The hidden cost of uncertainty for airspace users

Gurtner, G. and Cook, A.J.

The hidden cost of uncertainty for airspace users

Abstract

This article highlights the importance of uncertainty in day-to-day operations, and the need to take it into account to properly assess the cost of delay for airspace users. It defines a cost of uncertainty and estimates it using real data. It provides some easily computable models based on the average and standard deviation of delay to estimate the cost of delay in general. The article shows that uncertainty is also important in the formulation of buffers for airlines and provides a simple model to estimate the optimal assignment, further using real data to compute the optimal value at different airports.

1. Introduction

1.1. SESAR and industry context

SESAR, in its Master Plan (SESAR, 2019), establishes Performance Ambitions for 2035, relating to various ICAO key performance areas (KPAs). The capacity KPA includes a key performance indicator (KPI) target for average departure delay of 6.5 to 8.5 minutes by 2035, representing a relative improvement of 1 to 3 minutes per departure relative to the baseline value of 9.5, in 2012. These Ambitions are aligned with the main regulatory instrument in Europe for the delivery of ATM performance, the Performance Scheme (European Commission, 2019) of the single European sky. The latter initiative was launched by the European Commission, in 2004, particularly in response to the issue of increasing delay in Europe.

Despite low traffic following the 2008 economic crisis, there were six years of growth from 2014, with 2008 volumes recovered in 2016, and European traffic reaching its highest-ever levels in 2019. The general trend of departure delay has not been consistent with the SESAR Ambition, however[1]. 2018 was the fifth consecutive year that punctuality fell (average departure delay: 14.4 mins), with some improvement in 2019 (12.8 mins) (Performance Review Commission, 2020).[2]

The Master Plan (ibid.) cites that "greater predictability is expected to be a key outcome of the deployment of the SESAR target concept", and predictability is one of ICAO's KPAs in its Global ATM Operational Concept ('Doc 9854'). Predictability is captured in the SESAR Performance Framework (EUROCONTROL, 2017), which underpins the validation of SESAR Solutions and their contribution to European performance improvements through the Master Plan. This indicator measures the variance of the difference between actual and planned gate-to-gate durations.[3]

Various entities, including the Performance Review Commission and Performance Review Body (Performance Review Commission, 2020), an independent group of experts established under EU legislation assisting the Commission in the implementation of the SES Performance Scheme, monitor such delays and their associated costs. They are also used extensively in SESAR performance assessment for calculating the cost-benefit trade-offs of implementing various changes: of introducing SESAR Solutions, in particular.

[1] These analyses were performed up until early 2020, before the impact of the Covid-19 pandemic on European traffic. They were made in the context of capacity constraints up until that point in time, and whilst they assume a return to volume in the future, the text has been reviewed to moderate assumptions based on explicit forecasts prevalent in early 2020. Occasional observations are made to reflect the post-Covid-19 situation.

[2] En-route flow management delay improved in 2019 due in part to EUROCONTROL measures helping to reduce demand on constrained area control centres, through the use of level caps and lateral re-routings.

[3] In contrast, EUROCONTROL Performance Review Reports have historically reported variances between flights on given airport pairs.

1.2. Central motivation of this paper

Airspace users and passengers are most impacted by delays. In order to assign a cost to these delays for airlines, a few methodologies have been developed. The main approach in Europe (Cook and Tanner, 2015), most widely used by the industry, details the cost of delay broken down by phase of flight, type of aircraft, cost scenario (type of airline) and, importantly, delay duration.

However, the current use of these values remains often problematic. The values are often used without sufficient caution. In particular, since the main performance indicator for delay is the average, the cost of delay is also often computed as an average value, citing the average value of 100 euros/minute reported (Cook and Tanner, 2015), despite cautionary notes in this report against such insufficiently considered use. In situations whereby analyses focus on mitigation of higher than average delays, notably flight prioritisation mechanisms under demand-capacity imbalance, using empirical averages will result in a (large) underestimation of the true cost, since the cost of delay is non-linear as a function of delay duration.

While the report (ibid.) computes the details of the costs, which, up to the level of precision and the quality of the assumptions, contain all information needed to compute delay costs, we would like to highlight in this article that uncertainties (and, more generally, the shape of distributions) play a crucial role when computing high-level indicators, not least the 'true' average costs, and are sometimes counter-intuitive in their effects. (More subtly, the cost of delay is not negative for negative delay, i.e., airlines very rarely make money or save costs from early arrivals.)

We illustrate through detailed examples in this article how, taking into account the variance of the distribution of delays, and the associated non-linearities, changes the picture considerably. We show how this has significant impacts for cost-benefit analyses, the use and interpretation of performance indicators (which underpin the SESAR Performance Framework (EUROCONTROL, 2017)), the design of SESAR Solutions (e.g., regarding delay cost reduction estimates), and the high costs associated with the deployment of buffers in airline schedules.

1.3. The context of airline buffers

The SESAR Master Plan further explains that specifically, more predictable arrivals are expected from its expected technological changes, which will "have a beneficial effect in reducing the 'buffer time' that airlines factor-in to their schedules to add robustness to tactical time variations" (SESAR, 2019). Decreased buffers could then be taken into account during the strategic (planning) phase and included in the schedule creation process. This would allow for a better utilisation of the aircraft by the airlines, and hence reduce their associated opportunity cost.

Optimising buffers is a complicated task, but crucial for airlines, given the high costs of under-utilisation of an aircraft and of the delays that buffers are intended to mitigate. Finding the right level of buffering is part of the scheduling process for the airline, along with many other parameters. However, buffers are sometimes neglected (Welman et al., 2010; Xu et al., 2008) in delay propagation models, or taken into account only partially (Churchill et al., 2010; Arikan et al., 2013). Buffers are often considered more with regard to their positive contribution, i.e. the robustness afforded to operations (Wong and Tsai, 2012; Kafle and Zou, 2016; Fleurquin et al., 2013) rather than the corresponding cost of resource under-utilisation, although some more extensive studies exist (Ball et al., 2010), estimating the cost of under-utilisation to be 3.7B US dollars for airlines in 2007, excluding the lost demand. Either buffers are considered as fixed values (Campanelli et al., 2016; Gurtner et al., 2018; Pyrgiotis et al., 2013), or in schedule optimisers they are used as adjusting variables (Wu and Caves, 2004). But estimating actual buffers in use is difficult. One can compute the differences between scheduled and actual turnaround times, for instance as in Fleurquin et al. (2013), but it is challenging to infer the part of the difference due to buffers, actual delays, and interactions between flights (waiting for passengers, etc). Some authors consider certain percentage thresholds in the gate-to-gate time to infer nominal flight times and thus buffers (Kafle and Zou, 2016).

While buffers are key to controlling delay propagation, their exact formulation is still typically not well quantified. As is often the case in the deployment of buffers for personal travel planning, uncertainty often plays an important role (Bates et al., 2001). We here explore a simple mechanism for buffer formulation and, in particular, we consider the role of aircraft under-utilisation and delay uncertainty within a (rational) cost-minimisation framework. We show that uncertainty can have an even larger impact than its usual tactical disruption impact, when taking into account the strategic adaptation of airlines to an environment of changing uncertainty.

1.4. Overview of this paper

In this article, we first illustrate what we term the 'cost of uncertainty' in Section 2: why it is important and how to estimate it. Section 3 presents simple statistical models to be used for quick and easy computation, which also give an explicit value for the cost of uncertainty. In Section 4 we show that buffers may also be affected by uncertainty, and should be taken into account in principle to compute their cost. We draw some conclusions in Section 5, in particular regarding the scales of the different costs and their future likely trends.

2. The cost of variance

In this section, we explain the importance of various statistical effects in the estimation of the cost of delay in performance assessment.

2.1. Illustration

We begin with an illustration of how uncertainty can change the real cost of delay, compared to some makeshift computations using only the average delay[4].

As an example, we consider a synthetic normal distribution of arrival delay. In order to obtain the right magnitude for the results, we calibrate this distribution with parameters (mean, variance) from an empirical distribution of delays for a large European hub (see Section 2.2 for more details on the data we use) .

In order to estimate the tactical cost of delay for hypothetical flights experiencing these delays, we use a cost of delay function, derived from Cook and Tanner (2015):

$$\begin{aligned} c &= \alpha_1 \delta t + \beta_1 \delta t^2 + (\alpha_2 \delta t - \beta_2 \delta t^2)\sqrt{MTOW} \quad \text{if } \delta t \geq 0, \\ &= 0 \quad \text{otherwise.} \end{aligned} \tag{1}$$

This equation is from a quadratic fit on the delay δt and a linear fit on the root square of the maximum take-off weight $\sqrt{MTOW}$, the latter affording a particularly good fit. α_1, α_2, β_1, β_2 are coefficients from the regression. The fit using $\sqrt{MTOW}$ is discussed in extensive detail in Cook and Tanner (2015), whereby other fits are also investigated, such as with passenger numbers. These both give good fits, and are, of course, correlated with the size of the aircraft, and thus fuel burn costs and, to a less pronounced extent, crew costs (crew numbers being regulated by aircraft configurations and passenger numbers). It is thus logical that delay costs should be well modelled as a function of MTOW, or passenger numbers. In practice, there is relatively little to differentiate between these fits, although $\sqrt{MTOW}$ is marginally better at commoner (lower) delay values. (Fits therewith across delay durations from 5 to 300 minutes all demonstrate r^2 (sample coefficient of determination) values above 0.98 for at-gate and en-route, full tactical costs. Good fits are similarly illustrated for strategic (e.g. buffer) costs.) However, the more compelling argument for use of an MTOW fit is that MTOW data are typically more readily available in external sources (e.g. several from EUROCONTROL, as we shall see later) and are hence more useful to other researchers for dedicated fits and/or extrapolation or interpolation to other aircraft types, not directly modelled. The use of $\sqrt{MTOW}$ by such external sources is, of course, due to the fact that this is the basis of the formula for en-route charges in Europe. We

[4] For greater clarity, we specify here exactly what we mean by 'expectation' and 'average'. The expectation of the function of a stochastic variable is to be understood in the mathematical sense, as in Equation 2. The expectation of a quantity coincides with its empirical sample average when it is computed over the sample distribution. In this article, we thus reserve the term 'average' for empirical (sample) averages and 'expectation' for theoretical computations, although these terms can sometimes be exchanged.

choose $\sqrt{MTOW}$ to be the average observed at the European hub. We are then interested in computing the expected value of the cost over the distribution of delay, i.e.:

$$c^e(\sigma) = \int c(\delta t)\, p_\sigma(\delta t)\, d\delta t, \tag{2}$$

where we have thus highlighted the importance of the uncertainty, by explicitly defining this expected value as a function of the standard deviation σ of the distribution of delay. We then keep the mean of the distribution of delay fixed, and change the variance, while computing the expected cost. Figure 1 shows the result.

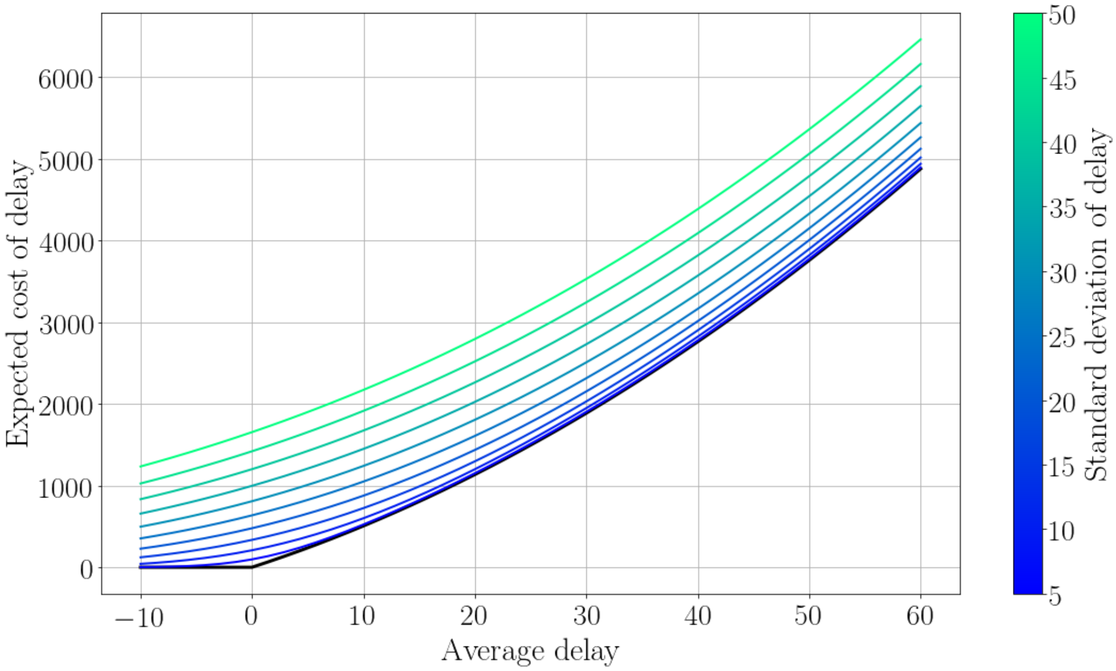

Figure 1: Cost of delay as a function of the mean delay for different standard deviations.

The black line corresponds to the case with zero variance. In this case, all the flights are delayed by exactly the same delay, and the corresponding expected cost is thus the cost of the expected delay. It is null for negative delays, and increases for positive delays. As the variance increases, the expected cost of delay increases with respect to the black line. For the variance observed in the data (upper curve), the expected cost is significantly higher than the cost of the expected delay. For very high delay, the cost of uncertainty (the part above the black line) is more than 30% of the total cost. It is easily above 50% even for moderate delays.[5]

2.2. Empirical relationship between cost of uncertainty and variance

The expected cost of delay thus depends on the delay probability distribution and the traffic mix. Thus, the cost of delay (driven in large part by late-arriving passengers) that an airline may expect depends on the destination airport. Given that airports are different in terms of delay distributions and traffic mix, this gives us an opportunity to infer a first relationship between uncertainty and the cost of delay.

In order to do this, we used empirical data. We have access to two types of data, which allow us to make a good estimation of the actual delay experienced by flights. Firstly, we used EUROCONTROL DDR data, in particular, the flown ('M3') data, containing the actual landing

[5]Note that in this example, the gap between the 'true' curve and the zero-variance one decreases with the average delay. This is not the case in general, and depends on the underlying distribution.

time (in addition to the type of aircraft, the airline, etc.). Unfortunately, DDR data do not contain information on the actual in-block time, so we used average taxi times (from EUROCONTROL CODA data) at the various airports to estimate them.

The second set of data is published schedules. These contain scheduled in-block times that the airlines planned. These data can be misleading, however, given that airlines may declare a schedule different from what they are planning to do, in order to game the system somewhat. However, we believe that these data represent a much better estimate of the intent of the airlines than other types of data, such as last-filed flight plans (DDR 'm1' data). In this paper, we use one full day of traffic data, for 12 September 2014. This decision is dictated by the fact that this particular day has been used several times by the team in past projects and is well understood. It is free of major disruptions (e.g. strikes) or anomalies (e.g. severe weather). However, one should bear in mind that the results of this article have been obtained on one day, and may differ if reproduced on other sets of data.

We then fused the two datasets. As a result, for each flight in the data (around 27 000 passenger flights, removing all-cargo flights, military operations, test flights, etc.), we have access in particular to the:

- destination airport;
- scheduled in-block time;
- (estimated) actual in-block time;
- type of aircraft.

As noted, since most delay costs are typically incurred due to **arrival delay**, we define 'delay' as arrival delay in the following discussion, i.e. the (estimated) actual in-block time, minus the scheduled time.

We start by computing the expected cost of delay for each airport, regardless of the type of aircraft. In order to highlight the role uncertainty, we also compute the cost of the average delay at each airport. Figure 2 shows the results of this computation.

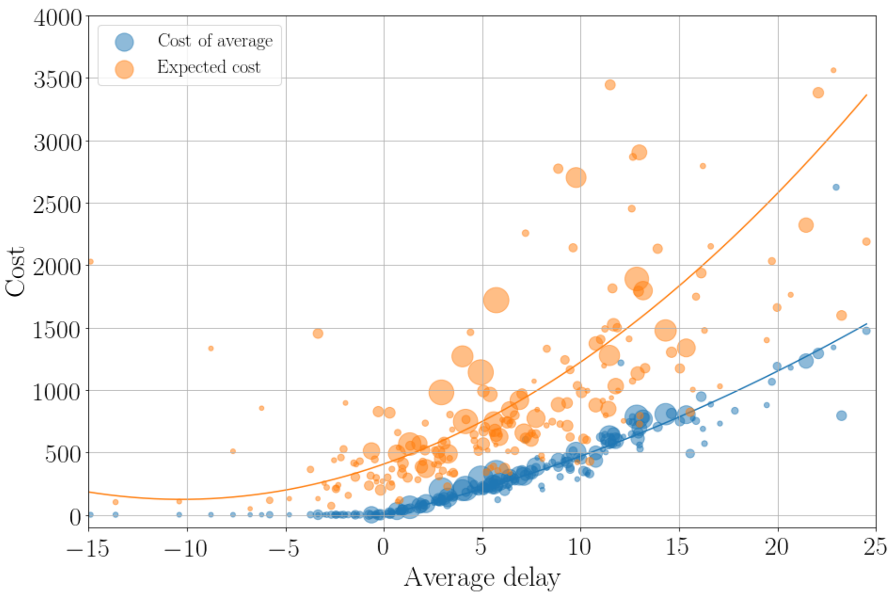

Figure 2: Expected cost of delay (orange) and cost of average delay (blue) for each airport (around 800). The size of the dot is proportional to the traffic at the airport (in the dataset). The solid lines are quadratic fits. The figure has been zoomed on the most interesting region; some (minor) airports lie outside the range presented.

First, it is clear that using the cost of the average delay severely underestimates the actual cost, and should not be used in general. The difference between the orange and blue dots can be thought of as the 'pure' cost of uncertainty in the system, i.e. the cost having a non-zero variance in the distribution of delay. Second, there is quite a lot of variety between airports, even for similar average delay.

In order to explore the exact effect of uncertainty, we define the cost of uncertainty as:

$$c^\sigma = c^e - c(\delta t^e), \tag{3}$$

i.e. the difference between the expected cost and the cost of the expected delay. In Figure 3, we compute this cost for each airport as a function of the uncertainty, i.e. the variance of the distribution of delay.

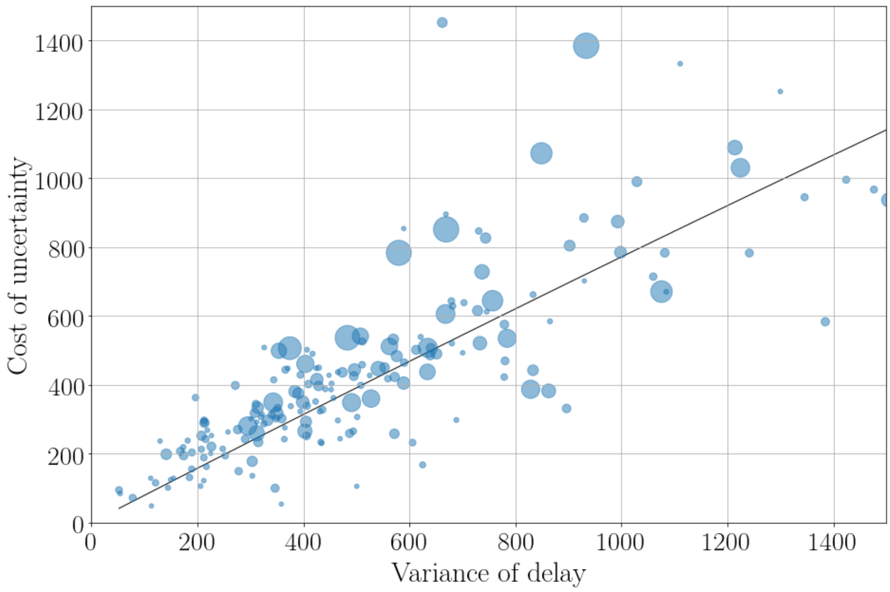

Figure 3: Cost of uncertainty for each airport against the variance of the delay at the airport. The black solid line is a quadratic fit ($R^2 = 0.86$).

The cost of uncertainty exhibits quite a linear behaviour with the variance of the delay. The slope (0.79 euros.min^{-2}) represents the empirical cost of a 1 minute square of uncertainty. Note that because the variance and mean are linked empirically, this value also encapsulates the effect of this relationship.

2.3. 'Pure' cost of uncertainty

Figure 3 demonstrates the empirical link between the variance and cost of uncertainty. However, because the relationship is inferred from different airports at the same time, it subsumes several phenomena. In particular, it does not answer the question: "For a given airport, what would be the cost of uncertainty if the latter decreased?". This is because every airport is different in terms of delay distribution: not only in terms of means and variances, but also of traffic mix.

From the performance monitoring point of view, this is very important. Indeed, for instance, one may consider a new SESAR Solution that increases by a given factor, X, the predictability KPI (decreases uncertainty). If the mean delay is constant, it is insufficient to compute the actual benefits.

In order to explore this issue, we define a simple model. For each airport, we have already access to its distribution of delay. In order to consider the effect of the variance only, we simply define a homothety for the distribution, in order to keep the mean constant. If x_i is a point of the distribution, we define the following transformation:

$$\forall i,\, x_i \leftarrow \mu_0 - \frac{\sigma}{\sigma_0}(\mu_0 - x_i), \tag{4}$$

where σ_0 and σ are the initial and new standard deviations, respectively, and μ_0 is the distribution mean. By sweeping the value of σ, we have access to a series of distributions for which we can compute the expectation of the delay, hence isolating the effect of the variance. Thus, we can reuse (2), building the p.d.f. $p_\sigma(\delta t)$ with the transformation (4). We call this the 'pure' cost of uncertainty, as it represents the extra cost incurred by the system under increasing uncertainty, discarding the effect of the average. For example, a SESAR Solution decreasing variance by a given percentage could claim a reduction of cost thanks to this formulation, even if its formal performance assessment is unsure regarding the exact impact of the Solution on the average delay.

In Figure 4 we show how the expected cost of delay varies when the transformation is applied to the various airports' distributions of delay. We also show the points where $\sigma = \sigma_0$, i.e. we evaluate the cost of uncertainty at the initial standard deviation (from the data), similarly to Figure 3.

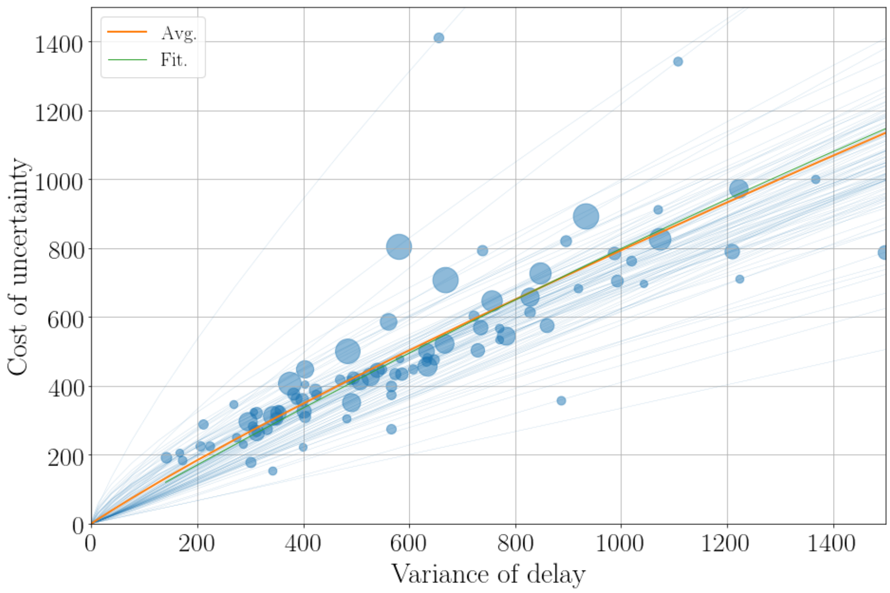

Figure 4: Cost of uncertainty for each airport. Each blue line is the result of the transformation (4) of the equation (2), for a given airport. The circles represent $c^e_d(\sigma_0)$ for each airport, i.e. the expected cost of delay estimated at the empirical standard deviation. The orange solid line represents the (weighted) average function among all airports. The solid green line is a quadratic fit. Only the largest 100 airports are represented.

This representation allows us to better see how the various airports are different and how their associated costs would change if the variance of delay were higher (but keeping the mean fixed). We also computed the average function, weighted by the airport traffic, and found that the average is very well fitted by a quadratic function, with a slope slightly steeper than for Figure 3 (0.91 euros.m^{-2}) and a small negative quadratic term ($-1.2\,10^{-4}$ euros.m^{-4}). 95% of the linear coefficients are within the range [0.72, 1.2].

3. Estimating the cost of delay in practice

In this section, we explore in greater detail the relationship between expected cost, mean delay, and variance. The goal is to provide some easy computational means for performance monitoring, as well as estimations of errors when using them. In particular, we point out the need to estimate the variance of distributions of delays to properly estimate the cost.

3.1. Link between mean and variance

As mentioned previously, there is an empirical link between the variance and mean of a delay distribution. Whilst changing one or the other independently may in some cases be possible, in general changing the underlying dynamic processes leads to changing both. Hence, it is important to bear in mind that, for example, KPIs for predictability (uncertainty) and KPIs for mean delay are interdependent.

Figure 5 illustrates this point. In this figure, we simply plotted the average delay for each airport as a function of the standard deviation. There is quite a weak, but significant, correlation between them (Pearson's correlation coefficient: 0.37). As a consequence, some information on the variance is included in the mean, and vice versa. In the figure, we also show how the cost of delay changes for each airport of the dataset.

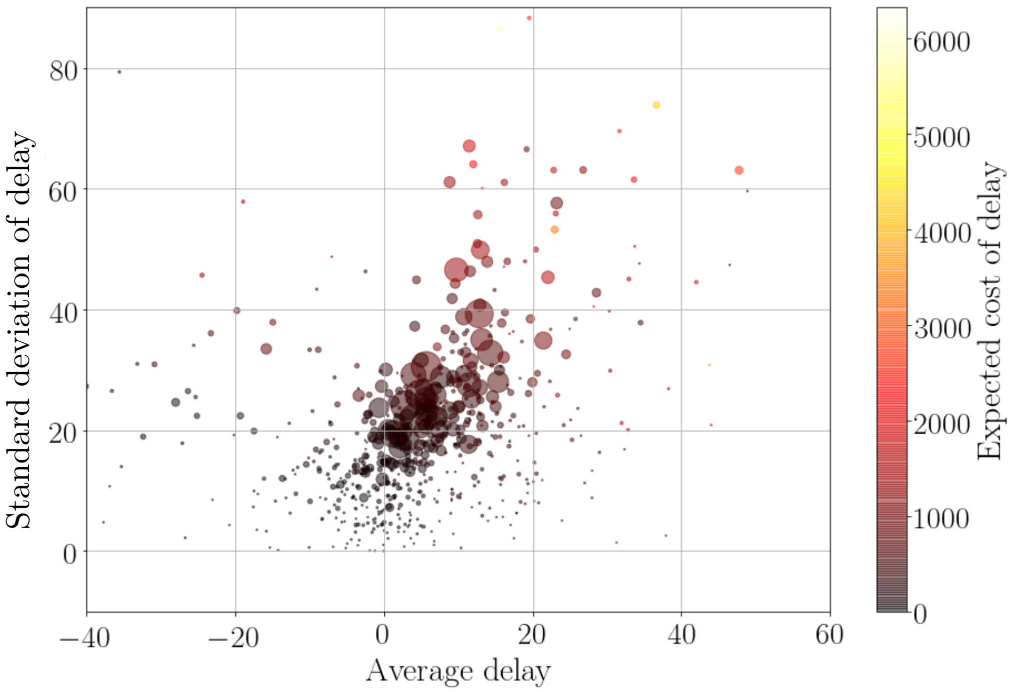

Figure 5: Standard deviation of delay for each airport as a function of the average delay. The circles are coloured according to the expected cost of delay at the airport (see key).

3.2. Empirical relationship between cost, mean, and variance

In order to provide simple and usable coefficients that can be used in various contexts, we consider ordinary least-squares regressions (OLS) with different formulas. Given the μ and V, the empirical mean and variance of the distributions of delay, we test the following formulas[6]:

$$c^e = a\mu \tag{5}$$

$$c^e = a\mu + b\mu^2 \tag{6}$$

$$c^e = a\mu + cV \tag{7}$$

$$c^e = a\mu + b\mu^2 + cV \tag{8}$$

Since we are usually interested in the positive mean delay area, we performed the regression only in this area[7].

The results are compiled in Table 1. Formulas (5) and (6), not taking the variance into account, perform poorly. The first is of particular importance, as it is how some researchers perform simple

[6] We omit the constant term because it is never significantly different from zero in practice.

[7] The results are similar when including the negative part, except for the fact that the simpler function c_d performs better in this area (as it is non-linear by construction) and that the linear law $c = a\mu$ performs better in the positive area.

calculations on the European cost of delay[8]. Both the other regressions perform reasonably well. The information contained in μ^2 does not much improve the goodness of fit, cf. the information on the variance, V. In both cases, we find a coefficient of around 0.62, smaller than found in sections 2.2 and 2.3. The coefficient of the mean delay is significantly smaller than for Equation (5).

	a	b	c	Adj. R^2
$a\mu$	244 [222, 266]	-	-	0.52
$a\mu + b\mu^2$	143 [99.1, 187]	0.58 [0.36, 0.81]	-	0.55
$a\mu + cV$	78.5 [64.3, 92.7]	-	0.62 [0.58, 0.65]	0.88
$a\mu + b\mu^2 + cV$	100 [77.4, 123]	-0.15* [-0.27, -0.03]	0.63 [0.60, 0.67]	0.88
$c_d(\mu)$	-	-	-	0.67

Table 1: Results of OLS regressions using different formulas. All parameters are significantly different from 0 to a 1% significance threshold, expect that marked '*'. The numbers in parenthesis represent the 95% confidence intervals. The last line shows the goodness of fit using only the cost of the average.

The goodness of fit as estimated by R^2 can hide different realities. In particular, using the above formulas on different levels of delay can lead to various degrees of error. In order to estimate the error, we compute the average overestimation $\bar{\delta t}_{fit} - \bar{\delta t}$ and the overestimation ratio $(\bar{\delta t}_{fit} - \bar{\delta t})/\bar{\delta t}$ computed on three levels of delay: 'small' (less than 5 minutes), 'medium' (between 5 and 10 minutes), and 'large' (greater than 10 minutes). We also compute R^2 on these restricted sets.

	$0 \leq \delta t < 5$	$5 \leq \delta t < 10$	$\delta t \geq 10$	All
$c = a\mu$	-1,190 euros	-370 euros	969 euros	0 euros
	-309%	-48.9%	32.0%	0%
	-16.5	-0.58	0.52	0.52
$c = a\mu + b\mu^2$	-446 euros	-93 euros	341 euros	0 euros
	-116%	-12.3%	11.2%	0%
	-2.32	0.058	0.53	0.55
$c = a\mu + cV$	-151 euros	7 euros	96 euros	0 euros
	-39.3%	0.94%	3.15%	0%
	0.33	0.63	0.87	0.88
$c = a\mu + b\mu^2 + cV$	-314 euros	-54 euros	235 euros	0 euros
	-81.9%	-7.18%	7.74%	0%
	-0.56	0.61	0.88	0.88
$c_d(\mu)$	-284 euros	-418 euros	-1,370 euros	-817 euros
	-74.1%	-55.4%	-45.3%	-48.3%
	-24.4	-12.9	-0.011	0.023

Table 2: Average overestimation (rounded euros, 3 s.f.), overestimation ratio, and R^2 for different formulas for small, medium and large delays, and for all data. The last line shows the same metrics computed on the cost of the average delay. Note that the last column shows 0 for the mean overestimation for all regressions because by construction the mean delay is zero at minimum MSE (which is the regression objective).

Table 2 shows the result of this computation. As expected, the linear law performs poorly everywhere. It tends to massively underestimate the cost when the mean delay is small, and it tends to overestimate it by a lot for the large delays. Adding μ^2 to the fit actually helps quite substantially in terms of not overestimating the cost, though one can expect an error of 10% at least. On the contrary, adding the variance into the equation allows one to estimate much better

[8]Usually using 100 euros/minute, derived from (Cook and Tanner, 2015).

the cost for medium and large delays, even if the fit is poor for small ones. Adding μ^2 seems to decrease the performance from this point of view. Estimating the cost using only the average delay leads to severe underestimation, as shown previously.

3.3. 'Pure' costs for mean delay and uncertainty

The same kind of analysis as performed in Section 2.3 can be made, using the mean delay and the variance at the same time. The process is the same: we modify empirical distributions via the transformation (4) and the corresponding one for the mean:

$$\forall i,\, x_i \leftarrow x_i - \mu + \mu_0, \tag{9}$$

where μ_0 is the initial mean delay and μ the new one. Sweeping the values of μ and σ, we sample the average cost of delay for each airport, compute a weighted average for airports, and use a linear regression to produce usable coefficients. We thus find that a regression $c = a\mu + cV$ yields $a = 58.1[57.6, 58.6]$ and $c = 0.701[0.694, 0.709]$, with $R^2 = 0.998$.

Interestingly, the coefficient for the variance is higher than those reported in Table 1. This is due to the fact that in the latter, the regression does not take into account the different traffic at the airports, in contrast to this procedure.

3.4. Recommendations regarding cost of delay estimations

We discuss in this section the appropriateness of different types of delay cost estimation, with particular regard to the capture of the cost of uncertainty. The reader may note that the specific values cited result from the (e.g. airport) data sampled, as explained, but the methodological philosophy is generic.

Given appropriate input cost data, such as those described in (Cook and Tanner, 2015), the estimation of the cost of delay is readily computable when one has access to fully disaggregated flight data, but it is more problematic when one has to extrapolate, for instance, the impact of a given change of procedure, at the European scale, without a microscopic model. Moreover, the primary impact of changes are sometimes non-independent, as is typically the case when new SESAR Solutions (for example) impact both mean delay and uncertainty. Hence, based on the calculations above, we form some recommendations on how to estimate the full cost of delay in different cases.

First, in the least accurate case, a change may initially be considered to only modify the mean arrival delay, without consideration of uncertainty. However, it is better to at least assume that the uncertainty is impacted as well, and thus the correct formula to use is either one such as the first row of Table 1, or the second row for added precision. 244 euros per minute would be used in the first case. In the case where one knows roughly the absolute level of delay, one can use a more precise formula, for example deploying the microscopic function c_d, but bearing in mind that the full cost is underestimated by some 45-75%, as indicated in the last row of Table 2.

In the (probably rare) case where a modification is definitively known to change only the mean delay, or the uncertainty, then values such as those from Section 3.3 should be used, i.e. 58 euros per minute and 0.70 euros per unit of variance.

The procedure used to obtained these two values assumes a constant shape for the distribution of delay when the variance is modified. One may, alternatively, use a relationship that is more robust to these shapes. The third row of Table 1 represents such a relationship, but this will underestimate the weighted cost at a European level, as it does not use traffic weights. Where more information is available, such as types of aircraft and passenger types and load factors, they should also be considered. The reader may also care to be reminded that the average value of 100 euros per minute often cited from (Cook and Tanner, 2015), was itself generated from an empirical distribution, and the corresponding disaggregated input costs detailed in this report should be used wherever possible in an appropriate microscopic model.

4. Uncertainty and airline buffers

When aircraft arrive at their destination airport, there are numerous turnaround processes (such as cleaning and refuelling) to complete before the aircraft is ready for its next rotation. When these processes take more time than expected, and generate new (or additional) delay to the next rotation of the aircraft, it is termed 'rotational' delay. The flight may of course itself arrive late in the first place, or need to wait for connecting passengers or a new crew from another aircraft. When propagated to the next rotation, the former is termed 'rotational' reactionary delay, the latter 'non-rotational' reactionary delay (Cook and Tanner, 2015). Reactionary delays can be very expensive to the airline, as they often propagate and compound throughout the operational day (ibid.). Reactionary delay was indeed the commonest delay type in 2019, in Europe, at 44.4%, followed by turnaround delay 32.6% (Performance Review Commission, 2020). Airlines often implement strategies to mitigate delay to the subsequent rotation. An obvious one is to plan turnaround buffer, i.e. extra time between flights. In the following we propose a simple model to estimate the optimal amount of such buffer that an airline should choose, based on the distribution of delay at arrival, and balancing strategic and tactical costs. In particular, we highlight the role of the uncertainty of the arrival time, in the formation of optimal buffers.

4.1. Simple buffer formation mechanism

First, we take the simple case of two flights (rotations) of an aircraft, A and B. Flight A is supposed to arrive at $t = t_a$. If the turnaround process takes τ_{ta}, then flight B can be ready to depart at $t = t_d^* = t_a + \tau_{ta}$. Thus, any extra scheduled time with respect to this can be considered as buffer.[9] In practice, in order to assign the buffer, one has to estimate both t_a and τ_{ta}, for example using their average values. In this case, the buffer is simply:

$$b = t_d^s - \langle t_a \rangle + \langle \tau_{ta} \rangle, \tag{10}$$

where the $\langle . \rangle$ denote empirical averages, computed, for example, for such flights based on past data.

This simple rule may, however, not give the best results for the airlines. Indeed, given that the cost of delay is quadratic (Cook and Tanner, 2015), a wide distribution of t_a will trigger significant delay costs. These should be integrated into the decision-making process for buffer assignment.

4.2. Cost-driven formation of buffers

A proper choice for the buffer should thus take into account costs and distributions of delay. For a given aircraft of a given airline, the tactical cost of departure delay is given by Equation (1), that we rewrite:

$$c(\delta t_d) \quad = \tilde{c}(\delta t_d) H(\delta t_d), \tag{11}$$

with $\tilde{c}(x) = \alpha x + \beta x^2$, $\alpha = \alpha_1 + \sqrt{MTOW}\alpha_2$, $\beta = \beta_1 + \sqrt{MTOW}\beta_2$, and H the Heaviside function: equal to 1 if its argument is positive and null otherwise. δt_d is the departure delay. Let us now assume that the previous flight arrives at time t_a, but was initially planned to arrive at time t_a^p, resulting in a delay $\delta t = t_a - t_a^p$. The departure delay of the next flight will thus be:

$$\begin{aligned} \delta t_d \quad &= 0 \quad \text{if} \quad \delta t < b \\ &= \delta t - b \quad \text{otherwise}, \end{aligned} \tag{12}$$

if we assume that the turnaround time is a constant. The corresponding expected cost is:

$$c_t = \int_{-\infty}^{+\infty} c(\delta t - b)\, p(\delta t)\, d\delta t. \tag{13}$$

[9] Here, we define (turnaround) buffer as this extra time. Note that this extra time may be present for various reasons, such as allowing for inbound connecting passengers, or even due to the limited availability of airport slots. Network carriers tend to add more buffer than 'low cost' or point-to-point operators.

This tactical cost is a decreasing function of the buffer b, i.e. the tactical cost of delay always decreases with the buffer.

As explained above, increasing a buffer is costly because of the under-utilisation of the aircraft, for which, following the empirical work of Cook and Tanner (2015), we choose a linear function, i.e.:

$$c_s = \eta b \tag{14}$$

As a consequence, the airline should balance one cost against the other, or, more precisely, try to minimise $c_t + c_s$ as a function of b.

Let us start with the slightly easier case where $\beta = 0$. In this case, it is easy to show that the minimum cost is reached for a buffer $b^* = F^{-1}(1 - \eta/\alpha)$, where F is the cumulative distribution function of the delay. This expression provides some important insights into the formation of buffer. Firstly, it tells us that when $\eta > \alpha$, there is no minimum greater than 0 (the lower bound for b). In other words, when the linear cost per minute of under-utilisation of the aircraft is greater than a linear tactical cost per minute, the best option for the airline is to forget about buffers: any extra buffer time will cost money. When $\eta < \alpha$, then there is a solution with a strictly positive buffer. In this case, it is worth keeping some buffer time to mitigate the effect of the bulk of the delay. However, decreasing returns in terms of extra buffer time apply, and a finite time is thus optimal. In the case where η tends to 0, there are only tactical costs, and the optimal strategy is to have a buffer as large as the highest tactical delay one can experience.

This simple law is, however, complicated by the appearance of a quadratic term in the proper cost function, because the airline has to take into account that high delays, although rare, are very costly. It is easy to show that the solution in this case is given by the implicit equation:

$$F(b) = 1 - \frac{\eta}{\alpha} + \frac{\beta}{\alpha} g(b), \tag{15}$$

where $g(b) = \int_b^{+\infty} x p(x+b) dx > 0$ is an expression that depends only on the distribution of delay (and b). Since $g(b)$ is always positive, the optimal buffer in this case is always greater than in the linear case. In other words, airlines have to increase buffer in order to mitigate non-linearity effects in the cost of delay. In general, b^* does not have an explicit form, and one needs to specify a corresponding distribution in order to proceed further. Note that this result does not take into account any kind of risk aversion that may be exhibited by the airline: this is a purely risk-neutral cost minimisation. Taking into account risk aversion would lead to even higher buffer values.

4.3. Empirical buffers

Using Equation (15), we can, in principle, compute the optimal buffers from the data we used previously. There is, however, an additional difficulty when using these data. Indeed, in the previous section we defined δt as being the difference between the actual arrival time and the 'planned' (filed) time. The latter may be naively considered as the scheduled time. However, it is known that for some airlines there is a significant difference between the filed plan and the intended outcome (e.g. to preempt departure delay). There is very little quantitative information on this sensitive behaviour, however, and practice varies by context and airline, even by time of day.

We thus try to derive the part of the buffer that is due to uncertainty. Indeed, by choosing $t_a^p = \langle t_a \rangle$, $p(\delta t)$ becomes the centred version of the distribution of arrival delay. In other words, we consider that the airline has already taken into account the average delay for its flights, and we are interested in the extra buffer due to the uncertainty.

An example is given in Figure 6, with the cost function of the A320 and the distribution of delay at Heathrow. As expected from Equation (15), buffers increase with the quadratic coefficient, β, of the cost function. The buffer is actually quite sensitive to β, and thus any proper buffer estimation should not use only a simplified linear equation.

Note that the right-hand part of Figure 6 is quite flat with respect to η, i.e. the buffer does not depend strongly on this variable in this area. Interestingly, the empirical value lies in this area, as also seen in Figure 7. All the buffers are within a small range, roughly 8-20 minutes, for this particular airport (Heathrow).

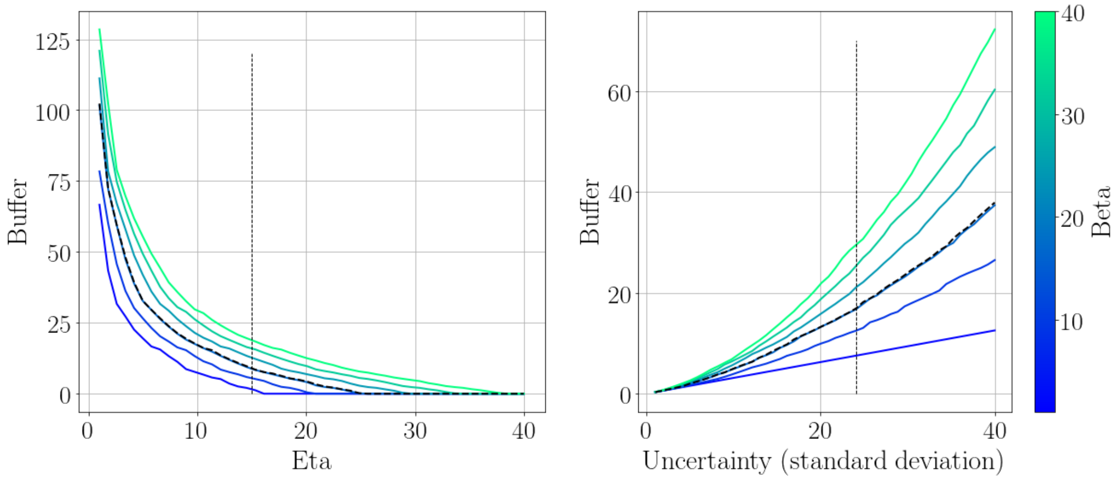

Figure 6: Example (for an A320 at London Heathrow) of the evolution of the optimal buffer as a function of η (left) and the uncertainty (right) for different values of β. The dashed black lines correspond to empirical values of β and the uncertainty.

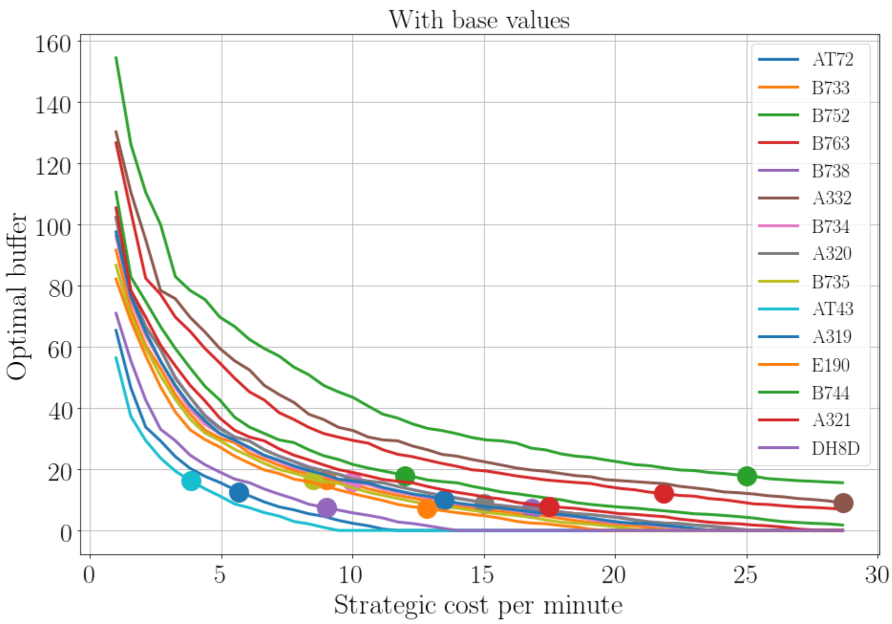

Figure 7: Evolution of the optimal buffer with respect to η for different aircraft cost functions of the Heathrow delay distribution. Circles represent the empirical values of the strategic cost.

It is also interesting to track the evolution of tactical and strategic costs at the optimal buffer value. Figure 8 shows the evolution of each type of cost at the optimal buffer value, estimated for different uncertainty values. In the right panel, one can see that the tactical cost takes up most of the total cost, but decreases in proportion with the standard deviation of the distribution of delay.

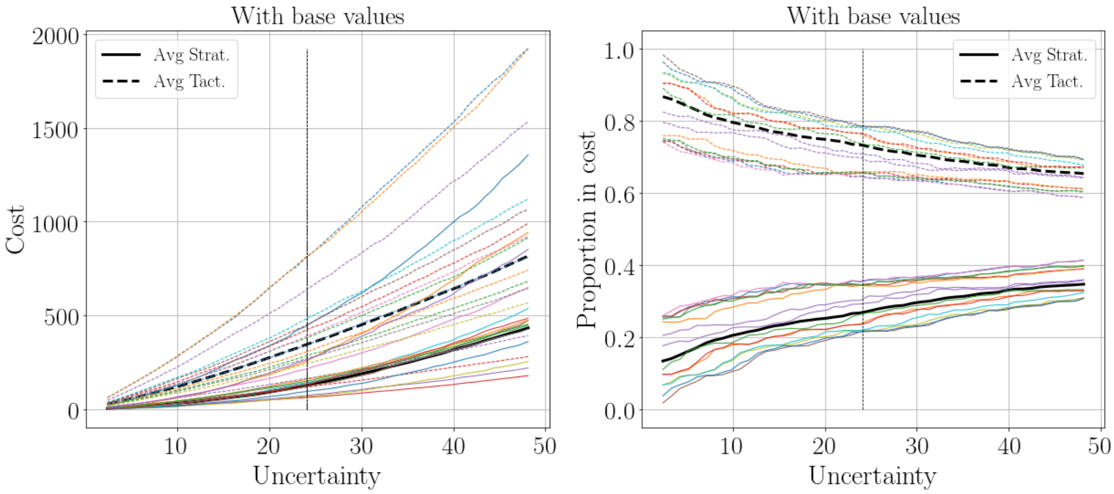

Figure 8: Evolution of the tactical (solid lines) and strategic (dashed lines) costs for different aircraft (see Figure 7 for the colour code of the aircraft) for Heathrow. The black lines represent averages, weighted by the traffic at the airport. The left panel shows absolute values and the right panel shows the proportion of each type in the total cost.

Finally, we compute the size of the average optimal buffers for different airports, weighted by the traffic mix at each airport. Figure 9 shows the results, where we plot the three different cost 'scenarios' described in Cook and Tanner (2015), where 'high' corresponds to a pessimistic scenario (high costs), 'low' to an optimistic one, and 'base' to the medium cost case. Note that the buffers in the 'base' scenario are higher than in the two other scenarios, even though the coefficients in the cost of delay function for the base scenario are systematically between those of the low and high cases. This is because of the non-trivial interplay between the (strategic) cost of under-utilisation and the (tactical) cost of delay. For low uncertainty levels, the buffer value is not very sensitive to the scenario, whereas for high values of uncertainty, substantial differences can be seen. This emphasises the need for airlines to have more reliable information when the operational environment is more volatile, as this can have an important impact on their (scheduling and buffer) decisions.

5. Conclusions

Understanding the full cost of delay is crucial in the delivery of a more efficient air transport system. While various indicators (such as punctuality) can be used as proxies for cost, the latter is paramount for airspace users – in tandem with safety and, the related issue of market (public) image. However, computing the full cost of delay is a complicated issue. Flights are highly heterogeneous, as are specific times of operations in terms of crew, passenger, and weather factors. A shift of paradigm from delay minimisation only to consider broader cost efficiencies, for example, is a core part of research initiatives such as SESAR.

Previous work on the cost of delay (Cook and Tanner, 2015) is routinely used in the industry, including in performance assessment, as we have discussed. However, whilst the cost of X minutes of average delay is relatively well understood, the cost of uncertainty, i.e. the cost of not knowing exactly which delay can be expected, is still largely a hidden cost. It is important that this is better quantified in future performance assessment.

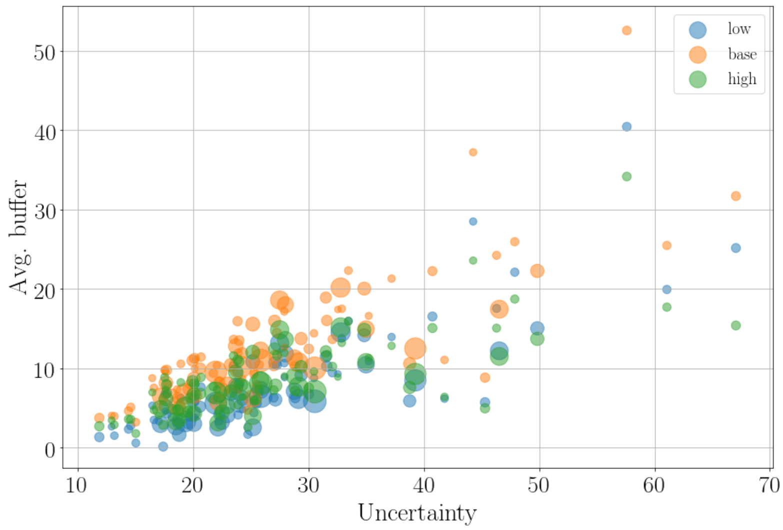

Figure 9: Optimal buffers for the 100 largest airports in our dataset. The circle size is proportional to the traffic at the airport. The three colours correspond to the three different cost scenarios in (Cook and Tanner, 2015).

This hidden cost should be understood and taken into account, since it is a significant part of the full cost of delay. ATM solutions that decrease the variance of delay may have a sizeable performance impact, even if the average delay is not significantly reduced. In this article, we have provided the means to carry out some relatively straightforward calculations on this, and highlighted the potential shortcomings of the method, and, moreover, of not applying any such consideration at all. This approach could be adopted by various applied and exploratory projects measuring reductions in average delay and uncertainty.

Uncertainty not only has the mechanical effect of inflating costs. It can also have some profound effects on the behaviours of different actors in the system. In particular, airspace users may react to uncertainty by modifying their buffers. Buffer formation is still very much a hidden facet of airspace user behaviour, but could be crucial to predict this, improve forecasting, and estimate the corresponding costs. In this paper, we provided a simple model for this formulation, underlining that, with real cost values, they may be considerable.

Note that the validity of our buffer formation model relies on four main assumptions:

- the distribution of delays sufficiently approximates airline experience;
- the cost of delay function sufficiently approximates airline's specific costs;
- the airlines are rational in their choice of buffers, i.e. they compute the expectation of their cost functions using the distribution of delay, and maximise their profit;
- other potential considerations for buffers are negligible with respect to buffer formation.

Clearly, the first two conditions depend on the quality of the delay data. If an airline wanted to use such distributions, it could seek appropriate input information. The fourth condition is more critical, since specific crew and passenger connectivities may play a large role in the flight turnaround time. What we are really computing is the minimum turnaround time needed to avoid losing money from delays (and, more specifically, from the variance of the delay), whereas in reality, specific conditions (such as connectivities) may result in a higher required turnaround time and buffers.

The third condition is particularly interesting. We believe that airlines usually use simpler approaches than a full cost minimisation per flight, using rules of thumb and generalised values for their fleet. While this reduces somewhat the power of our model, in the sense that airlines do not behave fully as expected, it presents them with the possibility of improving their methodology.

Indeed, the procedure explained here could be used by airlines to determine their optimal buffers in a data-driven manner.

It is clear that the above analyses may be enhanced with additional data. A particular barrier to this is the cost of schedule data, the availability of actual intended times, more detailed data relating to flown trajectories and flight phases, and the perennial issue of data cleaning. In future, the buffer models could also be extended to embrace wider network effects, i.e. further downstream delay protection.

References

Arikan, M., Deshpande, V., Sohoni, M., 2013. Building reliable air-travel infrastructure using stochastic models of airline networks. Oper. Res. 61.

Ball, M., Barnhart, C., Dresner, M., Hansen, M., Neels, K., Odoni, A., Peterson, E., Sherry, L., Trani, A., Zou, B., 2010. Total delay impact study. Technical report, National Center of Excellence for Aviations Operations Research (NEXTOR).

Bates, J., Polak, J., Jones, P., Cook, A., 2001. The valuation of reliability for personal travel. Transportation Research Part E 37 (2-3), 191–229.

Campanelli, B., Fleurquin, P., Arranz, A., Etxebarria, I., Ciruelos, C., Eguluz, V.M., Ramasco, J., 2016. Comparing the modeling of delay propagation in the US and European air traffic networks. Journal of Air Transport Management 56.

Churchill, A., Lovell, D., Ball, M., 2010. Flight delay propagation impact on strategic air traffic flow management. Transp. Res. Rec. 2177.

Cook, A., Tanner, G., 2015. European airline delay cost reference values, updated and extended values, version 4.1. https://www.eurocontrol.int/publications/european-airline-delay-cost-reference-values.

EUROCONTROL, 2017. PJ19 Performance Framework, D4.1, October 2017.
URL https://ec.europa.eu/research/

European Commission, 2019. Commission Implementing Regulation (EU) 2019/317 of 11 February 2019 laying down a performance and charging scheme in the single European sky and repealing Implementing Regulations (EU) No 390/2013 and (EU) No 391/2013.
URL https://eur-lex.europa.eu/

Fleurquin, P., Ramasco, J. J., Eguiluz, V. M., jan 2013. Systemic delay propagation in the US airport network. Sci. Rep. 3, 1159.

Gurtner, G., Delgado, L., Cook, A., Martín, J., Cristóbal, S., 2018. A multi-agent model for long-term KPIs alignment forecast. In: SESAR Innovation Days 2018.

Kafle, N., Zou, B., 2016. Modeling flight delay propagation: A new analytical-econometric approach. Transportation Research Part B 93.

Performance Review Commission, 2020. Performance Review Report, an assessment of air traffic management in Europe during the calendar year 2019 (draft).
URL https://www.eurocontrol.int/publication/performance-review-report-prr-2019-consultation

Pyrgiotis, N., Malone, K., Odoni, A., 2013. Modelling delay propagation within an airport network. Transportation Research Part C 27.

SESAR, 2019. European ATM Master Plan - Executive View, Ed. 2020.
URL https://www.atmmasterplan.eu/

Welman, S., Williams, A., Hechtman, D., 2010. Calculating Delay Propagation Multipliersfor Cost-Benefit Analysis. Research Report prepared by the Center for Advanced Aviation System Development (MITRE), MP 10 0 039.
URL https://www.faa.gov/regulations_policies/policy_guidance/benefit_cost/media/faabca.pdf

Wong, J., Tsai, S., 2012. A survival model for flight delay propagation. Journal of Air Transport Management 23.

Wu, C.-L., Caves, R., 2004. Modelling and optimization of aircraft turnaround time at an airport. Transportation Planning and Technology 27.

Xu, N., Sherry, L., Laskey, K., 2008. Multi-factor model for predicting delays at U.S. airports. Transp. Res. Rec. 2052.