# SESAR Engage KTN – PhD final report

| | |
|---|---|
| PhD title: | DiSpAtCH |
| Candidate's name: | Jonas Langer |
| Lead supervisor's name: | Prof. Peter Hecker |
| Co-supervisor's name: | Dr. Thomas Feuerle |
| Proponent institute: | Institute of Flight Guidance – University of Braunschweig |
| Thematic challenge: | N/A |
| Edition date: | 18 August 2022 |
| Edition: | 1.1 |
| Dissemination level: | Public |



This project has received funding from the SESAR Joint Undertaking under the European Union's Horizon 2020 research and innovation programme under grant agreement No 783287.

# 1. Abstract

The general objective of DiSpAtCH (Decision Support System for Airline Operation Control Hub Centre) is to elaborate on artificial intelligence technologies and how these technologies could efficiently support decision making in an Airline Operation Control Hub Centre (OCC) in unexpected or very complex situations.

The daily operation of airlines is often disrupted by unplanned events. As an airline it is therefore essential to operate an OCC to be able to react and mitigate any consequences from the initial disruption. The most challenging task is the information management task. This task includes monitoring, recognition and projection of relevant information out of all information available including current and future situations.

Today the decision making process mainly relies on the experience of the staff working in the OCC. Like in other industries, the desire of using Decision Support Tools (DST) based on machine learning (ML) algorithms is also increasing in the aviation industry. ML algorithms, like neural networks, need a large amount of data to be trained with. The focus of DiSpAtCH is to develop a DST which aims to help the staff in an OCC during disrupted situations. Therefore, three ML modules have been defined of which one aims to propose a suitable action/solution in a disrupted situation. To train the algorithm a database including information about disruptions as well as the implemented solutions from past disrupted situations is needed. Since these kinds of data are not available to researchers and often not recorded by airlines themselves, an approach was needed to get some data to start training algorithms and to validate that certain DST can be developed and support the disruption management process within an OCC. With a decision support system like DiSpAtCH the decisions within an OCC can be optimized which will result in fewer overall disruption cost.

DiSpAtCH provides an approach of using an airline simulation to generate generic operational data of an airline and its daily operations. Synthetic data are generated and ML algorithms are trained to predict actions/solutions for disrupted situations. A first validation shows that a four step classification process including two neural networks can be used to predict actions/solutions in disrupted situations with an accuracy of around 95% and therefore reduce the overall disruption cost by 61% compared to randomly selected actions/solutions.

## 2. Objective of the study

To better understand the objectives of this study, it is essential to be familiar with the current disruption management process of airlines. The airline industry is an industry with high potentials to get disrupted by many external influences (see Figure 1).

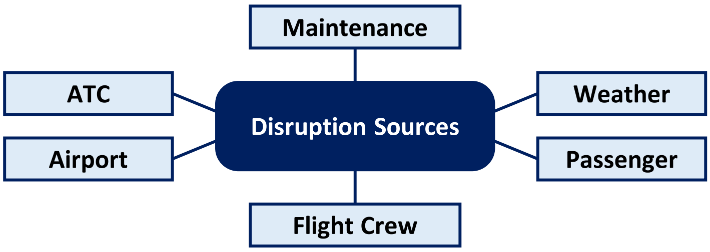


*Figure 1 Disruption Sources [1][2]*

To counteract disruptions, airlines operate so-called Operation Control Centre. The main tasks of OCCs are to control the operation continuously and aim to identify possible disruptions as early as possible to initiate actions to reduce the impact or even prevent any consequences at all. Finding a solution for a specific disruption scenario is often difficult since decisions are driven by several factors e.g. cost and available resources. A solution must not only address the obvious disruption, it must also be a feasible solution considering passengers, flight crew, airport, weather, maintenance, and ATC [1][2].

This makes it already a very complex decision-making process. Nevertheless, several forecasts (before COVID-19) showed a probable continuous increase in the total number of flights for the next 20 years [3][4][5][6]. OCCs must therefore, prepare themselves to handle not only more aircraft but also to solve disruptions in a much more complex environment, which makes the disruption management process more difficult [7]. The current situation with a worldwide restart of airline business will allow airlines to implement and use novel and innovative approaches like DiSpAtCH. Currently, the operational disruption management process of airlines is typically divided into five steps (see Figure 2).

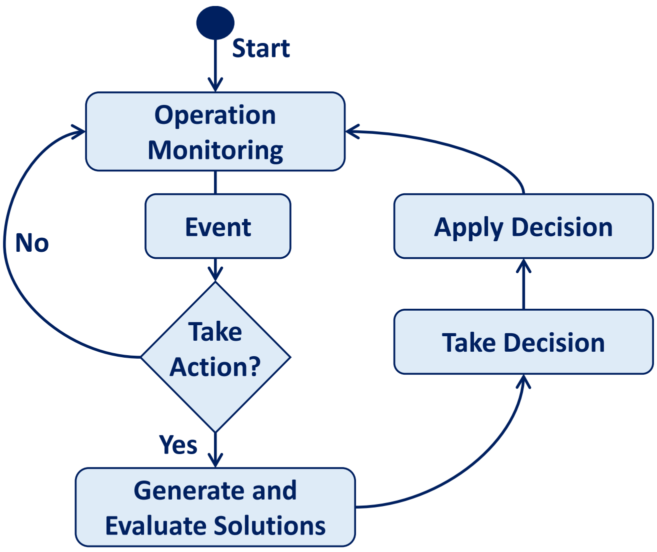


*Figure 2 Disruption Management Process [8]*

During the phase of operational monitoring, all flights are monitored to identify any deviation from the plan. As soon as an event happens, the situation is assessed and a decision for further action is taken. If no action is required, the operation monitoring continues again as normal. In case the event causes any disruptions, solutions are generated and evaluated to restore the initial plan. This step is often supported by computer tools and carried out by experienced controllers. To find the right action for a specific disruption can be challenging due to several restrictions that need to be considered (e.g. crew availability or airport resources) throughout the process (see Figure 3). After the decision was taken the action is applied to get back to the step of monitoring the operations [1][8].

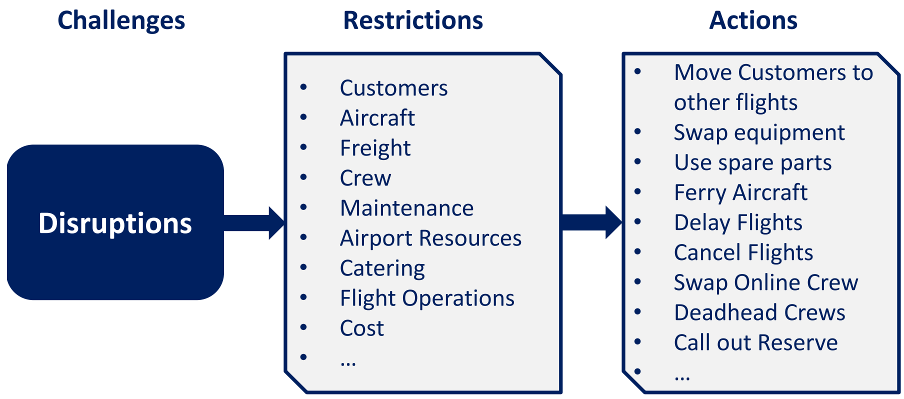


*Figure 3 Decision Making Model [1]*

If the disruption management process of airlines will get more challenging and complex in the future, more sophisticated computer tools are needed. Today, disruptions already cost the aviation industry billions of dollars annually [9]. Airlines therefore would benefit from a DST during the step of "generation and evaluation of solutions" and "take decision". With more advanced algorithms (e.g. machine learning) and more data that is available in a digital format in the aerospace sector the need for a new approach of data-based DST during the decision making process is growing.
ML algorithms can provide insights in patterns and structures within datasets. Furthermore, by using labelled data they learn to predict specific behaviours or events. The advantage of this completely data-driven approach is that it is not biased by individual experiences of humans and can therefore achieve more accurate results in many cases [10].

In summary, the objective of the DiSpAtCH project is to elaborate on ML and artificial intelligence technologies and how these technologies could efficiently support decision making in an OCC in unexpected or very complex situations. Airlines like Lufthansa and Swiss already started to work with Google to research on a very similar topic [13].

Based on the findings in the literature precise research questions for DiSpAtCH were defined:

- How can the needed data for the training of ML algorithms be generated?
- What inputs/resources are necessary?
- How can synthetic data be used to train, test and validate a DST which should later be used in the environment of real airline operations?
- Which ML algorithm seems most promising for this application?

# 3. Motivation

The use of DST based on ML algorithms is increasing in many industries. With the growing application possibilities of such tools, the need for training data is also growing [14][15]. ML algorithms like neural networks promise good results in some applications but without an extensive amount of training data, good results would not be possible [14][16]. In OCCs many decisions are taken daily which gives the chance to record a lot of data. Airlines also desire to use more sophisticated tools within the OCC but the needed training data are not available in sufficient quantity. Airlines often don't see the immediate benefit of recording e.g. all decisions taken to solve a disruption. As long as the benefits from recording such data are not clear to airlines, they most probably don't spend the additional time and money to record them [17]. Therefore, it is difficult for airlines and researchers to start working on novel DST based on ML algorithms. Furthermore, the current ongoing pandemic also makes the start of recording operational data more challenging since the daily operations are either not as repetitive and planned as before the pandemic or there is no willingness to spend the extra money. Also flight plans or the overall network structure of an airline may change frequently, that there is not enough time to capture enough data from real operations to train ML algorithms to achieve the needed prediction accuracy to be used as decision support.

This challenge can be addressed by helping airlines to see the benefit of recording the mentioned data. This can be achieved by developing DST and training algorithms by using synthetic data from an airline simulation. The airline simulation ensures that enough training data can be generated in a short time. If there is a change of flight plans or the overall network structure, the ML algorithms can be updated and trained again on synthetic data, so that the DST are still available immediately after any change in the daily operations of an airline.

As a result, not only possible applications of ML algorithm can be developed, tested and validated, but also a clear recommendation can be made to airlines including what applications of ML in the OCC seems most promising and what kind of data would be needed to be recorded. Based on the described benefits airlines may decide that the additional cost and time spent to record operational data are outweighed by the opportunities associated with novel ML algorithms.

But not only airlines would benefit from such an airline simulation tool. Researchers can use the generated data for their research projects where currently no real operational data are available, so that different research approaches are based on a standardized dataset. This increases the number of possible research projects as well as the comparability of the results from different research projects.

# 4. Advances this work has provided with regard to the state of the art

Since DiSpAtCH pursues the goal to establish novel ML algorithms to the disruption management process of airlines, several OCCs have been visited (1 holiday carrier and 4 legacy carrier) to not only identify the state of the art in disruption management, but also to ask staff in charge about their experience and needs regarding DST. Based on the visits and the interviews, three main needs were identified:

- Ability to analyze the current disruption situation and the restrictions regarding the available resources
  → enhance situational awareness
- Automated generation of preferred action/solution regarding the current disruption situation considering the available resources
- Comparison of disruption solutions regarding their overall costs and time impact on the operation

This initial field study ensured to get an overview of the current state of the art and therefore a base from where novel ML algorithms can be developed. Furthermore, based on the insights and the feedback during the interviews five hypotheses were defined which should be verified at the end of the DiSpAtCH project. The hypotheses (H) are the following (see Table 1):

*Table 1 Overview of the defined hypotheses*

<table>
<tr><th>Hypotheses ($H_0$)</th><th>Quantifiable (How to measure)</th><th>Testable (How to test)</th></tr>
<tr><td>H1: The DST contributes to increased situation awareness during disruption situations.</td><td>Several techniques are possible to assess situation awareness (freeze probe, real-time probe, post-trial self-rating, observer-rating, performance measures, and process indices). The selection of a technique will be done during the design of the validation campaign.</td><td rowspan="5">By using the developed airline simulation tool, in theory all hypotheses can be tested at the same time.<br>The developed visualisation will provide information about the current operation. Furthermore an input mask will provide the chance to select actions for each disruption.<br><br>The participants of the test and validation campaign should vary from experienced people from an airline to people with no connection to the field of aviation. They will be in the role of the duty manager and they have to decide which action should be implemented for each disruption.<br><br>To test and validate the impact of DiSpAtCH, each participant will have to simulate two runs (e.g. two operational days/weeks). The first run will be carried out with only the basic information about the operation and no additional support tool. The participants have to assess the disruption situation themselves and decide which solution to implement on their own. In the second run they will be supported by DiSpAtCH with all its developed algorithms. This time they can use the support tool to find and select a solution for each disruption situation.</td></tr>
<tr><td>H2: The DST helps to reduce the needed time to find feasible solutions for specific disruption situations.</td><td>This can be quantified by measuring the time needed to find a suitable solution after a disruption occurs.</td></tr>
<tr><td>H3: The DST helps to find solutions that minimize the time impact on the overall operation.</td><td>The airline simulation tool gives the opportunity to easily calculate the delay of each flight and therefore the delay of all flights of e.g. one day of operation.</td></tr>
<tr><td>H4: The DST ensures that the costs of actions to counteract disruptions are decreased in comparison to today's airline disruption costs.</td><td>For each action to counteract disruptions, a fixed cost is allocated. By adding up the individual cost of each action for a fixed operation period e.g. one day, the overall cost of all actions can be calculated.</td></tr>
<tr><td>H5: The DST enables to solve disruptions with fewer resources.</td><td>Resources from an OCC perspective are defined as aircraft, crews and fuel. Especially the use of backup crews as well as aircraft can be seen as additional use of resources and are measurable.</td></tr>
</table>

With these requirements the Framework for DiSpAtCH was developed (see Figure 4).To increase the acceptance of a new DST several interview partners emphasised the importance of the comprehensibility of the results proposed by the DST. Therefore, during the development of the framework the implementation of an optimiser was avoided. The focus is more on situational awareness and an overview of decisions in past similar situations. This should help to increase the acceptance of the developed DST described in the following chapter.

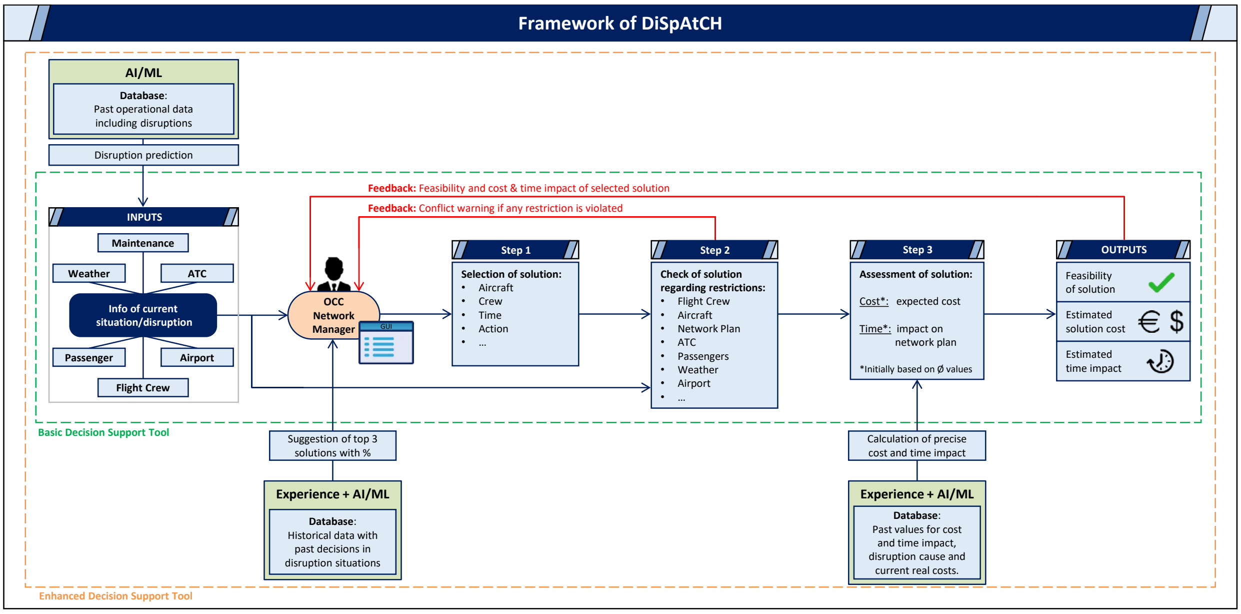


*Figure 4 Framework of DiSpAtCH*

The network manager is a central part of the framework and even with new ML based DST the network manager still has the final decision. Within the basic DST a three step process is presented with several feedback loops to the network manager. Step 1 represents the selection of a desired action/solution to counteract a current disruption. In step 2 the selected action/solution is checked against the restrictions/resources and in step 3 the cost and time impact are calculated. If an action/solution is feasible, a positive feedback is given to the network manager including the calculated cost and time impact. Otherwise a conflict warning is given after step 2 if any restriction/resource does hinder the implementation of the desired action/solution. Only if an airline has implemented the presented basic DST in any way and has started to record operational data, the enhanced DST including ML modules can be applied. Overall three possible applications for ML modules were proposed to satisfy the identified needs:

**Module 1:** Disruption probability prediction based on past operational data including disruptions.

**Module 2:** Suggestion of proposed action/solution for a specific disruption based on recorded decisions from past disruption situations.

**Module 3:** Prediction of precise cost and time impact of a selected action/solution and the current disruption.

All three ML modules only work if sufficient amounts of data are available to train and validate the algorithm of each module. As described earlier, such data are currently either not recorded by airlines or not accessible to researchers. Therefore, DiSpAtCH focusses on using synthetic data from an airline simulation to find out what kind of data are needed for each ML module to work and what results can be expected from each module. The research results of DiSpAtCH could also be valuable to airlines, since DiSpAtCH will be able to recommend which data are needed in which form and quantity in order to develop the corresponding ML modules. The final recommendation can be used by airlines to focus their efforts in data recording and later algorithm training.

By focusing on the goal of DiSpAtCH a research in the field of airline operation and DST was carried out. An analysis of related work, which included over 100 papers, shows an increasing interest in the field of airline disruption management. Especially in the past ten years the research focuses more on approaches of disruption management which includes more than one resource in the solution process. A resource could be e.g. the aircraft, crew or passengers. The analysis identifies airline disruption management as an ongoing research field and proposes that with higher accessibility of data more and more ML techniques could be applied to disruption management problem [14].
First applications of ML have proven that it can be used in the field of airline OCCs but only in a very specific field (point to point network and weather delay) since no other data were available. Missing data in general hinders the progress of new disruption management solutions [16]. In other industries the application of reinforcement ML algorithms seems promising. Reward functions can represent e.g. cost or performance which is then used to train the algorithms to increase the reward [19]. This approach can also be applied to the airline disruption management where daily Key Performance Indicator (KPI) could be used as reward function.
Because of the lack of complete and applicable data sets from an airline other approaches are needed. An airline simulation could be a solution as similar approaches in other applications have already proven to be successful. For example the following application in the field of wind farms shows how a simulation can be used to train an algorithm for power optimization of wind farms for later use in the real environment. The power optimization of wind farms is a complex process, since each wind turbine within a wind farm would impact other wind turbines due to the wakes it generates. To increase the overall power production an individual setting e.g. in yaw for each wind turbine is needed. Since real data in different wind conditions were difficult to record a high fidelity simulation was used to generate training data for the selected algorithms. By using defined performance parameter the training of the algorithms were improved over time. As a result an application for yaw control of each wind turbine to increase the overall power production was developed. The algorithms could be trained offline with data from the simulation and later be fine-tuned with real data from the wind farm as soon as it goes online [15].

In summary, examples from other industries like the optimization of the wind farm power production showed that simulations can be used to train algorithms offline without access to real data and later be fine-tuned as soon as the applications go live in the real environment Therefore this promising approach is also used in DiSpAtCH. The developed airline simulation gives new opportunities for airlines and researcher. Needed data for algorithms training can be generated relatively fast and in any desired quantity. Since there was no access to such data before DiSpAtCH, also other researcher or airlines can benefit from the airline simulation by using its data to train their own algorithms.

# 5. Methodology

The applied research methodology is shown in Figure 5 and gives an overview of the phases of the project. Overall it is divided into three main phases and each phase consists of two tasks.

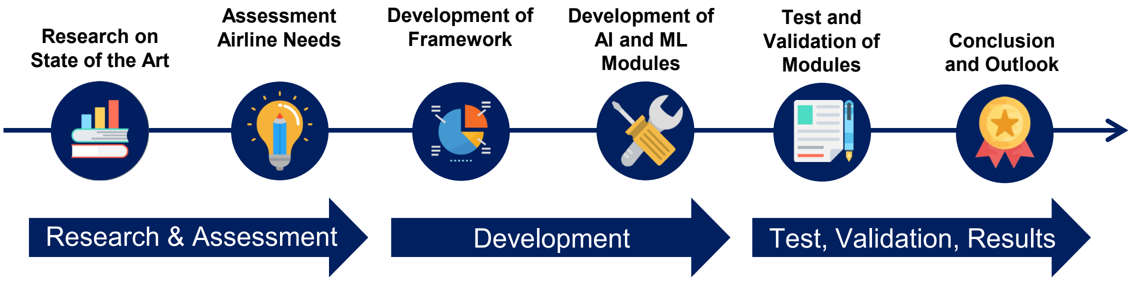


*Figure 5 Research Methodology for DiSpAtCH*

In the first phase **Research & Assessment** the main goal is to research the current state of the art in the area of ML, airline operation and DST as well as to assess the airline needs regarding support during the disruption management process. A comprehensive literature review and several interviews with people in charge of OCCs were carried out.

In the second phase **Development** a framework for a new decision support approach is designed and the airline simulation was developed. Furthermore, training data was generated, several algorithms are trained and modules for the DST are created.

The final phase **Test, Validation, Results** consists of tests and validations of the developed modules and the overall DST. Finally, a conclusion and outlook will summarize the main achievements of DiSpAtCH as well as the impact on the disruption management process from an airline perspective.

# 6. Description of the data the study relies on

Especially the developed airline simulation relies on external data. To setup the environment for running a simulation, which aims to be as realistic as possible, four external data sources were used. Table 2 gives an overview of what data was used.

*Table 2 Overview of used data*

| Data | Description |
|---|---|
| Airport database | Worldwide airport database including names, ICAO codes, UTC time zones and coordinates [20] |
| Disruption database | Disruption probabilities and delay distribution (in min) of a partner airline based on the IATA Delay codes of tracked flights between 2017 and 2019 [23] |
| Connection Statistics | Connection probability based on tracked flights by EUROCONTROL between 2015 and 2018 [24] |
| Cost Statistics | Cost Statistics for delays and e.g. flight cancellations based on cost report from EUROCONTROL [21] and Westminster University [22] |

Each of these databases is used to build the overall airline simulation and make the generated data as realistic as possible.

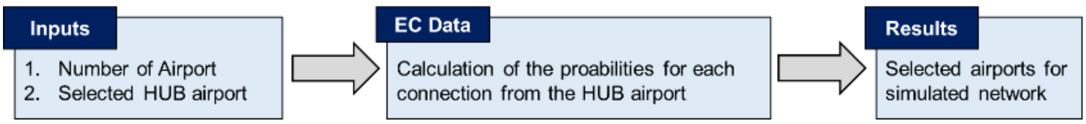


*Figure 6 Workflow for Airport Selection*

The **airport database** is used to select real airports and follows the workflow shown in Figure 6. First the input mask is used to select the number of airports within the desired network, as well as the selected HUB airport. In the second step, the past flights from the selected HUB airport are analysed and a table with all tracked connections (destinations) is generated. For each connection a probability is assigned based on the frequency of past flights. With the given probabilities the needed number of airports is drawn from the table. As a result, a network is generated which is based on real flown connections and should therefore represent real airline networks with the same selected HUB airport. Also, the connection distances are calculated by using their coordinates and in combination with the UTC time zones night curfews can be easily implemented. The **disruption database** is used by a disruption generator function, which is called for each flight in the simulation. With the given disruption probabilities and the corresponding delay distribution it is possible to assign disruptions and delays in minutes to flights based on IATA delay codes from a real airline. The data from the partner airline which is used in the basic simulation can of course be adapted by any user e.g. an airline by using their own experienced disruption probabilities and delay distributions. With the **connection statistics** from EUROCONTROL it is possible to generate network structures based on the selected HUB airport and the probabilities of connections based on past tracked flights. This further increases the realism of the generated airline networks. The **cost statistics** from Westminster University and EUROCONTROL are used to calculate the cost of disruptions and selected actions/solutions.

In a later application of DiSpAtCH to the operation of a real airline, all these data sources would preferably be provided by the airline.

# 7. Computational experiments

The use of an airline simulation was the main driver for this research and helped to generate data to train the ML algorithms as well as providing a platform for development and validation of the DST and its graphical user interfaces.

The airline simulation can be seen as an adjustable platform which allows the simulation of different airline network structures with its special disruption probabilities and available resources which are needed to implement actions/solutions.
To set up the airline simulation, several settings can be adjusted to generate the desired use case. Figure 7 shows the input mask which allows us to specify the settings of the desired airline to be simulated as well as the desired simulation time and steps. To create a specific airline, the following values can be adjusted:

- Number of airports overall
- Number of HUB airports
- Percentage of intercontinental airports
- Selected region of HUB airport
- Number of small, medium and large aircraft
- ICAO code of HUB airport (optional)
- Backup Reserve Crew Count
- Backup Aircraft Count

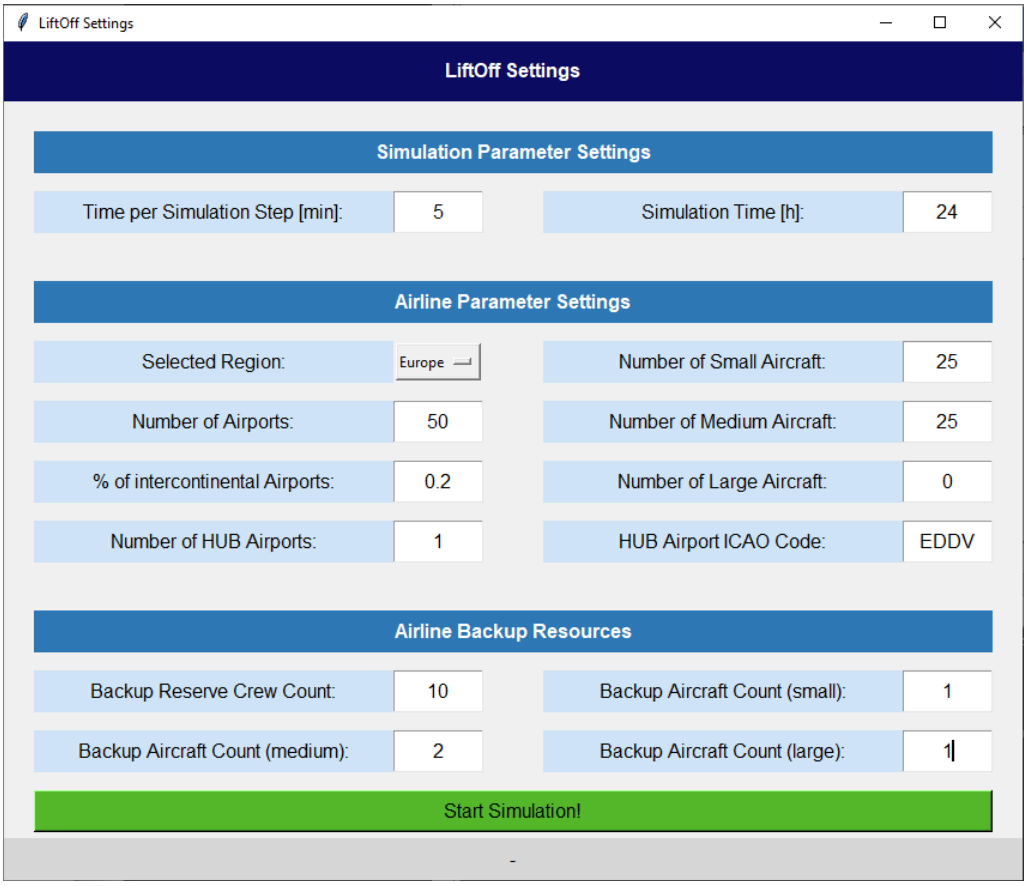


*Figure 7 Input Mask of the Airline Simulation*

Based on the given settings the tool creates a network by selecting airports from a worldwide airport database. Then the aircraft are allocated to an airport and the next destination is selected based on the probabilities of connections from past tracked flights by EUROCONTROL between the years 2015 to 2018. Different network structures can be created by the given number of HUB airports (no HUB airport = point-to-point, one or more HUB airports = HUB and spoke).

With some additional global simulation parameter like the definition of resources e.g. number of reserve crew and aircraft, the night curfew times and the maximum airport capacities per hour, all parameter are defined to run the airline simulation for the desired simulation time. Three aircraft types with assigned PAX capacity and the required number of crew members were also defined (small = 180 seats, 5 crew member; medium = 250 seats, 7 crew member; large = 400 seats, 9 crew member).

Each flight can be disrupted, and disruptions are randomly allocated based on a disruption probability statistic provided by an airline. The statistic includes probabilities for each disruption indicated by an IATA delay code, an average delay value and its standard deviation. Actions/solutions can be implemented as soon as a disruption is allocated to a flight. With defined values like delay cost, several KPIs are calculated for each simulated day. Additional KPIs can be added later if it turns out to be useful.

For now, the following 7 KPIs are used:

*Table 3 Selected KPI*

| KPI | Description |
|---|---|
| Number of Flights | completed flights per 24h (simulated time) |
| Number of PAX | passengers travelled per 24h (simulated time) |
| Overall Delay Minutes | sum of delay in minutes per 24h (simulated time) |
| Delayed Flights | number of flights with a delay >15 min per 24h (simulated time) |
| Overall Cost | cost of disruptions and solutions per 24h (simulated time) |
| Average Delay per Flight | average delay minutes per flight per 24h (simulated time) |
| Average Cost per Flight | average cost per flight per 24h (simulated time) |

Beside the daily KPIs a broad variety of parameters are recorded for each time step. These parameters will later be used as input for the ML modules and the KPIs can be used as a kind of reward function or filter to select the training dataset for the algorithms.

<u>Cost Calculation</u>

It is obvious that a selection of an action/solution will be a trade-off between the cost of the delay and the cost of implementing an action/solution. In addition, the appropriate resources (e.g. a backup crew) must be available for an action/solution to be implemented. With additional cost data about further actions/solutions, they can be easily implemented to the airline simulation as further options. The cost basis for the currently implemented actions/solutions is listed in Table 4.

For the calculation of delay cost, the delay cost statistic from Westminster University is used. The cost of cancellations is based on the cost report from EUROCONTROL. For the use of a backup crew or an aircraft, an organizational cost of 1.500€ is used as well as the delay cost statistic from Westminster University as well as changed aircraft operating cost based on the cost report from EUROCONTROL.

*Table 4 Cost basis for actions/solutions*

| Action/ Solution | Cost Basis |
|---|---|
| Delay a flight | Delay cost statistic from Westminster University [22] |
| Cancel a flight | Cancellation cost based on cost report from EUROCONTROL [21] |
| Use of backup aircraft | Fixed organizational cost (1500€), Delay cost statistic from Westminster University [22], Delay reduction (savings), changed aircraft operating cost based on cost report from EUROCONTROL [21] |
| Use of backup crew | Fixed organizational cost (1500€), Delay cost statistic from Westminster University [22], Delay reduction (savings) |

<u>Airline Simulation - Action/Solution Implementation</u>

As mentioned, the airline simulation does not only consider disruptions, but also actions/solutions. For these actions/solutions to be selected, a new GUI was developed, and it appears for each disrupted flight (see Figure 8). It gives an overview of the current disruption by showing the affected flight and aircraft, the disruption, and the allocated delay minutes. Information about the available resources at the current airport is also given.

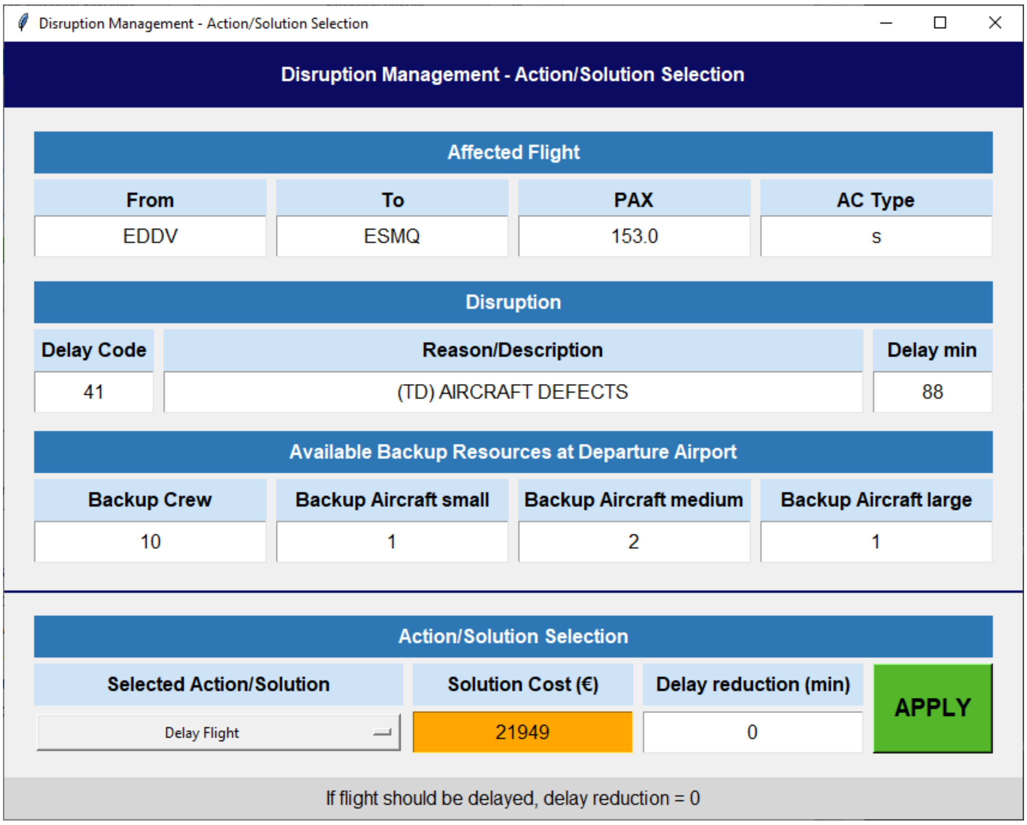


*Figure 8 Airline Simulation – Action/Solution Selection GUI*

With this information it is possible to select one of the 4 defined actions/solutions via the drop-down window. When the action/solution is selected via the drop-down window, algorithms in the background immediately calculate the expected cost as well as the probable delay reduction in min. This gives the opportunity to test several actions/solutions and to see which cost and time impact will result from the decision. The expected cost is also marked with a colour coding in the background for better feedback to the network manager in charge (see Figure 9). If the selected action/solution should be applied, the button “apply” can be pressed and the airline simulation will go on until the next disruption occurs and a new decision needs to be taken.

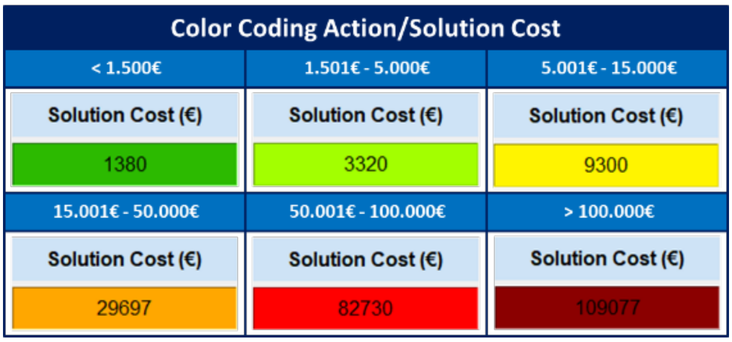


*Figure 9 Action/Solution Cost Color Coding*

# 8. Results

Based on literature research on Artificial Intelligence, ML, and the state of the art in OCC, a framework for the DST was developed. The DiSpAtCH framework shown in Figure 10 is divided in two versions, the “Basic” and the “Enhanced” version. The basic DST contains mainly the developed process without any ML or AI modules integrated. Only after completion of the basic DST the development of the enhanced DST can start. Both levels are defined as follows:

<u>Basic Decision Support Tool:</u>

**Step 1:** The preferred solution is selected by using a graphical user interface.

**Step 2:** The DST is using the information on the current availability of resources as well as the current information about disruptions and the input provided in step 1. The main goal of step 2 is to check if the selected solution is feasible regarding the resources and the current restrictions. If the selected solution is violating any restriction, a conflict warning is immediately shown to the user.

**Step 3:** If the selected solution is feasible, an estimation of the expected cost and time impact is carried out based on average cost and time values and shown to the user.

<u>Enhanced Decision Support Tool:</u>

**Before Step 1:** Use of past operational data and AI & ML algorithms to identify disruptions even before they occur or their impact on cost and time is already huge. Based on a database with taken decisions in past disruption situations an AI & ML algorithm is used to propose the top 3 most probable solutions to solve a specific disruption.

**Step 3:** Based on a database with past values for cost and time impact, disruption causes, and current real prices (e.g. fuel cost) an AI & ML algorithm is used to estimate the precise cost and time impact of that selected solution.

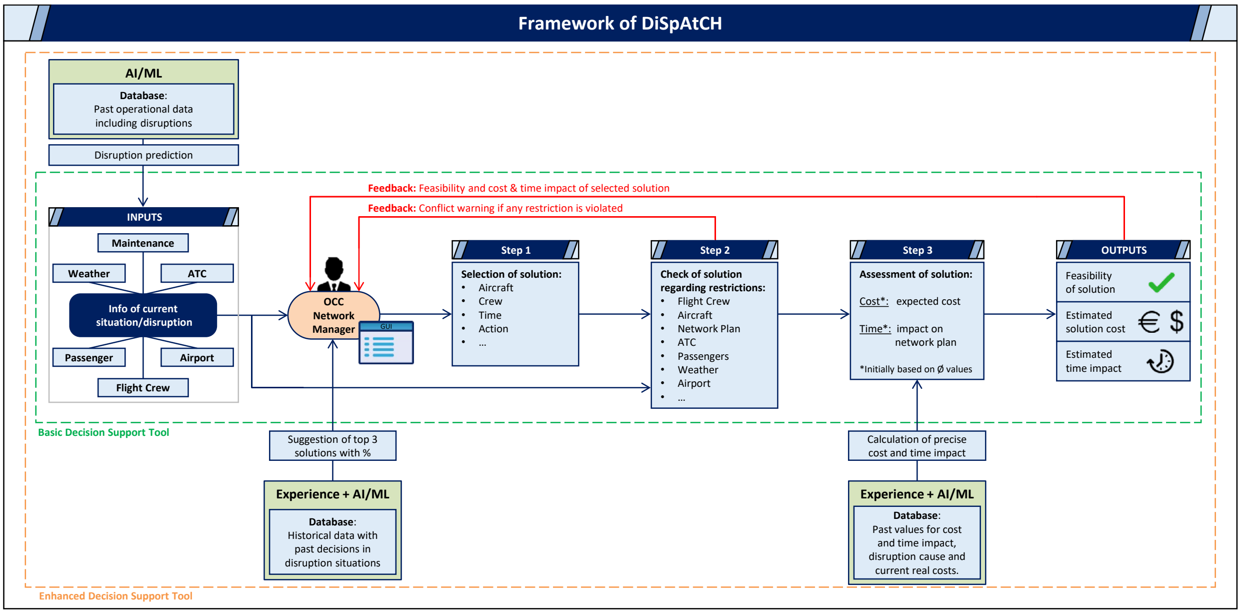


*Figure 10 Framework of DiSpAtCH*

Figure 11 shows an example of a simulated HUB and spoke network with Frankfurt selected as HUB airport (150 airports, 20% intercontinental airports, 1 HUB airport, 30 small aircraft, 25 medium aircraft and 20 large aircraft). The blue dots represent the 150 selected airports and the coloured markers represent aircraft either in flight (green) or on ground at a certain airport (red). Since the pictures are only screenshots of the running simulation, not all connections, only current flights in the simulation are visualized with a dotted line. The vertical red and green lines indicate the beginning (red) and the end (green) of the night curfew. While running they move from right to left. For the data generation the visualization is not needed. But currently the visualization is planned to be used for the final validation of DiSpAtCH. Staff of an OCC can use it to decide which action/solution they would implement, since the visualization provides a quick overview of the current operations and the disruption situation.
Since the airline simulation should be used to train ML algorithms some values are needed as a reward function. For now seven KPIs are calculated for each simulated day. For the cost calculation, the cost statistic is used in combination with some assumptions to calculate cost not only for delay minutes but also for each action/solution which an airline could implement to counteract disruptions.

With the airline simulation it is now possible to simulate a variety of different network structures and disruption scenarios. How this simulation can now be used to generate the needed training data for the ML modules is exemplarily outlined for the ML module 2, which goal is to use a database of information about past disruption situations and the implemented action/solution to train ML algorithms to learn from past decisions what action/solution should be proposed for future disruption situations which are not represented in the training data.
For the proposed action/solution of the algorithm to match the disruption situation and to achieve optimal KPIs, the training data must include datasets which represents decisions which achieved optimal KPIs in past simulation runs. This ensures that the algorithm can easily learn from past decisions what to propose in future. The main goal is to create a database with only optimal decisions regarding the selected KPIs. To select actions/solution the airline simulation is paused when a disruption is allocated to a flight and only continues after an action/solution was selected.
Because the airline simulation doesn't need a lot of processing power, many runs can be completed in a relatively short amount of time. This brings the benefit of using different approaches on generating training data. For now, the actions/solutions are selected randomly and several hundred days are simulated, so that in the end e.g. only the days with the best KPIs are used as training data. This approach follows the idea of Monte Carlo simulations. Another approach could be using the KPI as reward functions for reinforcement learning and each day is seen as an epoch of e.g. a recurrent neural network. If this approach will be used depends on the first results of the ML algorithms comparison and which ML algorithms will be selected as most promising for further research.

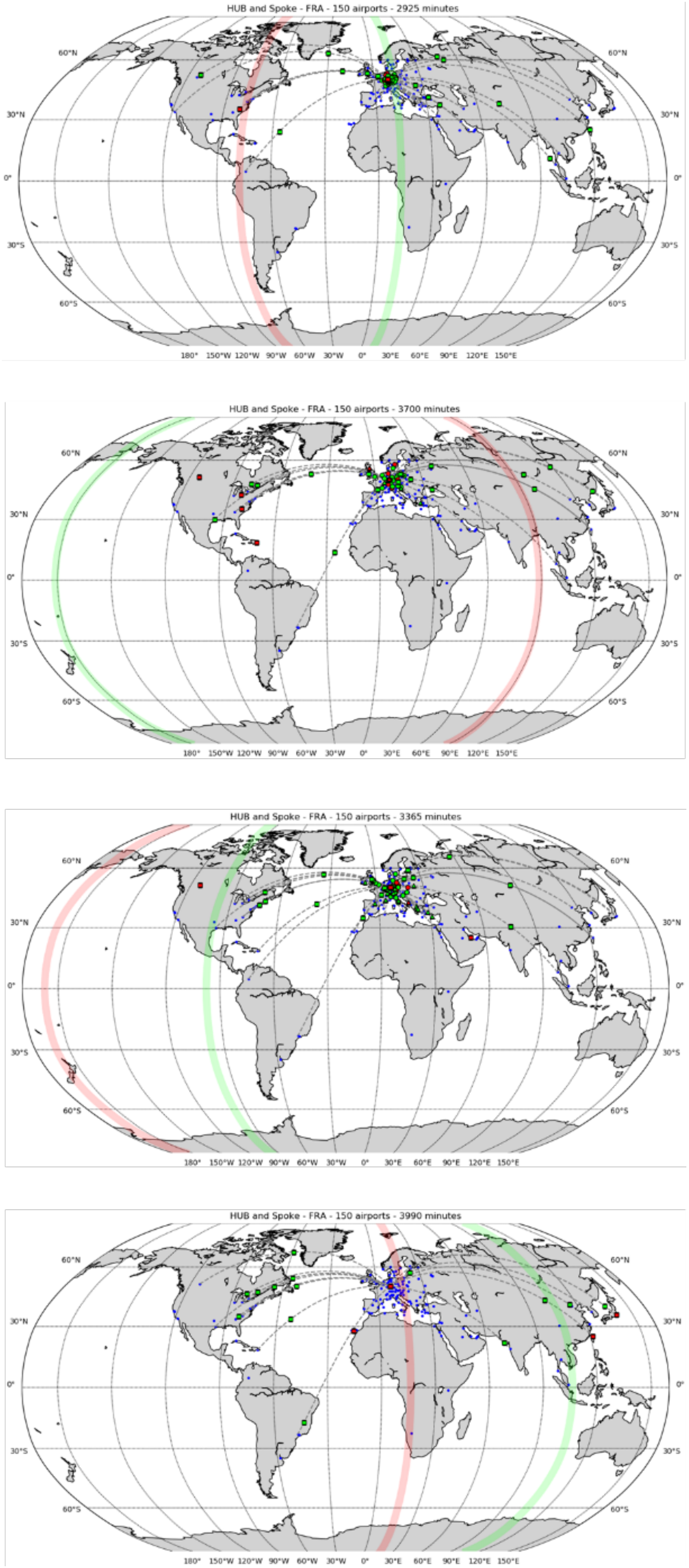


*Figure 11 Simulated HUB and Spoke network at different time steps*

For the analysis of several ML algorithms 100 days were simulated and the actions/solutions selected randomly. For each day all seven KPIs were calculated and for each disruption the taken action/solution were saved. Each data sample contains the information shown in Table 5, which include information about the affected flight as well as the disruption itself (input), the selected action/solution and its cost and effect on the flight (decision) and the overall KPIs of the operating day. In this case a flight was disrupted due to an aircraft defect (IATA delay code 41) which caused an additional delay of 16 minutes. As action/solution it was chosen to delay the flight [ID:1] which increased the time until liftoff from a normal turnaround time of 45 minutes to 61 minutes. Based on the delay cost statistics from Westminster University [22] the delay of 16 minutes for a small aircraft (e.g. A320) causes a cost of 588€.

*Table 5 Example Data Sample*

| | Description | Value |
|---|---|---|
| INPUT | departure airport [ID] | 3972 |
| | destination airport [ID] | 332 |
| | PAX | 153 |
| | aircraft type | 1 |
| | IATA delay code | 41 |
| | local time [h] | 17.75 |
| | delay minutes [min] | 16 |
| DECISION | solution [ID] | 1 |
| | delay reduction [min] | 0 |
| | time until liftoff [min] | 61 |
| | solution cost [€] | 588 |
| KPI | KPI flights count | 26 |
| | KPI number of PAX | 5820 |
| | KPI delayed flights | 8 |
| | KPI delay minutes [min] | 264 |
| | KPI overall cost [€] | 51649 |
| | KPI average delay [min] | 10 |
| | KPI average cost [€] | 1987 |

Before the data can be prepared for the training of the ML module 2, an approach of the desired action/solution prediction was defined (see Figure 13). A two-step classification process should be used to predict the action/solution. In the first decision step a binary classification is proposed. The ML algorithm is trained to predict if an individual action/solution is needed or should be implemented. Therefore, the prediction will either be class 0 (no action needed) or class 1 (individual action/solution needed). If class 1 is predicted in the first step, a second classification algorithm will predict the probabilities for each individual action/solution (class 0: Cancel Flight, class 1: Backup Crew, class 2: Backup Aircraft).

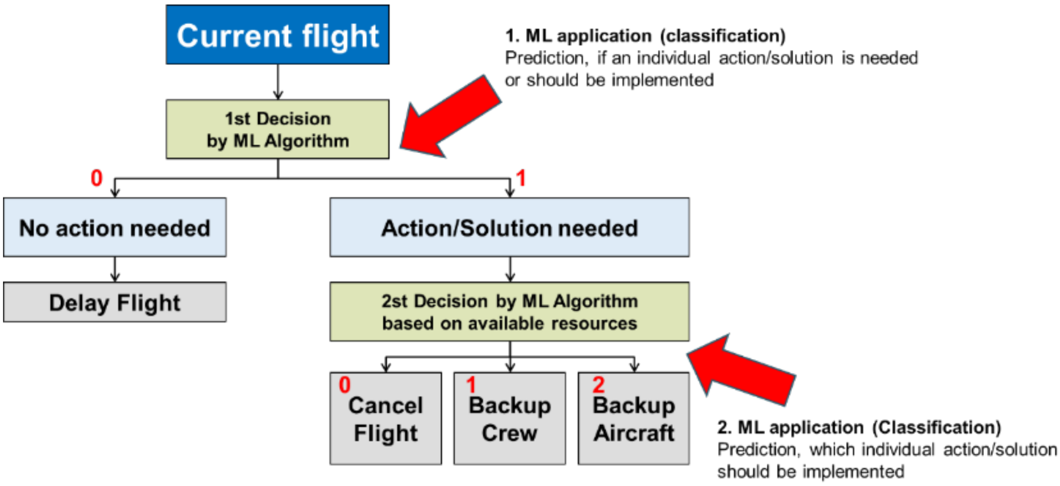


*Figure 12 Two-step classification for action/solution prediction*

The following example describes how the airline simulation and the calculated KPIs were used to generate the training dataset for this two-step classification process. Since each action/solution has an impact on the daily KPIs and the frequency and intensity of the disruptions occur equally over all days and flights, it can be assumed, that on days with best performing KPIs also the best possible decisions were taken. Therefore an approach to select the days with the best performing KPIs was used to define the training dataset.

Each KPI value was normalized between 0 (worst) and 1 (best) based on the overall minimum and maximum values. A new overall KPI value was calculated as the sum of all seven KPIs. A threshold value of 4 was then used to filter the overall data to only consider days with the best overall KPI value. Figure 13 shows an overview of the data pre selection. From around 25000 flights only 12738 flights were carried out on days with an overall KPI of 4 or higher. 11550 flights were delayed (class 0) and 1188 flights had an individual action/solution allocated during their disruption (class 1).

For each algorithm training all 1188 flights from class 1 and 1188 randomly selected flights from class 0 were used to get an equally balanced dataset.


*Figure 13 Data pre selection based on Overall KPI*

After the selection of the training dataset, the next step was further data preparation. Since normalized values (0 to 1) work well with many ML algorithms, a predefined range (min/max) was set for each Input category and the value was then normalized. Besides the normalization also new Input categories were created to get as many binary input values as possible. As an example, the information about the aircraft type was used to create a new input for each aircraft type. This enables a binary input for each category instead of an input with 3 possible classes.

Table 6 shows a data sample and how the initial value for each input changed after the preprocessing was carried out.

*Table 6 Pre Processing of a selected data sample*

| | Description | Value | New value | Range |
|---|---|---|---|---|
| Input | departure airport ID | 336 | 336 | |
| | destination airport ID | 347 | 347 | |
| | HUB indicator (1/0) | 1 | 1 | |
| | backup resource small aircraft | 3 | 1 | |
| | backup resource medium aircraft | 2 | 1 | |
| | backup resource large aircraft | 1 | 1 | |
| | backup resource crew | 3 | 1 | |
| | number of passenger onboard | 153 | 0.3825 | 0(min) to 400(max) |
| | aircraft type (1=small, 2=medium, 3=large) | 1 | 1 | small ac |
| | | | 0 | medium ac |
| | | | 0 | large ac |
| | ICAO code of delay category (0 = no delay) | 0 | 0 | |
| | local time | 5 | 0 | 5am to 11pm |
| | minutes of delay | 10 | 0.0334 | 0(min) to 300(max) |
| Decision | selected solution (1=delay, 2=cancel, 3=backup aircraft, 4=backup crew) | 1 | 1 | selected solution |
| | | | 0 | individual action (o=no, 1=yes) |
| | delay reduction | 0 | | |
| | time until liftoff | 46 | | |
| | solution cost (€) | 16 | | |
| KPI | KPI_daily_flight_count | 250 | | |
| | KPI_daily_pax_count | 49000 | | |
| | KPI_daily_flights_min_15m_delayed | 39.00 | | |
| | KPI_daily_delay_min | 1510.00 | | |
| | KPI_daily_dis_sol_cost | 835757.00 | | |
| | KPI_delay_per_flight | 6.04 | | |
| | KPI_cost_per_flight | 3343.03 | | |

With the completion of the preprocessing, the data is now ready to be used as training data. The dataset was further reduced to only include the needed input values and the information about the class that should be predicted. The following Table 7 shows five data samples of the final training database for the first classification algorithm.

*Table 7 Data samples of the final database for the first classification algorithm*

| | Sample 1 | Sample 2 | Sample 3 | Sample 4 | Sample 5 |
|---|---|---|---|---|---|
| departure airport ID | 336 | 2832 | 336 | 336 | 336 |
| destination airport ID | 2089 | 336 | 3482 | 3781 | 1987 |
| HUB indicator (1/0) | 1 | 0 | 1 | 1 | 1 |
| backup resource small aircraft | 1 | 0 | 1 | 1 | 1 |
| backup resource medium aircraft | 1 | 0 | 1 | 1 | 1 |
| backup resource large aircraft | 1 | 0 | 1 | 1 | 1 |
| backup resource crew | 1 | 0 | 1 | 1 | 1 |
| number of passenger onboard | 0.85 | 0.53 | 0.85 | 0.85 | 0.85 |
| aircraft_small | 0 | 0 | 0 | 0 | 0 |
| aircraft_medium | 0 | 1 | 0 | 0 | 0 |
| aircraft_large | 1 | 0 | 1 | 1 | 1 |
| ICAO code of delay category (0 = no delay) | 87 | 0 | 57 | 87 | 87 |
| local time | 0 | 0 | 0 | 0 | 0 |
| minutes of delay | 0.08 | 0 | 0.07667 | 0.05667 | 0.08 |
| action | 1 | 0 | 1 | 1 | 1 |

For a first selection of a suitable ML algorithm the software Rapidminer was used. An overall of 9 different ML algorithms were trained on the data and compared by their accuracy, class recall, classification error and running times (see Figure 14).

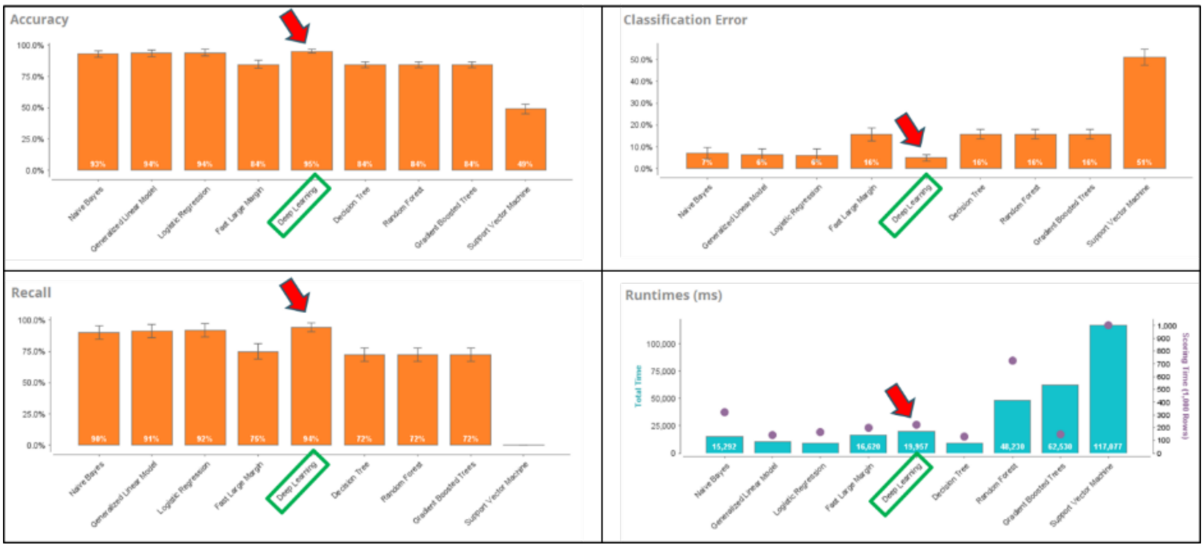


*Figure 14 ML algorithm comparison using Rapidminer*

With 95% the deep learning algorithm achieved the highest accuracy, followed by several algorithms with 93% to 94% accuracy. In the categories recall and classification error, the deep learning algorithm also achieved the best results. The comparison of the runtimes shows, that there are other algorithms that achieved faster runtimes but others needed more time. Overall the deep learning algorithm needed relatively low runtimes to achieve the best performance in all 3 categories. The low runtimes also enable daily/weekly updates of the algorithm during later use in the DST. Therefore, the DST will always include the latest decision taken during disruption and the performance will further improve over time.

With deep learning (neural network) selected as algorithm many setting can be adjusted to improve the performance even more. To tweak the hyper parameter of the algorithm a grid search approach was selected. This approach starts with typical basic neural network settings used in many other publications (see Table 8).

*Table 8 Settings of the Base Model Neural Network*

| | Input Layer | Hidden Layer | Output Layer |
|---|---|---|---|
| **Number of Layer** | 1 | 6 | 1 |
| **Neurons per Layer** | same as input parameter | 50 | 1 |
| **Activation function** | RELU | RELU | Sigmoid |
| **Epochs** | 50 | | |
| **Batch Size** | 25 | | |
| **Learning Rate** | automatic | | |
| **Data Dropout** | No Data Dropout | | |
| **Optimiser** | Adadelta | | |

The first neural network uses the setting defined in the Base Model. From here on a six step grid search was carried out to find good performing network setting. All six steps and the tested parameter settings are shown in Figure 15.

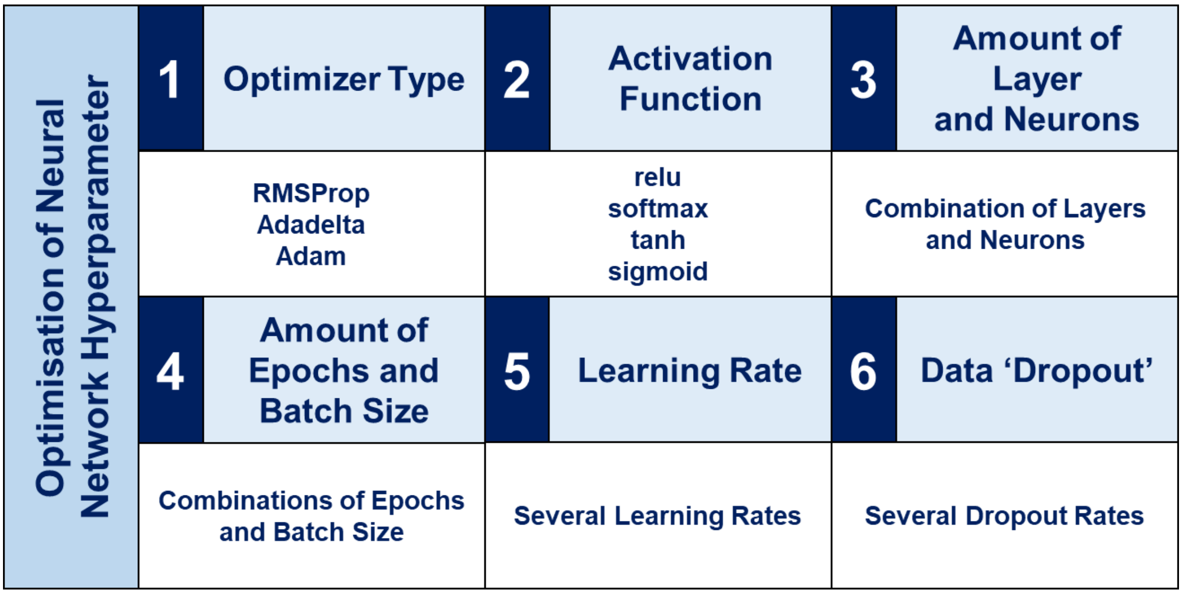


*Figure 15 Hyperparameter Optimization - Six Step Grid Search Overview*

The following Figure 16 shows the grid search results of step 3 where different number of neurons and layers are tested. In this example different network setting with 2, 3 and 4 layer and 100 to 300 neurons (in steps of 25 layers) are compared with each other. The visualization shows clearly, that a combination of 4 layer and 250 neurons per layer performed the best. These comparisons are repeated several time and the show results are average values, to get more valid results.

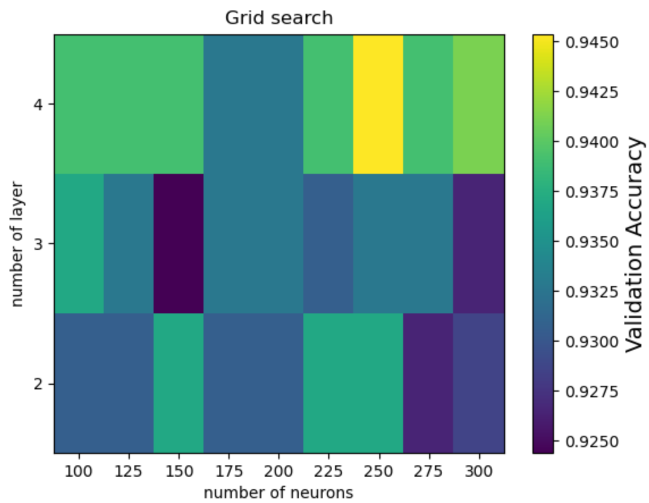


*Figure 16 Grid Search of number of neurons and number of layer*

This process was repeated for each of both ML algorithms shown in Figure 12. During the training process of the first ML algorithm, two rules regarding the delay minutes as input parameter were found. If the delay minutes are higher than 120 minutes, the preferred action/solution should be to cancel the flight and if the delay minutes are lower than 20 minutes the preferred action/solution should be to delay the flight. By adding these decision steps to the overall action/solution classification process (see Figure 17) the runtimes of the DST could be reduced, since in some cases the preferred action/solution can now be selected without calculating the prediction of the trained algorithms. In case the delay minutes are between 20 and 120 minutes the prediction of the first algorithm is calculated. As a result a probability is given, which indicates by which percentage the class 0 (no action needed – delay flight) is the preferred one. If the calculated probability is higher than 50% the preferred action/solution should be to delay the flight. In the case the first algorithms predicts, that an individual action/solution is needed, the second algorithm is run to calculate probabilities for the 3 classes of individual actions/solutions. As a result, probabilities for each action/solution are calculated and can be used to decide which should be the preferred one.

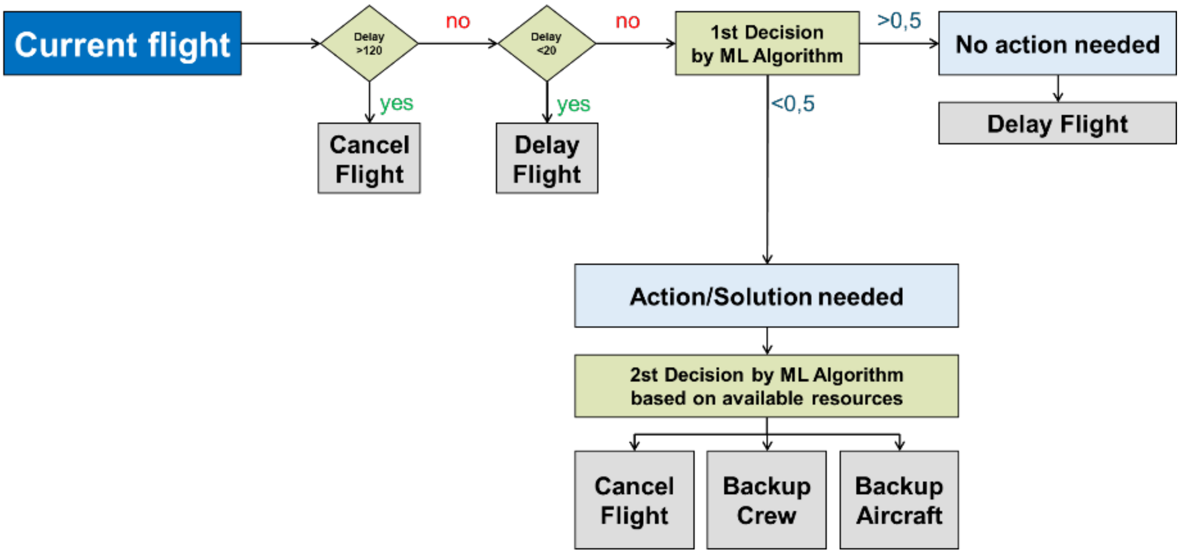


*Figure 17 Updated four-step classification for action/solution prediction*

With the finalized classification process and the trained algorithms, the action/solution selection GUI from Figure 8 was updated. The following Figure 18 and Figure 19 show the updated GUI including a section where a proposed action/solution is given, as well as the calculated probabilities for each action/solution. With these added information it is now possible to get a summarized overview of the disruption, test different actions/solutions and see what impact they would have on cost and delay times and to get a decision support based on the trained algorithm. All information is available in one GUI.

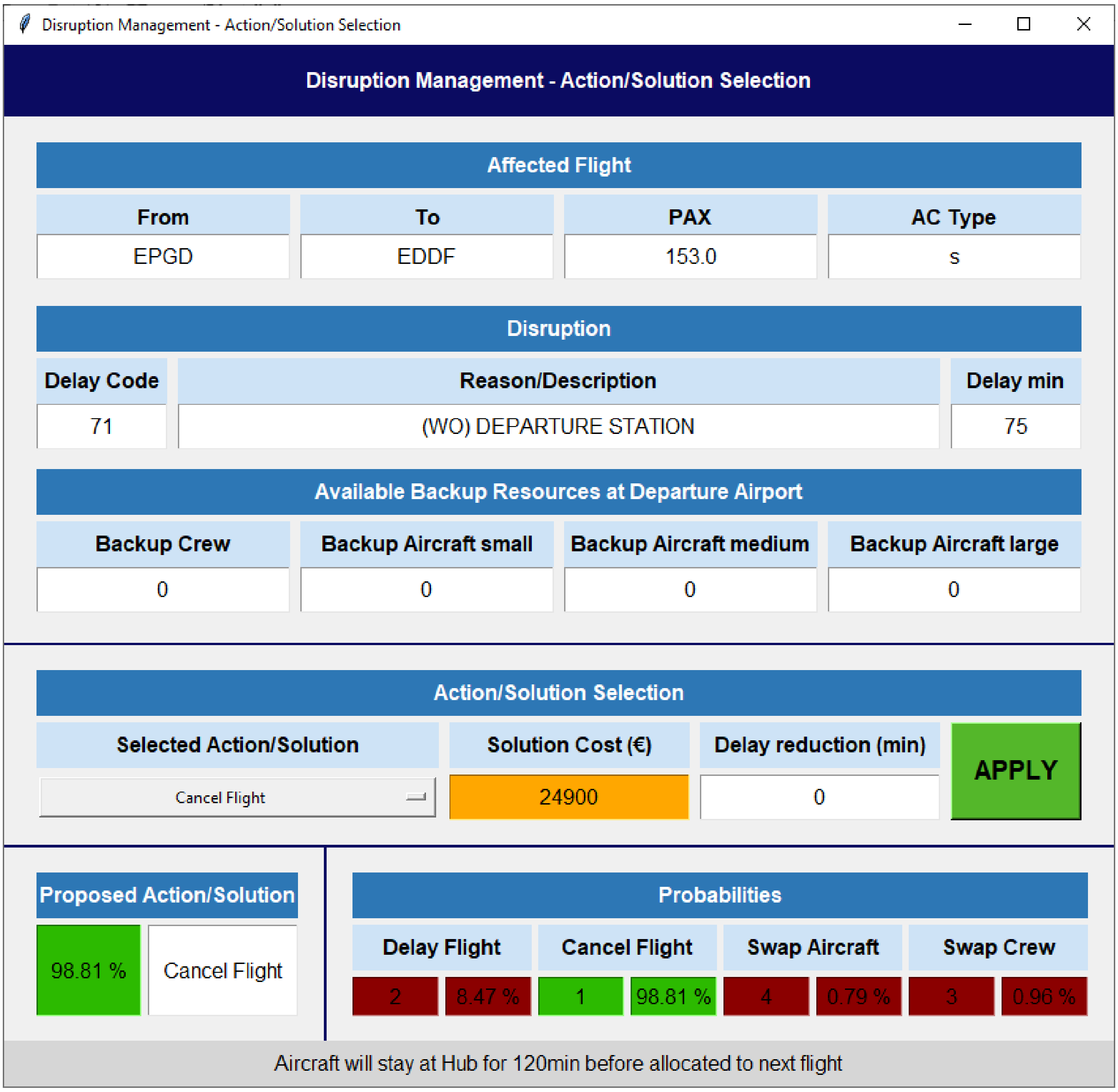


*Figure 18 Disruption Management GUI Prediction Example 1*

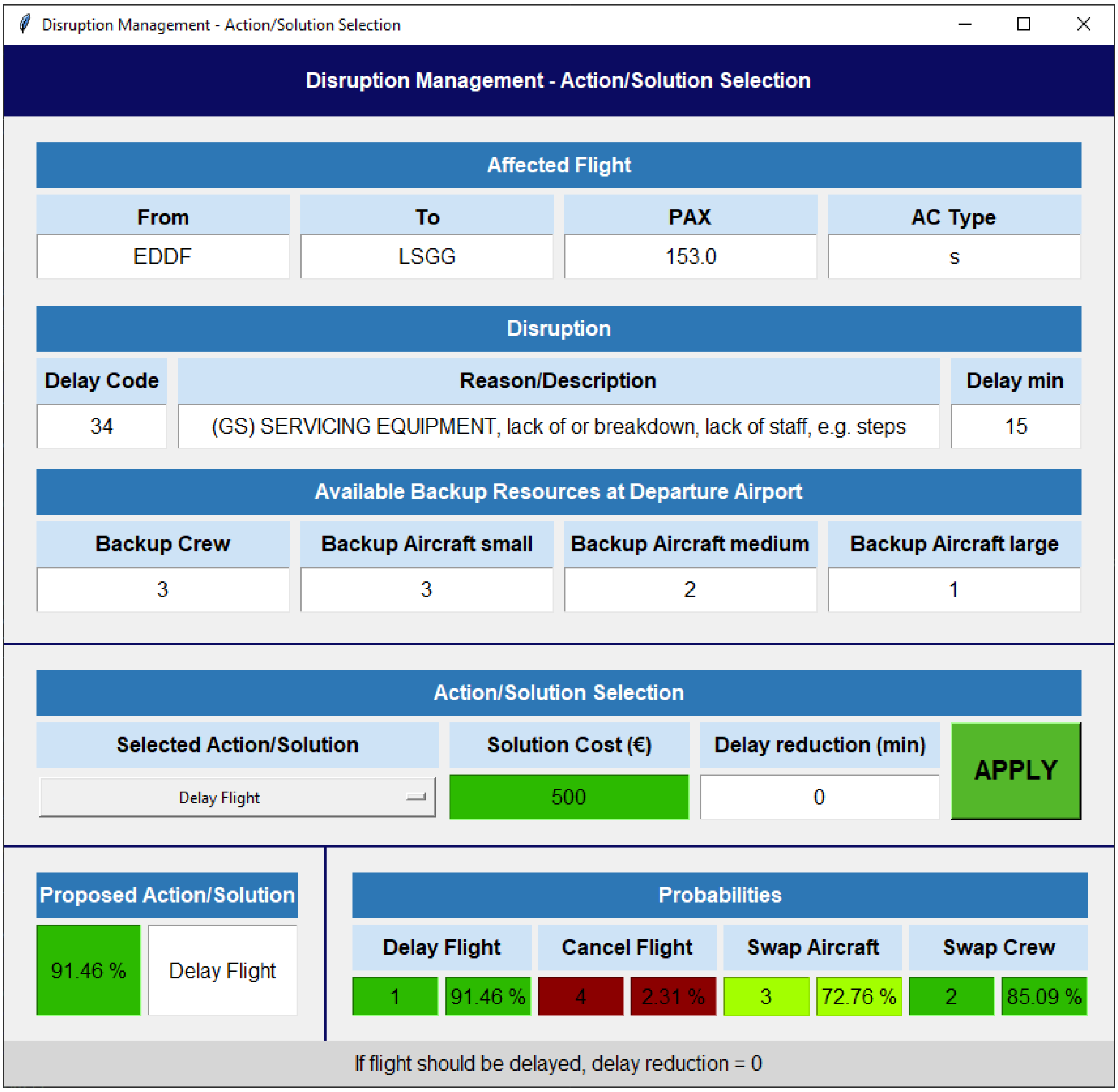


*Figure 19 Disruption Management GUI Prediction Example 2*

A first validation of the developed DST was carried out. KPIs over several days from randomly selected actions/solutions and algorithm based action/solution predictions were compared against each other. Table 9 gives an overview of averaged KPIs over several days of both approaches.

*Table 9 First validation of algorithm based decision*

| KPI | Random | Algorithm & Rules | Change |
|---|---|---|---|
| KPI_daily_flight_count | 263.78 | 269.64 | 2% |
| KPI_daily_pax_count | 51197.57 | 52315.45 | 2% |
| KPI_daily_flights_min_15m_delayed | 38.86 | 62.55 | 61% |
| KPI_daily_delay_min | 1337.61 | 1260.91 | -6% |
| KPI_daily_dis_sol_cost | 659892.97 | 255621.27 | -61% |
| KPI_delay_per_flight | 5.07 | 4.69 | -8% |
| KPI_cost_per_flight | 2509.40 | 947.35 | -62% |

The results show a decrease in overall cost by 61% compared to the base model (random action/solution selection). Also a decrease in average delay minute per flight of 8%. An increase by 61% can be seen in the overall amount of flights with at least 15 minutes delay. This can be interpreted as a more robust disruption management process, since the trained algorithms allow more flight to be delayed but still achieve to reduce the overall cost.

These are some promising first results. The next step of the validation will be a comparison of algorithm based decision against decision taken by people with experience in the field of airline operation, to find out how close the developed DST can come to the performance of human based decision.

# 9. Analysis of the results

Since DiSpAtCH is still in progress, the current results do not yet include the final validation of the DST.

The research began with a study of the state of the art by visiting airlines and their OCC. People in charge of the current disruption management were interviewed and their requirements for a novel DST were defined. Based on these requirements and the desire to use ML algorithms in the context of disruption management the framework of DiSpAtCH was developed.

With the developed framework data needs could be identified and since the needed data was not available to researchers, an airline simulation was developed. The airline simulation does now ensure that all data needed to train the desired ML algorithms are available and give a platform for validation of the final DST.

A dataset was generated by the developed airline simulation and then pre-processed and used to select an algorithm. The neural networks for the classification process were trained and a final four-step classification for action/solution prediction was developed.

The results of the trained algorithms were included in the Disruption Management GUI and do now provide sufficient information to be used as DST.

A first validation was carried out and showed promising result. The next step of the validation campaign will include people with experience in the field of airline operation, to find out how close the developed DST can come to the performance of human based decision.

# 10. Conclusions and look ahead

The following figure shows some main achievement of this project.

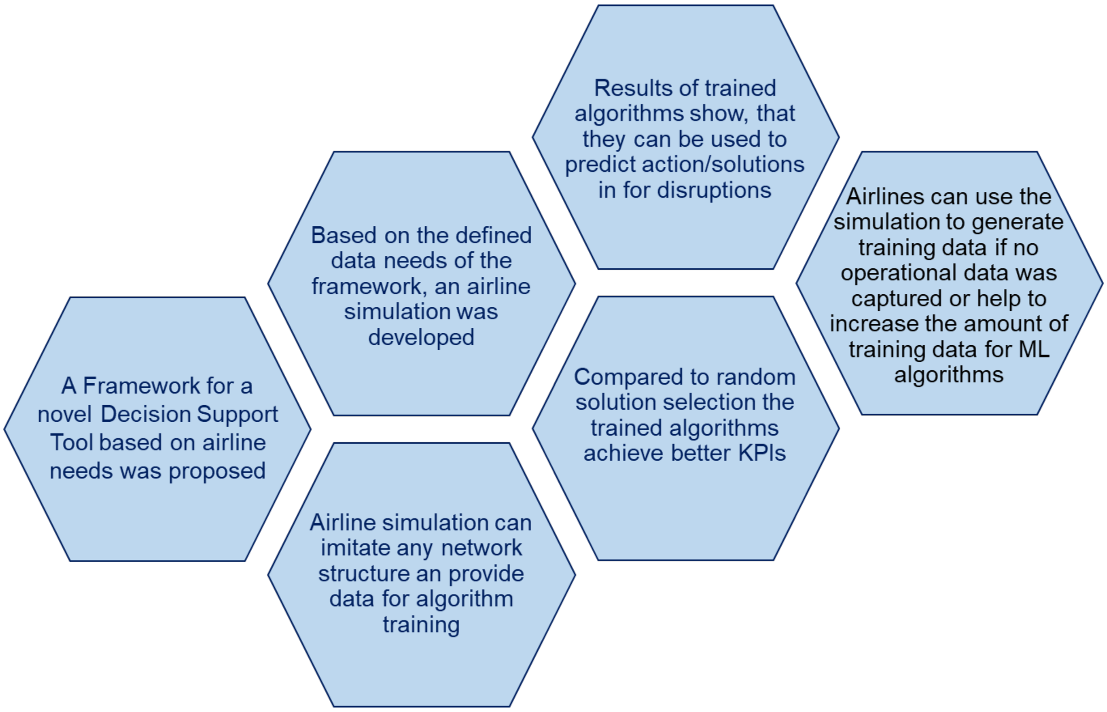


*Figure 20 Overview of some main achievements*

A first validation was completed, but for the final validation is still carried out. It is planned to compare three decision making processes against each other for disrupted operations:

- Airline simulation where actions/solutions are selected randomly
- Airline simulation where actions/solutions are selected based on the trained ML module
- Airline simulation where actions/solutions are selected by staff working in airline OCCs (experienced in disruption management)

By comparing the KPIs for each of these three options over a few simulated days it should be possible to see whether the ML module can keep up with the decisions of the experienced OCC staff as it already achieves better performance than the random selection of actions/solutions.

By setting up the airline simulation to imitate the operations of a specific airline it might be possible to generate training data which will lead to trained algorithms which can be directly used to support their real operations. This would be a great help for airlines, since they would not need to record operational data over months before they are able to use ML modules for decision support during daily operations.

If an airline would in addition record the proposed data during their daily operation the ML modules could be adjusted to their overall company goal in disruption management, e.g. reduction of cost or delay minutes. Also, a combination by prioritizing the KPIs could be implemented, e.g. 80% cost reduction and 20% delay reduction. In this example an airline would prioritize cost reduction over delay reduction and the ML algorithms can be trained taking into account the prioritization of the individual KPIs. This would allow us to train different algorithms with different prioritization and then selecting the algorithm with the desired overall goal for each flight, day, or season individually.

Every time a flight plan or network structure changes drastically the old trained algorithms might not perform as well as before. Therefore, a continuous training of the algorithms is needed to ensure that they can be used during daily operations. If a change in the flight plan or network structure has occurred, it may be necessary to start again with 100% synthetic data and then reduce the amount of synthetic data over time as more and more real operational data are available representing the new flight plan or network structure. The goal will therefore be to continuously use a mixture of synthetic and real operational data as training data and also to regularly train the algorithms. The best case would be if enough real operational data are available but since this is often not the case, DiSpAtCH provides a suitable solution to close the gap of missing data to start training ML algorithms from an airline perspective during disruption management in daily operations.

Since the developed airline simulation is highly adaptable to other needs its generated data might be helpful in other research areas in the context of airline operations where it is also difficult to get a sufficient amount of real data.

# 11. References

## 11.1 Link to PhD thesis / repository

When finished, the final results of DiSpAtCH will be uploaded to the University's Cloud.
https://cloud.tu-braunschweig.de/s/MmyQaWwt6ZzAN98

Already available is the paper from the doctoral symposium at ICRAT 2020, the presentation from the Engage KTN summer school in 2021, the poster from the SESAR Innovation Days 2021 as well as this report.

## 11.2 Associated outputs and publications

**ICRAT 2020, Paper at the Doctoral Symposium**

Paper:
https://drive.google.com/file/d/1uW-tUmcMYUW7zf9OANH-TJQkBMJGoDFe/view
Presentation:
https://www.youtube.com/watch?v=UDWjo9-uMnY&ab_channel=FAAWilliamJ.HughesTechnicalCenter

**Several internal technical presentation updates for Jeppesen / Boeing Global Services**
https://cloud.tu-braunschweig.de/s/MmyQaWwt6ZzAN98

**Poster at SESAR Innovation Days 2021**

**Journal Paper in progress**

**DLRK 2022 Presentation is planned**
**https://dlrk2022.dglr.de/**

## 11.3 References cited in this report


[1] P.J. Bruce, Y. Gao, and J.M.C. King, Airline Operations. A Practical Guide, Milton: Taylor and Francis, 2017.

[2] C. Barnhart, and V. Vaze, Irregular Operations: Schedule Recovery and Robustness in The global airline industry, P. Belobaba, A. Odoni and C. Barnhart (Eds.), Chichester: Wiley-Blackwell, 2016, pp. 263–286.

[3] FAA, FAA Aerospace Forecast – Fiscal Years 2019 – 2039, https://www.faa.gov/data_research/aviation/aerospace_forecasts/media/FY2019-39_FAA_Aerospace_Forecast.pdf (Accessed: 19.05.2022).

[4] EUROCONTROL, EUROPEAN AVIATION IN 2040 - CHALLENGES OF GROWTH - Annex1 Flight Forecast to 2040, https://www.eurocontrol.int/sites/default/files/2019-07/challenges-of-growth-2018-annex1_0.pdf (Accessed: 20.01.2020).

[5] EASA, European Aviation Environmental Report 2019, https://www.easa.europa.eu/eaer/system/files/usr_uploaded/219473_EASA_EAER_2019_WEB_HI-RES_190311.pdf (Accessed: 19.05.2022).

[6] IATA, IATA Forecast Predicts 8.2 billion Air Travelers in 2037, https://www.iata.org/pressroom/pr/Pages/2018-10-24-02.aspx (Accessed: 19.05.2022).

[7] J. Clausen, A. Larsen, J. Larsen, and N.J. Rezanova, Disruption management in the airline industry - Concepts, models and methods, Computers and Operations Research vol. 37, Lyngby: Elsevier, 2010, pp. 809–821.

[8] A. Castro, A.P. Rocha, and E. Oliveira, A New Approach for Disruption Management in Airline Operations Control, London: Springer, 2014.

[9] J. Rapajic, Beyond Airline Disruptions – Thinking and Managing Anew, 2nd ed., Abingdon: Routledge, 2019.

[10] C. Rebala, A. Ravi, and S. Churiwala, An Introduction to Machine Learning, Cham: Springer, 2019.

[11] E.C. Fernández, J.M. Cordero, G. Vouros, N. Pelekis, T. Kravaris, H. Georgiou, G. Fuchs, N. Andrienko, G. Andrienko, E. Casado, D. Scarlatti, and P. Costas, DART: A Machine-Learning Approach to Trajectory Prediction and Demand-Capacity Balancing, Seventh SESAR Innovation Days, 2017.

[12] NASA, C.S. Bosson, and T. Nikoleris, Supervised Learning Applied to Air Traffic Trajectory Classification, https://ntrs.nasa.gov/search.jsp?R=20180000801 (Accessed: 19.05.2022).

[13] aeroTELEGRAPH, T. Nowack, Künstliche Intelligenz hilft im Kampf gegen Verspätungen, https://www.aerotelegraph.com/lufthansa-swiss-austrian-kuenstliche-intelligenz-hilft-im-kampf-gegen-verspaetungen (Accessed: 19.05.2022).

[14] L.K. Hassan, B.F. Santos and J. Vink, "Airline disruption management: A literature review and practical Challenges", Computers and Operations Research, vol. 127, 2021.

[15] H. Dong, J. Zhang and X. Zhao, "Intelligent wind farm control via deep reinforcement learning and high-fidelity simulations", Applied Energy, vol. 292, 2021.

[16] K. Ogunsina, I. Bilionis and D. DeLaurentis, "Exploratory data analysis for airline disruption management", Machine Learning with Applications, vol. 6, 2021.

[17] J. Sandgard and J.E.B. Langner, "Interviews with airline operation control center", own communications, 2019.

[18] J.E.B. Langner, "Decision Support System for Airline Operation Control Hub Centre (DiSpAtCH) - Initial research results and developed framework", ICRAT 2020, 2020.

[19] K.M.C. Zielinski, L.V. Hendges, J.B. Florindo, Y.K. Lopes, R. Ribeiro, M. Teixeira and D. Casanocva, "Flexible control of Discrete Event Systems using environment simulation and Reinforcement Learning", Applied Soft Computing, vol. 111, 2021.

[20] OpenFlights.org, J. Patokallio, "Airport database", https://openflights.org/data.html#airport, (Accessed: 19.05.2022).

[21] EUROCONTROL, "EUROCONTROL Standard Inputs for Economic Analyses", Edition 9.0, December 2020, https://www.eurocontrol.int/sites/default/files/2021-03/eurocontrol-standard-inputs-economic-analysis-ed-9.pdf, (Accessed: 19.05.2022).

[22] A.J. Cook and G. Tanner, "European airline delay cost reference values", EUROCONTROL Performance Review Unit, Version 4.1, https://www.eurocontrol.int/sites/default/files/publication/files/european-airline-delay-cost-reference-values-final-report-4-1.pdf, 2015, (Accessed: 19.05.2022).

[23] Anonym Airline, "Disruption Database with recorded disruptions from 2017-2019", own communications, 2019.

[24] EUROCONTROL, EUROCONTROL R&D Data Archive from 2015 to 2018.

# Annex I: Acronyms

| **Term** | **Definition** |
|---|---|
| AI | Artificial Intelligence |
| DST | Decision Support Tool(s) |
| KPI | Key Performance Indicator |
| ML | Machine Learning |
| OCC | Operation Control Centre |