Fifteenth USA/Europe Air Traffic Management Research and Development Seminar (ATM2023)

Performance Characterization of Arrival Operations with Point Merge at Oslo Gardermoen Airport

Henrik Hardell
Communications and Transport Systems, ITN,
Linköping University (LiU), Norrköping, Sweden
Flight Procedure Design Group, Luftfartsverket (LFV),
Norrköping, Sweden, firstname.lastname@liu.se

Anastasia Lemetti, Tatiana Polishchuk and Lucie Smetanová
Communications and Transport Systems, ITN,
Linköping University (LiU)
Norrköping, Sweden
firstname.lastname@liu.se

***Abstract*—The paper focuses on the performance assessment of the arrival operations in Oslo Gardermoen airport implementing point merge (PM) procedures. We take a data-driven approach based on the open-source ADS-B data, and conduct a detailed performance assessment utilizing a diverse set of performance indicators, including newly developed metrics for better understanding of the PM specifics. The results of the performance evaluation indicate that the PM systems are currently underutilized in Oslo airport, and their increased usage may lead to the improved arrival performance, especially during the peak time periods.**

***Keywords*—arrival procedures, point merge, performance evaluation, continuous descent operations**

I. INTRODUCTION

PM is a systematised method for sequencing arrival flows developed by the EUROCONTROL Experimental Centre (EEC) in 2006, with the main purpose to facilitate greener arrivals, including Continuous Descent Operations (CDOs) and noise reduction [1]–[3]. Since then many airports around the globe applied the PM systems, including both smaller airports with a single runway and busy ones operating with multiple runways. Sweden follows the trend and considers PM implementation in several airports in the nearest future.

There is a number of variants of the PM implementation, i.e. with overlapping, partially overlapping or separated sequencing legs; position inside or outside Terminal Maneuvering Area (TMA); different geometry of the flows to PM or merging to a point, which depends on the design goals for each particular airport. Each design variant brings a set of benefits and drawbacks, contributing to the known trade-offs of the PM procedures. While providing unquestionable advantages in terms of improved controllability (reduced controllers workload, quantified in the number of instructions given to the pilots [4]–[6]) and better opportunities for greener descents, the PM usage often increases the distance flown by aircraft within TMA, and results in longer time aircraft spend in TMA. From the environmental perspective, on the one hand, the position of the sequencing legs on high flight level in general brings the procedures higher and, thus, higher from the ground; but on the other hand, higher concentration of the flights in the same designated area may lead to increased noise levels, which can be avoided by positioning the PM sequencing legs over the sea or over low-populated areas.

In this paper, we take a data-driven approach, targeting comprehensive investigation of the arrival flight performance resulting from the specifics of the PM implementation in Oslo Gardermoen airport, where PM is successfully operating since 2011. We conduct a detailed performance assessment of the arrival performance with PM procedures, applying a set of performance indicators previously proposed for efficient characterization of the arrival procedures, including the time-related metrics, horizontal and vertical efficiency, sequencing and metering for the arrival flows, and environmental efficiency quantified in additional fuel burn. We complement the set with new metrics for evaluation of the PM utilization. The work contributes to better understanding of the effect of the PM implementation on the arrival performance.

The rest of the paper is organized as follows. In Section III we overview the research related to the evaluation of arrival operations with PM systems. In Section IV we give the details about Oslo Gardermoen airport, describe the datasets and the performance metrics for evaluation of the airport performance. The results of the performance assessment are presented in Section V. Section VI concludes the paper.

II. THE POINT MERGE CONCEPT

The PM method was introduced by the Eurocontrol Innovation Hub in 2006, and since then has been implemented at several airports around the World (e.g., Oslo, Dublin, Istanbul, Tokyo). As of May 2023, there are 38 airports in 19 countries that have applied the PM procedure [7]. PM is a technique that simplifies merging of arrival-traffic flows and provides better opportunities for environmentally-friendly descents and increased terminal airspace capacity [8], and consists of arcs located at equidistant from a merge point (Figure 1). The sequencing legs may be either fully overlapping, partially overlapping or fully dissociated. The different design alternatives provide for a trade-off between capacity and environmental efficiency. In the case of fully or partially overlapping legs, the two opposite arcs are typically flown at level-flight and are vertically separated by 1000

Supported by the Swedish Transport Administration (Trafikverket) and in-kind participation of LFV within the ODESTA, ODESTA-PM and TMAKPI projects.

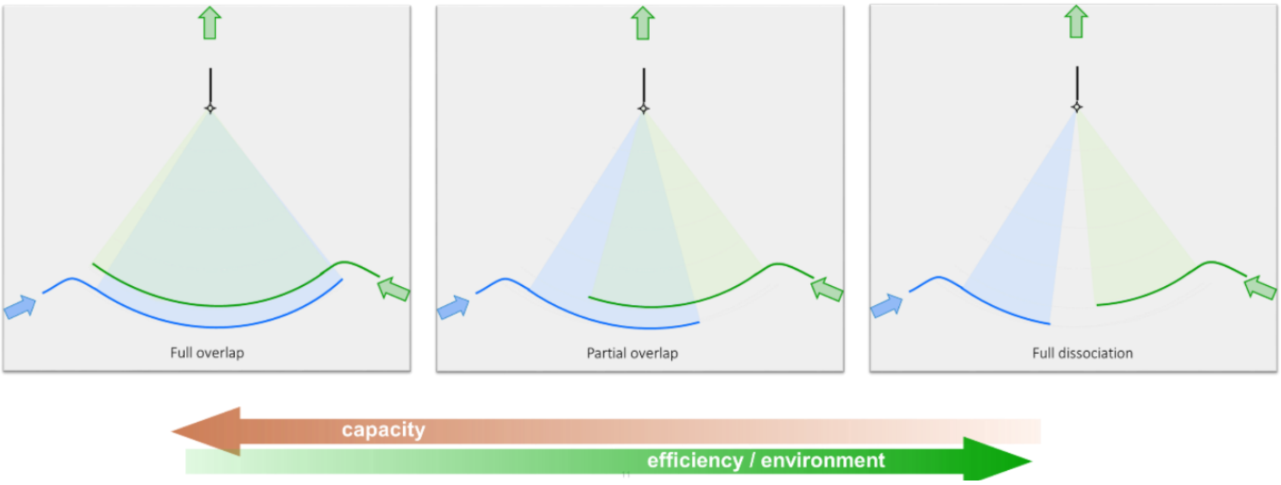


Figure 1: Main design variants for the PM systems with the associated trade-offs (Source: Eurocontrol [7]).

ft. The aircraft follow the sequencing legs until the desired separation with the preceding aircraft is obtained, and are then instructed by the air traffic controller (ATCO) to turn direct to the merge point. An airport may operate with either one, or multiple point merge systems (PMS) that are used simultaneously, serving traffic from different directions. In the latter case, traffic from different PMS may need to merge in an additional point if the aircraft are intending to land on the same runway.

III. RELATED WORK

Researchers investigated the PM systems from different perspectives. Several studies [4]–[6], [9], [10] evaluated the potential benefits provided by PM implementation in simulation environments. The authors of [5] presented the results of the fast-time simulations where they compared arrival operations with vectoring against the PM. The results showed the PM model reduces: mean controller task load (20±1%), the number of instructions to pilots (∼30%), and fuel consumption (170±14 kg), compared with vectoring.

Similarly, the real-time simulations for Istanbul International Ataturk Airport with 50 arrivals per hour [4] demonstrated that the total average number of instructions is about 33 percent less and the frequency occupancy is about 37 percent less for PM than for vectoring. In addition, in terms of trajectory dispersion, in PM, traffic is within a narrower triangular area, while in vectoring large traffic dispersion occurs. Better controllability and predictability of the trajectories gained with PM operations were reported during real-time simulation and validation exercises in Paris-Orly airport [10], as well as in the simulations at Beijing Capital International airport [6].

The authors of [11] proposed a data-driven computer vision approach for identification of the PM patterns in the large datasets containing historical flight tracks.

The studies of the PM usage are not limited to TMAs. In [12], the authors analysed how delays and fuel consumption can be reduced in the en-route sectors implementing PM.

A performance evaluation framework has been developed by the Performance Review Unit (PRU) of EUROCONTROL to characterise the arrival management process [13], [14]. EUROCONTROL Innovation Hub (formerly Experimental Centre) constantly works on investigation of the new metrics for better understanding of the reasons for performance inefficiencies within TMAs [15]–[18]. In addition, the authors of [19] suggested a set of metrics for comprehensive assessment of the arrival flight performance in TMA, and tested it on three European airports implementing different arrival procedures such as vectoring, trombone and PM. The proposed metrics help to assess the arrival operations and help to identify the areas of inefficiencies.

The majority of previous works were based on the simulated or artificially-generated data. Different authors applied various performance indications to evaluate the PM at each specific airport. To the best of our knowledge, there were no published attempts to comprehensively evaluate the overall performance of an airport operating with PM, focusing on characterization of PM design specifics, using open-access data. To fill in the gap, in this work we take a data-driven approach and characterize the performance of Oslo Gardermoen Airport operating with PM using ADS-B based historical data. We apply the methodology proposed in the aforementioned papers for arrival sequence in Oslo airport TMA and extend it with new metrics targeting better understanding of the specifics of the PM utilization.

IV. PERFORMANCE EVALUATION OF OSLO GARDERMOEN AIRPORT

In this section, we describe the Oslo Gardermoen airport, the datasets we created for the study, and review the metrics for evaluation of the airport performance, including the new ones developed in this work for the assessment of the PM utilization.

A. *Oslo Gardermoen Airport*

Oslo Gardermoen airport is the busiest airport in Norway, serving almost 28 million passengers and handling 255.000

aircraft movements per year (2019) [20]. The airport has two parallel runways (01/19 L/R) that are used in either segregated or mixed mode, and operates with PM procedures since its introduction in 2011. There are four PM systems in total (01 East and West, and 19 East and West), which are schematically illustrated in Figure 2, with overlapping sequencing legs at FL90, FL100 and FL110 in each system. Figure 3 shows the published PM procedures of the eastern system for the runways 01L/R. The merge point, located between 18 and 19 NM from the sequencing legs, of each system coincides with the Initial Approach Fix (IAF) of the Instrument Landing System (ILS) procedure, for landing on either the left or the right runway. The PM systems are provided with traffic from six main TMA entry points.

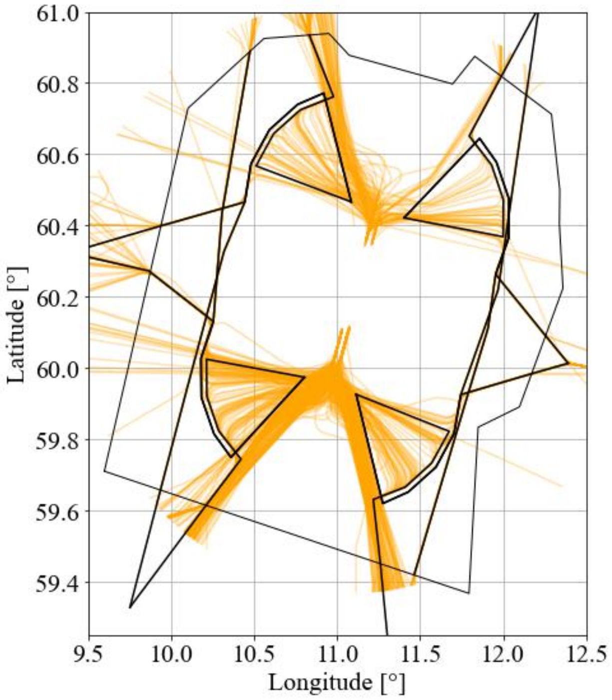


Figure 2: PM system layout at Oslo Gardermoen airport with trajectories of flights performing PM arrival procedure during one week in October 2019.

B. Data

For the aircraft tracking information we use the historical database of the OpenSky Network [22]. We use aircraft state vectors for every second of the trajectories within 50 NM circle around Oslo Gardermoen airport (ENGM) with the center between the runways. Note, that we refer to this circle area as TMA. The applicability of this type of data for performance assessment purposes is justified in [23]. The procedure of cleaning and preprocessing of the data is described in [19].

We consider the year 2019 and select four full weeks in October with 7829 arrival flights. Then we construct the following three datasets: TT (based on time in TMA), PM (Point Merge) and $nonPM$ (not in PM dataset), and split each dataset into the northern and southern parts, according to the landing direction (TT_{North}, TT_{South}, PM_{North}, PM_{South}, $nonPM_{North}$ and $nonPM_{South}$), that is flights from the $North$ subsets are landing from the north on the runways 19L and 19R, the southern part consists of flights landing from the south on the runways 01L and 01R.

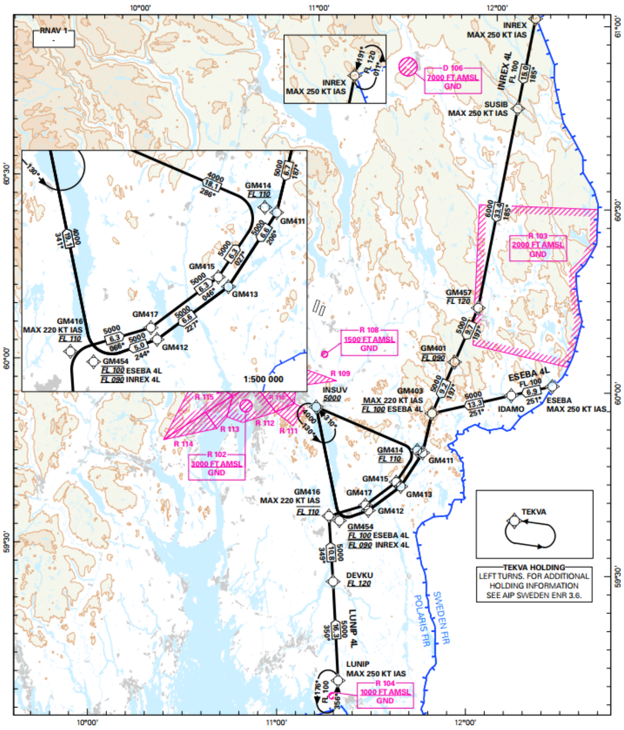


Figure 3: Parts of the published chart for PM procedures at Oslo Gardermoen airport, for the eastern system to runways 01L/R (Source: Norwegian AIP [21]).

TT Dataset. TT dataset is based on the Time in TMA KPI, and consists of the flights from the peak time periods. We calculated average per hour Time in TMA and removed 0.7 percentile from this set of values. The rest of the values correspond to the hours when aircraft spent significantly long time in TMA on average, and therefore represent the peak time periods. TT dataset consists of the flights arrived during these hours.

PM Dataset. PM dataset contains all the arrivals to Oslo airport during October 2019 which performed PM procedure. We use a circle catch area of about 3 NM around the starting points of the PM sequencing legs and filter out all aircraft which did not pass through any of these areas. We consider the following way-points from North-West PM to South-East: *GM429, GM432, GM418, GM423, GM405, GM410, GM416, and GM411*. Figure 4 illustrates the technique used for North-East PM system. The red circle catch areas on the edges are created around $GM418$ and $GM423$ waypoints which are the beginnings of PM legs. Colored curves in the figure illustrate the example flight trajectories performing PM procedure captured by the proposed technique.

NonPM Dataset. $NonPM$ dataset consists of all flights which are not included into PM dataset, that is the flights from this dataset do not perform PM procedure. The union of

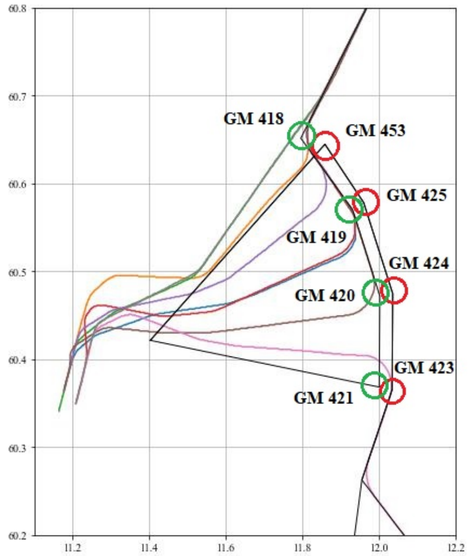


Figure 4: North-Eastern part of the ENGM PM system with catch areas used for identification of the flights following the PM procedure.

PM and $nonPM$ datasets is the whole set of our considered flights.

Table I shows the number of flights in each dataset and its corresponding subsets. Please, note that the datasets overlap, and 38 % of flights from the TT dataset follow the PM procedures and therefore belong to the PM dataset, and 62 % are in the $nonPM$ dataset.

TABLE I. TT, PM and $nonPM$ datasets

Dataset	# flights	% nonlevel flights
TT	3141	31
TT_{North}	1047	40
TT_{South}	2094	26
PM	2262	44
PM_{North}	681	66
PM_{South}	1581	34
$nonPM$	5567	49
$nonPM_{North}$	2683	63
$nonPM_{South}$	2884	35

We compare the direction of landing for all considered flights with the wind direction during the corresponding hours. Wind data are taken from the ECMWF ERA5 reanalysis dataset provided via the C3S Data Store [24] (calculated from 100m u- and v-components of the wind near the runways). For comparison we split all wind directions into two parts, northern and southern, turned by 16 degrees clockwise (the slope of the runways relative to the north direction). With these assumptions 85 % of all our considered flights are landing facing the wind.

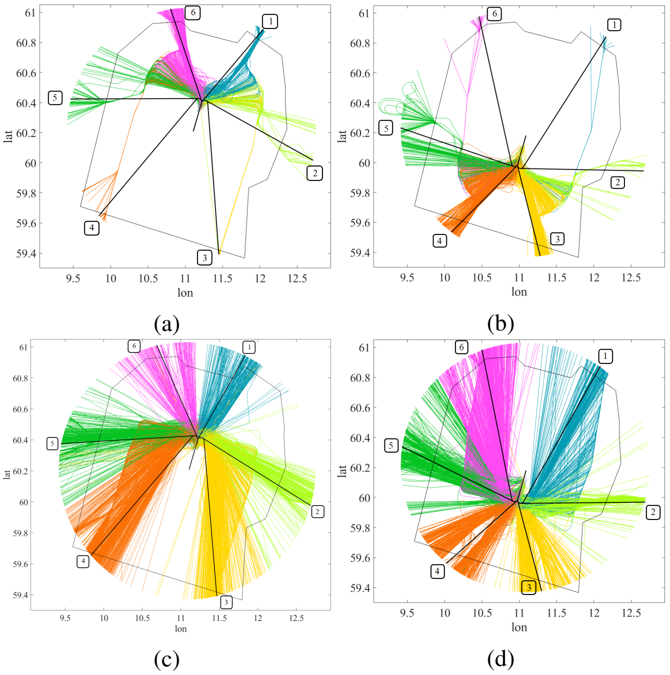
C. Performance Evaluation Metrics

We use a diverse set of metrics previously proposed in the related work (reviewed in Section III), most of which were applied earlier to evaluate the arrival performance at several airports including Dublin with PM procedures. The methodology for their calculation is described in more details in [19], [25], [26]. We complement the set with new metrics for efficient quantification of the PM utilization.

Distance represents the actual distance aircraft flown within TMA, calculated over the flight trajectories from our datasets.

Additional Distance is used to evaluate the *horizontal flight efficiency*. According to methodology described in [18], we cluster the trajectories by the points aircraft enter the TMA (50 NM circle). Then, inspired by [27], we choose *ideal* reference trajectory by constructing a user-preferred route tree inside the TMA. Next, for each cluster centroid, we find the closest point on the TMA border and determine it as the start of the horizontal reference trajectory. The reference trajectory goes directly to the interception of the localizer for an ILS approach, with a 3 NM straight segment before the Final Approach Point (FAP).

Reference trajectories are calculated separately for the northern and southern parts of our datasets, and they rely on the clustering which is sensitive to the choice of the runway. Figure 5 illustrates the clusters for PM_{North}, PM_{South}, $nonPM_{North}$ and $nonPM_{South}$ datasets.

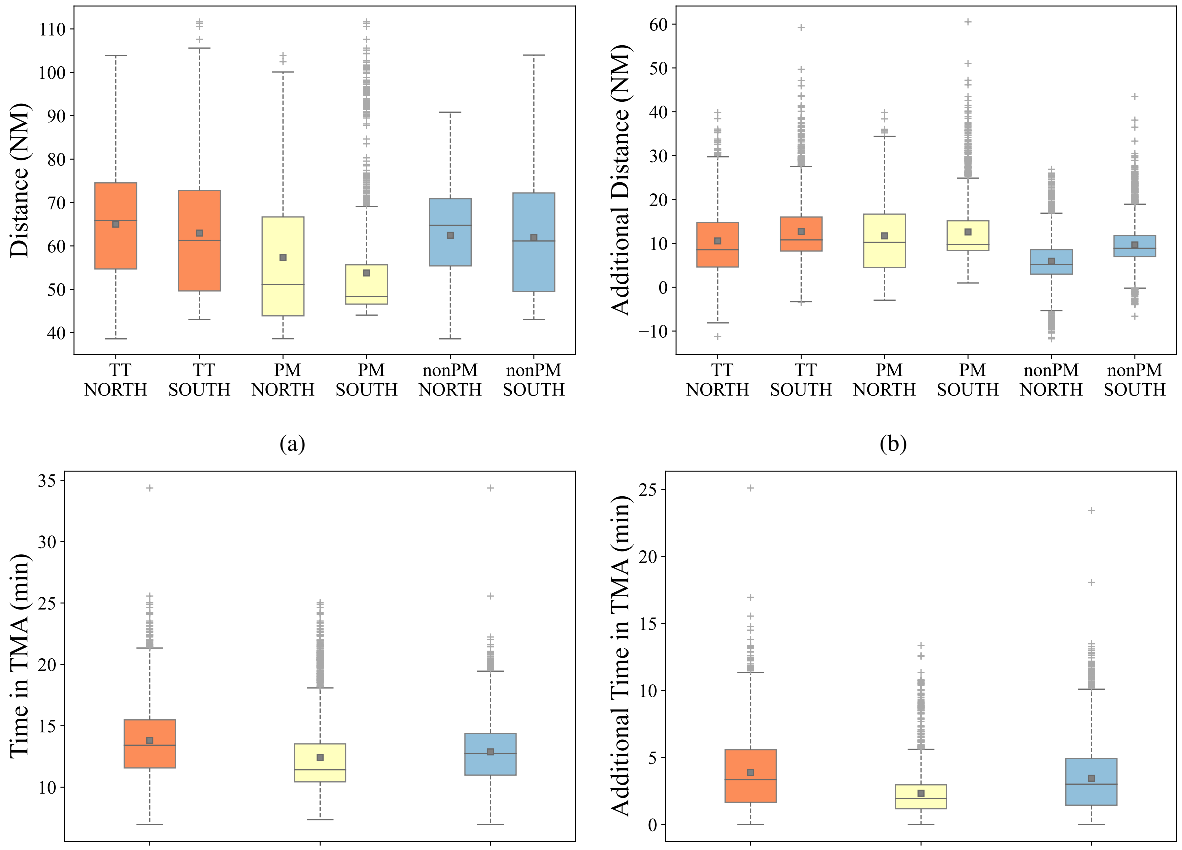


(a) (b)

(c) (d)

Figure 5: Reference trajectories (black lines) and the actual arrival trajectories colored by cluster for PM_{North} (a), PM_{South} (b), $nonPM_{North}$ (c) and $nonPM_{South}$ (d) datasets.

We use **Time Flown Level** in TMA for *vertical flight efficiency* evaluation. With small variation, we use procedure presented by EUROCONTROL in [28] for calculation. We assign as a starting point the point on TMA border (50 NM circle) crossed by aircraft. When the aircraft flies with the vertical speed below 300 feet per minute at least 30 seconds, we identify it as a level segment. As advised in [28] we subtract the first 30 seconds of level flight from each level



(a) (b)

(c) (d)

Figure 6: Distance in TMA (a), Additional Distance in TMA (b), Time In TMA (c) and Additional Time in TMA (d), calculated for TT, PM and $nonPM$ datasets.

duration. We do not consider as level flights under 1000 feet which corresponds to final approach.

For the **Vertical Deviation**, the reference profile is a continuous descent, constructed following the methodology proposed in [29], using EUROCONTROL Base of Aircraft Data (BADA) v 4.2 [30]. When calculating the vertical reference trajectories, we assume an unrestricted descent, hence, we do not respect any altitude restrictions that may apply in the TMA. However, we do not allow our vertical reference trajectories to cross the TMA border at a higher altitude than the cruise altitude for the flight, thus, we may have an initial level flight segment for flights that have a low cruise altitude. We calculate vertical deviation for the 10 final minutes prior to the final approach.

Additional Fuel Burn is calculated to evaluate the arrival *environmental efficiency*. The Additional Fuel Burn is computed as the difference between the real and the reference trajectory, corresponding to CDOs, fuel consumption. To detect the thrust force, we derive the thrust coefficient from the Total Energy Model (TEM) from BADA v.4.2 [30]. We also use the BADA technique for estimation of the idle-thrust CDOs for each flight. We take into account temperature and wind at the current position, obtained from ERA5 [24]. The methodology is detailed in [31]).

We calculate the **Minimum Time to Final** by overlaying a rectangular grid of cell size $\approx$ 1 NM over plotted flight trajectories in TMA. To each cell, we assign the time-wise best-performer's time needed from the cell location to the final approach. To the cells through which no flight trajectory pass during the chosen time period, we assign infinite (or very large) value. We visualize the resulting values with *heatmap* of the minimum time to final on a grid.

Horizontal Spread metric is a rough estimation of the percentage of the TMA area occupied by flights and a quantifier of the dispersion of the arrival flows. Horizontal Spread is calculated as ratio of the number of cells which contain at least one trajectory to the number of all cells covering the TMA. Arrivals following similar paths are indicated by lower value of the Horizontal Spread.

Spacing Deviation for an arriving aircraft pair is calculated as the difference between their respective minimum times to final. The aircraft in arriving pair are labeled as the leader and the trailer according to their arrival time to final point. Leader is the aircraft arriving earlier than the trailer aircraft. For spacing deviation calculation we use the following equation: $sd(t) = min\_time(\text{trailer}(t)) - min\_time(\text{leader}(t - s_{rwy}))$,

where s_{rwy} represents the time separation at the runway,

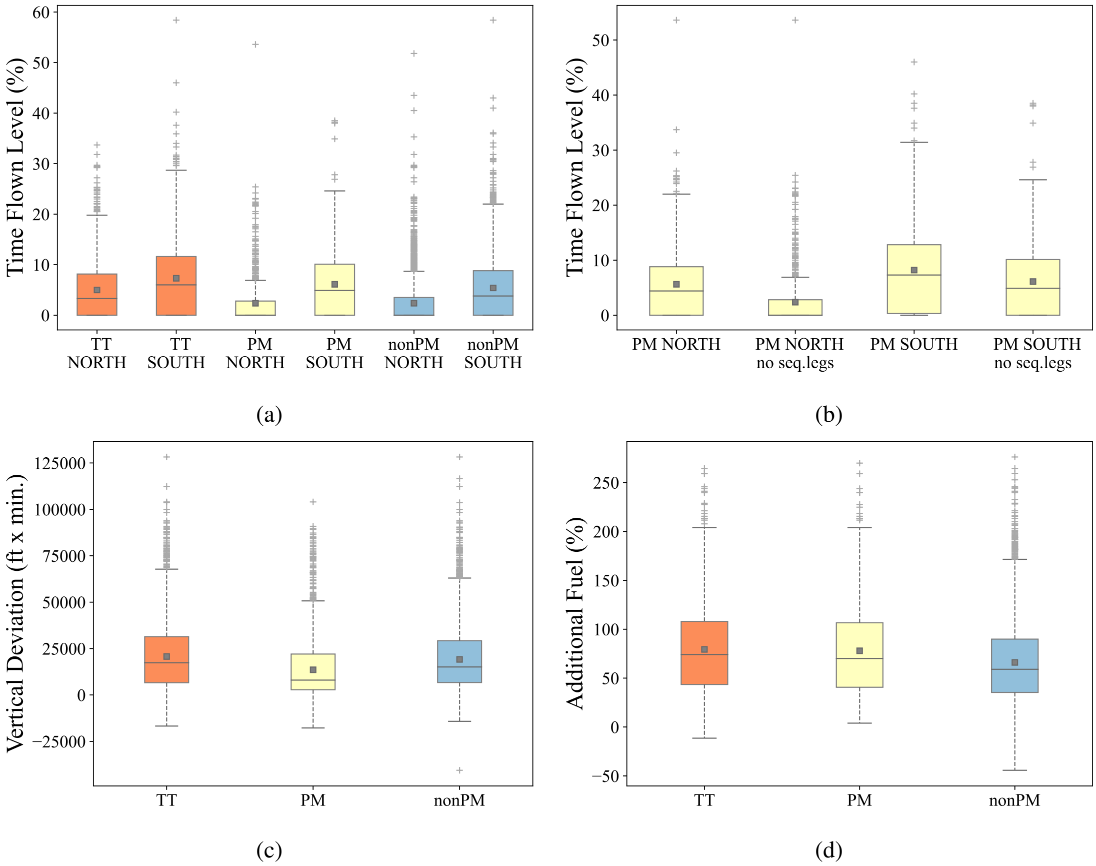


Figure 7: Time Flown Level (a), Vertical Deviation (c) and Additional Fuel (d) calculated for TT, PM and $nonPM$ datasets, Time Flown Level calculated for PM dataset with and without time aircraft spent on the PM sequencing legs (b).

and *min_time* is the minimum time to final. The spacing deviation reflects information about the control error, which is the accuracy of spacing around the airport.

We define **Throughput** at a given time horizon t as the count of the number of aircraft with the same minimum time to final within a given time window. We calculate the throughput crossing iso-minimum time lines from 900 to 30s to final, sampled at a 30s rate over 5-minute periods.

Metering Effort is calculated as the difference between the throughput value at the given time horizon and the one close to the final (30s). Metering effort is a quantifier of the controllers effort for metering and may be used as a proxy to controllers workload.

Time in TMA represents the actual time aircraft spent in TMA, calculated over the flight trajectories from our datasets.

Additional Time in TMA (ATT) is calculated for each aircraft as the difference between the Time in TMA and the Minimum Time to Final value assigned to the first cell in the grid, which the aircraft trajectory passes through after aircraft entered the TMA.

We evaluate the **PM usage** by identifying the flights adherent to the PM procedures, and calculating the proportion of these flights in the given dataset.

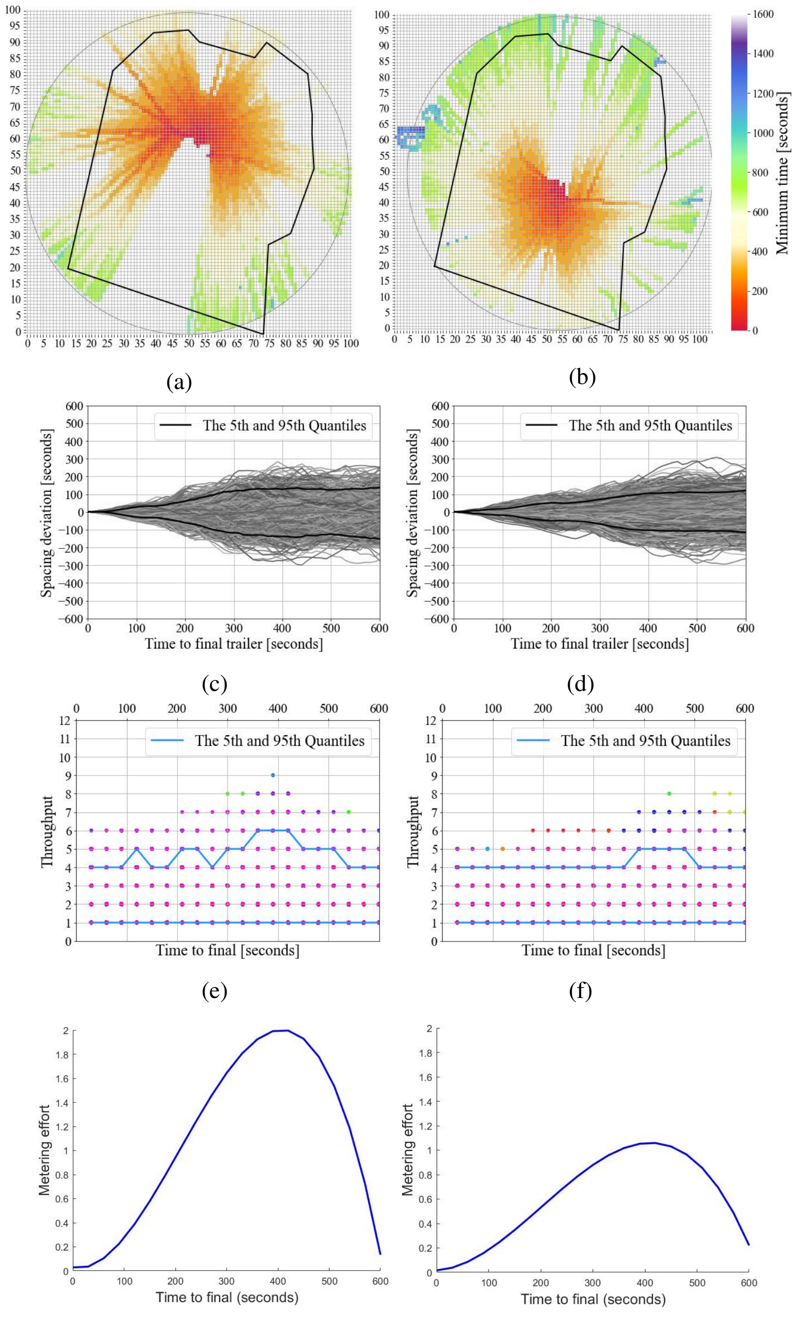
In order to evaluate what part of the PM sequencing leg is actually utilized by the flights, we introduce a new metric called **PM Utilization**, as the proportion of the length of the PM sequencing leg flown by the arriving aircraft to the full length of the corresponding PM sequencing leg, in percent. To estimate this, we measure the distance along the sequencing leg from the starting point to the point, when the aircraft was directed to turn towards the merge point and proceeded to the final approach. We use a small circle catch areas of ≈ 2 NM around each waypoint on sequencing legs for each PM system to capture that. Figure 4 illustrates the described method on example of the 19 East PM system of Oslo Gardermoen airport procedures. The green and the red circle catch areas correspond to two sequencing legs used by aircraft coming from opposite sides.

V. PERFORMANCE EVALUATION: RESULTS

First, we assess the arrival performance during peak-time periods using TT dataset. Then, we focus on the flights adherent to the PM procedures and evaluate their performance using the PM dataset, comparing it against the flights in the $nonPM$ dataset and highlighting the differences. Figures 6, 7, 8 and 9 illustrate the results, Tables II, III, IV and V contain the statistics for the metrics used in this section.

A. Arrival Performance During Peak Periods

Calculating the Time aircraft spent in TMA for the TT dataset, we observe, that during the peak time periods the



(a) (b)

(c) (d)

(e) (f)

(g) (h)

Figure 8: Minimum Time to Final heatmaps (a)-(b), Spacing Deviation (c)-(d), Throughput (e)-(f) and Metering Effort (g)-(h), for the TT_{North} and TT_{South} datasets, respectively.

aircraft spent from 6.95 to 34.37 min in TMA with the average 13.82 min (Figure 6 - (c), Table II). Average and median Additional Time values during peak time periods are noticeably small in this airport: 3.88 and 3.35 min respectively ((Figure 6 - (d), Table II)), which is an indicator of the outstanding arrival performance.

Additional time values for the TT dataset (Figure 6 - (d), Table II) vary between 0 and 25.1 minutes, with the average of 3.88 min and median 3.35, which is an indication of the outstanding arrival performance in general, with no significant delays even during the peak periods.

Additional Distance in TMA slightly varies for the northern and southern parts of the TT dataset (this KPI is calculated separately for the TT_{North} and TT_{South} because the reference trajectories for its calculation are obtained using the clustering technique applied to each runway separately). The corresponding values of the metrics calculated for TT_{North} and TT_{South} are described in Table III. We observe that for the northern part Additional Distance is lower on average but with higher deviation for both parts (Figure 6 - (b)).

TABLE II. Statistics on the Performance Metrics for TT, PM and $nonPM$ Datasets

	TT	PM	$nonPM$
Time in TMA (min)			
median	13.42	11.42	12.73
mean	13.82	12.41	12.88
std.dev.	2.97	3.02	2.55
min	6.95	7.35	6.95
max	34.37	25.02	34.37
Add. Time in TMA (min)			
median	3.35	1.95	3.02
mean	3.88	2.34	3.46
std.dev.	2.89	1.87	2.56
min	0.0	0.0	0.0
max	25.1	13.37	23.43
Add. fuel (%)			
median	74.13	70.09	59.03
mean	79.36	77.95	66.04
std.dev.	43.99	44.79	39.7
min	-11.43	3.95	-44.27
max	264.37	269.88	276.14
Vertical Deviation (ft x min.)			
median	17343.2	8010.09	15082.28
mean	20719.03	13571.7	19151.77
std.dev.	18166.65	16542.03	16308.39
min	-16751.8	-17785.15	-40586.97
max	128225.23	104006.41	128225.23

Average Time Flown Level expressed in percent of the flight time is also slightly smaller for the northern part (Figure 7 - (a)) with the average values of 4.99% and 7.31%, correspondingly. We also calculated that only 31% of the flights of the whole TT dataset have no levels (Table I, last column).

Figure 7 - (c) shows the vertical deviation compared to a reference CDO, for the last 10 minutes of flight, prior to the final approach, with median and average values of 17,300 and 20,700 $ft \cdot min$, respectively, for the TT dataset.

The results for the Additional Fuel Consumption are illustrated in Figure 7 - (d). We observe a median and average additional fuel of 74% and 79%, respectively, for the TT dataset.

The TT dataset was constructed in the way that it contains the complete sets of the arrival during the selected hours representing peak time periods, which enables the evaluation of the sequencing, metering and spacing for the arrivals during these time periods. Figure 8 shows the results for the sequencing and metering performance metrics, calculated separately for northern and southern parts of the TT dataset, because these metrics also depend on the location of the runway in use.

Figures 8 - (a), (b), illustrates the Minimum Time to Final heatmaps for TT_{North} and TT_{South} datasets, respectively. In Table IV we present the statistics corresponding to the performance metrics. We observe that the Minimum Time to Final values are slightly higher for the arrivals to the southern runways, this might be caused by different size of arrival flows as there is almost twice more aircraft arriving to the southern runways (1047 flights to the North and 2094 to the South). We also recognize stack pattern in some arrival trajectories from west in the heatmap for the southern arrivals.

Horizontal Spread values for these two datasets (63.32% for TT_{North} and 66.86% for TT_{South}) indicate that both parts leave enough space for the operations to the opposite runway direction. The Horizontal Spread for full TT is 80.41% which demonstrates that the TMA is fully utilized by the flight trajectories and does not leave too much space for abrupt changes and emergency maneuvers in the TMA. For comparison, previous studies [19] reported similar value for the Horizontal Spread in Vienna airport with trombone structure, and lower for Dublin airport (64%) with one PM system and one runway.

The Spacing Deviation is shown in Figures 8 - (c), (d). Even though the maximum absolute, average, and median values, are quite similar, the value of the 90th quantile width varies significantly, that indicates that the higher traffic volume on southern runway is still well-managed. PM systems enable smooth and continuous convergence of the arrival sequences to the final.

Figures 8 - (e), (f) illustrate Throughput metric. The maximum and average values of the throughput for TT_{North} and TT_{South} datasets are quite close (9 and 2.51, respectively, for TT_{North} vs. 8 and 2.08 for TT_{South}), which indicates stable Throughput if the TMA in both directions. Previously, in [26] we evaluated the performance of the Dublin airport (which is also implementing PM arrival procedures), based on TT dataset obtained using the same technique as in this work, and observed the maximum throughput of 12 flights. Given that there was only one runway in use in Dublin, the entry and final conditions at these two airports both operation PM, are quite different. This makes a fair comparison of the performance at two airports problematic.

Figures 8 - (g), (h) illustrate the Metering Effort. Noticeably higher values of the maximum and average obtained for TT_{North} dataset are indicating that for sequencing of arrivals more control effort is applied in the northern part of the procedures, despite the fact that the traffic intensity is much lower here. The slopes and the peaks of the metering effort curves illustrate when the sequencing and metering techniques are applied.

Looking at the Metering Effort figures obtained for Dublin airport (the results are presented in [19]), we can see that the shape of the metering effort slope and the time when it peaks are quite similar to the ones for Oslo, but the maximum value is significantly higher, i.e. 3 units, confirming that more control effort is to be applied to sequence the arrivals in this airport. Note, that for the comparison we choose similar time periods at these airports (October 2019) with the maximum amount of traffic observed before the Covid-19 pandemics. According to [32], in the Fall 2022 a second parallel runway is implemented and is getting gradually in use in Dublin airport. This is expected to increase the airport capacity and improve the throughput.

B. PM Utilization

Analysing the PM dataset, we observe that the PM systems are not utilized to the full extent, and they are also used quite irregularly. During the days of observation, the percent of flights performing PM per day varies between 5% and 30% (Figure 9 - (a)). In addition, the number of flights performing

TABLE III. Statistics on the Performance Metrics for the Northern and Southern Parts of the TT,PM, and $nonPM$ Datasets

	TT $North$	TT $South$	PM $North$	PM $South$	$nonPM$ $North$	$nonPM$ $South$
Distance (NM)						
median	65.85	61.29	51.15	48.34	64.76	61.14
mean	65.02	62.96	57.3	53.77	62.46	61.91
std.dev.	14.31	14.14	16.55	12.56	11.9	12.7
min	38.59	43.03	38.61	44.06	38.59	43.03
max	103.87	111.63	103.87	111.63	90.82	103.98
Add. distance (NM)						
median	8.54	10.79	10.24	9.72	5.14	8.88
mean	10.55	12.67	11.71	12.6	5.93	9.64
std.dev.	7.91	7.07	8.67	6.97	5.47	5.14
min	-11.26	-3.48	-2.96	0.96	-11.78	-6.61
max	39.87	59.21	39.88	60.51	26.9	43.5
Time on Levels (%)						
median	3.3	6.0	0.0	4.9	0.0	3.8
mean	4.99	7.31	2.36	6.11	2.37	5.39
std.dev.	6.07	6.99	4.99	6.4	4.54	6.18
min	0.0	0.0	0.0	0.0	0.0	0.0
max	33.7	58.4	53.6	38.5	51.8	58.4

TABLE IV. Sequencing and Metering Metric Statistics

	TT	TT $North$	TT $South$
Minimum Time to Final [seconds]			
median	430	441	528
mean	434	438	528
std.dev.	194	174	196
max	1400	1066	1399
Horizontal Spread [percent]	80.41	63.32	66.86
Spacing Deviation [seconds]			
median		0	0
mean		-0.04	-0.04
std.dev.		70.52	52.28
min		-376	-351
max		376	362
90^{th} Quantile width		322.25	250
Throughput [aircraft]			
median		2	2
mean		2.51	2.08
std.dev.		1.29	1.09
max		9	8
Metering Effort			
median		1	0
mean		0.7	0.2
std.dev.		0.73	0.41
min		0	0
max		2	1

TABLE V. PM Utilization

$PMSystem$	$No\ arc$	1/3	2/3	$Full\ arc$
19 West	72.59%	19.54%	4.32%	3.55%
19 East	84.27%	13.48%	2.25%	0%
01 West	84.18%	12%	3.14%	0.68%
01 East	85.27%	11.34%	2.65%	0.74%

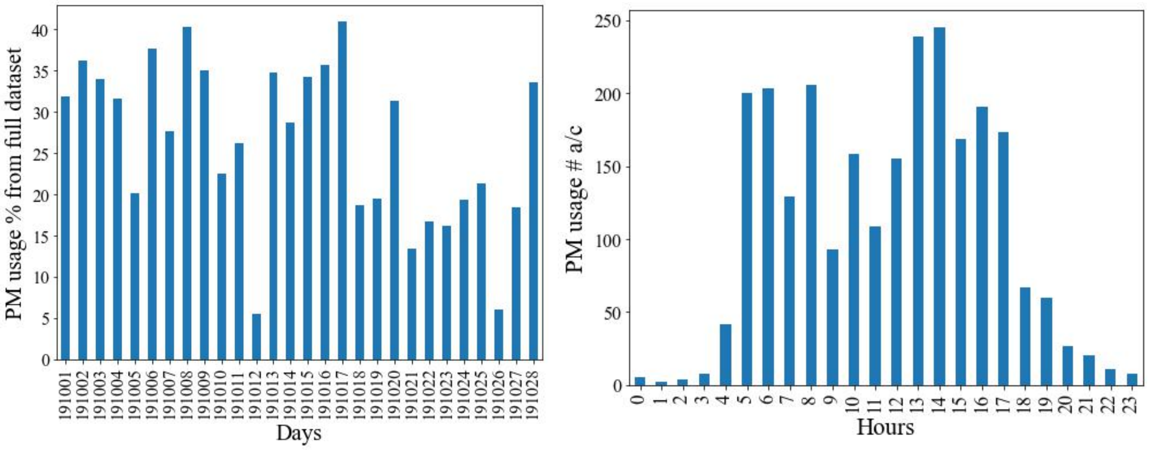
PM is particularly low during low-traffic hours, especially at night (between 10 pm and 4 am) (Figure 9 - (b)).

Clearly, the PM procedures are not always in use. In order to investigate the PM usage even further, we apply the newly developed metric PM Utilization, reflecting what part of the PM sequencing leg is flown by the aircraft. The resulting distribution is as follows: only 1.27 % of the arrivals performing PM stay on the sequencing legs until they reach the end of the arc, followed by 3.17 % of flights utilizing up to two thirds of the arc, 13.4 % of flights are utilizing only about one third of the arc, and the majority of the flights (82 %) are skipping the sequencing legs and fly directly to the merge point. Similar distribution is observed for all four parts of the PM systems (two associated with the runway 19 and two with runway 01) summarized in Table V. The PM system 19 West is used slightly more often, but still no more than 27% of flights are actually using the PM sequencing legs.

This adds to the picture that the PM systems are not utilized to the full extent. As a positive outcome of such result, we can assume that the airport has spare capacity to accommodate more incoming flights in the future. In what follows we analyse the performance of the flights following PM procedures in more details, in order to uncover the benefits and disadvantages of their usage.

C. PM Performance

Despite the expectation that the flights adherent to the PM procedures spent more time inside TMA, we observe (Figure 6 - (c)) that aircraft in PM dataset spent on average 12.41 min inside TMA, which is slighly lower than for the flights in $nonPM$ dataset (12.88 min). This might be explained by two factors. First, the low number of flights actually utilize the PM arcs, many flights skip the PM arc and fly directly to the merge point (see subsection V-B). Second, most of the flights from PM dataset use PM systems on the same side and land with the same direction they enter the TMA. Figure 5 shows that for the northern subsets of the datasets flights from the clusters number 3 and 4 are coming from the opposite direction (from the south), and for the southern subsets clusters 1 and 6 contain flights entering TMA from the opposite side (from the north). 44 % of flights from $nonPM$ dataset and only 7 % of flights from PM dataset are landing with the opposite direction they enter the TMA. Additional time calculated for PM dataset is also noticeably smaller than the one for $nonPM$ and TT datasets by all the



(a)

(b)

Figure 9: Percent of flights performing PM procedures per day in October 2019 (a), and the number of aircraft performing PM per hour (b).

.

statistics, which is a good indication that PM usage improves the flight performance (Figure 6 (d), Table II).

Next, to compare the *horizontal performance* of the flights adherent to PM against the non-PM flights, we calculate Additional Distance in TMA for the PM and $nonPM$ datasets (Figure 6 - (b), Table III). Average Additional Distance for PM_{North} is slightly lower than the one for PM_{South}, and the one for $nonPM$ datasets is noticeably lower. We suppose that adherence to the PM systems increases additional distance aircraft have to fly in TMA due to deviation from the direct reference trajectories.

Applying the other metric characterizing horizontal performance of the arrival flights, Distance in TMA, we observe the opposite trend ((Figure 6 - (a), Table III). The average and median values of this statistics are significantly lower for PM dataset when compared against the ones for both TT and $nonPM$, which shows that the flights actually fly shorter distanced when they perform arrival procedures with PM.

In addition, comparing the results of the Additional Distance for PM and TT datasets, we observe that there is no significant difference between them. And since TT dataset represents the peak-hour performance, we may assume that PM procedures are applicable for the high-traffic scenarios and do not necessarily lead to the degradation of the horizontal performance.

We evaluate the *vertical performance* of the PM dataset with Time Flown Level metric using two different methods. In the first method, we calculate it for the whole trajectories, the same way we do it for the TT and $nonPM$ datasets. In the second method, we exclude from our calculations the altitudes corresponding to the sequencing legs of the PM procedures, that is from 9000 to 11000 feet for the Oslo Gardermoen airport. The results of these two approaches are compared in Figure 7 - (b).

Comparing the results with excluded PM altitudes to the results obtained for $nonPM$ dataset (Figure 7 - (a)), we do not observe any significant difference between them (Table III). We can conclude that there is more significant difference in the *vertical performance* of the northern and southern subsets than the one between the PM and $nonPM$ datasets.

We observe that the Vertical Deviation (Figure 7 - (c)) for flights using PM is lower, compared to flights not using PM, which is in line with results for Time Flown Level. Median and average Vertical Deviation is 8,000 and 13,600 $ft \cdot min$, respectively, for the PM dataset, and 15,100 and 19,200 $ft \cdot min$, respectively, for the $nonPM$ dataset.

To confirm the result, we evaluate also the number of nonlevel flights using the second method described for the Time Flown Level metric, i.e. excluding the time aircraft spend on sequencing legs. The results are presented in the third column of Table I. The number of nonlevel flights in PM and $nonPM$ datasets are also very similar and differ mostly in their northern and southern subsets.

To summarize, northern parts of all the datasets (representing flights landing on 19R and 19L runways) demonstrate noticeably better vertical performance. When we calculate the time on levels excluding the time aircraft have to spend on sequencing legs, the flights utilizing PM perform similarly to the non-PM flights, and outperform the flights during peak hours.

Comparing the results of the Additional Fuel for the PM and $nonPM$ datasets (Figure 7 - (d)), we observe that both average and median values are lower for the flights which are not using PM (70% and 78% for PM, respectively, versus 59% and 66% for $nonPM$). Since the vertical efficiency of the PM flights is better than the one for the non-PM flights, we conclude that this fuel inefficiency is originated from the slightly increased Additional Distance in TMA.

However, comparison of the Additional Fuel results for the PM against the TT dataset, is in favor of the first. This adds to the picture that the increased PM usage may improve the environmental efficiency during peak time periods.

VI. CONCLUSIONS AND OUTLOOK

We discovered that the overall performance of the Oslo Gardermoen airport is outstanding from many perspectives, and that the design of its arrival procedures can be used as a good example for implementation of the PM in the future. Because of its geographical location and the flow configuration typical for many airports in Scandinavian region, the procedures may serve as a best-practice implementation example. Future research may look at applicability of such a design for implementation at some airports in Sweden.

Analysing the PM usage, we noticed that the systems are not utilized to the full extent, which may indicate that the airport has a spare capacity to accept more incoming flights. The adherence to the PM procedures does not result in the significant performance degradation (confirmed by all performance indicators except for the additional distance and the fuel efficiency connected to it). Based on the above, we believe that if more aircraft are directed to follow PM systems, the airport throughput can be increased. As a complement, the dynamic usage of the procedures can be introduced, according to the dynamic route planning idea outlined in [33]. It is a topic for future studies to quantify the capacity of the airport with PM procedures, and how to plan the dynamic operational regimes to maximize the airport throughput. New solutions are to be assessed with the extended set of performance metrics for comprehensive evaluation of the resulting performance, including other environmental indicators such as noise and emissions.

REFERENCES

[1] EUROCONTROL. *Guidance Material for the Design of Terminal Procedures for Area Navigation (DME/DME, B-GNSS, Baro-VNAV RNP-RNAV).* EUROCONTROL, Ed. 3.0, March 2003.

[2] L. Boursier, B. Favennec, E. Hoffman, A. Trzmiel, F. Vergne, and K. Zeghal. Merging arrival flows without heading instructions. In *7th USA/Europe air traffic management R&D seminar*, pages 1–8, 2007.

[3] B. Favennec, T. Symmans, D. Houlihan, K. Vergne, and K. Zeghal. Point merge integration of arrival flows enabling extensive rnav application and continuous descent–operational services and environment definition. 2010, 2010.

[4] Ö. S. Meric and O. Usanmaz. A new standard instrument arrival: the point merge system. *Aircraft Engineering and Aerospace Technology*, 2013.

[5] D. Ivanescu, C. Shaw, C. Tamvaclis, and T. Kettunen. Models of air traffic merging techniques: evaluating performance of point merge. In *9th AIAA Aviation Technology, Integration, and Operations Conference (ATIO) and Aircraft Noise and Emissions Reduction Symposium (ANERS)*, page 7013, 2009.

[6] M. Liang, D. Delahaye, and P. Maréchal. Potential operational benefits of multi-layer point merge system on dense TMA operation hybrid arrival trajectory optimization applied to Beijing capital international airport. In *Proceedings of the 7th International Conference on Research in Air Transportation*, 2016.

[7] EUROCONTROL, Point Merge. https://www.eurocontrol.int/concept/point-merge, last accessed on 11.05.2023.

[8] EUROCONTROL. Point Merge Implementation: A quick guide, 2020.

[9] S. Lee, Y. Hong, and Y. Kim. Optimal scheduling algorithm in point merge system including holding pattern based on mixed-integer linear programming. *Proceedings of the Institution of Mechanical Engineers, Part G: Journal of Aerospace Engineering*, 234(10):1638–1647, 2020.

[10] DSNA/EUROCONTROL SESAR. P05.06.02 – D57 - CDO at Orly Real-time Simulations Validation Report, 2015.

[11] C. Raphael, B. Favennec, E.G. Hoffman, and K. Zeghal. Detecting point merge patterns using computer vision. In *AIAA AVIATION 2021 FORUM*, page 2390, 2021.

[12] E. Aydoğan and C. Çetek. Point Merge Concept for En Route Air Traffic Flow Management. *JOURNAL OF AIRCRAFT*, 55(6), 2018.

[13] ATM EUROCONTROL. Airport performance (atmap) framework. *Measuring Airport Airside and Nearby Airspace Performance*, 2009.

[14] EUROCONTROL. Performance Review Report: An Assessment of Air Traffic Management in Europe during the Calendar Year 2021, 2021.

[15] R Christien, E Hoffman, A Trzmiel, and K Zeghal. Toward the characterisation of sequencing arrivals. In *ATM Seminar 2017*, 2017.

[16] R. Christien, E. Hoffman, and K. Zeghal. Spacing and Pressure to Characterise Arrival Sequencing. In *ATM Seminar*, 2019.

[17] R. Christien, E. Hoffman, and K. Zeghal. Towards a characterization of arrival metering, case study on a variety of european airports. In *AIAA Aviation Forum*, 2020.

[18] Pierrick Pasutto, Eric Hoffman, and Karim Zeghal. Vertical Efficiency in Descent Compared to Best Local Practices. In *ATM Seminar*, 2019.

[19] A. Lemetti, T. Polishchuk, H. Hardell, K. Zeghal, and L. Smetanova. Towards a Comprehensive Characterization of the Arrival Operations in the Terminal Area. In *SIDs*, 2021.

[20] Avinor. https://avinor.no/en/corporate/about-us/statistics/archive, last accessed on 05.02.2023.

[21] Norwegian AIP. https://avinor.no/en/ais/aipnorway/, last accessed on 05.02.2023.

[22] Opensky Network. https://opensky-network.org/, last accessed on 11.02.2022.

[23] T. Polishchuk, T. Lemetti, and Raul S. Evaluation of Flight Efficiency for Stockholm Arlanda Airport using OpenSky Network Data. In *OpenSky Workshop 2019*, volume 67 of *EPiC in Computing*, pages 13–24.

[24] Copernicus Climate Change Service (C3S) Data Store, European Centre for Medium-Range Weather Forecasts (ECMWF). https://cds.climate.copernicus.eu, last accessed on 11.02.2022.

[25] H. Hardell, T. Polishchuk, and L. Smetanová. Automated Traffic Scheduling in TMA with Point Merge to Enable Greener Descents. In *10th International Conference on Research in Air Transportation (ICRAT), Tampa, June 19-23, 2022*, 2022.

[26] H. Hardell, A. Lemetti, T. Polishchuk, and L. Smetanová. Evaluation of the Sequencing and Merging Procedures at Three European Airports Using Opensky Data. *Engineering Proceedings*, 13(1):13, 2022.

[27] V. Polishchuk. Generating arrival routes with radius-to-fix functionalities. In *ICRAT*. Drexel University Philadelphia, 2016.

[28] EUROCONTROL. Analysis of Vertical Flight Efficiency During Climb and Descent. 2017.

[29] H. Aksoy, E. T. Turgut, and Ö. Usanmaz. The design and analysis of optimal descent profiles using real flight data. *Transportation Research Part D: Transport and Environment*, 100:103028, 2021.

[30] EUROCONTROL. User Manual for the Base of Aricraft Data (BADA) Family 4, 2014.

[31] A. Lemetti, T. Polishchuk, and H. Hardell. Arrival Flight Efficiency in Numbers: What New the Covid-19 Crisis is Bringing to the Picture? In *SIDs*, 2020.

[32] O'Connor, F. Dublin airport slot capacity up just 7.4pc despite new runway., November 2022. https://www.independent.ie/business/irish/dublin-airport-slot-capacity-up-just-74pc-despite-new-runway-42155860.html, last accessed 05.02.2023.

[33] B. Favennec, K. Zeghal, and E. G. Hoffman. Introducing dynamicity in terminal areas: a new perspective. In *AIAA AVIATION 2020 FORUM*, page 2892, 2020.