5.4	Finding Suitable DME Ground Stations	67
5.5	Finding Suitable DME/DME Pair	69
5.6	Assignment of DME Ground Station to Aircraft	71
5.7	Assignment of Aircraft to DME Ground Station	
5.8	Machine Learning Application	73
5.8.1	Machine Learning Model Selection and Training	74
5.8.2	Direct Application of GBR	76
5.9	Model Retuning	77
5.10	Model Output	79
6 1	Model Validation	82
6.1	Model Validation Using Train Test Split	82
6.2	Model Validation Using Other Stations for Model Training	83
7 [Design of Testing Scenarios	86
7.1	Simulation of Aircraft equipment	
7.2	Optimization of DME infrastructure	
7.3	Generating Air Traffic Data	90
8 H	Hypotheses Testing Results	94
8.1	First Hypothesis Testing	95
8.2	Second Hypothesis Testing	
8.3	Third Hypothesis Testing	97
8.4	Hypotheses Conclusions	
9 F	Research Outcome Summary	101
Conc	lusions	
Refe	rences	
Attac	chment A – Model Validation Results	113



List of Figures

Figure 1 - Navigation systems overview [17]	5
Figure 2 - Difference between conventional, RNAV and RNP routes [21, edited]	6
Figure 3 - PBN specifications [19]	
Figure 4 - Plan of CNS Transformation [10]	10
Figure 5 - Deployment scenario for navigation systems [10]	11
Figure 6 - NDB and VOR decommission ambitions [27]	
Figure 7 - DME principle [30]	13
Figure 8 – Characteristics of DME interrogation pulse pair [31]	14
Figure 9 - Block Dlagram of DME [31]	15
Figure 10 - DME/DME RNAV principle [38]	21
Figure 11 - Histogram on DME Slant Range Error Distribution [9]	23
Figure 12 - Rationalised All ECAC DME/DME Performance FL95 [43]	25
Figure 13 - DME density map in FL 200	
Figure 14 - DME/DME model basic architecture	
Figure 15 - Chess Pieces Inheritance [46]	34
Figure 16 - SVM solution of a linear problem [60, edited]	
Figure 17 - Example of clusters each representing one mode of the system [64]	42
Figure 18 - ADS-B version 2 equipage [71]	
Figure 19 - Model Workflow	52
Figure 20 - Necessary EIRP to achieve the minimum power density as a function of	
height and distance from DME [26]	60
Figure 21 – Distribution of Flights Categories in 2019 [84]	67
Figure 22 - List of slant ranges, pp/s, aircraft objects assigned to VOR/DME AFI in	
Belgium on 2022-01-15 at 16:00 UTC using the rule-based model only	73
Figure 23 – Example of model output visualization	81
Figure 24 - The validation results include charts of comparison of the results of the ru	le-
based only model and the results of the model when ML applied with the real data on	۱
100 samples, histogram of the results, dependency of the result values, and results of	f
the used matrics for DME5	84
Figure 25 - Comparison of 100 samples of DME3 validation results with testing based	on
other ground stations (left), and with testing using train test split	
Figure 26 - Example of randomly removed DME stations (darker blue)	87
Figure 27 - Estimation of Eps parameter for DBSCAN clustering	88
Figure 28 - Clusters of DME ground stations	88
Figure 29 - Stations for decommission after DBSCAN clustering (darker blue)	89
Figure 30 - IFR movements forecast 2022-2024 [96]	
Figure 31 - IFR Movements Forecast from October 2018 [97]	91
Figure 32 - Evolution of latitude and longitude parameters of individual mixture	
components	92
Figure 33 - Histograms of original positions of aircraft (blue) and newly generated	
sample (orange) with traffic increase of 20% for the second cluster on Friday, June 28,	,
2019, 11:30 CET	93



List of Tables

Table 1 - NAVAIDS and Navigation Specification [22]	8
Table 2 – Pulse pair spacing and time delay for DME/N channels [23]	. 13
Table 3 - DME/DME RNAV FOM [39]	
Table 4 - Number of potential rationalised navaids [43]	. 24
Table 5 - Data in State Vector Table [68]	
Table 6 - ICAO Doc 8634 aircraft description [69]	
Table 7 - ICAO Doc 8634 aircraft engine description [69]	. 46
Table 8 - Parameters of DME ground station data [72]	
Table 9 - Parameters of DME ground station data [73]	. 49
Table 10 - Parameters of DME interrogator types	. 50
Table 11 - Sample of OpenSky Air Traffic Data	
Table 12 - Results of performance evaluation of tested ML models on prediction of A	С
description	
Table 13 – Data Frame of Air Traffic Data	
Table 14 – Sample of Data Frame of DME Ground Stations	
Table 15 - Results of evaluation of different distribution of DME interrogator types	. 66
Table 16 - Comparison of different ML algorithms using train/test split 70/30	. 74
Table 17 - Comparison of results using RF and GBR when trained on others	
Table 18 - Parameters of DME ground stations used for GBR training	
Table 19 - Sample of features for ML model	
Table 20 - Sample of output file with the aircraft list	
Table 21- Sample of output file with the DME ground stations list	. 80
Table 22 - Comparison of performance results for the rule-based and ML model with	
real data using train and test split	
Table 23 - Comparison of performance results for the rule-based and ML model train	ed
on others	. 84
Table 24 - Simulation results with default DME interrogator distribution, no traffic	
growth, and no removal of ground stations	
Table 25 Simulation results with 70% of aircraft equipped with dual DME interrogat	
"B", no traffic growth, and no removal of ground stations	
Table 26 – Simulation results with no traffic growth, and removal of ground stations	-
10%, 20% and 30%	
Table 27 - Simulation results with traffic growth by 10%, 20% and 30%, and removal	
ground stations by 10%, 20% and 30%	. 98



Abbreviations

4D	Four-dimensional Space
ADS-B	Automatic Dependent Surveillance-Broadcast
ANS	Air Navigation Service
ANSP	Air Navigation Service Providers
APCH	Approach
API	Application Programming Interface
APNT	Alternative Position Navigation and Timing
AR	Authorisation Required
ARB	Auxiliary Reference Burst
ATC	Air Traffic Control
ATM	Air Traffic Management
C2 link	Command and Control Datalink
CET	Central European Time
CNS	Communication, Navigation and Surveillance
DBSCAN	Density-Based Spatial Clustering of Application with Noise
DEMETER	Distance Measuring Equipment Tracer
DFMC	Dual-Frequency Multi-Constellation
DME	Distance Measuring Equipment
DOC	Designated Operational Coverage
Doc	Document
ECAC	The European Civil Aviation Conference
ECEF	Earth-Centred, Earth-Fixed
ED	EUROCAE Document
EECNS	Essential and Efficient Communication Navigation and Surveillance Integrated System
EIRP	Effective Isotropic Radiated Power
eLORAN	Enhanced Long Range Navigation
ENU	East, North, Up
EU	The European Union
EUROCAE	The European Organisation for Civil Aviation Equipment
EUROCONTROL	The European Organisation for the Safety of Air Navigation
FAA	The Federal Aviation Administration
FL	Flight Level
FMS	Flight Management System
FOM	Figure of Merit
FPL	Flight Plan
FTE	Flight Technical Error
GBR	Gradient Boosting Regression
GBR	Gradient Boosting Regression
GNSS	Global Navigation Satellite System
GPS	The Global Positioning System
ICAO	The International Civil Aviation Organization
iCNS	Integrated Communication, Navigation and Surveillance



ILS	Instrument Landing System
INS	Inertial Navigation System
IRS	Inertial Reference System
IRU	Inertial Reference Unit
ISA	International Standard Atmosphere
KNN	K-Nearest Neighbour
LDACS	L-band Digital Aeronautical Communications System
LOS	Line-of-Sight
MAE	Mean Absolute Error
MCS	Monte Carlo Simulation
ML	Machine Learning
MLS	Microwave Landing System
MON	Minimum Operational Network
MRB	Main Reference Burst
MSE	Mean Square Error
NASA	The National Aeronautics and Space Administration
NAV	Navigation
Navaid	Navigation aid
NDB	Non-directional Beacon
NextGen	Next Generation Air Transportation System
NSE	Navigation System Error
OOP	Object-oriented Programming
PBN	Performance-based Navigation
PDE	Path Definition Error
pp/s	pulse pairs per second
PRF	Pulse Repetition Frequency
RAIM	Receiver Autonomous Integrity Monitoring
RFI	Radio Frequency Interference
RMSE	Root Mean Square Error
RNAV	Area Navigation
RNP	Required Navigation Performance
RPAS	Remotely Piloted Aircraft System
SAFIRE	Spectrum and Frequency Information Resource
SARPS	Standards And Recommended Practices
SBAS	Satellite-based Augmentation System
SDR	Software-defined Radio
SESAR JU	The Single European Sky ATM Research Joint Undertaking
SIS	Signal in Space
SVM	Support Vector Machines
TACAN	Tactical Air Navigation
ТМА	Terminal Control Area
TSE	Total System Error
UHF	Ultra High Frequency
US	The United States
VHF	Very High Frequency
VOR	VHF Omnidirectional Range



VORTACVHF Omnidirectional Range Tactical Air navigationWGS84The World Geodetic System 1984WTCWake Turbulence Category



Introduction

The Global Navigation Satellite System (GNSS) has become the main source for providing position and timing information independently of terrestrial radio navigation aids (navaids). In the Communication, Navigation and Surveillance (CNS) domain in aviation, the usage of GNSS enables evolution of the Performance-based Navigation (PBN) concept with the ability to serve areas with high traffic density efficiently. Consequently, airspace users are encouraged to use procedures based on GNSS in all flight phases. The Commission Implementing Regulation (EU) 2018/1048 [1] encourages Air Traffic Management/Air Navigation Service (ATM/ANS) providers to use satellite-based navigation. [1,2]

The transmitted satellite signal is influenced by atmospheric attenuation on its way to the Earth, and it is received in the form of a weak signal. Therefore, it is susceptible to radio frequency interference (RFI). An occurrence of harmful GNSS RFI which may affect ATM operation has been reported. The number of similar incidents increases every year, not considering the reduced traffic caused by the COVID-19 pandemic. One of the most known cases is when the source of the interference represented by a jammer is placed in a vehicle moving on a highway or parking around an airport to avoid paying for the satellite-based toll. Due to the interference of such a jammer, the GNSS procedures cannot be used for PBN approach procedures, or various airport procedures may be negatively affected. Furthermore, there may be intentionally transmitted jamming or even spoofing signals for disabling the position determination of military vehicles in war zones. In case the alternative navigation source providing necessary performance is not available, e.g. in oceanic airspace, the concerned flights need to be vectored by Air Traffic Control (ATC) service relying on information from surveillance and communication systems. The GNSS outage may also be caused by a system error or space weather phenomena. In any case, the affected part of airspace may be quite extensive. [2,3,4,5]

It is necessary to ensure the resiliency of CNS systems, in order to sustain the safety level and efficiency of the ATM system when GNSS is unusable. The aviation stakeholders are fully aware of the urgent need to find measures to avoid any undesired impact which could be caused by harmful interference of the satellite-based navigation. This question was addressed at the 12th ICAO Air Navigation Conference, and it was also presented at the 40th ICAO Assembly, proposing individual measures. The measures included the maintenance of an independent minimum operational network (MON), ensuring the protection of the GNSS frequency spectrum by coordination with radio-regulatory,



protection of the safety of civil aircraft during military exercises and operations, and supporting the development of alternative positioning, navigation and timing (APNT) solutions. [4,6]

The individual GNSS components and the receivers are being improved in order to strengthen the system resilience. One of the means how to mitigate the GNSS RFI threat is using dual-frequency multi-constellation (DFMC) GNSS receivers. An amendment of Annex 10, Volume I of the Chicago Convention adding provisions of the DFMC receivers was introduced to ICAO Member states with applicability in 2023 [7]. Besides that, different Alternative Positioning Navigation and Timing (APNT) concepts are being developed to ensure adequate navigation performance to the air traffic as a GNSS back-up. At this moment, none of the systems has been chosen to be the standard for worldwide usage by ICAO. However, the standardization process of LDACS (L-band Digital Aeronautical Communications System) started years ago, and its navigation function might serve the air traffic as a GNSS back-up in the future. [3,7,8,9]

Even though the GNSS is the main mean for determining position of aircraft and it plays an important role in the time synchronisation used in other CNS systems, terrestrial navaids cannot be fully decommissioned and substituted by satellite-based navigation. In order to maintain the resiliency and ensure the safety of air operations, navaids have to support aircraft navigation in case of GNSS outage. Nevertheless, the current infrastructure of ground-based systems can be optimised to go in line with the PBN requirements. The current CNS network offers space for rationalisation and decommissioning of some ground beacons and maintaining just the MON. The shortterm solution for GNSS back-up is considered DME/DME navigation which can ensure required Area Navigation (RNAV) capability, complemented by VOR/DME where necessary. In the long-term, the APNT system could provide better performance than the current DME/DME and fits into the concept of integrated CNS (iCNS). [2,10]

The rationalisation of CNS infrastructure represents a significant part in the evolution of the European Sky. Not only Europe is focused on keeping a back-up system to mitigate for the loss of GNSS, but the FAA NextGen project also includes resilient navigation infrastructure where DME, VOR and TACAN (Tactical Air Navigation System) network are planned to be optimised and play a significant role during GNSS outage. In the case of the FAA, the installation of new DMEs is considered necessary. [10,11,12]

This research is focused on the rationalisation of the DME/DME infrastructure. In particular, it deals with a development of a software model of the DME/DME environment



to evaluate the DME network capability in the European airspace represented by load determination of individual ground stations.

The current role of DME in the CNS system and its importance in relation to the PBN concept is explained in the first part of the thesis. It is followed by the DME principle and technical description. Subsequently, the research background and existing approach is presented. Based on the current state-of-art, the main objective of the thesis and hypotheses are established. The research workflow is specified in the methodology, complemented by a description of selected methods for achieving the objectives. Afterward, the analysis of aircraft equipment capabilities creates one of the important inputs into the model. Other data sources are further explained, and parts of the model, as well as its functionality, are described in detail. Consequently, the model results are validated by comparison with the real data. Based on the established hypotheses, the testing scenarios are created, and the hypotheses are confirmed or rejected. Finally, the results are evaluated in the research summary, and the contribution of the thesis highlighted in the conclusion.



1 Research Background/Current State of the Art

This research is conducted in the ATM/CNS area. The ATM is defined as "The dynamic, integrated management of air traffic and airspace including air traffic services, airspace management and air traffic flow management — safely, economically and efficiently — through the provision of facilities and seamless services in collaboration with all parties and involving airborne and ground-based functions" [13]. In other words, the ATM includes all components that enable to operate air transport, and it connects airports with flights through the airspace. The role of the CNS infrastructure is to ensure required operational performance for air traffic and to enable necessary airspace capacity. Therefore, CNS creates a crucial part of the air transport. The essential functions of the CNS network are represented by aeronautical information exchange via voice or digital data communication systems, determination of the position of the aircraft, on-board for the flight planning in the airspace, and on the ground to provide air traffic control services. [14]

The main focus of this work is on the navigation component of the CNS systems, particularly on the DME and its role in the context of the evolving navigation environment in aviation. As a reason of the development of satellite-based navigation, the PBN concept has become a standard, and the need of operation has decreased concerning some conventional navaids. Besides the description of the DME principle, the following subchapters deal with the overview of navigation systems, the change that came with the PBN concept, the future plans in the ATM/CNS area, and the importance of maintaining the APNT navigation. In addition, research conducted within a SESAR JU project is presented to show the current state-of-art of the rationalisation of CNS infrastructure.

1.1 Navigation Systems

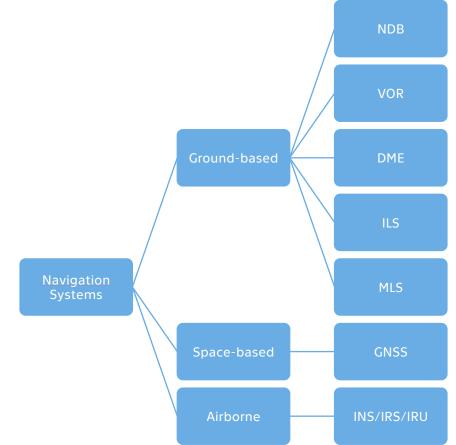
The main function of navigation systems is to determine the aircraft's position, in order to enable planning a flight route from a starting point to a target point through the airspace. Before the invention of the radio navigation systems, pilots could fly only using the visual navigation relaying on maps and dead-reckoning¹. The development of navigation aids increased the possibility of getting better orientated in the airspace up to using precise procedures enabling landing with zero visibility. The first terrestrial

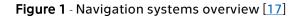
¹ Dead-reckoning is a method based on estimating of current track, ground speed and position based on the previous known positions. [15]



navigation systems started their operation in the World War II, such as ILS. They are usually based on simple technical principles using properties of electromagnetic waves, and thanks to that, they can provide necessary robustness and have been used with some technological enhancement until now. [15,16]

The position of aircraft can be determined following different methods, indicating the direction to a ground beacon, measuring the distance from the navigation system, their combination, or independently, the aircraft position can be determined measuring the changes in the movement of the aircraft with gyros and accelerometers. Besides that, the aeronautical navigation also included systems using the hyperbolic method measuring the time difference of arrival from several stations. GNSS replaced the last mentioned systems in aviation. The overview of the navigation system is specified in Figure 1.





The en-route position is usually not determined just from one standalone navigation system. However, the Flight Management System (FMS) is able to combine the position calculation from multiple sensors to provide the most accurate solution in order to ensure advanced flight planning and guidance systems in four-dimensions space (4D). That is to say, FMS and the airborne radio signal multi-mode receivers are mostly



doubled to have two independent inputs of the radio signal and its evaluation. Based on the availability of the sensors, the FMS may automatically select the navigation mode in the predefined order:

- GNSS
- DME/DME
- VOR/DME

The majority of airliners are equipped with an inertial system (INS/IRS/IRU) which allows integration with other navigation sensors. In that case, the INS can be used as the only source of navigation information when no other data is available. Moreover, different sensors may be used for correcting heading/altitude data. The airborne system capabilities that have to be provided to fly a specific route or in specific airspace are represented by performance requirements. This concept is known as Performance-based Navigation (PBN). The navigation specification is no longer sensor-based. However, the particular specification identifies which on-board equipment can be used to meet the required performance of the navigation systems. [17,18,19,20]

1.1.1 Performance-based Navigation

Deployment of PBN procedures and routes has brought airspace users a possibility of free navigation without constraints about the location of terrestrial navaids. The PBN concept enables to increase the airspace capacity and to optimise its utilization. In addition, it improves safety and efficiency in airspace and procedures design, as shown in Figure 2. The performance requirements applied to aircraft as well as aircrew are represented by two navigation specifications; RNAV and RNP. [19]

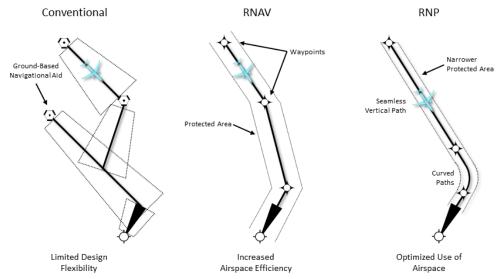
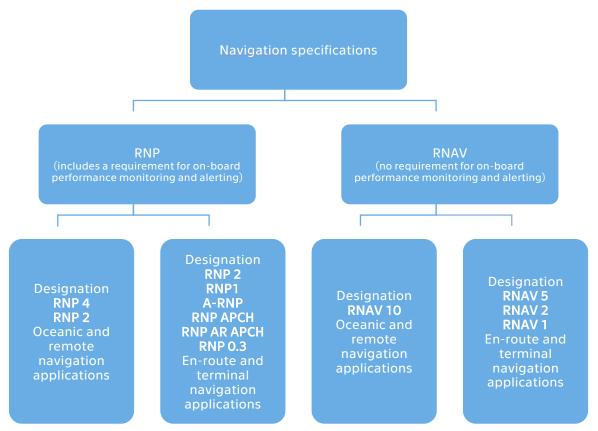


Figure 2 - Difference between conventional, RNAV and RNP routes [21, edited]



The difference between these two specifications is the inclusion of a requirement for onboard performance monitoring and alerting system defined for RNP specification. The other performance requirements are specified in terms of accuracy, integrity, continuity, and functionality needed for the particular operation. In addition, some navigation infrastructure also requires availability to enable navigation applications, such as GNSS SIS. Figure 3 represents the overview of the currently defined navigation specifications. They are divided based on their usage in oceanic or remote areas or in en-route and terminal navigation applications.





Even though the navigation performance is defined by requirements on the abovementioned parameters, it can be met by providing navigation function from one or more sensors. The purpose of the FMS is to choose an adequate navigation system or its combination to determine the best position solution in the required performance. The connection of the possible used navaids and navigation specification that can be achieved by their usage is represented in Table 1.



	NAVAID				
NAV SPEC	GNSS	IRU	DME/ DME	DME/ DME/ IRU	VOR/ DME
RNAV 10	\checkmark	\checkmark			
RNAV 5	\checkmark	\checkmark	\checkmark		\checkmark
RNAV 2 & 1	\checkmark		\checkmark	\checkmark	
RNP 4	\checkmark				
RNP 2	\checkmark		\checkmark	\checkmark	
RNP 1	\checkmark		\checkmark	\checkmark	
Advanced RNP	\checkmark		\checkmark	\checkmark	
RNP APCH APV Baro	\checkmark				
RNP APCH APV SBAS	√ With SBAS				
RNP AR APCH	\checkmark				
RNP 0.3	\checkmark				

Table 1 - NAVAIDS and Navigation Specification [22]

Note: DME/DME and DME/DME/IRU sensors used for RNP are subject to ANSP requirements and aircraft capability.

The SESAR JU research project Essential and Efficient Communication Navigation and Surveillance Integrated System (EECNS) [18] provided industrial research that included, among other things, an investigation of possible DME/DME support of RNP 1 reversion without modification of avionics. This solution assumed the sufficient integrity provided by the ground infrastructure and appropriate training of the flight crew, and current equipment capabilities of the ground station than in the current MOPS for DME Ground Equipment, EUROCAE ED-57 [23]. In 2020, EUROCONTROL issued the European GNSS Contingency/Reversion Handbook for PBN Operations [24], where the possible operational infrastructure is introduced when GNSS is considered unusable.

Taken into account the evolving role of GNSS and its position as the primary source of navigation information, the importance of some terrestrial navaids decreased. On the contrary, there is still a need to retain back-up navigation systems in operation regarding the GNSS vulnerability.



1.1.2 Future Evolution of Navigation

The priorities of the European ATM development and deployment are contained in the European ATM Master Plan [10]. The document is regularly updated, including contributions from SESAR JU research projects represented by all ATM stakeholders.

The CNS infrastructure is usually operated by the national Air Navigation Service Providers (ANSP) in Europe. Therefore, the network corresponds to the need of individual states but does not take into account the cross-border navaids. This led to inefficient mutual locations of the terrestrial navigation equipment. A similar situation arose from limited cooperation between civil and military aviation stakeholders, and both sides have been operating systems with substitutable functionality. The optimisation of the CNS network would not only reduce the operational costs of CNS, but it could also help to improve efficiency in the use of the radio spectrum. Besides the rationalisation of the already existing systems, the ATM Master Plan introduced a future CNS system relying on new digital integrated CNS solutions, together with GNSS and ADS-B. In addition, these systems are complemented by the minimum operational network of the terrestrial navaids, such as DME and ILS. The plan of the CNS systems transformation is shown in Figure 4. The important idea of the CNS future infrastructure highlights the orientation on the delivered services. In other words, it enables to operate the infrastructure not as separate physical navaids, but as one whole and integrated system with the required flexibility and capability to serve the future airspace needs. It can be supported by enhancement of the civil-military cooperation, mutual interoperability of the systems and willingness to share the data, with assurance of the data sharing, network protection and its resilience to cyber attacks. [10]

From the deployment point of view, the navigation systems optimisation enables rationalizing the network of conventional navaids, such as NDB, VOR and DME, with respect to the PBN concept. The ground-based systems should provide a suitable reversion capability for GNSS and support contingency function when GNSS is unavailable. The evolution of deployment scenario also includes future terrestrial ground-based technologies with the potential to substitute the current ones in the long-



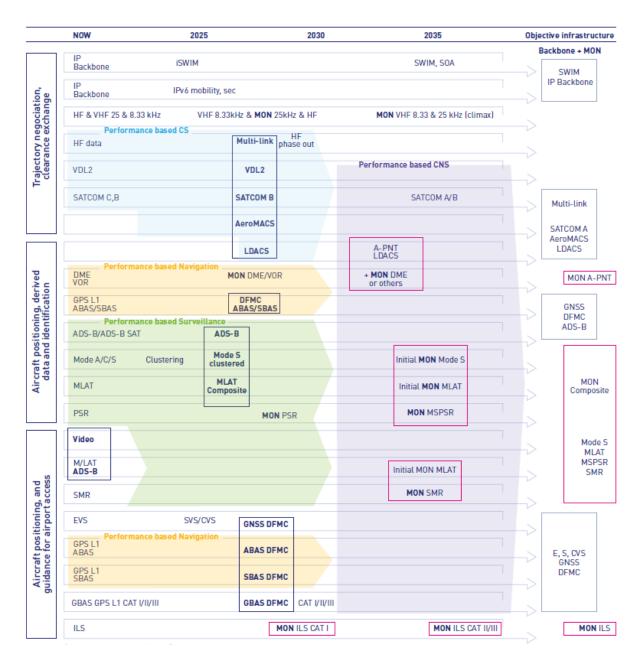
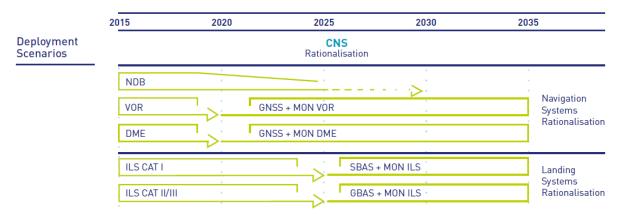


Figure 4 - Plan of CNS Transformation [10]

term. The network optimisation will support high performance and efficiency, taking into account also operational costs and optimal utilization of radio spectrum. The performance requirements allow service providers to choose the CNS technologies according to their needs considering local specificities. The airspace users have the option in adapting the airborne equipment of aircraft likewise. The rationalisation of the ground-based infrastructure is represented in Figure 5. [10,24]







The ATM Master Plan proposes to decommission some of the terrestrial navaids, nevertheless, the solution always meets the high level of safety and security, in particular cybersecurity, and resilience of the CNS network. The decommission of NDB raises from its unsuitable role in the PBN concept. In an exceptional case, an NDB can be used for missed approach procedures. The role of VOR is also limited in the PBN concept because of the supported specification of RNAV 5 used for en-route applications. However, existing VOR ground stations can be maintained in order to support some reverse operation when DME/DME is not available, in order to enhance situational awareness, etc. DME/DME is considered the most suitable solution for ensuring sufficient PBN capability in the short-term, as further described in the following subchapters. At the same time, the GNSS is supposed to increase its resiliency using DFMC in all phases of flight. [10,25]

The European plan for CNS network rationalisation in terms of PBN does not create an exception amongst the global strategies. The strategy is presented in an attachment of the ICAO Annex 10/Volume I [26], and the FAA also follows a similar evolution plan of CNS

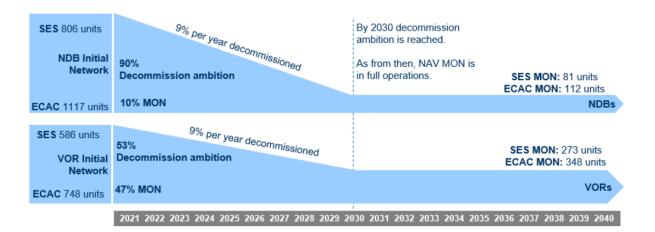


Figure 6 - NDB and VOR decommission ambitions [27]



infrastructure [12]. The basics of the deployment of the MON are based on the assessment of the current infrastructure. As represented in numbers, the navigation scenario contains an ambition to decommission 90 % of the current NDB network and 53 % of the current VOR network. There is no assumption that these navaids will be replaced by another type of navaid, see Figure 6. [27]

As resulted from the abovementioned, DME playes a significant role in the short-term horizon providing the necessary navigation performance to serve the air traffic as a GNSS back-up. Therefore, the attention of this research is devoted to DME/DME and its capability to provide PBN services.

1.2 Distance Measuring Equipment

Distance measuring equipment represents a system for slant ranging between an aircraft and a ground station. It was standardized by ICAO in 1952 and the operational standard of DME was last updated with development of the DME/P version in 1986. Since then, DME has become an essential part of air navigation, and with collocation with other systems such as VOR, ILS, or MLS, it can serve the air traffic in en-route, approach, landing, missed approach, and also departure phases of flight. The technology has been further improved, and the latest generation of DME transponders can overcome the level of the outdated operational standards and provide better navigation performance in line with the RNP specifications. A EUROCAE (The European Organisation for Civil Aviation Equipment) working group number 107 (WG-107) has been working on a revision of existing minimum operational performance specifications for the DME ground equipment and on development of a specification for avionics systems for RNP reversion using DME/DME positioning [16,18,28].

DME consists of two elements; a DME interrogator fitted on-board of an aircraft and a DME transponder located on the ground. DME is based on a simple principle of measuring the elapsed time from the moment of an interrogation transmission from the airborne interrogator to the reception of reply to this interrogation from the ground transponder. When the time is measured, a distance representing a slant range between aircraft and the ground facility can be determined as follows:

$$R = \left(\frac{\tau - d}{2}\right) \cdot c; \tag{2.1}$$

Where *R* is the slant range, τ is the measured time between sending the interrogation and receiving the reply, *d* is a predetermined delay in the ground transponder, and *c* is



the speed of radio wave propagation. The measurement of the slant range is shown in Figure 7. [29,30]

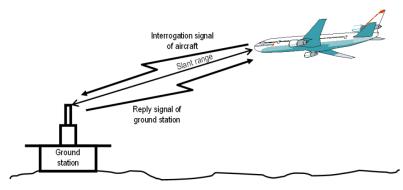


Figure 7 - DME principle [<u>30</u>]

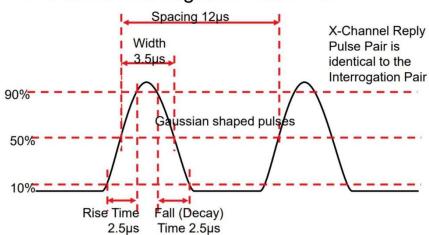
DME can be operated in two standards. The DME primarily used for en-route or TMA navigation is referred to as DME/N; the abbreviation "N" has a meaning of a narrow spectrum characteristic. The "precise" distance measurement represents the distance measuring element of MLS and it has two operational modes; a final approach mode for supporting flight operation in the final approach and runway areas and initial approach mode, which is interoperable with DME/N. For this research purpose, the further description deals with standards of the DME/N, the precision ranging element of MLS is specified in the EUROCAE documents ED-36 and ED-53. [23,26]

DME works in the UHF frequency band 960-1215 MHz which falls into the radio spectrum known as the L-band. The channel separation is 1 MHz, and the channels are divided into channels X and Y for DME/N; and channels W and Z intended for DME/P. The particular channels are distinguished by the pulse pair spacing; as shown in Table 2. DME/N system can be operated on one of the 252 channels; 126 X channels and 126 Y channels. The ground transponders have also defined different delay between the interrogation reception and their reply transmission. The reply on interrogation is transmitted with a difference of \pm 63 MHz from the interrogation frequency. [23,26]

DME/N Channel	Pulse Pair Spacing [μs]		Time delay
	Interrogation Reply		[µs]
x	12	12	50
Y	36	30	56

Table 2 – Pulse pair spacing and time delay for DME/N channels [2	3
---	---





X-Channel Interrogation Pulse Pair

Figure 8 – Characteristics of DME interrogation pulse pair [31]

The DME signal consists of a pulse pair. Each pulse has a smoothly rounded Gaussian shape, as represented in Figure 8. The pulse duration should take 3.5 μ s between the leading and trailing edge of the pulse at 50% of the maximum amplitude. The pulse rise from 10% to 90% of the maximum amplitude takes between 1.5 and 3 μ s. Similarly, the pulse decay time should reach between 1.5 and 3.5 μ s. The instantaneous amplitude of the pulse between points of 95% of maximum amplitude leading and trailing edge of the pulse should not fall below the 95% value. The pulse pair spectrum is within ±100 kHz of the assigned channel frequency. At least 90% of the radiated energy in each pulse shall be within 0.5 MHz. [23,29]

The block schema in Figure 9 further characterised the interrogator and receiver elements of the DME system. Both parts are consisted of an omnidirectional antenna with vertical polarization. The overall accuracy of the DME/N system is defined as of ± 0.1 NM. However, research showed that achieved accuracy in operation can reach better results. [23,30]

DME is often used to complete the VOR system providing measurement of bearing from or to a VOR ground station relative to magnetic north, also-called radial. Therefore, VOR/DME is capable of providing full information about the position of an aircraft in airspace relative to the ground station. The DME is preferably installed at the exact location as VOR. DME may be also associated with an ILS system in order to provide a source of the distance measuring function during landing as a substitute for marker beacons. The DME/P ensures a similar function for MLS. DME facilities can also be operated as a single facility without association with any systems. En-route ground stations are able to serve an area in a range of 200 NM up to altitudes until the path loss

14



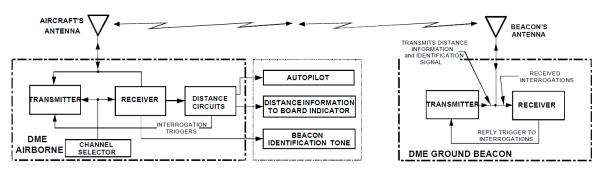


Figure 9 - Block Dlagram of DME [31]

causes an decrease of the performance according to the standard [23]. The terminal installation associated with ILS supports instrument approach procedures from a range of 25 NM in altitudes down to 1000 ft above the runway. The collocation of DME with other VHF systems or MLS has a specified frequency allocation in Annex 10/Volume I [26].

The distance measuring section of TACAN operates in the same frequency band, and it is based on the same principles as DME. In order to provide bearing measurement for civil aviation, TACAN can be associated with VOR as well; referred to as VORTAC. This technology is shortly described in Chapter 1.3.

1.2.1 DME Interrogator

The DME interrogator is installed on-board of an aircraft. It has a function of transmitting interrogation in the form of pulse pairs, recognising the reply for its interrogation, in order to calculate the slant range to the interrogated ground stations. The interrogator can scan simultaneously more than one DME transponder operated on different channels [29].

The interrogator can work in different modes. The search mode is intended to identify the ground station reply of own interrogations and set up the connection to the ground station. The track mode ensures the distance measurement. In case the interrogator loses connection with the ground station, the memory mode of the airborne equipment ensures retaining of the distance indication when it does not receive a suitable reply signal for a certain time period (usually 10 s) instead of switching immediately to the search mode. Provided that the signal is lost, the interrogator switches to the search mode. Otherwise, the indicated distance is acquisitioned again, and the interrogator continues interrogating the same signal in track mode. Within the memory mode period, the estimation of the slant range to the ground station is presented to the pilot. [29,30]

The average pulse repetition frequency (PRF) of the airborne interrogator, represented by the number of transmitted pulse pairs per second, is considered 16 pp/s at maximum



for each interrogated frequency, on the assumption that 95% of the occupation time of the interrogator is in track mode. In other words, 5% remains for the search mode. The maximum interrogation rate in the search mode can reach up to 150 pp/s. The airborne DME should acquire the correct range within 2 s in 80% and switch to the track mode. Provided that the interrogator has a 100 W or less peak pulse power, the average PRF can reach up to 30 pp/s. When the airborne DME interrogates more channels, the PRF for all frequencies should be lower than 48 pp/s. [29]

In order to recognize replies of own interrogations of an aircraft from other replies and squitter pulses from the ground station, the pulse repetition rate is randomly changed. This intentional change is called "jitter". The interrogator expects to receive the pulse pairs replies from the ground station with the same repetition frequency, and it can predict the time frame when the next reply should be received. While the DME interrogator identifies its own responses from the ground station, it locks on the ground station, and it interrogates in the track mode. The scanning DME can receive more than one reply on the same channel. Under those circumstances, the airborne equipment should track the station with the stronger signal when the condition is met that one signal is at least 8 dB or greater in amplitude than the next stronger signal. [26,30]

The aircraft equipment should correspond to the peak power of the second pulse at least 250 W when operating above 18,000 ft. Otherwise, the second pulse peak should be interrogated with a power of at least 50 W. The highest power should not exceed 2 kW. The first pulse should be transmitted with power of a maximum difference 1 dB from the second one. The receiver sensitivity level of the DME interrogator is set at maximum of -83 dBm. Accuracy of the DME/N interrogator is defined with 95% the total error of airborne equipment does not exceed \pm 0.17 NM or 0.25% of the calculated distance; whichever is greater. [26,29]

1.2.2 DME Transponder

Besides the replies on the aircraft interrogations, the DME transponder pseudo-randomly transmits squitter pulse pairs independently of whether or not it is interrogated by an aircraft. Therefore, the typical transmission rate reaches values about 800 pp/s; but at least 700 pp/s. If an aircraft interrogation is received by the ground station, a part of the squitter pulses is replaced by the pulse pairs replies. The DME ground station is capable of replying at a transmission rate of at least 2700 pp/s. The high-capacity transponders are able to reply up to 4600 pp/s. These values should correspond to the maximum



number of serving 100 aircraft simultaneously, respectively 200 for the high-capacity ground station. In other words, the replies capability of the transponder should be adequate to the peak traffic in the coverage of the ground station. Once this limit is achieved, further interrogations cannot be replied to, and the DME station is considered saturated. Furthermore, the additionally received interrogations force the ground station to reduce its sensitivity to maintain the maximum PRF. For his reason, the ground station replies only to aircraft located closer to the station, and it ignores interrogations from more distant aircraft. [23,26]

The coverage of the ground transponder is dependent on including both air and ground transmitter powers, receiver sensitivities, antenna characteristics, aircraft altitude and ground station antenna siting. For example, the omni-directional antenna of the transponder can achieve a gain of 9 dB of the main beam with an elevation of the $3 \pm 1^{\circ}$. Thus, the radiation pattern of the antenna in the vertical plane enables to receive the transponder signal also below the horizon. [23,32]

The interrogation pulse pairs with the correct spacing and nominal frequency shall trigger the transponder reply if the received peak power density at the transponder antenna is at least -103dBW/m² with the efficiency of 70% at minimum. The power density of the interrogation signal at the transponder antenna can vary from the just mentioned minimum up to a maximum of -22 dBW/m² when associated with ILS. This sensitivity level can vary by 1 dB between the station load up to 90%. The peak EIRP (Effective Isotropic Radiated Power) of the DME ground station should ensure the minimal peak pulse power density of - 89 dBW/m² in the coverage of the station. [23]

In order to avoid replying to echoes from multipath propagation, the DME dead time represents a time period when the ground beacon does not generate a reply on the received interrogations after reception of a valid interrogation immediately followed by its decoding. [26]

At least once in 40 seconds, the DME ground station will temporarily replace all pulses with the identification at a transmission rate of 1350 pp/s. This 1350 Hz Morse-code identifier enables to identify the DME ground station on-board the aircraft. In the meantime, the aircraft maintains the identifier in the memory mode to prevent a loss of the tuning. In case the DME is associated with a VHF navaid, the identification transmission is synchronized with the VHF facility identification code. In other words, the identification is transmitted three times from a VHF navaid at 1020 Hz and once from DME at 1350 Hz. [16,23]



Physical protection against weather conditions is necessary, and the reliability and continuity of the service can be ensured using monitors. The continuity of the service, measured by mean time between outages, can be improved using redundancy, such as dual monitors or dual transponders. [23]

The system accuracy may be affected by several factors, such as airborne equipment, propagation and ground equipment error. Nevertheless, the overall system should achieve accuracy 0.2 NM. [23]

1.3 TACAN

Although the research is focused on DME/DME navigation, it is necessary to mention also TACAN. Thanks to its distance measuring function identical to the DME function, TACAN can be used by civil aircraft. Military aviation considered the collocation of VOR/DME unsatisfactory for their usage, especially because of the large size and complex installation requirements in comparison to the low-frequency VOR antenna. Therefore, the US Air Force came up with a solution in the form of a new system, the tactical air navigation system. TACAN brought the integration of the bearing function with distance measurement into a single facility. Moreover, it enables better accuracy than VOR, and it is operated at a single frequency band. A new en-route navigation system VORTAC was introduced, in order to meet requirements not only for military, but also for the civil airspace users. In the case of VORTAC, the VOR bearing component satisfies the needs of the bearing determination for civil aircraft, and the distance component of TACAN replaced the function of civil DME. Military aircraft uses TACAN for the determination of both, azimuth and distance. It follows that TACAN is operated in the same frequency band as DME and uses the fixed pairing of operation channels when collocated with VOR. The TACAN compatible with DME was adopted by ICAO in 1959. [16]

Besides the DME pulse pairs, the TACAN ground station sends specifically coded pulses with one main reference burst (MRB) of 15 Hz and one auxiliary reference burst (ARB) of 135 Hz. The airborne receiver evaluates the phase relation between the two bursts and is able to determine the azimuth with an increased accuracy in comparison to VOR. TACAN ground facilities work with a higher transmission rate, and their output power can reach up to 5kW. The value of squitter pulse pairs reaches 2700 pp/s. The rise time of the pulse is defined 2 \pm 0.25 µs, and the decay time is defined 2.5 \pm 0.25 µs, which is slightly shorter than in the case of DME. [30]



Taking into account the identical function of distance measuring between aircraft and the ground station, there are no differences made considering TACAN or DME features in this research.

1.4 Alternative Position, Navigation and Timing

As abovementioned, even though the GNSS is evolving and the robustness and the reliability will increase by using multi-constellation and operating on more frequencies, the threat of GNSS unavailability with a significant impact on aviation remains. Therefore, the preservation of the current positioning systems functional is reasonable for ensuring the safety of aviation operations. In this context, the best candidate for GNSS reversion providing sufficient capability is DME/DME navigation. Some new alternative positioning, navigation and timing system concepts have been designed in order to ensure the long-term GNSS back-up, which are later mentioned. However, implementation such a system is a complex and time-consuming matter. The conventional navaids are usually operated by the states in the European region. For that reason, there is an opportunity to optimise the current DME network independently on national borders considering the possible increase in air traffic. [2,5,7,10]

Several research projects are ongoing in the area of APNT system concepts, as well as improvement of the current systems and their network optimisation. A short summary on APNT systems development is presented below. Although, the main focus lies on the DME/DME navigation solution and the following rationalisation of the DME infrastructure providing sufficient RNAV capability for the European Air Traffic network.

The standardization process of implementation of a new system in aviation always represents a complex task in the way of new technical requirements applicability on different aviation stakeholders worldwide. The compliance of a system with strict safety requirements and the development of the possible solutions is not only extremely time consuming but also very costly. Considering the current fast development in the area of digital technologies, the situation may occur when the new system is ready to be deployed, the technology is already outdated. Therefore, all APNT candidates aim to utilise the possible existing standards and equipment or the evolving as a part of a new system with a different primary intention. [33].

Furthermore, the limitation of the aeronautical frequency spectrum has to be considered by a new system implementation. Aviation spectrum users are criticized for the inefficient utilization of the aeronautical bands, and different entities are trying to gain



some allocation for sharing their services within aviation bands regardless of the safety threats. The most critical situation is on aeronautical frequencies 960-1215 MHz in the L-band, where a large number of civil and military navigation and surveillance systems operate on a spectrum sharing basis, including DME. The newly deployed system has to be accommodated with possibility to share frequencies with systems already in operation [33,34].

The most likely candidate for an APNT system, also considered as a long-term solution in the European ATM Master plan, is the L-band digital aeronautical communication system (LDACS). As stated in the name of the system, it is primarily intended for ensuring digital communication in aviation. However, LDACS has already demonstrated its capability to serve as an APNT system. Moreover, the research is conducted, in order to find a way, how LDACS could also support surveillance and RPAS (Remotely Piloted Aircraft System) C2 link (Command and Control link) applications. LDACS is intended to be operated in the Lband, and its interoperability with other systems in the band remains a priority. The specifications of the system are under development, and the standardization process within ICAO is being prepared. The required documentation including LDACS SARPS should be available till the end of 2022. [10,35,36]

The other APNT system concepts include eLORAN (Enhanced Long Range Navigation), Mode N with using the reverse multilateration principle, and further development of current DME. The other approach could also be to use a combination of the different existing and proposed systems to determine an aircraft position. Nevertheless, this potential solution of the so-called Modular-APNT requires data merging from various ranging sources with different performance levels and needs to be further investigated. Therefore, DME/DME is the only short term APNT. [34]



1.5 DME/DME Positioning

The DME/DME navigation is based on the principle of measuring slant ranges from two or more different ground stations. This enables determination of the horizontal position of an aircraft. This signal reception from several DME facilities with specific positions enables to reach RNAV capability for en-route as well as for departure and arrival routes. In order to generate position with the required performance, a position of the ground stations has to meet predefined conditions. For example, the FMS system must use DMEs with a relative angle between 30° and 150°, see Figure 10. Simultaneously, the range between aircraft and the two ground facilities must be greater or equal to 3 NM and lower or equal to 160 NM. The RNAV system must be less than 40° above the horizon when viewed from the DME facilities. The position is then determined using the triangulation method considering also previously calculated position in the FMS. [18,37]

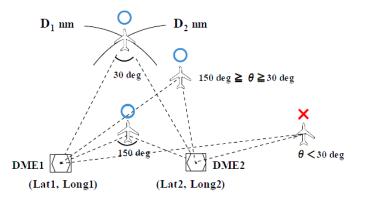


Figure 10 - DME/DME RNAV principle [38]

FMS considers Figure of Merit (FOM) value to identify the usable region of a DME facility, see Table 3. The DME ground station has to broadcast its identifier signal, satisfy the minimum field strength requirements and must be protected from interfering DME signal, otherwise, it will be considered invalid [37].

Table 3 - DME/DME RNAV	FOM [<u>39</u>]
------------------------	-------------------

FOM	DME/DME RNAV System of the aircraft must be:		
Value	Less than or equal to: (from the facility)	Less than: (above facility elevation)	
0	25 NM	12,000 ft	
1	40 NM	18,000 ft	
2	130 NM		
3	Beyond 130 NM		



In the PBN environment, it is assumed that DME can be used for navigation at a maximum range of 160NM. Information that indicates DME usable range (DOC) is stored in on-board Navigation database of FMS, which has a standardized format described in the Specification ARINC 424-20 Navigation System Database [39].

The area in which DME station ranging data is used by FMS is expressed by the following categories:

- Terminal usable within radius 25NM below FL 120
- Low altitude usable within radius 40NM up to FL 180
- High altitude usable within radius 130NM up to FL 600

The FOM parameter provides additional information about extended high altitude use beyond 130NM. Moreover, it can indicate whether the navaid is included in a civil international NOTAM system or is out of service. It is important to note that this database is provided by data houses that support FMS manufacturers database maintenance and is restricted due to its commercial value. [39]

1.6 RNAV Procedure Service Volume and DME Coverage Criteria

As abovementioned, DME has been evolving since the system standards were lastly updated, and they do not reflect the current system performance, including both, the onboard equipment and the ground facilities. As far as the DME/DME navigation solution has been considered a candidate for the short-term APNT in Europe, studies have been conducted on the current DME performance as well as their combination for DME/DME RNAV. The evaluation of data collection and processing proved that the current DME system is able to reach twice the better accuracy than the defined standard value. In other words, a DME ground station is able to provide in the 95% bounds (2 σ) lower error distribution than 0.1 NM, and a DME pair is able to provide the positioning accuracy at least two times better than current standards considering the aforementioned angle limitation; with a value of 0.15NM(~275m) \leq NSE¹ \leq 0.3NM(~550m). The data processing for the error distribution of a single DME is specified in Figure 11. [9,40]

The DME/DME positioning performance meets the accuracy requirements of RNP 1 specifications. However, the on-board monitoring and alerting system has to ensure the integrity for using RNP procedures, which is not provided for DME/DME navigation by the current avionics equipment. The FMSs are able to perform reasonableness checks to

¹ NSE (Navigation System Error) - the difference between the true position and the estimated position [41]



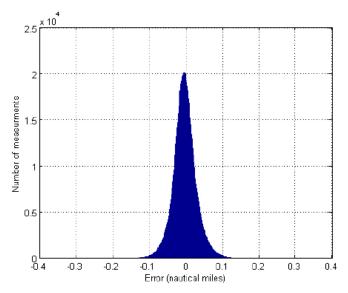


Figure 11 - Histogram on DME Slant Range Error Distribution [9]

verify the valid DME measurements to avoid database errors and erroneous system acquisition. Nevertheless, the DME/DME reasonableness checks are not considered as a full on-board monitoring and alerting, which can be ensured by Receiver Autonomous Integrity Monitoring (RAIM) in case of the GNSS. For this reason, stricter rules that compensate this lack should be set on DME transponders in order to enable RNP. [18,37,40]

The evolution of DME/DME positioning did not stop when it was referred as the shortterm APNT solution, but it is also expected that using multiple DME stations to get more precise position measurement will be further improved and put into operation. Besides the accuracy improvement by determining an aircraft position using all DME facilities in view, the multi-DME concept promises the system performance enhancement also by enabling the implementation of on-board integrity monitoring. [42]

Nowadays, research can be focused on the further development of improvement of the DME based on modern digital technology, in order to provide necessary performance including integrity requirements to serve the air traffic when GNSS unusable. However, this research is concerned with the current capabilities of the DME/DME RNAV positioning.

1.7 Navaids Infrastructure Rationalisation

The current conventional navaids are usually operated in the European region according to state needs. Therefore, the rationalisation of the VOR, DME, and TACAN ground stations infrastructure was assessed within a SESAR JU project with regards to support of the PBN



through ECAC (European Civil Aviation Conference) Member States; the project reference is D1.4.B ECAC – Wide Navigation Infrastructure Assessment [43]. The main objective of this project was to propose optimisation of the current navaids infrastructure and to find out whether decommission of some ground facilities would be possible.

The rationalisation of the navaids network is based on the results of the navaids Infrastructure assessment in the individual ECAC Member States, including also a crossborder network optimisation. The reports concluded the potential for reducing the navaids specified in Table 4.

Navigation Infrastructure Across All ECAC States		VOR	DME	TACAN
TOTAL	Current	754	793	169
	Rationalized	498	754	105

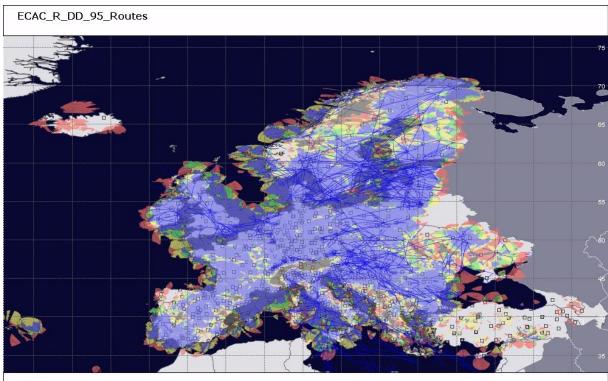
 Table 4 - Number of potential rationalised navaids [43]

The main objective of this research was to assess the navigation infrastructure capability across all ECAC states considering the state reports results in order to identify the possible solution for the navaids network optimisation, ensuring the sufficient RNAV capability. However, the rationalisation of VOR and DME was based on coverage estimation from limited information. Therefore, the results should not be considered as an implementation plan for the future navaids infrastructure, and a State ANSP is responsible for operation of the navigation facilities providing the required performance.

The research identified several issues which could influence the navaids rationalisation solution. For example, some of the states use a simulation tool considering antennas with their propagation characteristics for the navaids assessment. Nevertheless, the DEMETER (Distance Measuring Equipment TraceR) tool uses DOC of the facility, which may lead to different results. Another characteristic taken into consideration in the state reports was that the capability has to be ensured at all en-route levels, fully supporting the current route structures and traffic flows. In this case, the evaluation promulgated conventional routes were dependent on an evaluation where could have been used a different approach for each state.



The project report contains extensive studies on VOR/DME and DME/DME RNAV capability and navaids redundancy in different flight levels starting with the FL 95, assumed like a minimum en-route level for the ECAC Member States. In addition, it provides performance maps of the current state and the performance of the rationalised navaids network as seen from Figure 12. The DME/DME coverage evaluation also includes TACAN facilities and DME co-located with VOR. The individual maps of coverage were created in the DEMETER environment concerning the facilities DOC and a terrain.



Leaend

Figure 12 - Rationalised All ECAC DME/DME Performance FL95 [43] RED – No redundancy (1 valid DME pair) YELLOW – Limited Redundancy (at least 2 valid DME pairs having 1 common DME) GREEN – Full Redundancy (2 valid and independent DME pairs) BLUE – Excessive Redundancy (more than 2 valid and independent DME pairs)

Indeed, the SESAR project provides a complex assessment of DME/DME RNAV performance from the perspective of the DME ground facilities, further research is necessary to be carried out in order to evaluate the real impact of the DME stations decommission.



1.8 DME Network

As aforementioned, the current CNS network offers space for optimization and decommissioning of some ground beacons to keep only the MON. Even though the navaids optimization is often mentioned in connection with NDB and VOR, the DME optimization could also bring benefits while maintaining navigation performance, which was also proven in the Wide Navigation Infrastructure Assessment [43]. The necessity of the DME reduction results from the current congestion on DME channels [44]. Currently, States are encouraged to implement PBN capability in their airspace, which often causes the installation of new DME facilities in high-density areas. This can lead to leaving the original conventional navaids in operation, representing a more modular approach to navigation, and does no longer bring any additional benefit considering the network perspective. The growth of DME assignments and their sub-optimal distribution may represent an obstacle to introducing new systems, such as LDACS.

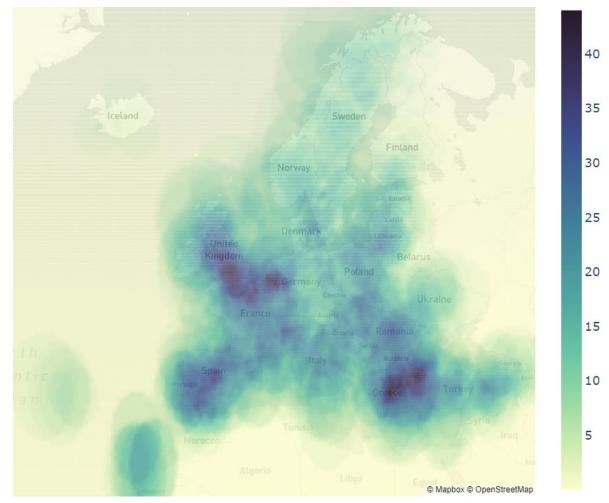


Figure 13 - DME density map in FL 200



The density map in Figure 13 shows that in many areas in Europe, more than thirty DME ground stations are available from a point in the airspace. The rationalization of DME infrastructure would reduce the load in the L-Band frequency spectrum and enable the operation of new systems.



2 Thesis Objectives and Hypotheses

Based on the above, it may be concluded that the short-term solution for ensuring required RNAV capability when GNSS is unusable is represented by DME/DME in the European Air Traffic network. The rationalisation of the current DME network allows further possibilities to improve the infrastructure and reduce the number of navaids in operation.

2.1 Thesis Objectives

The main objective of the thesis is to evaluate the optimised DME network capability to serve future air traffic in the European airspace by creating an appropriate model.

The current navaids infrastructure corresponds very well with the real air traffic density. In other words, the areas with a high air traffic, the number of the DME ground station, as well as other navaids, is also high, and vice versa. The SESAR study [43] shows that the RNAV capability might be accommodated with a lower number of DME facilities. This statement is considered positive in view of the utilization of frequency spectrum in Lband, of the compatibility with GNSS L5, and last but not least in terms of ANSP operational costs. However, the reduction of the ground transponders also causes a decrease of the reply capability and of the number of suitable stations for DME/DME positioning. A question remains whether the optimised DME network would be able to serve the increasing air traffic when a large scale GNSS outage occurs. The modern FMSs are able to scan and track up to five DME ground stations. Moreover, the aircraft is usually equipped with doubled systems to ensure resilience when there is a failure indicated in one of the systems. Other avionics manufacturers are already implementing the multi-DME logic, then on-board DME may interrogate up to 10 DME ground stations simultaneously. On condition that FMSs unpredictably need to rely on DME/DME navigation mode, it has to be ensured that the DME network is capable of serving all aircraft. In this situation, the ground stations may be approaching their saturation point, and they will no longer be able to send pulse pair replies to the more distant aircraft. Regarding this, the reduction of DME facilities in some high traffic areas could lead to an overload of the DME/DME network. Obviously, there is an effort to avoid this scenario. [45]

Therefore, the thesis aims to construct a suitable model of DME ground stations and air traffic to estimate the DME interrogation load with regards to technical standards of the current DME interrogators and transponders working in the DME/DME RNAV mode. The model then should provide the possibility to evaluate the eventual changes in the DME



infrastructure in the European Air Traffic network. In other words, the scenario of using the DME/DME back-up by an extensive GNSS unavailability is intended to test on the optimized DME network thanks to the constructed model.

2.2 Hypotheses

Given the thesis objective, individual hypotheses have been determined to be tested within the appropriate scenarios of the proposed model. The testing of the hypotheses may give an answer to the next evolution of the DME network in the European region.

Three hypotheses are defined to be confirmed or refused in the thesis:

- 1. The current DME infrastructure is able to serve the current air traffic on the assumption that 70% of aircraft are equipped with FMS enabling to interrogate up to five DME channels.
- 2. The optimised DME infrastructure, according to the SESAR project, is able to serve the current air traffic.
- 3. The optimised DME infrastructure, according to the SESAR project, is able to serve the estimated future air traffic.

2.3 Methodology

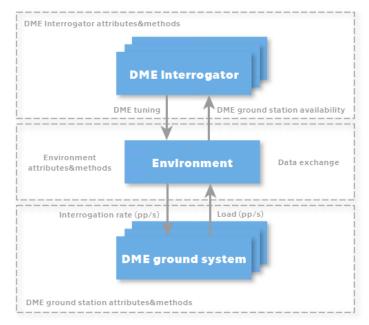
The proposed thesis aims at creating a software model corresponding to the real DME/DME RNAV environment with the possibility of testing scenarios that are not feasible in real traffic. The effort is put into modelling aircraft equipped with different DME interrogator types interrogating DME ground stations over Europe in order to evaluate possible changes in the current DME network. Therefore, several steps have to be carried out to construct the model with the required functionality. The model construction is based on the object-oriented programming (OOP) method, the data are processed and evaluated using statistical methods, and machine learning is implemented in several parts to fill in missing values or predict values based on the real measured data. All the methods used are presented in Chapter 3.

In the first step, suitable data analysis has to be carried out. For the input of the DME/DME model, it is necessary to obtain reliable data about air traffic, DME ground stations, and data about the equipment of aircraft. This data is complemented by other information about air traffic distribution or the on-board equipment capability. Since the model should simulate the real environment of the connection of aircraft DME interrogators



with the ground transponders, the important part of the input knowledge is represented by a proper understanding of an FMS tuning logic. It is assumed that DME facilities comply with the current technical specifications in Annex 10/Volume I [26] and DME operational standards [23,29] described in Chapter 1. Simultaneously, the interaction between these main entities for DME/DME RNAV is characterized by parameters, such as subtended angle, limited distance from the ground station, et cetera, contained in the EUROCONTROL RNAV guidelines [17]. Testing of all conditions has to be included in the model. Moreover, in the case of DME, the line-of-sight (LOS) has to be ensured between the transmitter and receiver. Therefore, radio wave propagation is necessary to be implemented into the model with respect to the radio horizon considering the atmospheric refraction.

Afterward, a model architecture is proposed to represent an interaction between individual objects of the model, see Figure 14. It is supposed that the model is able to show the situation from both perspectives; aircraft on-board DME interrogator and DME ground facilities. An aircraft is able to provide information about DME ground stations selected by the interrogator. In case more DME facilities are in line-of-sight than the channel tuning capacity of the interrogator, the model relies on the assumed FMS tuning logic and selects which of the available stations would be interrogated. Conversely, an individual DME facility is able to evaluate the number of aircraft interrogating the station in order to estimate the DME station load. Then, the DME ground station sensitivity is



adapted according to its capability to serve the defined number of aircraft **Figure 14** - DME/DME model basic architecture



simultaneously. The data exchange is proposed to be processed through an intermediary, an object representing the environment of the radio spectrum. The main task of this entity is to collect data and provide them to the appropriate DME ground station in one direction and then return the information about the ground station state back to the aircraft.

When all required data is obtained, the initial data processing is conducted to select suitable parameters for the model needs. All processes are constructed based on the rules known from the standards or modelled based on the knowledge of FMS experts. An important step is represented by an assignment of the individual aircraft to the ground stations that meet conditions set for possible tuning. A part of the tuning also includes the setting of the interrogation rate of the on-board equipment based on the DME interrogator type. The load of the ground station can be estimated knowing the interrogating aircraft. The exchange of information between aircraft and ground stations ends when no load of any of the ground stations overcomes the set limit of the maximum transmission rate. In other words, the process of retuning of overloaded DME ground station load, represented by the sum of the interrogation rate from all aircraft that have tuned the DME transponder. The output can also be visualised.

It is essential to realize that the model represents a static state of air traffic at one moment. However, the model processing takes time and changes dynamically before reaching this steady state. For the purpose of obtaining results from a different time, the input data of the air traffic from the required moment has to be obtained, and the whole model has to be run again. It is recommended to run a simulation for the same time several times in order to inhibit the effect of the randomly assigned variables in the model.

On the condition that the model is fully functional, it is considered a prototype of the rule-based model. In this step, the model functionality and the preliminary results are discussed with EUROCONTROL, and data from the real operation are compared with the results to identify possibilities for further improvement of the model in order to achieve more accurate results corresponding better to reality. At this stage, the implementation of a machine learning algorithm trained on real data can improve the rule-based model results and provide a better prediction of the final pp/s.



The model validation is necessary before the hypotheses are tested. For this purpose, data from the real operation are compared with the model results. An ANSP can provide this data from real measurements from the ground stations.

Once the model is created in a satisfactory form, the testing scenarios are defined in order to prepare the necessary conditions for testing the established hypotheses. The description of scenarios should contain the definition of data input and the possibility of evaluating the output data. Based on the scenarios testing, the results determine whether to confirm or reject the hypotheses. As indicated from the hypotheses definitions, the model will examine a specific constitution of the aircraft equipment, the optimised DME infrastructure with current traffic, and with increased future air traffic estimated in the European region.

The model is proposed to demonstrate the DME network capability considering an optimised DME infrastructure and the growth of air traffic. Based on the results of hypotheses testing, the optimisation of DME infrastructure may be proposed with a possibility of evaluating the impact of reduction of the number of DME ground stations. The model offers further possibilities to test the DME network optimisation feasibility. It can be used to identify areas where changes in the infrastructure may cause operational issues in the European network when aircraft need to rely on DME/DME RNAV navigation. Moreover, the model represents a tool for testing different issues related to DME/DME navigation, such as DME spectrum congestion, compatibility between other systems in L-band and DME channels, or insights into the difference between DOC defined for ground equipment and operational usage by on-board equipment using FOM. However, this is not the intention of this thesis.



3 Research Methods

The aim of the proposed thesis is to represent the real environment in a software model. Therefore, the method intended to use is based on one of the programming paradigms widely used in software engineering; the object-oriented programming. OOP offers a representation of real world entities like objects with mutual interaction. This concept is considered suitable to be applied to the specified environment of air traffic and the DME network. The model uses Python as the programming language.

Other methods include statistical methods as well as machine learning (ML) techniques that are mainly used in the data processing. In other words, it is used for preparing data for the model and creating the testing scenarios. Moreover, they are necessary for model validation and results evaluation. The following subchapters briefly introduce the techniques used in the research for data classification, clustering, and generating new data. Finally, the Monte Carlo Simulation (MCS) is also used in the model to avoid the model retuning to be deterministic.

3.1 Object Oriented Programming

The following subchapter is based on the book: Python 3 Object Oriented Programming [46].

As can be seen from the name of the software development concept, object oriented programming is based on an object corresponding to an entity from the real world. Then, the system models are typically represented by a collection of interacting objects with specific data set and behaviour. The individual objects may be grouped in a class with common characteristics. However, each object instance has its own data value of these characteristics known as an attribute. The behaviour can be explained like a procedure that the object is able to do. The set of actions that may occur on a specific class is referred to as methods that may be defined by parameters and return output values.

Once the object classes are designed with their attributes and methods, the determination of the object interface has to be set. In other words, the interface consists of a collection of attributes and methods of a subject which enable interaction with this object, while an internal procedure of the object is unknown. This principle of OOP is usually known as encapsulation. Nevertheless, the Python programming language does not apply this approach to data access. Unlike different object oriented programming languages that have implemented access control, all methods and attributes are



publicly available until they are marked with an indication with meaning for internal use only in Python.

The abstraction concept is related to software modelling, representing the fact that program objects are only a substitution of reality. It enables to ignore some properties of a real entity that are not relevant from the interface perspective to find appropriate level of detail between the public interface and internal object mechanism.

When creating a model, the proper setting of appropriate attributes, methods and considering the level of detail of knowledge that the external environment has to have for interacting with the object plays the key role in achieving the reasonable software model. It is recommended to always look at the requirements from the object's prospective, what it needs to know to may interact with its environment.

Composition and inheritance represent other basic principles enabling application of the level of abstraction used in OOP.

Object composition means that one object may be composed of several different objects. A typical example of composition is breaking down a mechanical system into smaller components, such as an aircraft is consisted of different parts like engines, fuselage, vertical stabilizer, horizontal stabilizer, or wings that can be further decomposed to aileron, flaps. The system models do not only have to represent a mechanical model where the composition is easy to imagine, but also the object may specify an abstract thing like the name or title of an object representing people. This relation between objects provides an easy way for setting levels of abstraction based on a different user perspective.

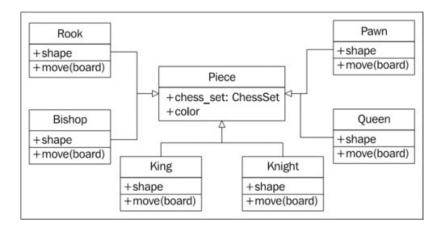


Figure 15 - Chess Pieces Inheritance [46]



One of the most significant tools of an OPP specification is the characteristics of relationships between objects by inheritance. Inheritance may be imagined as a hierarchical structure of a family tree with a superior parent class and its descendants. The oldest ancestor represents the structure root, a class in the programming environment. Similarly, to the real world, one class may inherit attributes and methods from the superior one. This enables application of the same definition of data sets and behaviour. However, the individual subclasses vary from the set values, and the same procedure may provide different result. For example, the piece classes may be defined in the chess representing system, see Figure 15.

In this case, the descendants inherited the colour of the chess set according to which they belong. The individual piece type is then constituted of an attribute of their shape and method of specific movement they can make on the chess board. Inheritance differs the OOP from previous programming paradigms and offers a wide ability to the extent the object properties. However, the programmers often tend to overuse this ability and connect objects with no behaviour in common that may create an unsuitable design.

Provided that the hierarchy principle is applied, the polymorphism is called ability which may cause different behaviour in one class. In relation to the chess board, the move function may be called on the piece, and the chosen subclass by a player will move on the board object. However, the result of the movement is dependent on the proper piece, and it may be done without the board object knowing the subclass.

In conclusion, the object oriented programming is used extensively in software engineering, offering the ability to model the real world into objects with attributes and functions. Moreover, it applies different concepts and properties to enhance the modelling process, such as encapsulation, abstraction, composition, inheritance or polymorphism.

3.2 Regression

The regression algorithms in Machine Learning (ML) enable to define of relationships between some model parameters, referred to as features, and output values, the target. It represents a type of supervised learning that is able to predict a target value based on the current features of a model for new data not used for the model training. Different ML algorithms were tested during the model creation, such as Linear Regression, Logistic Regression, Random Forrest, Support Vector Machines with different kernel functions, Lasso Regression, Polynomial Regression, or Extreme Gradient Boosting. The best



performance for the need of the thesis reaches the Gradient Boosting Regression (GBR) [47]. Therefore, it is described in detail. The metrics for evaluating the regression algorithm performance and model validation are also described below. [48,49]

3.2.1 Gradient Boosting Regression

Even though linear relations were expected between features and target in the created model, the best performance was reached by using GBR, representing an ensemble method in supervised machine learning that can usually outperform Random Forest. In other words, the final model is created as a combination of multiple weak learners represented by decision trees to provide the best total performance.

The algorithm tries to minimize the prediction error, represented by a loss function, to improve the prediction. Weak learners are created to minimize the loss of function. In each step, the prediction reduces its residuals for the model features and adds updated residuals, including a learning rate, to the previous step. The process repeats until the residuals are no longer improved. Thus, more trees are combined, and the final model can be split into several nodes before reaching the terminal node. [50,51]

The GBR is initialized by minimizing the loss function *L* of the model in the first step. The variable *x* represents model features with a dependency on *y* values. The constant value $F_0(x)$ represents the result:

$$F_0(x) = \underset{\gamma}{\operatorname{argmin}} \sum_{i=1}^n L(y_i, \gamma)$$
(4.1)

Considering the squared loss in the regression, the residual value γ which minimizes the loss function is equal to the mean of the γ value. The minimization step for γ is solved by finding the local minimum of the loss function, by taking a derivative of $\sum L$ with respect to γ . [50]

The following steps are repeated M times. The number of iterations M indicates the number of the trees, representing the weak learners, and m refers to the index of each tree. [50]

Firstly, the residuals r_{im} are computed by taking the derivative of the loss function regarding to the previous prediction F_{m-1} multiplied by -1 for each sample *i*. This step provides information about the direction and the magnitude where the loss function can be minimised [50].



$$r_{im} = -\left[\frac{\partial L(y_i, F(x_i))}{\partial F(x_i)}\right]_{F(x) = F_{m-1}(x)} for \ i = 1, ..., n$$
(4.2)

Secondly, the terminal nodes reasons R_{jm} are created in the next step by training a regression tree with features x against residuals r, where j represents a terminal node and J the total number of leaves [50].

Thirdly, the γ_{jm} value is computed in order to minimise the loss function on each terminal node [50].

$$\gamma_{jm} = \underset{\gamma}{\operatorname{argmin}} \sum_{x_i \in R_{jm}} L(y_i, F_{m-1}(x_i) + \gamma) \text{ for } j = 1, \dots, J_m$$
(4.3)

Lastly, the prediction F_m of the additive model is updated by picking the value γ_{jm} if the given features x belongs to the terminal node R_{jm} . The updated prediction is made by adding the corresponding γ_{jm} to the previous prediction F_{m-1} . The learning rate v corresponds to a proportion of the contribution of the additional tree prediction to the combined prediction.

$$F_m(x) = F_{m-1}(x) + \nu \sum_{j=1}^{J_m} \gamma_{jm} \mathbb{1} (x \in R_{jm})$$
(4.4)

To reach the best model performance, the number of iterations can often overcome one hundred steps, representing one hundred created trees [50].

The implementation of the GBR is made using the scikit-learn Python library [52] with a function GradientBoostingRegressor. The hyperparameters that were set up for the best model performance are the following:

- max_depth the maximum number of nodes in the tree;
- min_samples_leaf the minimum number of samples required to be at a leaf node; and
- subsample the fraction of samples used for fitting the weak learners.

3.2.2 Regression Validation Metrics

The model is evaluated based on the comparison of its with real data. The metrics used for validation are taken from the metrics regression section of scikit-learn Python library and the correlation coefficient. Specifically, the following parameters are used:

 Correlation coefficient – the ratio between the covariance of two variables and the product of their standard deviations:



$$r = \frac{\sum (x_i - \bar{x}) (y_i - \bar{y})}{\sqrt{\sum (x_i - \bar{x})^2 \sum (y_i - \bar{y})^2}};$$
(4.5)

where x_i represents the value of real pp/s and \bar{x} their mean, and y_i the value and \bar{y} mean of estimated pp/s in the model, either results of original the model or the prediction model.

• MSE (Mean Square Error) – the sum of square of prediction error, which is real output minus predicted output and then divided by the number of data points:

$$MSE = \frac{1}{N} \sum_{i=1}^{N} (y_i - \hat{y}_i)^2$$
(4.6)

• RMSE (Root Mean Square Error) – the square root of MSE:

$$RMSE = \sqrt{\frac{1}{N} \sum_{i=1}^{N} (y_i - \hat{y}_i)^2}$$
(4.7)

• MAE (Mean Absolute Error) – the sum of the absolute value of error:

$$MAE = \frac{1}{N} \tag{4.8}$$

For MSE, RMSE and MAE equations, the y_i represents the real value of pp/s, and \hat{y}_i the value estimated in the model. [53,54]

3.3 Data Classification

The implementation of machine learning techniques has become easy using the Python library scikit-learn [55]. For the purposes of the thesis, the supervised learning method is necessary to use to assign missing values to the input data of the model. In other words, based on the known values of the data category, the data with unassigned category can be also classified.

Several classification models enable to predict a category of an object were tested during the model construction, specifically logistic regression, decision tree, random forest, Support Vector Machines (SVM), K-Nearest Neighbour (KNN), naive bayes, and fuzzy decision tree. [56]

The results of each model were evaluated on the same data using cross validation which is considered a suitable estimator of the model performance. Based on the common practice, the data is divided into train and test dataset. The train set is intended for the model training, and the performance evaluation is conducted using the test dataset. The



category for both datasets needs to be taken from known data. The main purpose of the cross validation is to examine the capability of the model to predict a classification for un unknown data, the data that were not used for the model training. The scikit-learn function "cross_val_score" is chosen for the performance evaluation uses the k-fold cross-validation method [57]. The data are divided into k folds. The model is then trained on k - 1 folds, and the remaining fold is used for the validation of the model results. The used performance metric is represented by the mean of accuracies calculated for each fold. The accuracy is represented such as:

 $\frac{accuracy}{precision} = \frac{true \ positive + true \ negative}{true \ positive + false \ positive + false \ negative + false \ positive}$ (4.9) The Support Vector Machines method proposed by Vapnik [58] with the linear kernel

The Support Vector Machines method proposed by Vapnik [58] with the linear kernel function showed up the best performance. Therefore, SVM is described in the following subchapter.

3.3.1 Support Vector Machines

The SVM is a supervised machine learning technique based on constructing of hyperplanes in a high or infinite dimensional space. It can be used for classification as well as for regression. The support vectors represent the maximal margin of the closest points that separates the hyperplanes. In other words, the SVM method tries to find the optimal hyperplane which separates the different categories of the sample in the classification task. The SVM solution of a linear problem is represented in Figure 16. [59,60]

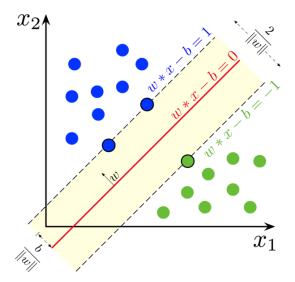


Figure 16 - SVM solution of a linear problem [60, edited]



Assuming the dataset of *n* instances, where $x_i \in \mathbb{R}^m$, i = 1, ..., n, and labels $y_i \in \{-1, 1\}$, i = 1, ..., n, the SVM method finds a (m - 1) dimensional hyperplane $w^T \cdot x + b = 0$ that has the maximum margin between the boundary points of the two categories, also defined as the minimum distance between the data instances and the decision boundary. It may be solved thanks to finding a solution for the optimization problem:

$$\min_{\boldsymbol{w},\boldsymbol{b},\boldsymbol{\xi}} \frac{1}{2} \cdot \boldsymbol{w}^T \cdot \boldsymbol{w} + C \sum_{i=1}^n \xi_i$$
subject to $y_i(\boldsymbol{w}^T \boldsymbol{x}_i + \boldsymbol{b} \ge 1 - \xi_i, \xi_i \ge 0, i = 1, ..., n)$

$$(4.10)$$

The parameter *C* characterizes the compromise between minimizing the error and maximizing the margin. The optimization problem can be solved using quadratic programming. The problem is to find local extremes of function constrained by other functions. These local extremes can be found by using Lagrange multipliers. The dual problem to equation 4.11 is then:

$$max \frac{1}{2} \sum_{i,j=1}^{n} \alpha_{i} \alpha_{j} y_{i} y_{j} K(\boldsymbol{x}_{i}, \boldsymbol{x}_{j}) - \sum_{i=1}^{n} \alpha_{i}$$

$$subject \ to \ \sum_{i=1}^{n} \alpha_{i} y_{i} = 0, 0 \le \alpha_{i} \le C, i = 1, ..., n$$

$$(4.11)$$

The Kernel function for a linear Kernel is defined as $K(x_i, x_j) = x_i^T \cdot x_j$, the dot product of x_i and x_j . Support vectors are the data instances x_s whose Lagrange multipliers are above zero. Finally, decision function for classifying a new data instance \hat{x} is:

$$\hat{y} = sgn\left(\sum_{i=1}^{n} y_i \alpha_i K(\boldsymbol{x}_i, \boldsymbol{\hat{x}}) + b\right)$$
(4.12)

[<u>58,59,60</u>]

3.4 DBSCAN Clustering

Clustering represents an unsupervised learning method aiming at grouping similar data points together. The density-based clustering is chosen as the most suitable method to ensure not only finding the required groups, but also to detect outliers. The Density-Based Spatial Clustering of Application with Noise (DBSCAN) algorithm [61] is used in the thesis model.



The main functionality of the DBSCAN is to find regions with high density and separate them from one another by regions of low density. The cluster creation is limited by two parameters. The first specifies the maximum distance between two points, circle of a radius $Eps(\epsilon)$. The second determines the minimum number of points within the radius to define a cluster and is referred to as MinPts. The neighborhood of a point p in the database D is defined as follows:

$$N(p) = \{q \in D \mid dist(p,q) \le \epsilon\}$$
(4.13)

Based on the parameters, each point can be determined as a core or a border point, or an outlier. The core point represents a point surrounded by minimum number of other points in the cluster in the maximum distance, if $N(p) \ge MinPts$. The border point is reachable from a core point, but it is surrounded by fewer than MinPts. The noise, or also referred as an outlier, is a point that is not reachable from any of the core points. [62]

The algorithm starts at a random point. The neighbourhood area is retrieved from the ϵ parameter. In the case, there is enough points in this radius, it is considered a core point and a cluster starts to create. In the opposite case, the point is marked as a noise. Even though the point is firstly labelled as a noise, it can be revisited in the following step and become a part of another cluster. If yes, they become a part of the cluster. All points included in the cluster are further examined to be core or border points. Afterwards, another point which has not been visited yet is chosen, and the previous steps are repeated. The process is finished when all points were examined. [62]

The distance between two points is usually the Euclidian distance [62].

3.5 Generating Samples

New samples of data are necessary to generate within one of the scenarios. For this purpose, mixture clustering is used, and the new data of latitude and longitude values are generated, considering them to have a normal distribution. The altitude values are generated from the inverse function because of its custom distribution function.

3.5.1 Mixture Clustering

In accordance with the Mathematical Methods for Data Analysis [63], mixture clustering is a suitable statistical method capable for describing systems with a finite number of different working modes. The model of the mixture consists of a set of individual models and a set of a model pointer represented by a sequence of discrete random variables. The values of the pointer point to the active component at any time.



The method is used to generate new data items from existing clusters. Therefore, it is necessary to estimate clusters, representing mixture components, from the existing data. Then, we can use the estimated model for generating new data. In the first step, priori parameters for the mixture components have to be set in the estimated centres of the clusters based on the distribution of the values. The visualization of the model data can also provide information about the necessary number of clusters. The two-dimensional mixture with two components is shown in Figure 17.

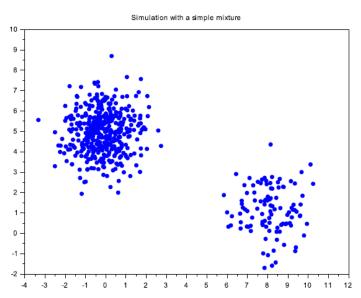


Figure 17 - Example of clusters each representing one mode of the system [64]

Each individual component of the model is assigned to one of the working modes before the mixture estimation. Once it is decided where the component belongs, the update of statistics of components and the mixture model can be processed. In reality, the working model that generated the data is estimated based on the assigned probabilities. When the component is assigned to the most probable working mode, the statistics of the model can be updated.

The current pointer value is estimated for each new data item. Its proximity to individual components is determined, and a weighting factor ω is computed as a product of the proximity and component probability which is constructed from relative frequencies of past activities of individual components. The proximity is computed by substituting into the component model the existing point estimates of its parameters from the last step of estimation together with the new data item. The value of proximity increases the closer it is to the component centre. In order to get values of the proximities, it is necessary to calculate in logarithm, subtracting the maximal value and taking the exponential. Otherwise, the values are too low for the program to distinguish them. The



weighting vector ω is equal to product of the proximities and the pointer model after normalization to sum equal to one.

Hence, the statistics of all components together with the pointer model are recomputed using the weighting factors. The update of statistics for the continuous model recomputes the information matrix and the counter. The discrete model adds the statistics item corresponding to the measured data multiplied by the corresponding item of the weighting vector ω .

In the last step, the point estimates are determined for all unknown parameters.

Once all data items are assigned to one of the components, the required number of new values may be generated from the mixture component by adding noise to the calculated pointer.

3.5.2 Inverse Transform Method

A possibility how to generate data samples for variables where the probability density function differs from the commonly used ones is using the Inverse Transform method. The principle is based on finding a link between a custom and uniform distribution. [65,66]

$$F(X) = U \tag{4.14}$$

Having *X* a set of realizations of a variable with a custom distribution function *F*, we can understand that *U* represents a set of equal realizations within the interval (0,1). Then it applies:

$$X = F^{-1}(U) (4.15)$$

In other words, *X* values can be got as generated values from *U* shown from the inversion of *F*. [66]

In a discrete model, the distribution function is mapped using a dense grid, represented by a sample (or empirical) probability function f. This function can be obtained as the normalized data histogram. Then the function F is constructed as a cumulative sum of f, represented by cumsum(F). In other words, an area delimited by the grid in a uniform distribution is measured, and the corresponding value delimiting the same area under in the histogram with the custom distribution is found. [66]



Then, a random value u belonging to U is generated, and the probability that the random value will fall into the intervals delimited by the specified bound is measured, that can be calculated in Scilab environment using the following command:

$$sum(cumsum(F) < rand_u) \tag{4.16}$$

Finally, we get ones in the left-hand interval and zeros in the right-hand interval, and the position of equality is found. [66]

3.6 Monte Carlo Simulation

The Monte Carlo Simulation (MCS) represents one of the most popular techniques for solving complex issues involving randomness. For the purposes of the thesis, the MCS represents a tool for avoiding dependency on the order and random assignments in the model. Thanks to the repetition of the model processes with randomly distributed variables, it helps to estimate the initially unknown distribution of the variables. The number of steps in the MCS can reach up to thousands of repetitions. The result of the MCS is calculated as the mean of the estimates in individual repetitions $\hat{\theta}$, and it determines the Monte Carlo unbiased estimate for the variable, represented by $\hat{\theta}_{MC}$ in the formula [67]:

$$\hat{\theta}_{MC} = \frac{1}{n} \sum_{i=1}^{n} \hat{\theta}.$$
(4.17)



4 Data Collection

Information about air traffic over Europe at a required time, DME ground stations, and types of possible DME interrogator capabilities create the essential dataset enabling the model functionality. In this Chapter, the sources of these model inputs are described in detail.

4.1 Air Traffic Data

The air traffic data are gathered from the OpenSky Network database [68]. The OpenSky Network represents a non-profit organisation providing open access data that are collected from volunteers, industrial, academic, and governmental organizations. A network of receivers on 1090 MHz operated by data suppliers ensures the provision of the data from ADS-B and Mode S technologies. The OpenSky Network enables downloading of the raw data, or it provides a possibility to gather complete datasets containing tables with already preprocessed data using SQL queries. This database is placed in a cloud, and it is accessible using the Impala Shell tool in order to filter, aggregate, and combine data collected in the database. The OpenSky Network offers an official repository enabling to access the data with Python API (Application Programming Interface). To be allowed to download data from the OpenSky database, the user has to be registered. The registration is subject to an explanation of the intent of the data needs.

For the needs of the thesis, a table containing the so-called state vector is used. The state vectors represent the most common tracking information that is decoded from the messages on 1090 MHz and is provided per one second for each aircraft. The data is provided in the metric system, unlike in aviation. The overview of the name and data type of data in the state vector table is shown in Table 5.

Name	Туре	Meaning
time	int	Unix timestamp of the state vector validity
icao24	string	24-bit ICAO address
lat	double	Latitude in WGS84
lon	double	Longitude in WGS84
velocity	double	Speed over ground of the aircraft $[ms^{-1}]$
heading	double	The direction of movement (track angle) as the clockwise angle from the geographic north
vertrate	double	Vertical speed of the aircraft $[ms^{-1}]$
callsign	string	Callsign that was broadcast by the aircraft

Table 5 -	Data in	State	Vector	Table [<u>68</u>]
	Dutum	June	VCCLOI	



Name	Туре	Meaning
onground	boolean	Whether the aircraft is broadcasting surface positions (true) or airborne positions (false)
alert	boolean	Special indicator used in ATC
spi	boolean	Special indicator used in ATC
squawk	string	Squawks assigned by ATC
baroaltitude	double	Altitude measured by the barometer $[m]$
geoaltitude	double	Altitude determined using the GNSS (GPS) sensor $[m]$
lastposupdate	double	Age of the position
lastcontact	double	Time when OpenSky received the last signal of the aircraft
serials hours	array <int></int>	Data is processed and partitioned in hourly batches - a mark of the beginning of the hour to which the data belongs

Besides the cloud-based database, the OpenSky Network provides different datasets for download. Additional information about an aircraft can be found in the metadata directory. More specifically, a file, which is updated monthly, contains aircraft database where a specific ICAO 24-bit address is complemented by other parameters, such as manufacturer, aircraft type, operator, and many other information describing the aircraft. The last dataset used for the air traffic data input is based on ICAO Doc 8634 [69] dealing with the Aircraft Type Designators. In this case, each aircraft type represented by an ICAO designator is described according to its manufacturer, model, engine count, engine type, and WTC (Wake Turbulence Category). Based on this description, each aircraft is represented by three symbols; for example, the "L2J" corresponds to a landplane with two jet engines. The meaning of individual categories is represented in Tables 6 and 7. The number of engines is represented by an integer.

Aircraft	Character
LandPlane	L
Amphibian	А
Gyrocopter	G
Helicopter	Н
SeaPlane	S
Tilt-wing	Т

Table 6 - ICAO Doc 8634 aircraft description	on [<u>69</u>]
--	------------------

 Table 7 - ICAO Doc 8634 aircraft engine description [69]

Engine	Character
Electric	E
Piston	Р
Jet	J
Turboprop/turboshaft	Т
Rocket	R



All abovementioned datasets create the input of air traffic data. Nevertheless, it is necessary to consider that the source of data is dependent on the equipment of aircraft with ADS-B out, respectively with Mode S transponder, that can transmit information derived from airborne systems. The requirement on ADS-B equipment is set out in the Commission Implementing Regulations (EU) No 1207/2011 [70] with regards to the aircraft with a maximum certified take-off mass exceeding 5 700 kg or having a maximum cruising true airspeed capability greater than 250 knots. According to the EUROCONTROL ADS-B equipage monitoring [71], 90.7% of aircraft and 91.3% of flights in the Network Manager area were equipped with ADS-B version 2 as of January 2022. The following Figure 18 represents the equipage rate of aircraft with ADS-B version 2 per market segment. In other words, it shows the percentage of equipped aircraft with respect to the total aircraft fleet.

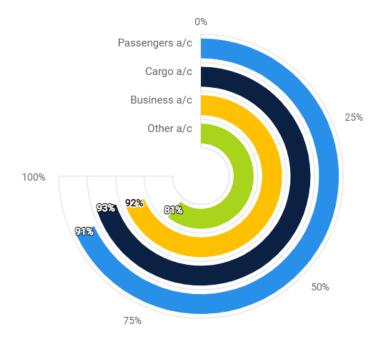


Figure 18 - ADS-B version 2 equipage [71]

Based on the percentage of the ADS-B equipage, it is assumed that the air traffic data obtained from OpenSky Network does not necessarily include all the traffic. Not only the figure does not show completed equipage, but also the requirements stated in the European regulation does not include smaller aircraft that could be equipped with a DME interrogator.

This issue can be solved by applying a machine learning algorithm, which is trained on the real data from DME ground stations provided by an ANSP. Thus, the interrogation rate from aircraft unequipped with ADS-B is compensated by adding interrogation pulse



pairs to other aircraft with the distribution corresponding to the real situation represented by the measured data.

4.2 **DME Ground Stations**

The data about DME ground stations is obtained from several sources; the Distance measuring equipment tracer (DEMETER) software tool [72], the Spectrum and frequency information resource (SAFIRE) database [73], and data from AIPs (Aeronautical Information Publication) provided by EUROCONTROL in a software readable version.

The DEMETER tool is primarily intended to support the implementation of PBN and navigation infrastructure rationalization. One of its functionalities includes assessing the DME/DME positioning performance in en-route or terminal areas and determining whether some changes in the ground infrastructure are needed to support the PBN application. DEMETER also supports the assessment of VOR/DME coverage and infrastructure evolution planning [72]. A DEMETER dataset containing the parameters of all navaids was provided directly by EUROCONTROL for the purposes of this thesis on 22 May 2022. The dataset includes parameters not only for DME operated in the ECAC States but also other navaids, such as TACAN, VORTAC, VOR, and NDB. Table 8 specifies the navaids parameters that are used as input in the model.

Parameter	Meaning	
navAidId	Identification of navaid	
Location	Name of the navaid location	
navAidType	Navaid service - DME, DME/ILS, TACAN, VOR/DME, VORTAC	
latitude	Latitude of the ground station in WGS84	
longitude	Longitude of the ground station in WGS84	
elevation	Elevation of the ground station [ft]	
antHeight	Height of antenna [ft]	
doc	Radius of DOC [NM]	
maxAltitude	Maximum altitude of DOC [ft]	
icaoCC	ICAO Country Code	
channel	Operational DME channel	
declaredEIRP	Declared EIRP	

 Table 8 - Parameters of DME ground station data [72]

The SAFIRE database represents the central register for radio frequency assignments in the EUR region. It enables frequency management for communication and navigation facilities. It is used by frequency managers of individual States for coordination of frequencies to ensure the protection of their assignment and avoidance of radio



interference [73]. The data in the SAFIRE database is customized to the needs of the frequency coordination criteria and contains similar parameters like the DEMETER dataset, see Table 9.

Parameter	Meaning
СТҮ	ITU Country Code
Ref	Unique assignment reference
Location	Name of the navaid location
Service	Navaid service - DME, DME/ILS, TACAN, VOR/DME, VORTAC
Areas	Latitude and longitude of the ground station in WGS84
MaxFL	Height of antenna [FL]
Radius	Radius of DOC [NM]
ld	Identification of navaid
DMEChannel	Operational DME channel
DMEPwr	EIRP [dBW]
Status	Assignment status – Assigned/Operational

Table 9 -	Parameters	of DMF	around	station	data	[73]
	i ulumeters		ground	Station	autu	

The AIP data were used for validation when the information from the two data sources differ.

4.3 DME Interrogator

Aircraft can be equipped with various types of avionics that can have different behaviour when tuning DME ground stations. Moreover, the DME interrogator can be doubled similarly to FMS in order to provide an independent data source for navigation information. The information about the equipment of the aircraft and capabilities of the DME interrogator and the FMS tuning logic are not accessible from any open source. Besides that, FMS manufacturers are unwilling to share detailed information about their products. Therefore, the input for DME interrogators has to rely on discussions with the experts, together with information found in general technical specifications.

It is assumed that a proportion of aircraft is equipped only with a single channel DME interrogator, which does not enable DME/DME RNAV positioning. Airliners and business jets are considered to use multiple channel DME interrogators. One type of DME interrogator uses multi-DME instead of using DME/DME pair to determine the aircraft position. In this case, the interrogator can scan up to ten DME ground stations at once. The standard equipment for most aircraft is considered a DME interrogator, which is able to scan four or five DME channels. Concerning the four-channel DME interrogator, two channels are used for tuning a suitable DME/DME pair, the one with the most appropriate



geometry for determining the DME/DME RNAV position. The other two channels are reserved for tuning to check raw data on the Navigation Display, VOR/DME position fixing, ILS/DME, procedure-specified or pilot-selected navaids. Similarly, the five-channel DME interrogator can tune a DME/DME pair for RNAV. However, it can tune a third channel related to DME/DME positioning to provide reasonableness checks. In other words, it compares the range measurements received by the navigation function with the expected radio range measurement to this station. This channel is not used for position computation. The remaining two channels have similar usage to the four-channel DME interrogator uses the multi-DME logic as "C", the five-channel interrogator as B, and finally, the four-channel interrogator as "A" type. The essential features of the interrogator types are specified in Table 10. [45]

system_i d	selection_poo I	horizontal_lo w	horizontal_u p	vertical_limi t	max_station s
Α	LOS	30	150	40	4
В	LOS	30	150	40	5
С	LOS	-	-	40	10
S	LOS	-	-	40	1

Table 10 - Parameters	of DME interrogator types
	of Drife interiogator types

The "selection_pool" parameter indicates that the FMS can tune only DME ground stations in the LOS of the aircraft. The "horizontal_low" and "horizontal_up" parameters represent limit values for the subtended angle to selected DME/DME pair which has to be between 30° and 150°. The vertical "vertical_limit" indicated that the elevation angle between the aircraft and the ground has to be 40° at maximum [17]. The number of maximum stations that the DME interrogator can tune is represented by "max_stations".

Another important value connected to the DME interrogator is its transmission rate. Based on the assumption that the interrogator is 5% of the time in search mode and 95% in track mode, the average PRF is used:

- Single channel interrogator: 16 pp/s;
- Multiple channels interrogator: 48 pp/s.

The transmission rate can differ for older equipment of general aviation and helicopters, where the PRF is considered 30 pp/s instead of 16 pp/s. In case of two similar interrogators are installed on-board an aircraft, the number of the interrogated pulse



pairs is also doubled. As the two interrogators work independently on one aircraft, it is assumed that both will tune the same DME ground stations. [29]

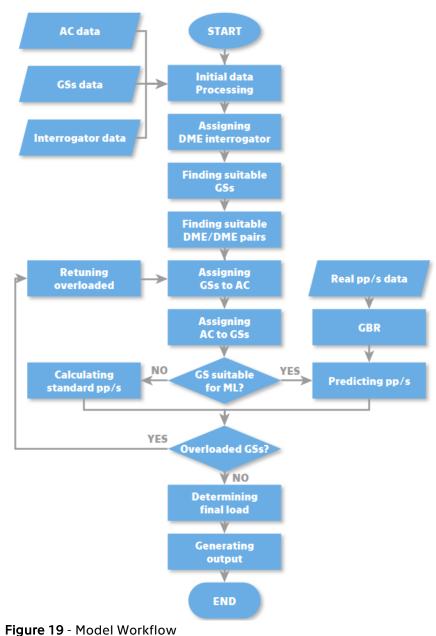
The abovementioned interrogation rate is used in the rule-based model. Nevertheless, the final estimated value of pp/s may not correspond to the standard values because of implementing a machine learning algorithm.



5 Model Construction

This chapter deals with the software model¹ created to simulate the connection of DME interrogators on-board of aircraft with the DME ground stations in order to estimate the ground stations load. The individual steps of the model construction and all parts of the model, consisting of Python scripts, are described in detail.

The model is run using the Python script "main.py" that is able to process all necessary functionalities thanks to calling the required parts of the model. The individual steps are shown in Figure 19.



¹ The model script is attached to the thesis in digital form, and it can also be provided upon request in GitLab [74] environment.



As aforementioned, the input data into the model are represented by the air traffic data, the DME ground stations data, and the DME interrogators capability data. At first, this data is initially processed to be sufficient and suitable as an input into the model. Subsequently, the input is transformed into predefined classes. In other words, each aircraft represents an instance of an "Aircraft" class, each DME ground station represents an instance of a "GroundStation" class, and each type of the DME interrogator represents an instance of an "FMS" class. DME interrogators are assigned to each aircraft when all data are loaded into the classes. Afterward, the "Environment" class is filled with all aircraft and ground stations.

In the following step, a list of ground stations that are possible to tune is constructed in relation to the DME interrogator equipage of the aircraft. Similarly, a list of all considerable DME/DME pairs is created. As soon as these lists are completed, the ground stations are assigned for each aircraft based on the assumed FMS logic. A list of interrogating aircraft is built for each DME ground station to achieve the connection from both perspectives.

Then, the load of the DME ground stations is calculated depending on the type of the DME ground station. The standard values of interrogation rate are used for all terminal stations. In case the DME facility is considered en-route, the final load is predicted using the GBR. For all DME ground stations with a DOC higher than 100 NM, the final load is taken as the higher value from the result of the rule-based model and predicted value by the GBR. On condition that some of the DME ground stations are overloaded after the first allocation, the process of retuning is run until all values of load for each ground station are estimated under a maximum load limit. Finally, the output of the model is generated, and the results can be visualized.

Additional script "utils.py" is used in the model processing. It ensures conversion of units and download of elevation values. Python APIs numpy [75] is used for calculations and matplotlib [76] for figures construction.

5.1 Initial Data Processing

The initial data processing ensures that the data of air traffic, ground stations, and DME interrogators is provided in the required form for the model. It includes data download, reading, cleaning, corrections, and filling in missing values. Moreover, the increased air traffic data can be generated if necessary for a testing scenario. The output is



represented by data frames that can be directly transformed into the appropriate classes.

5.1.1 Initial Processing of Air Traffic Data

The initial processing of air traffic data is contained in "opensky_download.py" script and, it consists of two main functions. Firstly, the air traffic data are downloaded from the OpenSky Network database. Secondly, the downloaded data are further processed and merged with additional files to get more detailed information about aircraft.

The connection to the OpenSky Network database uses a library pyopensky [77] which represents a Python interface to query and download ADS-B and Mode-S data using Impala Shell [68]. The input for the download query is characterized by the date and time of the required air traffic fingerprint optionally a limited area can be defined, setting up minimum and maximum latitude and longitude. By default, the area limitation is chosen to cover the uttermost ground station in, such as:

- Maximum latitude = 82°
- Minimum latitude = 23°
- Maximum longitude = 44°
- Minimum longitude = -35°

The downloaded data are automatically saved into a temporary csv file to create a backup in case the next step of processing is necessary to repeat, or the Impala Shell is unavailable. The data are collected not only for a specific second but also a time margin is created of \pm 10 seconds for cases data was not provided for an aircraft in the exact second. A few rows of the downloaded data are shown in Table 11, with values rounded to two decimals. As seen from the data sample, values of altitudes are not available for aircraft on the ground, and some aircraft provide values of barometric altitude only. The downloaded dataset also contains duplicate data for most aircraft due to the setup time frame. Therefore, further processing of the downloaded data is necessary.

time	icao24	lat	lon	onground	geoaltitude	baroaltitude
1627633790	4ca84e	46.39	1.27	false	11666.22	11277.6
1627633790	42434b	54.57	50.14	false	11315.7	10972.8
1627633790	440821	43.29	13.18	false	11102.34	10561.32
1627633790	34724e	36.81	-4.37	true		
1627633790	4a05a5	45.38	17.79	false		10675.62

Table 11 - Sample of OpenSky Air Traffic Data



The data processing starts with saving the downloaded data into a pandas data frame. Pandas [78] represents a Python data analyst library for efficient data manipulation. The reference time for the air traffic fingerprint is chosen as the most frequently occurring value from the time column. Subsequently, the list of aircraft is complemented by aircraft that miss the information from the reference time using another timestamp from the defined time frame reduced to \pm 5 seconds, and all other duplicated values of ICAO 24-bit addresses are deleted from the data frame. Thus, each aircraft has only one record in the data frame.

In the next step, the altitude is determined for each aircraft. In other words, geoaltitude is taken as a reference for aircraft in flight. Otherwise, the value of the final altitude value is retrieved from baroaltitude. The elevation of aircraft on the ground is gathered from open source APIs. In case the aircraft parameter on the ground is set at "True" value, the latitude and longitude of the aircraft is taken and the elevation is assigned from the Open Topo Data API [79] or the Open-Elevation API [80]. The choice of the API depends on the availability of data from the APIs. Sometimes, a situation may occur that one of the API does not provide elevation results during the initial data processing. For this reason, the elevation data can be downloaded from one or the other data source using the query containing aircraft latitude and longitude data. The results from both APIs are represented by a JSON object, from which the value of elevation is extracted, as shown in the example below.

In this example, the value 235 m is saved as an elevation for the given aircraft. The dataset SRTM 90m represents a digital elevation database with a resolution of approximately 90 m, originally produced by NASA [81].



If the aircraft in flight provides data only about barometric altitude, the geodetic altitude is recalculated from the average difference of all available barometric and geodetic altitudes related to the percentage difference between current barometric pressure and the pressure calculated according to the International Standard Atmosphere (ISA) [82]. For all not null values in the dataset, the average difference between barometric and geodetic altitudes is estimated as a mean of available altitude differences. In addition, outliers with altitude values over 20,000 m are not included into the average difference calculated using the pressure corresponding to the given altitude according to ISA is calculated using the pressure formula 6.1 [82] for each aircraft with only barometric altitude provided.

$$P = P_b \cdot \left(\frac{T_b + L_b \cdot (h - h_b)}{T_b}\right)^{\left(\frac{-g_0 \cdot M}{R \cdot L_b}\right)};$$
(6.1)

where P(Pa) is calculated pressure, $P_b(Pa)$ reference pressure, $T_b(K)$ reference temperature, h (m) height at which pressure is calculated, $L_b\left(\frac{K}{m}\right)$ temperature lapse rate, $h_b(m)$ height of reference level, R = 8.3144598 $\left(\frac{J}{mol\cdot K}\right)$ universal gas constant, $g_0 =$ 9.80665 $\left(\frac{m}{s^2}\right)$ gravitational acceleration, and M = 8.3144598 $\left(\frac{kg}{m}\right)$ molar mass of Earth's air. In the case of the air traffic data processing, the height is represented by the barometric altitude provided by the aircraft, and the reference height is set at zero. The percentage difference is then calculated as:

$$bar_alt_{diff} = \frac{1-P}{P_b}$$
(6.2)

The resulting difference which is added to the original value of the baro altitude, is determined as a product of the average difference between baro and geo altitudes and the percentage difference bar_alt_{diff} .

The data exploration also showed high differences between values of baro and geo altitudes in some cases or unexpectedly high values. If only geo altitude is available and the value is higher than 20,000 m, the unit of the value is considered in feet, and it is converted from meters to feet. When both values are available and higher than 15,000 m, the difference of baro and geo altitude is calculated. On condition that the values are similar, the geo altitude is kept for further processing. Otherwise, a test is carried out whether one or another value could be given in feet. If it is proven that one value is probably in feet, the geo altitude in meters is taken as the final value. In case both values are higher than 20,000 m, the geo altitude is again considered in feet, and it is converted in feet.



to meters. Consequently, the original barometric altitude values are dropped from the data frame, and only one column with the final altitude is kept for each aircraft. Although the air traffic data processing aims at keeping as much data as possible in case the aircraft in flight does not provide any information about altitude, it is deleted from the data frame.

The data frame is enriched by information about the ICAO code of the aircraft's operator and the aircraft type from the OpenSky Network aircraft database [68]. The data are merged with the created data frame based on the ICAO 24-bit address of the aircraft. Not each aircraft is assigned with this information. Therefore, the missing aircraft types are filled in. In the first step, the aircraft type is estimated for aircraft with a known operator when other aircraft types of the operator are assigned as the most common operator aircraft type. In other words, the aircraft operated by the same airline is assumed to be similar types. On condition that the aircraft is on the ground and no aircraft type was found for it, it is considered to be a ground vehicle at an airport, and it is excluded from the data frame.

For a reason the DME interrogator is later assigned based on the aircraft category according to the ICAO Doc 8643 [69], the dataset is further complemented by from the designators file from OpenSky Network [68] in dependency on the aircraft types. Albeit missing aircraft types are filled in when an aircraft operator has known other aircraft types, some of the aircraft types may still be missing. In order to provide each aircraft with a category for further DME interrogator assignment, the aircraft categories are completed using SVM classification with the linear kernel.

The SVM classification is chosen based on its best performance in the prediction of the description from the tested ML models. The ML model's performance was evaluated on ten different air traffic situations on randomly chosen dates and times between August and December in 2021. Only data with known final descriptions were taken as input to the model testing. The data was divided into the train and test datasets. The train set included three quarters of all data. The performance was evaluated using the k-fold cross-validation method by computing the score ten consecutive times with a different split each time. The resulting values represent the accuracy of the method in particular testing dates. Results of testing the different supervised machine learning models are shown in Table 12.



model / dataset		8/30/ 21 16:40	7/2/ 21 14:29	7/31/ 21 6:45	9/1/ 21 0:32	8/8/ 21 12:15	9/26/ 21 4:55	10/4/ 21 11:15	10/21/ 21 22:28	19/11/ 21 8:24	12/22/ 21 17:40	mean
decision tree		0.8177	0.7767	0.8000	0.6364	0.8587	0.9367	0.8429	0.8644	0.8567	0.8282	0.8218
randoi	random forest		0.7324	0.7406	0.7583	0.7364	0.9093	0.7286	0.8325	0.7551	0.8194	0.7804
SVM	linear	0.8865	0.8398	0.8404	0.8194	0.8891	0.8983	0.8902	0.8963	0.8793	0.8891	0.8793
SVM	rbf	0.8852	0.8422	0.8422	0.8194	0.8901	0.8983	0.8895	0.8909	0.8792	0.8901	0.8792
SVM	sigmoid	0.8504	0.7879	0.7850	0.8194	0.8493	0.8878	0.8758	0.8771	0.8529	0.8493	0.8529
SVM	poly	0.8840	0.8457	0.8457	0.7972	0.8869	0.8921	0.8887	0.8871	0.8761	0.8869	0.8761
KNN		0.8789	0.8252	0.8258	0.8194	0.8842	0.9513	0.8790	0.8983	0.8811	0.8779	0.8721
naïve l	oayes	0.8846	0.8166	0.8166	0.8194	0.8869	0.9480	0.8784	0.8900	0.8887	0.8856	0.8715
logistic regression		0.8827	0.8410	0.8404	0.8194	0.8858	0.9525	0.8808	0.8983	0.8887	0.8925	0.8782
fuzzy decision tree		0.8846	0.8410	0.8404	0.7958	0.8907	0.9536	0.8868	0.8983	0.8826	0.8863	0.8760

 Table 12 - Results of performance evaluation of tested ML models on prediction of AC description

In summary, the initial processing of the air traffic data includes data download from OpenSky Network, calculating or correcting altitude values, filling in missing values, simulating increased traffic, and complementing data by aircraft category. When all procedures are completed, the final data frame with air traffic fingerprint is ready as one of the model inputs. The sample of initially processed air traffic data is shown in Table 13. This initial processing runs every time the model is required to provide results at a different time.

Table 13 -	Data Frame	of Air Traffic Data
------------	------------	---------------------

icao24	lat	lon	on_ground	alt	description
40673b	52.45061	-1.63956	FALSE	5165.94	L4J
3d1d38	53.57629	7.136248	FALSE	1184.37	L1P
3ff9bb	50.41463	8.082226	FALSE	1199.56	L2J
40650b	51.50446	0.047731	FALSE	11548.78	L2J
4b846d	50.03484	8.564638	TRUE	30.36	L4J

5.1.2 Initial Processing of DME Ground Stations Data

The input data to the model are a result of the merge of the different sources of DME ground stations databases, DEMETER [72] and SAFIRE [73]. The data is further compared to AIPs in case the databases provide different information. The last version of the databases was updated at the end of May 2022. That is to mention, the navaids infrastructure does not change very often, and this data is considered sufficient for the thesis purposes. However, it needs to be updated when used later. The initial processing of the DME ground station data is contained in the "DME_init.py" script.



Firstly, data from both databases are loaded into a pandas data frame and only with the same required columns to describe the basic navaid parameters needed for the model. Several data preprocessing steps are carried out. Only service types DME, VOR/DME, ILS/DME, VORTAC, and TACAN are kept in the data frame. Duplicate records of the navaids are dropped in the next step, together with rows where crucial values are unknown. Only data for ECAC states are kept. The data merge is based on the DME identification and channel for each state. In case some data are not contained in both databases, the data from the corresponding AIP is loaded to provide information if the data in one or the other dataset are up to date. This step is carried out to clean the database from outdated data that were not removed from the dataset by a responsible person. All data that are contained in both databases, or where AIP information confirms that the ground station should be in operation, the corresponding parameters are saved into the final database. The antenna height is added to the elevation of ground stations, and just one column with the ground station elevation is kept and converted to meters. The maximum defined DOC of the ground station is converted to hundreds of feet, expressing the flight level. The FOM is assigned to each ground station before the data are saved into a final form.

Initially, the FOM was taken from the DEMETER database. However, it was shown that some assignments do not comply with the DOC values. This may be caused because the data may be inconsistent. For example, almost 10% of data with the assigned service ILS/DME has the FOM value of one and higher, even though the DOC should correspond to the ILS coverage of a radius of 25 NM. The comparison of the model output when considering only DOC or only FOM of the ground station differed significantly. For this reason, it was decided to assign FOM values based on the radius, maximum flight level, and declared EIRP in the database. Moreover, the subgroups are established for ground stations with the assigned FOM value two. These subgroups are created to cover a high number of ground stations in between the DOC radius higher than 40 NM and lower than 130 NM, representing one third of all ground stations and playing a significant role in the load distribution among the ground stations. The FMS is not capable of distinguishing between these subgroups in reality. Nevertheless, in the model, the subgroups may represent the DME ground station that may be tuned based on the FOM value but are not yet in range of the aircraft. In this case, the FMS may stay tuned to the channels interrogating so far. The subgroups are divided based on Figure 20 from Annex10/Volume I, Attachment C [26]. The figure represents an example of a necessary EIRP transmitted by DME ground stations in order to achieve the required power density

59



of -89 dBW/m² at the mid-band frequency; corresponding required power at the antenna of -111 dBW. The airborne received sensitivity is considered -120 dBW. Transmission line loss, mismatch loss, and antenna polar pattern variation is based on an isotropic antenna 9 dB. For different conditions, the EIRP may be necessary to make an appropriate increase, and vice versa, the coverage can reach higher values under suitable conditions [26].

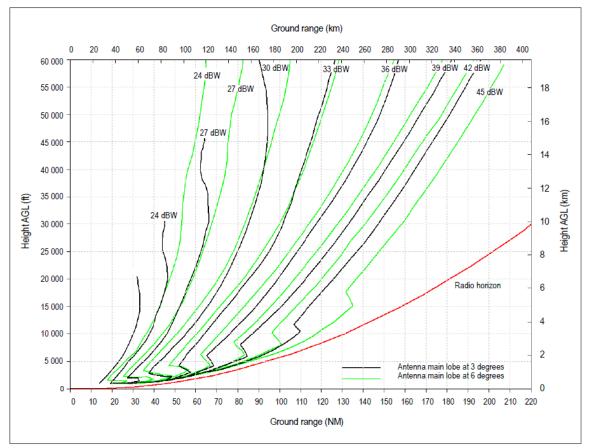


Figure 20 - Necessary EIRP to achieve the minimum power density as a function of height and distance from DME [26]

In case the service of the ground station is ILS/DME, the FOM value is set at zero. Moreover, DME facilities with DOC with a radius lower than 25 NM and maximum FL lower than FL 120 are considered terminal DME ground stations as well. The FOM value one is assigned to all other data with a radius lower than 40 NM or a maximum FL lower than FL 180. Even though the figure would look different for each ground station, it is taken as a reference for setting up the limitation for the range of the DME ground stations which would have the FOM value set at 2. The first subgroup is limited by an EIRP of 24 dBW, or radius of 50 NM, and a maximum FL 300. The second subgroup has an EIRP limitation of 27 dBW, radius 70 NM, and FL 450. The third subgroup under EIRP is limited by an EIRP of 30 dBW or a radius of 100 NM. The last subgroup corresponds to the FOM two value limits.



The DME ground stations with DOC greater than 130 NM are assigned the FOM value three.

Finally, the unnecessary columns for the model are dropped, and all values are saved into a csv file. The processing of the DME ground station does not have to be carried out every time the model is run. Unlike air traffic data, the data stays the same until a database update is provided. In other words, the model can only upload the saved csv file and skip the initial processing step.

cty	service	dme_id	dme_channel	radius	fl_max	fom
AUT	ILS/DME	OEZ	22X	25	100	0
DNK	ILS/DME	SN	24X	25	100	0
FIN	ILS/DME	UT	42X	40	150	1
BIH	VOR/DME	LAK	100X	40	250	2
HRV	VOR/DME	ZDA	23X	100	500	2
I	VOR/TAC	CDC	120X	200	600	3

Table 14 – Sample of Data Frame of DME Ground Stations

limit_radius	limit_fl	eirp	lat	lon	elevation
25	120	27	48.11754	16.58139	194
25	120	29	55.58699	12.13518	51
40	180	36	60.89657	26.95074	102
50	300	29	44.94165	17.29598	46
100	600	37	44.09532	15.36421	93
300	600	39	38.75587	16.36914	1003

In total, the model works with 1960 navaids ensuring coverage in the ECAC Member States. The parameters specified in Table 14 are converted to attributes of the GroundStation Class in the model.

5.1.3 Initial Processing of DME Interrogators Data

The initial data processing of the DME interrogators includes only a transformation of a csv file into the Python environment of the pandas data frame. The data is already prepared using parameters established based on information from FMS experts. The file with DME interrogators types is run through the "FMS_init.py" code to ensure the required input for the model.

5.2 Classes Definition

Once the model input data is preprocessed and cleaned, it is loaded into the predefined classes. Each aircraft and each DME ground station creates an instance of a class,



representing objects in the OOP concept. This step is carried out in the "main.py" file. The definition of classes corresponds to the input datasets. Hence, the model contains the Aircraft, GroundStation, and FMS classes. Moreover, a class Environment is constructed from Aircraft and GroundStation class to ensure a connection between these objects. Besides attributes, each class also has predefined methods used later in the model.

5.2.1 Aircraft Class

The attributes of Aircraft class are defined in the "class_AC.py" code as follows:

```
%python
class Aircraft:
    """ Aircraft object
    .....
    def __init__(self, icao24, lat, lon, on_ground, alt, description):
        self.icao24 = icao24
        self.lat = lat
        self.lon = lon
        self.alt = alt
        self.on ground = on ground
        self.description = description
        self.ac_coordinates = pm.geodetic2ecef(float(lat), float(lon),
(float(alt)))
        self.fms = None
        self.single_dual = None
        self.interrogation_rate = None
        self.ac dme list = []
        self.ac_dme_pairs = []
```

The first attributes are defined directly from the loaded data frame from the initial processing. The coordinates of an aircraft are converted to the ECEF (Earth-centered, Earth-fixed) coordinate system to enable calculations of ranges and angles between objects. A Python library pymap3D [83] is used for the conversion of coordinates and also for other calculations. In addition, the Aircraft class contains attributes that are later filled in with DME interrogator type, information on whether a single or dual interrogator is installed airborne, and the total interrogation rate. The last two attributes are intended to save a list of possible ground stations to tune and of available pairs of ground stations for DME/DME positioning. How these lists are created is explained later in the model description.

The Aircraft class contains only one method, "assign_fms", for assignment of an FMS type based on the description attribute of the object.



5.2.2 GroundStation Class

The GroundStation class is defined by attributes directly from the input dataset in the "class_GS.py" file. The only additional attribute consists of coordinates in ECEF coordinate system.

```
%python
class GroundStation:
    """ DME ground station object
    .....
    def __init__(self, cty, service, dme_id, dme_channel, radius,fl_max, fom,
limit_radius, limit_fl, lat, lon, elev):
        self.cty = cty
        self.service = service
        self.dme_id = dme_id
        self.dme_channel = dme_channel
        self.radius = radius
        self.fl_max = fl_max
        self.fom = fom
        self.limit radius = limit radius
        self.limit_fl = limit_fl
        self.lat = lat
        self.lon = lon
        self.elev = elev
        self.dme_coordinates = pm.geodetic2ecef(float(lat), float(lon),
float(elev))
```

The GroundStation class does not have any method implemented.

5.2.3 FMS Class

The FMS class is defined in the "class_FMS.py" script. At the same time, this script ensures the running of the initial processing of the DME interrogator input data. The FMS class contains the following attributes:

```
%python
class FMS:
    """ DME interrogator object
    """
    def __init__(self, system_id, selection_pool, horizontal_low,
horizontal_up, vertical_limit, max_stations, above):
        self.system_id = system_id
        self.selection_pool = selection_pool
        self.horizontal_low = horizontal_low
        self.horizontal_up = horizontal_up
        self.vertical limit = vertical limit
```



%python

self.max_stations = max_stations

The FMS class represents the tuning logic of the DME ground stations. For that reason, it contains several methods that decide whether the tested aircraft can tune a ground station based on the specified limits. The list of methods and its basic functionality looks as follows:

- check_range testing if the calculated range of an aircraft from a DME ground station is within specified limits representing the maximum coverage radius of a DME ground station;
- check_altitude testing if the altitude of an aircraft is within specified limits representing the maximum vertical coverage of a DME ground station;
- check_vertical_angle testing of the limiting angle between an aircraft and a DME ground station;
- check_los testing of LOS of an aircraft and a DME ground station;
- check_horizontal_angle testing of the limiting subtended angle between an aircraft and a DME/DME pair;

All used methods are further explained when used in the model.

5.2.4 Environment Class

The Environment class contains only two attributes represented by collecting all aircraft and ground stations objects. The Environment class is defined in the "class_envi.py" code.

```
%python
class Environment:
    """ Class for connection of Aircraft with GroundStation to ensure
exchange of information
    """
    def __init__(self, aircrafts, dmes):
        self.aircrafts = aircrafts
        self.dmes = dmes
```

Even though there are only two attributes, the Environment class represents the crucial part of the model thanks to its methods. They ensure the connection between aircraft and DME ground stations. The methods and their main functionalities are named below and later explained in detail when used in the model:

• get_dme_range - creating a list of DME ground stations that an aircraft can tune;



- get_dme_pair creating a list of DME/DME pairs that an aircraft can tune;
- assign_dmes assigning of chosen ground stations to an aircraft, simulation of the DME ground stations tuning;
- get_dme_ac_list assigning all aircraft that are tuned to the ground station;
- get_dme_features creating features for the ML algorithm;
- predict_ppps predicting load based on the applied ML algorithm;
- find_overloaded_dmes finding DME ground stations with load higher than a defined limit;
- retune_overload in case of one or more overloaded ground station exists, retuning of these stations;
- get_final_assignment in case one or more overloaded ground station exists, determining the final load of the DME ground stations.

Data is loaded into the Environment class after the assignment of DME interrogators in each aircraft as described in the following subchapters.

5.3 DME Interrogator Assignment

In the next step, one of the DME interrogator types is assigned to each aircraft. This process is enabled by the "assign_fms" method defined in the Aircraft class. It is called in the "main.py" code. Besides the DME interrogator type, the other data about single or dual interrogators and the interrogation rate are determined here.

The ICAO Doc 8634 [69] description is considered a crucial parameter based on which the type of DME interrogator is chosen for aircraft. Three categories were created from the selected descriptions as follows:

- multi-channel interrogator, doubled L2J, L4J, L2T, L4T, L3J, L6J
- multi-channel interrogator, single L2P, L1T
- single-channel interrogator L1P, L1J, H2T, H1T, H1P

The "multi_channel_dual_fms" category includes descriptions representing airlines where the equipment is expected to interrogate multiple channels with two independent DME interrogators. The "multi_channel_single_fms" category characterizes the representation of larger aircraft used in general aviation that are equipped with only one DME interrogator. The last category, "single_channel_fms", includes smaller aircraft in general and military aviation that may be equipped with a single DME interrogator to measure the DME range only. It is taken into account that the creation of three categories for all aircraft does not ensure that the equipment of the



specific aircraft corresponds to reality. Nevertheless, the information about the DME interrogation type is not available, and the aim of this distinction is to simulate the distribution of the DME interrogation rate in air traffic.

In case the aircraft belongs to the "multi_channel_dual_fms" category, the DME interrogator type is randomly assigned as an "C", "B", or "A" type. Only 2% of aircraft is chosen to be "C" type because the FMS using multi-DME logic occurs rarely. 69% can be assigned as "A" type, where the percentage is expected higher than the "B" type enabling to tune up to five channels with 29%. For all aircraft in this category, the equipment of DME interrogators is expected to double. When using the standard interrogation rate values, the average value for multiple channel interrogators is determined 48 pp/s. Therefore, 96 pp/s may be distributed over the interrogated channels. [45]

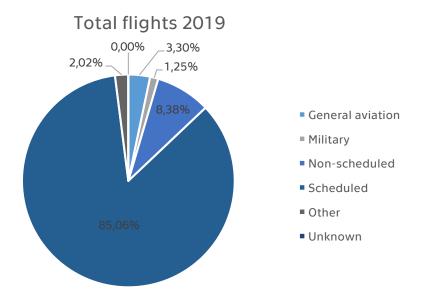
During the model validation, different distribution ratios between interrogator type "A" and "B" were tested. Specifically, two cases were evaluated. Where in the first case, the DME interrogator type "A" was assigned to 59% of aircraft and type "B" to 39%. The second case evaluated a lower representation of type "A" with 79% whereabouts "B" was assigned to 19%. Table 15 shows metrics of the GBR using a train test split results in comparison with real values from 10 known DME ground stations. More details on the method can be found in Chapter 6.

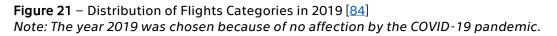
-	ator type oution	correlation coefficient	MSE (pp/s²)	MAE (pp/s)	
type A	type B				
59	39	0.68147	3349.42	35.28	
69	29	0.6967	3347.23	34.84	
79	19	0.68214	3426.91	35.98	

The same types of DME interrogator are assigned similarly for the "multi_channel_single_fms" category. The only difference is the standard interrogation rate which remains 48 pp/s, corresponding to the equipment only with one multiple channel interrogator.



The single channel DME interrogators can interrogate either 16 pp/s or 30 pp/s. As seen from the chart in Figure 21, general aviation and military traffic represents less than 5% of the total traffic. From this number, almost 38% is referred to as military traffic. Based on this information, it is assumed that military aircraft and a small part of the general aviation aircraft are equipped with DME interrogators with a higher transmission rate. For this reason, the interrogation rate of 16 pp/s is assigned to 60% of aircraft in this category, and the rest is assigned with the rate of 30 pp/s. All aircraft are considered to have only one DME interrogator.





In case the aircraft does not correspond to any of the defined categories, the single DME interrogator with a transmission rate of 16 pp/s is assigned to this aircraft.

5.4 Finding Suitable DME Ground Stations

When each aircraft has been assigned one of the DME interrogator types, the Environment class connects aircraft with ground stations complying with all tested conditions, such as range, flight level limit, elevation angle, and LOS. This step is processed using the "get_dme_range" method.

At first, the slant range is calculated from the ECEF coordinates of an aircraft and a ground station based on formula 6.3.

$$SR = \sqrt{(x_{AC} - x_{GS})^2 + (y_{AC} - y_{GS})^2 + (z_{AC} - z_{GS})^2}$$
(6.3)



The ground range is calculated similarly using only x and y coordinates of the aircraft and ground station in the ECEF coordinate system. Both ranges are converted from meters to nautical miles.

In order to avoid a high amount of tested ground stations, the basic slant range limit is limited to 300 NM. If the range is lower or equal to these values, further examination is processed in the FMS class method called "check_range". The calculated ground range, the limit value for radius, as well as the radius of the ground station are loaded to this function. The test is conducted to determine whether the ground range is lower or equal to the radius limit in case the limit exists. When the DME ground station has a FOM value of three, then the limit is taken as the radius of the ground station. If the ground range meets the tested condition, it is considered valid, and the ground station undergoes the next testing. In the opposite case, the ground station is excluded from the condition's examination, and the next ground station is taken into the testing loop. Similarly, the ground station is excluded if it does not meet any of the following conditions.

The following parameter to check is the maximum flight level limit. For this purpose, the method "check_altitude" is called from the FMS class. The aircraft altitude converted to feet, the flight level limit of the ground station, and the flight level limit set in DOC represent inputs of the method. Firstly, the FL values are multiplied by one hundred to reach the value in feet. Afterwards, the aircraft altitude is compared to the maximum flight level limit from FOM or DOC values of the ground station. When the aircraft altitude is lower or equal to the limit, the testing of this ground station continues.

The vertical angle is tested in the "check_vertical_angle" method of the FMS class. The elevation angle is determined using the ecef2aer function from the pymap3d library [83] that calculates azimuth, elevation, and range. The results of the library were verified by comparison of the elevation angle calculation based on formula 6.4:

$$evelation_{angle} = \sin^{-1} \frac{Z_{ENU}}{SR};$$
(6.4)

where the z_{ENU} represents the up coordinate of the aircraft in the ENU (East, North, Up) coordinate system. In case the vertical angle is lower than the limit, the ground station is considered suitable for the last test.

The LOS between the aircraft and the tested ground station is the last condition that has to be met to save the ground station to a possible list of candidate stations to interrogate. The "check_los" method takes inputs of the aircraft altitude, the ground station elevation, and the slant range in meters. The LOS assumes an effective 4/3 Earth



radius. Considering the reference coordinate system WGS 84, both heights are put into the formula 6.5 in meters. Therefore, the resulting distance is in kilometres.

$$LOS = 4.12 \cdot \sqrt{alt_{AC}} + 4.12 \cdot \sqrt{elev_{GS}}$$

$$(6.5)$$

Before the values of the aircraft altitude and the elevation of the ground station are placed in the formula, the negative values are converted to positive values. A negative value can occur when the airport is situated under the mean sea level, such as in the case of the Schiphol airport in Amsterdam. If the negative values were kept, the results would be shortened based on the formula. However, the higher difference in the heights should lead to a higher LOS distance between the transmitter and receiver. On the condition that the determined LOS is lower or equal to the slant range between the aircraft and the ground station, the ground station is considered a candidate for interrogation.

All ground stations that meet the above defined conditions are saved into a list. This list contains the slant range in nautical miles and the object of the ground station. The sorted list from the closest to the furthest ground station is saved into a predefined variable in the Aircraft class, representing a selection pool for assigning DME transponders to the aircraft.

5.5 Finding Suitable DME/DME Pair

The list of ground stations from the previous step creates a basis for finding suitable DME/DME pairs for aircraft. For each aircraft with some DME ground stations in range and a DME interrogator type "B" or "A", a list of DME/DME pairs possible to tune is created. The DME/DME RNAV positioning rules follow the EUROCONTROL Guideline [17]. The slant range of the ground station from the aircraft has to be higher than or equal to 3 NM and 160 NM at maximum, and the facilities coupled with ILS should be excluded as well. If these conditions are met, the aircraft object, together with the list of ground stations, is loaded into a "get_dme_pair" method of the Environment class.

The DME/DME pair testing takes the first ground station from the list of suitable ground stations and examines all combinations with the other stations. In case when the second ground station is also between 3NM and 160 NM, and not coupled with ILS, the distance between the ground stations is calculated. This range is determined in order to avoid testing pairs of the same DME transponders. In other words, if the distance is zero, the loop does not continue further testing and takes the following DME ground station. Although there should not be two DME ground station on the co-channel within the range of 160NM, the ground stations operating on the same channel would be excluded.



The subtended angle between aircraft and the DME/DME pair has to be between 30° and 150°. The angle is determined based on formula 6.6 derived from the law of cosines:

$$\alpha = \cos^{-1} \frac{SR_1^2 + SR_2^2 - R_{12}^2}{2 \cdot SR_1 \cdot SR_2};$$
(6.6)

Where SR_1 is the slant range between aircraft and the first DME transponder of the pair, SR_2 the slant range between aircraft and the second DME transponder of the pair, and R_{12} the range between the first and the second DME transponder, the result is converted to degrees. Whether or not is the angle between the limits is tested in the "check_horizontal_angle" method of the FMS class.

In case the DME/DME pair is in a suitable geometrical position to the aircraft, the uncertainty of the calculation is determined using the following DME RNAV accuracy formula [17]:

$$2\sigma_{DME1/DME2} = 2\frac{\sqrt{\left(\sigma_{DME1,air}^2 + \sigma_{DME1,SIS}^2\right) + \left(\sigma_{DME2,air}^2 + \sigma_{DME2,SIS}^2\right)}}{\sin(\alpha)};$$
(6.7)

where σ_{SIS} is equal to 0.05 NM (or larger value if required – transponders first installed prior to 1989), σ_{air} is the maximum value from {0.085 NM, (0.125% of distance)}, and α is the subtended angle calculated above. The result of the DME RNAV accuracy formula is then compared to the value of Navigation System Error (NSE) 4.33 NM for RNAV 5 that is calculated from the formula 6.8, used for calculation of Total System Error (TSE) [19].

$$TSE^2 = PDE^2 + FTE^2 + NSE^2 \tag{6.8}$$

The PDE (Path Definition Error) is considered negligible, and Flight Technical Error (FTE) 2.5 NM, corresponding the half of the full-scale deflection [85]. In case the DME/DME pair meets the last condition, it is considered a suitable candidate for DME/DME RNAV positioning.

Finally, the appropriate DME/DME pairs are sorted based on the best geometry. In other words, they are sorted according to the results of the DME RNAV accuracy formula. This list of DME/DME pairs is saved into a predefined variable in the Aircraft object, creating a selection pool for DME/DME pair assignment.



5.6 Assignment of DME Ground Station to Aircraft

Each aircraft has an available list of DME ground stations in its range sorted from the closest to the furthest, and aircraft with DME interrogator type "A" or "B" has also saved a list of DME/DME pairs in range sorted based on the better geometry. Consequently, the simulation of interrogating DME ground stations follows, thanks to the "assign_dmes" method in the Environment class. The DME ground station assignment to an aircraft depends on the DME interrogator type of the aircraft.

The easiest solution is for aircraft equipped with one channel DME interrogator. In this case, the closest DME is assigned to the aircraft. The number of interrogating pp/s can be 16 or 30 pp/s determined during the assignment of the DME interrogator. In case the aircraft has a DME ground station in range, the chosen station is saved to a list "dme_only" variable.

The "assign_single" function assigns the ground stations used only to measure slant range, and the "get_one_channel_ppps" calculate the number of the standard value of pp/s for one channel. These functions are used within the "assign_dmes" regardless of the type of the DME interrogator.

When an aircraft is enabled to use a DME interrogator, the "C" type, with multi DME positioning logic, the closest stations from the list of DME ground stations in its range are interrogated. In this case, the maximum possible number of stations is ten. If fewer ground stations is available, the aircraft takes all in the list. The transmission rate is distributed equally among the interrogated channels, and it also depends on whether or not is the number of DME interrogators doubled. According to the standard [29], the maximum number of pp/s per channel is set at 16 pp/s. All DME stations possible to tune are saved into the DME "dme_only".

The DME interrogator type "B", capable to tune up to five channels, provides DME/DME positioning with a complementary DME channel intended for integrity or reasonableness checks. In addition, two channels can be tuned independently for reasons named in Chapter 4.3. The first DME/DME pair is taken from the DME/DME pairs list if the aircraft is in coverage of such a pair. The integrity DME is also assigned from the DME/DME pairs list. It has to create a pair with one of the ground stations from the tuned DME/DME pair. The one with the best geometry from the pair with the integrity DME is considered interrogated. The remaining two channels are taken as the closest ground stations. In other

71



words, one DME channel can be used in the DME/DME pair as well as a single DME for measuring the slant range. An exception creates the integrity DME. If one ground station is considered suitable as an integrity DME as well as a DME station for only the slant range measurement, the aircraft does not take this DME twice. The aircraft then tunes only one other complementary DME for the slant range measurement.

The abovementioned logic works only when a suitable DME/DME pair is in the range of the aircraft. In the opposite case, the DME interrogator tries to find a ground station for DME/DME positioning. Therefore, it interrogates all possible stations in range. In the case of the type "B", it means five different DME channels.

When available ground stations are assigned to this aircraft, the interrogation rate is determined for each channel. When the ground station is tuned in the DME pair, as well as used in the single DME use, it is considered to be interrogated only once. To put it differently, it has the same value of pp/s as the other channels tuned by the DME interrogator. Once the interrogation rate is determined, the DME/DME pair is saved into a "dme_pairs" list and all other tuned ground stations into the "dme_only" variable.

Similarly to DME interrogator type "B", an aircraft equipped with type "A" interrogates DME/DME pair and the corresponding number of DME ground stations for measuring the slant range, based on their availability. The exception represents the fact that this DME interrogator type does not enable to tune the integrity DME.

The variables "dme_only" and "dme_pairs" contain the corresponding Aircraft and GroundStation objects, together with the standard value of pp/s, and the slant range value, representing the output of the "assign_dmes" method. This ensures that each aircraft, as well as each DME ground station, can be investigated independently, and it is considered a great advantage of the OOP concept.

5.7 Assignment of Aircraft to DME Ground Station

As soon as the first allocation of ground stations to aircraft is finished, each ground station is complemented with a list of aircraft that interrogate this ground station. This step is conducted only for the needs of the model, where the created list is essential for retuning when some of the ground stations are overloaded. The list of aircraft for each ground station is also used in the model output.



This list is created for DME ground stations one after the other in the method "get_dme_ac_list" of the Environment class. The input of the function is the output of the DME ground station allocation from the previous step. The content of the allocation is only reorganized so that a sorted list of slant ranges with appropriate aircraft is assigned to the ground stations. The calculation of total pp/s is part of this step, and it is saved into the output. Naturally, this step has to be carried out each time there is a new allocation, or there is a change of the load value. In other words, the method is called with each retuning and prediction of values by the GBR. Then, the output contains the ground station object, the total load, and the corresponding list of aircraft interrogating this station, respectively, the slant range and the aircraft object, as represented in Figure 22.

```
v ac range list: [[14.704494803748, 24.0, <class AC.Aircraft 0...9B52242E0>]...
 > special variables
 > function variables
 > 00: [14.704494803748, 24.0, <class_AC.Aircraft o...9B52242E0>]
 > 01: [16.191112253934953, 24.0, <class AC.Aircraft o...9B51F4FD0>]
 > 02: [21.56398582535168, 24.0, <class AC.Aircraft o...9B515C9A0>]
 > 03: [23.574213591300204, 24.0, <class AC.Aircraft o...9B523E070>]
 > 04: [26.309078761046415, 24.0, <class AC.Aircraft o...9B52CEF10>]
 > 05: [26.917958554116673, 24.0, <class AC.Aircraft o...9B51C4D60>]
 > 06: [53.3291368444381, 19.2, <class AC.Aircraft o...9B51AB790>]
 > 07: [60.52691646037613, 24.0, <class AC.Aircraft o...9B52003A0>]
 > 08: [64.29622642975102, 9.6, <class AC.Aircraft o...9B52B5C70>]
 > 09: [71.37567727379435, 24.0, <class_AC.Aircraft o...9B5262790>]
 > 10: [72.31454717088769, 19.2, <class AC.Aircraft o...9B52E5130>]
 > 11: [80.01356412228448, 9.6, <class AC.Aircraft o...9B51B81C0>]
 > 12: [88.85446949789582, 24.0, <class AC.Aircraft o...9B52CE730>]
   len(): 13
```

Figure 22 - List of slant ranges, pp/s, aircraft objects assigned to VOR/DME AFI in Belgium on 2022-01-15 at 16:00 UTC using the rule-based model only

5.8 Machine Learning Application

The outputs of the estimated load of the rule-based model were considered unsatisfying when compared to real values of pp/s as further described in the validation part. Therefore, a ML model is used in this step to predict the load of en-route DME ground stations.



5.8.1 Machine Learning Model Selection and Training

Initially, the most suitable ML algorithm was chosen, and the features were selected to achieve the best model performance. Among all tested ML algorithms, the Gradient Boosting Regression (GBR) algorithm [47] proved the best performance among other tested. Results of the comparison between the ML algorithms are shown in Table 16. The table represents data with an old test dataset that does not include values of pp/s that were equal to zero. In this case, the performance of Random Forest and Gradient Boosting regression are comparable. The hyperparameters of the model were set at default values of regressors provided by the sci-kit library [52] for comparison.

ML algorithm→ Evaluation Metric↓	Gradient Boosting Regressio n	Random Forest Regressio n	Linear Regression	Polynomia I Regression	Logistic Regression	Linear Forest Regression
correlation coefficient	0.6930	0.6633	0.5529	0.6292	0.4742	0.6564
MSE (pp/s ²)	2623.82	2619.34	4789.25	3369.19	6638.57	2642.67
MAE (pp/s)	34.89	33.89	46.66	40.36	48.36	34.08

 Table 16 - Comparison of different ML algorithms using train/test split 70/30

ML algorithm→ Evaluation Metric↓	SVR Linear Kernel	SVR RBF Kernel	SVR Polynomial Kernel	SVR Sigmoid Kernel	Lasso Regression	RANSAC Regression
correlation coefficient	0.5860	0.6145	0.5318	0.5533	0.5735	0.5218
MSE (pp/s ²)	5099.36	5188.57	5866.97	5630.59	4411.53	5790.35
MAE (pp/s)	41.25	41.62	44.36	44.00	44.59	45.65

The GBR predicts a more precise load when the data are trained on a different ground station. That was the main reason why GBR was decided to be used in the end, as seen in Table 17. Moreover, further testing proved that the GBR reached better performance on a different dataset where also data with real pp/s equal to zero are included.

Table 17 - Comparison of results using RF and GBR when trained o	n others
Table 17 - Companson of results using RF and OBR when trained o	nouners

Random Forest Regression				
correlation coefficient	0.657991			
MSE (pp/s ²)	11164.04			
MAE (pp/s)	61.71689			

Gradient Boosting Regression				
correlation coefficient	0.68091065			
MSE (pp/s ²)	10293.08			
RMSE (pp/s)	59.63565			

Then, the hyperparameters of the GBR were tuned to achieve the best ML model performance using grid search [86]. The final hyperparameters are the following:



- max_depth = 4
- min_samples_leaf = 5
- subsample = 0.5

The ML model is trained on ground stations data provided by an ANSP where the real load of ground stations was known. Data from 10 ground stations with parameters in Table 18 are used for the GBR training. Five hundred values were randomly chosen for each transponder, representing a dataset from two weeks of load measurement. In total, the GBR model is trained on 5000 rows of data. Finally, it was possible to train the ML model and use the trained model parameters to estimate the load that better corresponds to the real data.

service	id	radius	fl_max	fom	limit_radius	limit_fl	eirp
VOR/DME	DME1	90	250	2	100	600	37
VOR/DME	DME2	40	250	2	50	300	37
VOR/DME	DME3	100	500	2	100	600	37
VOR/DME	DME4	40	260	2	50	300	29
VOR/DME	DME5	60	500	2	100	600	37
VOR/DME	DME6	50	250	2	50	300	37
VOR/DME	DME7	40	250	2	50	300	37
VOR/DME	DME8	60	500	2	100	600	37
VOR/DME	DME9	60	500	2	100	600	37
VOR/DME	DME10	60	500	2	100	600	37

 Table 18 - Parameters of DME ground stations used for GBR training

For each DME ground station that has a FOM value different from zero, features for the GBR are created based on the ground station attributes and the rule-based model steps where aircraft are assigned to ground stations. The list of all features is represented in Table 19. They contain time in seconds within 24 hours, radius and flight level declared in DOC, EIRP, number of each type of DME interrogators tuned to the ground station, percentage of aircraft equipped with dual FMS, and number of aircraft within a given distance. The distance categories are determined with the help of the OptBinning Python library [87].

Saving the GBR output parameters and import into the workflow of the model is ensured by the Python pickle library [88]. The model using the same features is trained separately in a "gbr_train_model.py" script and can be improved when new data is available.

Table 19 - Sample of features for ML model

time	doc_radius	doc_fl	eirp	num_S	num_C	num_B	num_A
30020	100	500	37	1	1	3	3



perc_	num_range_	num_range_	num_range	num_range_	
dual	30	60	_120	more	
0.88	8	0	0	0	

5.8.2 Direct Application of GBR

In the workflow, the ML predictions are made just after the list of interrogating aircraft is created for each ground station. Only ground stations with FOM higher than zero are further processed. The terminal ground stations are not considered to be potentially overloaded. Moreover, no data for training the ML model are available for this type of ground station.

The "predict_ppps" method in the Environment class is called, and the initial list with interrogating aircraft for each ground station and the first allocation is loaded together with the time of the air traffic fingerprint in seconds, representing one of the features of the ML algorithm.

The features are created in the first step using the "get_dme_features" method. The method calculates numbers of every of the DME interrogator types, aircraft in a specific range, and the percentage of aircraft equipped with dual FMS that are tuned to the ground station by the rule-based model. In addition, it completes the data with the DOC values for the radius and maximum flight level. Afterwards, the parameters of the GBR previously trained are loaded using the pickle library, and the created features are used to estimate a more precise value of the ground station load.

The goal of the thesis is to evaluate the most critical values for the DME network load. Therefore, for cases where the DME radius is higher than 100 NM, the predicted load by the ML algorithm is used only in cases where it is higher than the load estimated by the rule-based model. This approach was also decided because of the unsatisfying results for GBR predicted values for ground stations that have a higher range than ground stations in the train dataset. For facilities with DOC equal to or lower than 100 NM, the predicted value is always considered as the final load because of the availability of a similar range of ground stations used for the model training.

In case the predicted value by the ML model is taken as the valid result, it is necessary to recalculate the interrogation rate for all aircraft that tuned the ground station in both lists, in the list of all aircraft assigned to the ground station and the list of ground stations interrogated by the aircraft. The recalculation is easily made using the proportion to the



original load value (7) and multiplying this coefficient by the original value. Thus, the proportion of the interrogation rate remains the same as in the rule-based model.

$$recalculation_coeff = \frac{predicted_load}{original_load};$$
(6.9)

Once the final pp/s is determined for all DME ground stations, it is tested whether some station has a higher load than a limit value. In case all values are below the limit, the model skips the step of model retuning and creates directly output files that can be visualized.

5.9 Model Retuning

If the interrogation rate of a DME ground station exceeds the allowed maximum of pp/s, it reduces its sensitivity to reduce the ground station load, respectively number of received pulses. In this case, the weaker signals are ignored, and the most distant aircraft loses its information from the ground station. This process is simulated in the "retune_overload" method of the Environment class. The limit for the maximum load of the ground station is set at 2700 pp/s by default. This value corresponds to DME transponders enabling to serve up to 100 aircraft. However, the currently installed ground facilities are able to provide double the capacity. The load limit can be changed according to the needs.

The "retune_overload" method takes the list of aircraft with assigned ground stations as well as the list of ground stations with assigned aircraft as its input. Firstly, it examines whether some of the ground stations have a higher load than the set-up limit. This is enabled by the "find_overloaded_dmes" function. It tests each DME ground station, and if it is overloaded, it is saved into a list of overloaded DME stations. When one or more ground stations are overloaded, the while loop is created to reduce the load of these stations. It continues running until all DME facilities have the load value below or equal to the limit.

The retuning algorithm takes the list of aircraft of an overloaded DME transponder and removes the last aircraft in the list, representing the most distant aircraft from the ground station. At the same time, the load is reduced by the number of pp/s of the removed aircraft. The furthest aircraft are dropped until the load of the DME ground station is below or equal to the interrogation rate limit. The removed aircraft are saved in a separate variable. For each removed aircraft, the overloaded ground station is deleted from the selection pool of DME ground stations assigned to the aircraft.



all DME/DME pairs containing the overloaded ground station are excluded from the pairs selection pool. This ensures that the interrogation of the overloaded DME transponder is not simulated again. Consequently, all removed aircraft are reassigned with new interrogated DME ground stations. The reassignment is done by repeating the whole assignment process, including the ML model prediction once again. This time, the overloaded ground station is not available in any of the selection pools, and it is no longer possible to interrogate it.

The described process of retuning may cause different distributions of all DME ground stations loads. Therefore, a new list of overloaded DME transponders is created before the next round of DME ground station retuning starts.

A situation occurred when the interrogation rate limit was set too low, and the model did not offer any possibility of replacing the overloaded station with another not overloaded station. In other words, the removed DME ground station excluded from the reassignment was tuned again when other overloaded DME stations were reassigned. In this case, the while loop became infinite. The algorithm was adjusted to eliminate this situation. If the number of overloaded ground stations does not decrease, all overloaded DME ground stations are being removed from the selection pool. Therefore, they cannot be tuned again, and the while loop ends while the reassignment for all aircraft concerned is finished. It follows that aircraft may not tune all the available channels or reach the original number of interrogated stations such as in the initial assignment. This retuning possibility is not usually triggered in the normal operation. For recognition of this process from the standard retuning, the model prints an information about it.

The result of retuning of overloaded DME ground stations can be dependent on the order of stations in the overloaded list. The most distant aircraft is removed from the ground station list to simulate the sensitivity reduction. Simultaneously, this ground station creates a DME/DME pair with a different ground station. In this case, the aircraft is retuned and takes the next DME/DME pair not containing the first overloaded DME. However, the new DME/DME pair does not necessarily contain neither the second DME ground station, which was in the original pair. Therefore, the removal of one ground station also affects the load of other ground stations.

In order to avoid the model dependency on the order of the list with overloaded DME transponders, the Monte Carlo Simulation is carried out. The number of simulations is set at 100 by default. However, the number of loops can be reduced to save time consumption by MCS. Several testing simulations showed that one static situation could



lead to a limited number of solutions on the order of units. The MCS is applied using random shuffling of the order of the list with overloaded DMEs. It is applied in every step, excluding the first retuning where the original order remains.

The final assignment is processed in the "get_final_assignment" method of the Environment class. This function takes all the possible solutions of the MCS simulations as an input. The first result of the MCS is considered the final assignment of DME ground stations to each aircraft. This step is carried out to enable the connection of aircraft and ground stations and their visualization. The final DME load is then calculated as the mean of loads for each assignment in the MCS.

5.10 Model Output

The model output consists of two files. The first file describes the results from the aircraft point of view, and the second file from the ground station point of view. Both results can be adjusted to the needs of a specific phenomenon investigation. The basic files can be used as input into the final visualization of the modelled air traffic situation at a given time. The form of the final files is created within the "get_final_assignment" method.

The file describing the model results from the aircraft perspective includes information about the aircraft, such as its ICAO 24-bit address, coordinates, and the DME interrogator type of the aircraft. In addition, it shows the number of tuned stations. They are divided into a DME/DME pair and the remaining stations, including the integrity DME and other tuned stations that are not used for DME/DME RNAV positioning, complemented by the information about the interrogation rate in pp/s, as seen in Table 20.

icao24	ac_lat	ac_lon	ac_alt	fms_id	num_gs	dme_pair	only_list
48ca63	50.44089	22.38819	2924.05	А	4	['RVN', 19.67, 'UZH', 19.67]	[['RSW', 19.67], ['RZW', 19.67]]
50101e	44.00093	12.41263	3610.36	В	4	['GHE', 16.66, 'ELB', 16.66]	[['RIM', 16.66], ['CEV', 16.66]]
407838	38.04305	23.45527	10111.74	В	5	['SUD', 21.4, 'TRL', 21.4]	[['AML', 21.4], ['TGR', 21.4], ['ATV', 21.4]]
495306	38.86336	-8.37075	7117.08	С	10		[['MOJ', 8.09], ['ARR', 8.09], ['SRA', 8.09], ['ESP', 8.09], ['CAS', 8.09], ['NSA', 8.09], ['FTM', 8.09], ['BEJ', 8.09], ['MTR', 8.09], ['TBC', 8.09]]
3e623e	49.33932	8.945477	802.56	S	1		[['MND', 16.0]]

 Table 20 - Sample of output file with the aircraft list



The output file based on the list of DME ground stations contains information about the ground stations, their load, and the list of aircraft interrogating the specific DME facility with a value of pp/s for each aircraft. The second file is characterized in Table 21.

dme_id	dme_lat	dme_lon	elev	service	radius	fl_max	fom	limit_radius	limit_fl
BEL	54.66114	-6.22991	79	VOR/DME	200	500	3	300	600

Table 21- Sample of output file with the DME ground stations list

eirp	dme_load	ac_list
39	455	[['40712b', 27.911248889540776, 32.0], ['46866b', 28.11139363793055, 32.0], ['4ca86e', 32.97831810210451, 32.0], ['gen485', 34.89641329504079, 20.57], ['4ca293', 44.51419616314169, 5.21], ['gen105', 51.53196119795507, 32.0], ['896210', 84.74867816888408, 24.0], ['3c64c5', 86.0410402770368, 4.82], ['gen606', 108.92915879377576, 8.23], ['407131', 109.64348036063174, 32.0], ['4ca788', 111.95917801206197, 8.23], ['3c4b24', 128.47471969369698, 20.57], ['484416', 130.71477122791788, 12.05], ['gen161', 135.94501344203536, 24.0], ['4ca7b4', 141.4572505037107, 12.05], ['4d227d', 150.7740151944143, 6.64], ['4bb146', 156.08334419248033, 16.07], ['abca08', 167.73140396454548, 2.1], ['406ecc', 181.60054321113012, 24.0], ['aa9321', 201.84122018411082, 24.0], ['ab0e54', 211.0221980436268, 24.0], ['4bb143', 216.03146167457513, 10.61], ['485342', 237.71179508502695, 24.0], ['4b1921', 252.6777356359859, 24.0]]

Both files and the visualization output include the date and time of the air traffic fingerprint in their names.

The output data can be visualized on the map. The visualization of data is optional, and it is enabled by the folium library [89]. This Python library provides possibilities to create an interactive map that represents a suitable graphical user interface of the model output.

The visualization is carried out in the "dme_map.py" script. The input of the script consists of both output files of the model. Therefore, keeping the output consistent and not removing existing columns is necessary, even though some research requirements may include adding new variables.

Icons of DME facilities and an aircraft icon ensure the recognition of two main model elements. These icons are loaded into the visualization module. New classes GroundStation and Plane are explicitly defined for the needs of the vizualization. These classes are filled in with data from the output file. The background map is chosen from the available maps in the folium library and the aircraft and ground stations are placed on the map based on their coordinates. An example of visualization of the aircraft assigned to a ground station and the ground station load is shown in Figure 23.



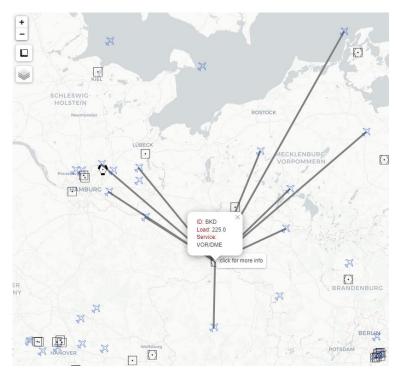


Figure 23 – Example of model output visualization

Moreover, the interactive connection and markers with an object description are enabled by adding this functionality to the visualization. When the user clicks on an aircraft, its ICAO 24-bit address is shown. In case the user clicks on a ground facility, the marker display ID, load, and type of service of this facility. The visualization also enables to use of different colours of the ground station icon to highlight the higher load. The connection is symbolized by lines between aircraft and ground stations and vice versa in case the user mouses over an object.

The visualization script creates an html file that can be opened in any web browser based on user preferences.



6 Model Validation

The model is validated based on the comparison of the results of the model with the real data obtained from ANSP. In the presented figures, the rule-based model results are referred to as original, represented by orange colour. The results with the GBR implementation are referred to as predicted and shown in green colour. The metrics used for validation are taken from the metrics regression section of scikit-learn [54], namely MSE, RMSE, and MAE. The correlation coefficient is also added to the assessed parameters.

6.1 Model Validation Using Train Test Split

The first validation method of the model uses splitting the dataset into random train and subset [90]. The size of the test subsets is set at 30% out of 5000 rows of real data. The random state integer is set at 42 according to the best practices with reference to the Hitchhikers guide to galaxy book [91] with no other significance [92]. For each ground station, 500 rows are randomly selected from a dataset containing data from the first two weeks in March 2022.

For the whole dataset, the features are created adequately to the feature creation process in the model. The target of the model represents the real value of interrogation rate from aircraft in pp/s corresponding to the exact time. Thus, 70% of the data can be used for model training. Once the model is trained, the load is predicted for the remaining 30% of the data.

The data are then divided based on the identification of the ground station, and the results from the model simulation are assigned to the testing data. Afterwards, the validation metrics are calculated for each of the tested ground stations separately. The value is given as the total result of the model validation and the mean of their results is calculated as an average from the intermediate results for individual ground stations. Hence, the square root of the MSE does not have to match the value of RMSE.

Table 22 shows the result of the cases where the ML algorithm was not implemented in the model and the results of the model after the ML algorithm implementation. It can be observed that the difference is significant, especially while considering the value of MSE. The implementation of the GBR brings not only improvement in the model precision but also its results better correlate to the variability of the real values. The prediction error of 46.5 pp/s is acceptable considering the purpose of the model and the DME ground



station transmission rate limit of values 2700 pp/s at the lowest value of the standard and more than 5000 pp/s at modern transponder equipment [23]. This prediction error value is also the lowest achieved among all other ML algorithms tested.

Table 22 - Comparison of performance results for the rule-based and ML model with the realdata using train and test split

Note: The validation dataset also includes rows where the real values correspond to 0 pp/s. Therefore, it differs from the results presented in Subchapter 5.8.1, where different ML models are compared using a dataset without including 0 pp/s values.

Rule-based model only (train/test)				
correlation coefficient	0.5955			
MSE (pp/s ²)	12564.20			
RMSE (pp/s)	97.4102			
MAE (pp/s)	76.0245			

ML model applied (train/test)				
correlation coefficient	0.6967			
MSE (pp/s ²)	3347.23			
RMSE (pp/s)	46.5085			
MAE (pp/s)	34.8411			

For the train and test split the validation results for each DME ground station are located in the Attachment A. Figure 24 visualizes the validation result for one of the tested DMEs. The visualization contains different figures representing comparison of values from one hundred data samples, histogram of the values, and scatter plot with the dependency of the results on the real values. In addition, the values of used metrics are calculated for each ground station.

6.2 Model Validation Using Other Stations for Model Training

The second validation of the model is aimed at evaluating results for ground stations that are not part of the train subset. In other words, having total data from 10 ground stations, the validation is performed using one DME ground station dataset as the test one and the remaining nine DME ground station data as the train dataset to predict the final load of the test station.

In this case, the 4500 rows of training data from all other ground stations with the known real values of load are used for the GBR. The metrics are then used for comparison of the 500 predicted values of the missing ground station in the train dataset. Similarly to the previous validation, the results are also calculated as the mean value from each station's results.



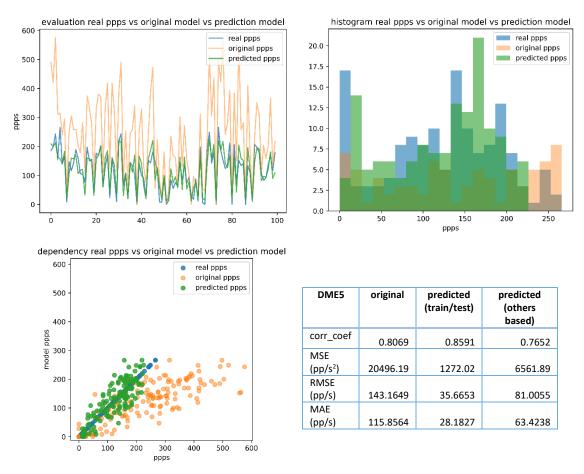


Figure 24 - The validation results include charts of comparison of the results of the rule-based only model and the results of the model when ML applied with the real data on 100 samples, histogram of the results, dependency of the result values, and results of the used matrics for DME5

Table 23 confirms that even though the prediction error increases using this validation method, the performance remains with applying the ML algorithm. The values for each ground station are contained in Attachment A. The total error value is increased by a specific case described below.

Rule-based model only:				
correlation coefficient	0.5955			
MSE (pp/s ²)	12564.20			
RMSE (pp/s)	97.4102			
MAE (pp/s)	76.0245			

ML model applied (others based)				
correlation coefficient	0.6794			
MSE (pp/s ²)	11221.99			
RMSE (pp/s)	76.5689			
MAE (pp/s)	61.6883			



Nevertheless, the validation based on other ground stations showed unsatisfying results for DME3 with DOC 100/500. Considering DOC of other ground stations, the DME3 has the highest range, and there is no corresponding ground station with similar parameters in the train dataset. Figure 25 compares the rule-based model only and the model with ML

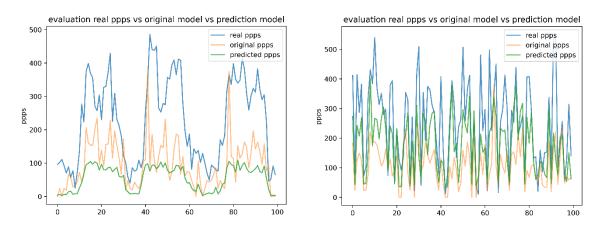


Figure 25 - Comparison of 100 samples of DME3 validation results with testing based on other ground stations (left), and with testing using train test split

implementation with real values for both validation methods and 100 testing samples. Based on the results, it can be concluded that the GBR is not able to predict values for the cases where DOC parameters of the ground station are higher than the ones used in the train dataset. However, when the train dataset contains a ground station with a similar DOC, the predicted load using GBR is more precise than the estimation made by the rule-based model only.

This model validation method resulted into using the predicted load by the GBR only for cases where a ground station with maximum DOC parameters is contained in the train dataset. In this specific case, it represents the value of a radius of 100 NM and a maximum altitude of 50 000 ft. Otherwise, the model selects the higher value of load resulting from the comparison of the rule-based model value and the predicted value.



7 Design of Testing Scenarios

The working hypotheses are defined in Chapter 2.2. The constructed model has to be adjusted to the testing of hypotheses. The testing scenarios require changes in the assignment of the DME interrogator, optimization of DME infrastructure, and simulation of increased air traffic. The methods used for these processes are described below.

7.1 Simulation of Aircraft equipment

The DME interrogator capable of tuning up to five DME channels is represented by the FMS type "B". Two changes are necessary to be made in the model to create the testing scenario where 70% of aircraft are equipped with this interrogator type. Firstly, 70% of all aircraft are randomly chosen in the "main.py" script. Secondly, a new method, "assign_fms_first_hypothesis", is created that assigns only interrogator type "B" in the Aircraft class. Moreover, the aircraft within 70% are considered to be equipped with doubled FMS each with its DME interrogator. The remaining 30% of aircraft are considered equipped according to the common aircraft distribution used in the model. In other words, the total number of aircraft equipped with the DME interrogator capable of tuning up to five stations can overcome 70% because a corresponding part of the remaining aircraft can still be considered equipped with the "B" type interrogator.

7.2 **Optimization of DME infrastructure**

Even though the hypotheses contain the infrastructure optimization according to the SESAR project, a different approach had to be chosen because of the unavailability of the required dataset where the potential DME candidates for decommissioning in Europe would be listed. Therefore, the optimization of DME infrastructure can be simulated by removing the required number of ground stations on a random basis. However, it is important to take into account areas where the density of ground facilities is low, and the decommission could cause a loss of RNAV capability in the surrounding area at lower flight levels. Another key parameter is considered the type of ground station. Whether it is collocated with ILS or VOR, it is a standalone DME, TAC, or VORTAC. The only possible stations that can be candidates for removal are ground stations specified as DME or VOR/DME. The last condition that has to be met for the optimization facilities is the FOM value higher than zero. Thus, only en-route navaids are included. Figure 26 represents the situation where the ground stations are removed randomly, not considering the density of their placement. The circles highlight the ground stations that would most likely stay in operation because their removal could cause operational issues, as



mentioned above. For this reason, the DBSCAN clustering is implemented to enable the removal of stations only in areas with a high density of ground facilities.

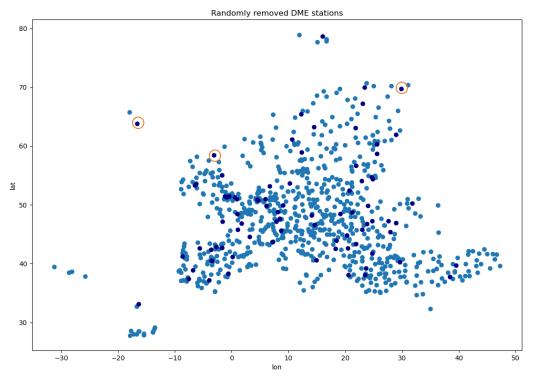


Figure 26 - Example of randomly removed DME stations (darker blue)

The optimization of infrastructure can be run using the "remove_dmes" function of the "ac_data_sim.py" script. This function is automatically run in the model if a "remove_limit" variable is set at a number higher than 0%. Then, the model uses the optimized dataset as an input where the number of DME stations is reduced by the required percentage defined in the "remove_limit" variable.

Only ground stations with service types DME or VOR/DME and FOM value higher or equal to one are chosen from the original dataset representing all DME ground station. In total, 828 ground stations are considered to be en-route navigational aids. The latitude and longitude values are taken from the dataset, and the data are normalized using the StandardScaler function of the scikit-learn library [93]. The NearestNeighbors [94] method was used to find the optimal value of the DBSCAN parameter Eps (ϵ) [95]. The optimal value was determined from the line, representing the distance to the nearest two points for each point, in Figure 27, where the point of maximum curvature is estimated between 0.12 and 0.13. After the testing of DBSCAN clustering, the final value of ϵ was determined 0.12, representing the maximum distance between two points in



the cluster. The minimum number of points in the cluster was set at five to enable finding smaller groups in areas with higher density and exclude noise points.

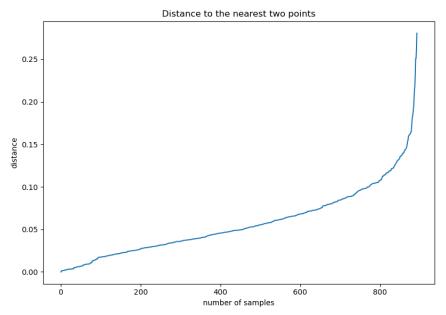


Figure 27 - Estimation of Eps parameter for DBSCAN clustering

When both parameters are set, the clustering algorithm process can start to create groups and mark outliers. The results with clusters highlighted with different colours are drawn in Figure 28. The outliers are represented by orchid colour.

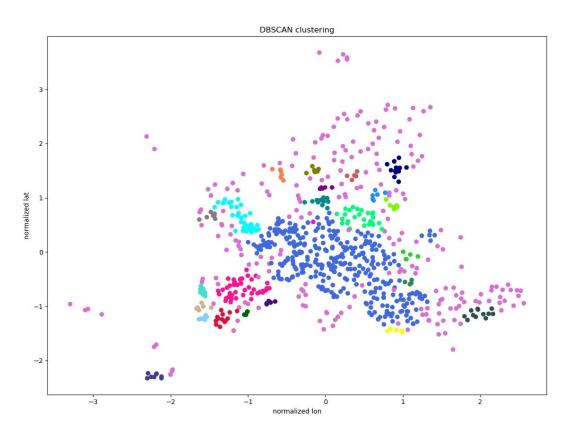


Figure 28 - Clusters of DME ground stations



After the construction of clusters, the candidates for decommissioning are chosen from the clusters of ground stations. In other words, all outliers remain in the dataset untouched. In case the cluster is higher than ten stations, the required percentage is removed from the cluster. This condition is made on the assumption the removal will be in multiples of tens and can be adjusted if necessary. On the condition that the cluster includes fewer stations, it is joined with a different cluster with less than ten stations, and the candidates for decommissioning are chosen from the joined group. Figure 29 characterizes the situation where 10% of ground stations are removed only from the areas with higher density, with a minimum of 5 ground stations close to each other being identified. The chosen stations are dropped from the original dataset of DME ground stations, and the model works with the optimized DME infrastructure. It is important to mention that the stations for decommissioning are chosen randomly within the clusters. Therefore, the final dataset can vary from case to case.

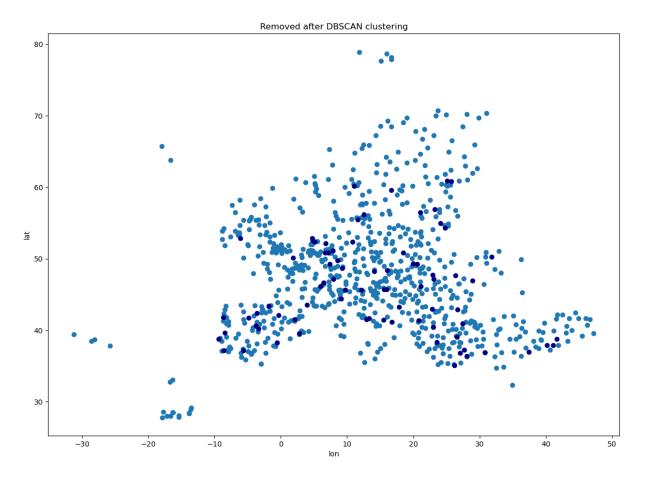


Figure 29 - Stations for decommission after DBSCAN clustering (darker blue)

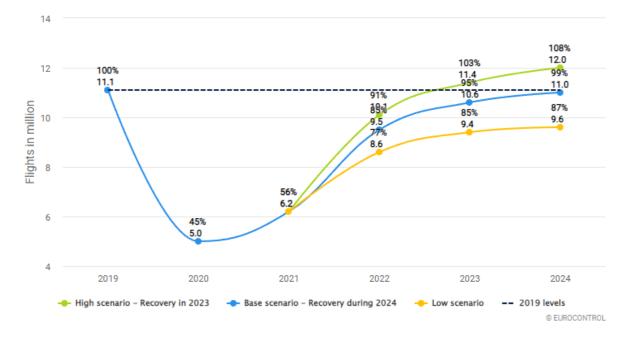
As aforementioned, the optimization of DME stations can be run by changing the "remove_limit" variable to the required percentage of facilities to be removed. For the



thesis purposes, the removal of ground stations was tested for values of 10%, 20%, and 30%.

7.3 Generating Air Traffic Data

It is essential to mention that the working hypotheses were defined before the COVID-19 pandemic. Therefore, the estimated future air traffic cannot reach higher values than in 2019. The IFR movement in 2022 is expected to be approximately 85% of the traffic levels in 2019, even though the traffic is affected by the global economic challenges related to the consequences of the COVID-19 pandemic and invasion of Ukraine, see Figure 30. Based on the forecast, air traffic is expected to recover in 2027 at the latest. [96]. Before the COVID-19 pandemic, the air traffic forecast counted the growth up to EUROCONTROL 3-year forecast for *Europe 2022-2024 Actual and future IFR movements, % traffic compared to 2019



^{*}Europe = ECAC 44 Member States

Figure 30 - IFR movements forecast 2022-2024 [96]

13,635 thousand IFR flight movements in 2025, representing an increase of approximately 19% since 2019 [97], as characterized in Figure 31. Even though the traffic has not yet reached the levels before the crisis, it is assumed that the air traffic will grow and overcome 2019 values.



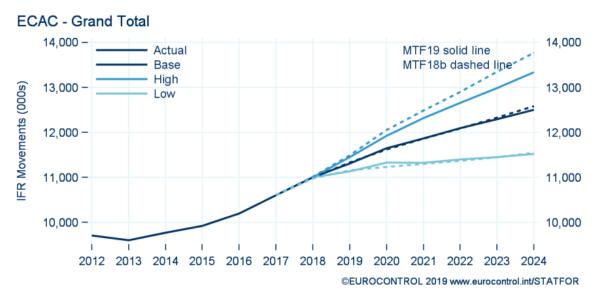


Figure 31 - IFR Movements Forecast from October 2018 [97]

The purpose of the third hypothesis is to examine whether is the optimized DME ground infrastructure able to serve also increased air traffic. Based on the abovementioned development of air traffic, the generated air traffic corresponds to a growth of 10%, 20%, and 30% of the number of aircraft on the tested days.

New aircraft positions were generated using mixture clustering for generating new values of latitude and longitude and inverse transform sampling for altitude values. The generating of new data was created with the support of the Scilab environment [98] in collaboration with the Department of Applied Mathematics of the Czech Technical University in Prague, Faculty of Transportations Sciences, and is processed separately from the model.

The mixtures created from the two-dimensional values of latitude and longitude contain four components. In other words, four clusters are created from the datasets considering the values have a multinomial normal distribution. The mixture clusters are created separately for aircraft indicated in the ADS-B message to be on the ground or in flight. This should ensure generating aircraft on the ground in the location of an airport. The initial centres of the clusters were estimated from a scatter plot for latitude and longitude, and for each position of the aircraft, data is decided on which component it belongs to. Simultaneously, the mixture statistics are updated by multiplying the corresponding data item for the given component. Thus, the resulting values of latitude



and longitude of centres of the clusters are evolving in accordance with the component statistics update considering the weighting factor of individual data items, see Figure 32.

Once all positions are assigned to one of the components, the required number of new values of latitude and longitude may be generated from the mixture component

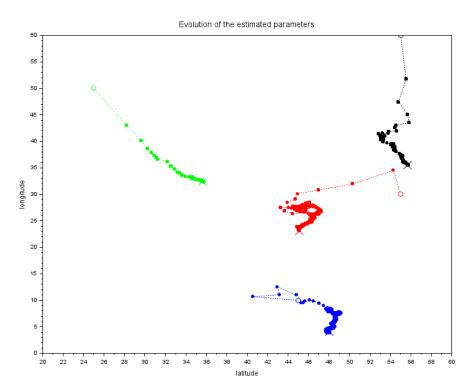


Figure 32 - Evolution of latitude and longitude parameters of individual mixture components

distribution by adding noise to the calculated pointer. The proportion of the number of newly generated data corresponds to the size of the component.

A similar approach was not suitable for generating the altitude values because the probability density function of altitude values differs from the commonly used ones. Therefore, new samples are generated using the inverse function of the Gaussian cumulative distribution. Once the distribution function for the altitude data is made using a cumulative sum, new values can be generated through the inverse distribution function. Identically to the latitude and longitude data generation, the altitude values are generated separately for each of the clusters. For aircraft on the ground, the values of altitude are set at zero by default. The results of the data growth for individual position values are represented in Figure 33.



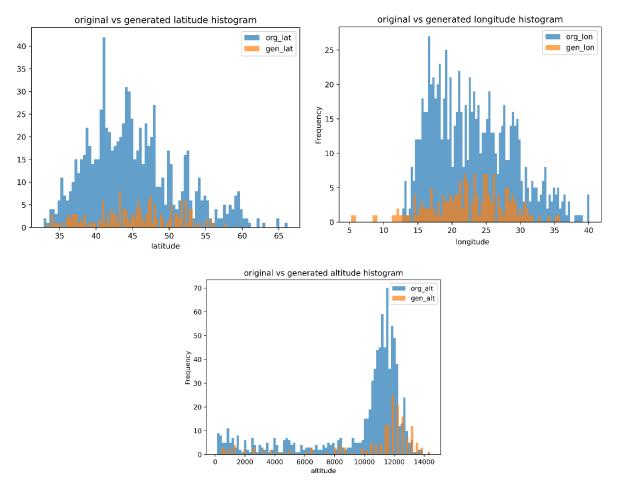


Figure 33 - Histograms of original positions of aircraft (blue) and newly generated sample (orange) with traffic increase of 20% for the second cluster on Friday, June 28, 2019, 11:30 CET

For each generated aircraft position, an ICAO 24-bit address is assigned in the form of "genXXX" where XXX represents the order of the generated aircraft, for example, "genO34". In case the number of generated aircraft exceeds one thousand, the ICAO 24bit address will be in the form of "geXXXX". The description based on the ICAO Doc 8643 [69] is allocated using the same approach when filling in the missing values of the description, so SVM classification is used.

The simulation of air traffic growth does not consider the separation of aircraft, or the different increases forecasted for different parts of European airspace. Generating air traffic, taking into account all the details, is beyond the scope of the thesis, and the current approach is found sufficient for the purpose of evaluating the DME network.



8 Hypotheses Testing Results

All hypotheses are tested using the same reference dates for the air traffic fingerprint. The simulation for each hypothesis is run hundred times because several processes, including assignment of the DME interrogator type or the reduction of DME, are made on a random basis. The results are represented by mean or maximum values of the individual simulation results. In other words, the whole model uses principles of the MCS.

The reference dates for hypotheses testing are taken in accordance with the highest number of flights in the Network Manager Area. The busiest day since now is considered Friday, June 28, 2019 [99]. In the morning peak hour, 11:30 CET (Central European Time), 3151 aircraft were transmitting messages on 1090 MHz captured by OpenSky Network between latitudes of 23° to 82° and longitudes of -35° to 44°. The second testing date for all scenarios is taken from 2022. The busiest day is dated on Friday, July 8, with the total number of 32,392 flights, reaching 87% of June 28, 2019 [100]. However, on July 8, 2022, there were received ADS-B messages from 3580 aircraft at 13:00 CET, which is more than at the busiest day in 2019. It can be caused by better equipment of the aircraft with ADS-B technology as well as better coverage of the OpenSky Network receivers compared to 2019. Therefore, air traffic for these two dates was chosen to test the DME network load.

In order to have a reference for comparison of the hypotheses testing results, the first simulation is run without applying any of the testing scenarios. The values that are evaluated include the following parameters:

- Mean of the estimated load of the ground stations;
- Maximum load from all simulations;
- Number of overloaded ground stations;
- Percentage of cases when the retuning function was run; and
- Occupancy of DME interrogator channels from the total number of channels available.

The results of the model in the basis form, using the predefined distribution of DME interrogator types and not changing number of ground stations or aircraft, are presented in Table 24.



Table 24 - Simulation results with default DME interrogator distribution, no traffic growth, and noremoval of ground stations

Parameters	Load mean (pp/s)	Load max (pp/s)	Number of overloaded Mean	Retuning necessary in % cases	Interrogator channels occupancy %		
Air traffic growth → ↓DME GS reduction	+ 0%						
28.06.2019 11:30							
- 0%	126.27	1220	0	0	97.12		
08.07.2022 13:00							
- 0%	127.71	1455	0	0	97.20		

With the default setting of the model after one hundred simulations, the model results correspond to the mean load of the ground station at around 127 pp/s and the maximum load of 1455 pp/s. The interrogator channel occupancy reaches over 97% of the total capacity.

8.1 First Hypothesis Testing

For the first hypothesis, 70% of aircraft were simulated to be equipped with the DME interrogator that enables to interrogate up to five DME channels. In other words, 70% of aircraft were assigned with DME interrogator "B" and the equipment was considered to be doubled. The remaining part of air traffic was kept in the default distribution used in the model. In this scenario, no air traffic growth and no reduction in the number of ground stations were simulated.

Table 25 shows the result of the simulation. In comparison with the default distribution of DME interrogator equipment, the average load of the ground stations increased by approximately 6 pp/s in both cases. The maximum value of the load reaches the value of 1642 pp/s. Nevertheless, the reference date from 2019 results in a lower maximum value of load than obtained in the reference simulation, which is most likely an impact of the different equipment distribution. The percentage of interrogator channels occupancy decreased slightly. This may be due to the fact that the number of possible channels to tune has increased, and at the same time more DME interrogators have tuned as complementary DME some of the stations already selected in the DME/DME pair.



Table 25 - Simulation results with 70% of aircraft equipped with dual DME interrogator "B", notraffic growth, and no removal of ground stations

Parameters	Load mean (pp/s)	Load max (pp/s)	Number of overloaded Mean	Retuning necessary in % cases	Interrogator channels occupancy %	
Air traffic growth → ↓DME GS reduction	+ 0%					
28.06.2019 11:30						
- 0%	132.3	1176	0	0	97.05	
08.07.2022 13:00						
- 0%	133.6	1642	0	0	97.16	

8.2 Second Hypothesis Testing

In the second scenario, the default DME interrogator distribution is used. To simulate the decommission of ground station, the number of ground stations is reduced with the help of the DBSCAN. The ground stations are removed by 10%, 20%, and 30% only in areas with dense coverage. The air traffic remains at the reference value.

Table 26– Simulation results with no traffic growth, and removal of ground stations by 10%,20% and 30%

Parameters	Load mean (pp/s)	Load max (pp/s)	Number of overloaded Mean	Retuning necessary in % cases	Interrogator channels occupancy %			
Air traffic growth	+ 0%							
→	3151 aircraft							
↓DME GS								
reduction								
	28.06.2019 11:30							
- 10%	132.61	1866	0	0	97.09			
- 20%	133.78	2089	0	0	97.05			
- 30%	135.24	2296	0	0	97.01			
08.07.2022 13:00								
- 10%	135.17	1823	0	0	97.18			
- 20%	135.72	2076	0	0	97.15			
- 30%	136.64	2560	0	0	97.12			

The simulation results prove that the optimized DME network is still able to serve the current air traffic. Table 26 shows that the reduction of DME ground station has a significant impact on the maximum load of ground station. This parameter reaches the



value of 2560 pp/s in one of the simulations with the reduced number of ground stations by 30%. In other words, in none of the cases, the default limit of 2700 pp/s where the retuning process would be run did not occur. As expected, the mean value of pp/s increases with each reduction of DME facilities. Similarly to the previous scenario, the interrogator channel occupancy moved around the value of 97% with a tendency to decrease with each reduction in the number of DME facilities.

8.3 Third Hypothesis Testing

The last simulation uses the reduction of DME ground stations as in the previous case. On top of that, the air traffic increases for both reference days by 10%, 20%, and 30%. It is represented by the highest number of 4096 aircraft on 28 June 2019 and 4654 aircraft on 8 July 2022. Positions of new simulated aircraft are dependent on the situation from the reference date.

Table 27 presents the results of the simulation for each growth of the air traffic in combination with the reduction in the number of DME ground stations for both testing dates. The maximum load value when DME ground stations are reduced by 20% is lower than in the 10% reduction on 28 June 2019. This shows the importance of the selection of the DME facilities for decommissioning. It is assumed that with a reduction of 10%, crucial stations for the infrastructure were simulated as decommissioned. Therefore, the value can reach a higher maximum in one of the simulation cases. In all other cases, the maximum load of DME ground stations increases with the growth of the air traffic. Nevertheless, the increase is more significant when the percentage of removed DME ground stations grows.

The growing number of interrogating aircraft and removal of ground stations have an expected impact on the mean load of the ground stations. However, the value does not change significantly and stays with the biggest difference of 6 pp/s in case of traffic increase in 2019 when 10% of stations are considered decommissioned.

The simulation of air traffic growth by 20% and 30% on the busiest day in 2022 shows that the retuning process of the model can be run in case of removing crucial DME ground stations. In other words, the limit of the load of 2700 pp/s can be reached. In the case of 20%, the limit was reached in five simulations out of one hundred. In one of the cases, two ground stations were considered overloaded, and their sensitivity reduction was simulated. Only one ground station had to be retuned in all the remaining cases. When



Table 27 - Simulation results with traffic growth by 10%, 20% and 30%, and removal of ground stations by 10%	10%, 20% and 30%
---	------------------

Parameters	Load mean (pp/s)	Load max (pp/s)	Number of overloaded Mean	Retuning necessary in % cases	Interrogator channels occupancy %	Load mean (pp/s)	Load max (pp/s)	Number of overloaded Mean	Retuning necessary in % cases	Interrogator channels occupancy %
Air traffic growth 										
	28.06.2019 11:30									
10%	137.50	2125	0	0	97.01	140.85	1986	0	0	96.61
20%	138.93	2006	0	0	96.98	141.61	2520	0	0	96.56
30%	140.69	2472	0	0	96.92	142.82	2559	0	0	96.50
				08.0	07.2022 13:00					
10%	144.90	2182	0	0	96.75	144.82	2365	0	0	96.83
20%	144.73	2345	0	0	96.73	145.00	2509	0	0	96.80
30%	144.91	2586	0	0	96.69	145.46	2671	0.06	5	96.77

Air traffic growth → ↓ DME GS reduction	30%				
	28.06.2019 11:30				
10%	143.39	2057	0	0	96.26
20%	144.30	2682	0	0	96.19
30%	145.49	2679	0.06	6	96.11
		08.07.2	2022 13:00		
10%	145.69	2103	0	0	96.54
20%	146.10	2452	0	0	96.50
30%	146.69	2682	0.06	6	96.45



DME ground stations were reduced by 30%, six simulations out of one hundred ran the retuning process for one DME facility. The same result corresponds to the case of the removal of 30% of stations for the busiest day in 2019.

The channels occupancy decreased with each increase in air traffic and decrease in the number of stations in the network. On both tested dates, it stayed over 96% of tuned channels on both tested dates.

8.4 Hypotheses Conclusions

The above tested scenarios correspond to the definition of the three working hypotheses. Based on the testing results, the hypotheses can be confirmed or refused, as described below.

1. The current DME infrastructure is able to serve the current air traffic on the assumption that 70% of aircraft are equipped with FMS enabling to interrogate up to five DME channels.

Based on the results of the first hypothesis testing, this working hypothesis can be confirmed. Considering the maximum value of the DME ground station load of 2700 pp/s, none of the ground stations reached this value in the model simulations.

2. The optimised DME infrastructure, according to the SESAR project, is able to serve the current air traffic.

The second working hypothesis can be confirmed. The simulation of DME infrastructure optimization decreasing the number of ground stations by 10%, 20%, or 30% did not cause an estimation of any ground station load higher than 2700 pp/s when testing the busiest days in 2019 and 2022.

3. The optimised DME infrastructure, according to the SESAR project, is able to serve the estimated future air traffic.

The future air traffic growth by 10%, 20%, and 30% was estimated from the distribution of the aircraft in the selected days in 2019 and 2022 while decreasing the number of DME ground stations by 10%, 20%, and 30%. The worst case, considering the capability of the DME network to serve future traffic, resulted in the necessity to reduce the sensitivity of one DME ground station in the whole network in 6% of the simulations. Taking into account the randomness of the DME ground stations removal that might be crucial and, in reality, would not be decommissioned, and the default value of 2700 pp/s, which is in



reality much higher, the third working hypothesis can be confirmed. In other words, the optimised DME network can also serve the future air traffic.



9 Research Outcome Summary

The research brings insights not only related to the results of the hypotheses testing but also several interesting facts discovered during the process of the model development and comparison of the estimated values of the model with the real values.

The validation of the rule-based model output showed several surprising facts. When the output of the rule-based model was compared to the real values, the model determined values differed in most of the cases. In some cases, they resulted in a much higher load. In other cases, the model determined values were almost corresponding to or slightly lower than the real values. In total, the number of pp/s of the rule-based model was almost double the total pp/s in real data. In general, it can be assumed that real values do not correspond to values in the technical specification of DME interrogators and may not reach the limit pp/s. Therefore, the ML model was chosen to eliminate this difference. The improvement of the load estimation by implementing GBR is limited by the number of available real data. The real decoded pp/s received from aircraft interrogation was obtained only from one ANSP for ten ground stations. This data was used for the GBR training. Usually, the ANSPs can provide only data about transmitted pp/s by ground stations that include squitter pulse pairs where it is impossible to further distinguish between aircraft interrogation replies and the squitter. After the ML algorithm was applied, the results of the model became comparable to the real values, and the DME network was possible to evaluate.

However, the gap between model output and the real load can have many different causes. One can be that the real DOC of the ground stations may differ from the declared values in data sources. Moreover, the FMS tunes the ground stations based on the FOM. Therefore, in some cases, the ground stations may have a higher range than declared, and the real distribution of pp/s is changed in the whole DME ground network.

In general, it is difficult to estimate and validate load values for DME ground stations with a higher DOC than 100 NM. In reality, the behaviour of such stations may differ from the assumption made during the model construction. These values are also considered the most critical from the network point of view. If real data were available for similar stations, the results of the DME network capability would be considered more reliable. Nevertheless, the current model takes the highest possible value estimated by the rulebased model or predicted by the GBR. As aforementioned, usually, the estimation of the



rule-based model indicated higher values than real ones. Therefore, it is not assumed that it would make radical changes to the evaluation results.

The evaluation of the DME network capability shows that the only case where a DME ground station was considered overloaded occurred when air traffic was increased by 30%, and the number of DME ground stations was reduced by 20% or 30%. In just one case, two ground stations were considered overloaded in the whole DME network. Then, the most distant aircraft interrogating this aircraft had to tune a different ground station. In all other cases, representing at a maximum of 6% of all simulations, only one ground station in the whole DME network overcame the load value of 2700 pp/s.

As expected, the mean values of load increase with each growth of the air traffic and with each reduction in the number of DME ground stations. However, the change in the number of DME ground stations has a more significant impact than the number of aircraft. The maximum load values in the current air traffic can be around 1500 pp/s. Taking into account the RMSE of the model of 46.5 pp/s, this value is still far away from the standard value of 2700 pp/s defined for the DME transponder in 1986.

The increased capability of DME interrogators to tune a higher number of channels can increase the demand for the available ground stations. This should be taken into account when setting the maximum values for interrogation rate when creating new MOPs for DME interrogators.

The simulation of decommissioning of DME ground stations has a significant impact on a load of individual facilities. Even if the air traffic did not change, the removal of 30% of stations caused an increase in the maximal load up to approximately 2500 pp/s in one of the simulations. This highlights the necessity to investigate the possibility of decommissioning each ground station in order to maintain the DME network capability without causing over-interrogation of nearby stations. Unlike the increase in air traffic, whose demand is dependent on many factors related to the socio-economical local and global situations, the decommission of a DME ground station can be analysed, and its feasibility and operational impact can be further investigated. Fortunately, the growth of air traffic has a lower impact on the increase in load of the ground stations.

The percentage of interrogated channels occupancy decreases with the air traffic growth which may be caused by the unsuitable placement of newly generated aircraft positions where no DME ground stations are available. The decrease in values is also observed in reducing the number of ground stations. In this case, it indicates that not all aircraft find



as many ground stations to interrogate as they can, or they often choose a DME ground station from DME/DME pair to be also used as a complementary station. This decrease might be also caused by reducing the number of available facilities for aircraft equipped with the DME interrogator type "C".

In general, the evaluation using the created model proved that the DME network in ECAC states is robust and can serve the increased traffic even if the number of DME ground stations is reduced, also considering the possible error values of the model. Moreover, the transmission rate limit was set at 2700 pp/s. It is assumed that the ground stations with the greatest DOC are the high-capacity DME transponders capable of providing a rate of more than 5000 pp/s.

Considering the role of DME navigation for the GNSS backup, it is not expected the reduction of DME ground stations would reach the simulated 30%. Together with the SESAR study [43], where the number of DME stations were proposed to be lowered just around 5%, this thesis contributes to the possibility of reducing the CNS infrastructure in the European region.



Conclusions

The threat of satellite-based navigation unavailability has to be minimised by ensuring a resilient ATM system. The short-term solution for alternative positioning systems is represented by DME/DME RNAV navigation. The terrestrial navaids infrastructure in Europe corresponds to national ANSP needs for conventional navigation from one navaid to another navaid. Therefore, it offers an opportunity for network rationalisation, taking into account the needs of PBN, resulting in a network, cross-border approach. The thesis addresses the question of the DME network capability to serve current and future air traffic and evaluates the current as well as an optimised DME network in the European region.

In the theoretical part of the thesis, the background of the research is presented. In addition, the current studies are described in connection with the content of this work. The role of DME and its evolution is explained in the context of PBN and its role as the short-term solution for GNSS background. The current technical specification for the DME interrogator and transponder is explained in detail. Moreover, basic information about TACAN is provided. Based on the shortcomings of current studies output and the role of the DME navigation, the working hypotheses of the thesis are defined.

In order to test the established hypotheses, the methodology is proposed, containing the basic architecture for a model construction for evaluating the DME network. The proposed model is based on OOP using the Python programming language. The main entities of the model are represented by aircraft equipped with DME interrogators and DME ground stations and the mutual interaction between them. Other research methods used for the model construction are described, including crucial machine learning methods that help to achieve the goal of the thesis.

The model development is described in detail from the collection of the input data, though all model functions, to the model output. The developed model for DME network evaluation is based on a rule-based model derived from the actual usage of DME in navigation and the provided information about FMS logic aboard. On top of that, it is combined with a prediction using the GBR for a better estimation of the final load of DME transponders. The validation of the model is made using the comparison of the model output with the real data obtained from an ANPS. The process of increasing the number of air traffic and reduction of DME ground stations is described before testing the defined hypotheses.



Finally, the results of testing scenarios can be presented and the working hypotheses are confirmed. The simulation results prove the robustness of the DME network and contribute to addressing the question of whether and how it would be possible to rationalize the DME infrastructure. The European DME infrastructure provides necessary capability to be ground based complementary position source when GNSS RFI occurs.

To sum up, the developed model enables evaluation of the DME network capability considering the reduction of DME ground stations in the infrastructure and the growth of air traffic which fulfils the main objective of the thesis.

The main limitation of the model is considered the FMS types, number of channels, and the logic of the assignment of ground stations to the aircraft are based only on the best available information provided by FMS experts. To improve the rule-based part of the model, the real process of the DME stations selection needs to be known. Similarly, the distribution of the FMS types was estimated based on the discussion with experts and comparison with real measured data. The model does not include the dependency of flight procedures on the usage of specific DME ground stations. It should also be noted that in the context of the model detail, other technical specifications, such as the dead time of the DME transponder or the suppression bus of the DME interrogator installed onboard, are negligible. The terrain is also not considered in the model. Thus, aircraft cannot be interrogating ground stations with terrain obstacles in the real environment.

In order to achieve more precise results from the model, its limitations need to be eliminated. In other words, more data from the real operation would help to better understand the DME ground station load, particularly data from en-route DME with a declared radius of more than 130 NM. Of course, better information about the FMS logic and distribution of different DME interrogator types would ensure a more accurate connection between the aircraft and the ground station. It is essential to realize that the quality of the model results with the implementation of the GBR is dependent on the quality of the rule-based model because the features are created from it. Likewise, it would be interesting to investigate the log of tuned DME ground stations in individual flights. That would bring additional knowledge to FMS logic and the dependency of the station selection on specific flight procedures.

The presented model is a static model considering the current distribution of air traffic at one given moment. However, it is assumed that the selection of the DME channel to interrogate will depend on the previous stations tuned. For this purpose, obtaining the FMSs log is crucial. It would also bring the possibility to consider tuning the ground



stations with respect to time. In case enough data is gained, the dependency of the channel selection could be simulated based on the flight route. It could have a significant impact on the distribution of the load of the ground stations.

The developed model enables the evaluation of the DME network capability with an accuracy represented by RMSE of 46.5 pp/s. This was the best accuracy achieved using the combination of the rule-based model with implementation of the GBR. Moreover, other methods were applied for testing purposes to increase air traffic and reduce the number of DME ground stations. The whole work brings the possibility to analyse the DME network from the individual DME facilities perspective. It provides additional information about the mean values of pp/s and their distribution over the network. Such an analysis would not be possible without the construction of a similar tool.

To conclude, besides evaluating the DME network capability, the model provides a useful tool for validating the possible DME infrastructure optimization. For this purpose, it is necessary to further assess the operational usage of individual ground stations with respect to flight procedures. As shown in the model results, it may cause a decrease in the available number of ground stations. The network evaluation can also be enhanced by using more complex simulations of the increased air traffic. The model can be further adjusted according to the future needs of PBN requirements using DME transponders.



References

- [1] Commission Implementing Regulation (EU) No 2018/1048 of 18 July 2018 laying down airspace usage requirements and operating procedures concerning performance-based navigation. *Official Journal L 189/3*, 2018. URL: <u>https://eur-lex.europa.eu/legal-content/EN/TXT/PDF/?uri=CELEX:32018R1048&from=EN</u>.
- [2] Think Paper #9 1 March 2021. EUROCONTROL, 2021. URL: https://www.EUROCONTROL.int/sites/default/files/2021-03/EUROCONTROL-think-paper-9-radio-frequency-intereference-satellite-navigation.pdf
- [3] Duchet, D. and Berz G. GNSS RFI Mitigation: International Efforts to Protect Aviation. EUROCONTROL. *Presentation at the 58th Civil GPS Service Interface Committee Meeting, Miami*, 24 September 2018. URL: <u>https://www.gps.gov/cgsic/meetings/2018/duchet.pdf</u>
- [4] ICAO Assembly 40th Session: Technical Commission. WP/188: An Urgent Need to Address Harmful Interferences to GNSS. URL: <u>https://www.icao.int/Meetings/a40/Documents/WP/wp_188_en.pdf</u>
- [5] Vitan V. GNSS Compatibility and Interferences. EUROCONTROL. *Joint FMG CM 2019/RAFT15*. 2019.
- [6] ICAO Assembly 40th Session: Technical Commission. WP/082: Interference-Resilient Satellite-Based CNS Systems. URL: <u>https://www.icao.int/Meetings/a40/Documents/WP/wp_082_en.pdf</u>
- [7] ICAO State letter: Ref.: AN 7/62.1.4-21/41. ICAO, 2021.
- [8] SESAR JU: Alternative Position, Navigation and Timing (A-PNT). URL: https://www.sesarju.eu/sesar-solutions/alternative-position-navigation-and-timing-pnt. [Accessed 2019-12-17].
- [9] Vitan, V., Berz G., Saini L., Arethens J-P., Belabbas B. And Hotmar P. Research on alternative positioning navigation and timing in Europe. 2018 Integrated Communications, Navigation, Surveillance Conference (ICNS). IEEE, 2018, 2018, 4D2-1-4D2-17. doi: 10.1109/ICNSURV.2018.8384887. ISBN 978-1-5386-5679-2. URL: https://ieeexplore.ieee.org/document/8384887/
- [10] European ATM Master Plan Executive view: Digitalising Europe's Aviation Infrastructure. SESAR Joint Undertaking, 2020. URL: <u>https://www.sesarju.eu/masterplan2020</u>
- [11] Lawrence, D. FAA Navigation Administration Programs Update. FAA, Civil GPS Service Interface Committee, 2020. URL: https://www.gps.gov/cgsic/meetings/2017/lawrence.pdf
- [12] Performance Based Navigation Strategy 2016. FAA, U.S. Department of transportation, 2016. URL: <u>https://www.faa.gov/nextgen/media/pbn_nas_nav.pdf</u>
- [13] ICAO Doc 4444: Procedures for Air Navigation Services: Air Traffic Management. 16th ed. 2016.
- [14] Communications, navigation and surveillance: At the heart of the air traffic management network. EUROCONTROL. URL: <u>https://www.EUROCONTROL.int/communications-</u> navigation-and-surveillance. [Accessed: 2020-05-08].
- [15] Dead Reckoning (DR) *SKYbrary*. URL: <u>https://skybrary.aero/articles/dead-reckoning-dr</u>. [Accessed: 2021-10-16].
- [16] Kelly, Robert J. and Cusick D.R. Distance Measuring Equipment and its Evolving Role in Aviation. Advances in Electronics and Electron Physics Volume 68. Elsevier, 1986, 1986, s. 1-243. ISBN 9780120146680. doi:10.1016/S0065-2539(08)60854-9
- [17] EUROCONTROL Guidelines for RNAV 1 Infrastructure Assessment. 2021, 2nd ed. URL: https://www.EUROCONTROL.int/sites/default/files/2021-07/EUROCONTROL-guidelinesrnav-1-infrastructure-assessment-20.pdf



- [18] SESAR PJ.14-03-04 Short Term APNT Solution TS-IRS: PJ14 EECNS Essential and Efficient Communication Navigation and Surveillance Integrated System. 2019.
- [19] ICAO Doc 9613: Performance-based Navigation (PBN) Manual. 2013, 4th ed.
- [20] Advanced Flight Management Computer System: ARINC Characteristic 702A-3. 2551 RIVA ROAD, ANNAPOLIS, MARYLAND 21401-7435: AERONAUTICAL RADIO, 2006.
- [21] CASPER Noise Lab: Performance-Based Navigation (PBN). URL: https://syv.noiselab.casper.aero/content/1/procedures_101/1. [Accessed: 2022-01-15]
- [22] European Airspace Concept Handbook for PBN Implementation: PBN HANDBOOK No. 1. EUROCONTROL, 2021, 4th ed. URL: <u>https://www.eurocontrol.int/publication/airspace-concept-handbook-implementation-performance-based-navigation-pbn</u>
- [23] EUROCAE ED-57: Minimum Operational Performance Specification for Distance Measuring Equipment (DME/N and DME/P) (Ground Equipment). 1986, 2nd ed.
- [24] European GNSS Contingency/ Reversion Handbook for PBN Operations: PBN Handbook No. 6. EUROCONTROL, 2020, 1st ed. URL: <u>https://www.EUROCONTROL.int/sites/default/files/2021-10/euroconontrol-gnsscontingency-reversion-handbook-6.pdf</u>
- [25] European Navaid Infrastructure Planning Handbook including Minimum Operational Network (MON): PBN Handbook No. 4. EUROCONTROL, 2021, 1st ed. URL: <u>https://www.eurocontrol.int/publication/european-navaid-infrastructure-planning-handbook-including-minimum-operational-network</u>
- [26] ICAO Annex 10/Volume I: Aeronautical Telecommunications Radio Navigation Aids. ICAO, 2018, 7th ed.
- [27] CNS infrastructure evolution opportunities: An economic assessment on behalf of the CNS Advisory Group. EUROCONTROL, 2021, 1st ed. URL: https://www.EUROCONTROL.int/publication/cns-infrastructure-evolution-opportunities
- [28] EUROCAE: WG-107 / DME Infrastructure supporting PBN Positioning. URL: https://www.eurocae.net/about-us/working-groups/. [Accessed: 2021-10-16].
- [29] RTCA DO-189: Minimum Operational Performance Standards for Airborne Distance Measuring Equipment (DME) Operating within the radio Frequency Range of 960-1215 Megahertz.1985.
- [30] Testing of DME/TACAN Ground Stations: Application Note. *Rohde & Schwarz*, 2017. URL: https://www.rohde-schwarz.com/cz/applications/testing-of-dme-tacan-groundstations-application-note_56280-126537.html. [Accessed: 2021-10-17].
- [31] Basic principles of DME. International Institute of Aviation Safety and Technology, 2017. URL: <u>http://iiast.com/</u>. [Accessed: 2021-10-17].
- [32] Specification Sheet: DME Antenna, OMNI Directional Model d Bs 5100A. *dB Systems*, 2005. URL: <u>https://www.dbsant.com/antenna-type/dme-antenna/</u>. [Accessed: 2021-10-17].
- [33] Marquard, S. Mode N A new Navigation System Concept & A-PNT for Aviation. DFS Deutsche Flugsicherung, 2018.
- [34] Berz G., Vitan V., Saini L. And Spanner M. Alternative PNT: What comes after DME?. Research paper at the 20th International Flight Inspection Symposium (IFIS), Monterey 16 – 20 April 2018.
- [35] LDACS: Future aeronautical communications. URL: <u>https://www.ldacs.com/</u>. [Accessed: 2021-10-23].
- [36] L-band digital aeronautical communication. EUROCONTROL. URL: https://www.EUROCONTROL.int/system/l-band-digital-aeronautical-communicationsystem. [Accessed: 2021-10-23].



- [37] Advisory Circular: U.S. Department of Transportation Circular. Subject: Airworthiness Approval of Date: 3/28/14 AC No: 20-138D Positioning and Navigation Systems. FAA. *Federal Aviation Administration*, 3/28/14 (AC 20-138D). 2014.
- [38] RNAV Training for ATC. *Air Traffic Control Association Japan*, 2007. URL: https://www.icao.int/safety/pbn/Documentation/States/Japan_RNAV%20Training%20fo r%20ATC.pdf
- [39] ARINC Specification 424-20: Navigation System Database. AERONAUTICAL RADIO, INC. 2011.
- [40] Vitan, V., Berz G. And Solomina N. Assessment of current DME performance and the potential to support a future A-PNT solution. 2015 IEEE/AIAA 34th Digital Avionics Systems Conference (DASC). IEEE, 2015, 2A2-1-2A2-18. doi: 10.1109/DASC.2015.7311357. ISBN 978-1-4799-8940-9. URL: http://ieeexplore.ieee.org/document/7311357/
- [41] EUR PBN TF. Guidance Material: Subject: aircraft and operator approval for RNP 4 operations. 2013, rev 1.0. URL: https://www.icao.int/EURNAT/Other%20Meetings%20Seminars%20and%20Workshops/P BN%20TF/PBNTF%20-%20EUROCONTROL%20RAISG1/App%20B%20Doc029%20%20RNP%204-EN.pdf
- [42] Vitan V. Research on A-PNT in Europe. SESAR JU presentation to the PNT Advisory Board by: Osechas O. at 2018 iCNS Conference, April 2018. URL: <u>https://www.gps.gov/governance/advisory/meetings/2018-12/osechas.pdf</u>
- [43] SESAR 15.03.02 D06. Navigation Infrastructure Rationalisation: D1.4.B ECAC-Wide Navigation Infrastructure Assessment. 2013, ed 00.01.00.
- [44] WP017: Update on reducing DME congestion. 28th Meeting ICAO Frequency Management Group (FMG/28) combined with 22nd Meeting Network Manager Radio Frequency Function (RAFT/22), 2022.
- [45] EUROCONTROL meeting: Evaluation of DME network for ensuring RNAV capability in the European region by GNSS unavailability. Meeting with Gerhard Berz and Valeriu Vitan on 7th February 2020 in EUROCONTROL, Brussels, Belgium.
- [46] Dusty, P. Python 3 Object Oriented Programming. Birmingham, UK: Packt Publishing Ltd., 2010. ISBN 978-1-849511-26-1.
- [47] Friedman, J.H. Greedy function approximation: a gradient boosting machine. *Annals of statistics*, 2001,1189–232.
- [48] Sharma, G. 5 Regression Algorithms you should know Introductory Guide!. *Analytics Vidhya*, 2021. URL: <u>https://www.analyticsvidhya.com/blog/2021/05/5-regression-algorithms-you-should-know-introductory-guide/</u>. [Accessed: 2022-06-04].
- [49] Ohri, A. 10 Popular Regression Algorithms in Machine Learning Of 2022. Jigsaw Academy, 2022. URL: <u>https://www.jigsawacademy.com/popular-regression-algorithms-ml/</u>. [Accessed: 2022-06-04].
- [50] Masui, T. All You Need to Know about Gradient Boosting Algorithm Part 1. Regression: Algorithm explained with an example, math, and code. *Towards Data Science*, 2020. URL: <u>https://towardsdatascience.com/all-you-need-to-know-about-gradient-boosting-algorithm-part-1-regression-2520a34a502</u>. [Accessed: 2022-06-20].
- [51] Dhiraj, K. Implementing Gradient Boosting in Python. 2019. URL: https://blog.paperspace.com/implementing-gradient-boosting-regression-python/. [Accessed: 2022-06-20].
- [52] Gradient Boosting for regression: Sklearn.ensemble.GradientBoostingRegressor. Scikitlearn. URL: <u>https://scikit-</u> learn.org/stable/modules/generated/sklearn.ensemble.GradientBoostingRegressor.html #sklearn.ensemble.GradientBoostingRegressor. [Accessed: 2022-06-20].



- [53] Wu, S. 3 Best metrics to evaluate Regression Model?: R Square, Adjusted R Square, MSE, RMSE, MAE. *Towards Data Science*, 2020. URL: <u>https://towardsdatascience.com/what-are-the-best-metrics-to-evaluate-your-regression-model-418ca481755b</u>. [Accessed: 2022-06-20].
- [54] Base classes and utility functions: Sklearn.base. *Scikit-learn*. URL: <u>https://scikit-learn.org/stable/modules/classes.html#module-sklearn.metrics</u>. [Accessed: 2022-06-20].
- [55] Pedregosa, F. et al. Scikit-learn: Machine learning in Python. *Journal of machine learning research*, 12 October 2011, pp.2825–2830.
- [56] Gong, D. Top 6 Machine Learning Algorithms for Classification: How to Build a Machine Learning Model Pipeline in Python. *Towards Data Science*, 2022. URL: <u>https://towardsdatascience.com/top-machine-learning-algorithms-for-classification-2197870ff501</u>. [Accessed: 2022-03-19].
- [57] Cross-validation: evaluating estimator performance. *Scikit-learn*. URL: <u>https://scikit-learn.org/stable/modules/cross_validation.html#cross-validation</u>. Accessed: 2022-03-19].
- [58] Vapnik V. The nature of statistical learning theory. New York: Springer, 1999, 2nd ed.
- [59] Ma, Yuxin, Wei Chen, Xiaohong Ma, Jiayi Xu, Xinxin Huang, Ross Maciejewski and Anthony K. H. Tung. EasySVM: A visual analysis approach for open-box support vector machines. *Computational Visual Media*. 2017, 3(2), 161-175. ISSN 2096-0433. doi:10.1007/s41095-017-0077-5
- [60] Kumar, V. So why the heck are they called Support Vector Machines?. Towards Data Science, 2020. URL: <u>https://towardsdatascience.com/so-why-the-heck-are-they-called-support-vector-machines-52fc72c990a1</u>. [Accessed: 2022-03-19].
- [61] Ester, M. et al. A density-based algorithm for discovering clusters in large spatial databases with noise. 1996, pp. 226–231.
- [62] Bhattacharyya, S. DBSCAN Algorithm: Complete Guide and Application with Python Scikit-Learn: Clustering Spatial Database. *Towards Data Science*, 2019. URL: https://towardsdatascience.com/dbscan-algorithm-complete-guide-and-applicationwith-python-scikit-learn-d690cbae4c5d. [Accessed: 2022-02-13].
- [63] Uglickich, E. A I. Nagy. Mathematical Methods for Data Analysis. Faculty of Transportation Sciences Czech Technical University, Prague, 2022. URL: <u>https://www.fd.cvut.cz/personal/nagyivan/MatMetAnalDat/TextbookMMAD.pdf</u>
- [64] Nagy I., Suzdaleva E., Karny M., Mlynarova T. Bayesian estimation of dynamic nite mixtures, International Journal of Adaptive Control and Signal Processing vol. 25, 9, p. 765-787. 2011.
- [65] Nguyen, K. How to generate Gaussian samples: Part 1: Inverse transform sampling. MTI Technology, 2020. URL: <u>https://medium.com/mti-technology/how-to-generate-gaussian-samples-347c391b7959</u>. [Accessed: 2022-01-15].
- [66] Moufad, B. Deriving the inverse transform sampling method from first principles: Insights on generating samples from a custom probability density function. *Towards Data Science*, 2021. URL: <u>https://towardsdatascience.com/an-insight-on-generating-samples-from-acustom-probability-density-function-d0a06c290c54</u>. [Accessed: 2022-01-15].
- [67] Aslanyan, T. K. Monte Carlo Simulation and Variants with Python: Your Guide to Monte Carlo Simulation and Must Know Statistical Sampling Techniques With Python Implementation. *Towards Data Science*, 2021. URL: <u>https://towardsdatascience.com/monte-carlo-simulation-and-variants-with-python-43e3e7c59e1f</u>. [Accessed: 2020-03-06].
- [68] The OpenSky Network: Free ADS-B and Mode S data for research. URL: <u>https://opensky-network.org/</u>. [Accessed: 2020-06-02].



- [69] ICAO Doc 8643: Aircraft Type Designators. ICAO. URL: https://www.icao.int/publications/doc8643/Pages/default.aspx [Accessed: 2020-02-07]
- [70] Commission Implementing Regulation (EU) No 1207/2011 of 22 November 2011 laying down requirements for the performance and the interoperability of surveillance for the single European sky. *OJ L 305*, p. 35. 2011.
- [71] ADS-B Equipage: Automatic dependent surveillance broadcast airborne equipage monitoring. EUROCONTROL. URL: <u>https://www.EUROCONTROL.int/service/adsb-equipage</u>. [Accessed: 2020-06-02].
- [72] DEMETER: Distance measuring equipment tracer. EUROCONTROL. URL: https://www.EUROCONTROL.int/online-tool/distance-measuring-equipment-tracer. [Accessed: 2020-06-02].
- [73] SAFIRE: Central register of European radio frequency assignments and the European frequency request coordination system. EUROCONTROL. URL: <u>https://www.EUROCONTROL.int/online-tool/central-register-european-radio-frequencyassignments-and-european-frequency-request</u>. [Accessed: 2020-06-02].
- [74] GitLab: The One DevOps Platform. URL: <u>https://about.gitlab.com/</u>
- [75] Harris, C.R., Millman, K.J., van der Walt, S.J. et al. Array programming with NumPy. *Nature* 585, 357–362, 2020. doi: 10.1038/s41586-020-2649-2
- [76] Hunter, J.D. Matplotlib: A 2D graphics environment. *Computing in science & engineering*, 9(3), pp.90–95. 2007.
- [77] Pyopensky: Python interface for OpenSky historical database. *GitHub*. URL: <u>https://github.com/junzis/pyopensky</u>. [Accessed: 2020-06-02].
- [78] McKinney, W. et al. Data structures for statistical computing in python. *In Proceedings of the 9th Python in Science Conference*, pp. 51–56, 2010.
- [79] Open Topo Data: Open Topo Data is a REST API server for your elevation data. URL: https://www.opentopodata.org/. [Accessed: 2021-07-28].
- [80] Open-Elevation: A free and open-source elevation API. URL: <u>https://open-elevation.com/</u>. [Accessed: 2021-07-28].
- [81] SRTM 90m DEM Digital Elevation Database. URL: <u>https://srtm.csi.cgiar.org/</u>. [Accessed: 2021-07-28].
- [82] U.S. Standard Atmosphere. Washington, D.C.: National Oceanic and Atmospheric Administration, National Aeronautics and Space Administration, United States Air Force, 1976. URL: <u>https://www.ngdc.noaa.gov/stp/space-weather/online-</u> publications/miscellaneous/us-standard-atmosphere-1976/us-standardatmosphere_st76-1562_noaa.pdf
- [83] Module pymap3d. URL: <u>https://geospace-code.github.io/pymap3d/</u>. [Accessed: 2021-08-01].
- [84] Aircraft communication, navigation and surveillance dashboard: Monitoring fleet capabilities and preparing PBN deployment plans. EUROCONTROL. URL: <u>https://www.EUROCONTROL.int/dashboard/communication-navigation-and-surveillancedashboard</u>. [Accessed: 2020-06-14].
- [85] Navigation Infrastructure Assessment in Support of PBN. ICAO. URL: https://www.icao.int/safety/pbn/PBN%20references/Navigation%20Infrastructure%20As sessment%20in%20Support%20of%20PBN.pdf. [Accessed: 2022-03-05].
- [86] Exhaustive search over specified parameter values for an estimator: Sklearn.model_selection.GridSearchCV. Scikit-learn. URL: <u>https://medium.com/mti-technology/how-to-generate-gaussian-samples-347c391b7959</u>. [Accessed: 2022-07-04].
- [87] Navas-Palencia, G. OptBinning: The Python Optimal Binning library. URL: http://gnpalencia.org/optbinning/. [Accessed: 2022-07-04].

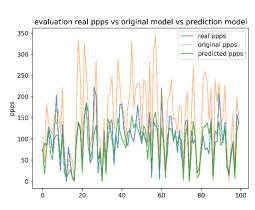


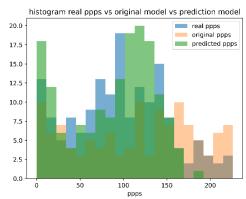
- [88] Van Rossum G. The Python Library Reference, release 3.8.2. Python Software Foundation. 2020.
- [89] Python-visualization, 2020. *Folium*. URL: <u>https://python-visualization.github.io/folium/</u>. [Accessed: 2021-12-14].
- [90] Split arrays or matrices into random train and test subsets: Sklearn.model_selection.train_test_split:. *Scikit-learn*. URL: <u>https://scikit-learn.org/stable/modules/generated/sklearn.model_selection.train_test_split.html</u>. [Accessed: 2022-02-12].
- [91] Adams, D. The Hitchhiker's Guide to the Galaxy. London: Del Rey, 1995. ISBN -10: 0345391802.
- [92] Sahagian, G. What is Random State? And Why is it Always 42?: An inquiry into what the number 42 has to do with machine learning and the universe. *MTI Technology*. URL: https://grsahagian.medium.com/what-is-random-state-42-d803402ee76b. [Accessed: 2022-02-12].
- [93] Standardize features by removing the mean and scaling to unit variance: Sklearn.preprocessing.StandardScaler. Scikit-learn. URL: <u>https://scikit-learn.org/stable/modules/generated/sklearn.preprocessing.StandardScaler.html</u>. [Accessed: 2022-02-12].
- [94] Nearest Neighbors. *Scikit-learn*. URL: <u>https://scikit-</u> learn.org/stable/modules/neighbors.html. [Accessed: 2022-02-13].
- [95] Maklin, C. DBSCAN Python Example: The Optimal Value For Epsilon (EPS). Towards Data Science.URL: https://towardsdatascience.com/machine-learning-clustering-dbscandetermine-the-optimal-value-for-epsilon-eps-python-example-3100091cfbc. [Accessed: 2022-02-13].
- [96] EUROCONTROL Forecast Update 2022-2024: Recovery from COVID-19 and Russian invasion of Ukraine. EUROCONTROL. URL: https://www.EUROCONTROL.int/publication/EUROCONTROL-forecast-update-2022-2024#:~:text=EUROCONTROL%20has%20released%20a%20short,Ukraine%20and%20glo bal%20economic%20challenges. [Accessed: 2022-09-15].
- [97] EUROCONTROL Seven-Year Forecast February 2019: Flight Movements and Service Units 2019-2025. EUROCONTROL, 2019 [Accessed: 2022-10-04]. URL: <u>http://www.atceuc.org/uploads/docs/EUROCONTROL-7-year-forecast-february-2019-main-report.pdf</u>
- [98] Scilab. [Accessed: 2022-10-04]. URL: <u>https://www.scilab.org/</u>. [Accessed: 2022-01-16].
- [99] New traffic record set: 37,228 flights in one day: New traffic record set: 37,228 flights handled across the European aviation network. EUROCONTROL, 2019. URL: <u>https://www.EUROCONTROL.int/news/new-traffic-record-set-37228-flights-one-day</u>. [Accessed: 2022-06-04].
- [100] Daily Traffic Variation States. Aviation Intelligence. EUROCONTROL. URL: <u>https://www.EUROCONTROL.int/Economics/DailyTrafficVariation-States.html</u>. [Accessed: 2022-09-15].

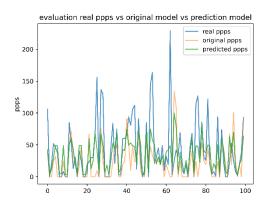


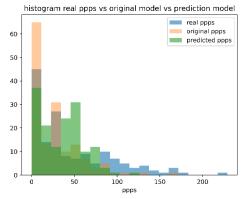
Attachment A – Model Validation Results

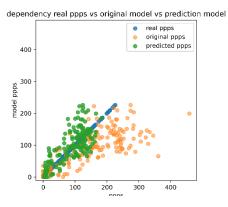
DME1



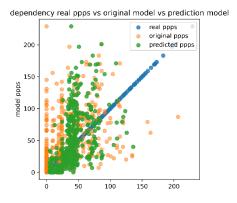








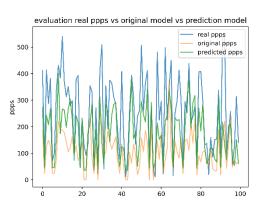
DME1	original	predicted (train/test)	predicted(others based)
corr_coef	0.7280	0.7570	0.7357
MSE	7739.72	1397.93	2101.66
RMSE	87.9757	37.3889	45.8439
MAE	66.8264	28.5685	34.3712

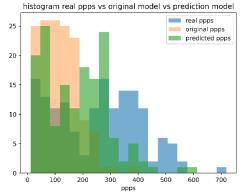


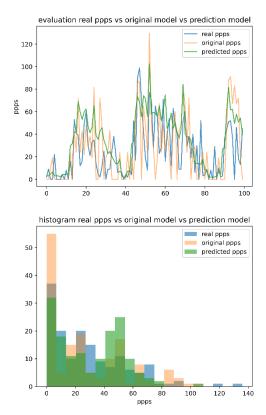
DME2	original	predicted (train/test)	predicted(others based)
corr_coef	0.4497	0.6664	0.5861
MSE	1962.56	1407.06	1364.89
RMSE	44.3008	37.5108	36.9444
MAE	29.7356	24.4217	25.2967



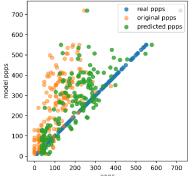




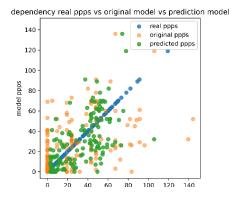




dependency real ppps vs original model vs prediction model



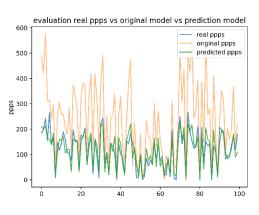
DME3	original	predicted (train/test)	predicted(others based)
corr_coef	0.6742	0.7793	0.8085
MSE	34947.48	14326.61	51352.03
RMSE	186.9425	119.6938	226.6099
MAE	146.7604	85.9323	186.5686

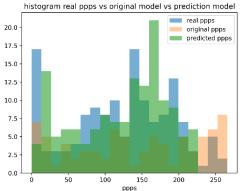


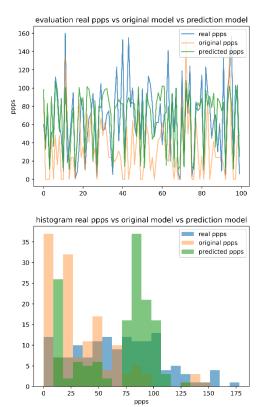
DME4	original	predicted (train/test)	predicted(others based)
corr_coef	0.4667	0.6480	0.6393
MSE	976.73	424.43	551.82
RMSE	31.2526	20.6017	23.4908
MAE	22.0268	15.7989	16.4772

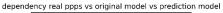


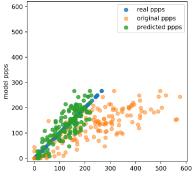




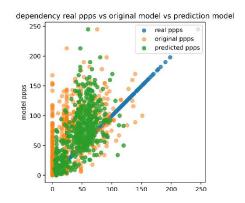








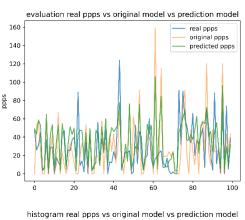
DME5	original	predicted (train/test)	predicted(others based)
corr_coef	0.8069	0.8591	0.7652
MSE	20496.19	1272.02	6561.89
RMSE	143.1649	35.6653	81.0055
MAE	115.8564	28.1827	63.4238

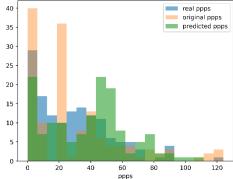


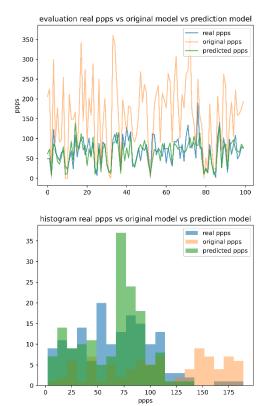
DME6	original	predicted (train/test)	predicted(others based)
corr_coef	0.4902	0.6588	0.6654
MSE	2784.93	933.63	1755.34
RMSE	52.7724	30.5553	41.8968
MAE	40.1228	23.4587	30.4642



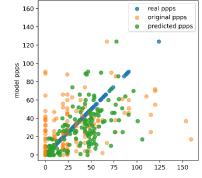




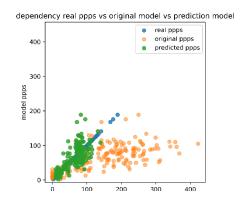




dependency real ppps vs original model vs prediction model



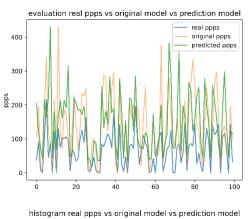
DME7	original	predicted (train/test)	predicted(others based)
corr_coef	0.4055	0.5920	0.6067
MSE	1022.48	585.51	1067.63
RMSE	31.9762	24.1974	32.6746
MAE	23.2228	18.3460	25.3394

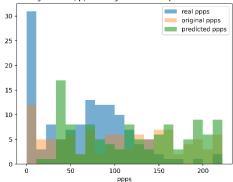


DME8	original	predicted (train/test)	predicted(others based)
corr_coef	0.6492	0.7172	0.6366
MSE	14754.60	671.42	975.28
RMSE	121.4685	25.9117	31.2295
MAE	97.1848	18.8177	23.4952

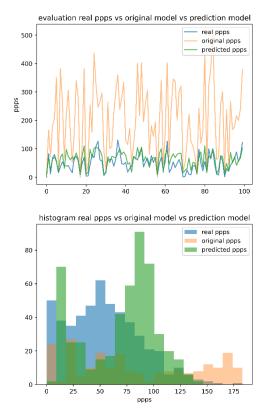


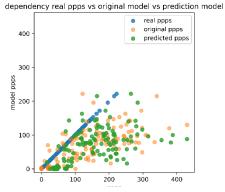




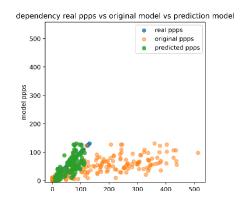








DME9	original	predicted (train/test)	predicted(others based)
corr_coef	0.6599	0.5680	0.7036
MSE	9252.65	11841.52	45406.65
RMSE	96.1907	108.8188	213.0884
MAE	71.9680	85.5105	185.3718



DME10	original	predicted (train/test)	predicted(others based)
corr_coef	0.6250	0.7212	0.6477
MSE	31704.71	612.13	1082.76
RMSE	178.0582	24.7413	32.9053
MAE	146.5408	19.3735	26.0758