Koninklijk Nederlands
Meteorologisch Instituut
Ministerie van Infrastructuur en Waterstaat

Hot-air balloon forecasts at KNMI, a validation study for the application of ECMWF and HIRLAM data

Cisco de Bruijn, Bas de Haan, Paul de Valk, and Sander Tijm

De Bilt, 2021 | Technisch rapport; TR-392

Hot-air balloon forecasts at KNMI, a validation study for the application of ECMWF and HIRLAM data

Cisco de Bruijn[1], Bas de Haan[2], Paul de Valk[2], and Sander Tijm[2]

[1]KNMI, RDWD, Department of Research and Development of Observations and Data Technology
[2]KNMI, WKD, Department of Weather and Climate Services

September 14, 2021

Abstract

This validation study focuses on the input data quality for the special hot-air balloon forecasts at KNMI of two models HIRLAM and ECMWF. The HIRLAM model will be out-phased in 2021 and therefore replaced by ECMWF model as data source. It is shown that the ECMWF model slightly outperforms HIRLAM in predicting the wind speed and direction. This has been assessed subjectively by eighteen on-duty forecasters and objectively by analyzing statistical scores where model outcome and observations are compared. Regular wind observations at 10 m height in the Netherlands have been used as well as upper-air wind observations from the Cabauw meteorological tower and from Hot-air Balloon (HAB) tracks. The validation period is August - October 2020.

Figure 1: Ascending recreational hot-air balloons

Contents

1 Introduction

For more than a decade KNMI provides special wind-forecasts for hot-air balloonists, because safety is important [1]. KNMI is in contact with Hot-air Balloon (HAB) organizations like PBN and KNVvL. Hot-air ballooning depends heavily on accurate forecasts and adversary weather might lead to accidents. Therefore reliable forecasts and guidance by experts are crucial. A dedicated application creates automatically bulletins which are disseminated among the end users, see [2]. This application uses data from NWP models. At first HIRLAM provided the data and since this model is going to be decommissioned, replacement is necessary. This report is structured as follows. In Section 2 the surface and upper-air wind measurements are described. In Section 3 an overview of the NWP models is given. In Section 4 and 5 the subjective and objective verification results are reported, respectively. A discussion on the validation results can be found in Section 6, followed by conclusions and outlook. On the final page of this report a list of acronyms can be found.

2 Observations

In this study we have used surface and upper-air observations for verification. At surface level we have applied the wind information from the AWS's (Automatic Weather Station). The measurements are from cup anemometers and vanes which are placed on 10 m tall mast and the collection process is fully automatic. There are 38 land stations spread over the The Netherlands, see[2]. The instruments are maintained and calibrated by KNMI and comply to WMO regulations. For our validation study we have 10 minutes averaged data.

For upper air observations we have used data from the meteorological tower at Cabauw site [3]. At the mast cup-anemometers and vanes are installed at 10,20,40,80,140,200 m. The observations are continuously available as 10 min averages and have a high quality stamp.

Other upper-air observation data is from the HAB itself. This is an indirect wind measurement based on Global Navigation Satellite System (GNSS) positions and elapsed time, which is the same principle as the weather balloon (radiosonde) measurement. HAB's have other characteristics than weather balloons; they are bigger and have a larger inertia and respond differently on a wind change, see [4] for more details. HAB wind observations sample the Atmospheric Boundary Layer (ABL) and are suitable for the validation of the special balloon forecast [4].

	HIRLAM 7.2	ECMWF
Domain:	Europe and North Atlantic	Globe
Hor. res.:	11×11 km^2	11×11 km^2
Vert. res.:	60 layers; surface – 10 hPa	85 layers; surface – 1 hPa
Data assimilation:	every 3 h, 3DVAR	every 6 h, 4DVAR
Observations:	conventional, Mode-S EHS, AMDAR, satellites	conventional, AMDAR, satellites
Lateral boundaries:	every 3 h, from ECMWF model	not required, continuous domain
Physical parameterisation:	TKE-l, ISBA surface scheme	EDMF, HTESSEL surface scheme

Table 1: Model characteristics of the HIRLAM and the ECMWF model

3 NWP models

In Table 1 the model characteristics are given. ECMWF is a global model that is run 4 times a day. The analysis is done every 6 hours and every 12 hours a complete forecast is made. HIRLAM runs 4 times a day and covers a limited area (Western Europe and the Atlantic Ocean). Every analysis is accompanied with a +48h forecast. In a 3DVAR Data Assimilation (DA) scheme, it is assumed that all observations have been measured at analysis time. This is generally true for observations from radiosondes, synops, buoys and ships. However, aircraft and satellite observations are asynoptic, introducing a time shift between observation and model background state [6]. This timing issue can be resolved by using 4DVAR DA scheme which has been implemented in the ECMWF model.

Conventional observations are synop, buoy, ship and radiosonde data. AMDAR and Mode-S EHS are aircraft data. AMDAR is world wide available, but is commercially exploited and delivers directly wind and temperature. Mode-S EHS are third party data from Air Traffic Control (ATC) and after processing wind and temperature data can be obtained. Satellites are remotely sensing instruments. Satellite data currently used by the ECMWF model are IASI(infrared), MODIS(visible light), METOP(scatterometer), AEOLUS(profiles of a single wind component). HIRLAM uses only scatterometer satellite data. It should be noted that HIRLAM receives large scale information via the lateral boundaries and as such benefits indirectly from the world wide satellite observations of the ECMWF model. HIRLAM has a boundary layer scheme that is based on the evolution of the Turbulent Kinematic Energy (TKE) equation. The interaction with the surface is described by the Interaction Surface Biosphere Atmosphere (ISBA) scheme. ISBA has two layers and responds rather slowly to the atmospheric forcing. HTESSEL, which is implemented in the ECMWF model, has 7 layers and a so-called skin layer which allows a quick response to the atmospheric forcing. In the ECMWF model the turbulence and convection scheme are integrated in the Eddy Diffusivity Mass Flux (EDMF) scheme. The eddy diffusivity is parameterized by K-diffusion(first-order closure).

	bad	so-so	good
How useful was the more recent HIRLAM-run compared to observations?	10	14	27
How useful was the older ECMWF-run compared to observations?	3	13	35
Were specific phenomena present in ECWMF, while they were absent in HIRLAM or vice versa?	5	3	43
Was the older ECMWF-run good enough to substitute the more recent HIRLAM-run at the moment of issuing the bulletin?	17	10	24
	yes	**no**	
Were there relevant differences between the older ECMWF-run and the more recent HIRLAM-run which had significant impact on the quality?	24	27	
	yes	**no**	
conclusion: Was the ECMWF-model good enough to substitute the HIRLAM-model?	44	7	

Table 2: Results of the subjective evaluation during 17 August - 19 October 2020

4 Subjective verification

The evaluation was carried out by duty forecasters during 17 August - 19 October 2020. Forecasters compared observations with ECMWF and HIRLAM model data. They used 10 m wind observations from 38 AWS's, but also upper air data from the meteorological mast at Cabauw and the SODAR at Schiphol. In Table 2 it is explained how the models were compared. During the shift a questionnaire form was filled in and in Table 2 the results are summarized. The following conclusions may be drawn:

- The usefulness of ECMWF (35 good/3 bad) is better than HIRLAM (27 good/10 bad).
- In 50% of the cases the results are similar, in the other 50% model results deviate
- ECMWF has 24 cases of out-performance against 10 cases of HIRLAM
- The majority of the meteorologists prefers ECMWF above HIRLAM (44 from the 51 cases)

5 Objective verification

For the objective verification we have used surface and upper-air observations which have not been used in the analysis. This means that only observations not used by the model are used for verification.

5.1 Surface verification with AWS data

Daily verification takes place on regular basis to keep track of the model performance. Hereto the 10 m wind data from the Automatic Weather Stations (AWS) are used. The AWS's are equipped with a 10 m wind mast. It includes a wind vane to measure the wind direction and a cup anemometer to measure wind speed. The response length of the anemometer is defined as the length of the passage of wind (in meters) required for the output of the wind speed sensor to indicate 63 percent of a step-function change of the input speed. The response length is about 3 m. The corresponding response time is the response length divided by the wind speed. The lowest model levels (HIRLAM level 65, ECMWF level 91) both lie close to this height and it would seem simplest to use these data directly as the forecast 10m wind. We present results from HIRLAM(CIS), which is almost the same version as HIRLAM(H11). We show results in

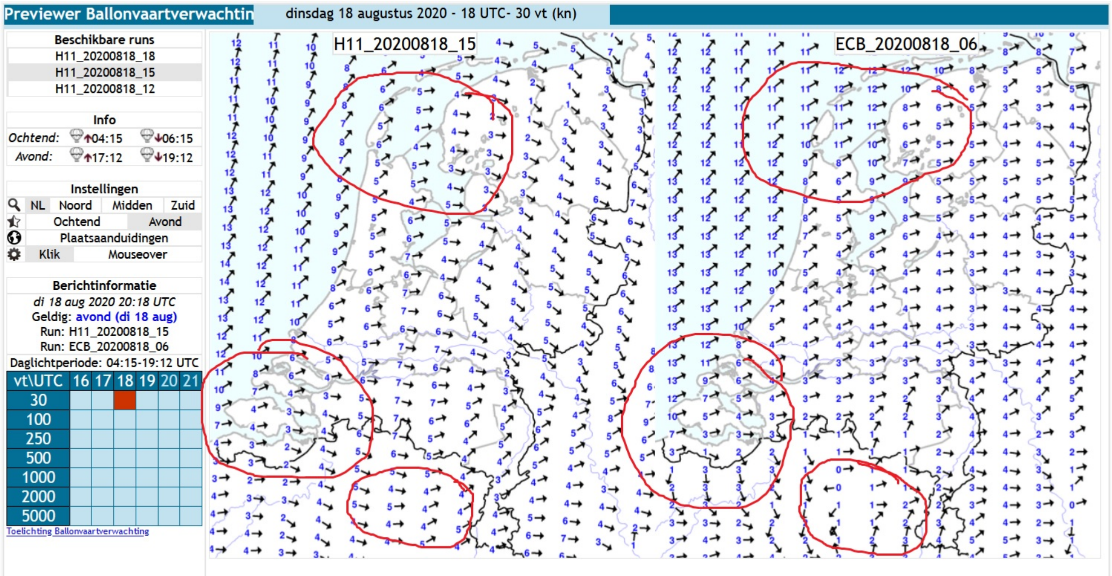


Figure 2: Example of the inter comparison, left panel HIRLAM 2020081815 +03, right panel ECMWF 2020081806 +12, areas with significant differences are marked in red

scatter diagrams where we focus on +00-06h (HIRLAM(H11)) and +06-12h (EC) lead times, because they are typically used for the special balloon forecast.

Here we present results for lead times varying from 0 to 48 hours. HIRLAM is labeled as CIS. Note that also HARMONIE results are shown. HAP1 is the operational HARMONIE cy40h1.1.1, HAP2 is cy40h2.1tg1 with modifications in the physics which brings the performance closer to HARMONIE cy36 and HAP3 is HARMONIE cy40 with modifications in shallow convection and clouds which will be incorporated in cy43.

It should be noted that the HAB forecast objective is now-casting and the first 12 hours are the most important. The longer lead times are useful for flight planning and the organization of the trip. In the scatter plots in Figure 4 it is obvious that the ECMWF and HIRLAM wind direction have a bias, which means that the wind direction is approximately too much veered. Note that HIRLAM has a smaller bias, which is probably caused by the shorter lead time. In the time series, which are depicted in Figures 5,6,7, it is revealed that ECMWF has the smallest standard deviation in wind direction, but at the same time a larger bias. It is well-known that ECMWF wind fields are smoother than either HIRLAM and HARMONIE and this is also found during the subjective verification (see section 4). Finally it should be noted that the AWS verification comprises all cases occurring during the verification period. This implies that also strong wind cases are included in the scores. It is noted that a HAB pilots are not interested in extreme wind events, because they fly only when wind speeds are below 10 knots.

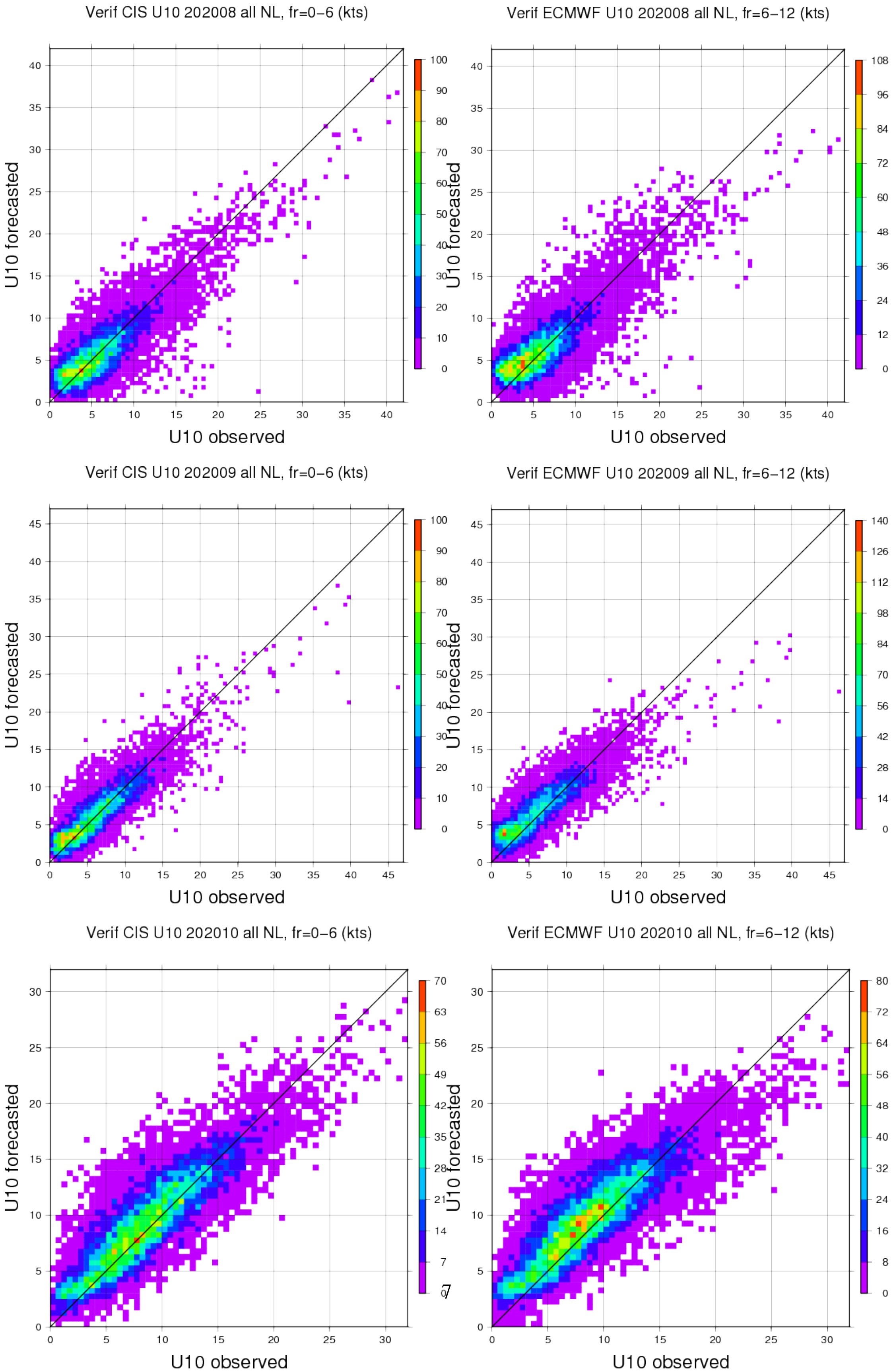


Figure 3: Scatter diagrams 10m wind August-October 2020, ECMWF +06-12, HIRLAM +00-06, units are knots

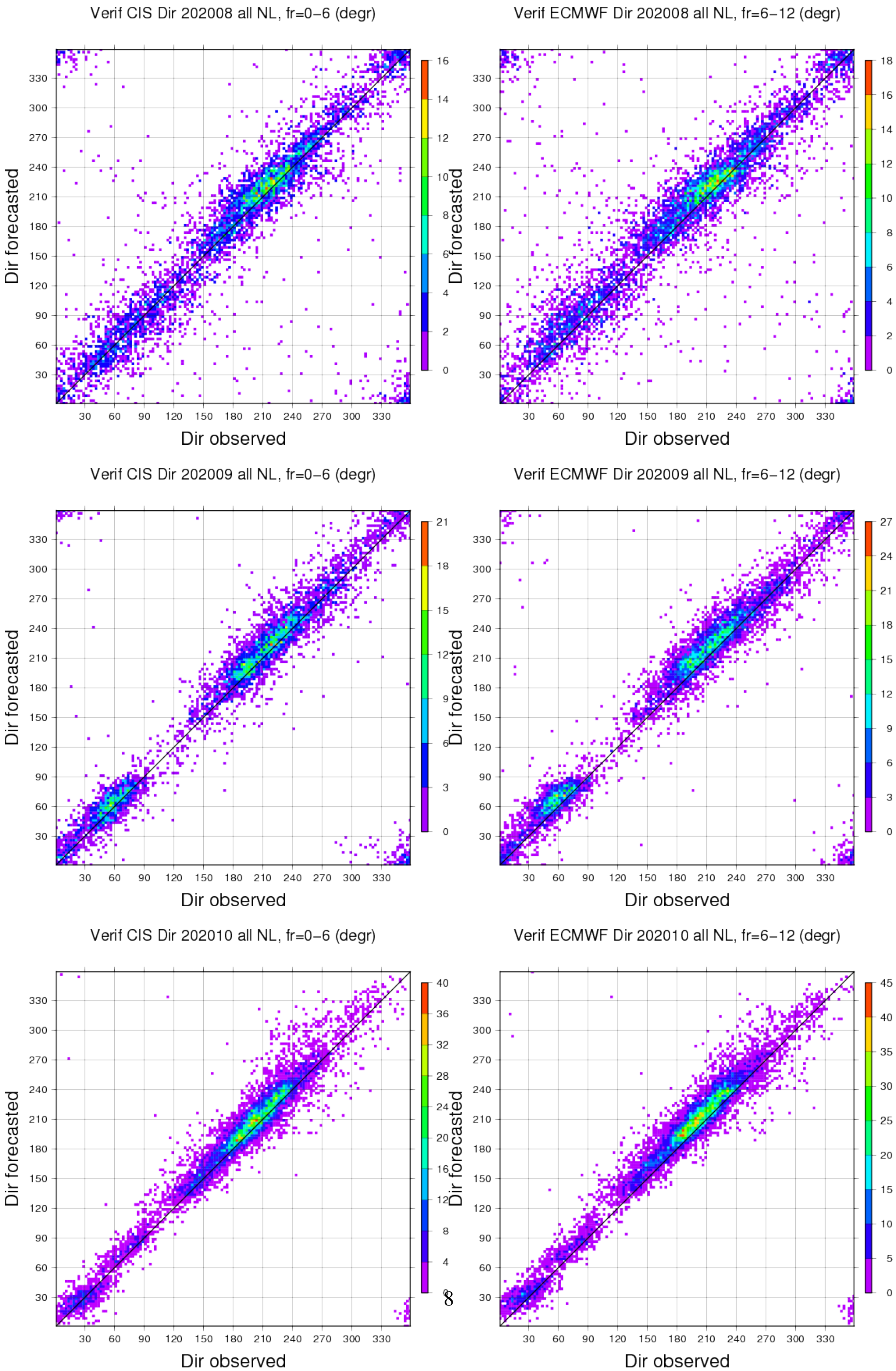
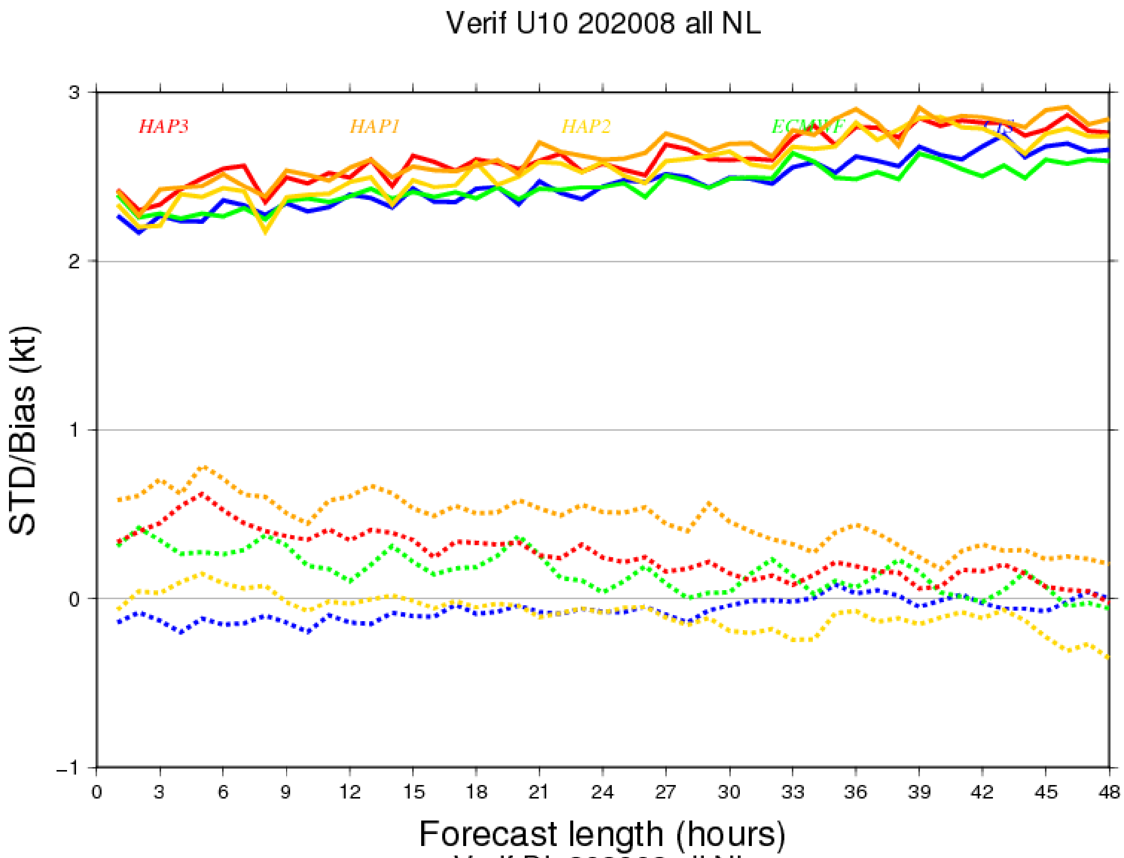


Figure 4: Scatter diagrams 10m wind direction August-October 2020, ECMWF +06-12, HIRLAM +00-06, units are degrees



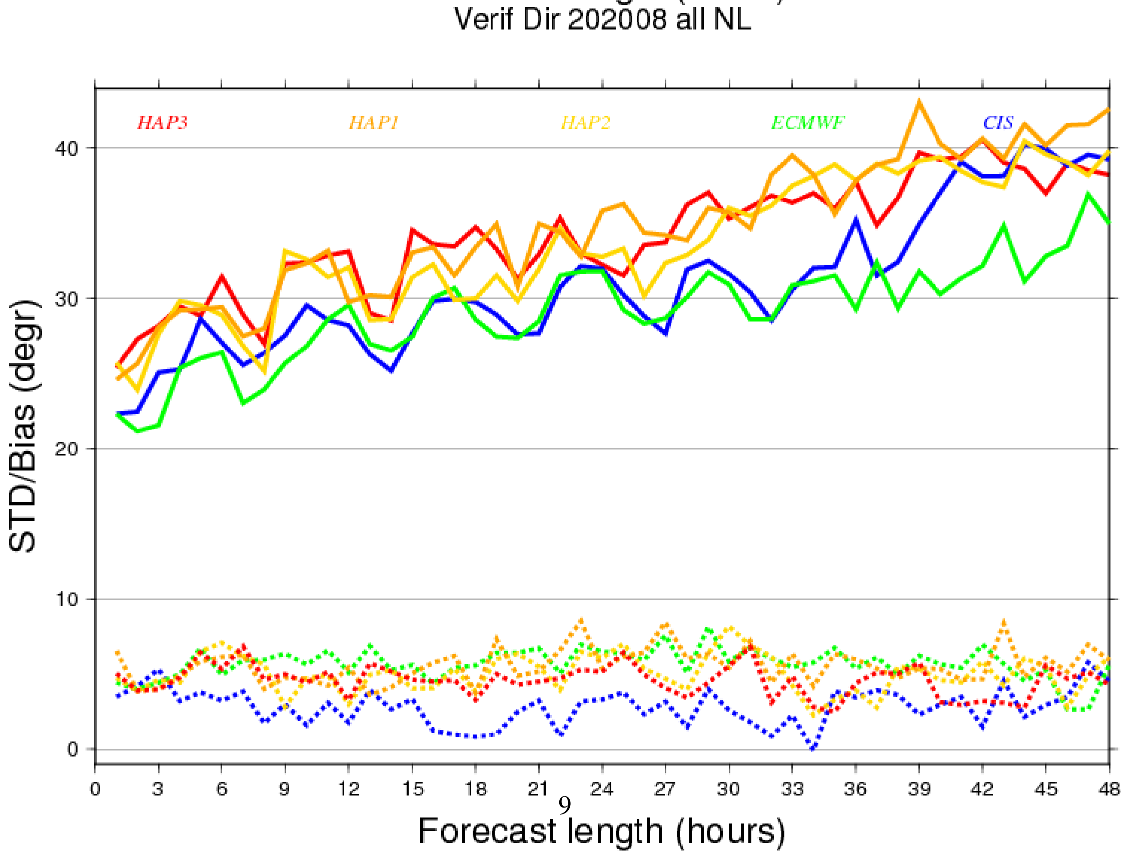


Figure 5: Observation-minus-forecast (o-f) statistics, standard deviation and bias as function of forecast length during August 2020, for 10m wind from SYNOP (o) and model forecast (f). HAP1=HARMONIE(cy40), HAP2=HARMONIE(cy36), HAP3=HARMONIE(cy43), ECMWF and CIS(HIRLAM)

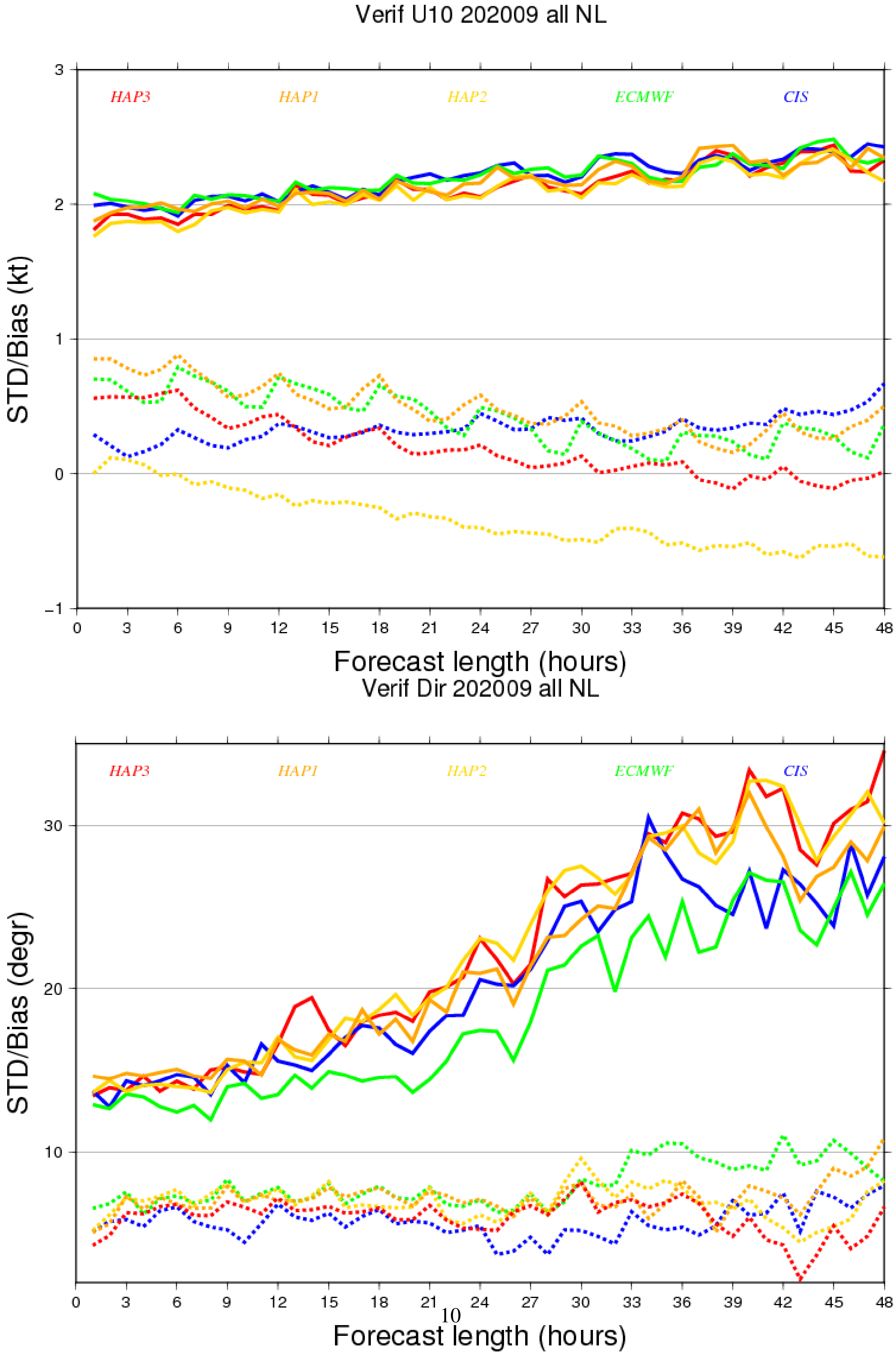
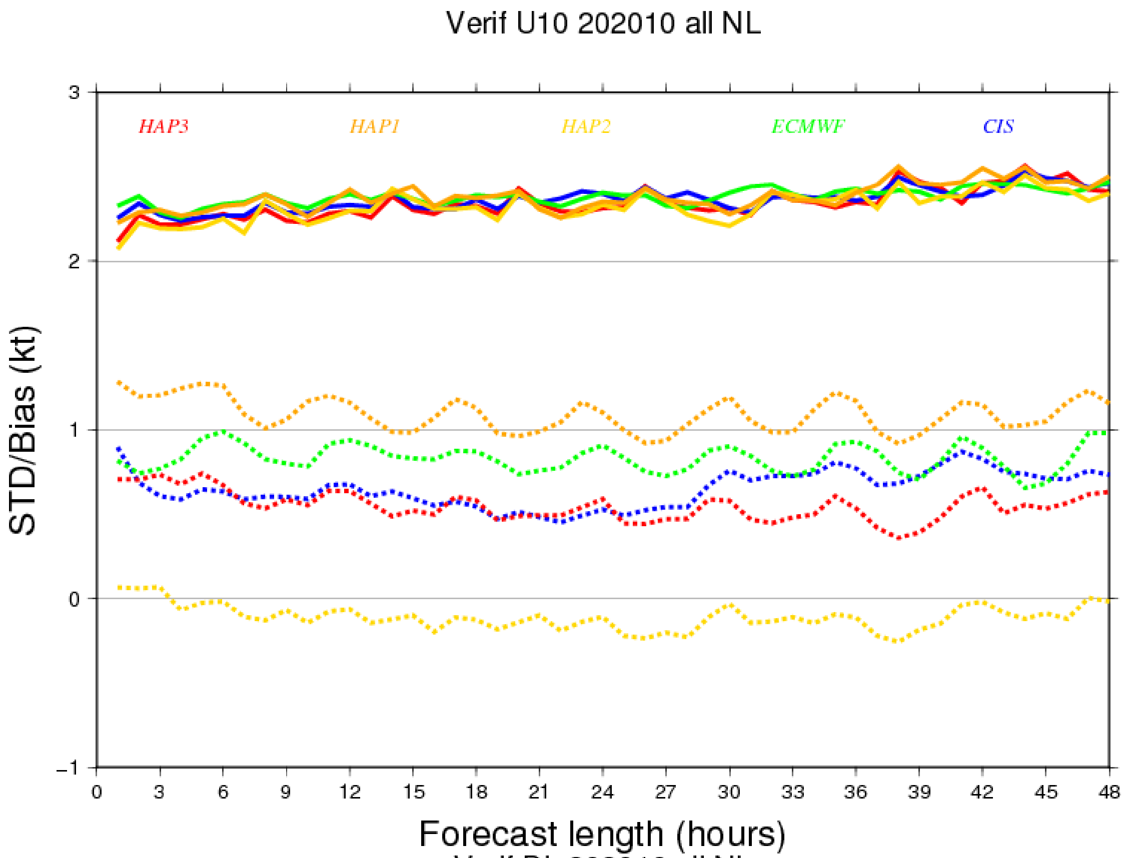


Figure 6: as previous Figure but now for September 2020



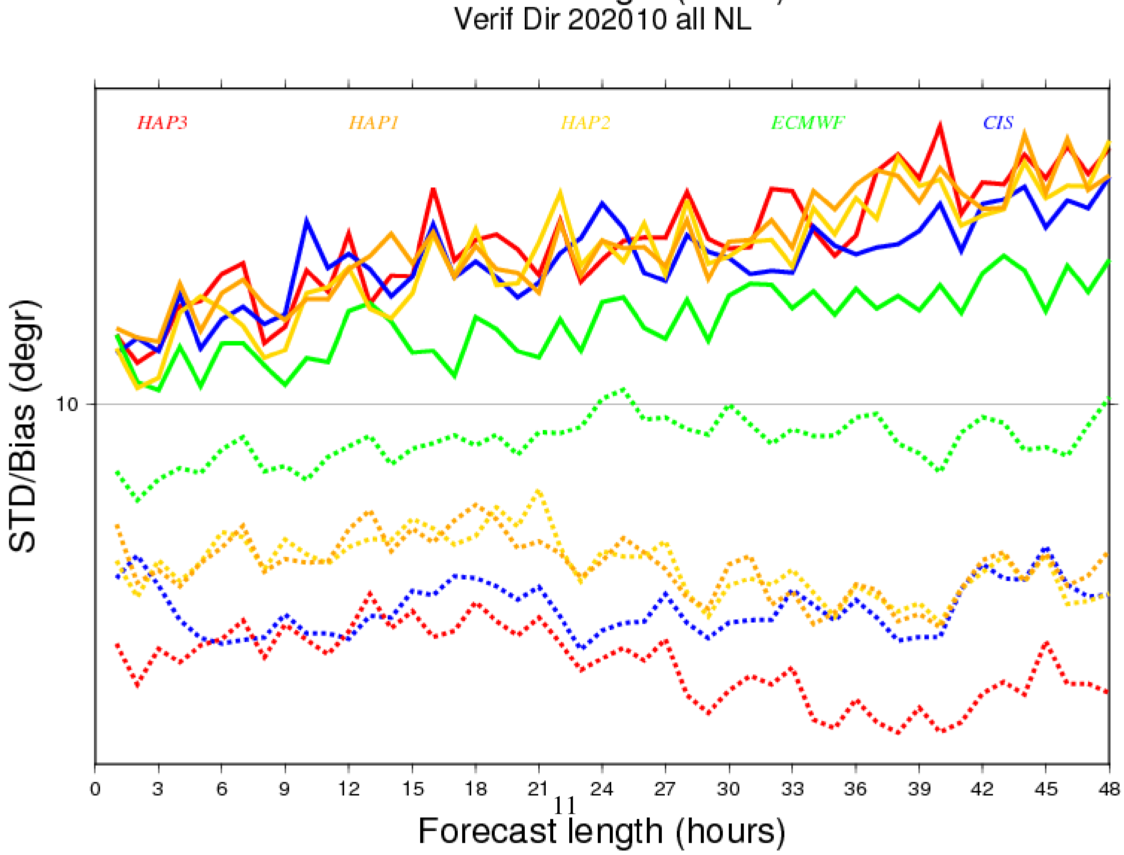


Figure 7: as previous Figure but now for October 2020

5.2 Upper air verification with HAB data

Due to representation mismatches between model, grid-box average, roughness and synop station roughness, interpretation of model 10m wind speed bias is not straightforward. Synop stations are normally situated in relatively flat terrain and consequently one could state that a NWP model usually over-predicts the 10m wind speed, resulting in a negative bias. (see [7], [8]). To overcome this problem, an upper-air verification has to be carried out. However upper-air observations are sparse and not always available.

Nonetheless with the help of the balloonist Peter Kelder who kindly supplied his flight data of 2020, we were able to perform an upper-air validation. We have processed the HAB tracks to observations [4], taking into account that a HAB has a response length of about 100 (m) (see [5]) and we have focused on the hours before sunset when balloons are flying. HAB's usually fly when there is daylight and good visibility. Further the wind speed should not exceed values of 10 knots and there should not be convective thermals and these conditions usually occur just after sunrise or just before sunset. In our data set only evening flights are available and the time-slot in our study is from 18:00-20:00 UTC. A drawback of this approach is that the data set is significantly reduced, but the advantage is that more tailor-made observations can be used.

For a first check of the quality of the HAB wind, we compare the HAB observations directly with high quality mast observations at Cabauw. We have applied the following two constraints. Firstly, the HAB winds, which are depicted in Figure 8, should be in a 30 km radius from Cabauw, and secondly the HAB wind observation should be within the vertical range of the mast, i.e. between 10 and 200 m. Note that due to the first constraint the HAB flights in North-Holland are excluded.

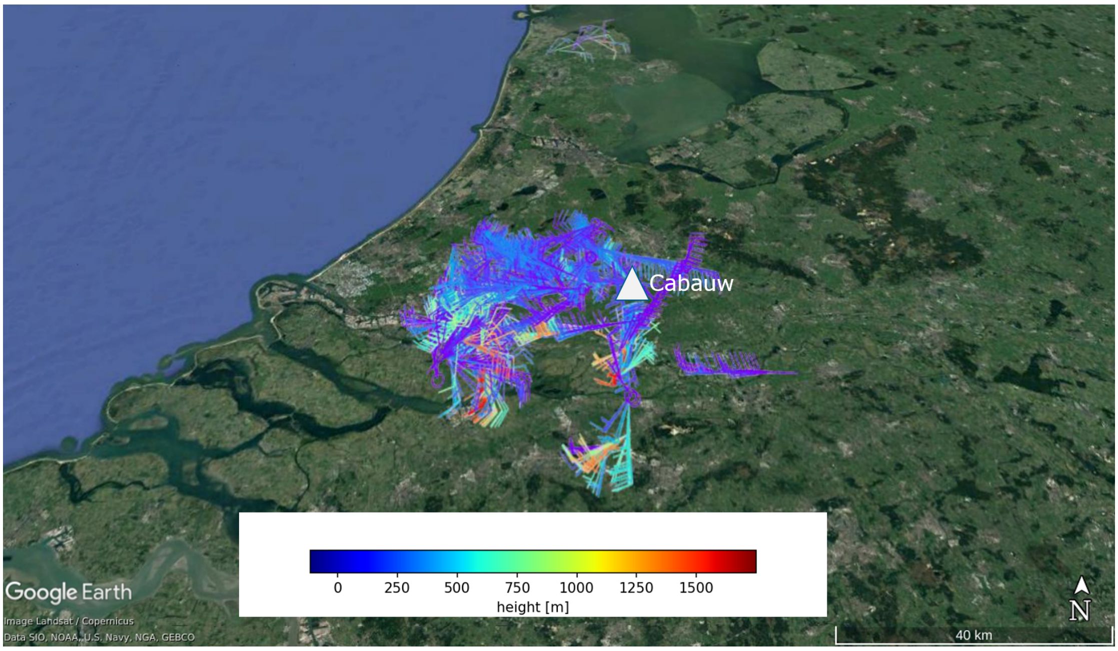
In Figure 9 a scatter plot with HAB observations and the Cabauw mast observations is shown. The deviation of the u-component is slightly smaller than the v-component, but all in all the match is good, which implies that the HAB observations make sense. There is a typical outlier in the v-component and the reason might be the blocking effect of obstacles. After this encouraging assessment we continue with the validation of the NWP models, applying HAB data. We limit our research to the first guess or model background. Statistics of observation minus background shortly denoted as (o-b), is an important diagnostic for NWP models to check for model and/or observation biases. In Figure 9 we recognize a clear bias in the v-component of HIRLAM. This bias is absent in the 6 hourly ECMWF output and this can be attributed to the 4DVAR assimilation scheme and a better physical parameterisation scheme.

So far, we have validated the NWP model in the vicinity of the Cabauw tower. Now we use all available HAB observations for the validation of HIRLAM and ECMWF and we present them in a scatter diagram Fig. 10. Clearly, ECMWF outperforms HIRLAM with less scatter and the best statistical scores.

5.3 Upper air verification with Cabauw mast data

To reassess the results from the previous section we carry out a cross-validation by applying another data set. An obvious choice would be to apply Cabauw tower data. Wind observations are available on 10,20,40,80,140,200 m heights and have a high quality. In Fig. 11 the statistical scores are depicted as a function of height. Just as in Fig. 9 the same bias is recognized for HIRLAM(H11). Further a significant smaller bias is revealed for u and v-components of the ECMWF model.

HAB flights take place when the decoupling begins and the ABL becomes more stratified. HIRLAM has difficulties with a too strong decoupling in the ABL, which is a well known problem. However HIRLAM has become end-of-life software and there will be no development on the physical parameterisation scheme any more. The ECMWF model is still being developed and has a updated boundary layer scheme. Further HIRLAM performs the analysis more frequently with a 3DVAR scheme and as a result the model is not able to adapt to the stabilization process which might lead to spin-up phenomena.



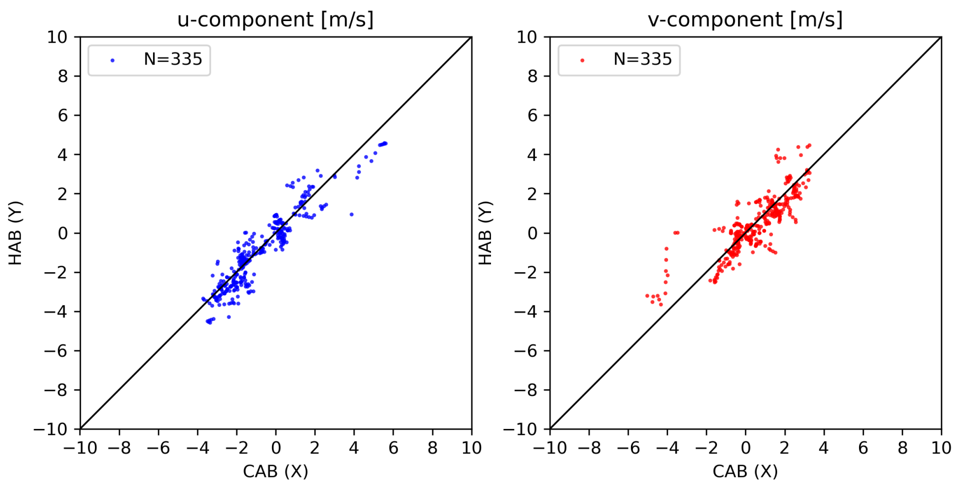


Figure 8: top panel: HAB wind observations nearby Cabauw between 17 July and 22 September 2020, bottom panel: Direct comparison of HAB wind observations with Cabauw mast wind observations

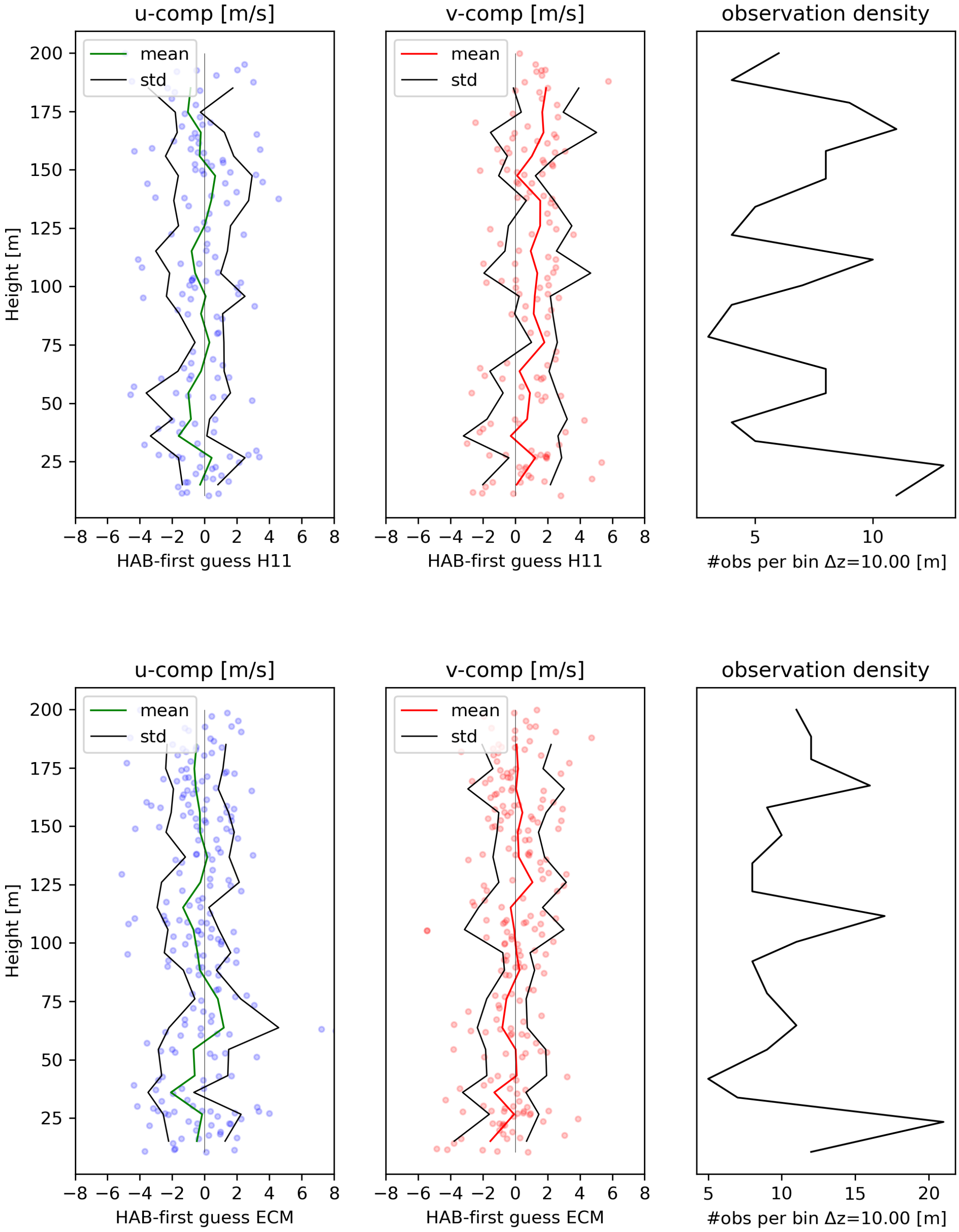


Figure 9: Observation - background (o-b) statistics of HIRLAM(H11) upper panel and ECMWF(ECM) bottom panel with HAB data during July - September 2020, b is first guess, previous model run valid at analysis time

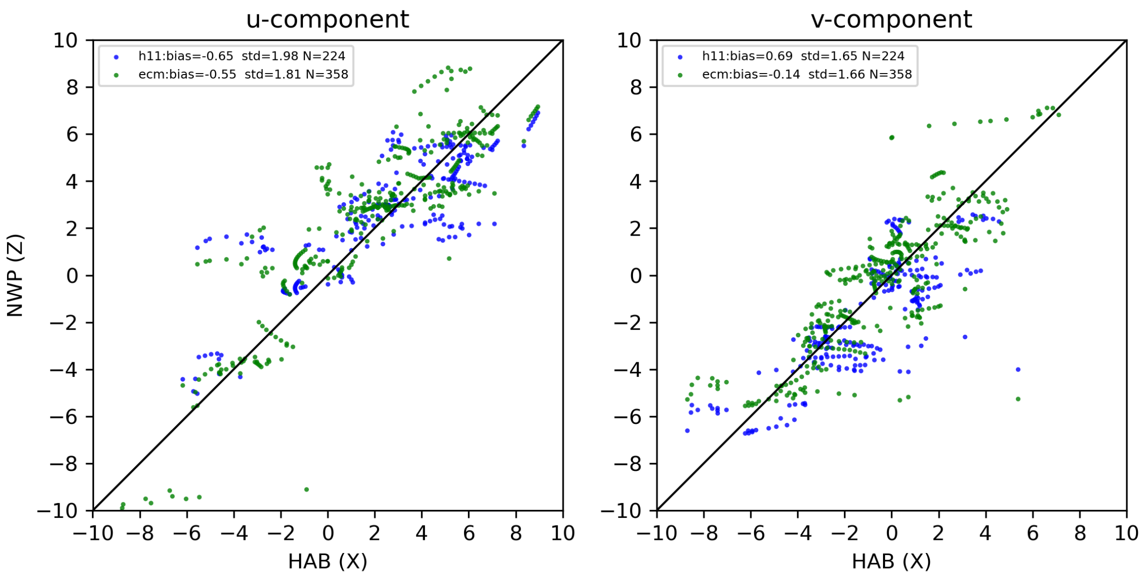


Figure 10: (o-b) statistics of HIRLAM(H11) and ECMWF(ECM) with the complete HAB observation data set during July-September 2020

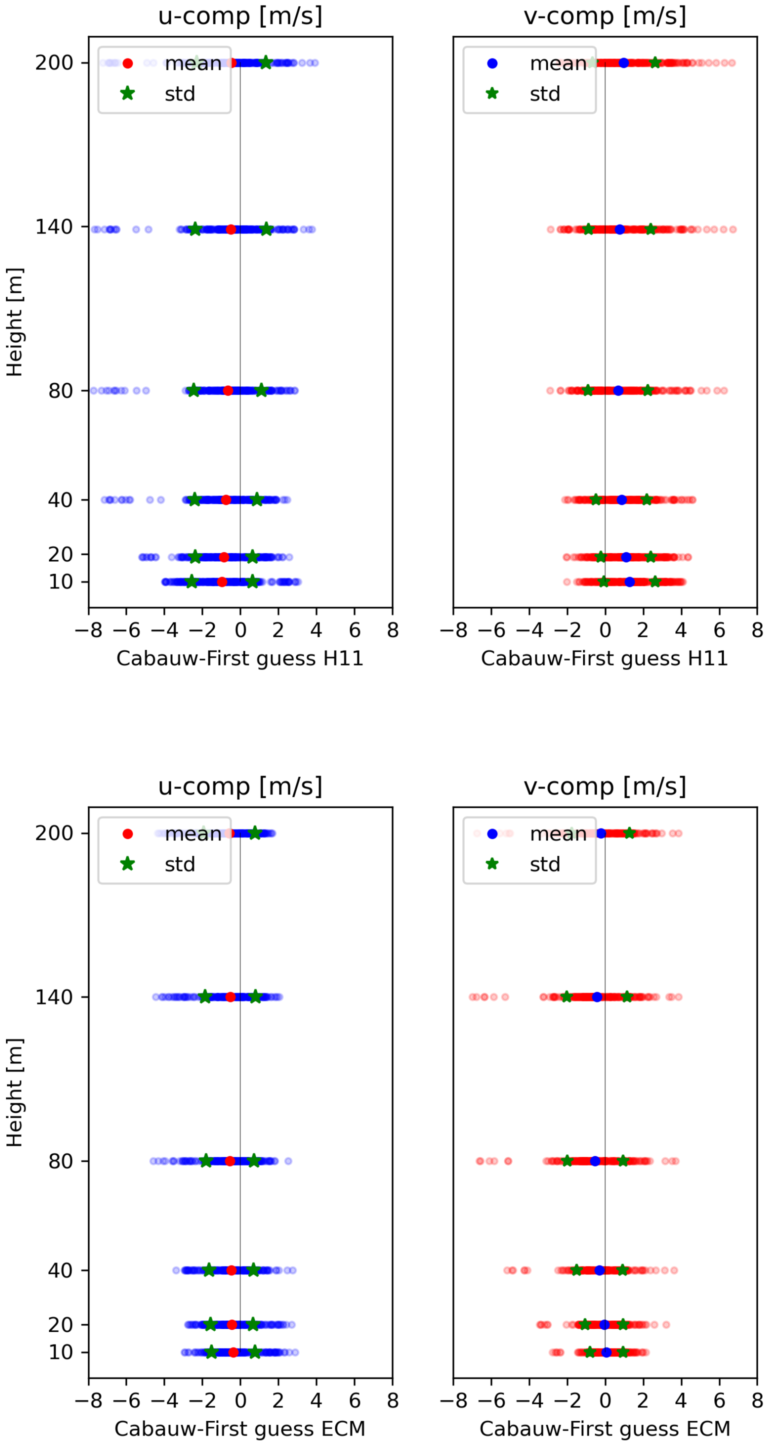


Figure 11: (o-b) statistics of HIRLAM(H11) upper panel and ECMWF(ECM) bottom panel with Cabauw mast observations during July-September 2020

6 Discussion

In this work we have used 10 m wind observation as well as special upper air observations for the objective verification. The 10 m observations were continuously present, while HAB winds are only available in narrow time-slot (16:00- 20:00 UTC). For a cross-validation we have used the Cabauw mast data in the same time-slot as the HAB wind data. From operational feed-back in 2021 it is known that HIRLAM has difficulties with stable conditions, especially at the interface of the collapsing ABL. We did not address specifically this shortcoming in our validation, because the development of HIRLAM has finished.

Due to the covid-19 pandemic only a period in the late summer/autumn is chosen. Spring and early summer are not considered. In this period sea breezes often occur due to the temperature contrast between land and sea and these local circulations have a substantial impact on the wind forecast, but are not taken in account here. In this study we did not investigate the performance of the gustiness which is an important parameter for balloonists, but this issue is a subject for further research.

7 Conclusions

In this report we have investigated the performance of the input NWP data for the special HAB forecast. The question was if HIRLAM could be replaced by ECMWF. A subjective verification by duty forecasters revealed that the the ECMWF model was the preferred choice. Subsequently an objective verification was carried out. Based on a surface verification during August-September 2020 we conclude that ECMWF is slightly worse in predicting wind direction than HIRLAM. Further ECMWF and HIRLAM have a similar quality in predicting wind speed. This means that the subjective and objective verification are not contra dictionary.

We have also applied upper air observations to check if this conclusion could be confirmed. Based upon first guess validation with HAB and Cabauw mast observations we conclude again that the 6-hourly ECMWF has a smaller bias and standard deviation for the u- and v-component of the wind than the hourly HIRLAM.

Our final conclusion and recommendation is that ECWMF can be used as input for the special balloon forecast.

8 Outlook

The high resolution HARMONIE (2.5×2.5 km^2) is a promising model, but in its current state not suitable for providing wind data for the balloon bulletin. The output fields contain lots of details, for example the rapidly evolving wind pattern around thunderstorms. These data should not be interpreted deterministically, but probabilistic-ally. An Ensemble Prediction System (EPS), which provides wind probabilities, could be helpful to tackle this problem.

A promising development is that HARMONIE will assimilate observations more frequently. There are even plans for continuous assimilation of observations.

We conclude this outlook with a promising development in the infrastructure to collect HAB data. HAB data can be collected using smartphones [5], but alternatively transponders can also be applied. Currently more and more HAB's are equipped with transponders, so that they are under surveillance of Air Traffic Control (ATC). These data are also used by www.luchtballonradar.nl, a website where HAB's can be tracked real-time. Interestingly this website offers an archive for completed flights as well.

References

[1] Anonymous, “Balloon flying handbook,” *U.S. Department of Tranportation Federal Aviation Administration*, vol. FAA-H-8083-11A, p. 252, 2008.

[2] R. Boonstra, “Verificatie ballonvaart verwachting, grondwind,” *KNMI-Publication*, p. 37, 2010.

[3] Fred C. Bosveld, Peter Baas, Anton C.M. Beljaars, Albert A.M. Holtslag, Jordi Vilà-Guerau de Arellano, and Bas J.H. van de Wiel, “Fifty Years of Atmospheric Boundary-Layer Research at Cabauw Serving Weather, Air Quality and Climate,” *Boundary-Layer Meteorology*, 2020.

[4] E.I.F. de Bruijn, F.C. Bosveld, S. de Haan, B.G.J. Wichers Schreur, and A.A.M. Holtslag, “Observing Boundary Layer Wind from Hot-air Balloon Flights,” *Wea. Forecasting*, vol. 31, pp. 1451–1463, 2016.

[5] E.I.F. de Bruijn, F.C. Bosveld, S. de Haan, and B.G. Heusinkveld, “Measuring Low Altitude Wind with a Hot-Air Balloon and Their Validation with Cabauw Tower Observations,” *J. Atmos. Oceanic Technol.*, vol. 43, pp. 263–277, 2020.

[6] G.-J. Marseille and A. Stoffelen, “Toward Scatterometer Winds Assimilation in the Mesoscale HARMONIE Model,” *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, vol. 10, pp. 2383–2393, 2017.

[7] W.C. de Rooy and K.A. Kok, “A combined physical-statistical approach for the downscaling of model wind speed,” *Weather and Forecasting*, vol. 19, pp. 485–495, 2004.

[8] Wim de Rooy and Hylke de Vries et al., “Harmonie verification and evaluation,” *HIRLAM Technical Report*, vol. 77, pp. 1–84, 2017.

Acknowledgments

Gert-Jan Marseille (KNMI, department RDSW) is gratefully acknowledged for a critical reading of this report and for his valuable comments, which have improved this report significantly. Steven Knoop (KNMI, department RDWD) is thanked for the LATEX template of this report. We are also indebted to the EMADDC-project at KNMI for providing the software for the interpretation of the HAB flight information.

Fenneke Overes (KNMI, department WKD) is thanked for setting up the subjective verification. The following meteorologists are thanked for their contribution to the subjective evaluation: Yorick de Wijs, Bas de Haan, Dorothé Lucassen, Aline Kraai, Niels de Groot, Jakob Kuiper, Hans van Bruggen, Adri Huiskamp, Rini Hoevenaren, Dirk Wolters, Cyrill Taminiau, John Brouwer, Michou Baart de la Faille, Sandra Niessink, Anika Hoekstra, John Kambeel and Patrick van den Hoek. Finally the balloonist Peter Kelder is kindly thanked for sharing his flight data in 2020, which allowed us to do the HAB validation.

List of acronyms:

- ABL = Atmospheric Boundary Layer
- AMDAR = Aircraft Meteorological Data Relay
- ATC = Air Traffic Control
- AWS = Automatic Weather Station
- ECMWF = European Center for Medium Range Weather Forecasts
- EMADDC = European Meteorological Aircraft Derived Data Center
- GNSS = Global Navigation Satellite System
- HAB = Hot-air Balloon
- HARMONIE = High Resolution Limited Area Model in Europemed, non-hydrostatic model
- HIRLAM = High resolution Limited Area Model, hydrostatic model
- KNVvL = Koninklijke Nederlandse Vereniging voor de Luchtvaart
- MODE-S EHS = MODE-S Enhanced Surveillance
- MODE-S MRAR = MODE-S Meteorological Routine Airport Report
- NWP = Numerical Weather Prediction model
- PBN = Professionele Ballonvaarders Nederland
- RDSW = Department of Research and Development of satellite observations
- RDWD = Department of Research and Development of observations and data technology
- UTC = Universal Time Corrected
- WKD = Weather and Climate Services
- 3DVAR = three dimensional data assimilation (x,y,z)
- 4DVAR = four dimensional data assimilation (x,y,z,t)

Koninklijk Nederlands Meteorologisch Instituut

Postbus 201 | 3730 AE De Bilt
T 030 220 69 11 | www.knmi.nl