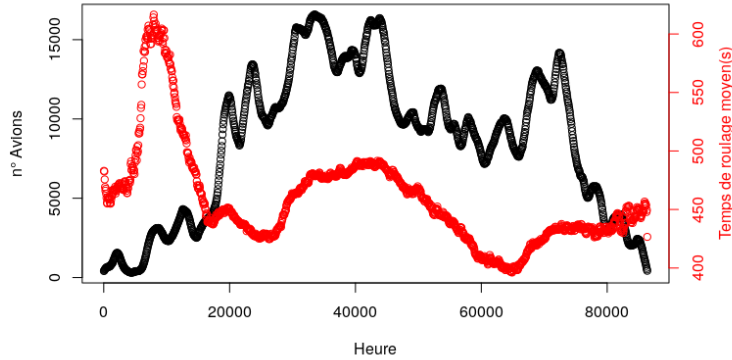


# Data Analysis Results

May 26, 2020

## 3 a.m. peak

If we plot the taxi time averaged over all the flights as a function of the hour of the day we see a peak that cannot be explained by the number of aircraft that are crossing the taxiway (Figure 1).

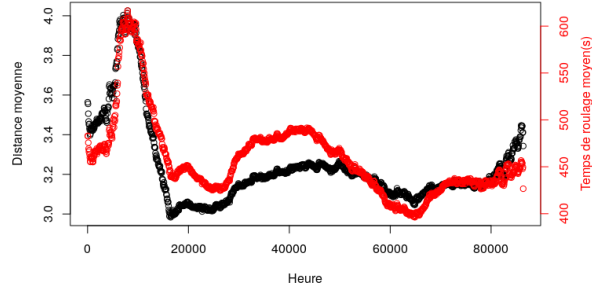


**Figure 1:** Comparison between number of aircrafts in the taxiway (black line) and taxi time (red line).

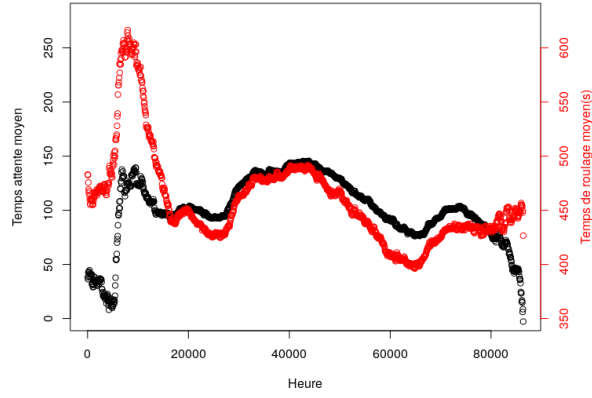
So we performed an analysis of all the variables that can contribute to the unexplained peak of 3 a.m.

We tested the distance ( $d$ ) between the parking and the runway, the average speed during the taxiing ( $v$ ), the waiting time ( $t_w$ ) of the potential stops of the aircraft during the taxiing (in the middle of the taxiway or at the runway). We chose those variables because we can roughly model the taxi time ( $t_t$ ) as:

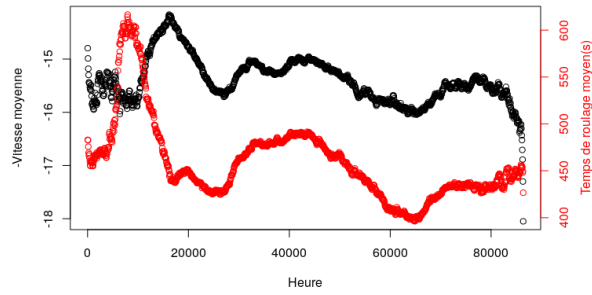
$$t_t = \frac{d}{v} + t_w$$



(a) Comparison between distance of the parking from the runway (black) and taxi time (red).



(b) Comparison between waiting time during the taxiing (black) and taxi time (red).



(c) Comparison between the opposite (since we expect the taxi time to be inversely proportional to the speed) of average speed during the taxiing (black) and taxi time (red).

**Figure 2:** Behavior of average taxi time as a function of the hour of the day compared to the behavior of parking-runway distance, waiting time and speed.

Figure 2 shows the results.

As we can see from the pictures, what explains better the 3 a.m. peak is the larger distance traveled by the aircrafts on average at that hour. Also the waiting time gives a little contribution, while the speed is on average larger in correspondence of the peak. This means that it gives a negative contribution to the taxi time, with respect to the other hours of the day.

We can go deeper in the analysis of the reasons of a larger taxi time if we hypothesize the dependencies of waiting time, speed and distance from some other variables that we can find in our data set.

Our first hypothesis about dependencies is the following:

$$d = d(\text{gate}, \text{company}, \text{number of aircrafts on platform}, \text{weather})$$

$$v = v(\text{aircraft type}, \text{weather})$$

$$t_w = t_w(\text{number of aircrafts on platform})$$

In fact:

- *Distance*: it depends on the company because each company is assigned a terminal or a parking zone, that can be further or closer to the runway (that's why it depends on the gate also). It depends on the number of aircrafts because in case of intense traffic an aircraft parked in the north terminal can leave from the southern runway, if there are less aircrafts waiting there. It depends on the weather because if de-icing procedures are needed, the aircraft necessarily has to pass through de-icing zones.
- *Speed*: it depends on the aircraft type because heavier aircrafts are slower on the ground. It depends on the weather because in case of low visibility conditions the pilots must drive slower.
- *Waiting time*: it depends on the number of aircrafts in the airport because if there are many aircrafts, the probability of stopping to make another aircraft pass is higher.

For what it concerns the 3 a.m. peak, we looked into the distance dependencies since it's the most influential variable for the taxi time in that time interval.

We explored the company variable, comparing the frequencies of flights of a specific company at all hours of the day with the ones in the interval from 1:35 a.m. to 3:40 a.m.

**Table 1:** Average company frequencies for the whole day (left) and for the time interval 1:35 a.m. - 3.40 a.m. (right)

<b>Company</b>	<b>Frequency</b>	<b>Company</b>	<b>Frequency</b>
AFR	48,3%	FDX	26,8%
EJU	6,8%	AFR	26,4%
FDX	2.25%	ABR	20,3%

We see that in the time interval of interest there are many flights from FedEx and ASL Airlines Ireland (ABR) (that provides for aircrafts to FedEx), in opposition to what happens during the rest of the day, where the flights are mainly from Air France and Easy Jet.