Responses to anonymous reviewer #2

We thank the referee for highlighting the importance of this study and for helpful advice and constructive comments about our paper. The suggestions have led to a revised manuscript with two main goals: (1) While we confirm the minimal autocorrelation in the IAGOS data, we have reconstructed the training, validation and test datasets to ensure that model building and testing do not occur on the same day or overlap in any way. We now give additional details on the selection of the study region, considering humidity uplift from lower atmospheric levels, and describe the RHi distribution within the dataset, as well as provide further information on the development procedure of the algorithm. (2) We offer further insights into humidity, dynamics and temporal dependence, along with updated contributions of input variables to humidity prediction, and refined explanations regarding the importance of dynamical variables in RHi prediction. Additionally, we have included additional maps presenting humidity patterns. To this end, we have written concise explanations and modified pictures accordingly.

In the following we enumerate the referee's comments (RC) and our replies (R) to each, referencing the corresponding tracked changes in the manuscript.

RC: Summary of paper:

In this study, the authors train an artificial neural network to predict the distribution of relative humidity over ice in the UTLS over Western Europe. The network is trained on a mixture of thermodynamical and dynamical variables, although the former explain most of the prediction skill. The network is better than ERA5 at predicting RHi, and its inputs lead to better contrail prediction from the Cocip model in one case study.

The paper deals with an important topic. It is very well written. The introduction is excellent. The figures illustrate the discussion well, although I would have preferred to see more maps because scatterplots only give an incomplete indication of the ability of the network to reproduce patterns of humidity.

Others have commented in the online discussion on the need to better separate the training dataset from the validation dataset. I will not elaborate further on that aspect but revisions to the method are clearly needed there.

Ra) author's response

Thank you for the positive feedback on our proposed RHi improvement method. To better illustrate the accurate and enhanced spatial variability of RHi patterns over Western Europe, in addition to presenting the RHi distribution in the original Figure 6 of the manuscript, we have also added equidistant latitude-longitude maps from ERA5 and ANN RHi data at 200 hPa, taken at 08:00 UTC on July 21, 2021, corresponding to the flight of the HALO aircraft during the CIRRUS-HL campaign on that date.

Refer to Figure A1 below and Figure S5 in the revised supplement.

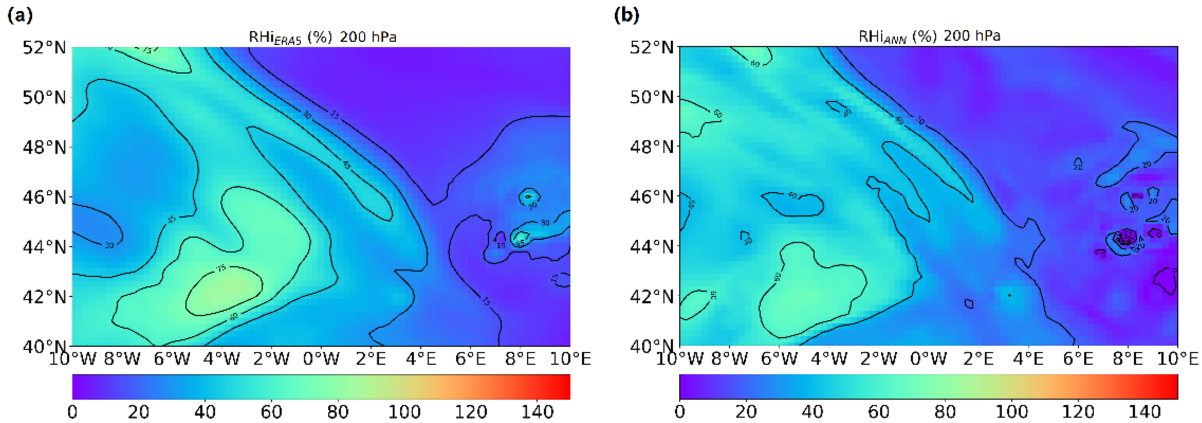

Figure A1 (S5): Patterns of (a) RHi_{ERA5} and (b) RHi_{ANN} at 200 hPa at 08:00 UTC on 21 July 2021.

The online discussion highlighted the needs to separate the training and validation datasets due to the autocorrelation present in the IAGOS data within a single ERA5 grid box during the IAGOS and ERA5 collocation. We appreciate this feedback and calculated the autocorrelation function of the IAGOS data series following the method of Dotzek and Gierens (2008). The results, displayed in Figure A2 here, show that autocorrelation reaches up to 0.04 at the original resolution (4 s, 1 km) before gradually declining to near zero at a resolution of 0.25 degrees (25 km), which corresponds to the ERA5 grid box size. It can be explained by the nature that the water vapor field is quite chaotic with steep gradients. Diao et al (2014, 2015) showed that in-situ RHi measurements taken with 1 Hz or 10 Hz instruments reveal small-scale structures in the RHi time series, in agreement with our finding. Therefore, we assume that autocorrelation can be disregarded for the averaged IAGOS RHi measurements within the ERA5 grid box.

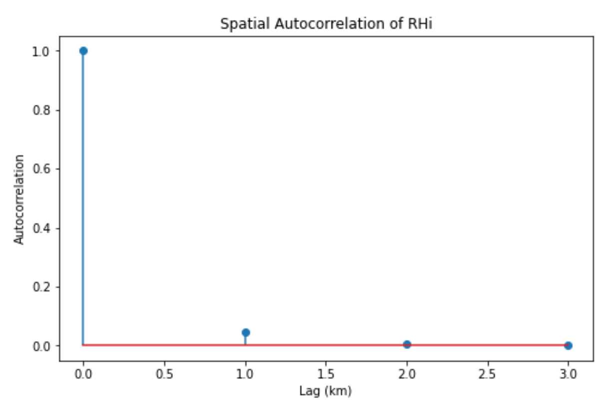

Figure A2: Autocorrelation function of the IAGOS measurement series (4s, 1km).

In response, we have still revised our methodology to ensure that the training set consists of data from time periods that do not overlap with those used for either validation and testing. Specifically, we are now using 4 days of data for building the ANN model, followed by a 1-day gap, and then 1 day of data for validation or testing. The complexity of the ANN model has also been increased to better match the growing complicacy of the training data, including the addition of hidden layers and neurons, as well as weight initialization and batch normalization for each layer to enhance generalization. Revisions to the text, along with updated Figures 4, 5, 6, 7, and S6 for RHi using the new models, are provided below. Updates on the validation of T_{IAGOS} and q predictions are presented in Figures S1, S3, and S4 in the revised supplement.

Rb) manuscript changes

L174-176: "To ensure that the training set consists of data from time periods that do not overlap with those used for either validation and testing, we now use 4 days of data to build the ANN model, followed by a 1-day gap, and then 1 day of data for validation or testing."

L291-294: "In Sect. 2.3, four consecutive days of samples from the ERA5-IAGOS collection are allocated for model training, with the following day excluded to avoid overlap with the continuous weather system, and another day reserved either for validation, to evaluate the model's generalization to unseen data during training, or for testing."

L43-45 in the supplement: To ensure model robustness "and construct an independent test data set, we now use a sequence-based split: four consecutive days of data are used to build the ANN model, followed by a 1-day gap, with the subsequent day's data reserved for validation or testing".

L286-287: "We use 3 hidden layers, each with 100 neurons and He weight initializer (He et al., 2015), along with batch normalization between layers to improve generalization. The humidity output is referred to as RHi_{ANN}."

L298: "an ANN for *q* is implemented, with 300 neurons in each hidden layer and…"

L18-20 in the Abstract: "The ANN shows excellent performance and the predicted RHi in the UT has a mean absolute error MAE of 5.7% and a coefficient of determination R^2 of 0.95, which is significantly improved compared to ERA5 RHi (MAE of 15.8%; R^2 of 0.66)."

L348-350: "The MAE decreases significantly from 15.82% (ERA5) to 5.71% (ANN), the R^2 values increase from 0.66 (ERA5) to 0.95 (ANN), and the root mean spare error RMSE decreases from 20.52% (ERA5) to 7.88% (ANN)."

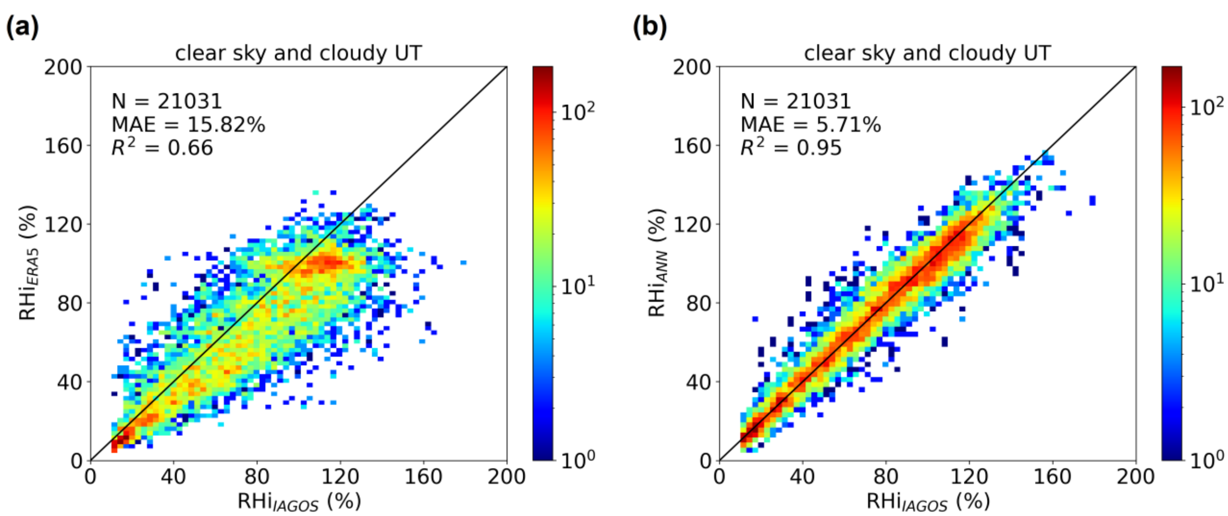

Figure A3 (4): Distribution of (a) RHi_{ERA5} and (b) RHi_{ANN} versus RHi_{IAGOS} in the UT in all sky (clear and cloudy conditions) in the test data set. The number of data sets N, the mean absolute error MAE and the coefficient of determination R^2 are shown in the panels.

L354-356: "In the cloudy (Fig. 5a-b) and clear sky (Fig. 5c-d) conditions in the UTLS, the MAE of the RHi decreases from 16.28% (11.21%) to 5.95% (4.28%), respectively. Also, the R^2 increases by 0.30 (0.23) to 0.95 (0.95) for the two scenarios."

L361-363: "The ANN model also has strong skills of RHi correction in the LS, see Fig. 5e and f. R^2 values increase from 0.59 (ERA5) to 0.95 (ANN), similar to the UT region in Fig. 4. The improvement of RHi prediction by the ANN is also documented by the decrease of MAE by 6.07%."

L377: "with a MAE of approximately 5.8%..."

L485-486: "Using this ANN humidity correction, the MAE of RHi_{ANN} when comparing to RHi_{ERA5} both against RHi_{IAGOS} is reduced from 15.82% to 5.71%, 16.28% to 5.95%, 11.21% to 4.28%, and 9.78% to 3.71%..."

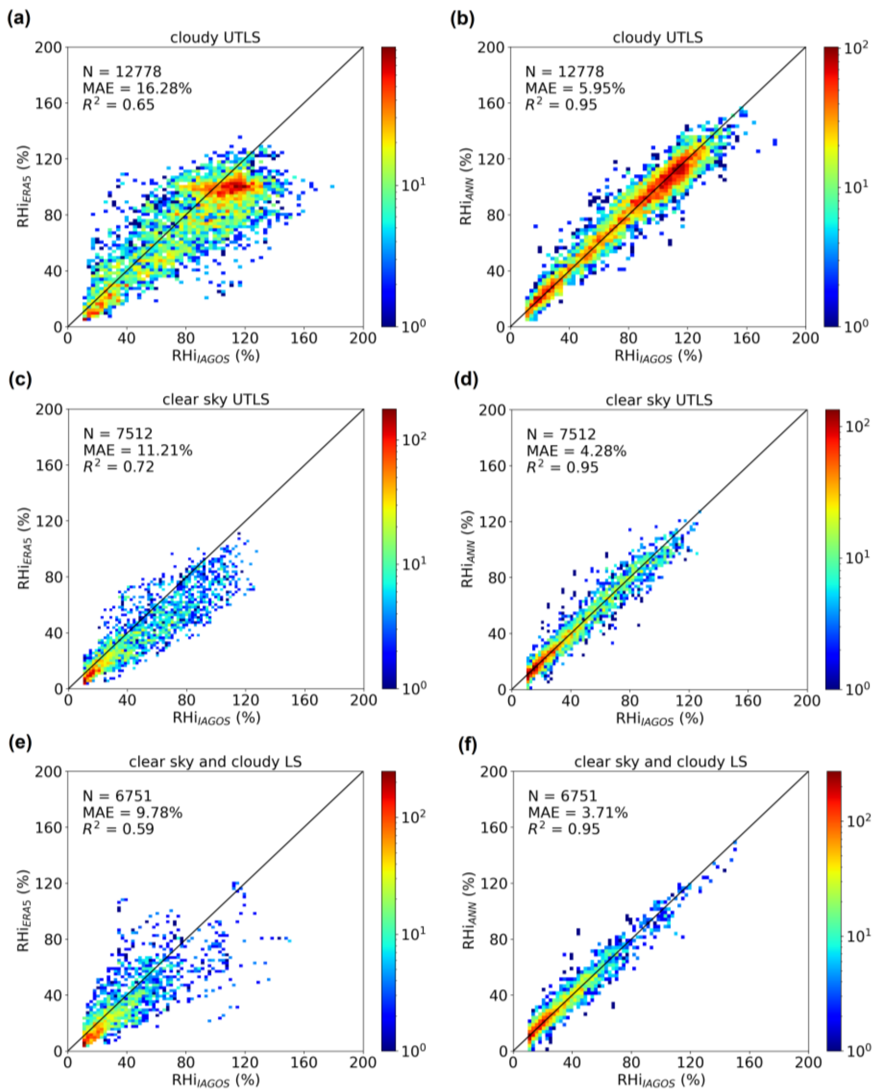

Figure A4 (5): Comparison of RHi_{ERA5} (left column) and RHi_{ANN} (right column) against RHi_{IAGOS} in the (a) and (b) cloudy UTLS, (c) and (d) clear sky UTLS, and (e) and (f) clear sky and cloudy (or all sky) LS regions in the test data set. The number of data sets N, the mean absolute error MAE and the coefficient of determination R^2 are indicated in the individual panels.

L369-371: "As opposed to this, Fig. 6d shows that RHi_{ANN} and RHi_{IAGOS} have a closer agreement, with an MBE of about ±11% for all UT measurements up to 140%. The RHi between 80% and 130% in the important range for cirrus clouds is well represented by the ANN with an MBE better than +/- 7%."

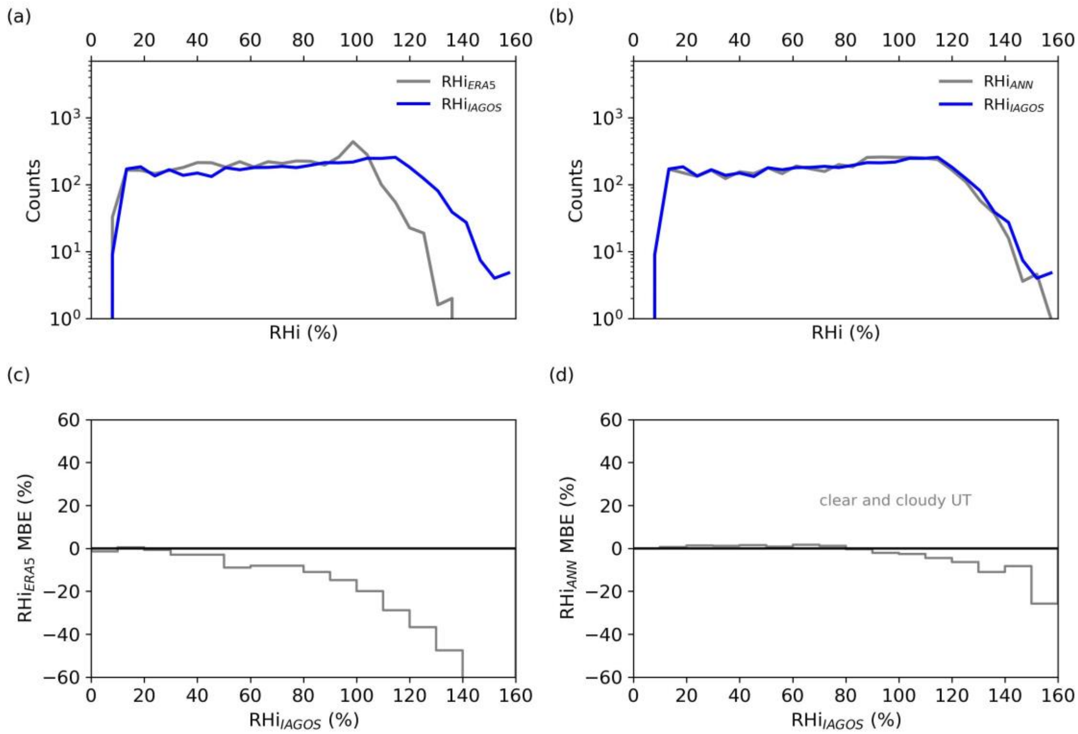

Figure A5 (6): Frequency distribution (a and c) and overall mean biased error MBE (%) (b and d) of RHi_{ERA5} and RHi_{ANN} against RHi_{IAGOS} in the clear and cloudy UT (grey) in the test data set.

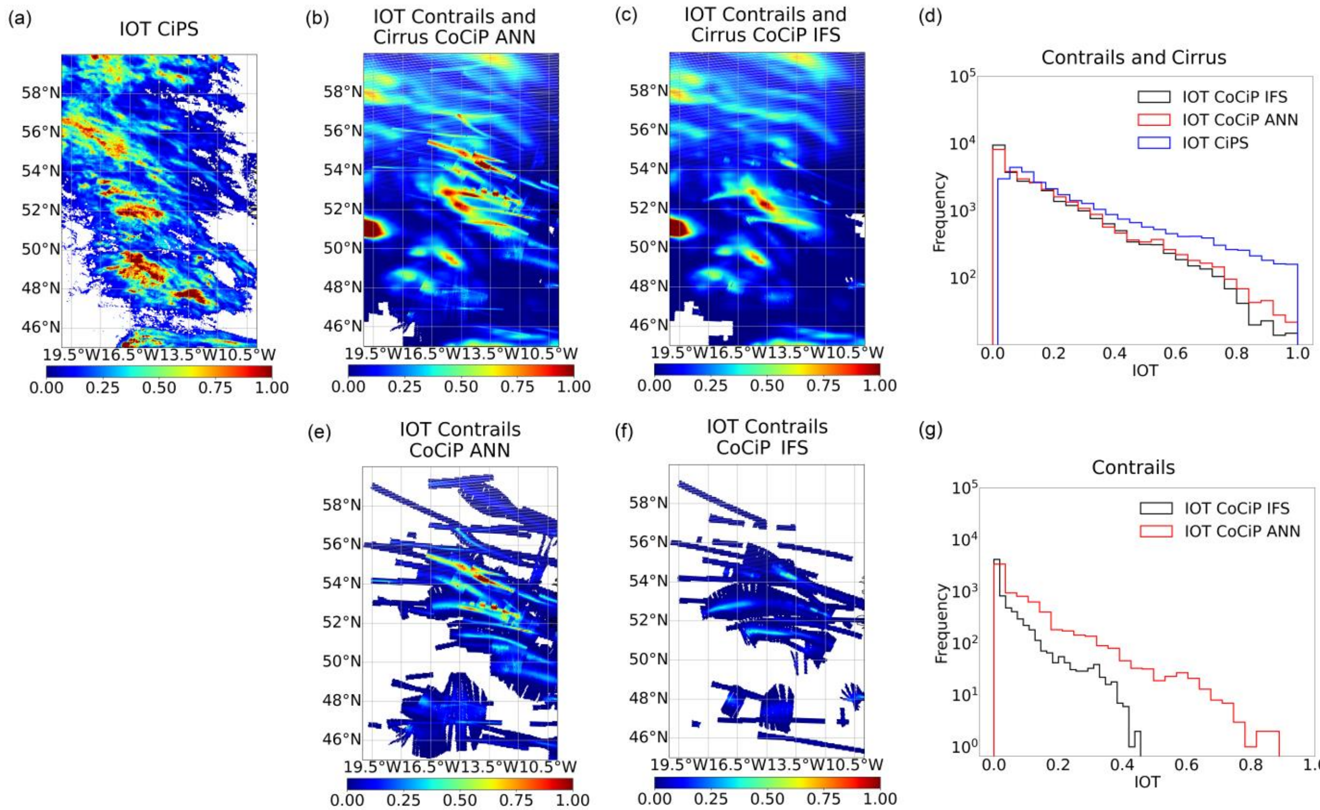

Figure A6 (7): Distributions of IOT for contrails and cirrus retrieved from (a) MSG observations using the CiPS algorithm, and simulated using the CoCiP model with (b) q_{ANN} or (c) q_{IFS} at 10:00 UTC on 14 April 2021. The IOT distribution for contrails from CoCiP simulations is shown in (e) and (f). The IOT frequencies (histograms) for contrail cirrus and contrails are shown in (d) and (g), respectively.

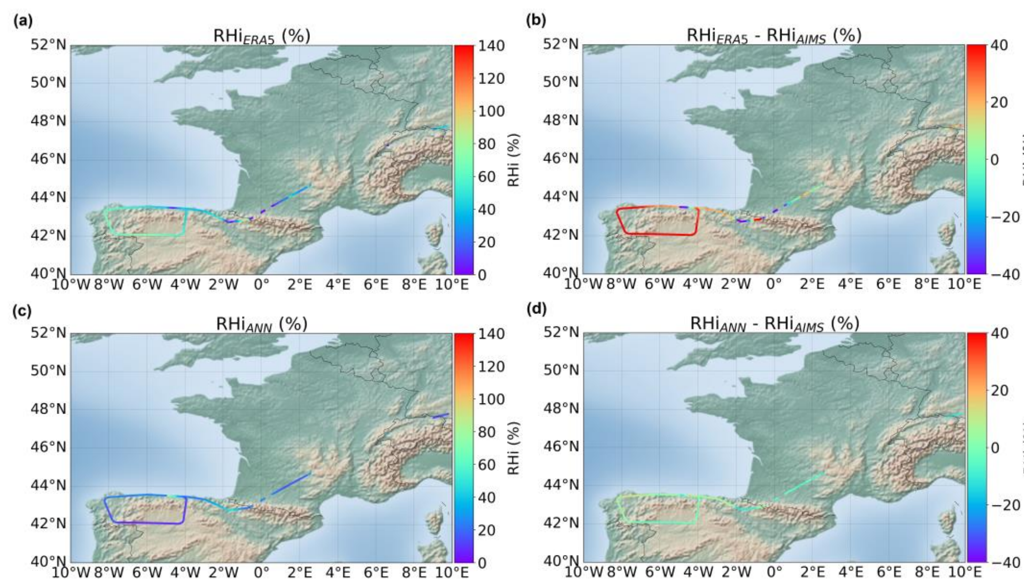

Figure A7 (S6): RHi derived from (a) ERA5 or (c) the ANN model and the differences relative to AIMS measurements in (b) and (d) obtained from the HALO aircraft on 21 July 2021 during the CIRRUS-HL campaign.

References

Diao, M., Zondlo, M. A., Heymsfield, A. J., Avallone, L. M., Paige, M. E., Beaton, S. P., Campos, T., and Rogers, D. C.: Cloud-scale ice-supersaturated regions spatially correlate with high water vapor heterogeneities, Atmos. Chem. Phys., 14, 2639–2656, https://doi.org/10.5194/acp-14-2639-2014, 2014.

Diao, M., Jensen, J. B., Pan, L. L., Homeyer, C. R., Honomichl, S., Bresch, J. F., and Bansemer, A.: Distributions of ice supersaturation and ice crystals from airborne observations in relation to upper tropospheric dynamical boundaries, J. Geophys. Res.-Atmos., 120, 5101–5121, https://doi.org/10.1002/2015JD023139, 2015.

Dotzek, N. and Gierens, K.: Instantaneous fluctuations of temperature and moisture in the upper troposphere and tropopause region. Part 2: Structure functions and intermittency, Meteorol. Z., 17, 323–337, https://doi.org/10.1127/0941-2948/2008/0292, 2008.

He, K., Zhang, X., Ren, S., and Sun, J.: Delving Deep into Rectifiers: Surpassing Human-Level Performance on ImageNet Classification, 2015 IEEE International Conference on Computer Vision (ICCV), Santiago, Chile, 1026-1034, https://doi.org/10.1109/ICCV.2015.123, 2015.

I have a couple of additional comments that would need to be addressed before the study is published.

RC1: First, I am surprised by the selection of the study region, as shown in Figure 1. Why doesn't it extend further west? Given that the network relies on the temporal evolution of humidity it seems it would make sense to include the regions where most of the humid regions are either formed or advected from. There is plenty of IAGOS data over the North Atlantic and Eastern US too.

R1a) author's response

Thank you for highlighting this aspect. In selecting the study region, we aimed to account for both horizontal advection and vertical air mass transport, as these factors contribute significantly to the inclusion of humid regions. The moisture formed and transported from lower atmospheric levels to cruise altitude is also represented in our analysis. The distribution of RHi_{IAGOS} is already included in Figure S2 of the supplement, showing that RHi density remains high, decreasing gradually from 100% and notably only as it approaches 150%, a trend consistent with findings in Wolf et al. (2023, Fig. 3) and Teoh et al. (2024, Fig. S4).

Our focus on the Eastern Atlantic, Europe, and Africa stems from the higher density of air traffic in these regions compared to others. This includes the concentrated morning eastbound and afternoon westbound transatlantic flights between the U.S. and Europe, which result in uneven sampling times over a 24-h period over the Atlantic Ocean (Schumann and Graf, 2012). Our regional selection enables us to establish more precisely the relationships between meteorological input variables and measured humidity across Europe, capturing temporal and regional variations effectively.

R1b) manuscript changes

L129-131: "We did not include the western regions of NAR, from where humid air is often advected, but instead focused on moisture originating and transported from lower atmospheric levels up to cruise altitude."

L136-138: “The distribution of RHi_{IAGOS} shows high-density values that gradually decrease starting from 110% and drop significantly as they approach 150% (see Fig. S2 in the supplement), a trend consistent with findings in Wolf et al. (2023, Fig. 3) and Teoh et al. (2024, Fig. S4).”

References

Schumann, U. and Graf, K.: Aviation-induced cirrus and radiation changes at diurnal timescales, J. Geophys. Res.-Atmos., 118, 2404–2421, https://doi.org/10.1002/jgrd.50184, 2013.

Teoh, R., Engberg, Z., Schumann, U., Voigt, C., Shapiro, M., Rohs, S., and Stettler, M. E. J.: Global aviation contrail climate effects from 2019 to 2021, Atmos. Chem. Phys., 24, 6071–6093, https://doi.org/10.5194/acp-24-6071-2024, 2024.

Wolf, K., Bellouin, N., Boucher, O., Rohs, S., and Li, Y.: Correction of temperature and relative humidity biases in ERA5 by bivariate quantile mapping: Implications for contrail classification, EGUsphere [preprint], https://doi.org/10.5194/egusphere-2023-2356, 2023.

RC2: Second, the lack of importance of the dynamical variables in explaining the prediction is surprising. The explanation proposed by the authors, that of a strong correlation between thermodynamical and dynamical variables, is plausible. But time scales are crucial in that correlation, so I wonder whether the study design somehow maximises the correlation. By the choice of the study region for example, which excludes the North Atlantic where dynamics might affect the evolution of humidity more clearly? Or by the choice of lead times? On that point, the question on temporal dependence asked in Section 3.1 on lines 227-228 is never really answered. How much does including distributions 6hr before current time improve the network prediction, for example?

R2a): author’s response

Thank you very much for your insightful questions and for raising the issue of the limited apparent importance of dynamical variables in the RHi prediction model. Based on further investigation, we now provide a clearer explanation of their impact.

1. Updated K_x values for thermodynamic and dynamical variables in RHi prediction

After ensuring the separation of training, validation, and test data on different days, we recalculated the relative contributions of variables to the new prediction of RHi (denoted as K_x, equation (1) in the manuscript, not necessarily less than 1). Following the suggestion from community comment #4, we set the investigated ERA5 feature to its mean value while keeping the other variables unchanged for the computation of K_x. The importance of the dynamical variables - vertical velocity (*w*), divergence (*d*), and especially horizontal wind components (*u*, *v*) and relative vorticity (*vo*) - has now increased.

For instance, the K_x values for *RHi*, temperature (*T*), and geopotential (*z*) remain high, at 1.85, 1.24, and 1.39. K_x values for *w* and *d* have now risen to 0.29 and 0.26, respectively, and for *u*, *v*, and *vo* to 0.98, 0.75, and 0.74. We acknowledge that thermodynamic variables (e.g., *RHi*, *T*) may inherently capture some of the dynamical trends. Hofer et al. (2024) also shows that the predictor RHi_{ERA5} has the greatest impact on humidity predictions, while the explanatory power of the dynamical proxies is insufficient when only using data from the current time and level. However, this updated analysis confirms that incorporating a broader vertical region and the historical (2 and 6h) times into the dynamical variables do indeed play a more prominent impact in the ANN model than previously thought, contributing to the understanding of humidity evolution. Revisions to the text and Figure A8 (updated Figure 3) are provided below.

2. Moisture uplift rather than advection from the North Atlantic

You rightly point out that our study region does not include the whole North Atlantic, where dynamical processes could have a clear influence on humidity evolution. As stated in response to your suggestion in RC1, we instead focused on moisture uplift from lower atmospheric levels rather than horizontal advection from the North Atlantic.

3. Temporal dependence of RHi on meteorological variables

As for the temporal dependence mentioned in Sect. 3.1, we now provide a more detailed analysis of the impact of past meteorological variables on RHi predictions. We have before calculated the Pearson correlation coefficients between thermodynamic and dynamical variable from ERA5 at various lead times (up to 24 hours prior) and measured RHi_{IAGOS} at the time and location of IAGOS data acquisition, but wasn't explained in the discussion version of the manuscript.

Based on our calculations, for this 6-hour lag range, the correlation between RHi_{ERA5} and RHi_{IAGOS} decreases by about 5.4% from 0.49 from 6 hours prior to the current time, while T_{ERA5} and z show constant significant correlations with RHi_{IAGOS} (around -0.5 and 0.4). *w* consistently demonstrates negative correlations with RHi, with an absolute correlation decreases of 86% from -0.11 from the 6-hour lag to the current time. For horizontal winds (*u*, *v*), correlations fluctuate around 0.34 and 0.44, while *d* exhibits the higher correlation at the 6-hour lag, decreasing from 0.18 to the current time by about 83%. In contrast, *vo* continues to exhibit negative correlations with RHi_{IAGOS}, with an increasing absolute correlation coefficient that approaches -0.2. These results demonstrate the correlation between current RHi and meteorological conditions at preceding times and have been added as the new Sect. S2 in the revised supplement.

4. Effect of including lead times on model performance

We have tried before that including meteorological distributions from 6h before current time slightly improves the performance of the RHi prediction model, as seen in a reduction of both MAE (from 2.31% to 2.21%) and RMSE (from 4.01% to 3.64%) in the validation test. Additionally, we assessed the influences of introducing different time lags (1h, 2h, 3h, 6h), and observed that the degree of the MAE and RMSE decrease increases with larger time lags. We then chose to include data from current time, 2h, and 6h intervals to balance prediction accuracy and computational efficiency. These results, along with supporting metrics, are presented in Table A1 below and Table S1 in the new Sect. S2 in the supplement.

5. Literature about airmass transport on humidity evolution from trajectory analysis

We have also referred to the trajectory analysis in Dyroff et al. (2015) to indicate that the RHi bias is linked to air masses from approximately 230 hPa in high northern latitudes, likely affected by both vertical intrusions and horizontal transport. Thank you for your feedback, we now add more detailed explanations mentioned above.

Reference

Hofer, S., Gierens, K., and Rohs, S.: How well can persistent contrails be predicted? An update, Atmos. Chem. Phys., 24, 7911–7925, https://doi.org/10.5194/acp-24-7911-2024, 2024.

R2b) manuscript changes

L18 in the Abstract: "while other dynamical variables are of low to moderate or high importance."

L321-323: “they provide a moderate and non-negligible contribution to the accuracy of the RHi prediction model. In fact, *w* and *d* show K_x of 0.29 and 0.26 and those for *u*, *v*, and *vo* even higher, which are 0.98, 0.75, and 0.74. There is generally less importance in the contributions of the variables representing dynamical quantities…”

L325-326: “The fact that dynamical variables for instance particularly u, v, vo are closely as important as RHi_{ERA5}, T_{ERA5}, and *z* for the description of the physical processes that lead to the decrease/increase of relative humidity in Sect. S2 in the supplement...”

L329-332: “Hofer et al. (2024) shows that RHi_{ERA5} is the most influential predictor for humidity predictions, while the explanatory power of dynamical proxies is insufficient when only using data from the current time and level. However, our updated analysis confirms that incorporating a broader vertical region and the historical time into the dynamical variables has a more significant impact on the ANN model and contributes to the understanding of humidity evolution.”

L483-484: “while other meteorological variables, including horizontal wind speed, relative vorticity, vertical velocity, and divergence, have a high or moderate to low but measurable influence.”

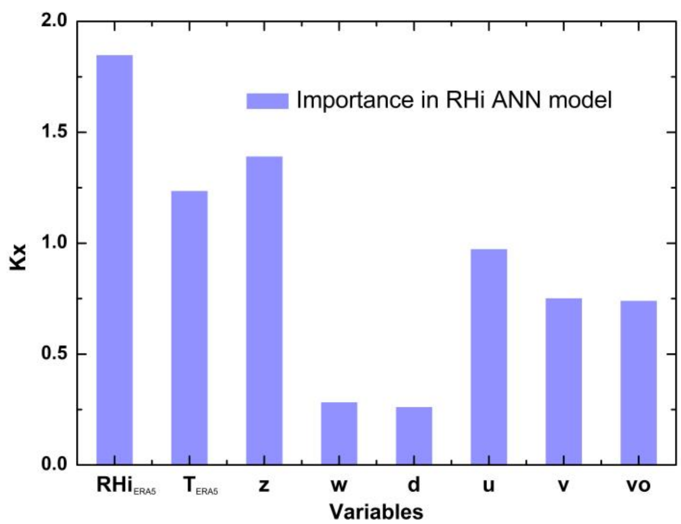

Figure A8 (3): Relative importance of the individual variables to the ANN model for predicting RHi.

L254-256: “Further computations of the Pearson correlation coefficient between RHi_{IAGOS} and temporal meteorological variables from ERA5, and the impact of including data distributions from hours before the current time on improving the network prediction, are explained in Sect. S2.”

New Sect. S2 in the supplement: “The correlation between RHi_{IAGOS} and ERA5 temporal meteorological variables”.

“We have determined the temporal dependence of measured RHi_{IAGOS} at the time and location of IAGOS data acquisition on meteorological variables at the preceding time up to 24 hour prior through the calculation of the Pearson correlation coefficient. Based on the calculations, compared to the time 6 hour before, the correlation of RHi_{ERA5} and RHi_{IAGOS} from 0.49 decreases by about 5.4% at the current time. The correlations for T_{ERA5} and *z* with RHi_{IAGOS} are also statistically significant and almost constant, with coefficients of about -0.5 and 0.4. *w* consistently demonstrates negative correlations with upward motion, resulting in cooling and an increase in RHi. The absolute correlation decreases from the 6-h time lags to the current time from -0.11 by about 86%. The correlation for *u* and *v* tends to fluctuate around 0.34 and 0.44. *d* generally exhibits positive correlations, with the highest value occurring around the 4-h to 5-h time lag, are 0.18 at the 6-h time lag higher than that of 0.03 at the current time by about 83%. In contrast, *vo* continues to exhibit negative correlations with RHi_{IAGOS}, with an increasing absolute correlation coefficient that approaches -0.2.

Including meteorological data from 6 hour prior improves the accuracy of the RHi prediction model on the validation dataset, reducing the MAE from 2.31% to 2.21% and the RMSE from 4.01% to 3.64%. The effect of time lags on model accuracy is calculated and presented in Table S1. As meteorological variables from 1, 2, 3, and 6 hours before the current time are introduced, the decrease in MAE and RMSE gradually becomes more significant. To balance information richness with computational efficiency, we choose the combination of current time, 2 hour, and 6 hour."

Table A1 (S1): Impact of including data distributions from 6 hours prior on network prediction accuracy.

Scenarios	MAE (%)	RMSE (%)	R^2
current	2.31	4.01	0.99
current, -1 h	2.21	4.17	0.98
current, -2 h	2.3	4.01	0.98
current, -3 h	2.33	4.01	0.98
current, -6 h	2.21	3.64	0.98
current, -2h, -6h	2.23	3.78	0.99

Other comments:

RC3: Lines 188-189: I am not sure that the RHi peak is always artificial. Sanogo et al. (2024) https://doi.org/10.5194/acp-24-5495-2024 suggest that the peak is seen in IAGOS in cloudy conditions. See their Figures 4 and 5

R3a): author's response

Thank you for pointing out this ambiguous description. The RHi peak reported by Sanogo et al. (2024) indeed reflects the maximum observed RHi values from IAGOS, as you suggested. Meanwhile, the RHi = 100% peak in ERA5 data mainly results from saturation adjustments inherent in numerical weather prediction models. We agree with your feedback and have revised "artificial" to "partially artificial" for clarity.

R3b) manuscript changes

L197-200: In addition, a "partially" artificial occurrence accumulation peak exists in the ERA5 data set at RHi_{ERA5} = 100%. "In RHi_{IAGOS}, a small peak is observed between 100% and 110% under cloudy conditions (Sanogo et al., 2024). However, much of the accumulation peak in the ERA5 data is attributed to the cloud saturation adjustment in NWP models".

We have also removed terms such as "artificial" when describing the occurrence peak in the following text.

Reference

Sanogo, S., Boucher, O., Bellouin, N., Borella, A., Wolf, K., and Rohs, S.: Variability in the properties of the distribution of the relative humidity with respect to ice: implications for contrail formation, Atmos. Chem. Phys., 24, 5495–5511, https://doi.org/10.5194/acp-24-5495-2024, 2024.

RC4: Lines 308-311: Can you clarify how that statement relates to the statement on correlation made earlier in the paragraph?

R4a) author's response

Thank you for highlighting this ambiguous description. The statement about connections between ANN layers relates only to the calculation of importance metrics and not to the correlation between different meteorological variables mentioned earlier in the paragraph. We have removed the statement you pointed out and expanded the discussion on the impacts of dynamical variables on RHi predictions, incorporating insights from the literature and our analysis. The text has been revised accordingly for clarity, and please refer to our responses to RC2 for further details.