

Prediction of Atmospheric Noise Temperature at the Deep Space Network with Machine Learning

Longtao Wu¹, David D Morabito¹, Joaquim Teixeira¹, Lei Huang², Hai Nguyen¹, Hui Su³,
Melissa Soriano¹, Lei Pan¹, and Daniel S. Kahan¹

¹Jet Propulsion Laboratory

²California Institute of Technology

³Jet Propulsion Laboratory, California Institute of Technology

November 24, 2022

Abstract

Ka-band (32 GHz) communications links utilized by the National Aeronautics and Space Administration (NASA) flight missions for science downlink are susceptible to degradation due to weather. In this study, a customized real-time forecast system has been developed to predict zenith atmospheric noise temperature (Tatm) at the Deep Space Network (DSN) tracking sites using machine learning (ML). A random forest model is trained with the Global Forecast System (GFS) forecast and analysis datasets in addition to the Tatm measurements derived from on-site advanced water vapor radiometers (AWVR). The real-time forecast uncertainty is quantified for different error regimes using the Self-Organizing Map method. The results show that the Root Mean Square Error (RMSE) of the 24-hour Tatm prediction at Goldstone, CA increases with the increase of Tatm. Ninety percent of the forecasts have RMSE (bias) of less than 3.50 K (0.22 K) for fair-weather conditions with Tatm < 17 K. In comparison to the current approach in designing Ka-band communications links, application of weather forecasts can increase data return to the downlink for 80% of the time. A downlink gain of up to 1.61 dB (45% more data) can be realized at 20° elevation angle when Tatm = 9 K.

26 **Abstract**

27 Ka-band (32 GHz) communications links utilized by the National Aeronautics and Space
28 Administration (NASA) flight missions for science downlink are susceptible to degradation due
29 to weather. In this study, a customized real-time forecast system has been developed to predict
30 zenith atmospheric noise temperature (T_{atm}) at the Deep Space Network (DSN) tracking sites using
31 machine learning (ML). A random forest model is trained with the Global Forecast System (GFS)
32 forecast and analysis datasets in addition to the T_{atm} measurements derived from on-site advanced
33 water vapor radiometers (AWVR). The real-time forecast uncertainty is quantified for different
34 error regimes using the Self-Organizing Map method.

35 The results show that the Root Mean Square Error (RMSE) of the 24-hour T_{atm} prediction
36 at Goldstone, CA increases with the increase of T_{atm} . Ninety percent of the forecasts have RMSE
37 (bias) of less than 3.50 K (0.22 K) for fair-weather conditions with $T_{atm} < 17$ K. In comparison to
38 the current approach in designing Ka-band communications links, application of weather forecasts
39 can increase data return to the downlink for 80% of the time. A downlink gain of up to 1.61 dB
40 (45% more data) can be realized at 20° elevation angle when $T_{atm} = 9$ K.

41 1 Introduction

42 Ka-band (32 GHz) communications links utilized by the National Aeronautics and Space
43 Administration (NASA) deep space flight missions for science downlink are more susceptible to
44 degradation due to weather than lower frequency S-band (2.3 GHz) and X-band (8.4 GHz) (Slobin,
45 2009). The current approach in designing Ka-band communications links employs conservative
46 assumptions, such as designing at the lowest elevation angle of a tracking pass and with high
47 probability of weather constraints (Davarian et al., 2004). For example, the Europa Clipper mission
48 assumes a 3 dB margin for the data downlink. Such assumptions can result in inefficient downlink
49 capacity. By making use of real-time weather forecasting one can greatly increase data return
50 efficiency and the reliability of the communications links (Shambayati, 2004; Biscarini et al.,
51 2016; Morabito et al., 2016; Nikoukar et al., 2020).

52 Weather service agencies, such as the National Centers for Environmental Prediction
53 (NCEP), provide operational weather forecasts globally using numerical weather prediction
54 (NWP) models. Raw NWP output, however, is not adequate to resolve the fine-scale topographical
55 features and mesoscale processes that govern local weather at specific locations, such as at the
56 Deep Space Network (DSN) tracking sites. In addition, the NWP raw output contains biases
57 relative to local observations. Morabito et al. (2016) showed that the NWP model captures local
58 weather conditions at the DSN tracking sites most of time but significant discrepancies occur due
59 to cloud and precipitation processes.

60 A variety of (dynamical or statistical) downscaling methods have been developed to
61 provide robust weather predictions at local scales for various applications (Barsugli et al., 2013).
62 Dynamical downscaling extrapolates the effects of large-scale processes to regional or local scales
63 of interest using high-resolution NWP models. Physical laws are explicitly represented in the NWP
64 models. However, dynamical downscaling can be sensitive to biases in the boundary and initial
65 conditions from large-scale coarse-resolution models. Dynamical downscaling can be
66 computationally expensive, which is not practical for use in mission operations such as for
67 command and on-board control. Statistical downscaling determines relationships between large-
68 scale patterns simulated by the global NWP models and responses on the local scale using
69 statistics-based techniques, including Machine Learning (ML) methods. These relationships are
70 applied to the NWP outputs and are transformed into statistical inferences of local responses with
71 bias corrections. Statistical downscaling requires a relatively low number of computing resources
72 and can be readily applied to real-time or near real-time space mission operations.

73 In this study, a customized weather forecast system with uncertainty quantification (UQ)
74 is developed for the DSN tracking sites using a ML method. The workflow of the ML forecast
75 system is shown in Figure 1. We use global NWP datasets and in-situ observations to train the
76 ML-forecast model. A UQ model is also trained to characterize the bias and standard error of the
77 real-time forecasts. Thus, real-time weather forecasts with error characteristics can be provided for
78 the DSN tracking sites. More details of the real-time ML forecast system will be described in the
79 following sections. We use a 24-hour forecast of zenith atmospheric noise temperature (T_{atm}) at
80 Goldstone, CA as a demonstration case of the ML forecast system. This ML forecast system will
81 be adopted to other tracking sites and expanded to predict other atmospheric variables when in-
82 situ observations become available.

83 Section 2 describes the datasets used for the training and testing of the ML forecast system.
84 The ML forecast model and the forecast results are presented in section 3. Section 4 presents the

85 UQ model and results. Section 5 discusses the assessments of the ML forecasts used to quantify
86 improvements in Ka-band downlink efficiency for the Europa Clipper mission. Section 6 provides
87 summary and discussion.

88 **2 Training and testing datasets**

89 2.1 Atmospheric noise temperature

90 This study aims to apply ML to NWP model forecast data and generate real-time forecasts
91 of zenith T_{atm} at the DSN tracking sites, which can be ingested into the DSN communication link
92 protocol. The increase of T_{atm} can be used to assess other weather-induced effects such as
93 attenuation due to gas (e.g. water vapor and oxygen), cloud liquid and rain liquid (Slobin, 2009;
94 Morabito, 2014). The Advanced Water Vapor Radiometer (AWVR) has been used to measure
95 weather effects at the DSN Goldstone tracking site for over 20 years and at the Madrid site for ~11
96 years. The AWVRs were primarily developed to remove atmospheric effects from radio science
97 data for the Cassini mission and are currently being used for the Juno mission (see references in
98 Morabito et al., 2022). T_{atm} and precipitable water vapor (PWV) can be derived from sky brightness
99 temperatures observed by the AWVR (Slobin, 2005; Morabito, 2007; Morabito et al., 2015; 2016;
100 2022). The AWVR data acquired in 2019 encountered problems related to inadequate temperature
101 control (Morabito et al., 2022), thus these data are not used in this study.

102 We use hourly-averaged T_{atm} from 2015 to 2020 (except 2019) as observations to train and
103 test the ML forecast system. The AWVR observations (Figure 2) show that the minimum value of
104 T_{atm} at Goldstone is 7.15 K. The maximum value of T_{atm} can be up to 132.10 K (or higher), mainly
105 due to the impacts of cloud liquid and rain precipitation. 90% (95%) of the T_{atm} observations are
106 less than 16.50 K (19.11 K).

107 2.2 NWP analysis and forecast

108 NCEP Global Forecast System (GFS) provides global analysis and forecasts (up to 16 days)
109 of the weather at 0.25° horizontal resolution every six hours (NCEP, 2015). The GFS data archive
110 is available online with data spanning from 2015/01/15 to the present. Data samples numbering
111 5435 from both the GFS and AWVR are available to train and test the ML forecast model. Both
112 GFS analyses and forecasts are used as predictors, in order to retain information from current
113 observations and physical model forecasts.

114 We analyze the correlation of $\ln(T_{atm})$ at 24-hours after the forecasting time (T24) with
115 the coarse-grid-scale variables from both GFS analysis and 24-hour forecasts at the forecasting
116 time (T0). Figure 3 shows that GFS 24-hour forecast (PWV_f024) nicely predicts the magnitude
117 and variation of PWV observed at Goldstone (PWV_AWVR) for year 2017. The variation of
118 $\ln(T_{atm})$ is following the variation of PWV most of the time, with a correlation coefficient of 0.71
119 between the $\ln(T_{atm})$ at T24 and the 24-hour forecasted PWV at T0. In this study, those GFS
120 variables with correlation > 0.30 are selected as the predictors to forecast $\ln(T_{atm})$ at T24 in the
121 ML forecast model (Table 1). In general, GFS 24-hour forecasts have better correlation with T_{atm}
122 than GFS analysis, because GFS 24-hour forecasts are valid at the same time of T_{atm} while GFS
123 analysis is valid at 24 hours before T_{atm} . All of the moisture, cloud, precipitation and surface
124 meteorological data from the 24-hour forecasts provide useful information to predict T_{atm} . Only
125 PWV and surface meteorological data from analysis have correlation with $\ln(T_{atm})$ that are larger

126 than 0.30. This suggests that including NWP forecasts as predictors will provide better forecast
127 skills than using in-situ measurements alone.

128 **3 Machine learning forecast**

129 Random forest (RF) regressor in the standard Python scikit-learn package ([https://scikit-
130 learn.org/stable/modules/generated/sklearn.ensemble.RandomForestRegressor.html](https://scikit-learn.org/stable/modules/generated/sklearn.ensemble.RandomForestRegressor.html)) is used to
131 build the ML-based forecast system. Historical AWVR and GFS data (a total of 5,435 samples)
132 from 2015 to 2020, except 2019, are used for model training and testing. Pipeline in the scikit-
133 learn package (<https://scikit-learn.org/stable/modules/generated/sklearn.pipeline.Pipeline.html>) is
134 applied for the standardization of the predictors. Hyperparameters are optimized by a grid search
135 with the 10-fold cross-validation method ([https://scikit-
136 learn.org/stable/modules/generated/sklearn.model_selection.GridSearchCV.html](https://scikit-learn.org/stable/modules/generated/sklearn.model_selection.GridSearchCV.html)). Model
137 performance is evaluated by inspecting the root mean square error (RMSE) of T_{atm} relative to the
138 AWVR measurements. We evaluate the model performance using the “leave-one-year-out” cross
139 validation method (Kaplan et al., 2015; Su et al., 2020), which ensures the independence of the
140 training and testing data. That is, for a single model, four years’ worth of data are used to train the
141 ML forecast model and the withheld year is used to evaluate the forecast performance.

142 The forecasted T_{atm} has a similar distribution to that of the AWVR observations (Figure 2).
143 However, the data range in the forecast is smaller than that of the observations. The minimum
144 (maximum) value is 7.80 K (43.19 K) in the forecast vs. 7.15 K (132.10 K) in the AWVR
145 observation. Figure 4 shows that the forecasted T_{atm} has good agreement with the observations,
146 especially when the observed T_{atm} is less than 17 K. The RMSE (bias) of all of the forecasts is 4.72
147 (−0.27) K. Ninety percent of the forecast T_{atm} has a bias within 0.25 K (Figure 5a). The RMSE for
148 $T_{atm} < 9.50$ K (20% of the forecasted T_{atm}) is 0.63 K (Figure 5b). The RMSE increases with the
149 increase of T_{atm} . Ninety percent of the forecasts has a RMSE less than 3.50 K (or 25% relative to
150 the mean T_{atm}) for fair-weather conditions with $T_{atm} < 17$ K. The forecast RMSE increases
151 significantly for extreme-weather conditions with $T_{atm} > 17$ K.

152 The PWV_f024 from GFS is the most important predictor of T_{atm} in the RF forecast model
153 (Figure 6), consistent with the highest correlation between PWV_f024 and $\ln(T_{atm})$ (Table 1).
154 The 24-hour forecasts from GFS are more important than the analysis from GFS in the RF forecast
155 model, consistent with the correlations in Table 1. The relative importance of predictors are similar
156 in different years with slight differences in the magnitude of the relative importance.

157 **4 Uncertainty quantification (UQ)**

158 Real-time error characterization is provided by a UQ methodology similar to that in
159 Teixeira et al. (2021), wherein a clustering algorithm is used to identify regimes of prediction error.
160 Teixeira et al. (2021) showed that error characterization of a geophysical prediction of atmospheric
161 motion vectors benefitted from separating those predictions into different geophysical regimes
162 because the error characteristics of the predictions can vary significantly between the regimes. The
163 cursory analysis (Figure 5) shows that the bias and RMSE of the RF prediction is significantly
164 greater for predictions of $T_{atm} > 17$ K. We enhance upon this simple observation by clustering with
165 predictor variables in addition to the predictions themselves in the UQ analysis.

166 A schematic of our UQ approach is detailed in Figure 7. We denote the predictors of the
167 RF model as X , the RF prediction as \hat{Y} , and the observed T_{atm} as Y . We characterize the error as

168 both bias (the mean of $[\hat{Y} - Y]$) and standard error (the standard deviation of $[\hat{Y} - Y]$). After
 169 randomly splitting the dataset into a training (75%) and a testing (25%) set, we first train a
 170 clustering algorithm (a self-organizing map [SOM], as will be discussed in a subsequent
 171 paragraph) on the set of predictors X , predictions \hat{Y} , and observations Y from the training dataset
 172 (Panel 1 in Figure 7). Because it includes both the RF prediction and the observed T_{atm} , the
 173 clustering algorithm captures regimes of prediction error better than it would if the observed T_{atm}
 174 were not included. Of course, this implies a trade-off in which the clustering algorithm itself cannot
 175 be used operationally, since we have no knowledge of the actual T_{atm} in real applications; instead,
 176 we develop a classification algorithm (specifically a classification RF) which maps each set of
 177 predictors and prediction (X and \hat{Y}) to its assigned cluster (Panel 2 in Figure 7). As such, any RF
 178 prediction and its associated predictors can be assigned a set of probabilities of belonging to each
 179 cluster. Subsequently, the bias and standard error for a particular RF prediction is the weighted
 180 mean of the biases and standard errors from each of the clusters (when applied to the testing set),
 181 where each value is weighted by the probability mapped from the RF prediction to that respective
 182 cluster. Because the clustering approach can be sensitive to the initial conditions of the training
 183 dataset, we run an ensemble of 100 identical models with different randomly-sample training
 184 datasets. The final model is the mean across the ensemble of predicted bias and standard error
 185 values.

186 We employed a self-organizing map in the clustering portion of the error characterization
 187 model. SOM is an unsupervised learning technique in which a neural net fits a grid of neurons to
 188 the topological shape of a dataset. We choose SOM due to their repeated success in clustering
 189 atmospheric data into identifiable and interpretable geophysical regimes (Marques and Chen,
 190 2003; Richardson et al., 2003; Liu et al., 2006). As in Teixeira et al. (2021), we validated our
 191 trained model using the Continuous Rank Probability Score (CRPS), a scoring rule for
 192 probabilistic forecasts. CRPS is as a function of a cumulative distribution function F and an
 193 observation \mathbf{x} as follows:

$$194 \quad 195 \quad \text{CRPS}(F, \mathbf{x}) = \int_{-\infty}^{\infty} (F(\mathbf{y}) - H(\mathbf{y} - \mathbf{x}))^2 d\mathbf{y}$$

196 where $H()$ is the Heaviside step function (Gneiting and Katzfuss, 2014). In this context, F is the
 197 cumulative distribution function associated with the bias and standard error that the UQ model
 198 provides, while \mathbf{x} is a scalar observation of the difference between T_{atm} and the model prediction.
 199 In essence, CRPS attempts to measure how close an observation is to the center of a distribution,
 200 while penalizing distributions that are unnecessarily wide. We analyze the mean CRPS value for
 201 the test population of T_{atm} errors, with a lower CRPS value reflecting a more calibrated
 202 probabilistic forecast. Conveniently, the CRPS also serve as a metric for selecting the optimal
 203 number of clusters for SOM. Figure 8 shows CRPS values evaluated under different choices of
 204 cluster number compared against the CRPS value for a naïve model, which is equivalent to a
 205 single-regime model in which the error distributions are given by the population mean and
 206 population standard deviation of T_{atm} . The figure illustrates two key observations. Firstly, for all
 207 SOMs, the CRPS of forecasts is significantly lower than that of the naïve model; secondly, that as
 208 the number of clusters increases, the CRPS decreases in an inverse-logarithmic fashion.

210 The second observation forms the basis for our choice of the cluster number. Our final
 211 model consists of SOM with a 30×30 feature grid which clusters the training data into 50 distinct

212 regimes. The UQ values applied on the testing dataset are illustrated in Figure 9. For 80% of the
 213 data, the estimated bias ranges between with -1.00 K and 1.76 K. However, for the 10% of data
 214 lying in the extremes, the estimated bias can range as low as -11.30 K and as high as 38.32 K. The
 215 RMSE estimates are similarly skewed: 50% of the data has an estimated RMSE estimate between
 216 0.67 K and 1.99 K; a further 30% range between 2.00 K and 3.28 K; and the highest 20%, however,
 217 ranges between 3.28 K and 12.59 K.

218 We note that the results of the UQ analysis a bit different from that of Figure 5. Specifically,
 219 Figure 5 indicates that the first 40th percentile of data have a RMSE < 1.00 K, while the UQ result,
 220 as illustrated by the right-most panel of Figure 9, predicts that 40th percentile of data have a RMSE
 221 < 1.76 K. We note that this is because Figure 5 is derived from validation versus withheld data,
 222 and in general we would not know the prediction uncertainty of an estimate that arises from the
 223 forecast T_{atm} from the ML model in section 3. This gap is filled by the methodology in section 4,
 224 and the corresponding analog to Figure 5 is illustrated in the right panel of Figure 9. In general,
 225 the UQ methodology tends to overestimate the uncertainty of low T_{atm} percentiles, while
 226 underestimating the uncertainty of high T_{atm} percentiles. However, the similar upward convex
 227 shapes of the two plots and their similar magnitudes are encouraging, especially since estimation
 228 of uncertainties (second moments) tend to be much more difficult than estimation of the means
 229 (first moments).

230 **5 Link budget analysis**

231 The link budget analysis used by flight projects and telecom engineers makes use of the
 232 received signal-to-noise (SNR) as a key parameter related to data rate (Yuen, 1983). Received
 233 SNR is related to the received signal power divided by the observed noise power. The downlink
 234 signal power consists of contributions from the spacecraft transmit power, antenna gain, and
 235 various losses encountered as the signal propagates down to the receiving antenna, followed by
 236 contributions such as the gain of the receiving antenna. One of the loss contributors is due to the
 237 atmosphere, consisting primarily of the atmospheric attenuation due to oxygen and water vapor
 238 absorption, as well as liquid water such as in clouds and rain. Observed noise power includes the
 239 contribution due to the atmosphere, which is related to the elevation angle of the antenna, the
 240 optical depth of the atmosphere at zenith, and the effective radiating temperature of the
 241 atmosphere.

242 Telecom engineers typically use two types of SNR to analyze a link: carrier SNR and data
 243 channel SNR (we neglect ranging). For all link budgets, we ensure that the carrier SNR is adequate
 244 for signal lock. The data channel SNR is converted to the bit SNR (E_b/N_o) using the data rate. For
 245 a given elevation angle, we can define the downlink gain from the application of weather forecasts
 246 (forecast gain) as being the difference between the E_b/N_o using the forecast-provided zenith T_{atm}
 247 (adjusted to the elevation angle of interest) and the E_b/N_o using the nominal DSN
 248 Telecommunications Link Design handbook, 810-005 (Slobin, 2009) provided zenith value at 90%
 249 weather availability (and then adjusted to the elevation angle of interest).

250 The link budget approach made use of nominal Europa Clipper link budget parameters at
 251 Ka-band taken from the Europa Clipper Telecommunications Design Control Document (Babuscia
 252 et al., 2020). A link budget was first generated to verify agreement with Europa Clipper project
 253 Ka-band downlink link budget from Babuscia et al. (2020), which assumed a range distance of 6.4
 254 AU, an elevation angle of 10° , no Jupiter hot-body noise and the Canberra, Australia DSN tracking
 255 site.

256 Next, the link budget was run to reference Goldstone as the downlink site using a Ka-band
257 (32 GHz) capable downlink 34-m diameter antenna. The links were run with both the DSN
258 Telecommunications Link Design handbook's assumptions and weather forecast assumptions
259 provided by the forecasting results. We assumed no ranging in any of the links, just a carrier channel
260 and a downlink data channel with Binary Phase-shift Keying (BPSK) modulation. The link budgets
261 were run for elevation angles of 90°, 60°, 40° and 20° with zenith $T_{atm} \pm \sigma$ values of 8.96 ± 0.63 K,
262 10.02 ± 0.95 K, 11.10 ± 1.98 K, 12.75 ± 2.99 K, 14.57 ± 2.35 K, 15.76 ± 3.30 K, 17.57 ± 9.78 K
263 and 22.55 ± 16.48 K. The results were examined and provided to the forecasting system.

264 The results of forecast gain versus mean T_{atm} is shown in Figure 10 for each elevation angle
265 case. At the low-end of $T_{atm} = 8.98$ K, one can realize a forecast gain of 1.61 dB (45% more data)
266 at 20° elevation angle with an uncertainty of 0.16 dB. This forecast gain is significant and would
267 result in higher data return for missions with Ka-band telecommunication links. The forecast gain
268 decreases with the increase of T_{atm} . At 80% of the time, higher data return can be realized when
269 weather forecasting is used in telecommunication operations in place of using 90% weather
270 availability in the link assumptions. All of the curves intersect at 0 dB for 15.76 K, as this is the
271 90% weather availability point to which projects nominally design. Below this point at the high-
272 end of T_{atm} , most of the curves lie above the typical telecom link analysis margin point (-3 dB),
273 implying that using 3 dB margin is adequate here. If a project opts to, they can lower the data rate
274 appropriately during these adverse weather conditions to ensure lock-up with added margin.

275 **6 Summary and discussion**

276 In this study, we have developed a customized real-time weather forecast system for the
277 DSN tracking sites using a ML method. A RF forecast model is trained using the global GFS
278 forecast and analysis datasets and the T_{atm} derived from the AWVR measurements at Goldstone,
279 CA. Twenty-four hour forecasts of T_{atm} are provided at Goldstone, CA. Error characterization of
280 the forecasts is calculated using a SOM clustering and a RF classification method. Downlink gain
281 of the telecommunications links is estimated by comparing the differences between using the real-
282 time weather forecasts and using the 90% weather availability in a telecommunication link budget
283 model.

284 The forecasts have a RMSE of 0.63 K for $T_{atm} < 9.50$ K. Forecast error increases when the
285 T_{atm} increases. Ninety percent of the forecasts have a RMSE < 3.50 K and a bias within 0.25 K.
286 The RMSE for all the forecasts is 4.72 K. The UQ methodology allows real-time ML forecasts to
287 have unique RMSE values and bias estimates; 90% of forecasts will have an estimated RMSE $<$
288 4.49 K and an estimated bias within ± 2 K. Forecast gain can be as high as 1.61 dB (45% more data
289 downlink) when the forecasted T_{atm} is at the low-end at 20 deg elevation. Higher data return can
290 be realized for 80% of the time if the real-time weather forecast is applied in flight project
291 operations.

292 The real-time ML forecast system can provide forecasts of up to 16 days ahead for the DSN
293 tracking sites. Real-time commands would be used to change data rates downlinked by the
294 spacecraft for short time span forecasts (~6 hours to 2 days, which depends on round-trip light
295 time and project operations). Forecasts over a longer time span may be used in the preparation of
296 upcoming tracks such as defining data rate profiles in command loads that are uplinked to the
297 spacecraft. For projects that uplink command loads on reasonably short periods such as within two
298 weeks, these forecasts can be used to provide canned-in downlink rates to the uplink command
299 loads. Real time commands would then be used to override these in case of significant weather

300 changes from the 16-day forecasts versus the 1-2 day forecasts. This ML forecast system will be
301 adopted to other tracking sites and expanded to predict other atmospheric variables in future
302 studies.

303 With enhancement of knowledge and advancement of technology, technical issues for deep
304 space missions have become increasingly more complex and comprehensive. Larger volumes and
305 more diverse types of data are and will be collected. As the capacity of on-board data storage
306 becomes stressed, efficient transmission of data to Earth in a timely fashion is desired. The ML
307 model for predicting T_{atm} can be generalized to many other missions in which data communications
308 are essential. It could serve as a component of future onboard data prioritization protocol.

309 **Acknowledgments**

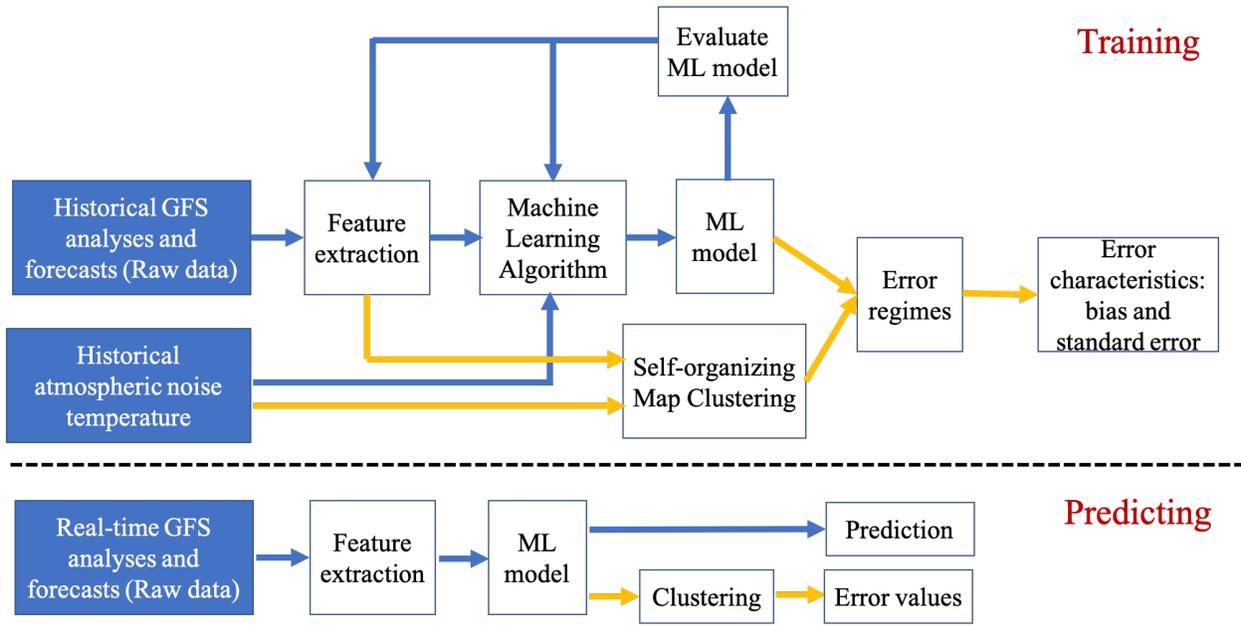
310 We would like to acknowledge Alan Tanner and Stephen Keihm for their methodological
311 recommendations for calibration of the AWVR data. The research was carried out at the Jet
312 Propulsion Laboratory, California Institute of Technology, under a contract with the National
313 Aeronautics and Space Administration (80NM0018D0004).

314 **References**

- 315 Babuscia, A., M. Soriano, and P. Illott (2020), Europa Clipper Telecommunications Design
316 Control Document, Revision A, JPL D-101604, Jet Propulsion Laboratory, Pasadena, California,
317 October 22, 2020.
- 318 Barsugli, J. J., Guentchev, G., Horton, R. M., Wood, A., Mearns, L. O., Liang, X.-Z., Winkler, J.
319 A., Dixon, K., Hayhoe, K., Rood, R. B., Goddard, L., Ray, A., Buja, L., & Ammann,
320 C. (2013). The practitioner's dilemma: How to assess the credibility of downscaled climate
321 projections. *Eos, Transactions American Geophysical Union*, 94(46), 424–
322 425. <https://doi.org/10.1002/2013EO460005>
- 323 Biscarini, M. et al. (2016), Optimizing data volume return for Ka-band deep space links
324 exploiting short-term radiometeorological model forecast, *IEEE Trans. Antennas Propag.*, vol.
325 64, no. 1, pp. 235-250.
- 326 Gneiting, T. and M. Katzfuss (2014), Probabilistic forecasting, *Annu. Rev. Stat. Appl.*, 1, 125–
327 151.
- 328 Kaplan, J., Rozoff, C. M., DeMaria, M., Sampson, C. R., Kossin, J. P., Velden, C. S., et al.
329 (2015). Evaluating environmental impacts on tropical cyclone rapid intensification predictability
330 utilizing statistical models. *Weather and Forecasting*, 30, 1374–1396.
331 <https://doi.org/10.1175/WAF-D-15-0032.1>
- 332 Liu, Y., R. H. Weisberg and R. He (2006), Sea surface temperature patterns on the West Florida
333 Shelf using the growing hierarchical self-organizing maps. *Journal of Atmospheric and Oceanic*
334 *Technology*, Vol. No. 2, 23, 325-328, ISSN 0739-0572.
- 335 Marques, N. C. and N. Chen (2003), Border detection on remote sensing satellite data using self-
336 organizing maps, *Proceedings of EPIA 2003 – 11th Portuguese conference on Artificial*
337 *Intelligence*, pp. 294-307, ISBN 3-540-20589-6, Beja, Portugal, December 2003, Springer-
338 Verlag, Berlin.
- 339 Morabito, D. D. (2014), A Comparison of Estimates of Atmospheric Effects on Signal
340 Propagation Using ITU Models: Initial Study Results, *The Interplanetary Network Progress*
341 *Report*, vol. 42-199, pp. 1–24. http://ipnpr.jpl.nasa.gov/progress_report/42-199/199D.pdf
- 342 Morabito, D. D., S. Keihm, and S. Slobin (2015), A Statistical Comparison of Meteorological
343 Data Types Derived from Deep Space Network Water Vapor Radiometers, *The Interplanetary*
344 *Network Progress Report*, vol. 42-203, pp. 1–21. http://ipnpr.jpl.nasa.gov/progress_report/42-203/203A.pdf
- 346 Morabito, D. D., L. Wu, and S. Slobin (2016), Weather Forecasting for Ka-band Operations:
347 Initial Study Results, *The Interplanetary Network Progress Report*, vol. 42-206, pp. 1–24.
348 https://ipnpr.jpl.nasa.gov/progress_report/42-206/206C.pdf
- 349 Morabito, D. D., D. Kahan, M. Paik, L. Wu, E. Barbini, D. Buccino and M. Parisi (2022), A
350 Study of Twenty Years of Advanced Water Vapor Radiometer Data at Goldstone, California,
351 *The Interplanetary Network Progress Report*, vol. 42-228, pp. 1-18.
352 https://ipnpr.jpl.nasa.gov/progress_report/42-228/42-228A.pdf

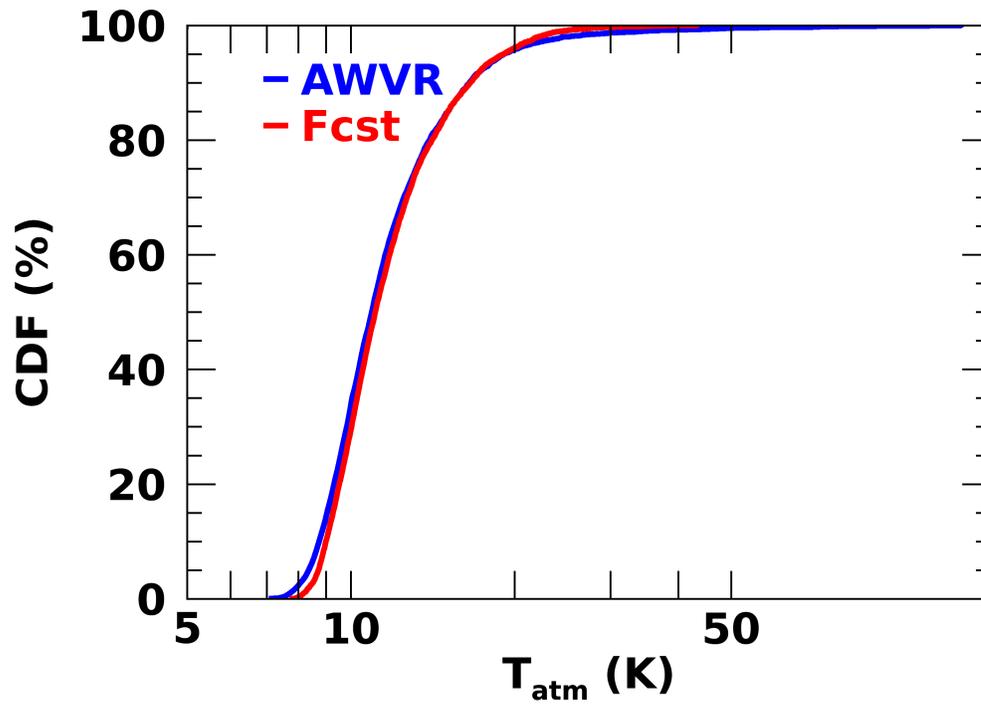
- 353 Nikoukar, R., D. Copeland, S. Sprouse, M. Cox and K. Kufahl (2020), Quantifying Weather
354 Effects on Ka-band Communication Links: A Parker Solar Probe Study, *2020 IEEE Aerospace*
355 *Conference*, 2020, pp. 1-9, doi: 10.1109/AERO47225.2020.9172786.
- 356 National Centers for Environmental Prediction/National Weather Service/NOAA/U.S.
357 Department of Commerce. 2015, updated daily. NCEP GFS 0.25 Degree Global Forecast Grids
358 Historical Archive. Research Data Archive at the National Center for Atmospheric Research,
359 Computational and Information Systems Laboratory. <https://doi.org/10.5065/D65D8PWK>.
360 Accessed 29 Mar 2022.
- 361 Richardson, A. J., C. Risien and F. A. Shillington (2003), Using self-organizing maps to identify
362 patterns in satellite imagery, *Progress in Oceanography*, Vol. 59, No. 2-3, 223- 239, ISSN 0079-
363 6611.
- 364 Shambayati, S. (2004), On the Benefits of Short-Term Weather Forecasting for Ka-band (32
365 GHz), *Proceedings IEEE Aerospace Conference*, 3, 1498.
- 366 Slobin, S. D. (2009), Atmospheric and Environmental Effects, *DSN Telecommunications Link*
367 *Design Handbook*, DSN No. 810-005, Space Link Interfaces, Module 105, Rev. D,
368 <http://deepspace.jpl.nasa.gov/dsndocs/810-005/105/105D.pdf>
- 369 Su, H., L. Wu, J. H. Jiang, R. Pai, A. Liu, A. J. Zhai, P. Tavallali and M. DeMaria (2020),
370 Applying Satellite Observations of Tropical Cyclone Internal Structures to Rapid Intensification
371 Forecast with Machine Learning, *Geophys. Res. Lett.*, <https://doi.org/10.1029/2020GL089102>.
- 372 Teixeira, J. V., H. Nguyen, D. J. Posselt, H. Su and L. Wu (2021), Using machine learning to
373 model uncertainty for water vapor atmospheric motion vectors, *Atmos. Meas. Tech.*, 14, 1941–
374 1957, <https://doi.org/10.5194/amt-14-1941-2021>.
- 375 Yuen, J. H. (Ed.) (1983). *Deep Space Telecommunications Systems Engineering*. New York:
376 Plenum Press.

377 **List of figures**



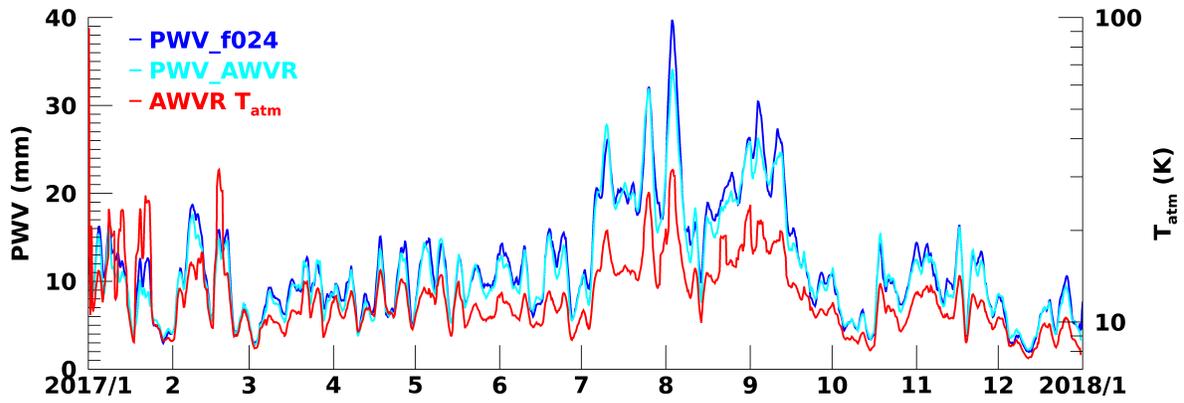
378

379 **Figure 1.** Workflow of the Machine Learning-based forecast system.



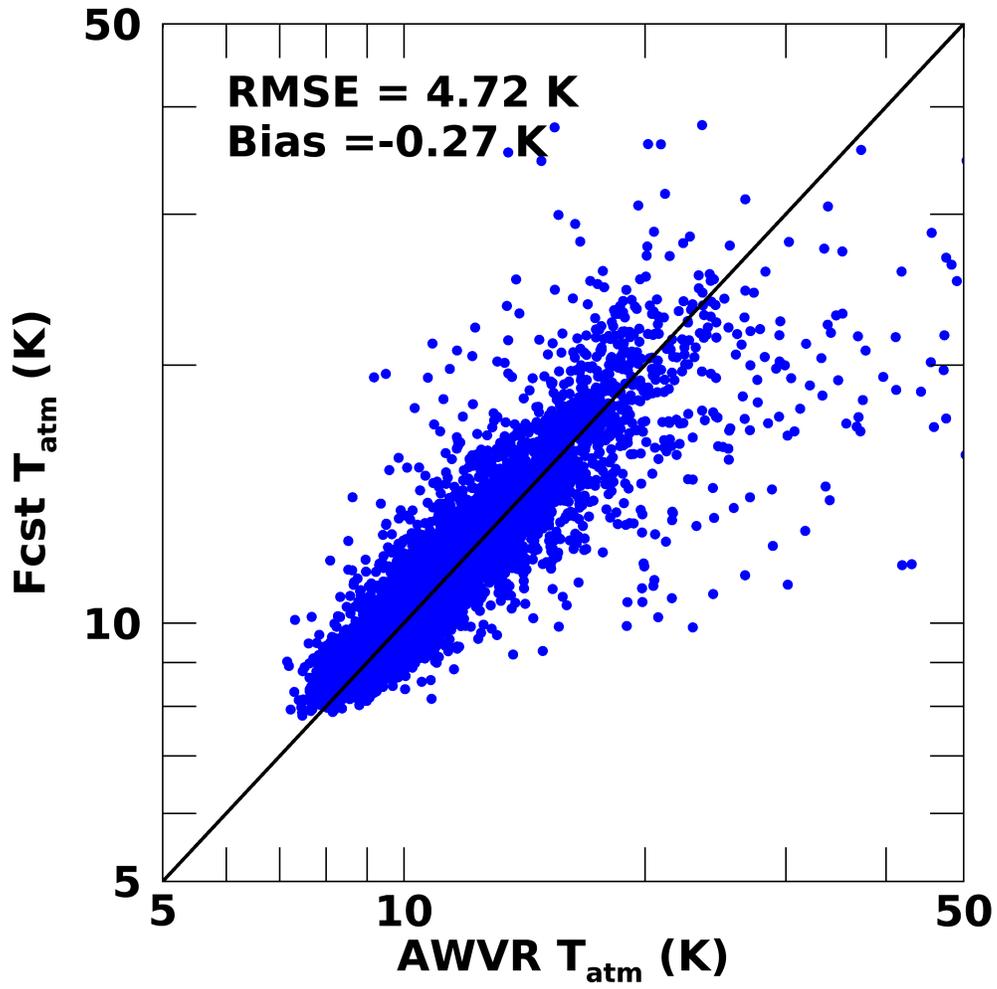
380

381 **Figure 2.** Cumulative distribution function (CDF) of T_{atm} (K) at Goldstone, CA.



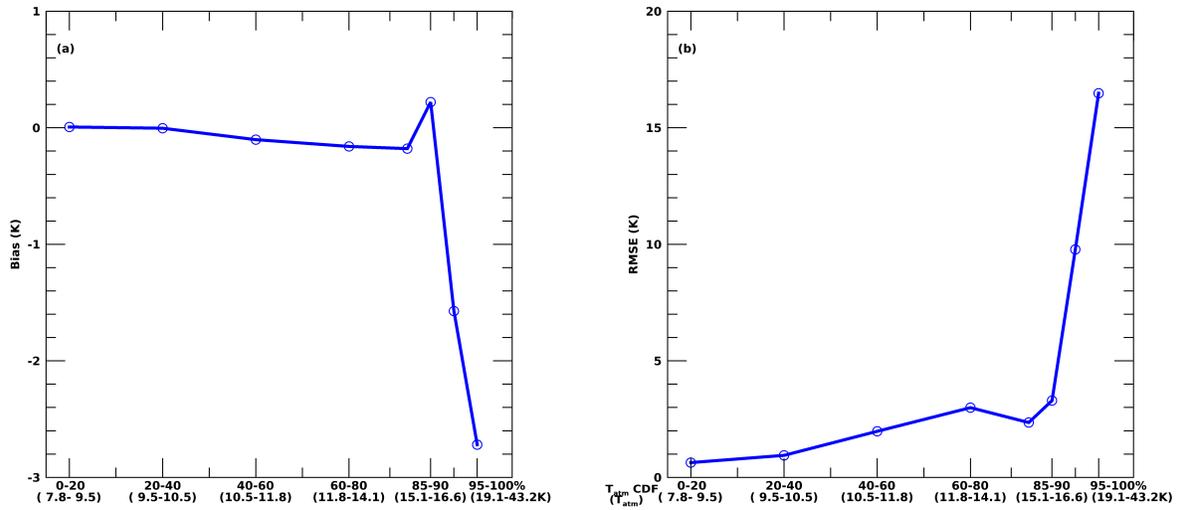
382

383 **Figure 3.** Time series of precipitable water vapor (PWV) from GFS 24-h forecast (PWV_f024)
384 and AWVR measurements (PWV_AWVR), as well as the AWVR T_{atm} in 2017 at Goldstone,
385 CA.



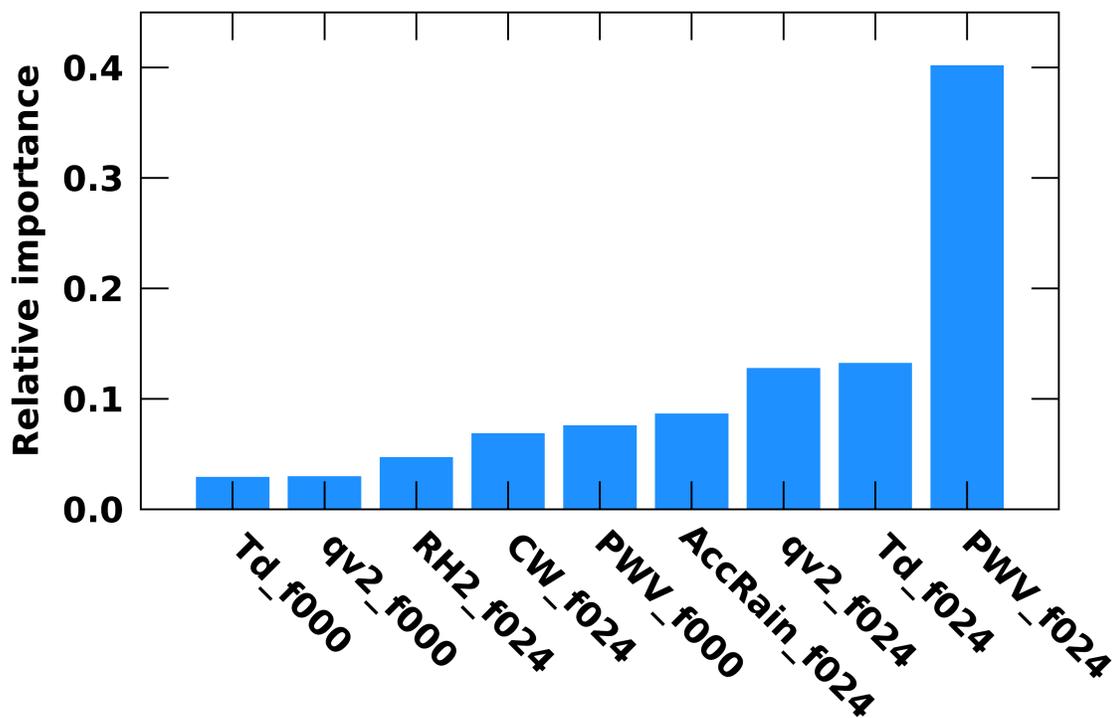
386

387 **Figure 4.** The observed (AWVR) vs. forecasted (Fcst) T_{atm} (K).



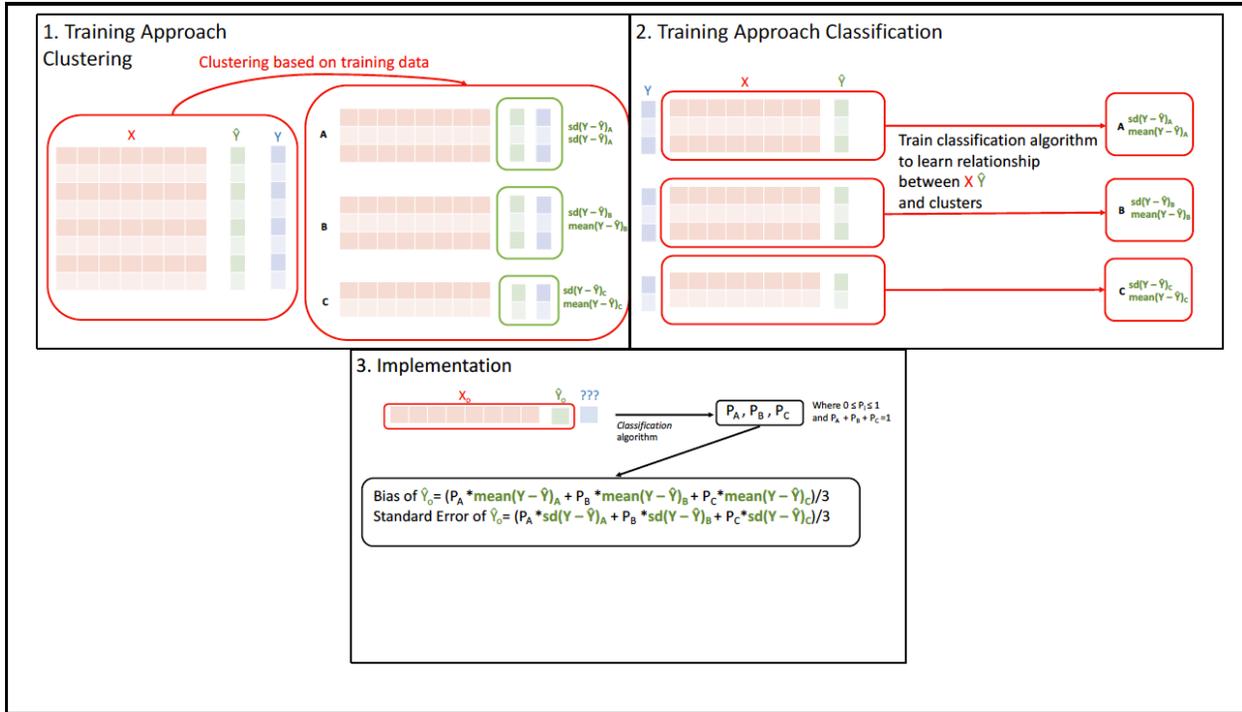
388

389 **Figure 5.** Forecast (a) bias (K) and (b) RMSE (K) sampled by the forecasted T_{atm} . X-axis is the
 390 percentile range (%) of the forecasted T_{atm} ; values inside parentheses are the corresponding
 391 ranges of the forecast T_{atm} (K).
 392



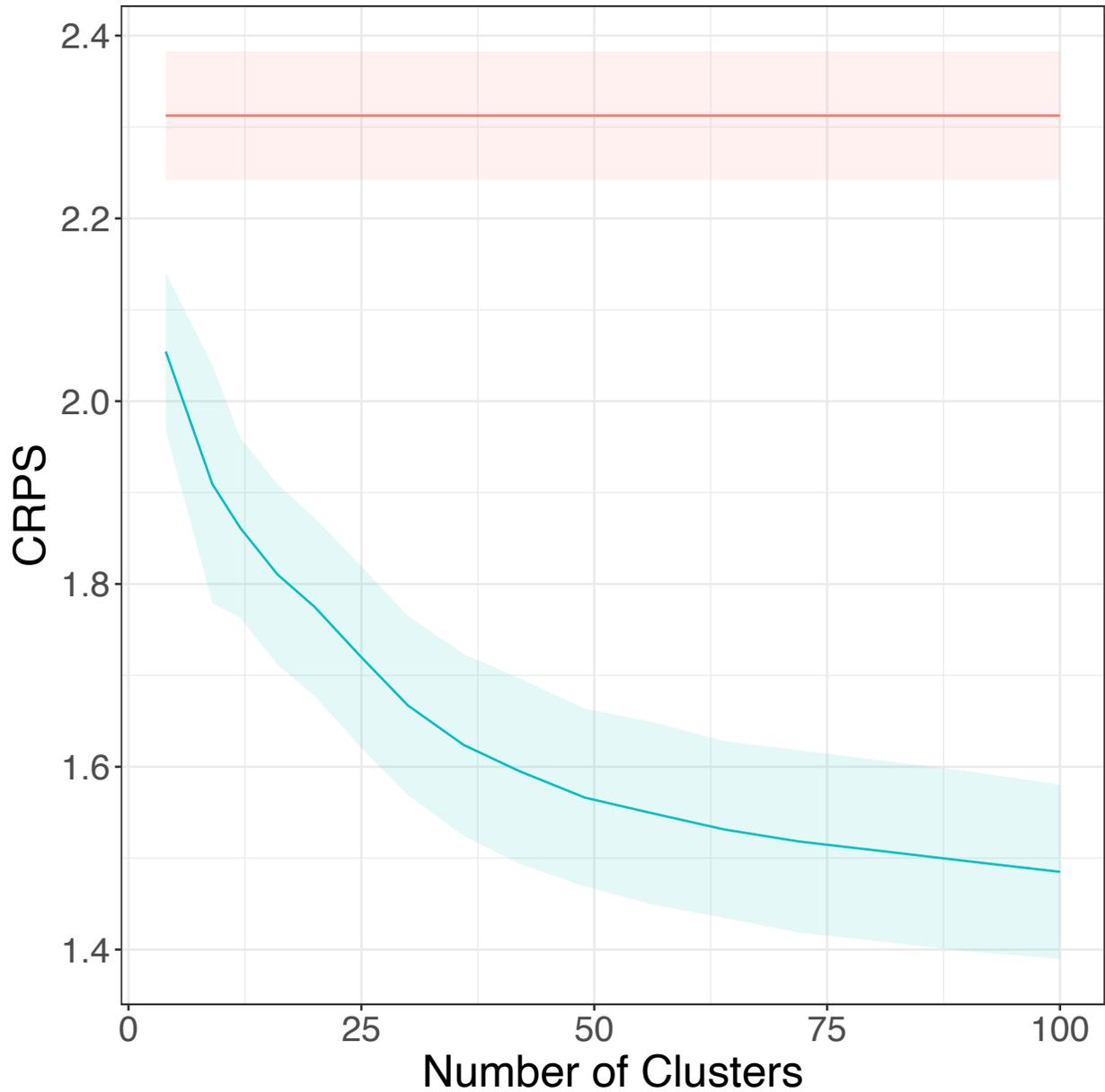
393

394 **Figure 6.** Relative importance of predictors for predicting T_{atm} in 2017 using the ML model.



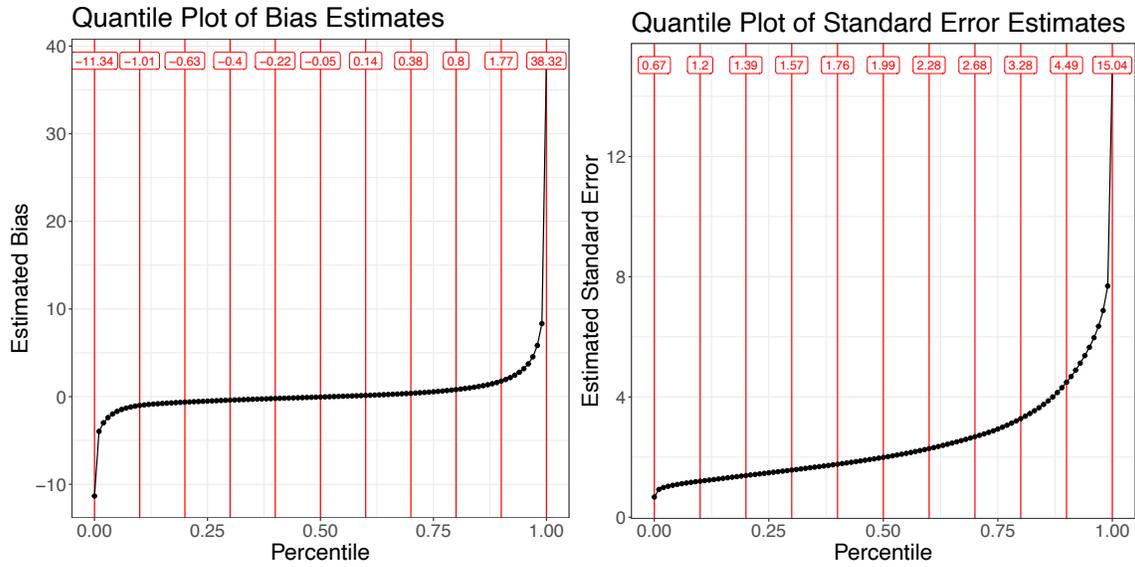
395

396 **Figure 7.** A schematic representation of the Uncertainty Quantification model training and
 397 implementation framework. Panel 1 illustrates the clustering step, where a clustering algorithm is
 398 trained on the predictors (X), the predict and (\hat{Y}), and the true values (Y). Panel 2 demonstrates
 399 the subsequent step, where a classification algorithm models the relationship between X , \hat{Y} , and
 400 the defined clusters, circumventing the need for Y in the implementation step (Panel 3).



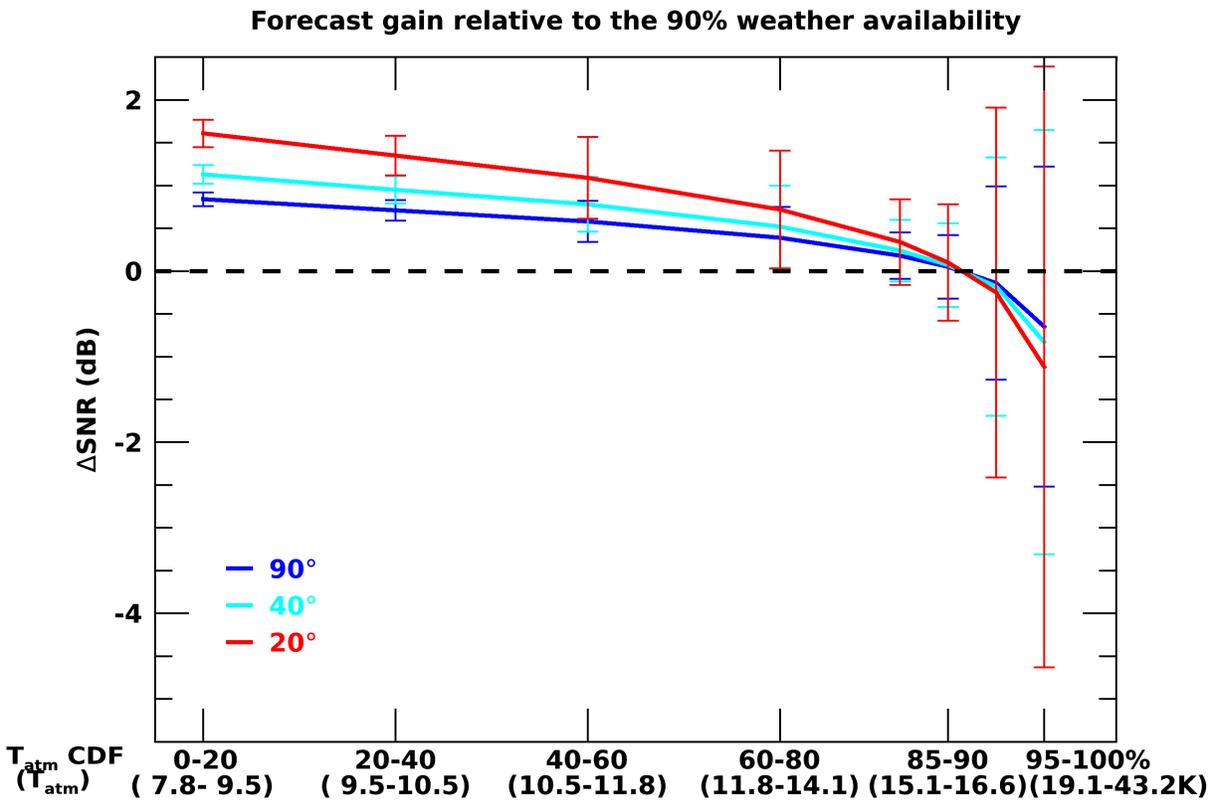
401

402 **Figure 8.** Line plot of CRPS values under different choices of cluster number. The blue line
403 represents the mean CRPS for ensemble predictions at different specified cluster numbers for the
404 Self-Organizing Maps, while the blue shadow area represents the standard deviation of the CRPS
405 for the ensemble. The red line represents the CRPS value for a “naïve” model which considers
406 only the training population mean and standard deviation, with the red shadowed area
407 representing the standard deviation of the CRPS ensemble for different training datasets.



408

409 **Figure 9.** Quantile plot of bias (K) (left panel) and standard error (K) (right panel) from the
 410 Uncertainty Quantification model. Each point represents the respective quantile to the nearest
 411 one-tenth of a percentile. The red numbers represent the value for each decile (in K), including
 412 the maximum and minimum values.



413

414 **Figure 10.** Forecast Gain ($\Delta E_b/N_0$, dB) relative to the 90% weather availability versus mean T_{atm}
 415 for different elevation angle cases (90°, 40° and 20°). Error bars represent the forecast errors. X-
 416 axis is the percentile range (%) of the forecasted T_{atm} ; values inside parentheses are the
 417 corresponding ranges of the forecast T_{atm} (K).

418 **List of tables**

419 **Table 1.** List of predictors to forecast 24-hour T_{atm} in the machine learning forecast model. GFS
 420 analysis data is valid at the forecasting time (T0); GFS 24-hour forecast data provides 24-hour
 421 forecasts at T0; T_{atm} is valid at 24 hours after the forecasting time (T24). Third Column is the
 422 correlation of predictors with $\ln(T_{atm})$.

Variable name	Variable type	Correlation	Description
PWV_f024	24-h forecast	0.71	Precipitable water vapor
qv2_f024	24-h forecast	0.59	2-m specific humidity
Td_f024	24-h forecast	0.57	2-m dew point temperature
PWV_f000	analysis	0.52	Precipitable water vapor
qv2_f000	analysis	0.40	2-m specific humidity
RH2_f024	24-h forecast	0.39	2-m relative humidity
Td_f000	analysis	0.38	2-m dew point temperature
CW_f024	24-h forecast	0.37	Cloud water
AccRain_f024	24-h forecast	0.32	6-hr accumulated precipitation

423