Improving the Schiphol wind forecast

By Edwin Kok

24/12/2021

Internship for LVNL — Supervisor: Ferdinand Dijkstra

With guidance from KNMI — Supervisor: Dirk Wolters

As part of MEE at WUR — Supervisor: Jordi Vila

Introduction

The weather highly determines Schiphol operations. The choice of which runway(s) to use is mainly determined by wind speed and direction. Each runway has a capacity of number of aircrafts being able to take off and land per hour, which is highly influenced by a set of important meteorological variables. The most important ones are visibility, cloud ceiling, wind, and the occurrence of severe weather. For proper planning it is vital that these are forecasted as accurately as possible. If they are forecasted too 'negative' capacity is left unused, and thus fewer income is realized. If they are forecasted too 'positive' the impact is even greater, as this will result in crowded holding/taxiing areas and delays. This report will focus on testing the performance of wind speed forecasts. The previous intern (Soraya van Beek) has investigated low visibility phases.

Meteorological operations

KNMI issues a separate weather forecast for Schiphol airport (runway 27), called the 'Schiphol Kans Verwachting' (SKV). The steps to arrive at this forecast are explained in short below.

- The first step is to use model output from operational NWP. Currently HIRLAM is used for this, but they are in the process of transferring to HARMONIE.
- Numerical Weather Prediction (NWP) output is improved based on a statistical algorithm to produce a TAF Guidance product. TAF stands for Terminal Aerodrome Forecast and is a uniform way to interpret the most relevant weather variables in aviation.
- The most relevant variables are grouped in the SKV. Adjustments and remarks are made by the aviation meteorologist. Runway capacities based on forecasted weather are included.

See Figure 1 for a visualization of the steps involved in making the SKV.

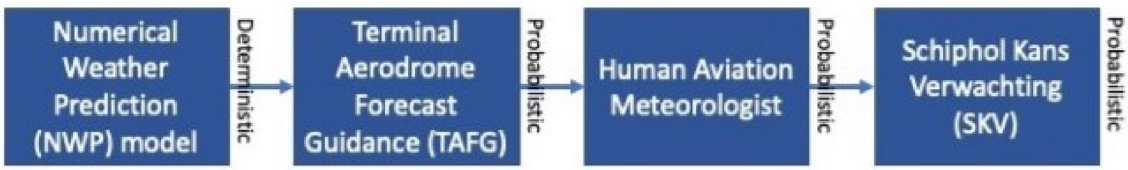


Figure 1: Schematic of the steps involved in making the SKV. From Soraya's report on low visibility phases.

Goal/research questions

The ultimate goal is to improve the SKV. In order to achieve this the performance of current products and methods needs to be evaluated. This report aims to accomplish this for wind speed.

Research questions

- How do different model inputs for the TAF Guidance product perform when evaluating surface winds, and how does the SKV compare?
- What is the effect of forecast lead time?

Methods

The data used in this report consist of:

- HARMONIE TAFG, ranging from 2014/12/30 to 2017/12/31.
- HIRLAM TAFG, ranging from 2015/01/01 to 2017/12/31.
- ECMWF TAFG, ranging from 2015/04/03 to 2017/12/31.
- SKV, ranging from 2015/01/01 to 2017/09/18.
- METAR observations every 30 minutes over the entire period.

Most of the analysis was done for forecast lead times of 3 and 6 hours, with some additional statistics for 9, 12, and 24 hours. All data, excluding observations, have been filtered first for these forecast times. Hereafter, they have been merged on the time for which the forecast was valid, along with observations for the corresponding time.

The BIAS has been calculated according to:

$$BIAS = Forecast - Observations$$

Additional filters have been applied for speed ranges of:

- 15-19 kts
- 20-24 kts
- 25-29 kts
- 30-34 kts
- 35+ kts

With these speed filters, histograms of the error distribution were made. Additionally, bar plots were made that show the percentage of time in which the error exceeds |5| kts per season of the year.

The mean observed wind speed for each day of the year was subtracted from the mean forecasted wind speed for each day of the year. 30-Day moving averages of these values have been plotted alongside 30-day moving averages of the RMSE per day, according to:

$$RMSE = \sqrt{\sum_{i=1}^{n} \frac{(Forecast - Observations)^2}{n}}$$

Lastly the error as a function of time of day has been plotted.

Results

As a first indication of product performance the BIAS is shown in Table 1 for several forecast lead times. Although the order of magnitude is small, a few things can be noted. For the 3-hour forecast HARMONIE performs best, for all other forecast times HIRLAM has the lowest BIAS. The 3-hour forecast for HIRLAM differs significantly in that the absolute value (0.19 kts) is comparatively high, and in that it is the only negative BIAS (underestimation of wind speed) from all products and lead times. Overall, ECMWF has the worst performance when judging by the BIAS. SKV seems to improve the 3-hour HIRLAM forecast, while the 6-hour forecast is a deterioration.

Table 1: BIAS (forecast - observations) in wind speed (kts) over the entire time period for the HARMONIE, HIRLAM, and ECMWF TAFG, and the SKV.

	HARMONIE	HIRLAM	ECMWF	SKV
3-hour forecast	0.03	-0.19	0.12	0.06
6-hour forecast	0.09	0.00	0.24	0.17
9-hour forecast	0.09	0.02	0.22	-
12-hour forecast	0.10	0.08	0.25	-
24-hour forecast	0.09	0.07	0.30	-

Operations are impacted to a greater extent when wind speeds are high. Therefore, the data were filtered for higher wind speeds, starting at 15 kts, and the error distribution was examined. Since the sample size becomes small, 5 kts intervals were determined. Figure 2 shows one example of such an error distribution for a wind speed range of 15-19 knots. Assuming a standard distribution there is a 95.45 % chance the actual wind speed will be within the range of forecasted speed ± mean ± 2 *

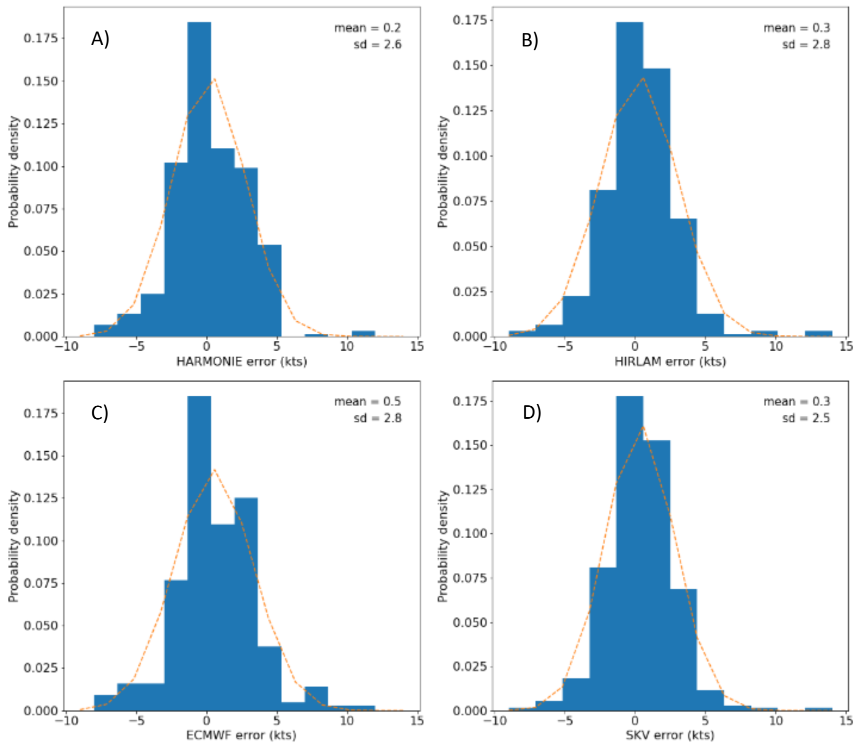


Figure 2: Histograms of error (forecast - observations) distribution (blue bars) for forecasted wind speeds of 15-19 knots for a lead time of 3 hours. The orange dotted line shows the fictitious standard distribution based on the mean and standard deviation. A) HARMONIE, B) HIRLAM, and C) ECMWF TAFG, D) SKV based on the HIRLAM TAFG.

standard deviation (sd) (indicated in the figure). For example, if the HARMONIE TAFG forecasts a wind speed of between 17 knots, there is a 95.45 % probability that the actual wind speed will be within 12 and 22.4 kts (17.2 ± 5.2). See Appendix A for more wind speed ranges and forecast lead times. Especially for higher wind speeds, take care in relying on this information. Sample sizes are small, meaning the assumption of a standard distribution does not hold and the effect of outliers may be significant.

Figures 3 and 4 show similar information as Figure 2, visualized differently. The bottom part shows the percentage of time in which there is an exceedance of an error greater than|5|kts based on season for 15-19 kts, and 20-24 kts, in Figure 3 and 4 respectively. Both Figures are for a forecast lead time of 3 hours. In Appendix B more wind speed ranges and forecast times are shown. Remarkable is that for the range of 15-19 kts HIRLAM performs exceptionally well in the summer, with 0 cases of the error being greater than|5| kts. In winter all products overestimate wind speed more than they underestimate it. For the range of 20-24 kts HARMONIE outperforms all other products in winter and spring. All products underestimate the wind speed in summer for the 20-24 kts speed range. The SKV generally (with few exceptions) performs slightly better than HIRLAM TAFG, as seen by the lower red bars compared to the orange bars in both Figures. The reduction of sample size with increasing wind speeds can be seen by comparing the top parts of Figures 3 and 4.

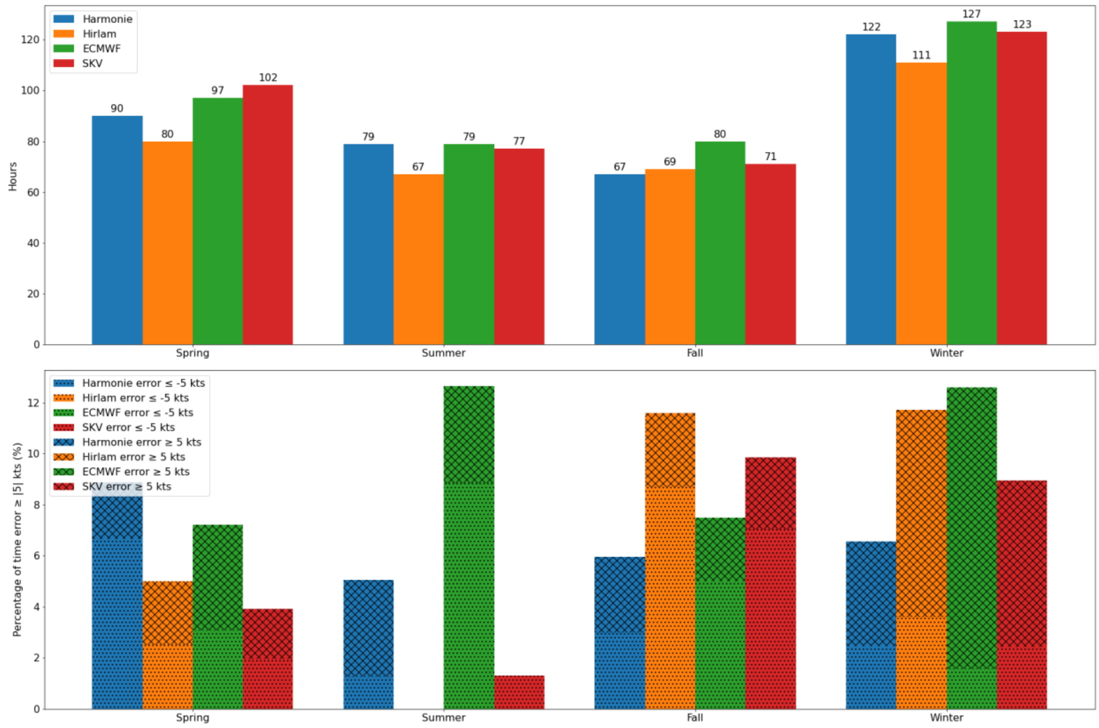


Figure 3: Top: number of hours that a wind speed of 15-19 kts is forecasted per season for HARMONIE, HIRLAM, and ECMWF TAFG, and SKV with a lead time of 3 hours. Bottom: percentage of time that a forecast is more than |5| kts difference from the measured wind speed. In blue HARMONIE, orange HIRLAM, and green ECMWF TAFG, and in red SKV. Dotted patterns show an underestimation of wind speed, crossed patterns an overestimation.

Figures 5 and 6 show the mean error (top) and RMSE (bottom) per day of year for a forecast lead time of 3 hours and 6 hours, respectively. Analogous to Figure 2, the reliability of forecasts could be determined, if one assumes RMSE to be (nearly) the same as sd, and with the difference that no speed filter is applied. As with Figure 2, the assumption of a standard distribution is debatable. For both the 3-hour and 6-hour forecast wind speed is generally overestimated in winter and underestimated in (late) summer. The 3-hour HARMONIE TAFG forecast shows a large BIAS in fall, with a similar shape (less extreme) apparent in the ECMWF TAFG and SKV. Generally, adjustments made from HIRLAM to achieve the SKV are good for the 3-hour forecast, with the exception of April and the winter months. SKV performance is worse for the 6-hour forecast compared to HIRLAM.

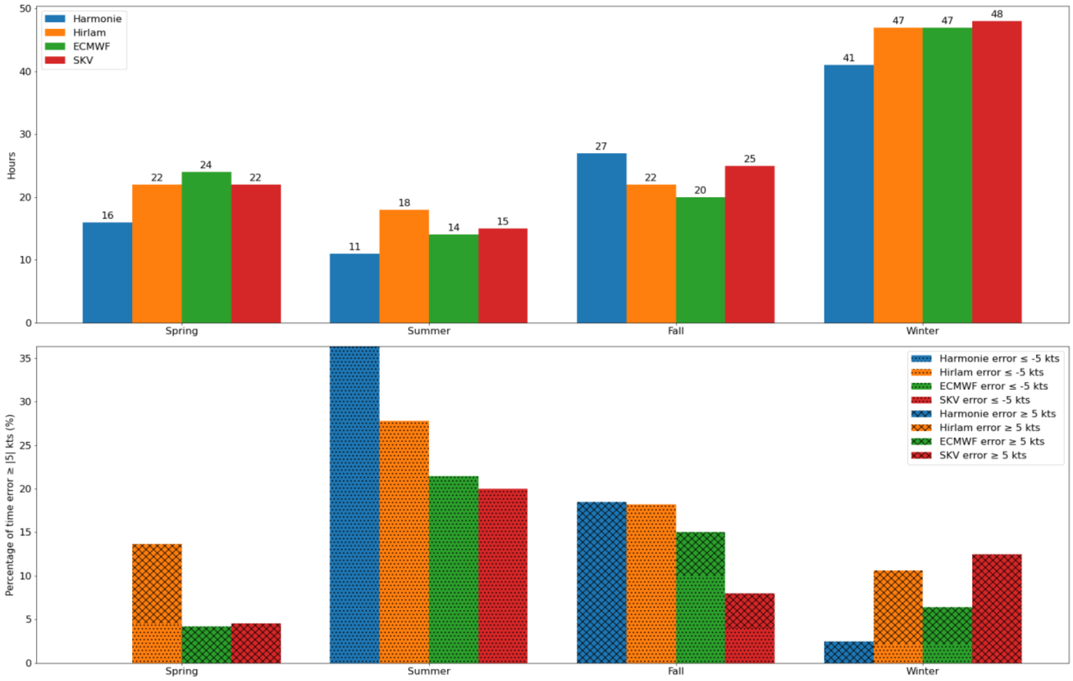


Figure 4: Top: number of hours that a wind speed of 20-24 kts is forecasted per season for HARMONIE, HIRLAM, and ECMWF TAFG, and SKV with a lead time of 3 hours. Bottom: percentage of time that a forecast is more than |5| kts difference from the measured wind speed. In blue HARMONIE, orange HIRLAM, and green ECMWF TAFG, and in red SKV. Dotted patterns show an underestimation of wind speed, crossed patterns an overestimation.

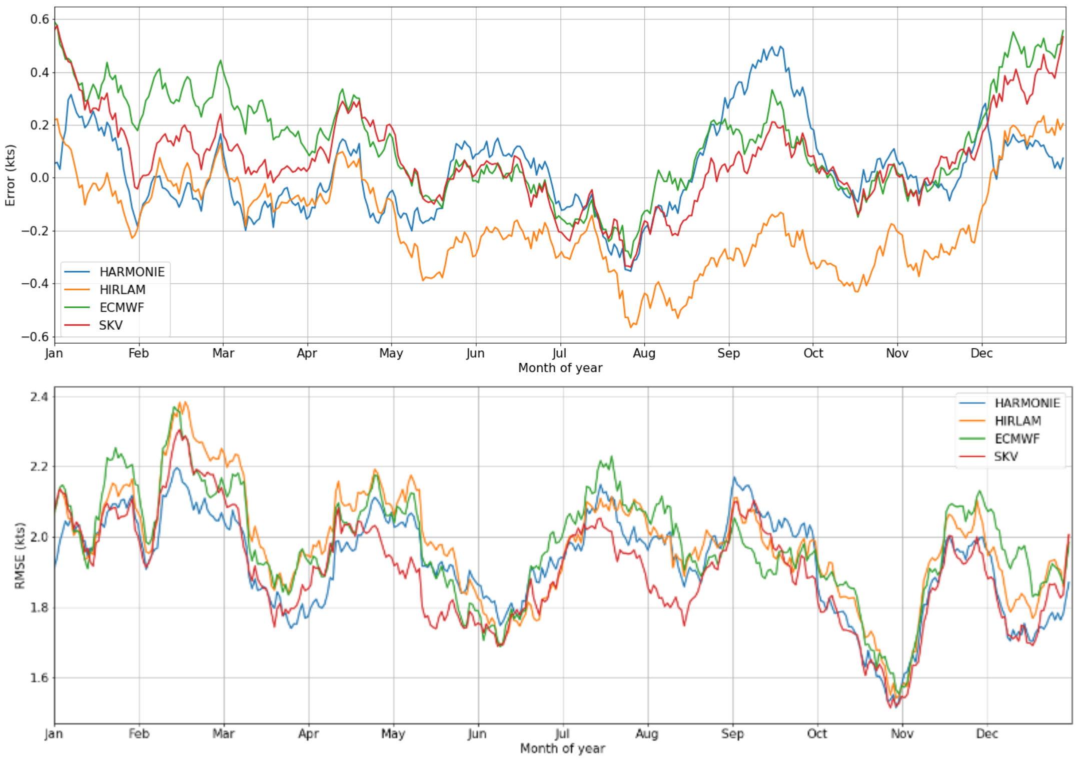


Figure 5: Top: error (forecast - observations) of the daily averaged wind speed (kts) for the 3-hour forecasts. Bottom: the RMSE (kts) corresponding to the top Figure. In blue HARMONIE, orange HIRLAM, and green ECMWF TAFG, and in red SKV. All lines are a 30-day moving average of the actual vales.

Figure 6: Top: error (forecast - observations) of the daily averaged wind speed (kts) for the 6-hour forecasts. Bottom: the RMSE (kts) corresponding to the top Figure. In blue HARMONIE, orange HIRLAM, and green ECMWF TAFG, and in red SKV. All lines are a 30-day moving average of the actual vales.

Lastly, I investigated the effect of time of the day where the forecast was valid (Figure 7). From the left Figure it is noted that there is a tendency to overestimate (or underestimate less) wind speed during the day for the 3-hour forecast for all products. For the 6-hour forecast, this pattern only holds true for the SKV.

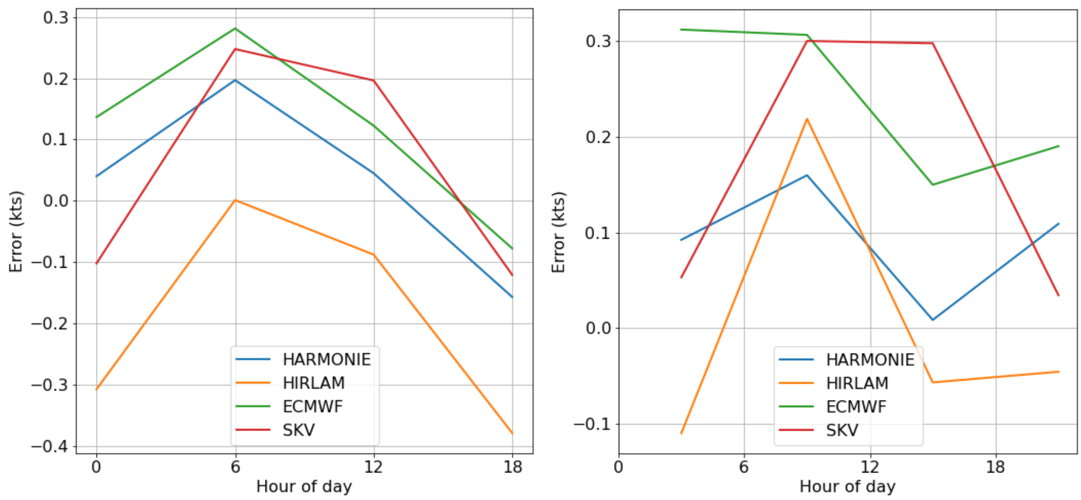


Figure 7: Error (forecast - observations) of the forecasted wind speed (kts) as a function of time of the day. Left: 3-hour forecast. Right: 6-hour forecast. In blue HARMONIE, orange HIRLAM, and green ECMWF TAFG, and in red SKV.

Discussion

All attempts to provide reliabilities on forecasts in this report are based on the entire data set of roughly 2.5 years, without attention to specific weather conditions. Although it is clear that forecast quality decreases with higher wind speeds, it remains unclear under what conditions specifically this is the case. For example, the highest errors could occur due to frontal passage, formation of isolated showers, sea breeze circulation, etc. It could also be that a certain phenomenon's magnitude is forecasted correctly, but that the timing is off. In this case closely monitoring of the actual situation and adjusting the forecast in response might drastically improve the quality. Unfortunately, I did not have access to the right data, nor did I have sufficient time, to pursue the physics as to why results are the way they are. Without this knowledge (or TAFG physics/algorithms), it is not possible to recommend improvements for the SKV.

Especially at the start of my internship there was a lack of a clear, well defined, specific goal. For LVNL the ultimate goal is clear, improve the weather forecasts for Schiphol. But for an intern with no experience and no access to the relevant processes involved in making the forecasts this is a difficult task. After a few weeks I received additional guidance from Dirk Wolters (KNMI), from which point onwards progress was being made. Getting the right data, and thereafter reading it into data frames, also took at least a month. Furthermore, it is my understanding that KNMI is working on improving forecasts for Schiphol, but that the specifics are not necessarily being clearly communicated to LVNL (or even internally to some extent).

I recommend for LVNL and KNMI to start discussing specific goals more intensively, so that every student (or other person involved) does not have to invest valuable time in rediscovering what needs to be done. I also believe there are great benefits in showing students early on in detail every step involved in making the Schiphol forecast. Furthermore, I suggest keeping a well-organized database of relevant meteorological data, and scripts to interpret them. I hope that this database can grow, starting with the work provided by me.

Conclusions and recommendations

For the 3-hour forecast, the HARMONIE TAFG generally shows the lowest BIAS and the highest reliabilities. Therefore, KNMI's decision to switch to this forecast as opposed to HIRLAM TAFG for the SKV seems justified. It should be noted that HARMONIE TAFG tends to overestimate wind speeds in the fall. Moreover, forecast reliabilities are worse than those for the HIRLAM TAFG in summer (Figures 3 and 4). Generally, forecast quality decreases slightly with increasing lead time, with the exception of HIRLAM TAFG, which performs best at a 6-hour lead time.

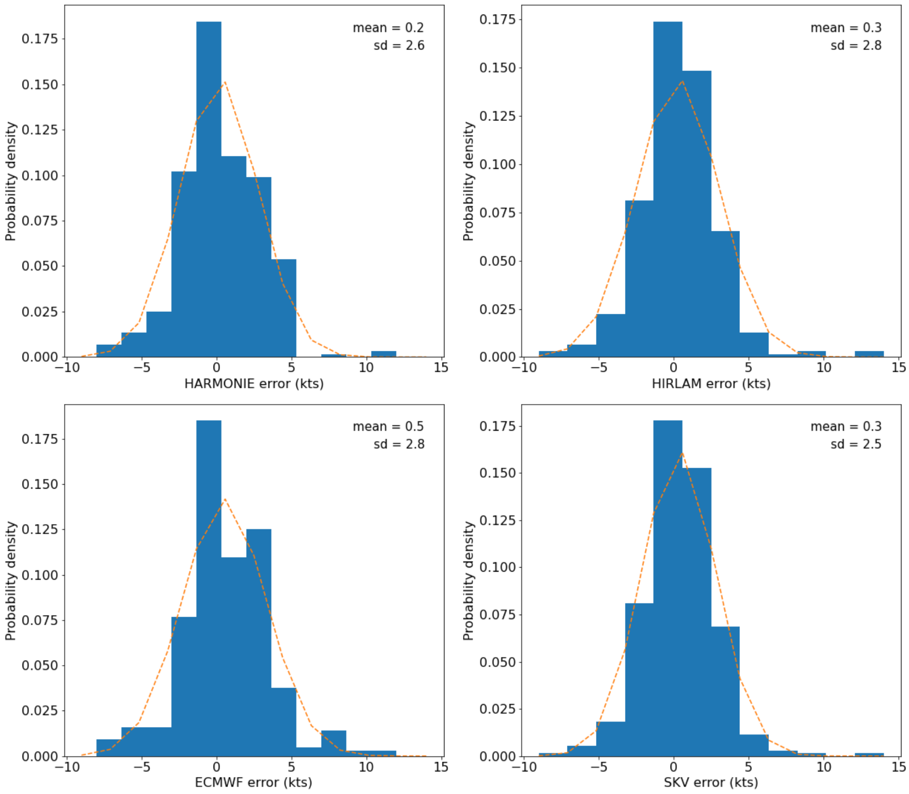
A logical next step would be to continue the reasoning from this report, extending it to include wind direction and gusts. Combining those results with Appendices A and B could provide LVNL with valuable information regarding reliabilities of the wind forecasts.

With the availability of historical SKV data, Soraya's work on TAFG visibility performance can also be extended to include SKV performance.

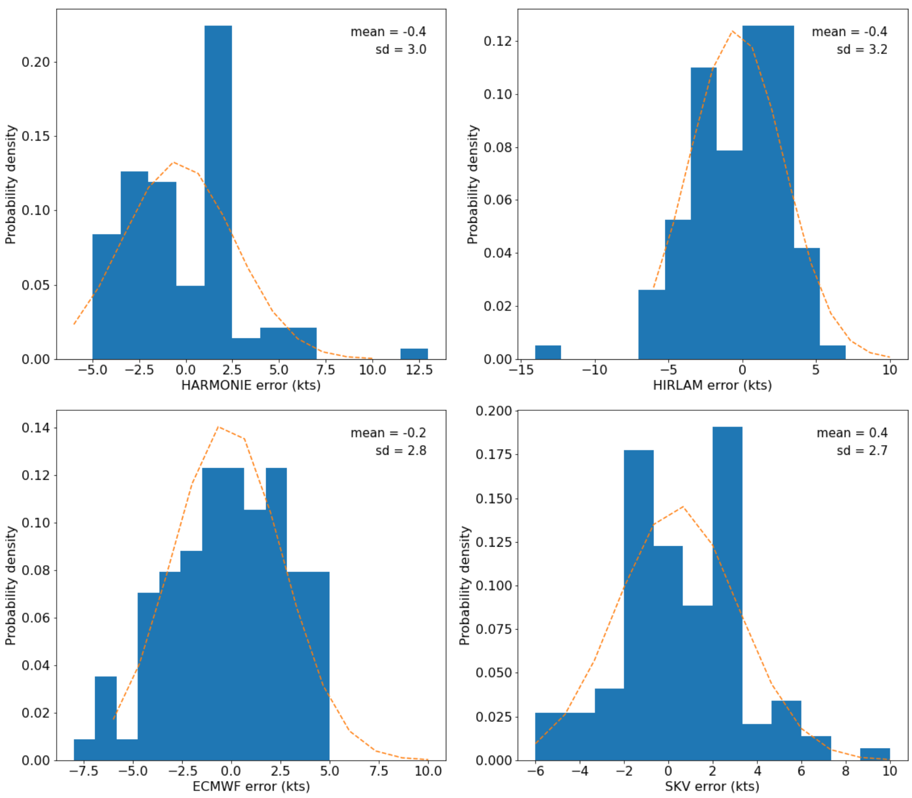
Currently, the SKV is made for the entirety of Schiphol based on RWY 27. With the high local variability of wind, the coarse resolution of NWP models is not ideal. Therefore, I also recommend being more involved in incorporating Large Eddy Simulation (LES) models for Schiphol.

Appendix A: Histograms of error distribution

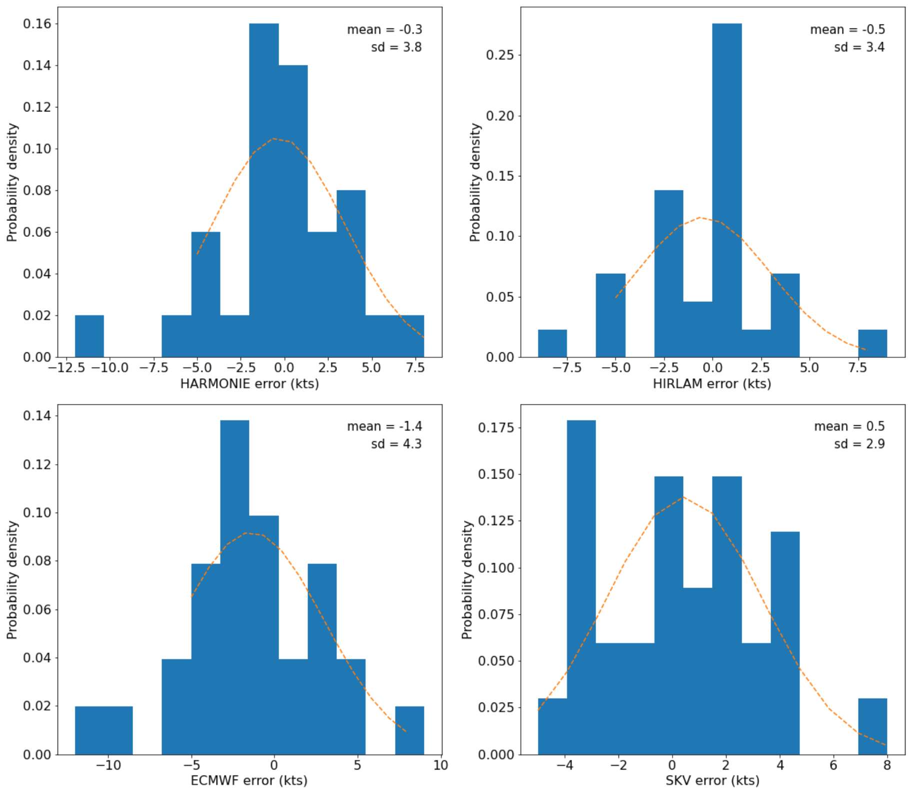
3-hour forecast, 15-19 kts



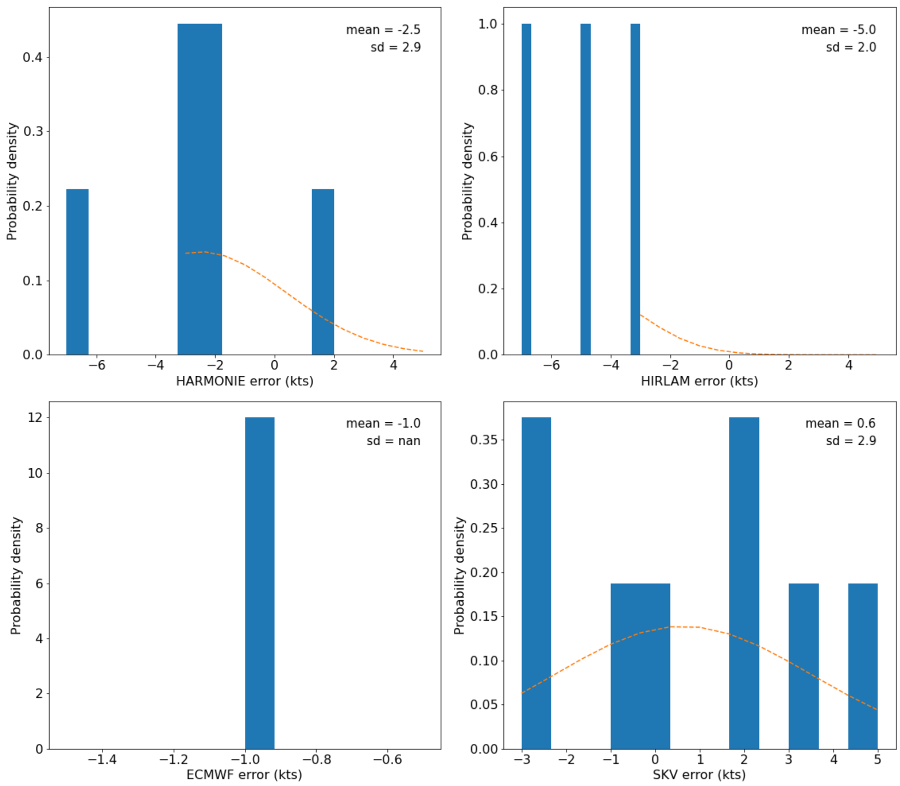
3-hour forecast, 20-24 kts



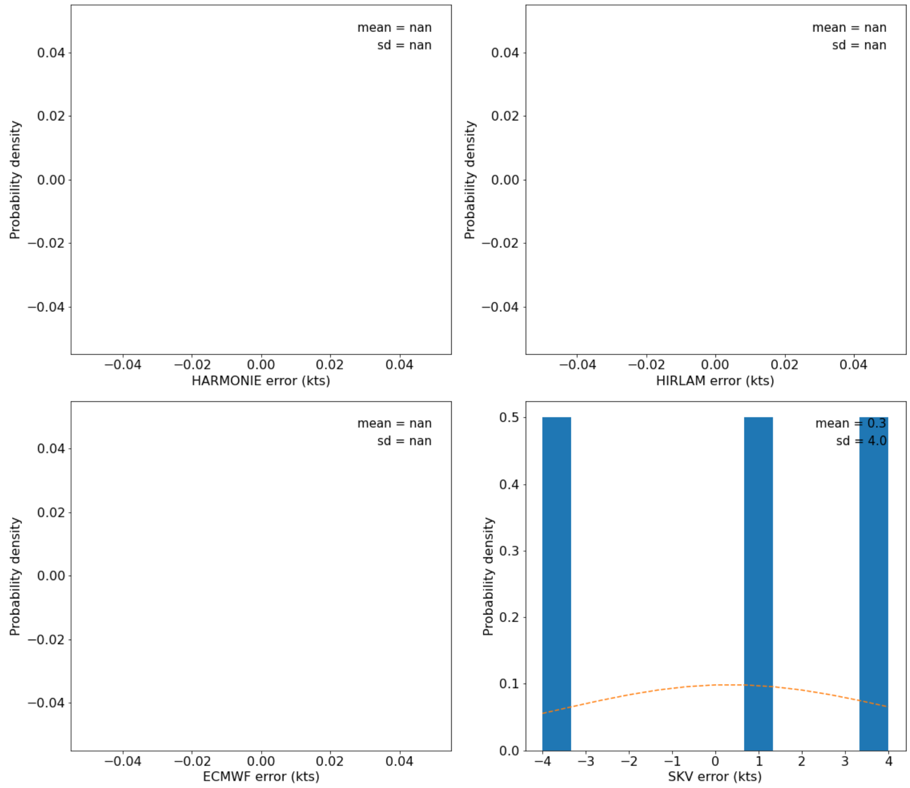
3-hour forecast, 25-29 kts



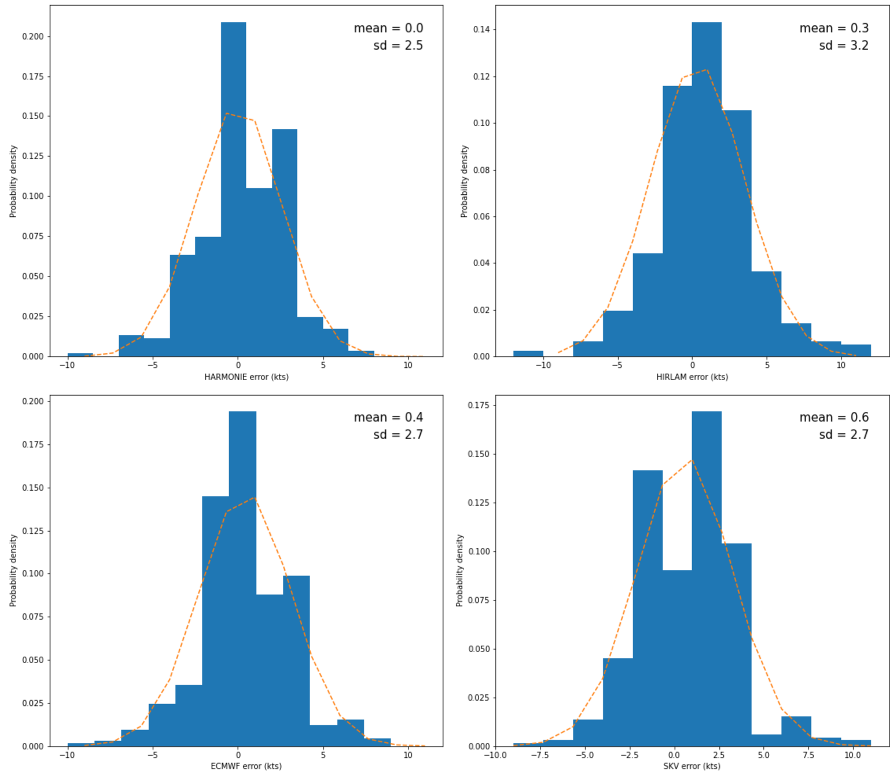
3-hour forecast, 30-34 kts



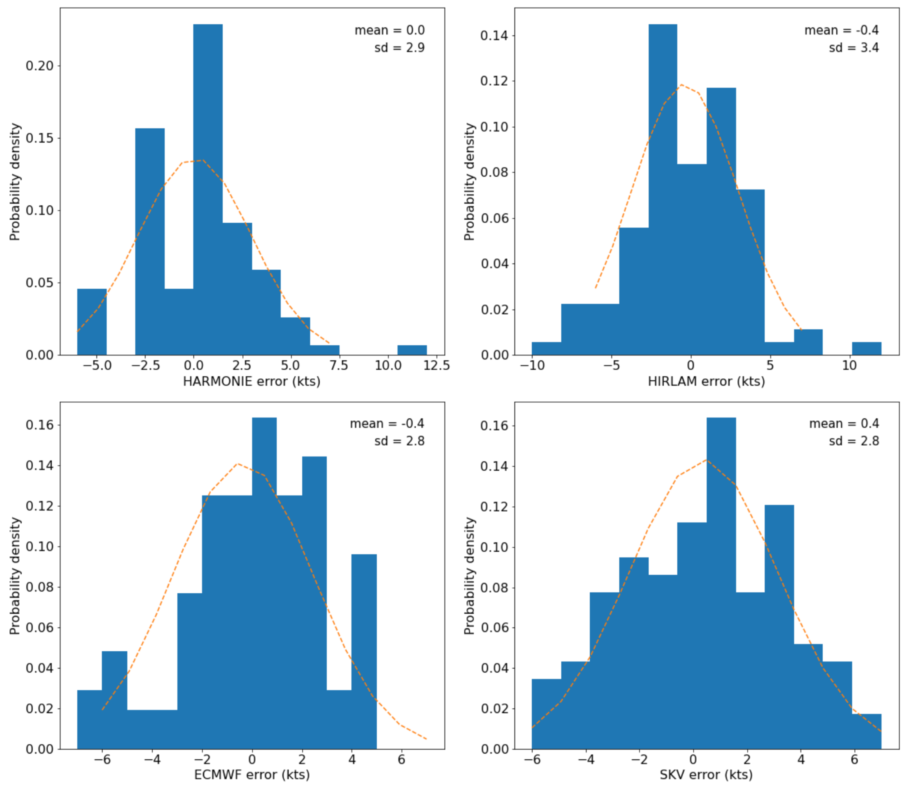
3-hour forecast, 35+ kts



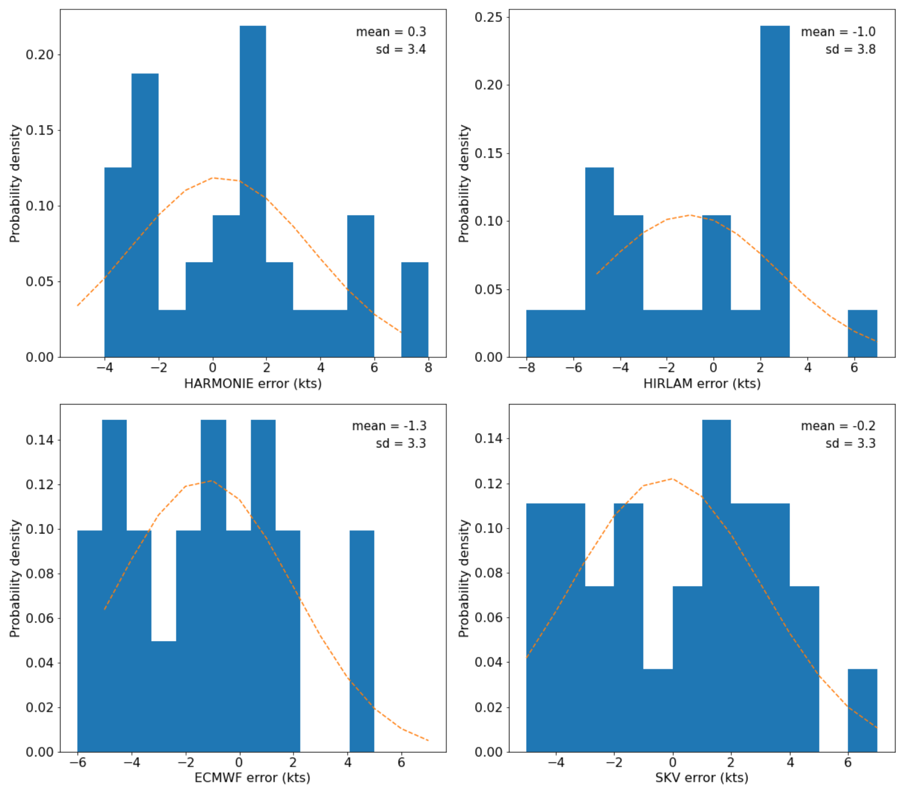
6-hour forecast, 15-19 kts



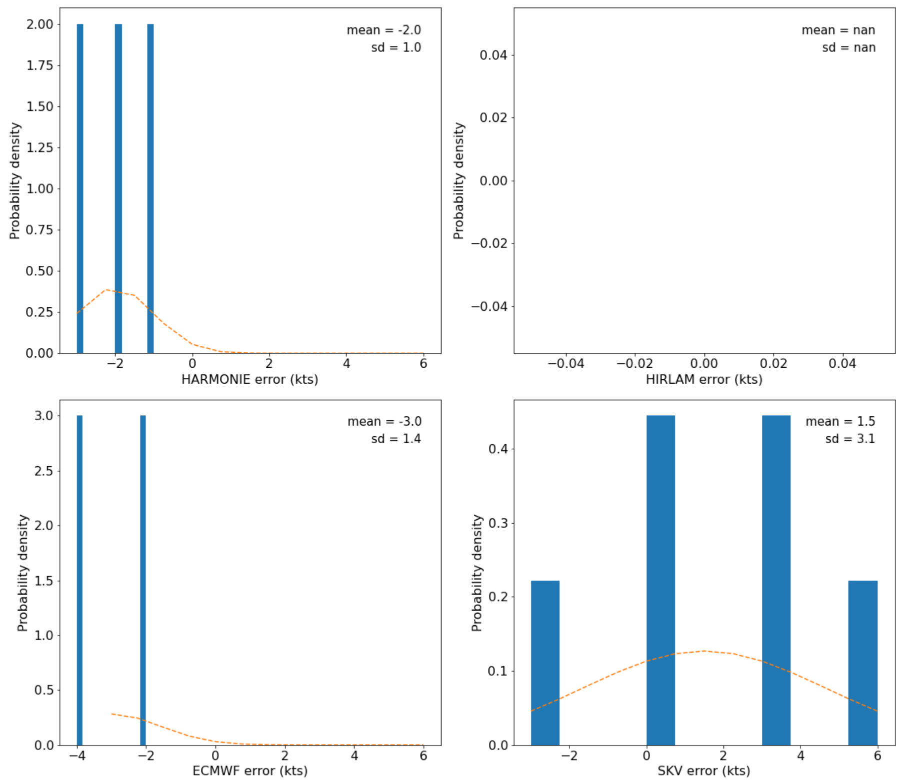
6-hour forecast, 20-24 kts



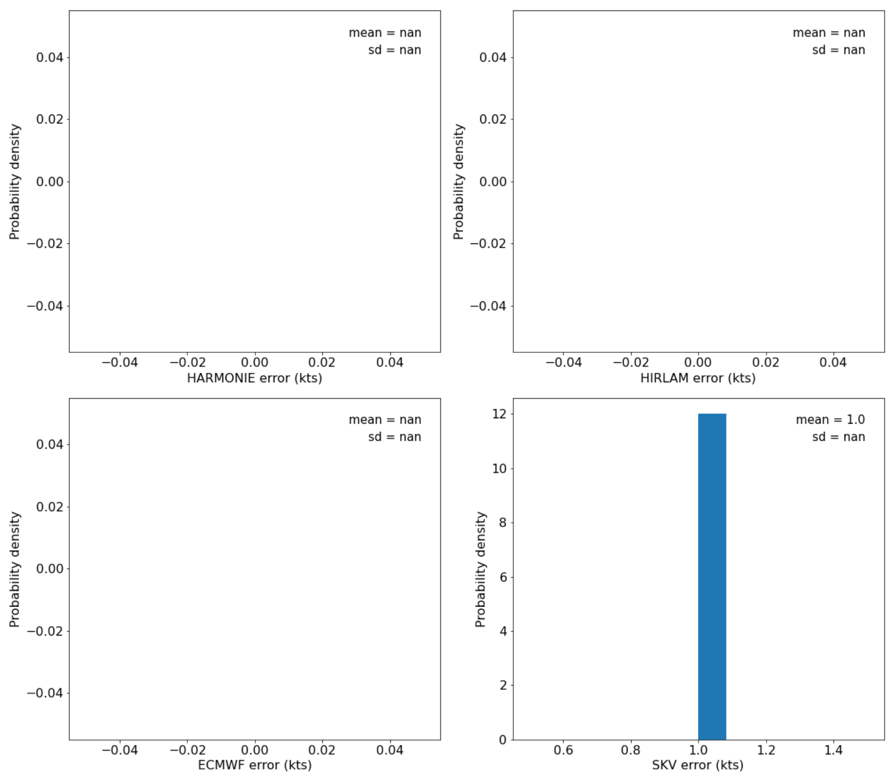
6-hour forecast, 25-29 kts



6-hour forecast, 30-34 kts

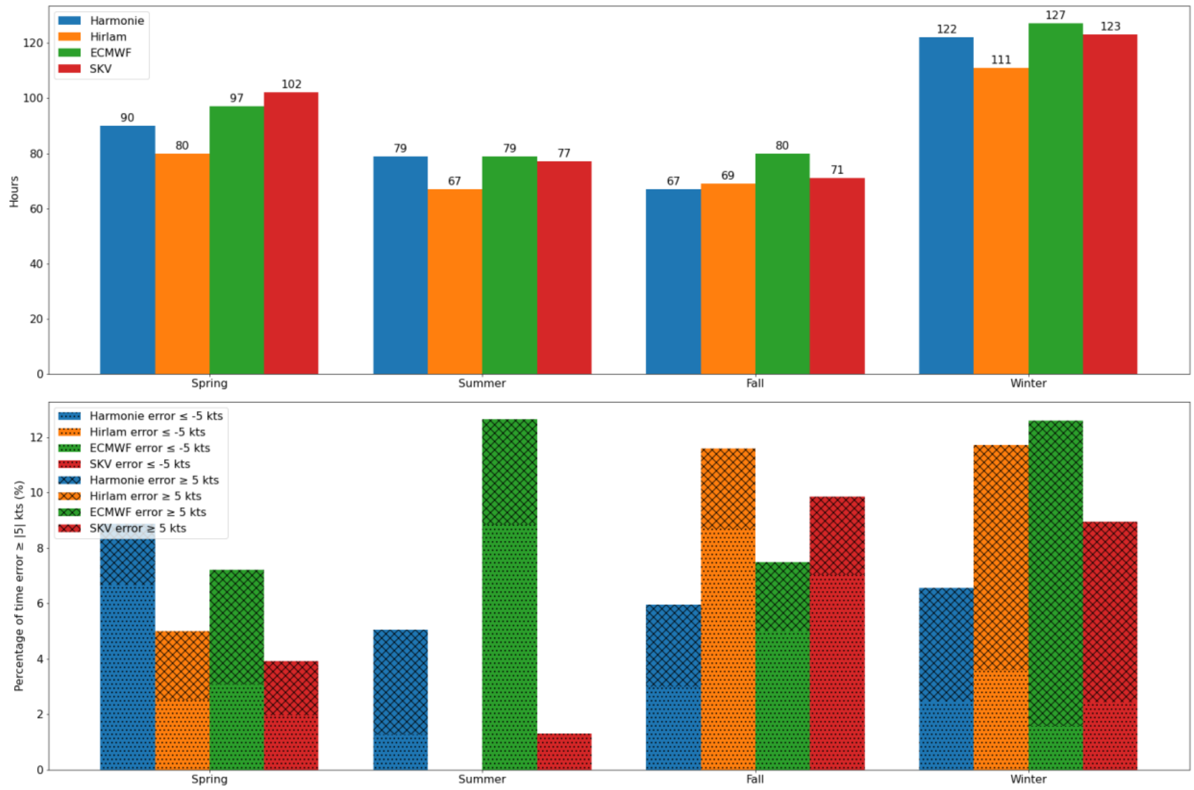


6-hour forecast, 35+ kts

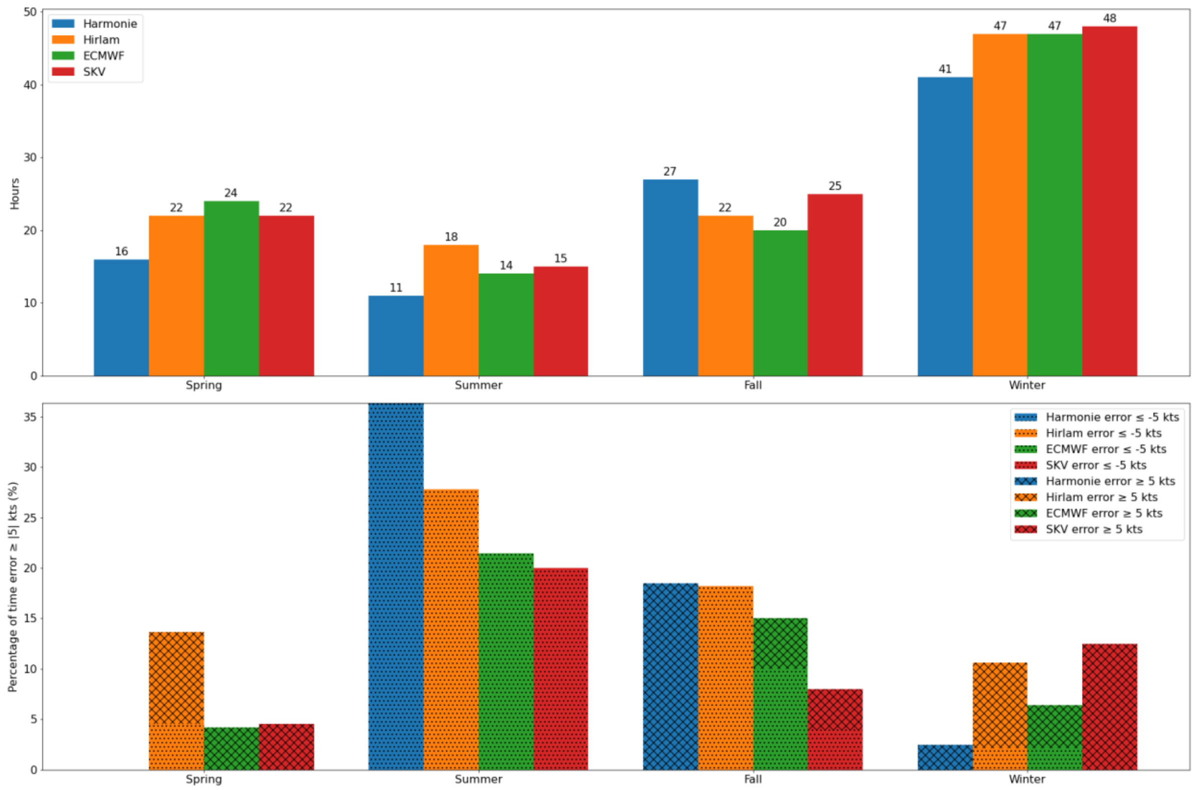


Appendix B: Bar plots of error exceedance ≥ |5|

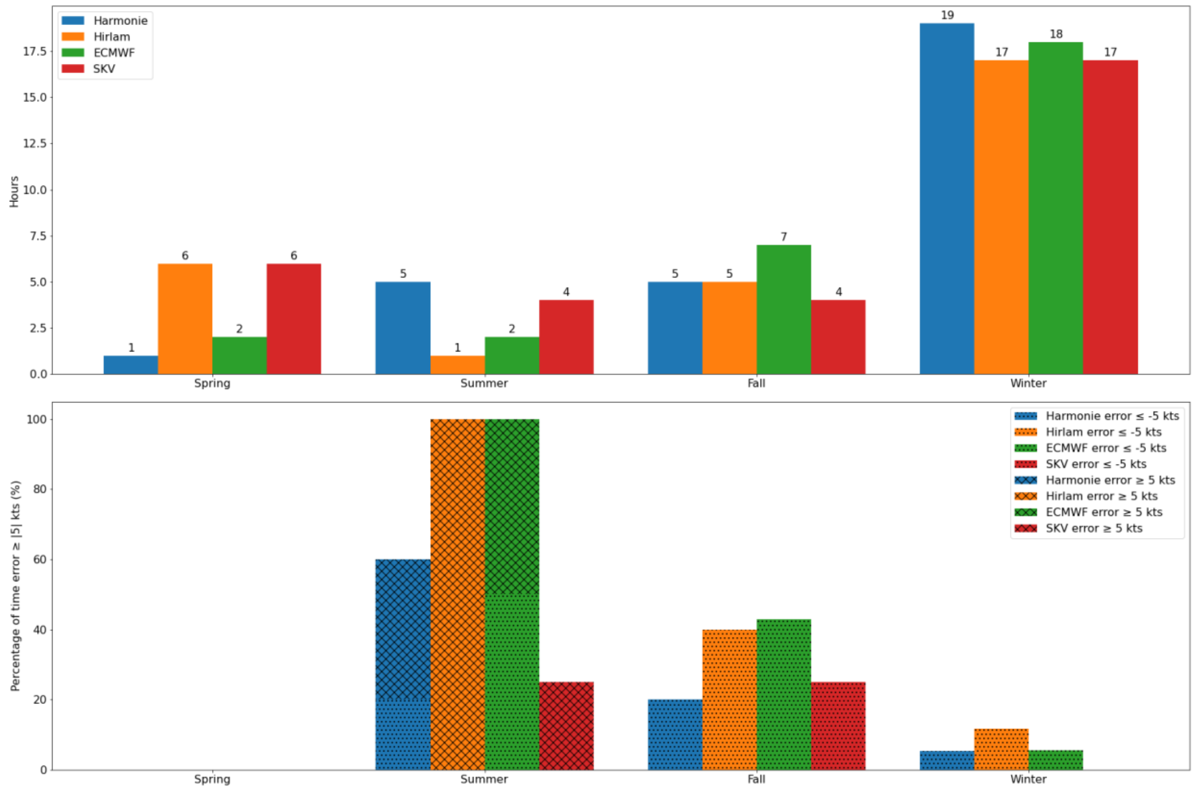
3-hour forecast, 15-19 kts



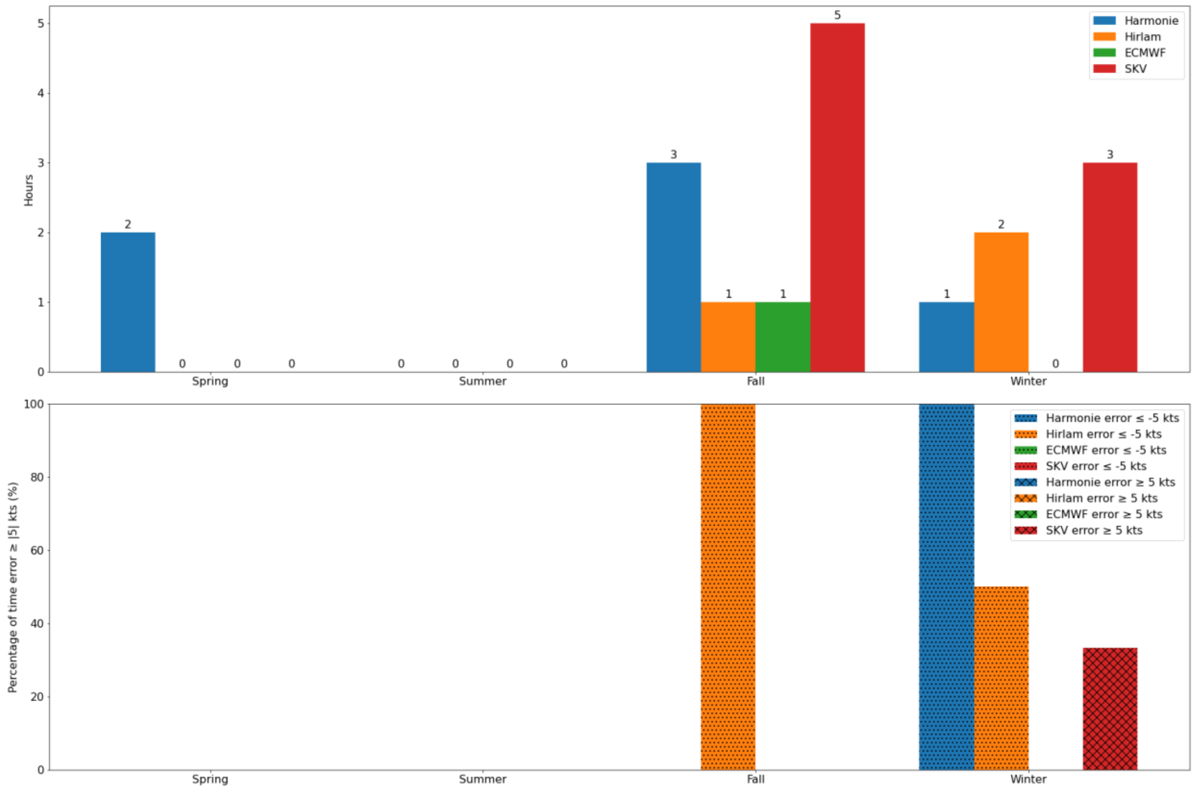
3-hour forecast, 20-24 kts



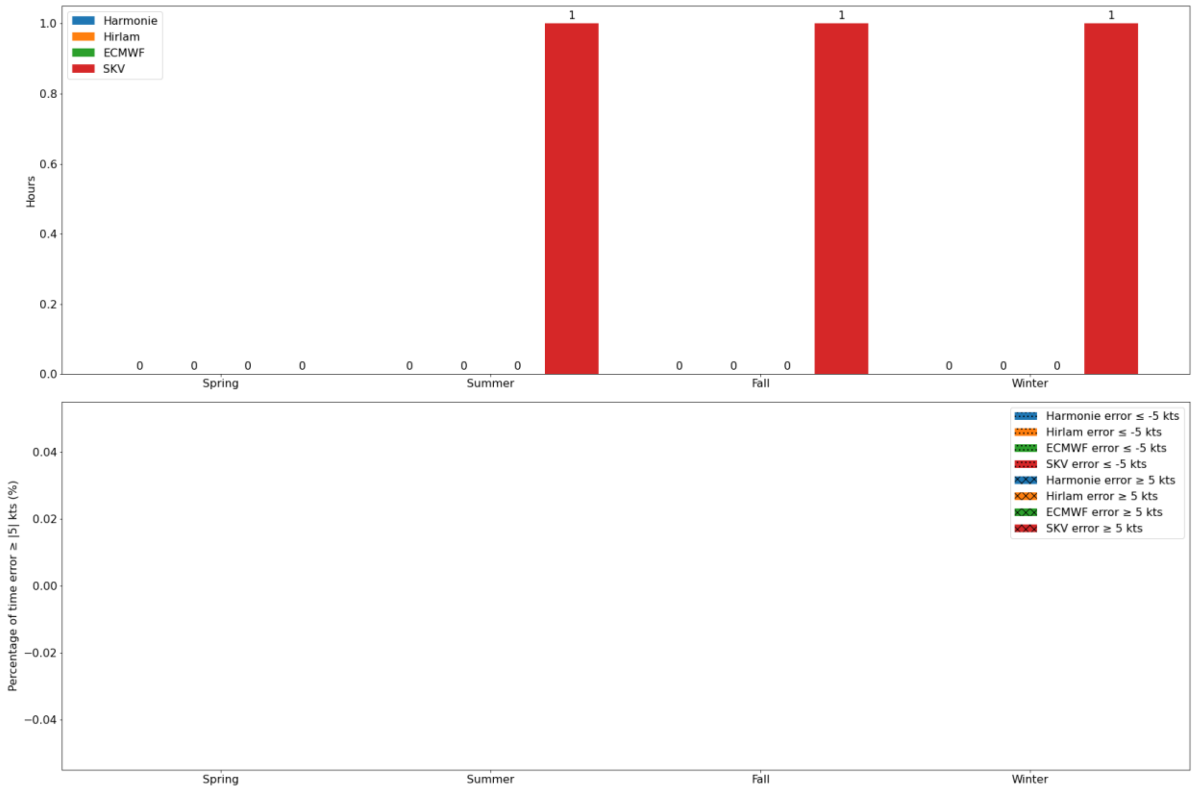
3-hour forecast, 25-29 kts



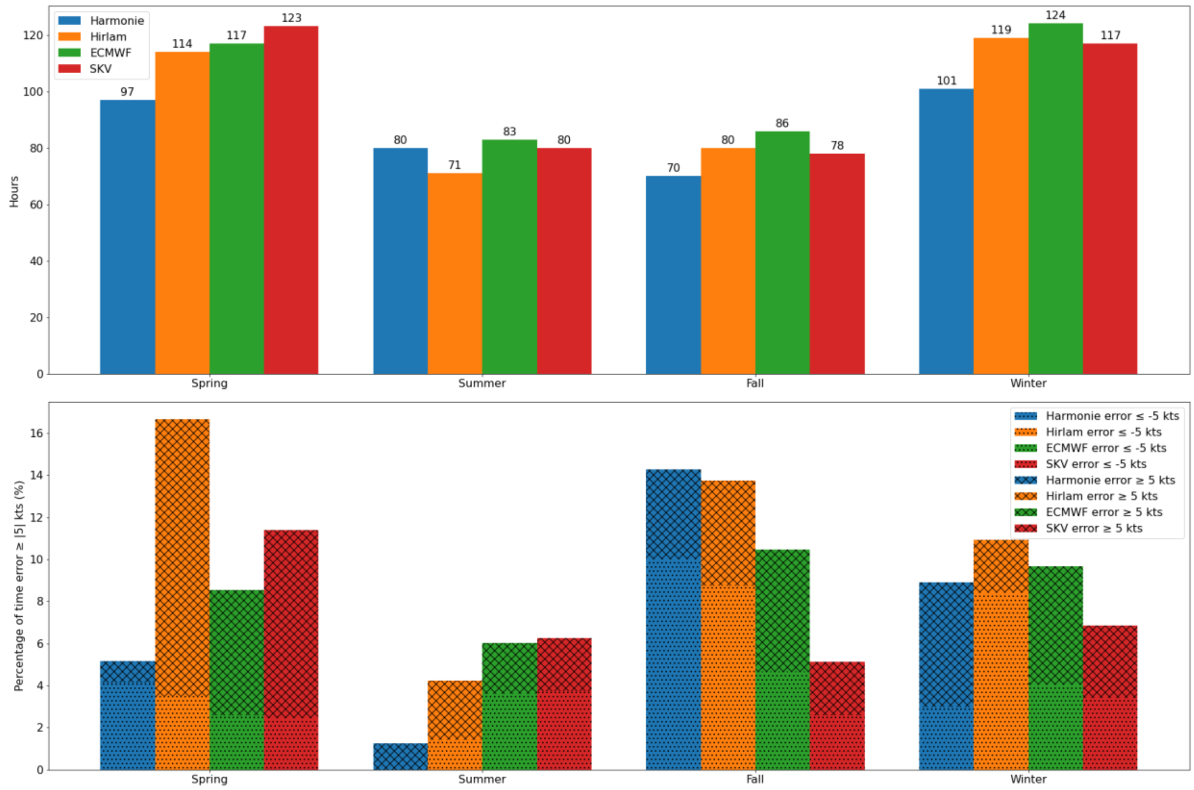
3-hour forecast, 30-34 kts



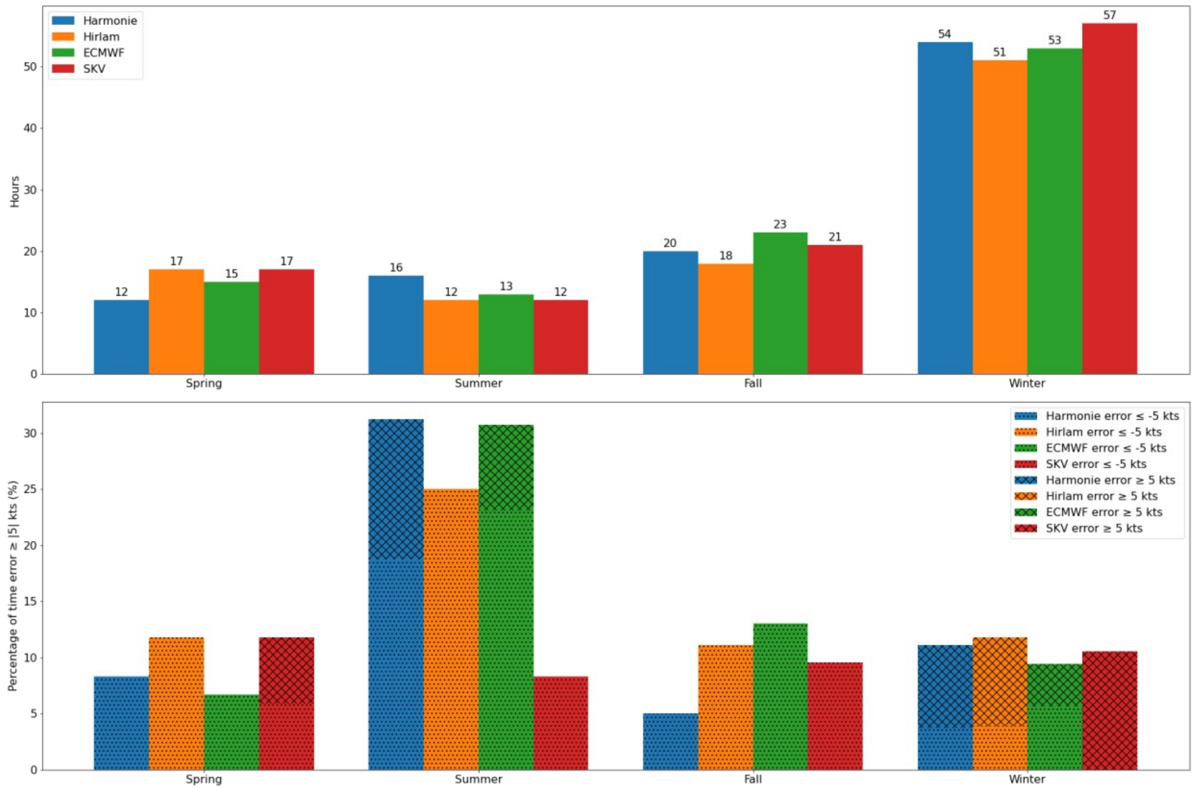
3-hour forecast, 35+ kts



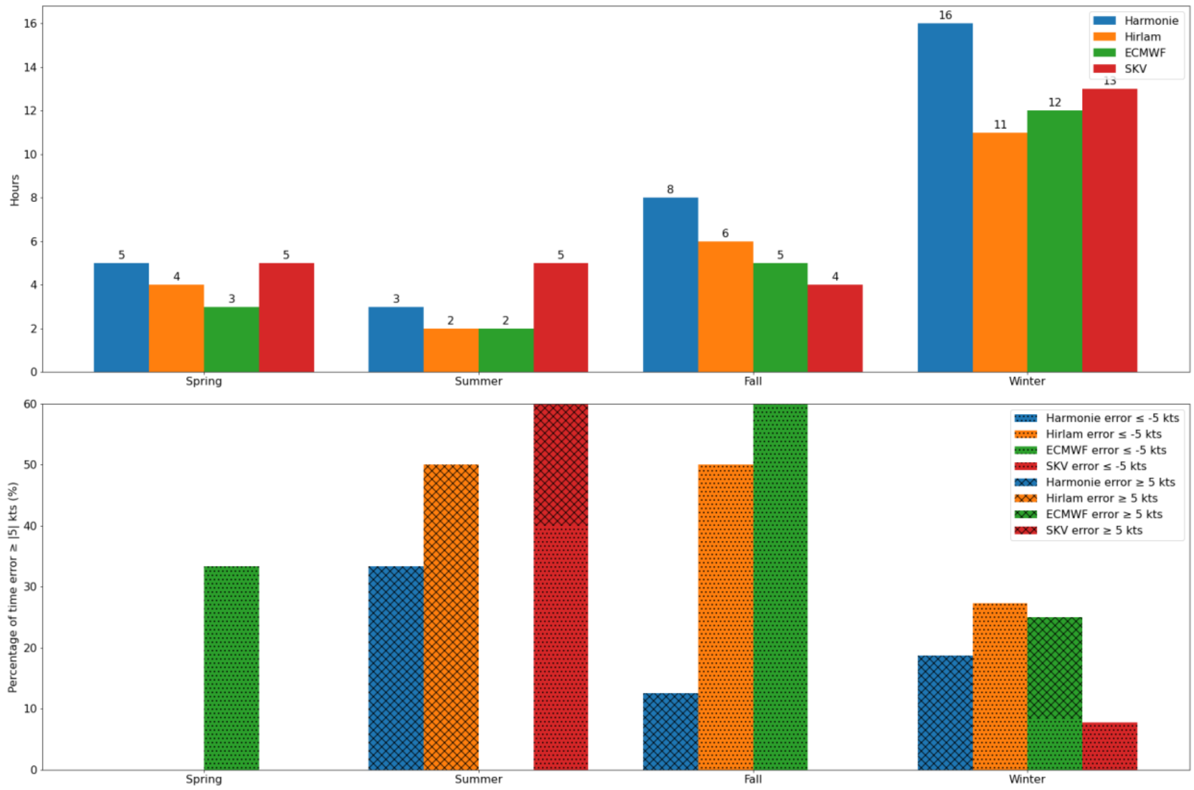
6-hour forecast, 15-19 kts



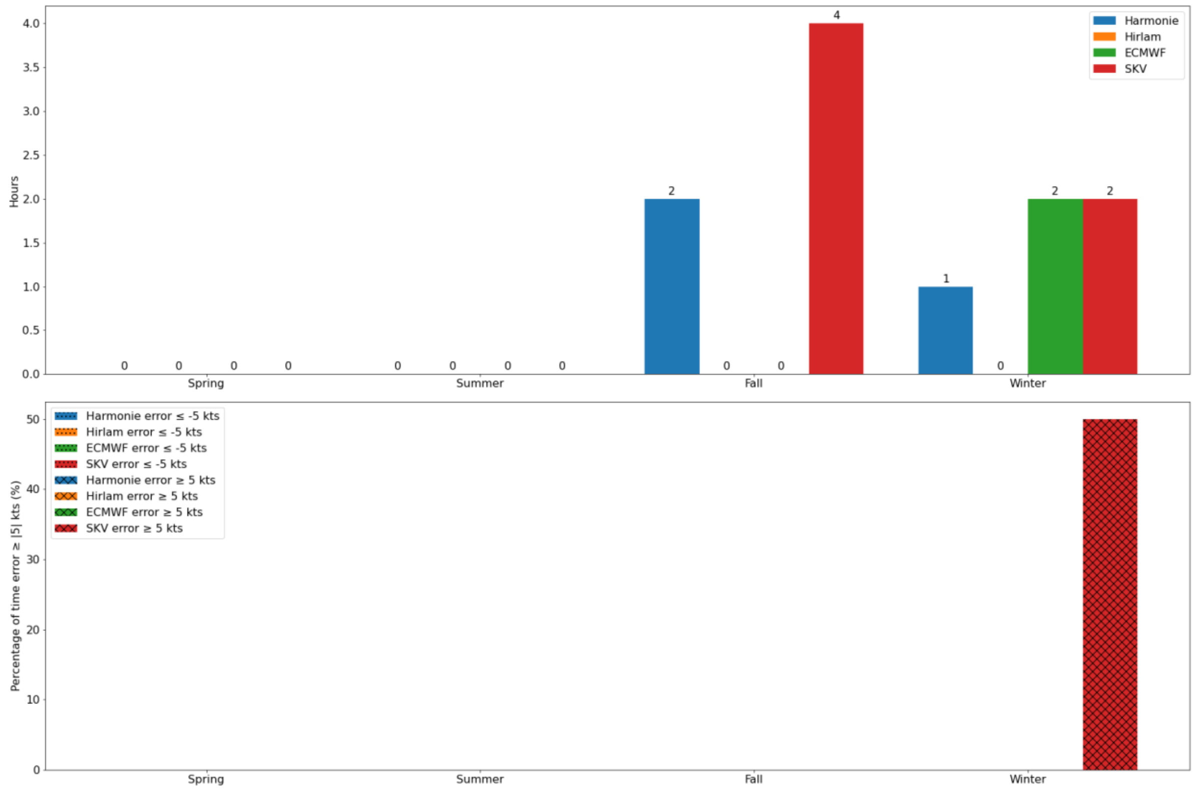
6-hour forecast, 20-24 kts



6-hour forecast, 25-29 kts



6-hour forecast, 30-34 kts



6-hour forecast, 35+ kts