A Multi-layer Model for Long-term KPI Alignment Forecasts

Gérald Gurtner, Luis Delgado,
Andrew Cook
School of Architecture and Cities
University of Westminster
London, United Kingdom

Jorge Martín, Samuel Cristóbal
The Innaxis Foundation
and Research Institute
Madrid, Spain

***Abstract*—This article presents a new holistic model for the air traffic management system built by the Vista project. The model studies the alignment and trade-offs of key performance indicators in the 2035 and 2050 horizons. It is based on three layers modelling the strategic, pre-tactical and tactical phases of ATM. It heavily features multi-agents, is highly data-driven, and includes highly microscopic models. It is runnable as a 'what-if' tool and has been applied to different scenarios, including long-term forecasts for 2035 and 2050. The results obtained with the model so far show clear trends, including surging emissions, an important reduction in delay uncertainty, and increases of flight plan buffers.**

I. Introduction and objectives

The air transportation system is a complex system that is increasingly performance driven. Following the work of ICAO [1], SESAR has adopted and developed high-level KPIs (key performance indicators) in order to monitor its evolution and drive it in the desired direction.

One of the issues encountered is the lack of ability to forecast trade-offs or alignments between KPIs. Whilst it is clear that these KPIs are not independent, it is less clear to what extent. A difficulty often facing KPI definition is finding sufficiently simple metrics that are readily understandable.

The reason why KPIs are so intertwined is that air transportation is a complex socio-economic system with multiple levels of dynamical processes, led by multiple actors with idiosyncratic rules and behaviours. As a consequence, a change in the system can have far-reaching consequences. Several changes at the same time can be interrelated and produce unforeseen impacts. This is a typical instance of an emergent system, as noted multiple times in recent years [2], [3].

This implies that partial models might not capture overall system performance, partially because the sum of their behaviour is not the behaviour of their sum. As a consequence, holistic models are required, which usually trade a high level of detail against a better integration of several mechanisms. The Vista project has built such a model to address this issue. The model simulates typical days of operation, representing different futures, at the current (2014), 2035 and 2050 horizons.

The first period of Vista was dedicated to defining the scope of the model. In particular, the forces behind the potential changes in the system were identified (and called 'factors' in the following), using an extensive literature review. Scenarios were also defined to prioritise the simulations and obtain knowledge on the most important factors. This has been described in [4] and is only briefly recalled here. In this article, we present an overview of the full model, the methods and data used for calibration, and some examples of the results obtained.

The paper is organised as follows. Section II explains the scope of the model, the main challenges it aims to address, and the scenarios chosen to be run. Section III describes the different models, while Section IV explains how they have been calibrated. Section V presents various results obtained with the model from the different scenarios. Finally, we draw some conclusions in Section VI.

II. Operational scope

The Vista project is looking far in advance by trying to predict how the envisioned changes for 2035 and 2050 will likely interact. Given the extreme depth of the time horizon, Vista's aim is not to forecast exactly what will happen. Instead, Vista built a model that simulates a typical day of operations given a set of external forces (i.e., parameters that affect the system) such as macroeconomic, technological or regulatory parameters. By analysing the different indicators under different factors, Vista is able to identify trade-offs and trends across different timeframes and across different stakeholders for a given time horizon.

Five stakeholders are modelled in Vista: ANSPs, airports, airlines, passengers, and the environment. Vista considers KPIs for each of them in order to analyse the trade-offs when different factors are considered.

ANSPs are heavily regulated and have traditionally provided the full scope of air navigation services. The evolution to more performance-driven operations leads to a tendency towards the unbundling of services and technological innovation. In addition, almost all ANSPs have become engaged in one or

This project has received funding from the SESAR Joint Undertaking under grant agreement No 699390 under European Unions Horizon 2020 research and innovation programme. The opinions expressed herein reflect the authors views only. Under no circumstances shall the SESAR Joint Undertaking be responsible for any use that may be made of the information contained herein. The authors thanks DLR for their support with the NO_x emission model.

more strategic alliances and industrial partnerships [5]. ANSPs are modelled in Vista as non-profit driven and, therefore, the capacity provided is based on the expected delay. The operating costs, and hence unit rates, are an outcome of these capacities and demands.

Large airports' current business models rely heavily on non-aeronautical revenues (parking, shopping, etc.) [6]. Congestion is a major issue for most, and different strategies are implemented to increase their capacity, such as soft management procedures or heavy changes in infrastructure [7], or improvements from airport expansion programmes and technological enhancements. For small airports, aeronautical revenues are more significant and low-cost, point-to-point operations are more relevant to their income strategies. The relationship between airport operations and airlines business models results in the fact that the evolution of airports relies heavily on the business models of airlines and their future traffic. A spectrum of private and public ownership exists, but nearly all are heavily regulated, in particular regarding aeronautical charges [8]. Vista is able to capture both economic (e.g., revenues) and performance (e.g., delay) KPIs for airports.

Competition among airlines is an important factor when considering future traffic evolution. Airlines are highly market-driven as it is relatively easy to reassign aircraft to more suitable routes according to their business needs adjusting to changes in demand. Low-cost carriers (LCCs) generally have lower yields compared with 'legacy' operators. LCC expansion has mainly been based on point-to-point (P2P) strategies, aiming at higher utilisation by using a homogeneous fleet, and lower costs by using secondary airports [9]. However, more recently, some LCCs have shifted to the legacy model to some extent by opearting from primary airports or feeding long-haul flights. Legacy carriers have been forced to adapt their model by lowering their costs, sometimes trying to gain market share with an 'in-house LCC', and by unbundling services provided following an LCC approach to pricing. Vista can capture different indicators for this stakeholder at different ATM planning phases.

For the passenger, price, travel time, comfort and convenience constitute some of the factors influencing their choice [10]. The literature often defines archetypal profiles for passengers, usually taking into account socio-economics and travel purpose (often simply 'business' or 'leisure'). Some profiles have been defined by the project DATASET2050 [11] and have been loosely adopted in the Vista model. This more detailed passenger profiling allows us to model the door-to-gate and gate-to-door phases providing door-to-door metrics, in addition to more classical gate-to-gate indicators.

The final 'stakeholder' modelled in Vista is the environment. This is a passive stakeholder, regarding which interest lies on the reduction of the impact of aviation on the environment. In particular, metrics related to the emissions of (CO_2 and NO_x). At this stage, noise is not included in Vista.

The selection of scenarios run is based on a consultation, a dedicated, expert workshop and the current model's capabilities. The way scenarios are used in Vista has been explained in [4]. Here, only the main features are highlighted.

TABLE I
EXAMPLES OF SCENARIOS SIMULATED IN VISTA.

Scenario	Short description
Current	'Current' situation (SEP 2014)
L35 baseline	Baseline environment in 2035 (slow economic growth and slow technological advancements)
H35 baseline	Baseline environment in 2035 (high economic growth and high technological advancements)
Non-supportive 2035	Using L35 baseline plus a poor emphasis on environmental and passenger protection and very a high price for fuel
Supportive 2035	Using L35 baseline plus a poor emphasis on environmental and passenger protection and very a high price for fuel
L50, H50, Non-supportive 2050, Supportive 2050	As per above, for 2050

Scenarios are built by fixing exogenous variables to the model. In Vista, these variables are called business or regulatory 'factors', and include macro-economic considerations such as GDP, but also SESAR solutions such as free routing. These factors are gathered into sets and fixed at certain values for a given scenario.

Nine scenarios, which are presented in Table I, have been modelled in Vista. The first scenario represents the 'current' situation (2014), used for calibration and comparison. We then used two main baselines: the 'Low' baseline is built around slow economic growth and slow technological improvement. The 'High' baseline, in contrast, includes high economic growth and technological improvement.

In addition to the baseline scenarios, four factors have been selected in order to assess their impact on the system. These factors are grouped into 'supportive' and 'non-supportive' cases. Loosely speaking, the former depicts a world where passenger rights are protected, environmental issues are tackled and airlines benefit from a favourable economic situation (low price of fuel). The latter describes the opposite situation. In summary:

- price of fuel: low in supportive; high in non-supportive;
- implementation of passenger reaccomodation tools and provision schemes: 'on' in supportive; 'off' in non-supportive;
- implementation of 4D trajectories ECAC-wide: current operations in non-supportive; fully implemented in supportive;
- price of emissions allowances: high in supportive; low in non-supportive;

Note further that NO_X emissions are taxed (based on their equivalent radiative impact compared to CO_2) in the supportive case, and are not taxed in the non-supportive case.

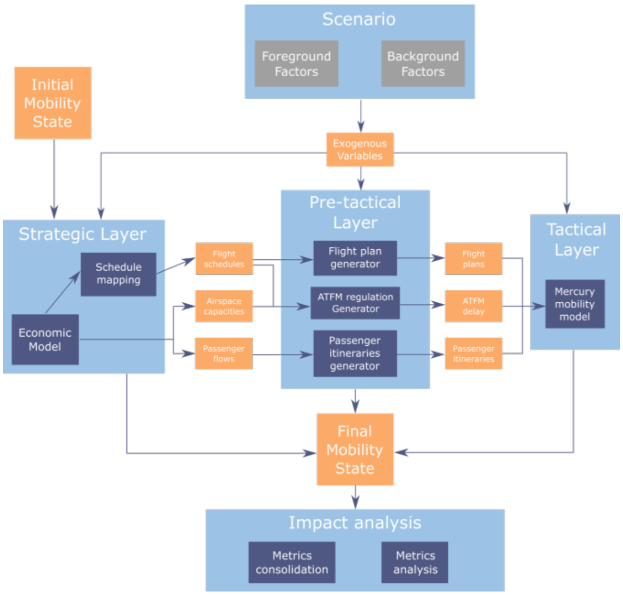

Figure 1. Vista model

III. Description of the Model

Figure 1 presents an overview of the Vista model. Vista models the three temporal phases of ATM (strategic, pre-tactical and tactical) for each scenario investigated, with the objective of generating a representative (busy) day of operations for each given scenario. The various factors define the scenario to be modelled. The strategic layer considers the factors and the economic environment to provide the outcome of strategic decisions made by the stakeholders, the capacities provided, demand, flight schedules and passenger flows for a typical day of operations. These flows, schedules and capacities are transformed into individual flight plans, passenger itineraries and ATFM regulations by the pre-tactical layer. Finally, the tactical layer executes the flights and passenger itineraries at a flight and passenger level, tracking the evolution of delay, passenger connections and the tactical decisions carried out by the actors. Among these layers, only the tactical one is based on a prior model, 'Mercury'. Mercury has been re-implemented and enhanced for Vista, but is based on previous models, used during the last ten years across several research projects such as POEM and ComplexityCosts [12].

One characteristic of Vista is that each layer may be executed integrated into the full model, or as a stand-alone model, which allows us to test the different phases independently, by providing the appropriate inputs.

A. Strategic

1) Economic model: The economic model is the first block of the strategic layer. It has the task of creating appropriate levels of supply and demand in Europe based on different scenarios. The model provides a high-level view of the number of flights and passengers for each origin-destination pair.

The model is based on a common environment, in which different agents evolve. The types of agent implemented in the model are summarised below.

- Airline: one per actual airline. Used to compute and handle flights and decide if new OD pairs should be opened.
- Flight: one per OD pair (without connection) per airline. Notionally represents all the flights operated by an airline between the OD pair. Computes the marginal profit of the leg and chooses the supply (number of seats) for the next simulation turn. Also chooses between the possible flight plans based on their fuel and ATC charges costs.
- ANSP: one per actual ANSP. These set their capacities (number of controllers) based on a target delay. They then set the unit rates to have zero profit based on the expected demand.
- Passenger: one per initial origin and final destination (with connections). Notionally represents all the passengers going from 'O' to 'D' by any legitimate itinerary (see below).

Note that there are no hard-coded archetypes of agents within each type. What defines the different behaviours of the agents is their cost structure and their initial conditions (the initial network for airlines, for instance). As noted above, a (dynamic) network underlies the structure of the model, where the airports are nodes and flights are edges. In this network, passengers use collection of edges from their origin to their final destination.

Supply and demand interact in this network in an intricate way. On the one hand, the supply is leg-based. To compare the supply and demand, the latter is aggregated based on all the passengers going through this leg for this airline, for whatever itinerary (an "itinerary" is a distinct permutation of legs between a given OD pair, operated by a given alliance). The supply for the leg for this airline is given by the corresponding flight agent. A price variable attached to the leg plays an adjusting role, as we will see in the following. On the other hand, the demand reacts only to price of itineraries. These prices are thus aggregated from each leg in the corresponding itinerary.

In each turn, the agents perform a number of actions depending on their nature. See [13] for more details. In brief:

- based on the scenario, the parameters of the agents are changed;
- ANSPs predict the traffic for the round; they compute the capacity needed to be under a target delay per flight and set this value for their capacity, limited by their maximum capacity; they set their unit rate so that their profit is null;
- flights predict delay at airports, en-route (ATFM) delay, and the price of their leg for this turn; they set their supply based on these predictions and their own cost function;
- passengers take the price of the last round for each leg and sum them to form the price for the itineraries; they weight each itinerary based on their utility function (see below) and set their demand level for each of them based

on this weighting;

- demand is aggregated for each leg;
- supply and demand are compared for each leg - prices evolve on each leg based on the discrepancy;
- all agents record variable values for this round, i.e., prices, delays, etc;
- a new turn is initiated.

2) Schedule mapper: The schedule mapper is the second block in the strategic layer. It converts the high-level flows of the economic model into individual schedules, to be used by the flight plan generator. Planning schedules based on expected demand is a highly demanding task, even at the single airline level. Airlines usually have dedicated tools for this. The complexity of assigning schedules is due to the high number of possibilities and the multiple constraints. These constraints include hard constraints such as crew, aircraft, and airport slots, plus soft constraints such as the cost of operating an OD pair. It is out of the scope of Vista to reassign completely from scratch all the schedules created by the economic model. In particular, this would imply capturing very complicated slot management behaviours from the airlines (including 'irrational' behaviours such as endowment effects). Moreover, the number of flights considered is too large to ensure a sufficiently fast, stable solution from a computational perspective. Vista simplified the problem by relying on an initial state, set to be 12SEP14 (see later).

The schedule mapper compares, for each airline, the new flow to the old one for each OD pair. If the new flows are high enough compared to the old ones, each airline tries to add a new aircraft and optimise its route so that most of the new flow is covered. If the new flows are small, it removes one of its aircraft. The airline takes crew and airport slots into account only indirectly through the possible patterns (routes) available to the new aircraft, and the corresponding turnaround times (see later, under 'Calibration'.).

More specifically, the mapper goes through the following steps:

- load data on airports, historical schedules, pattern data (see Section IV-B), and strategic flows;
- compute average travelling times between every OD pair;
- compute likely departure times;
- load the decision tree for the turnaround times;
- for each airline:
  - trim its network by removing aircraft which are in excess;
  - grow the network by adding aircraft to meet demand;
- compute the new schedules and add them to the database.

B. Pre-tactical

1) Flight plan generator: The flight plan generator transforms OD schedules into possible flight plans. For each flight plan, an estimated operating cost is computed including fuel, en-route airspace charges and emissions costs. These operating costs are taken into consideration when prioritising the different flight plan options for each schedule. This block is based on historical (possible) trajectories and on aircraft performances.

The flight plan generator is divided into different sub-blocks; these are supported by historical data analysis on flight trajectories and aircraft performance, as summarised below.

- Route generator: for each scheduled OD pair, a set of possible routes is computed. These routes are based on the clustering of historical routes. If the OD pair does not exist in the historical dataset, new routes are estimated.
- Trajectory generator: for each route of each schedule, a possible trajectory is computed including, *inter alia*, climb, cruise and descent distances, selected altitudes, air speeds, estimated fuel usage and average cruise wind. A trajectory optimiser could have been used to generate the different flight plans. However, with that approach idealised optimal trajectories would have been obtained. Instead, in Vista, the trajectory generator relies on historical data on flight plans and aircraft performance (BADA), thus obtaining realistic trajectories.
- Flight plan generator: computes, for each trajectory, its estimated direct operating costs: en-route airspace charges, cost of fuel and emissions. The en-route airspace charges are computed modelling the 39 regions managed by EUROCONTROL CRCO plus the airspaces of Egypt, Belarus, Morocco, Uzbekistan and Ukraine. Other surrounding countries which might follow different charging schemes are also modelled: Algeria, Iceland, Russia, Tunisia. This ensures that the cost of all trajectories is estimated accurately.
- Flight plan selector: the trajectory generator produces as many trajectories as exist for the OD pairs. However, the tactical layer requires a flight plan per schedule: this is also required by the ATFM regulation generator to estimate the demand on each ANSP and, from this, the probability of having an ATFM regulation. Airspace users often change their flight plan prior to departure as a function of the tactical situation. These tactical changes to the flight plans cannot be decided at the pre-tactical layer. For this reason, the flight plan selector prioritises all the flight plan options of the flights based on their operating cost. A logit rule, which considers the cost of fuel, of airspace charges and of emissions is used to estimate the probability of selecting a specific flight plan, in a similar manner to the airline computing the cost of the flight plan in the strategic layer.

All the blocks of the flight plan generator might be affected by the various factors of the environment. For example, the length of the routes used by the route generator will be affected by the introduction of free routes.

2) ATFM regulation generator: The ATFM regulation generator estimates the probability of being affected by ATFM regulations and the corresponding amount of delay. As presented in Figure 1, the input of the ATFM regulation generator is the capacity of the ANSPs (for the scenario to be processed, and for the baseline 2014 case) and the traffic (demand of the scenario to be processed, and of the baseline 2014 case).

The ATFM generator thus requires the outcome of the flight plan generator (and flight plan selector) and the 2014 demand and capacity, in order to be able to compute the variation of demand and capacity with respect to the 2014 baseline case, which is used as a reference.

ATFM regulations are divided between regulations due to capacity issues (i.e., regulations flagged as “C”, being implemented by an FMP) and all the other regulations. It is assumed that regulations due to capacity are affected by the demand on the ANSP, while the regulations that are not due to capacity remain homogeneous across the ECAC region. This assumption allows us to modify the probability of having ATFM delay due to the demand expected at different ANSPs and their expected capacity, while maintaining delay, which is not directly linked with capacity/demand imbalances (e.g., regulations due to weather). From historical data (analysis of AIRAC 1313-1413), regulations due to capacity represent 43.3% of all the regulations issued, followed by weather as the main causes. Weather, by its nature, might occur at wider locations across the ECAC and might not be related directly to demand.

3) Passenger itinerary generator: The strategic layer generates the passenger flows and flight schedules (including their aircraft type and number of seats available). The objective of the pre-tactical layer is to transform the flows of passengers into individual itineraries, i.e., to assign the passengers to specific flights. This is done in a three-stage process:

- computing the possible options available for the passengers in each flow;
- optimising the assignment of passengers among their options considering aircraft capacities and minimum connecting times at airports;
- creating additional passengers’ itineraries to ensure that the load factors of the aircraft are realistic;
- assigning which passengers are ‘premium’ and which are ‘standard’ (which features in passenger reaccommodation rules in the tactical layer).

C. Tactical layer

The tactical layer models delay propagation between flights and the adaptability of the system during disruption (cancellations, background and foreground delay) and with limited resources (e.g., airports and en-route capacity). The modelling is performed both per flight and per passenger to allow us to capture both flight and passenger indicators. The model used for the tactical layer is the Mercury event-driven simulator, as introduced earlier. One of the characteristics of this tactical model is its door-to-door capabilities. Passenger types are disaggregated into more detailed profiles (e.g., ‘environmental traveller’, ‘culture seeker’) and mapped to different transport choices regarding access/egress to/from the OD airports.

There are various feedback loops considered in this layer. One example is the expected arrival time of flights, which is updated to the airline several times during the simulation. The airline operator can use this information to adjust the behaviour of outgoing connecting flights, for example, by waiting for connecting passengers. Several airline costs are considered during the simulation: fuel, emissions, CRCO, crew, maintenance and passenger delay costs. Those have an impact on the behaviour of the system such as when selecting the flight plan.

The tactical layer simulation is a sequence of two processes: the gate-to-gate simulation and the door-to-door simulation, which are executed for each flight. This requires information from previous layers, such as the flight schedules, flight plan options, airspace data and passenger itineraries, but also some other inputs, such as connecting times for passengers and taxi times for flights at different airports.

IV. Calibration

The full calibration of the model required substantial data acquisition and analysis. In the following, we highlight the main sources and the main steps carried out. For current operations, the model is calibrated based on a day of traffic (12SEP14) selected due to being a busy, nominal day (e.g., not disrupted by industrial actions or severe weather events).

A. Input data

One of the main characteristics of the Vista project is the of use different sources of data to inform the model. The main data source is EUROCONTROL’s DDR data [14]. The data were used extensively, in particular to:

- set the initial state of the economic model;
- extract the distribution of delays for airports and ANSPs, which helped to:
  - infer delay-traffic relationships (for airports);
  - perform a mean-variance analysis on airport delay;
  - perform an analysis of ATFM regulations;
- compute the length of trajectories in each ANSP’s area;
- cluster possible routes between origin and destination airports;
- model flight plan preferences (flight level and speed requests);
- model flight trajectories (characteristics of climb and descent phases);
- estimate average wind distributions between regions by comparing ground speed with requested air speed.

Other flight-related data have been used, in particular the BADA 4.2 model in order to estimate fuel consumption, both in the planning (pre-tactical) and executive (tactical) phase.

Since DDR data is flight-centric, we used other sources of data to compute passengers’ information. We used a mixed database from global distribution system and IATA (‘PaxIS’) data to obtain detailed information on passengers’ itineraries for 12SEP14, including the fare price and the class of the passenger. Passenger itineraries are also based on previous computations from the ComplexityCosts project [12]. Passengers elasticities have been sourced from the literature ([15] for price and income, and [16] for frequency). To estimate the increase of passenger income and thus the increase in demand, we used projections on future GDP from [17]. These projections have been extrapolated for some countries, using decreasing growth over time.

In order to build the available itineraries, Vista also used data relating to airline alliances and partnerships. To simplify the model, we considered that the flights of any airline in a partnership (or alliance) can be used in combination with any other flight from the partnership or airline to create a valid itinerary.

In order to estimate the different operational costs of the airlines, we used the computations of [18]. This includes both strategic (planning, buffer) and tactical (delay) costs, and is used as a standard reference, e.g. by the Performance Review Body to estimate the cost of delay.

Some financial data have also been used for airports. In particular, we used these data to estimate their costs and the landing and departure fees for the airlines.

For the ANSPs, we used only the assumption that they were supposed to be at zero profit and that they have to be under a target value for their delay per flight. Using the initial delays, we are able to compute some efficiency metrics for them (cost per unit of capacity). This efficiency is then changed based on the different scenarios considered.

Other data sources have been required in order to model the distribution parameters for the different flight and passenger phases such as taxi times (CFMU datasets [19]), minimum turnaround times (estimated in previous projects [12]), minimum connecting times [12] and non-ATFM delays [20]. Door-to-gate and gate-to-door modelling relies on distributions calibrated in the DATASET2050 project [21].

B. Calibration analyses

Various analyses were performed to calibrate the model. The most important are highlighted here. Firstly, several linear and non-linear regressions have been performed to examine pairs of variables. For instance, the link between the mean delay at an airport and the level of traffic has been established using linear regression over different time windows and at different airports. The standard deviation of delay (i.e., the unpredictability in the model) has been computed using correlation analysis between the mean delay and its standard deviation. This analysis has produced some interesting results *per se*, for example that mean delay is more correlated with traffic when the traffic is high at the airport.

To choose among a small number of routes, we used a clustering algorithm (kernel density estimation) on each OD pair to reduce the number of potential routes. The algorithm then chooses the flight level, flight speed, and climb and descent times. This is done by classifying the different possible values per type of aircraft and flight plan distance. The result is that individual flights have highly specific trajectories, but still compliant with the typical features observed in the data.

Similarly, the probability of having a regulation based on the time of the day, the ANSPs the flights are crossing, the capacity of the ANSPs, etc, is computed based on historical data. Different adjustments have been tested on the probabilities to adjust them with macroscopic quantities, such as delays.

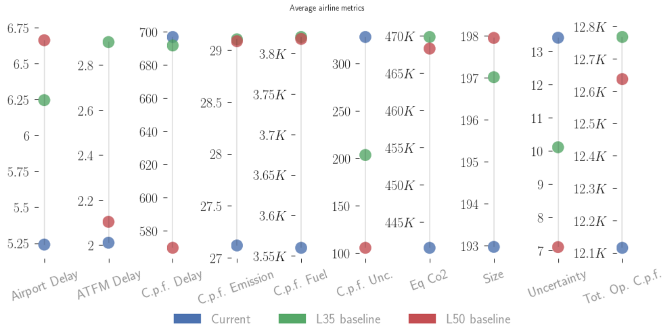

Figure 2. Strategic averages of microscopic metrics for flights

Possible itineraries for passengers are directly computed based on alliance structures and historical itineraries. However, the itinerary generator also needs some calibration regarding the number of connecting passengers and the load factors of the aircraft. This is done by generating new itineraries to ensure that load factors are within target windows.

Calibration was also performed to assure realistic schedules for the additional flights created by the economic model. We first used a pattern analysis to study which kinds of sequences of airports the aircraft were following. We then predicted, for a given pattern, the initial time of departure of the first aircraft and all the subsequent times during the day. This was done by using a classification and regression tree (CART) trained on a month of data, with five predictors.

V. Results

In this section we show only a selection of the results. The deliverables [22] and [13] present further results, in which the impact of some parameters are explored more systematically. In particular, the important factor of the price of fuel is analysed in detail. Very different values of the price have been tested. The results indicated that airlines are greatly impacted by the price of fuel, but also that they transfer part of this extra cost to the passengers, which moderates the increase in traffic. This has a positive impact on the environment and is added to the benefits arising from flights using less fuel-intensive routes. Dedicated papers, with different focus areas regarding the results, are also planned.

A. Strategic

The strategic layer outputs more than 60 microscopic metrics, counting only average values. In Figure 2 we show the value of several flight-centric metrics in different scenarios.

The graphs compare different metrics in three scenarios: the current, the 2035 'Low' baseline and the 2050 'Low' baseline. It is interesting to observe that the total operational cost per flight increases in 2035 but decreases again in 2050. The increase is mainly driven by the cost of fuel, which increases significantly in 2035. This is mainly due to the shift to the use of larger aircraft in the future, as also reflected in 'size' metrics, such as average passengers per flight. This increment in size also implies higher costs of maintenance and

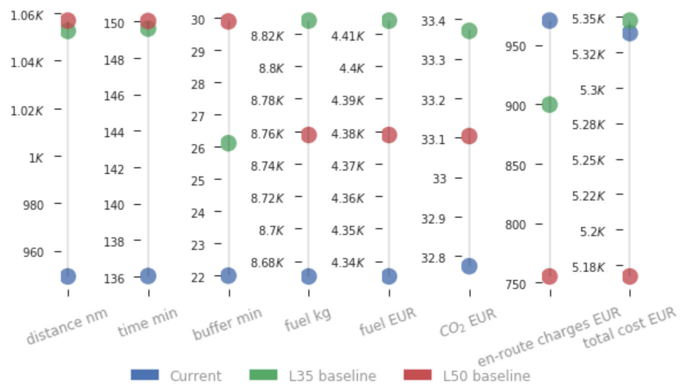

Figure 3. Pre-tactical average flight plan metrics

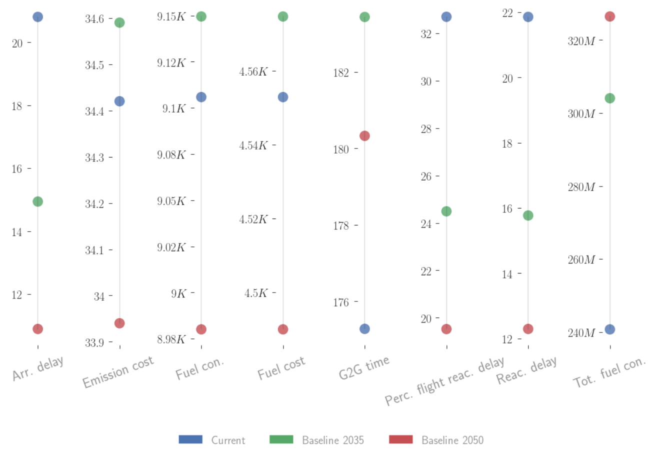

Figure 4. Tactical average flight metrics

higher crew costs. Due to increases in particular in airport fees and various technological improvements leading to flat fuel consumption, the operational cost decreases between 2035 and 2050. It is also interesting to note that the share of the uncertainty in the cost of delay decreases rapidly, due to the implementation of various SESAR solutions.

Airport delay increases monotonically from the current situation, to 2050. This is mainly due to the vast increase of traffic and the inadequacy of the airport capacity increase during the same period. This situation is in fact much worse in the 'High' scenarios (not shown here), since demand increases much faster than in the 'Low' ones. Note also how ATFM delay behaves: from a sizeable increase between the current situation and 2035, to sharp decrease in 2050. This is due to the fact that demand increases substantially before 2035, whereas technological improvements do not support efficiency to catch up with it. Across the whole period, the growth in traffic is smaller and the gains in efficiency much higher, which leads to smaller delays.

B. *Pre-tactical*

Figure 3 presents the results for the flight plan indicators computed pre-tactically. It is interesting to compare them with the strategic results. We can observe that the flight plan distance increases for all the 2035 and 2050 scenarios driven by an increase in OD distances. This increment in distance is aligned with an increment in fuel usage. The average en-route airspace charges tend, however, to decrease over time as ANSPs set a lower unit rate. The increment in OD distance, followed by a reduction in the corresponding trajectory lengths due to SESAR improvements, leads to an increment in the average buffer time.

C. *Tactical*

We show in Figure 4 some metrics obtained from the tactical layer. In terms of delays, the model depicts a positive picture. Arrival delays decrease monotonically. In particular, reactionary delays, the main delay drivers, decrease significantly. This result is in contradiction with some results obtained for the strategic layer. Interestingly, it points to an important inner mechanism of the model. For a given OD pair, the gate-to-gate time decreases due to various improvements in ATM (e.g., free routing). This is associated with higher buffers for the flights as they are based on historical schedules that we did not modify. These additional buffers lead mechanically to a decrease in the average reactionary delay. Note that the economic model assumes that buffers will be adjusted and therefore reactionary delay would be similar to historical levels. This leads to discrepancies in the estimation of delay at the strategic and tactical levels. Agents may decrease their buffers in the future to benefit from more efficient OD times. However, the extent of this buffer reduction is complicated and depends on the business model of the airline. This could be an interesting line of research for the future.

The gate-to-gate time does *not* decrease on average, because airlines tend to operate longer routes in the future. This is reflected in higher fuel consumption and fuel costs in 2035. In 2050, gains in efficiency are sufficient to decrease fuel consumption, even with heavier aircraft and longer routes operated.

D. *Key performance indicators*

In Figure 5 we show some key metrics for stakeholders and for the most complex scenarios, showing the baseline, 'Supportive' and 'Non-supportive' scenarios.

Overall, in both scenarios, over the period we observe smaller costs of delay for the airlines, partly driven by smaller costs of uncertainty. This is mainly due to reactionary delays decreasing, as noted above, this being in turn due to larger buffers. The situation is also quite positive for passengers: lower fare-to-income ratios, at least in the supportive scenario, and shorter door-to-door times. The situation is less positive for the environment. We predict a great increase in the total emissions in the future. This is mainly driven by the increase in traffic, but is also due to airlines using heavier aircraft and operating longer routes. Note that the model did not include any gains in efficiency due to better aircraft design (or switches to alternative power sources), so there is some scope

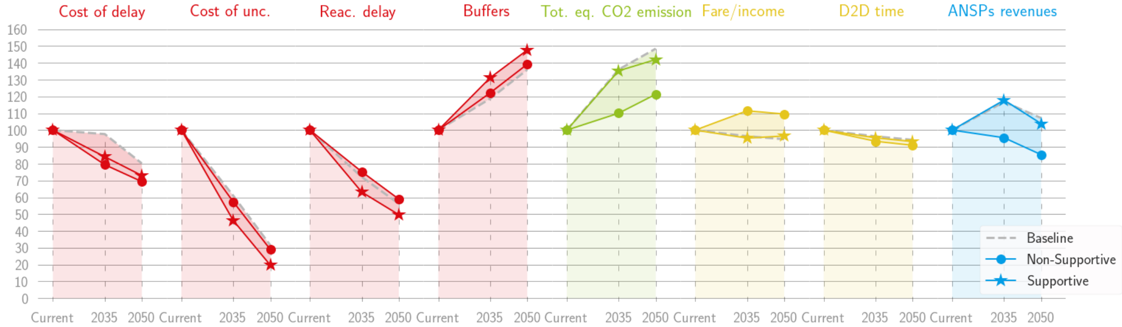

Figure 5. Key indicators for stakeholders by scenario. From left to right: cost of delay per flight; cost of delay uncertainty per flight; reactionary delay per flight; buffer per flight; total equivalent of CO2 emissions; fare-to-income ratio; door-to-door travel time; ANSP revenues.

for improvement here. It is also interesting to track ANSPs' revenues, even though they are not an indicator *per se*: the ANSPs are not in competition in the system and they are profit neutral.

VI. Conclusions and Future Work

The Vista project built a holistic model in order to capture high-level alignments and trade-offs between key indicators. The model is composed of three layers, aligned with the strategic, pre-tactical, and tactical ATM phases.

The core of the strategic layer is the economic model, which allows us to capture complex feedback loops in the supply and demand interplay within the system. It is highly granular compared to standard economic models, using more than 45k interactive agents with their own objective functions and imperfect behaviours to build estimations of key features of potential futures. In particular, the model is able to capture hub vs. point-to-point competition, without assuming prior archetypes for the agents. The pre-tactical layer is heavily based on historical data analysis and allows us to efficiently sample them to produce likely flight plans. The tactical layer is able to use these to execute a typical day of operations, including details at the passenger level in a door-to-door context.

The model has been run on several scenarios, carefully built on various sources of data to represent various potential futures. In particular, scenarios including two degrees of support to the system have been chosen to highlight potential inconsistencies and side-effects.

The results show that the model is able to describe a wide array of metrics at the same time. The strategic layer predicts moderately higher delays, a decrease in delay uncertainty, but overall increasing operational costs. This is partly due to airlines operating longer routes with heavier aircraft. Specific ANSPs' revenues and flight plan characteristics are captured by the pre-tactical layer including an increase in flight-plan buffers. The tactical layer shows that the increase in buffer times translates into smaller reactionary delays. This calls for a more detailed study of the interactions between the layers.

Vista was originally envisioned with an additional 'learning loop' block, which would extract information from the tactical and pre-tactical layers and feed them back to the strategic layer. This would have allowed for a better alignment between the layers, even if one of the features of the model is indeed to have different levels of information between different blocks, as is the case in reality (when tactical execution needs to be thoroughly - but imperfectly - analysed to feed the strategic view of an airline, for example). This loop has not been built in Vista, but is an open possibility for the future of the model. In particular, it could use one of the multi-agents' learning paradigms which have recently been developed. The modular nature of the model also renders it open to future development and the flexible data input structure allows ready modification of the scenarios.

Another area of consideration concerns the sensitivity of the model to initial conditions and parameters. Whilst an analysis of the impact of various parameters has been performed, a more systematic approach is needed in order to fully explore the behaviours of the model(s), in particular the economic model.

As it is, the model can be used to produce high-level results on the potential trade-offs for the future of the ATM system. As highlighted, some system modifications tend to improve part of it and degrade others (e.g. with regard to the environment). Such quantitative analyses can be used to support evidence-led policy making.

Acknowledgements

We thank the Deutsches Zentrum für Luft- und Raumfahrt (DLR) for kindly providing confidential formulae and valuable advice for the evaluation of emissions impacts, based on [23].

References

[1] ICAO, "Manual on global performance of the air navigation system (Doc 9883) Ed. 1," Montreal, 2009.
[2] S. Bouarfa, H. Blom, Curran, and M. R., Everdij, "Agent-based modeling and simulation of emergent behavior in air transportation," *Complex Adaptive Systems Modeling*, vol. 15, no. 1, 2013.
[3] H. Blom, A. Cook, F. Lillo, R. Mantegna, S. Micciché, D. Rivas, Vázquez, R., and M. Zanin, "Applying complexity science to air traffic management," *Journal of Air Transport Management*, no. 42, pp. 149–158, 2015.

[4] G. Gurtner, L. Delgado, A. Cook, J. Martín, S. Cristóbal, and H. Plets, “Building a holistic atm model for future kpi trade-offs,” in *Proceedings of the Seventh SESAR Innovation Days*, 2017.
[5] Helios, “New dawn?” *Air Traffic Management*, pp. 22–24, 2015.
[6] T. D’Alfonso, C. Jiang, and Y. Wan, “Airport pricing, concession revenues and passenger types,” *Journal of Transport Economics and Policy*, no. 1, pp. 71–89, 2013.
[7] P. Berster, M. Gelhausen, and D. Wilken, “Is increasing seat capacity common practice of airlines at congested airports?” *Journal of Air Transport Management*, vol. 46, pp. 1–25, 2013.
[8] N. Adler, P. Forsyth, J. Mueller, and H.-M. Niemeier, “An economic assessment of airport incentive regulation,” *Transport Policy*, vol. 41, pp. 5–15, 2015.
[9] R. Doganis, *Flying off course – airline economics and marketing*. Routledge, 2010.
[10] A. Paul, U. Kluge, A. Cook, and G. Tanner, “Dataset2050 d3.2 - future passenger demand profile,” http:/dataset2050.com, pp. 1–80, 2017.
[11] U. Kluge, A. Paul, A. Cook, and S. Cristóbal, “Factors influencing european passenger demand for air transport,” in *Air Transport Research Society World Conference*, 2017.
[12] ComplexityCosts project, “D4.5 – Final Technical Report,” 2016.
[13] Vista project, “Final assessment report,” Deliverable 5.2 of the Vista project, available on demand, 2018.
[14] EUROCONTROL, “DDR2 reference manual for generic users,” 2015.
[15] M. Smyth and B. Pearce, “Air Travel Demand, IATA economics briefing,” https://www.iata.org/whatwedo/documents/economics/air_travel_demand.pdf, 2008.
[16] N. Adler, “Competition in a deregulated air transportation market,” *European Journal of Operational Research*, vol. 129, no. 2, pp. 337–345, 2001.
[17] STATFOR, “Challengers of growth,” 2013.
[18] A. Cook and G. Tanner, “European airline delay cost reference values, updated and extended values, version 4.1,” 2015.
[19] Cassiopeia project, “DCI-4HD2D - D3.2 Final technical report,” 2016.
[20] EUROCONTROL, “CODA Digest: All-Causes Delay and Cancellations to Air Transport in Europe – 2014,” https://www.eurocontrol.int/sites/default/files/publication/files/coda-digest-annual-2014.pdf, 2015.
[21] DATASET2050, “D5.2 - assessment execution,” http:/dataset2050.com, 2017.
[22] Vista project, “Initial assessment report,” Deliverable 5.1 of the Vista project, available on demand, 2018.
[23] K. Dahlmann, A. Koch, F. Linke, B. Lhrs, V. Grewe, T. Otten, D. Seider, V. Gollnick, and U. Schumann, “Climate-compatible air transport system – climate impact mitigation potential for actual and future aircraft,” *Aerospace*, vol. 38, no. 3, 2016.