Article

Early Night Fog Prediction Using Liquid Water Content Measurement in the Monterey Bay Area

Steven Kim [1,*], Conor Rickard [2], Julio Hernandez-Vazquez [1] and Daniel Fernandez [3]

1   Department of Mathematics and Statistics, California State University, Monterey Bay, Seaside, CA 93955, USA
2   Department of Biology and Chemistry, California State University, Monterey Bay, Seaside, CA 93955, USA
3   Department of Applied Environmental Science, California State University, Monterey Bay, Seaside, CA 93955, USA
*   Correspondence: stkim@csumb.edu

Abstract: Fog is challenging to predict, and the accuracy of fog prediction may depend on location and time of day. Furthermore, accurate detection of fog is difficult, since, historically, it is often carried out based on visual observations which can be biased and are often not very frequent. Furthermore, visual observations are more challenging to make during the night. To overcome these limitations, we detected fog using FM-120 instruments, which continuously measured liquid water content in the air in the Monterey, California (USA), area. We used and compared the prediction performance of logistic regression (LR) and random forest (RF) models each evening between 5 pm and 9 pm, which is often the time when advection fog is generated in this coastal region. The relative performances of the models depended on the hours between 5 pm and 9 pm, and the two models often generated different predictions. In such cases, a consensus approach was considered by revisiting the past performance of each model and weighting more heavily the more trustworthy model for a given hour. The LR resulted in a higher sensitivity (hit rate) than the RF model early in the evening, but the overall performance of the RF was usually better than that of the LR. The consensus approach provided more robust prediction performance (closer to a better accuracy level between the two methods). It was difficult to conclude which of the LR and RF models was superior consistently, and the consensus approach provided robustness in 3 and 2 h forecasts.

Keywords: liquid water content; fog prediction; logistic regression; random forest

1. Introduction

One of the most unpredictable meteorological phenomena is fog, which is formed for various reasons depending on seasons and locations. Due to the unique conditions of the California coast (cold air near the surface and hot temperatures at higher altitudes), there is much to learn about the fog events that occur in the Monterey Bay area located in California, the United States [1,2]. Early references describe the frequency of fog (which is defined as conditions on the ground with less than one-kilometer visibility) along the Pacific coast at about 40 days per year on average [3]. In the Monterey Bay area, the annual frequency is about 25 to 35 days on average. Unpredictable fog events have caused flight delays for a local airport, affected road conditions for commuters [4], and increased the risk of accidents for fishermen [5]. In addition, the presence of fog is of relevance for ecosystem processes and, potentially, for the capture of water for various purposes, including agriculture, reforestation, and even human consumption [6]. Frequent fog events affect the lives of many, and researchers have attempted various ways of predicting fog events; the predictability of observed factors depends on the geographical location, the season of the year, and the time of day; thus, predictive models should be localized.

Along the coast of California, advection fog is formed as a result of a cold ocean in conjunction with warmer air temperatures [7]. Cool ocean waters from the north move

south along the west coast of the United States [8]. Furthermore, northwesterly winds induce upwelling, particularly in the late spring, which provides additional surface water cooling at some locations [9]. The western coastline of the United States has a uniquely dependable inversion of warm air above cool air, which is in contact with the ocean [10]. The cooling of air below its dew point close to sea level and the containment of this by the warm air above often results in fog that can be advected toward the land in the evening by inland breezes [2].

Despite the relatively straightforward scientific explanations of fog generation, the prediction of fog is challenging because of the details associated with fog formation and its persistence. Important factors include the number and size of condensation nuclei available coupled with the ongoing balance and sufficient scale of temperature gradients and available humidity to maintain the liquid droplets that constitute fog. Therefore, one approach for fog prediction, rather than modeling the physical details of fog formation, is to apply regression-based models to forecast seasonal trends of weather conditions and fog. A multiple regression approach was used to estimate the probability of marine fog at 24 and 48 h intervals using variables such as evaporative heat flux and surface relative humidity [11]. Though the models outperformed the other models compared in the study, the r-square was only 0.30. Another regression analysis was performed to predict the ceiling height (the height of the lowest layer of overcast clouds or broken clouds) one, three, and six hours ahead based on the current ceiling height and other meteorological variables [12]. The researchers compared the predictive model with an observational model and concluded that the observational model was more effective for fog prediction. For binary outcomes, logistic regression was used to estimate the probability of when a low ceiling would no longer exist, one, two, and three hours ahead of 3 pm and 6 pm [13]. The logistic regression model was then compared to persistent climatology forecasts, showing an improvement in predicting when a low ceiling no longer exists.

More advanced models have been implemented for fog prediction. (Here, "advanced" does not necessarily imply better predictions, but it implies the ability to handle computations efficiently in automated procedures.) Interestingly, the use of artificial intelligence for fog prediction appeared in the literature of the late 1980s [14]. In Perth, Australia, a study was conducted using a fuzzy logic model to predict fog, and it compared the fuzzy model with two other forecasting models [15]. Recently, Markov chains and machine learning algorithms have been used to predict fog using multiple predictors, such as temperature, relative humidity, and other meteorological measurements [16]. Other advanced methods used to predict fog or fog dissipation, such as decision trees, support vector machines, ensemble methods (e.g., gradient boosting, random forest models), and artificial neural networks (e.g., extreme learning machine, multilayer perceptron), can also be found in the literature [17–20]. A recent study has evaluated the sensitivity of machine learning techniques for visibility forecasting and showed that ensemble-based models perform the best [21] and that tree-based methods perform better with multiple predictors [22]. These studies showed the importance of location, in that one location considered in their study proved more unpredictable than another, regardless of prediction techniques, and the random forest and similar methods performed better relative to the other methods considered. They also discussed the possibility that the inclusion of other models within different frameworks (e.g., logistic or linear regression, time-series, lazy learning) may further improve prediction.

The primary focus of this article is the 1–3 h forecasting in the Monterey Bay area for the times between 5 pm and 9 pm, which is when fog often begins in this region. Our predictions were binary (fog or no fog) and verified based on observed liquid water content (LWC), as measured by an FM-120 optical spectrometer. Other meteorological factors, including, among other variables, dew point depression, wind speed, and previous values of LWC, were used to drive the predictions at each hour. Our aim was not to exhaust and compare all existing prediction algorithms but rather to demonstrate the different results between the logistic regression method and the random forest method for each

hour. We also demonstrate how we can obtain a more robust prediction performance when the two methods have disagreeing predictions given the same predictors. To address the disagreement between the prediction models, a consensus approach was considered by reviewing their past prediction performances and more heavily weighting the one with better performance at a given time. We note that the validation of a prediction model requires data observed over a long period of time. However, the LWC measured by the FM-120 was available for only one season in this study. Thus, this research work is considered as a proof of concept.

This article is organized as follows. The LWC data are introduced in Section 2.1. Accompanying meteorological variables are described in Section 2.2. In addition to hourly averaged LWC, the slope of LWC is used as a predictor and is explained in Section 2.3. The prediction models considered in this study are explained in Section 2.4 and their evaluation criteria in Section 2.5. As an exploratory analysis, correlations between LWC and other predictors are presented in Section 3.1. The performance of each model is reported in Section 3.2 and the performance of the consensus approach in Section 3.3. The discussion is contained in Section 4.

2. Materials and Methods

2.1. Data

A regular, automated means of fog detection can be challenging. One method has involved the use of standard fog collectors (SFCs)—apparatuses used for the collection of fog droplets as fog passes through a mesh. Other methods employ unidirectional (conical) or planar fog harps, which use vertical strands of thread to collect the tiny fog droplets [23,24]. Once the SFCs or harps collect fog droplets, the water accumulates and drips down to a trough and into a rain gauge that records the time when the rain gauge tips [25]. While standard fog collectors are effective in collecting volumes of water associated with fog events, they are not perfect in recording the actual times of fog events and they can potentially miss some fog events that result in little or no accumulation of water due to the varying nature of fog and wind [26]. One issue, for instance, is that a sufficient number of fog droplets are required to pass through the mesh of an SFC, coalesce upon the mesh, and fall into the rain gauge in order to detect when a fog event has occurred, so there is a gap between the actual fog event and the timestamp, which may be over an hour [27]. An alternative approach, which we employ in this study, is to use an FM-120 optical spectrometer to measure the liquid water content (LWC) within the air.

For this study, the LWC was measured in grams per cubic meter at 5 m above ground level using an optical spectrometer device known as an FM-120 manufactured by Droplet Measurement Technologies. This optical spectrometer continuously samples droplet-laden ambient air to report size distributions in the range of 2 μm to 50 μm. Two FM-120 units were deployed at the time of this experiment to measure the efficiencies of standard fog collector devices, which were not used in the present study. While the FM-120 units produce droplet spectra in the range specified above, this study only utilized the LWC, which is an integrated measure of droplet numbers and sizes that is used in order to determine the presence of fog. The two FM-120 units were consistent in their detection of LWC values indicative of fog events, which further supports the accuracy of their data. Furthermore, an FM-120 unit was used in a different study with standard fog collectors in Chile [27]. This study also illustrated consistency in the detection of fog events in conjunction with a standard fog collector, with the expected lag of up to 60 min evident between the detection of sufficiently large LWC values by the FM-120 and the collection of water from the standard fog collector [27]. Both LWC and weather data were recorded at a site known as Fritzsche Field, near the Marina Airport, located at $36.695486°$, $-121.757475°$.

2.2. Exploratory Data Analysis

The response variable that we applied in this paper was the LWC (grams per cubic meter), and it was log-transformed base-10. The LWC is derived from the measurements of

airborne droplets, and values of LWC greater than 0.01 g per cubic meter constitute fog, corresponding to values of log(LWC) > −2. Note that, based on this threshold, we observed that fog rarely existed between 10 am and 3 pm at this location (as shown in the upper right panel of Figure 1). We note that when the observed value of LWC was numerically zero, which led to an undefined value of log(LWC), we added 3.91×10^{-8} to avoid a numerical error in the statistical modeling.
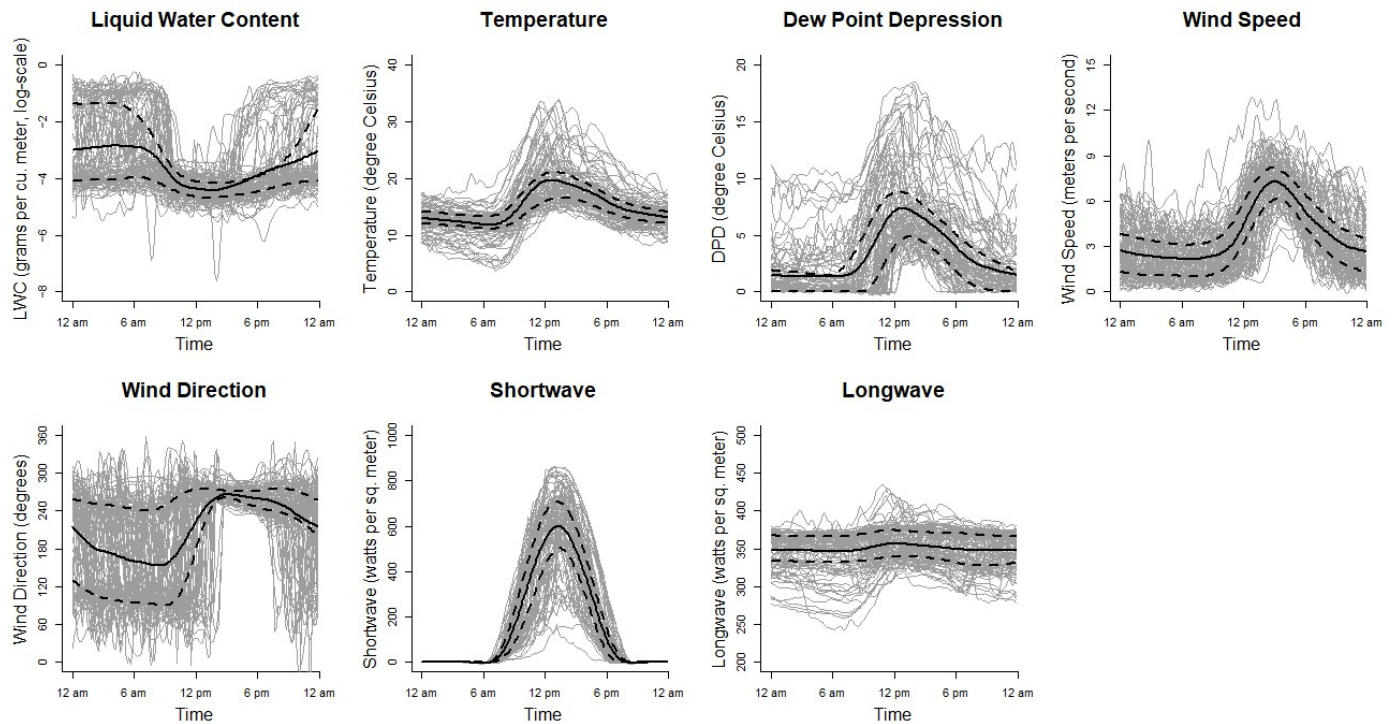

Figure 1. Observed daily trends (gray) and smoothing splines for overall average (black): liquid water content (LWC), temperature (T), dew point depression (DPD), wind speed (WS), wind direction (WD), shortwave (SW), and longwave (LW), observed from 29 July to 6 November 2020. Solid curves are average trends; dotted curves are the first and third quartiles.

The explanatory variables of interest were the temperature (T; degrees Celsius), dew point depression (DPD; degrees Celsius), wind speed (WS; meters per second), wind direction (WD; degrees), shortwave (SW; watts per square meter), and longwave (LW; watts per square meter). These variables were recorded every ten minutes, and the observation period was from 29 July to 6 November 2020. Figure 1 presents these variables with respect to time of day, and smoothing splines were applied for average daily trends. Table 1 presents descriptive statistics (mean, median, standard deviation, minimum, and maximum) of these variables observed at 0 am (midnight), 3 am, 6 am, 9 am, 12 pm (noon), 3 pm, 6 pm, and 9 pm.

As shown in Figure 1, the LWC, T, DPD, SW, and LW have inflection points near 12 pm (noon), and WS and WD have inflection points near 2 pm. The information contained in these inflection points (e.g., the rate of change in DPD) may play an important role in forecasting a fog event in the early evening. We explored the relationships between the response variable (LWC) and the explanatory variables with a 3 h interval to see if a 3 h forecast would be plausible. For hours t = 17, 18, 19, 20, 21 (i.e., from 5 pm to 9 pm), we calculated the hourly averages of all variables and regressed the LWC of hour t on each explanatory variable of hour t − 3 (simple regression) and all explanatory variables simultaneously (multiple regression). A variable X averaged at hour t is denoted by X(t) hereafter.

Table 1. Descriptive statistics of liquid water content (LWC), temperature (T), dew point depression (DPD), wind speed (WS), wind direction (WD), shortwave (SW), and longwave (LW), observed at 0 am (midnight), 3 am, 6 am, 9 am, 12 pm (noon), 3 pm, 6 pm, and 9 pm from 29 July to 6 November 2020.

		0 am	3 am	6 am	9 am	12 pm	3 pm	6 pm	9 pm
LWC (log base-10)	Mean	−2.98	−2.89	−2.91	−3.86	−4.37	−4.29	−3.82	−3.46
	Median	−3.67	−3.52	−3.50	−3.98	−4.37	−4.33	−4.16	−3.97
	SD	1.37	1.35	1.25	0.70	0.38	0.51	1.13	1.17
	Minimum	−4.67	−5.32	−4.47	−5.01	−5.24	−5.30	−6.09	−5.04
	Maximum	−0.50	−0.47	−0.38	−1.54	−3.10	−1.78	−0.74	−0.59
T	Mean	12.94	12.38	11.84	14.86	19.61	18.22	15.50	13.84
	Median	12.95	12.62	12.13	14.02	18.18	17.58	15.08	13.62
	SD	1.89	2.09	2.47	3.52	4.33	3.14	2.82	1.97
	Minimum	8.37	5.97	4.55	7.68	14.08	13.23	11.25	8.63
	Maximum	16.87	17.52	18.60	25.68	31.83	28.98	24.22	20.25
DPD	Mean	1.46	1.38	1.41	3.26	7.36	6.06	3.54	2.04
	Median	0.71	0.25	0.09	1.92	5.97	5.60	2.93	1.28
	SD	2.21	2.29	2.42	3.67	3.91	3.16	3.07	2.42
	Minimum	0.00	0.00	0.00	0.00	1.18	0.72	0.00	0.00
	Maximum	10.58	10.38	10.24	12.51	17.92	16.39	12.74	11.84
WS	Mean	2.66	2.31	2.14	2.47	5.06	7.22	4.88	3.22
	Median	2.08	2.02	2.10	2.20	5.13	7.28	4.82	2.63
	SD	1.72	1.36	1.23	1.39	2.08	1.30	1.56	1.88
	Minimum	0.17	0.42	0.15	0.48	0.92	2.90	1.72	0.42
	Maximum	7.45	5.73	5.05	7.45	11.93	11.52	9.23	7.97
WD	Mean	208.03	176.99	160.83	158.19	236.36	265.05	258.20	239.32
	Median	240.20	178.68	127.40	112.33	264.12	263.72	256.75	249.52
	SD	67.49	75.84	70.91	77.37	66.00	9.84	19.04	55.51
	Minimum	56.85	47.70	18.77	51.10	35.32	238.67	227.08	46.92
	Maximum	304.87	319.97	297.38	332.30	304.35	295.65	317.65	319.22
SW	Mean	0.00	0.00	0.84	211.95	571.27	442.13	59.76	0.00
	Median	0.00	0.00	0.00	225.47	596.88	441.30	43.77	0.00
	SD	0.00	0.00	1.79	106.06	161.48	145.77	60.07	0.00
	Minimum	0.00	0.00	0.00	17.20	110.67	60.70	1.50	0.00
	Maximum	0.00	0.00	10.93	445.85	841.13	731.62	239.77	0.00
LW	Mean	348.59	347.66	346.44	350.20	357.50	353.28	349.10	347.71
	Median	358.37	359.28	357.23	358.78	355.02	351.15	352.47	355.35
	SD	23.93	26.36	28.90	27.65	22.67	21.76	24.80	23.86
	Minimum	275.72	266.23	245.38	264.83	313.80	298.33	293.48	280.82
	Maximum	377.13	388.17	382.10	401.78	416.98	402.27	389.88	387.18

2.3. Derived Variables

DPD is known to be a strong predictor of the LWC. For instance, we hypothesize that LWC(18) is strongly correlated with DPD(15). In addition, we also hypothesize that LWC(18) is correlated with the rate of change in DPD during the afternoon. For instance, we attempted to explain LWC(18) by the slope of the hourly averages of DPD observed from t = 12 to t = 15 using ordinary least squares estimation (OLSE). This derived variable is denoted by ΔDPD(12,15). In addition, since the DPD was observed over 10 min windows, we could approximate the instantaneous rate of change at t = 15 using six data points. The slope of DPD estimated by OLSE during the 1 h period, t = 15, is denoted by ΔDPD(15). The slope of an hourly averaged variable X observed from hour t to hour u is denoted by ΔX(t, u), and the instantaneous slope of X at hour t, estimated by 10 min data, is denoted by ΔX(t) hereafter. The correlations between LWC(t) and the derived variables and the correlations between LWC(t) and the explanatory variables (introduced in Section 2.2) are provided in Section 3.1 (Table 2) for t = 17, 18, 19, 20, and 21.

Table 2. Correlation between LWC(t) and the variables observed t − 3. Statistical significance is indicated by * for $p < 0.05$, ** for $p < 0.01$, or *** for $p < 0.001$.

	LWC(t = 17)	LWC(t = 18)	LWC(t = 19)	LWC(t = 20)	LWC(t = 21)
LWC(t − 3)	0.39 ***	0.71 ***	0.86 ***	0.81 ***	0.79 ***
ΔLWC(t − 3)	0.41 ***	0.39 ***	0.49 ***	0.43 ***	0.35 ***
ΔLWC(t − 6, t − 3)	0.21 *	0.56 ***	0.77 ***	0.74 ***	0.73 ***
DPD(t − 3)	0.14	0.33 ***	0.47 ***	0.52 ***	0.52 ***
ΔDPD(t − 3)	0.29 **	0.26 *	0.21 *	0.05	0.25 *
ΔDPD(t − 6, t − 3)	0.22 *	0.37 ***	0.43 ***	0.39 ***	0.13
T(t − 3)	0.20 *	0.38 ***	0.49 ***	0.52 ***	0.49 ***
WS(t − 3)	0.05	0.04	0.09	0.09	0.08
WD(t − 3)	0.19	0.05	0.19	0.21 *	0.22 *
SW(t − 3)	0.28 **	0.37 ***	0.51 ***	0.45 ***	0.32 **
LW(t − 3)	0.23 *	0.18	0.01	0.06	0.15
All (Adjusted R-Squared)	0.41	0.66	0.79	0.72	0.67

2.4. Prediction Model for the Binary Outcome Fog

The left panel of Figure 2 presents the daily variability of LWC measured by the FM-120, and the right panel presents the distribution of hourly averaged LWC from 5 pm to 9 pm. We clearly see multimodal distributions which distinguish between foggy evenings and not-foggy evenings. We used a threshold of log(LWC) = −2.5 to define a fog event, and about 13% and 20% of the observed days were foggy as of 5 pm and 9 pm, respectively, according to the threshold. We note that this threshold corresponds to an LWC of about 0.003 g/m^3. While a threshold of 0.01 g/m^3 (or log(LWC) = −2) is generally taken as the threshold for fog, we chose the slightly lower threshold of 0.003 g/m^3 in order to allow for more cases for the predictive model based on the data observed from 29 July to 6 November 2020. Fog events were observed more frequently later in the evening during the period of observation (13.0%, 14.0%, 14.9%, 17.8%, and 19.8% for 5 pm, 6 pm, 7 pm, 8 pm, and 9 pm, respectively).

Using a similar method to that described in Section 2.3, we derived ΔLWC(t − 6, t − 3) and ΔLWC(t − 3) and used these as potential predictors for LWC(t) for t = 17, . . . , 21. We then considered four levels of prediction models using logistic regression (LR). The baseline prediction model only considered DPD(t − 3), which is referred to as LR0. LR1 considers LWC(t − 3) in addition to LR0. LR2 considers ΔLWC(t − 3) and ΔDPD(t − 3) in addition to LR1. LR3 considers ΔLWC(t − 6, t − 3) and ΔDPD(t − 6, t − 3) in addition to LR2. For instance, in order to predict fog at 6 pm, LR0 uses the DPD value at 3 pm; LR1 adds the LWC value at 3 pm; LR2 further considers the instantaneous change (slope) of LWC and of DPD at 3 pm; and LR3 further considers the slope of hourly averaged LWC and DPD from 12 to 3 pm. All statistical analyses were performed in R. To compare the parametric LR to a nonparametric machine learning algorithm, we considered the random forest (RF) with all the predictors used in LR3. The randomForest package was used in Ref. [28].

The LR uses a mathematical function, known as the logistic function, which receives observed predictors (e.g., change in LWC from 12 pm to 3 pm) and returns the probability of a binary outcome (e.g., the probability of a fog event at 6 pm). It may receive one or more predictors to estimate the probability given the observed data, and a binary outcome is predicted based on the estimated probability. Unlike the LR, the RF does not assume a specific functional relationship between the probability of a binary outcome and observed predictors. Instead, the RF builds decision trees based on bootstrap samples, randomly

selects a few predictors (instead of trying all available predictors at once) to evaluate predictability, and repeats the random process a large number of times to determine a final model. The RF is considered as a black box because it is difficult to know why the final model predicts the binary outcome (e.g., a fog event) given input predictors [29]. A large-scale numerical experiment showed that RF performed better than LR with about 69% of the datasets tested in the study [29]. In general, if the true relationship between predictors and the probability of a binary outcome is close to the assumed logistic function, the LR will outperform the RF. Otherwise, the RF will outperform the LR, as it does not rely on the same mathematical assumption and is not sensitive to outliers.
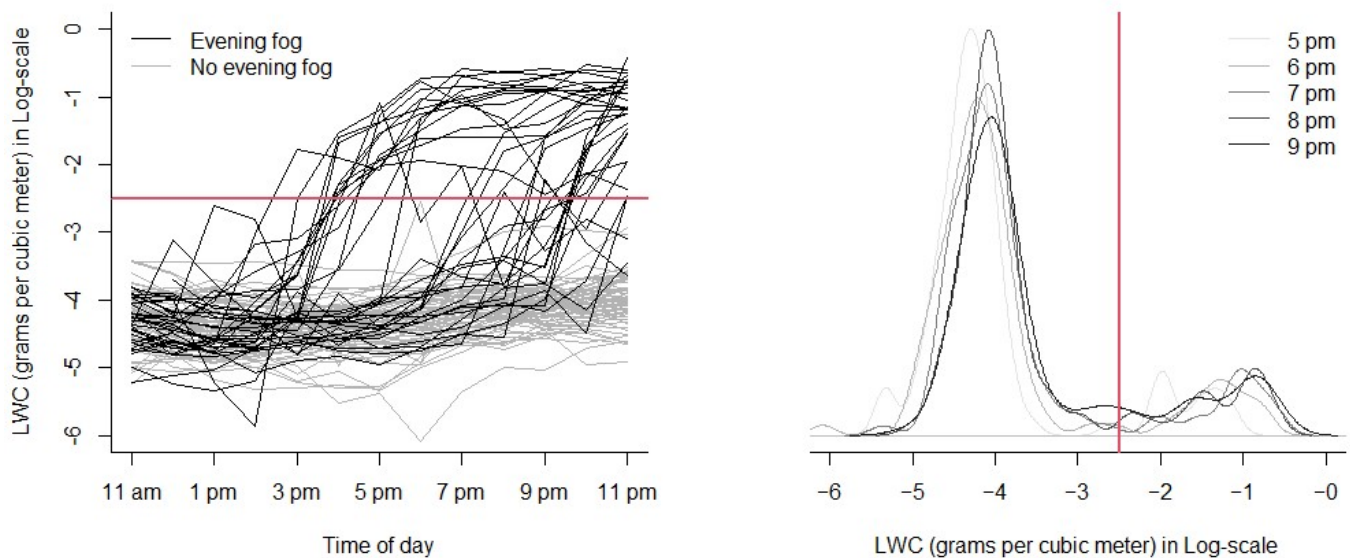
Figure 2. The daily trend of LWC (**left**) and the observed distribution of LWC(t) from t = 17 to t = 21 (**right**), observed from 29 July to 6 November 2020. The red line represents the threshold above which fog is defined to be present. This paper used a slightly lower threshold of log(LWC) = −2.5.

2.5. Evaluation of Prediction Models

The models were tested to predict fog at 5 pm, 6 pm, 7 pm, 8 pm, and 9 pm, and we evaluated the predictive performance by the hit rate (HR), false alarm rate (FAR), and critical success index (CSI), which are defined in Equations (1)–(3), respectively. The HR is defined as the percentage of times that the fog was correctly predicted during the times that there was actual fog (also known as the sensitivity, true positive rate, or recall score). The FAR refers to the percentage of times that the fog did not occur when it was predicted to occur (100% minus precision). The CSI refers to the ratio of occasions when fog was correctly predicted to the number of times when there was actual fog plus the number of times that it was incorrectly predicted that there would be fog. If we let TP, TN, FP, and FN denote true positive, true negative, false positive, and false negative, respectively, the three criteria are defined as:

$$HR = TP/(TP + FN), \tag{1}$$

$$FAR = FP/(TP + FP), \tag{2}$$

$$CSI = TP/(TP + FP + FN). \tag{3}$$

All values for HR, FAR, and CSI, which are referred to as Equations (1)–(3), respectively, are values between zero and one; high values for HR and CSI and low values for FAR are desired. In addition, the CSI will always be less than or equal to the HR.

To be more realistic in the evaluation of prediction models, we considered the following procedure. Starting with data for 27 September (about 60% of all available data), we trained a prediction model (model parameters and decision threshold), predicted the fog (presence or absence) for the next day between 5 and 9 pm, and recorded the prediction result (correct

or incorrect). Then, adding the data for 28 September, we updated the model and made the prediction for the next day. We continued this process until we reached 6 November, the last observation day. We summarized and compared the prediction performance of the five models considered (LR0, LR1, LR2, LR3, and RF).

3. Results

3.1. Data Exploration

Table 2 presents the correlations between LWC(t) and each variable observed at $t - 3$ for $t \geq 17$. In the univariate analysis, the slope ΔLWC(t − 3) has a higher correlation than the hourly average LWC(t − 3) at t = 17 (5 pm), and LWC(t − 3) is a stronger predictor than ΔLWC(t − 3) for $t \geq 18$ (6 pm or later). The relationship between DPD(t − 3) and LWC(t) is weak in the early evening, but a moderate correlation is observed in the late evening. The slope ΔDPD(t − 3) has a higher correlation with LWC(t) than DPD(t − 3) at t = 17, and DPD(t − 3) has a higher correlation with LWC(t) than ΔDPD(t − 3) for $t \geq 18$.

In the table, statistically significant variables are marked as *, **, or *** depending on the level of significance. LWC(t − 3) has a strong relationship with LWC(t) for $t \geq 18$ (6 pm and later), but the quick change ΔLWC(t − 3) has a moderate relationship with LWC(t) for $t \geq 17$ (5 pm and later). In addition, the gradual hourly change ΔLWC(t − 6, t − 3) has a stronger relationship with LWC(t) for $t \geq 18$ than the quick change ΔLWC(t − 3), and vice versa, at t = 17. Using multiple regression, predicting fog at 5 pm and 9 pm seems more challenging than predicting fog at 7 pm.

The DPD(t − 3) has a moderate relationship with LWC(t) for $t \geq 18$, but ΔDPD(t − 3) and ΔDPD(t − 6, t − 3) appear to have a higher correlation with LWC(t) rather than DPD(t − 3) at t = 17. The gradual change ΔDPD(t − 6, t − 3) seems to be associated with LWC(t) until t = 20.

The temperature T(t − 3) has a moderate relationship with LWC(t) for $t \geq 18$, but not at t = 17. The wind speed WS(t − 3) and the wind direction WD(t − 3) have relatively weaker associations with LWC(t). It is interesting that the shortwave SW(t − 3) has a moderate relationship with LWC(t) for $t \geq 18$, but was found to be not a useful predictor in the multiple linear regression. Additionally, multiple linear regression showed that, given the information of LWC and DPD, the other meteorological factors did not make meaningful contributions to the explanation of LWC.

3.2. Prediction Model Performance

The best prediction model depended on the time between the forecast and the observed event (e.g., 1 h, 2 h, or 3 h forecast), as shown in other studies [16,18], and it further depended on the time of evening (5 pm to 9 pm) in the Monterey Bay area. The critical success index, CSI = TP/(TP + FP + FN), is summarized hourly in Table 3. For the 3 h forecast, LR3 resulted in the highest CSI = 0.61 at 5 pm, and the simpler LR1 tended to perform the best from 7 pm to 9 pm, with CSI = 0.64 to 0.89. While it may seem surprising that the LR3, which utilized information from between 3 and 6 h prior, performed more poorly than the LR1, which utilized information from only 3 h prior, we surmise that the added information associated with the LR3 distracted the prediction from 7 pm to 9 pm. The RF and LR1 showed similar results from 6 pm to 8 pm. For the 2 h forecast, LR1 and RF were competitive, and RF was superior to all logistic regression models for the 1 h forecast, except for LR3 at 6 pm. It was difficult to suggest one model for all evening hours. This motivated the following section, which involved the consensus approach.

Table 3. Hit rate (HR = TP/(TP + FN); also known as the sensitivity, true positive rate, or recall score), false alarm rate (FAR = FP/(TP + FP); 100% minus precision), and critical success index (CSI = TP/(TP + FP + FN)) of each prediction model by evening hours.

		1 h Forecast					2 h Forecast					3 h Forecast				
Time	Criterion	LR0	LR1	LR2	LR3	RF	LR0	LR1	LR2	LR3	RF	LR0	LR1	LR2	LR3	RF
	HR	0.75	0.77	0.88	1.00	0.88	0.38	0.50	0.38	0.50	0.50	0.16	0.32	0.65	0.81	0.33
5 pm	FAR	0.33	0.14	0.22	0.20	0.00	0.40	0.00	0.24	0.33	0.00	0.50	0.33	0.20	0.29	0.00
	CSI	0.55	0.68	0.70	0.80	0.88	0.30	0.50	0.34	0.40	0.50	0.14	0.28	0.55	0.61	0.33
	HR	0.75	0.88	0.75	0.88	0.88	0.75	0.88	0.88	0.88	0.88	0.25	0.50	0.63	0.63	0.50
6 pm	FAR	0.40	0.22	0.00	0.00	0.13	0.33	0.13	0.22	0.22	0.13	0.50	0.00	0.17	0.50	0.00
	CSI	0.50	0.70	0.75	0.88	0.78	0.55	0.78	0.70	0.70	0.78	0.20	0.50	0.56	0.38	0.50
	HR	1.00	0.89	1.00	1.00	1.00	0.67	0.89	0.67	0.78	0.89	0.67	0.89	0.89	0.89	0.89
7 pm	FAR	0.18	0.00	0.10	0.10	0.00	0.40	0.20	0.00	0.00	0.00	0.33	0.00	0.11	0.27	0.00
	CSI	0.82	0.89	0.90	0.90	1.00	0.46	0.73	0.67	0.78	0.89	0.50	0.89	0.80	0.67	0.89
	HR	1.00	0.90	0.90	0.90	0.90	0.80	0.80	0.80	0.70	0.80	0.70	0.70	0.70	0.70	0.70
8 pm	FAR	0.33	0.18	0.18	0.25	0.18	0.38	0.00	0.11	0.13	0.20	0.30	0.13	0.36	0.22	0.13
	CSI	0.67	0.75	0.75	0.69	0.75	0.53	0.80	0.73	0.64	0.67	0.54	0.64	0.50	0.58	0.64
	HR	0.90	0.90	0.90	0.90	0.90	0.90	0.90	0.90	0.90	0.90	0.80	0.80	0.80	0.70	0.80
9 pm	FAR	0.53	0.18	0.18	0.31	0.10	0.40	0.18	0.25	0.31	0.25	0.38	0.00	0.11	0.22	0.20
	CSI	0.45	0.75	0.75	0.64	0.82	0.56	0.75	0.69	0.64	0.69	0.53	0.80	0.73	0.58	0.67

3.3. Consensus Prediction

For the 2 h forecast from 5 pm to 9 pm, the RF showed the same or better performance (higher HR, lower FAR, and higher CSI) than LR3 except for the FAR at 8 pm. For the 1 h and 3 h forecast, the relative superiority of LR3 and RF depended on the time of evening and the evaluation criterion (HR, FAR, or CSI), even though the two models utilized the same set of predictors. Furthermore, there were many cases where LR1 outperformed the logistic regression models. The RF is useful for utilizing multiple predictors without parametric assumptions, while LR1 is simple and performs well at certain hours. We attempted a consensus approach between LR1 and RF. When the predictions of LR1 and RF agreed (i.e., both predicted fog or no fog), the prediction was straightforward. When the predictions of LR1 and RF disagreed, consensus was determined by choosing the model which was correct more often per all prior disagreements of that type. For example, if LR1 predicted fog and RF did not, yet each time this occurred in the past one model was correct more than 50% of the time, the prediction of that model was chosen as the consensus. The same method for consensus applied when RF predicted fog and LR1 did not. We note that the consensus approach does not guarantee the performance of the better model. Instead, it is devised to achieve a performance level that is close to that of the better model.

The CSI = TP/(TP + FP + FN) of the consensus prediction is compared with those of individual LR1 and RF in Figure 3. For the 3 h prediction, the consensus prediction resulted in a CSI score that was sometimes the same as that of the better model (at 9 pm) and sometimes the same as that of the worse model (at 5 pm). For the 2 h prediction, it never resulted in a lower CSI score than the worse model for the 2 h prediction. For the 1 h prediction, however, the consensus prediction did not seem helpful, and it was found that it could have a slightly lower CSI than the worse model.
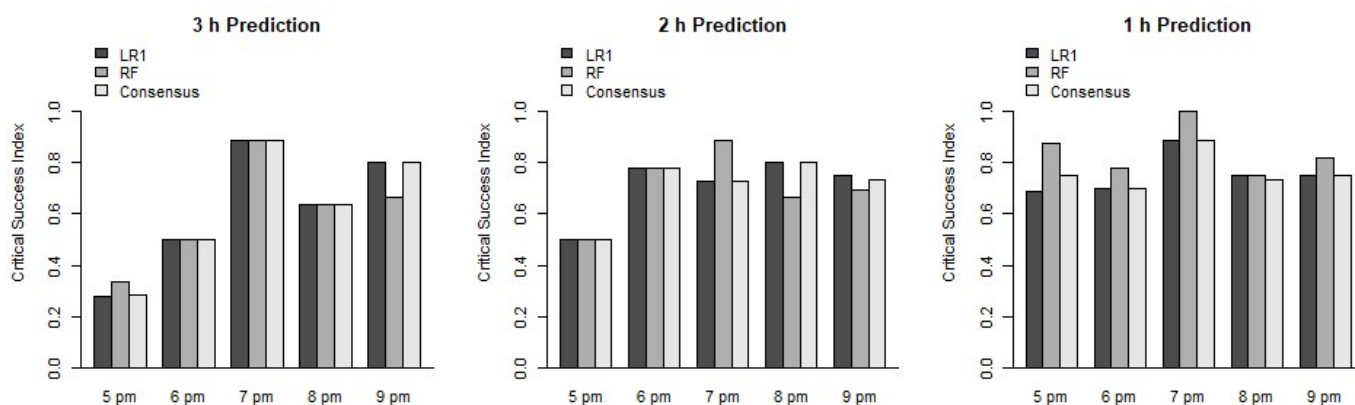
Figure 3. CSI of LR1, RF, and the consensus prediction; CSI = TP/(TP + FP + FN).

4. Discussion

Radiation fog, such as the Tule fog in California's Central Valley, is often associated with problems caused within some transportation systems, such as roadways in the Central Valley. This, coupled with the slightly more predictable nature of radiation fog, since it occurs in still conditions and does not advect, results in more studies considering and measuring its presence and persistence [30]. This study, however, focused on advection fog, given its importance in coastal ecosystems, its greater potential for water capture (particularly when it is accompanied by orographic lift), and its added impacts on aircraft and maritime transportation. The FM-120 instrumentation allowed us to accurately and effectively determine the presence or absence of fog droplets and how their presence is influenced by other meteorological factors.

Other studies have generated a variety of different metrics for assessing effectiveness in fog prediction (generally of radiation fog). We note some of the challenges associated with comparisons of results from different models applied in different fog regions.

Some recent studies have reported the F1 score (F1S), which combines hit rate (HR = TP/(TP + FN); also known as the sensitivity, true positive rate, or recall score) and false alarm rate (FAR = FP/(TP + FP); 100% minus precision) as a single metric, and is defined as F1S = 2 × (1 − FAR) HR/(1 − FAR + HR). F1 scores ranged between 0.53 and 0.77, depending on predictive models, for 1 h forecasts of low visibility at Valladolid Airport, Spain [16]. In a recent study of South Korea, twelve models were compared, and they resulted in a median F1S of 0.81 in Incheon Port and of 0.66 in Haeundae Beach in South Korea for a 1 h forecast of fog dissipation [18]. Another study considered various simulation settings of hourly fog assessments which resulted in an average F1S from 0.25 to 0.81 [31]. In this study, for the 1 h forecast between 5 pm and 9 pm in the Monterey Bay area, the F1S ranged from 0.81 to 0.94 under LR1 and from 0.85 to 1.00 under RF. Despite the common metric, it is very challenging or even impossible to compare or relate fog prediction studies, for several reasons. The outcome of interest differs from study to study (e.g., fog, fog dissipation, low visibility), and the definition of a binary outcome to be predicted varies. In this study, instead of using a threshold of visible distance, a common metric associated with the presence of fog, we used the LWC measured by the FM-120 to define a fog event. Furthermore, available predictors, causes of fog or fog dissipation, and statistical methods vary from study to study. Additionally, we demonstrated in this study that fog prediction is more or less challenging at certain times, and we showed that the performance of models depends on the time of day. We focused on the critical period of the day (between 5 pm and 9 pm) when fog often starts to appear in the Monterey Bay area. Instead of using cross-validation, we updated the models as we moved forward for fog prediction in the process of evaluating model performance. The limitation of our study is that we observed the LWC data for only one season; we shall therefore extend this study to multiple seasons and locations to determine whether one model is consistently superior.

Though the random forest (RF) method is efficient for testing a large number of combinations of predictors and evaluating predictive performance, it seems that human subjectivity and experiences must be involved in this process. In our RF model, we derived the slope variables of DPD and LWC and used the estimated slopes as predictors of RF. It performed better than integrating all observed DPD and LWC information in the model-building process. In other words, deriving variables based on scientific rationale seemed to improve fog prediction as opposed to letting an automated algorithm dictate the prediction process based purely on raw data.

The time- and location-specific performances of multiple candidate models motivates the use of a consensus approach, and it has been used to account for model uncertainty [32–34]. In the presence of model uncertainty, the multiple-model approach seems reasonable by weighing one model more based on the history of performance at a certain time, location, and other factors. Having too many models would complicate the process without meaningful gain, and we can benefit from combining a few superior models selected by careful monitoring and continual evaluation of past prediction performance.

Unlike related research which compared multiple prediction models separately, we attempted time-specific consensus prediction to address the challenge posed by the outperformance of one model by another depending on the time of day. The consensus approach was applied to achieve performance close to the performance of a better model. In this study, we attempted 1 h, 2 h, and 3 h forecasts of early night fog events in the Monterey Bay area and found that the consensus approach was robust for the 2 h and 3 h forecasts but not for the 1 h forecast. Our findings are limited by the scope of the specific location (the Monterey Bay area), and they are based on data collected for only one season. The performance of consensus prediction should be validated when a larger collection of data collected over a longer period by FM-120 instruments becomes available.

5. Conclusions

This paper implemented a consensus-based approach toward the modeling and prediction of advective fog at a central California coastal location. Use of FM-120 instruments provided accurate, reliable, and consistent measurements of air liquid water content, which allowed for regular measurement of the presence of fog. The authors note that, given the cost of FM-120 instruments, other lower-cost methods of fog detection, including the use of standard fog collectors, could be considered in revised versions of the proposed methods, with the caveat that there is some delay between the onset of a fog event and the collection of water by a standard fog collector.

In conjunction with the measurements for LWC, we also integrated regular measurements of key meteorological variables into the modeling and prediction process. The variables we found most important to include were wind speed, wind direction, dew point depression, previous hours' liquid water contents, and the slopes of dew point depression and LWC. The results presented in this paper utilized 1 h to 3 h forecasts based on these variables in order to predict the presence of advective fog in the window from 5 to 9 pm.

Higher correlations existed between some of the more strongly related variables and LWC at later times than at the earliest time (5 pm), possibly because, once the fog formed at 5 pm, it was a good indicator of fog existing at later times. Additionally, both the random forest and the logistic regression models indicated some impressive predictive behaviors for 1 h to 3 h forecasting at various times between 5 and 9 pm. The comprehensive critical success index (CSI) exhibited values from a low of around 25% at 5 pm for the more difficult 3 h forecast to a high of 100% for the 1 h forecast at 7 pm. The specific critical success rate values were dependent upon the time of the predicted fog and the hours ahead of the prediction.

Furthermore, we applied a consensus approach to resolve those cases where the results of the random forest and linear regression models disagreed. For all but one of the times between 5 pm and 9 pm and predictions from 1 to 3 h that we examined, the consensus

approach provided robust results in terms of CSI for agreement between model outputs and the actual LWC data derived from the FM-120 instruments for the 2 h and 3 h forecasts.

References

1. Blake, D. The subsidence inversion and forecasting maximum temperature in the San Diego area. *Bull. Am. Meteorol. Soc.* **1948**, *29*, 288–293. [CrossRef]
2. Leipper, D. Fog on the U.S. West coast: A review. *Bull. Am. Meteorol. Soc.* **1994**, *75*, 229–240. [CrossRef]
3. Ward, R.D. Fog in the United States. *Geogr. Rev.* **1923**, *12*, 576–582. [CrossRef]
4. Taylor, D.L. Monterey County Commute Traffic a Challenge to Cutting Greenhouse Gasses. 5 August 2021, Monterey Herald. Available online: https://www.montereyherald.com/2021/08/05/monterey-county-commute-traffic-a-challenge-to-cutting-greenhouse-gasses/ (accessed on 29 April 2020).
5. Churm, S.R. Treacherous Surf, Thick Fog Blamed in Aanglers' Deaths. 18 February 1985, Los Angeles Times. Available online: https://www.latimes.com/archives/la-xpm-1985-02-18-mn-3152-story.html (accessed on 29 April 2020).
6. Klemm, O.; Schemenauer, R.S.; Lummerich, A.; Cereceda, P.; Marzol, V.; Corell, D.; Van Heerden, J.; Reinhard, D.; Gherezghiher, T.; Olivier, J. Fog as a fresh-water resource: Overview and perspectives. *Ambio* **2012**, *41*, 221–234. [CrossRef]
7. Koračin, D.; Dorman, C.E.; Lewis, J.M.; Hudson, J.G.; Wilcox, E.M.; Torregrosa, A. Marine fog: A review. *Atmos. Res.* **2014**, *143*, 142–175. [CrossRef]
8. Marchesiello, P.; McWilliams, J.C.; Schcheptkin, A. Equilibrium structure and dynamics of the California current system. *J. Phys. Oceanogr.* **2003**, *33*, 753–783. [CrossRef]
9. Tseng, Y.H.; Chien, S.H.; Jin, J.; Miller, N.L. Modeling air-land-sea interactions using the integrated regional model system in Monterey Bay, California. *Mon. Weather. Rev.* **2012**, *140*, 1285–1306. [CrossRef]
10. Blake, D. Temperature inversions at San Diego, as deduced from aerographical observations by airplane. *Mon. Weather. Rev.* **1928**, *56*, 221–224. [CrossRef]
11. Koziara, M.C.; Renard, R.J.; Thompson, W.J. Estimating marine fog probability using a model output statistics scheme. *Mon. Weather. Rev.* **1983**, *111*, 2333–2340. [CrossRef]
12. Vislocky, R.L.; Fritsch, J.M. An automated, observations-based system for short-term prediction of ceiling and visibility. *Weather. Forecast.* **1997**, *12*, 31–43. [CrossRef]

13. Hilliker, J.L.; Fritsch, J.M. An observations-based statistical system for warm-season hourly probabilistic forecasts of low ceiling at the San Francisco International Airport. *J. Appl. Meteorol.* **1999**, *38*, 1692–1705. [CrossRef]

14. Peak, J.E.; Tag, P.M. An expert system approach for prediction of maritime visibility obscuration. *Mon. Weather. Rev.* **1989**, *117*, 2641–2653. [CrossRef]

15. Miao, Y.; Potts, R.; Huang, X.; Elliot, G.; Rivett, R. A Fuzzy Logic Fog Forecasting Model for Perth Airport. *Pure Appl. Geophys.* **2012**, *169*, 1107–1119. [CrossRef]

16. Cornejo-Bueno, S.; Casillas-Pérez, D.; Cornejo-Bueno, K.; Chidean, M.I.; Caamaño, A.J.; Sanz-Justo, J.; Casanova-Mateo, C.; Salcedo-Sanz, S. Persistence analysis and prediction of low-visibility events at Valladolid Airport, Spain. *Symmetry* **2020**, *12*, 1045. [CrossRef]

17. Gultepe, I.; Tardif, R.; Michaelides, S.C.; Cermak, J.; Bott, A.; Bendix, J.; Muller, M.D.; Pagowski, M.; Hansen, B.; Ellrod, G.; et al. Fog research: A review of past achievements and future perspectives. *Pure App. Geophys.* **2007**, *164*, 1121–1159. [CrossRef]

18. Han, J.H.; Kim, K.J.; Joo, H.S.; Han, Y.H.; Kim, Y.T.; Kwon, S.J. Sea fog dissipation prediction in Incheon Port and Haeundae Beach using machine learning and deep learning. *Sensors* **2021**, *21*, 5232. [CrossRef]

19. Cornejo-Bueno, S.; Casillas-Pérez, D.; Cornejo-Bueno, L.; Chidean, M.I.; Caamaño, A.J.; Cerro-Prada, E.; Casanova-Mateo, C.; Salcedo-Sanz, S. Statistical analysis and machine learning prediction of fog-caused low-visibility events at A-8 motor-road in Spain. *Atmosphere* **2021**, *12*, 679. [CrossRef]

20. Castillo-Botón, C.; Casillas-Pérez, D.; Casanova-Mateo, C.; Ghimire, S.; Cerro-Prada, E.; Gutierrez, P.A.; Deo, R.C.; Salcedo-Sanz, S. Machine learning regression and classification methods for fog events prediction. *Atmos. Res.* **2022**, *272*, 106157. [CrossRef]

21. Bari, D.; Ameksa, M.; Ouagabi, A. A comparison of datamining tools for geo-spatial estimation of visibility from AROME-Morocco model outputs in regression framework. In Proceedings of the 2020 IEEE International Conference of Moroccan Geomatics (Morgeo), Casablanca, Morocco, 11–13 May 2020; pp. 1–7.

22. Bari, D.; Ouagabi, A. Machine-learning regression applied to diagnose horizontal visibility from mesoscale NWP model forecasts. *Springer Nat. Appl. Sci.* **2020**, *2*, 556. [CrossRef]

23. Goodman, J. The Collection of Fog Drip. *Water Resour. Res.* **1985**, *21*, 392–394. [CrossRef]

24. Shi, W.; Anderson, M.J.; Tulkoff, J.B.; Kennedy, B.S.; Boreyko, J.B. Fog Harvesting with Harps. *ACS Appl. Mater. Interfaces* **2018**, *10*, 11979–11986. [CrossRef]

25. Schemenauer, R.S.; Cereceda, P. A proposed standard fog collector for use in high-elevation regions. *J. Appl. Meteorol. Climatol.* **1994**, *33*, 1313–1322. [CrossRef]

26. Fernandez, D.M.; Torregrosa, A.; Weiss-Penzias, P.S.; Zhang, B.J.; Sorensen, D.; Cohen, R.E.; McKinley, G.H.; Kleingartner, J.; Oliphant, A.; Bowman, M. Fog water collection effectiveness: Mesh intercomparisons. *Aerosol Air Qual. Res.* **2018**, *18*, 270–283. [CrossRef]

27. Montecinos, S.; Carvajal, D.; Cereceda, P.; Concha, M. Collection efficiency of fog events. *Atmos. Res.* **2018**, *209*, 163–169. [CrossRef]

28. Breiman, L. Random forests. *Mach. Learn.* **2001**, *45*, 5–32. [CrossRef]

29. Couronné, R.; Probst, P.; Boulesteix, A.L. Random forest versus logistic regression: A large-scale benchmark experiment. *BMC Bioinform.* **2018**, *19*, 270. [CrossRef]

30. Bergot, T.; Lestringant, R. On the Predictability of Radiation Fog Formation in a Mesoscale Model: A Case Study in Heterogeneous Terrain. *Atmosphere* **2019**, *10*, 165. [CrossRef]

31. Kim, W.; Yum, S.S.; Hong, J.; Song, J.I. Improvement of fog simulation by the nudging of meteorological tower data in the WRF and PAFOG coupled model. *Atmosphere* **2020**, *11*, 311. [CrossRef]

32. Chmielecki, R.M.; Raftery, A.E. Probabilistic visibility forecasting using Bayesian model averaging. *Mon. Weather. Rev.* **2011**, *139*, 1626–1636. [CrossRef]

33. Roquelaure, S.; Bergot, T. A local ensemble prediction system for fog and low clouds: Construction, Bayesian model averaging calibration, and validation. *J. Appl. Meteorol. Climatol.* **2008**, *47*, 3072–3088. Available online: http://www.jstor.org/stable/2617279 1 (accessed on 29 April 2020).

34. Clemen, R.T. Combining forecasts: A review and annotated bibliography. *Int. J. Forecast.* **1989**, *5*, 559–583. [CrossRef]