PRC Performance Insight #5

Open performance data

- *an enabler to further performance enhancements?*

Measuring European Aviation Performance

Powered by the Aviation Intelligence Unit

Performance Review Commission

Rue de la Fusée 96, B-1130 Brussels, Belgium

pru-support@eurocontrol.int
www.ansperformance.eu

Introduction

This Performance Insight is published by the Performance Review Commission (PRC) to foster the dialogue and understanding on **PRC's Open Performance Data Initiative**. The PRC promotes a performance-based approach. This revolves around establishing independent and reproducible recommendations for policy making, strategic decision-making, and planning (c.f. Figure 1).
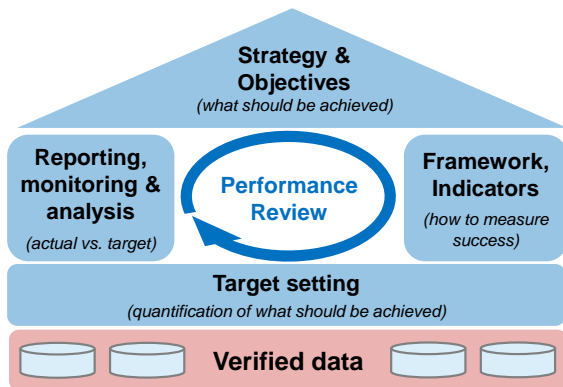

Figure 1 – PRC framework – robust data for monitoring

Independent performance monitoring is based on verified and robust performance data and a properly defined performance data processing framework supporting reproducibility. Throughout the recent years, data literacy and capabilities increased significantly. Community-based data collection (crowd-sourcing) emerged and offers to enhance and augment today's performance data environment.

The PRC launched the Open Performance Data Initiative in 2022 with a view to establish an open data platform, to support higher levels of transparency and reproducibility of performance related monitoring and associated analyses. This initiative addresses the requirement to establish a level-playing field for all stakeholders, ranging from the interested public, operational and strategic planners, to political decision-makers. Pursuing an open data-based approach will allow all interested parties to tap into harmonised data environment to build their assessments, validate published results by other parties, and reproduce findings to support the political and operational discussion about observed performance.

The upcoming challenges to our industry require that stakeholders can embrace and tap into a harmonised and consolidated data set to drive the discussion on higher performance levels, performance expectations in terms of targets, or realised performance benefits following the deployment of new capabilities.

This paper provides the rationale and background of the Open Performance Data Initiative, benefits and barriers of implementing an open data approach and examples of applications of open performance data in the form of a select "snapshot" of recent research. As an independent body promoting performance review, the PRC hopes to inspire interested parties to join and support the next steps of the Open Performance Data Initiative.

Rationale and Background

Independent (operational) air navigation system (ANS) performance builds on transparent and robust performance monitoring and reporting.

Within the European context, EUROCONTROL's PRC is committed to provide impartial analyses of ANS performance and promotes an open and transparent dialogue between stakeholders, policy-makers and strategic decision-makers.

A principal enabler for such a transparent dialogue is a level playing field for all stakeholders to access the underlying data, reproduce results, or build their own analyses to validate findings and recommendations. Open performance data can therefore remove the barrier of information asymmetry amongst the parties and foster an open dialogue and mutual understanding.

Air transportation and ANS faces a variety of political priorities and societal challenges. For example, traffic recovery reached 85-90% of pre-pandemic traffic levels in summer 2022, and the summer season was hugely turbulent. Staff shortages on the airport and airspace user side created a perfect storm for the travelling public, increasing Air Traffic Flow Management (ATFM) delays questioned the planning on the Air navigation service provider (ANSP) side. These factors amplified pre-COVID patterns and lead to a decline in operational performance questioning efficiency, punctuality and predictability. At the same time ripple effects of Putin's war of

aggression in Ukraine, such as increasing fuel costs and inflation, pose additional challenges to air transport operations. Discussions on climate change and the environmental footprint of aviation are accelerating in light of the observed unprecedented extreme weather phenomena (rain/flooding in 2021, drought in 2022). To ensure an evidence-based dialogue, there is a need to provide impartial and reproducible analyses on the basis of accepted and transparent data.

Open data is not per se a new topic. Many national and pan-regional programmes or policies exists with a focus on generating value for economy and society through the reuse of public sector information. The PRC is committed to independent and impartial recommendations on performance related matters to policy-makers and strategic planers based on a performance-based approach. A critical enabler for this approach is the availability of robust data to support the performance monitoring. Independent and impartial analyses will support the identification of priorities or policies to address the emerging challenges and ensure that air transportation will meet its societal expectations and air navigation demonstrates the required higher levels of performance and facilitates the envisaged change over the coming years.

Aviation in general, air transportation and air navigation in particular are data-intensive sectors. Practical experience and research has shown that a huge portion (about 70%) of the data is collected and curated by national government organisations or – within their focus of operations – by different stakeholders.

Open and wider access to these data is limited, and typically constraint by operational processes and data product policies. Despite pan-European activities and a gradual move towards open data across air transportation research, no consolidated platform or dataset is available for open validation of future concepts and modes of operations.

Based on this apparent lack of open access to air transport data for research and development, operational performance analyses, and policy-making, the emergence of crowd-sourced, open access and open data activities have gained a higher visibility. These initiatives have benefited from like-minded communities of interest spanning hobbyists, researchers and open-minded professionals, internet and cloud-based crowd-sourcing and data sharing, associated low-cost sensor technologies, and emerging data analytical approaches (e.g. artificial intelligence/ machine - learning) requiring substantial higher volumes of data for validation.

Within the operational air transportation domain, ADSB and Mode S has become the key source of air traffic movement information. Open data products for weather and climate-oriented research, but also social media derived data sets on incidents or passenger observations frame the data universe to advance the state-of-the-art. To pave the way forward, it is therefore essential to bundle initiatives, drive the adaptation of the open data and open access approach within the air navigation community and establish a platform for the exchange of relevant data for current and future performance measures and analyses.

Benefits of Open Access and Open Data for Performance Monitoring
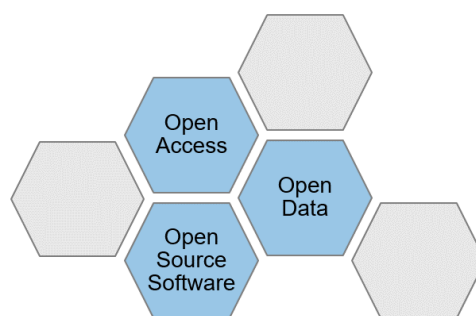

Figure 2 – Principal definitions – Open Access, Open Data

To understand the depth of the proposed Open Performance Data Initiative it is vital to define the key terminology[1] upfront.

---

[1] BE OPEN project (2021): Impact assessment of Open Science in Transport.

Open Access: Open Access refers to the process of managing copyright and licensing terms and access to the data. This access mode is vital to overcome today's limitation and remove barriers such as price (including subscriptions, licensing fees, user-pay-principles, or even view fees), and permission (e.g., licensing restrictions) to enable free access and use of the data.

Open Data: The term Open Data refers to data – typically managed and available online – which is free of cost and accessible data. Open data can be freely used, reused, and distributed provided that the data source is attributed.

Open-Source Software: Open-Source Software is software which source code can be openly accessed, inspected, modified, and enhanced by anyone.

The focus of the Open Performance Data Initiative is on open data, in particular open access and open use of data for non-commercial purposes. It is understood that potential developments of novel algorithms and techniques can produce tools that are made available to the community as open-source software.

There exists a huge body of literature highlighting the impact of open data. Across the different sources the following impact areas emerge:

- Transparency and accountability
- Innovation, participation, and joint development
- Stakeholder performance, service delivery and actor effectiveness
- Public and stakeholder engagement and empowerment

Open data and performance analytics are often viewed as separate practices, however, both are inextricable when it comes to the operations of modern governance. To ensure transparent and independent review, open data provides the underlying "raw material" for the analysis and monitoring of performance. In that respect, PRC's Open Performance Data Initiative aims at increasing transparency, supporting higher levels of data access and information availability for decision-making, engaging and increasing accountability of involved stakeholders.

Recent political priorities and calls for action require for a plain and level playing field. This also include access to the underlying data to validate and verify results and performance measures.

Transparency and accountability

A key tenet of the performance-based approach is independent and impartial monitoring to provide the evidence for achieved performance levels and to identify and share best practice for further performance enhancements in ATM.

Higher levels of transparency and accountability provide the foundation for a democratisation of air traffic management. Public and stakeholder access to data is also seen as a key ingredient to establish accountability. This also requires a culture change that is implicitly built into the performance-based approach. Observed levels of performance are publicly communicated and tracked. Associated performance measurement algorithms are published and the monitoring results are reproducible.

The recent years have also seen a higher level of data analytical skills and data processing capabilities with a wider portion of the stakeholder population and interested public. The higher level of engagement to tackle political priorities or provide decision-making input to organisational management is evidenced by a higher frequency of "underlying" data requests to establish or reproduce monitoring results.

Conceptually, reproducibility is not a new topic. It is as old as the scientific method and at the heart of the verification/validation of results. The consistent and robust interpretation of results by independent researchers/analysts forms the evidence and the inherent body of knowledge in support of the applied methods or conclusions. This principle is also relevant for the further development on the basis of such results.

Innovation, participation, and joint development

Open data from both public and private sources are adding a new dimension[2] to the data revolution. It also is a key ingredient to leverage the promises of "big data", or supporting data-analytical advances in artificial intelligence/ machine-learning.

A clear driver for innovation is the opportunity to understand how systems work, where bottlenecks emerge or certain patterns manifest. At the same time, it is important to understand the parameters and conditions for good and best-in-class performance. Sharing and access to data therefore fosters to have another view on performance by observing new developments with different priorities/ perspectives, or other organisational /institutional factors limiting the use and value creation from data, including its combination with additional data and resources.

The ability of public, interested researchers, policy-makers/think tanks, or strategic planners to access the data and exploit it creates joint communities of interest. This may enhance the value of the data to drive joint developments ranging from data standardisation, including access and storage, but also to build on the data. The latter may range from methodological approaches, operational solutions and other tools.

Effectiveness

At the heart of the performance-based approach is the potential to achieve higher levels of performance and associated decision-making through data-driven decision-making. This may affect day-to-day decision-making, but also long-term decision such as the need for adopting new operational concepts or deploying operational capabilities.

Open data goes beyond the classical – potentially limited – sharing of data in support of use-case analyses, projects or activities. It is a continual process which allows to champion and demonstrate observed performance levels. In that respect open performance data can also serve as a

tool to improve communication and partnerships in terms of data sharing.

Engagement and Empowerment

The most direct impact of open data for engagement and empowerment is related to its transparency effects. Air transportation and air traffic management are at a crossroad. Future political goals and expected performance levels require higher levels of transparency and the existence of a data basis in support of verification and validation/reproduction of results. The transparency discussion will result from a higher pressure on the industry to demonstrate – for example – its contribution to climate change and greenhouse gas emissions. There is also a discussion on the success of air navigation system modernisation (i.e., transformation) in general. This requires demonstration of operational performance benefits derived from funding and modernisation activities.

The need to address the societal challenges and political goals will require a wider engagement of air navigation with the public and policy level, but also facilitate such an engagement with the industry and interested public. The PRC is confident that the practice of the past to provide a platform to validate and reproduce performance observations will facilitate higher levels of engagement and empowerment with and between stakeholders and public. Open data can be a key enabler to facilitate this engagement and ensure the appropriate level of independent transparency to drive and inform the public debate, political and strategic decision-making.

---

[2] McKinsey, Open data: Unlocking innovation and performance with liquid information, October 2013.

Obstacles and Blocking Points

The previous sections showed benefits of adopting an open performance data stance to meet the challenges and expectation of our industry. It goes without saying that this also requires a cultural change. While air transportation and air navigation is a data rich environment, there are also concerns that need to be addressed.

Initial discussions have revealed a set of typical barriers:

- Business criticality, e.g. confidentiality, secrecy, or legal restrictions
- National security and defence

Business criticality (confidentiality, secrecy, or legal restrictions)

Open data per se shall not be confused with the release of confidential or other markings of data. It is important to recognise existing frameworks (e.g. privacy or personal data, GDPR). Air transportation related data recognises markings such as confidential and secret. Such data will not be available to the open data community. The performance-based approach embraces transparency on the basis of comparing national and local unit level performance, sub-population of airspace users, procedure oriented application of concepts. It is envisaged that the Open Performance Data Initiative will not comprise data or data sources that would conceal individuals, or expose individual organisations. On an international level, for example, the Organization for Economic Cooperation and Development has outlined a set of eight basic privacy principles that can serve as a blueprint to ensure privacy or confidentiality issues.

National security and defence

Concerns about disclosure of data that can impact national security or defence is a further consideration. A fundamental issue is the potential disclosure and localisation of security assets (e.g. military or law enforcement, first responders). An open data approach is complementary to pre-existing legislation and directives, and – of course – national security or defence. While the processing, storage, and sharing of marked material is out of scope of open performance data, the question about how combinations and integration of air transportation data with other data sets may support the disclosure of such sensitive information. Proper data governance or principles (e.g., privacy principles mentioned in previous section) provide to ensure proper treatment of the data.

PRC's position on blocking points

A key point in the context of the "blocking issues" relates to the context of how open performance data is expected to be used. It is important to recognise that it is not intended to support any real-time or operational use, nor to support commercial activities. Rather it is intended to support post hoc / after the fact performance analysis to enable improve future performance monitoring and improvement. By understanding and agreeing the practicalities of how data would be collected, stored and made available under this form of usage many of the blocking factors can be removed or significantly mitigated, allowing sharing of data without causing material commercial or security concerns.

The technological transformation and digitalisation in ATM will undoubtedly further foster the availability of data. The sharing of a set of defined data for analysis and research is expected to exploit further performance synergies for the benefit of the entire industry.

Embracing open performance data to drive higher levels of performance and transparency requires a cultural change and a new look on air transportation data. It is understandable that airspace users, air navigation service providers, or airports may fear liability or costs associated with inadvertently disclosing commercially sensitive or *confidential* information. The same applies for violating privacy rules. A related concern is linked to national security and defence. The underlying issue is the absence of appropriate data sharing mechanisms and policies.

With the exemption of specifically marked information or data sets, public air transportation data represents data free from markings. Operational procedures are established to conceal the identity of critical assets or operations. In consequence, open performance data will not comprise such marked data.

There is a thrust that air transportation data is funded by airspace users and represents public data. Thus, operated flights and data related to these flights, including the infrastructure to sense and process these data qualify for open data when managed through appropriate rules / processes. Accordingly, the benefits of new insights and the measurement or analysis of aspects that were never measured before to address the future challenges of traffic growth, climate footprint, and operational efficiency can be built on a basis that reflects privacy, commercial/organisational, and national security and defence concerns.

Snapshots – Application Cases of Open Performance Data

The previous sections highlighted the Open Performance Data Initiative from a generic high-level perspective. It provided the motivation and background to establish and promote an open data approach for performance monitoring in Europe.

The following application cases build on identified priority areas and comprise an overview of the potential – both in terms of advancing the state-of-the-art of performance monitoring, but also providing and fostering a dialogue to advance the data-analytical approach to performance measurement or develop jointly accepted algorithms.

It is not intended to provide a complete literature review of work within the performance domain based on open data. However, the snapshots provide an overview on the many avenues to pursue the further development of the data analytical framework (potentially resulting in supporting open-source software) to facilitate the discussion about performance.

The selected application cases comprise:

- Scope of application
- Climate change, emission estimation of air transportation
- Event detection of flight trajectories

- Operational performance monitoring at airports.

"Summary of papers using OSN data as input"[3]

This work presents a comprehensive air traffic dataset, derived and enriched from the full OpenSky Network data and made publicly available[4]. It spans all flights seen by the network's more than 3500 members since 1 January 2019.

The archive is being updated every month and for the first 18 months includes 41 900 660 flights, from 160 737 aircraft, which were seen to frequent 13 934 airports in 127 countries. The dataset has proven to be popular with about more than 67,000 downloads as of end of May 2022.

The use of the dataset is not limited to air transportation related topics. Some authors used it as an indicator of economic activity (at a given airport or region or globally). Examples of such use of data provided by OpenSky Network can be found in the following: Bank of England, Monetary Policy Committee (2020), International Monetary Fund (2020), or United Nations Department of Economic and Social Affairs (2020). The data has also been applied in a wide variety of domains as highlighted in Figure 3 spanning Artificial Intelligence/ Machine-Learning, Security (signal security, data security), Earth Science (climate impact of aviation), COVID (air transport developments during COVID phase), and other aviation related topics.
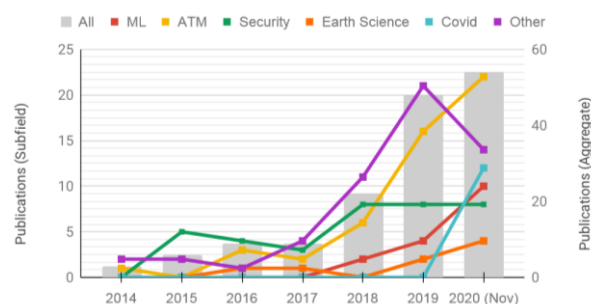
Figure 3 – Fields of application of Open Performance Data

---

[3] "Research Usage and Social Impact of Crowdsourced Air Traffic Data", Martin Strohmeier, Multidisciplinary Digital Publishing Institute Proceedings 59.1 (2020): 1.; "Summary of selected papers using OSN data as input",

Martin Strohmeier, Xavier Olive, Jannis Lübbe, Matthias Schäfer and Vincent Lenders

[4] https://zenodo.org/record/6411336#.YoTuY5NBzS4 last access: 18 May 2022.

It goes without saying that an open data-based approach will enable a wider use of aviation related performance data beyond today's narrow application. At the same time there is the potential that additional perspectives and datasets can be discovered to foster the performance view on air transport and air navigation.

"Evaluation of Aviation Emissions and Environmental Costs in Europe Using OpenSky and OpenAP"[5]

Authors propose a data-driven approach that estimates cruise-level flight emissions over Europe using OpenSky ADS-B data and OpenAP emission. Flight information, including position, altitude, speed, and the vertical rate are obtained from the OpenSky historical database and aircraft performance models have been obtained from OpenAP, an open-source aircraft performance model developed and maintained by Delft University of Technology.

This study makes use of the first four months of flights in 2020 over the major part of Europe and covers the period before and at the start of the COVID-19 pandemic. The aggregated results show cruise-level flight emissions by different airlines, geographic regions, altitudes, and timeframe (e.g., weeks). In addition, environmental costs associated with aviation in Europe by using marginal cost values from the literature is estimated. The paper is an example how crowd sourced flight data can be employed to rapidly assess aviation emissions at varying spatio-temporal resolutions on a continental scale.
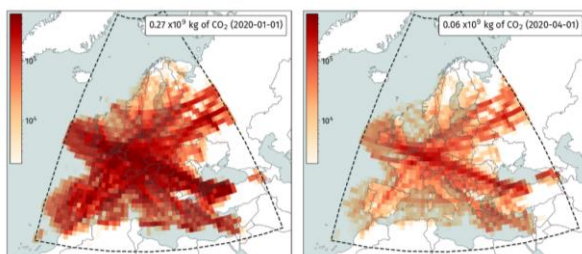

Figure 4 – CO_2 estimation before / after COVID (taken from Ref 5, Fig. 7)

The statistics in this paper were centred around overall CO_2 and other types of emissions during the beginning of the COVID-19 pandemic. Aggregated emission statistics per airline, flight information regions, and flight levels were also provided. Based on the aggregated emission results, it was found that, on average, between 0.25 and 0.30 billion kg of CO2 was emitted daily over Europe before COVID-19 in 2020. Authors further estimated the environmental cost of such emissions based on related global warming and air quality degradation. The combined cost is estimated to be around 70 million USD (in 2015 value) every day for the European region, which is mostly driven by air quality impacts.

Estimations were performed for almost all commercial flights, accounting for around 96% of flights in the dataset obtained from OpenSky. The processing time required for estimating emissions is approximately 30 min for computing flights during one day of all European operations (pre-COVID-19), with a single regular desktop computer.

This example indicates that open data approach is promising for providing near-real-time estimates and monitoring of aircraft emissions. It also combines open air transport data with an open emission model.

"Detecting Events in Aircraft Trajectories: Rule-Based and Data-Driven Approaches"[6]

This paper proposes a taxonomy of significant events, including usual operations such as take-off, Instrument Landing System (ILS) landing and holding, as well as less usual operations like firefighting, in-flight refuelling and navigational calibration. Based on the taxonomies a set of rule-based and statistical methods for detecting a selection of these events are developed. The automatic identification of operational events associated with trajectories will present a wide range of possibilities for monitoring and post-operational analysis of air traffic performance.

---

[5] "Evaluation of aviation emissions and environmental costs in Europe using OpenSky and OpenAP." Sun, Junzi, and Irene Dedoussi (2021)

[6] "Detecting Events in Aircraft Trajectories: Rule-Based and Data-Driven Approaches." Olive, Xavier, et al. (2020).
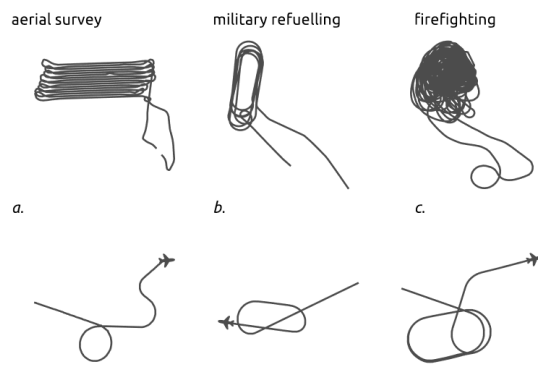
Figure 5 – Pattern detection (taken from Ref 6, Fig. 8)

The authors show that external structured knowledge about flight plans, airspace structure, procedures and operational practice is essential to implement relevant rule-based detection mechanisms, whereas statistical machine-learning based methods are of great help when access to such information is sparse. Rule-based methods are easy to describe. They present robust and explainable results but are subject to many parameters to adjust in order to select corner cases only fitting part of the implemented definition, albeit widespread in the operational world. ML-based statistical methods on the other hand excel at extracting knowledge without or with poorly specified operational input; however, significant expertise is essential to interpret the results.

This is an example of work that supports the identification of methodological alternatives to investigate drivers for operational performance. It could for example help to support the identification of filtering criteria or stimulate to assess how non-nominal operations influence the current performance metrics.

"Towards Measuring Operational ANS Performance at airports with Open Data"[7]

The GANP and PRC measurement framework of operational performance at and around airports identifies performance indicators for runway system capacity/throughput, ground movement efficiency, and efficiency during the arrival phase. These indicators can be complemented by traffic

and movement statistics and characterisations. For example, the current indicator for the efficiency in the arrival phase uses a triplet of aircraft type, landing runway, and observed travel time from entry into the arrival sequencing area around an airport to the landing of the aircraft. The assessment of taxi-times requires identification of stand/gate positions and the used runway.

The metrics are characterised by establishing a flight-by-flight record with the required milestones, e.g., movement events such as actual off-block time, actual take-off time, entry into arrival airspace. To date, these data are collected through the EUROCONTROL standard for airport operator data. The latter process requires airport operators under the EUROCONTROL Performance Review System to provide such flight-by-flight records by the end of the month following the month of operation. This introduces a significant time gap between the operation and the performance monitoring.
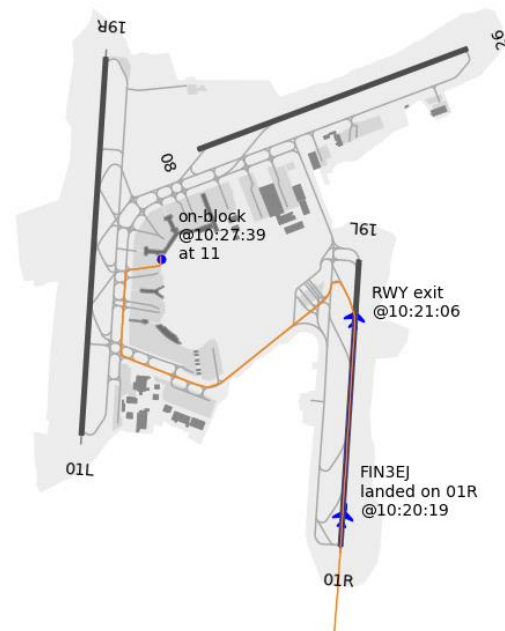

Figure 6 – Identification of key movement milestones

The feasibility study conducted at Stockholm Airport (ESSA) entailed the integration of low-cost

---

[7] "Towards Measuring Operational ANS Performance at Airports with Open Data." Selistrov, Dean, et al. (2022) – unpublished.

ADSB/Mode S receiver data in the OpenSky Network data. With the use of an open-source software package, the data were extracted from the database and processed. The analytical software was extended to establish an APDF conform flight-by-flight data record and calculate the current operational performance measures for the additional taxi-in, additional taxi-out, and additional time in terminal airspace. Additionally, the processed data set was used to provide movement related statistical data.

This example demonstrates the inclusion of data sources into the crowd-sourced network, the use of open-source software, and associated expansion of analytical tools to establish near-real-time operational performance monitoring for European airports. Next to reducing the reporting time, this approach can also serve to reduce the human effort in the data collection under the APDF standard.

PRC Position – Facilitating Change through Open Data

This PRC Performance Insight provides an initial view on the Open Performance Data Initiative. It builds on the general move to open data and address the need for higher levels of transparency within the air transportation and air navigation communities. The societal challenges on our industry, ambitious political goals, and discussion about the required modernisation of air navigation will require open access to the underlying data, reproducibility of the reported results.

The PRC is committed to establish this level playing field by establishing a platform via which the public, policy-makers, strategic planners, practitioners, and researchers have access to data (i.e. open access for non-commercial use), jointly support the collection and curation of the data (e.g. crowd-sourcing, integration of further datasets), and harmonise on the refinement and development of performance metrics (joint data analytics, open source tools).

The PRC invites others aviation stakeholders willing to explore the benefits of this initiative to join in the provision of additional data.

The Open Performance Data Initiative requires a cultural change within the industry to embrace the benefits of open data for performance analytics and governance. The initiative is therefore a longer-term project and will be implemented in stages.

To implement the initiative a set of initial activities have been identified:

- establish a community wide cooperation for the collection of underlying operational data through a combination of crowdsourcing and organisational data;

- strengthen the access, usability, of data;

- promote the utilisation of open performance data through use-case analyses, further development of PRC's performance monitoring data products;

- expand the surface coverage for airports; and,

- a web-based landing page for stakeholder interested in utilising the open performance data for monitoring and analyses but also for researchers and data providers to disseminate their results and achievements.

How can you help?

The PRC invites interested stakeholder to join and support the next steps of the Open Performance Data initiative. The PRC is looking for technical collaboration to help:

- **augmenting the existing open data coverage at airports**. This can include the establishment of an interface to local surveillance data feeds or the placement of low-cost passive ADSB/ Mode S receivers.

- **developing use cases for supporting performance-oriented developments**. This can studies based on open data on a variety of performance monitoring related matters (e.g., efficiency of arrival sequencing, fuel burn estimation for specific flight phases, special operations at/around airports). The PRC is also happy to discuss other practical application cases.

Open Performance Data Initiative

Develop technical framework, platform, and cooperation to leverage open performance data

Facilitate evidence-based dialogue with stakeholders to expand the scope of open performance data in support of all key performance areas

Drive participatory process for further uptake of open performance data and to further joint development and harmonisation of data analytical approaches, tools and monitoring products