Lindner, B. L., P. J. Mohlin, A. C. Caulder, and A. Neuhauser, 2018: Development and testing of a decision tree for the forecasting of sea fog along the Georgia and South Carolina Coast. *J. Operational Meteor.*, **6** (5), 47-58, doi: https://doi.org/10.15191/nwajom.2018.0605

Journal of Operational Meteorology
Article

Development and Testing of a Decision Tree for the Forecasting of Sea Fog Along the Georgia and South Carolina Coast

BERNHARD LEE LINDNER
College of Charleston, Charleston, South Carolina

PETER J. MOHLIN
NOAA/National Weather Service Forecast Office, Charleston, South Carolina

A. CLAYTON CAULDER
College of Charleston, Charleston, South Carolina

AARON NEUHAUSER
College of Charleston, Charleston, South Carolina

(Manuscript received 11 December 2017; review completed 16 April 2018)

ABSTRACT

A classification and regression tree analysis for sea fog has been developed using 648 low-visibility (<4.8 km) coastal fog events from 1998–2014 along the South Carolina and Georgia coastline. Correlations between these coastal fog events and relevant oceanic and atmospheric parameters determined the range in these parameters that were most favorable for predicting sea fog formation. Parameters examined during coastal fog events from 1998–2014 included sea surface temperature (SST), air temperature, dewpoint temperature, maximum wind speed, average wind speed, wind direction, inversion strength, and inversion height. The most favorable range in SST for sea fog formation was 10.6–23.9°C. The most favorable gaps between air temperature and SST, dewpoint temperature and SST, and dewpoint temperature and air temperature were found to be –1.7–2.2°C, 0°C, and 0–2.2°C, respectively. The most favorable range in maximum wind speed was 11.1–20.4 km h-1, and the most favorable wind directions were parallel to the coast or SST isopleths. The most favorable range in inversion height was 70.6–617.2 m, and the most favorable inversion strength was anything >6°C. Utilizing these eight predictors, a forecasting decision tree was created and beta tested during the 2016/2017 sea fog season. The decision tree successfully predicted sea fog on 17 of the 18 dates that it occurred (94%) and successfully predicted a lack of sea fog for 189 of the 194 days where sea fog did not occur (97%). Two of the six incorrect predictions appear to have extenuating circumstances.

1. Introduction

Sea fog is primarily a type of advection fog that forms over the ocean when relatively warm, humid air moves over a cooler sea surface resulting in a cooling of the bottom layer of air below its dewpoint (e.g., Taylor 1917) (Fig. 1). Although it is also possible for advection fog to occur via cold air moving over warm water (generally referred to as sea smoke or steam fog), the warm air/cool water variety tends to be much more prevalent. Sea fog can severely impair visibility in coastal areas, causing significant economic losses from delays in aviation and shipping (Garmon et al. 1996). Sea fog is particularly hazardous for mariners, pilots, and motorists in coastal areas; it is estimated that 40% of all accidents at sea in the Atlantic occur during

Corresponding author address: B. Lee Lindner, College of Charleston, 66 George St, Charleston, SC 29424
E-mail: lindnerb@cofc.edu

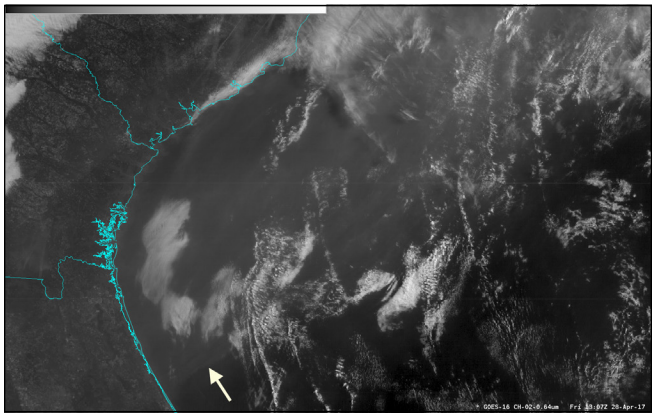

Figure 1. The 1307 UTC 28 April 2017 Geostationary Operational Environmental Satellite-16 Red Visible 0.64 µm (Channel 2) image depicting a typical sea fog event along and offshore of the SC, GA, and FL coasts. The sea fog is depicted as a milky area in the center of the image (arrow points out the edge of the fog bank, which covers most of the middle of the image); reflectances are in the 6–7% range. In comparison, reflectances are in the 3–4% range in the southeastern portions (lower right) of the image, which is more typical of clear sky over water. Note the inhomogeneity of the sea fog, with patches of denser fog. Patches of cirrus, stratus, and altostratus cloud overlay the sea fog, with some shadows of these clouds visible on the sea fog. Image courtesy of Chad Gravelle (NOAA Cooperative Institute for Mesoscale Meteorology Studies). *Click image for an external version; this applies to all figures and tables hereafter.*

dense fog (Trémant 1987). Indeed, one of the first rigorous studies of sea fog followed the Titanic sinking (Taylor 1917).

There are many challenges associated with sea fog forecasting (Koračin 2017). Unlike radiation fog, sea fog is not restricted to light winds; sea fog also tends to affect an area far from where it originated, and it persists for longer (Roach 1995). Sea fog also is difficult to model because of the complexity (Li et al. 2012) and massive differences of scale among its governing processes, ranging from cloud formation processes at the microscale to the synoptic and mesoscale weather events that are conducive to sea fog formation (Koračin et al. 2014).

The two main approaches to sea fog prediction are dynamical and statistical in nature (Koračin 2017). Numerical models for sea fog have greatly improved in recent decades, but there is still somewhat of a disconnect between dynamical simulations and operational applications (Lewis et al. 2004). Provided that a large archive of fog parameters and fog observations exists, statistical approaches are useful and have the advantage of computational efficiency for nowcasting (Koračin 2017). Although limited to the short range, a localized statistical approach to sea fog forecasting is a simple, yet effective, way for operational forecasters to account for local effects and variations that contribute to sea fog formation in a given area. By correlating atmospheric and oceanic parameters with local sea fog events, forecasters can generate reasonably effective forecasting algorithms that can be incorporated into existing systems for that locality.

Artificial neural networks (Fabbian et al. 2007) and fuzzy logic (Hansen 2007; Miao et al. 2012) are statistical techniques that have been used to improve fog forecasting, and other statistical techniques, such as random forests, support vector machine classification, gradient boosting, and K-neighbor clustering have been used for cross-validation of sea fog forecasting (Herman and Schumacher 2016). Classification and regression tree (CART) analyses (e.g., Lewis 2000; Almuallim et al. 2001; Benz 2003) have been increasingly used when standard statistical analyses have failed to find any predictive patterns for forecasting or classifying fog (Wantuch 2001; Lewis 2004; King 2007; Tardif and Rasmussen 2007; Van Schalkwyk and Dyson 2013). CART is a binary recursive partitioning process using a decision tree where the root node splits into two child nodes and then repeats for each child (Breiman et al. 1984; Song and Lu 2015). The number of child nodes is dependent on the number of predictors. Thus, CART analysis uses a decision tree to focus on relevant relationships among parameters (Song and Lu 2015). Decision trees have the advantage of handling complex data and being easy to interpret but have the disadvantage of being prone to overfitting and underfitting (Song and Lu 2015).

There are certain meteorological and oceanic parameters clearly associated with sea fog formation, although the range and the strength of the correlation of these parameters to the formation of sea fog varies significantly with location (Koračin et al. 2005). Sea surface temperature (SST), dewpoint temperature, air temperature, wind speed, and wind direction all play an important role in sea fog formation off the Atlantic coast of Georgia and South Carolina. Particularly close attention must be paid to the interplay between dewpoint temperature and SST (Tang 2012), as sea fog conditions tend to be most favorable when the dewpoint

temperature is greater than the SST (Cho et al. 2000).

The vertical temperature profile, particularly inversion strength and height, also is important for sea fog formation (Koračin et al. 2014). For the purposes of this study, inversion strength is defined as the change in temperature between the top of the inversion layer and the surface, and that distance is the inversion height. The observed link between inversions and fog has been used as a forecasting tool at least since the time of Aristotle and Aratus (Neumann 1989). Although stability is not necessarily a requirement for advection sea fog formation, its likelihood is increased significantly when the air over the ocean is stable; generally, lower altitude inversions are more favorable for sea fog formation (Lewis et al. 2003). Wind speed and direction also play a key role in advection sea fog formation. Although some wind is necessary for advection sea fog formation, too much wind can cause it to dissipate, so sea fog events in a given region are generally most favorable in the presence of weaker winds (Li et al. 2012).

This study focuses on the South Carolina and Georgia coastal region (Fig. 2) for which no existing sea fog studies have been conducted. Sea fog in this region is primarily advective during the late autumn, winter, and early spring, which potentially may be due to differences in temperature between the cooler near-shore or shelf waters and those in the Gulf Stream farther offshore (Fig. 1). Croft et al. (1997) also noted the dominance of advective fog in winter along the coast of the Gulf of Mexico, but the potentially stronger temperature contrast of the waters offshore of the Georgia and South Carolina coast potentially could make advective fog even more dominant for this region. Furthermore, sea fog in this region occurs predominately during the overnight and morning hours, although it has been known to last all day. Commercial and recreational maritime navigation is the primary consideration for the forecasting of sea fog in this region. Thus, the objective of this study is the forecast of sea fog with surface visibilities <4.8 km (3 n mi), the level at which maritime navigation is impacted. Note that, although it is common among mariners and meteorologists to refer to all these events as sea fog, technically visibilities <1 km (0.5 n mi) are sea fog, while visibilities of 1–10 km (0.5–5 n mi) are sea mist. However, most sea mist events in this region contain pockets of sea fog (Fig. 1), and thus in the present study the term sea fog will be loosely applied to both for the sake of brevity.

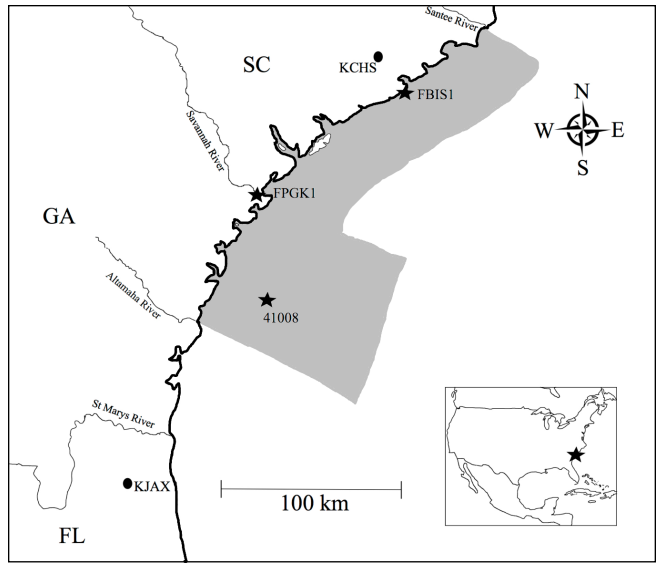

Figure 2. Map of the southeastern United States with buoys, stations, and sounding locations marked (dark line represents the coastline). The area of responsibility for the Charleston, SC NWS forecast office in the coastal waters is shaded light gray. SST data were obtained from buoy 41008 (Grays Reef). Stations FBIS1 (Folly Island, SC) and FPGK1 (Fort Pulaski, GA) were used to obtain temperature, dewpoint, wind speed, and wind direction. Soundings used to obtain upper-air data came from KCHS (Charleston, SC) and KJAX (Jacksonville, FL).

Due in part to the success of King (2007) in forecasting sea fog in the Gulf of Mexico, a CART analysis was used in the present study to identify favorable ranges in predictors for sea fog in the South Carolina and Georgia coastal region, with the goal of developing and testing a decision tree for forecasting sea fog. Flow charts and look-up tables also were examined as potential approaches, but the decision tree was deemed the better option. The following sections detail the study methods, the climatology of sea fog in the region as well as the correlation of predictors to its occurrence, the development of a decision tree based on these predictors, the testing of the decision tree using the 2016/2017 sea fog season, and a discussion of the results.

2. Methods

Plymouth State University Weather Center archives of surface visibility maps of the southeastern United States were used to observe and document low-visibility [<4.8 km (3 n mi) for this application] coastal fog events

along the Georgia and South Carolina coastline (Fig. 2). Meteorological Aerodrome Report (METAR, also known as Aviation Routine Weather Report) surface observations of the present weather and visibility are obtained by Plymouth State through the NOAAPORT data stream, and then Plymouth State archives them with local software. NOAAPORT is a broadcast system that provides a one-way communication of NOAA environmental data and information in near real-time to NOAA and external users. No specific numerical models are used. Specific symbols on the visibility maps distinguish fog from anything else that impairs visibility. If fog covered more than half of the area of responsibility for a minimum of 3–6 h, it was considered a coastal fog event. In addition, the historical record for the National Weather Service Weather Forecasting Office in Charleston, South Carolina, (NWS WFO CHS) was examined to see if a dense fog advisory was issued for those events. The primary weakness with using METAR data for determining coastal fog events is that land-based observations are used with this system, which results in some uncertainty for fog conditions in coastal waters. However, it is likely that sea fog is denser over the neighboring ocean waters than it is over the land-based stations.

The coastal waters for which the NWS WFO CHS is responsible extend from the coast to 37 km (20 n mi) offshore between the South Santee River, South Carolina and Savannah, Georgia, including Charleston Harbor, and from the coast to 111 km (60 n mi) offshore from Savannah to Altamaha Sound, Georgia (Fig. 2). These waters form the area of focus for this study. From 1998–2014, 648 coastal fog events were noted during the local sea fog season of October–April, when the coastal SST is often cooled below the dewpoint of the over-riding air. Croft et al. (1997) noted October–March as a distinct fog season for the coast of the Gulf of Mexico, very similar to what is used in the present study. Prior to July 1998, surface visibility maps were not archived, so data gathering began with that date. For the dates of these 648 coastal fog events, average wind speed, wind direction, temperature, and dewpoint were taken from the National Data Buoy Center for the C-MAN station at Folly Island, South Carolina (FBIS1) and the Georgia Water Level Observation Network station at Fort Pulaski, (FPKG1), and SST data were taken from Station 41008 [Grays Reef, about 74 km (46 mi) southeast of Savannah, Georgia; Fig. 2]. Average values of atmospheric and oceanic parameters taken from FBIS1, FPKG1, and 41008 were calculated by taking the mean of the hourly measurements from 1000–1800Z on the morning of the coastal fog event. It also should be noted that data from FPKG1 were not available from 1998, 2001–2005, and 2010–2014, so only data from FBIS1 were used during these periods. Additionally, dewpoint temperatures were not available during the entire period from FPKG1 because the apparatus that measures dewpoint temperature at that station seems to have been broken. The archives of the National Centers for Environmental Information provided 1200Z Charleston International Airport (KCHS) atmospheric soundings for the morning of each coastal fog event from which inversion strength and inversion height were extracted. On dates when soundings were unavailable at KCHS, soundings at the Jacksonville, Florida, airport (KJAX) were used. KJAX is near the southern border of the area of responsibility for the NWS WFO CHS (Fig. 2) and is thus particularly relevant to the southern part of the area of responsibility. Furthermore, KJAX typically has similar soundings to KCHS on days when sea fog is present, as the prevailing winds travel essentially from Jacksonville to Charleston or vice versa. Because sea fog events have been defined as having a parallel or onshore wind component, it should be safe to assume that sounding data from these locations accurately reflect the upper air profile off the coast.

Because not all coastal fogs are necessarily advection sea fog events, the coastal fog events were divided into three categories: pure sea fog (having winds parallel to the coastline or an onshore wind component), non-sea fog (radiation fog, descending stratus cloud, advection fog, etc., that formed onshore and later moved offshore), and a mixture of pure sea fog and non-sea-fog (having winds that were onshore or parallel to the coast during only a portion of the time analyzed). Coastal fog events were categorized by the direction of surface winds at FPKG1 and FBIS1 on the morning of the fog event (1000–1800Z). Winds approximately parallel to the coastline and SST isopleths (160–200° azimuth along Georgia coastal waters and 180–220° along South Carolina coastal waters) were determined to be pure sea fog events. For a few events that advected from the northeast, winds were taken to be approximately parallel to the coastline and SST isopleths for 360–10° azimuth along Georgia coastal waters and 30–40° along South Carolina coastal waters. Although less frequent, fog events with an onshore wind component (20–150° azimuth along the Georgia coast and 50–170° azimuth along the South Carolina coast) also were taken to be

pure advection sea fog events.

Decision trees are prone to overfitting and underfitting when using a small dataset (Song and Lu 2015). To increase sample size, this present study used a large visibility threshold for sea fog, included incomplete years of data and mixed fog events, and used KJAX and FPKG1 data to supplement KCHS and FBIS1 data. These measures could contribute noise to the data. However, it is worth noting that aside from the benefit of increasing sample size, there also are physical reasons for these measures. The distance from the northern to southern ends of the area of responsibility for the NWS WFO CHS is approximately 250 km (155 mi), and the KCHS and FBIS1 stations are near the northern end (Fig. 2). Thus, the addition of data from KJAX and FPKG1 (both near the southern end of the area of responsibility) helps minimize the bias that would occur if only KCHS and FBIS1 data were used, particularly as the KJAX and FPKG1 data are underweighted. In addition, marine navigation is impacted by mixed sea fog as well as pure sea fog. Because it is desirable to forecast both types, mixed sea fog events were included in the climatology. Furthermore, using a 4.8 km threshold for sea fog events matches the standard used for marine navigation. Moreover, sea fog is likely denser over the neighboring ocean waters than it is over the METAR stations. Finally, excluding cases with missing values is inefficient and runs the risk of introducing bias (Song and Lu, 2015).

Two sea fog decision trees that could be used for operational forecasting were developed based on the results obtained for the parameters mentioned above. The first decision tree included the additional parameter of average wind speed in the inversion layer. However, this parameter was determined to have only a weak correlation to the occurrence of sea fog. Average wind speed for sea fog events ranged from 2 kt to 40 kt (2–46 mph). Even if extremes were eliminated, there was no mode or most frequent values. In other words, the standard deviation was too high for average wind speed to be a useful predictor. Thus, the decision tree was pruned to remove that parameter, which resulted in the second and final decision tree. The decision tree was then beta tested during the 2016/2017 sea fog season.

3. Data

Out of the 648 coastal fog events, 224 events were determined to be pure sea fog, 179 events were non-sea fog, and 245 events were mixed sea fog. Based upon fifteen years of experience forecasting sea fog in the NWS WFO CHS, eight meteorological and oceanic parameters were examined to see which showed correlation with either pure or mixed sea fog events. As discussed above, average wind speed in the inversion layer, was discarded because of a weak correlation with sea fog events. Of the remaining seven potential parameters, some showed stronger correlations if only pure sea fog events were used, while other parameters showed stronger correlations if both pure and mixed sea fog events were used. In addition, as noted above, data availability was often compromised for some parameters, and using both pure and mixed sea fog events in those cases was deemed necessary to increase sample size. The seven parameters were plotted, and outliers were discarded. The ranges in relationships that showed the strongest correlations were used in creating the decision tree.

Favorable ranges of temperature-related predictors were calculated as the mean value plus or minus one standard deviation of the 224 pure sea fog events. In terms of SST, the most favorable ranges for sea fog formation were 10.6–23.9°C (51–75°F). This range is primarily a reflection of the time of the year that sea fog tends to occur. This range is similar to findings for the Northern Gulf of Mexico (King 2007), although the present study includes some warmer SST. Sea fog formation was found to be more likely when air temperature was >1.7°C (3°F) below the SST and <2.2°C (4°F) above the SST (Fig. 3). Air temperature was not available at both stations for all sea fog events; hence this parameter was not computed for every pure sea fog event. As also noted by King (2007), fog formation was most favored when the SST was less than a fraction of a degree from the dewpoint temperature, although a significant number of fog events did occur with SST as much as 2°C above or below the dewpoint temperature (Fig. 4).

Favorable ranges for dewpoint depression, maximum wind speed, inversion height, and inversion strength were calculated as the mean value plus or minus one standard deviation of the 469 pure sea fog and mixed sea fog events. Sea fog formation is favored when the dewpoint depression (the difference between the air temperature and the dewpoint temperature) ranges from 0°C to 2.2°C (0–4°F) (Fig. 5), slightly larger than that noted for the Northern Gulf of Mexico (King 2007). Maximum wind speeds were typically light to moderate, with a favorable range of 11.1–20.4 km h^{-1} (6–11 kt) (Fig. 6). The King (2007) study used a similar

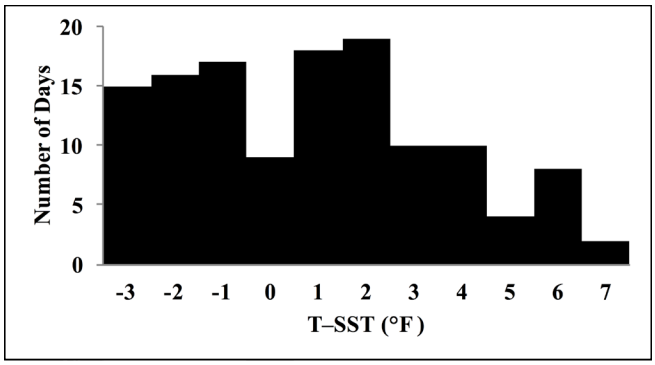

Figure 3. Average temperature (T, °F) recorded at FPGK1 and FBIS1 minus average SST (°F) recorded at buoy 41008 (Grays Reef) from 1000–1800Z on the morning of a given pure sea fog event. Air temperature was not available at both stations for all sea fog events; hence there are fewer than 224 data points. The mean value is –0.8°F, and the standard deviation is 3.8°F.

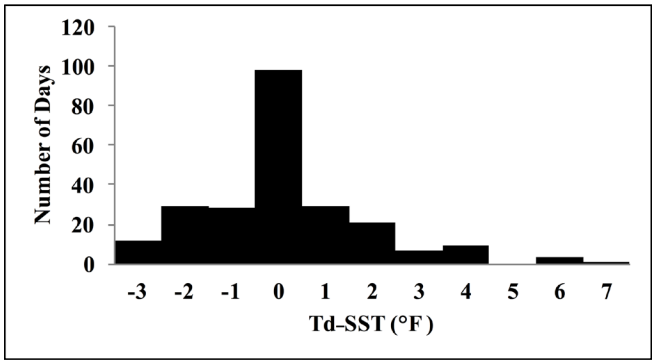

Figure 4. Average dewpoint temperature (Td, °F) recorded at FPGK1 and FBIS1 minus average SST (°F) recorded at buoy 41008 (Grays Reef) from 1000–1800Z on the morning of a given pure sea fog event. The mean value is 0.1°F, and the standard deviation is 1.8°F.

upper limit but had no lower limit. The most favorable inversion strength was anything <6°C (42.8°F) (Fig. 7), and the most favorable range for inversion height was 71–617 m (233–2024 ft) (Fig. 8). However, sea fog events with inversion height below 70 m (230 ft) may have been undercounted as such low opacity fogs are difficult to detect remotely and are frequently patchy. Nonetheless, such shallow fog events aren't a major concern in general. Means, standard deviations, and ranges for noteworthy predictors of sea fog are shown in Table 1.

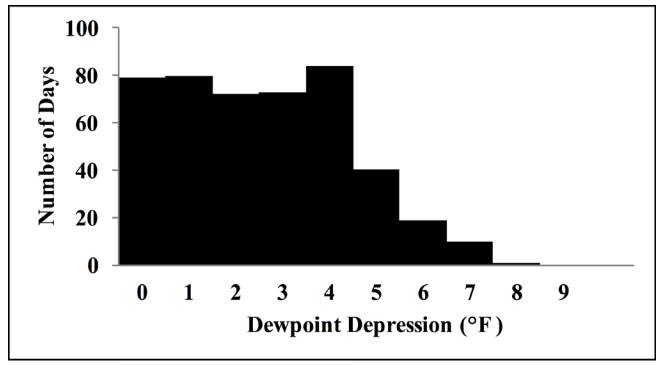

Figure 5. Average dewpoint depression (in °F) recorded at FPGK1 and FBIS1 from 1000–1800Z on the morning of a given pure or mixed sea fog event. The mean value is 2.6°F and the standard deviation is 1.9°F.

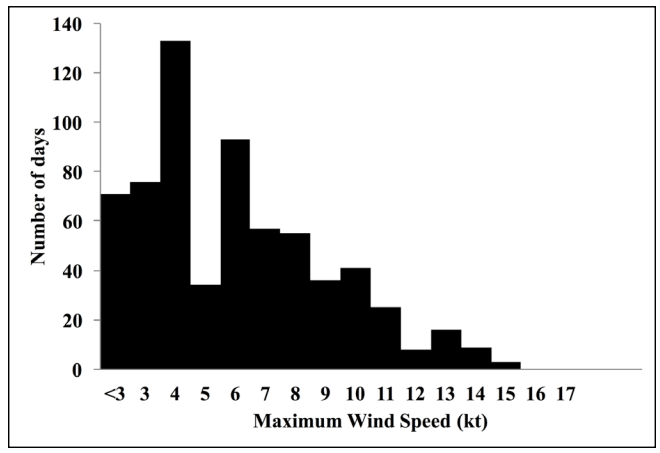

Figure 6. Maximum wind speed (kt) recorded at FPGK1 and/or FBIS1 from 1000–1800Z on the morning of any given coastal fog event. The mean value is 6.0 kt, and the standard deviation is 5.4 kt.

4. Analysis

Pure sea fog events off the coast of Georgia and South Carolina are most prevalent from February through April, while mixed sea fog events (those that had pure sea fog for part of the coastal fog event) are more evenly distributed over the season. The higher frequency of pure sea fog events from February through April may be because, throughout the study period, a higher frequency of south and southwest winds occurred from February through April, while October through January had a more normal distribution of wind directions (resulting in more mixed fog events). It is difficult to draw further climatological conclusions because of the missing or limited data from 1998–2000 and 2006–2009 and missing data from the surface

Table 1. Mean, standard deviation (SD), and range in noteworthy predictors of sea fog during the period 1998–2014 off the coast of Georgia and South Carolina. Also shown are the ranges in predictors that were used in the decision tree for a highly likely (HL) prediction of sea fog and a likely (L) prediction of sea fog, along with the percentage of sea fog events (POE) that fell within the ranges listed for HL (POE-H) and for L (POE-L).

Variable	Mean	SD	Range	HL	POE-H	L	POE-L
Td–SST (°F)*	0.1	1.8	–3–7	–2–2	85%	–4–4	98%
Inversion Height (ft)	1260	851	100–6075	275–2000	86%	275–3000	98%
Inversion Strength (°C)	4.0	2.8	1–11	1–6	76%	1–10	98%
SST (°F)*	62.5	8.1	41–82	51–75	89%	47–83	98%
T–SST (°F)*	–0.8	3.8	0–15	–3–3	60%	–5–5	91%
Dewpoint Depression (°F)	2.6	1.9	0–8	0–4	66%	0–7	100%
Maximum Wind Speed (kt)	6.0	5.4	0–15	6–11	52%	0–15	99%

*Td=dewpoint temperature, SST=sea surface temperature, T=atmospheric temperature

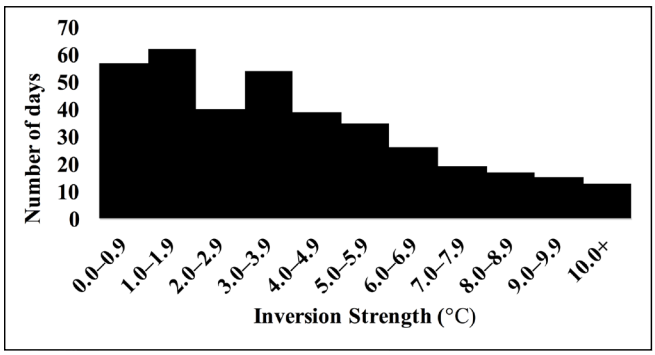

Figure 7. Inversion strength (°C) from the 1200Z sounding at KCHS or KJAX on the morning of a given pure or mixed sea fog event. Soundings were not always available, so there are fewer than 469 data points. The mean value is 4.0°C, the standard deviation is 2.8°C, and the range is 1–11°C.

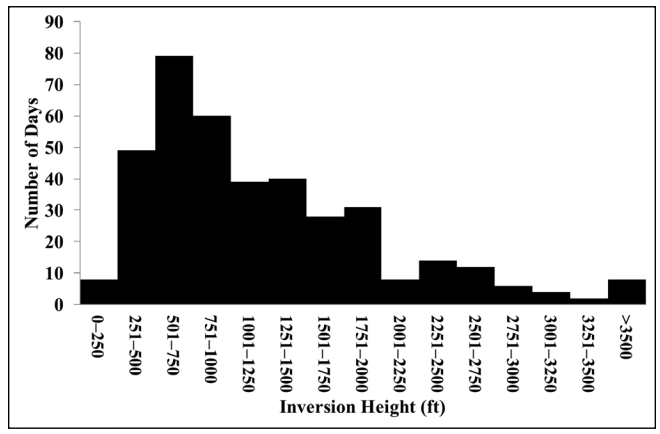

Figure 8. Inversion heights (ft) from 1200Z sounding at KCHS or KJAX on the morning of a given pure or mixed sea fog event. Soundings were not always available, so there are fewer than 469 data points. The mean value is 1260 ft, the standard deviation is 851 ft, and the range is 100–6075 ft.

visibility and buoy databases. Thus, any climatological conclusions (particularly ones regarding the annual frequency of sea fog) should be made with caution.

The favorable ranges of predictors of sea fog formation were used to create a decision tree (Fig. 9) that can be used to determine the general likelihood of sea fog occurring somewhere within the NWS WFO CHS forecast region. The decision tree has three branches, the first of which has favorable ranges targeted towards pure sea fog with winds parallel to the coast; if conditions fall within this range, sea fog is highly likely (probability of sea fog occurring is 80% or higher). The second branch has a broader set of favorable ranges targeted towards pure or mixed sea fog; conditions within this range suggest that sea fog is likely (probability of sea fog occurring is 50% or higher). The third and final branch is for parameters outside of the favorable ranges; in this case, sea fog is not likely (probability of sea fog occurring is <50%). These probability-based definitions for "highly likely," "likely," and "not likely" follow the protocol used at the NWS WFO CHS when using these terms in the issuance of dense fog advisories. Experience in forecasting sea fog in the region has noted that parameters such as wind direction, the difference between dewpoint temperature and SST, inversion height and strength, and SST are more important predictors. Thus, the selected ranges in these parameters were more correlated to sea fog events, and the highly likely branch of the decision tree used ranges that included 75–90% of the sea fog events

(Table 2). Other parameters are less correlated to sea fog events; thus, the highly likely branch of the decision tree used ranges in these parameters that included 50–70% of the sea fog events. For the likely branch of the decision tree, ranges included more than 90% of sea fog events for each parameter, with most including more than 98% of sea fog events (Table 2).

a. Beta testing during the 2016–2017 sea fog season

The decision tree (Fig. 9) was tested during the sea fog season of 1 October 2016–30 April 2017, a total of 212 days. Overall, it was an interesting sea fog season with a few unusual events. Six pure sea fog events and six mixed sea fog events were recorded, which is significantly below that recorded in recent years. A typical year from the prior 15 years had 15–20 events of pure and mixed sea fog, each. Below-normal occurrences of favorable south to southwest winds and above-normal occurrences of unusually dry air seem to have been responsible for the reduced number of sea fog events during the period for the beta test. The season also got off to a slow start with no events until the middle of November, and one event occurred right at the end of the season. There were approximately 204 total hours of sea fog during the season within the NWS WFO CHS forecast area. Half of the sea fog events during this season lasted at least 22 h, with the longest continual sea fog event lasting 36 h. The beta test included all sea fog events that lasted a minimum of 3–6 h, with either a sea fog covering a minimum of half of the area of responsibility (Fig. 2) or any localized sea fog that significantly impacted marine navigation, particularly Charleston Harbor or the Port of Savannah. A more detailed description of each event follows. "Nearshore waters" extend from the coast to 20 n mi offshore of the South Carolina and Georgia coastlines and "all waters" includes Charleston Harbor, the waters along the South Carolina coastline extending from the coast to 37 km (20 n mi) offshore, and the waters extending from the coast to 111 km (60 n mi) offshore of the Georgia coastline:

14 November 2016: A mix of pure sea fog with land fog that advected over the sea and impacted all waters with visibility of 2–6 km (1–3 n mi) and patches where visibility was less.

30 November 2016: The first pure sea fog event impacted all waters with visibility of 2–6 km (1–3 n mi) with patches where visibility was less.

Table 2. The dates within the 2016/2017 sea fog season when sea fog was forecasted by the decision tree and the dates when sea fog was observed (occurrence noted by an X and green highlighting; lack of occurrence noted by red highlighting and the lack of an X). Sea fog occurrences were widespread except for the 1 February and 27–28 April events when sea fog was patchy.
* Sea fog occurred initially but dissipated sooner than expected; classified as a correct forecast because sea fog existed in the morning.
** Very isolated and brief event near the mouth of the Savannah River that lasted 4 h; classified as a missed forecast because some sea fog did occur.
*** Sea fog did not occur initially because of active thunderstorms. Once the thunderstorms dissipated (around 1700 UTC 28 February), sea fog did form; classified as a missed forecast because sea fog was absent in the morning.

DATE IN 2016/17	SEA FOG FORECASTED	SEA FOG OCCURRED
Nov 14	X	X
Nov 29	X	
*Nov 30	X	X
Dec 12	X	X
Dec 14	X	X
Dec 17	X	X
Dec 18	X	X
Dec 19	X	X
Dec 25	X	X
Dec 26	X	X
Jan 11	X	
Jan 12	X	
Jan 14	X	X
Jan 15	X	X
Jan 16	X	X
Jan 17	X	X
**Feb 1		X
Feb 7	X	
Feb 25	X	X
***Feb 28	X	
Mar 1	X	X
Apr 27	X	X
Apr 28	X	X

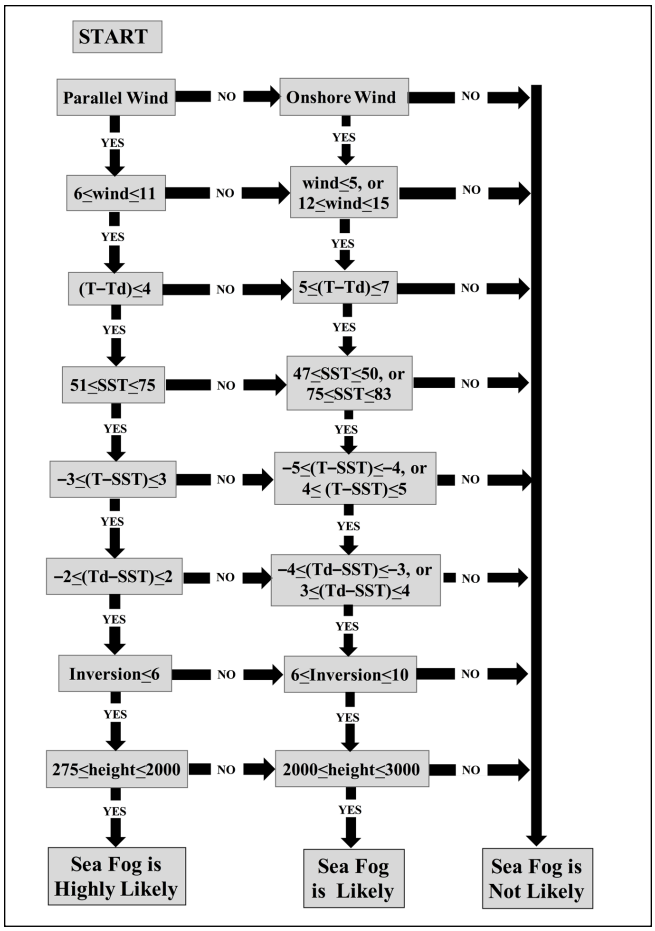

Figure 9. A decision tree for the prediction of sea fog occurring somewhere within the coastal region between Altamaha Sound, GA and South Santee River, SC (the area of responsibility for the NWS Charleston Forecasting Office). Beginning at the upper left of the tree ("START"), a prediction of the existence of sea fog is produced based upon measurements of several parameters such as the wind velocity [wind direction is "Parallel Wind" if the azimuth is 180–220° along SC or 160–200° along GA (30–40° or 0–10°, respectively, if advected from the northeast); "Onshore Wind" if the azimuth is 50–170° along SC or 20–150° along GA; "wind," maximum wind speed rounded to nearest kt)]. Other decision parameters include temperature ("T," rounded to nearest °F), dewpoint temperature ("Td," rounded to nearest °F), sea surface temperature ("SST," rounded to nearest °F), inversion strength ("Inversion," °C) and inversion height ("height," measured in ft). Sea fog is considered "highly likely" if all the conditions in the first column are met and is "likely" if all the conditions in the second column are met or a mix of conditions from the first and second columns are met. If any of the conditions are not met in either the first or second column, sea fog is "not likely."

12 December 2016: Sea fog impacted all nearshore Atlantic waters with visibility of 1 km (0.5 n mi) or less.

14 December 2016: A mix of pure sea fog with land fog that advected over the sea and impacted Charleston Harbor and the nearshore waters off Charleston County with visibility of 1–2 km (0.5–1 n mi) and patches where visibility was less.

17–19 December 2016: 29 h of continual sea fog that impacted all waters with visibility of 0.5–1 km (0.25–0.5 n mi) and patches where visibility was near 0.

25–26 December 2016: 28 h of continual sea fog arriving from the north and northeast impacted all nearshore waters and Charleston Harbor with visibility of 2–4 km (1–2 n mi) and patches where visibility was less.

14–15 January 2017: 22 h of continual sea fog impacted all nearshore waters and Charleston Harbor with visibility of 0.5 km (0.25 n mi) or less and patches where visibility was near 0.

16–17 January 2017: 36 h of continual sea fog impacted all waters with visibility of 0.5 km (0.25 n mi) or less with patches where visibility was near 0. There was a 10-h gap between the prior sea fog event and this event.

1 February 2017: Very isolated sea fog that occurred in and near the mouth of the Savannah River and impacted the waters off the Georgia coast out about 9 km (5 n mi) with visibility of 1–2 km (0.5–1 n mi). Sea fog only lasted about 4 h.

25 February 2017: Sea fog that impacted all nearshore waters and Charleston Harbor with visibility of 1–2 km (0.5–1 n mi).

28 February–1 March 2017: 26 h of continual sea fog impacted all nearshore waters and Charleston Harbor with visibility of 2– 6 km (1–3 n mi). Sea fog began late in the day (1700 UTC) on 28 February owing to active thunderstorms in the morning.

27–28 April 2017: 22 h of continual sea fog impacted nearshore waters of South Carolina with visibility 0.5 km (0.25 n mi) or less. However, sea fog was patchy in extent. This event is pictured in Fig. 1.

Although the decision tree can be used to make a forecast every hour, for the purposes of this beta test only daily forecasts were used. For each date from

1 October 2016–30 April 2017, data from the three surface monitoring stations and atmospheric soundings were used to test each step in the sea fog prediction algorithm (Fig. 9), and a forecast was made as to the occurrence of sea fog for that day. For the purposes of this beta test, if the decision tree predicted either that sea fog was highly likely or likely, this was taken as a prediction that sea fog would occur. If the decision tree predicted that sea fog was not likely, this was taken as a prediction that sea fog would not occur. As a general forecasting practice, the decision tree could be used more effectively by issuing forecasting statements to the public that "sea fog is highly likely," "sea fog is likely," or "sea fog is not likely," but for the purposes of this beta test, the decision tree was turned into a yes-or-no prediction to facilitate the test. As shown in Table 2, the decision tree forecast sea fog 22 times during the season; 17 of those 22 days had an occurrence of sea fog (77% success). On the remaining 190 days, the decision tree predicted no sea fog; 189 of those days had no sea fog (99% success rate). It is worth noting that the lone missed event was small in areal coverage (near the mouth of the Savannah River) and only lasted 4 h, making it difficult to forecast under any scenario. In total, 212 forecasts were made with the decision tree, and 206 forecasts were correct (97% success rate). The decision tree successfully forecast 17 of the 18 days on which sea fog occurred (94%) and successfully forecast 189 out of the 194 days when no sea fog occurred (97%). It also is worth noting that the sea fog event on 27–28 April 2017 was initially not forecast by the NWS WFO CHS, but the decision tree did predict it.

5. Discussion

Decision trees have increasingly come into use for forecasting fog (Wantuch 2001; Lewis 2004; King 2007) and differentiating types of fog (Tardif and Rasmussen 2007; Van Schalkwyk and Dyson 2013). Although all sea fog forecasting decision trees utilize many similar predictors, each tree is quite distinct as the underlying geography and meteorology is unique to each location. Unsurprisingly then, the decision tree for coastal Georgia and South Carolina (Fig. 9) is similar to, but different from, earlier sea fog forecasting decision trees. Wantuch (2001) developed decision trees for fog in Budapest using dewpoint depression, wind speed, and inversion strength as predictors (although that was not a sea fog study). King (2007) developed decision trees for sea fog in the northern Gulf of Mexico using the same predictors as used in this study except for the difference between temperature and SST and using no numerical value for inversion strength or height. Lewis (2004) developed decision trees for sea fog in Korea using the same predictors used in this study except for adding temperature and dewpoint temperature as predictors and removing inversion height. In addition to using slightly different predictors in their decision trees, each study uses different numerical cutoffs for those predictors. Hence, even though similar parameters may be relevant, each region will likely have different correlations and a unique decision tree that is best suited to forecasting sea fog in that region.

Although the decision tree in this study had success in forecasting sea fog, recommendations are suggested to improve upon it. Additional years of climatological data would increase the sample size and could allow for better correlation of the underlying parameters to the occurrence of sea fog. More coastal and offshore stations would provide better data upon which to build the analysis. An analysis as to the reasons the decision tree failed to successfully predict sea fog for a few dates in 2016/17 might result in improved predictor selection. New parameters could be explored as potential predictors. For example, the base height of the inversion layer or using only sea fog events that had southwest winds could be important predictors. A more extensive statistical examination of the current predictors could help optimize the decision tree. The relative importance of predictors could then be assessed. A study could determine if the decision tree is more accurate with a smaller dataset that omits cases where there were missing data or mixed fog events. The decision tree could be pruned by eliminating weaker predictors, which may increase efficiency. Multiple decision trees could use different isolines for visibility to potentially create gradated fog predictions. The use of the parameters of inversion strength and inversion height could be explored for potentially forecasting sea fog duration.

Additional testing in future years could be more definitive in determining the accuracy of the decision tree. Autocorrelation tests and cross validation with other statistical techniques also could help determine the accuracy of the decision tree. Contingency tables and associated forecast verification metrics could be useful. Alternate statistical techniques using the same data could be tested to see if they perform better than a decision tree. Real-time relationships might be extendable into short-term forecasts. The climatology of

sea fog could be more closely examined for correlations with month, time of day, or duration of the fog event. Finally, the decision tree could be integrated into the NWS Graphical Forecast Editor forecast process.

The forecasting of sea fog is just one part of ongoing research to improve the efficacy of the NWS WFO CHS. Previous investigations of the sea breeze (Frysinger et al. 2003) and the synoptic climatology of severe weather (Alsheimer and Lindner 2011) for the Charleston area resulted in improved forecasting of those phenomenon. Similarly, studies of the climatology of tropical cyclones in the Charleston area has found some patterns that may be useful for long-term forecasting (Lindner and Neuhauser 2018). The sea breeze and tropical cyclones also have impacted precipitation received in the Charleston area (Lindner and Frysinger 2007). Lindner and Cockcroft (2013) and Lindner et al. (2018a, 2018b) have shown that indexical images can enhance public recognition of the concept and hazard of hurricane storm surge for the Charleston area. Similarly, indexical images could be used to better demonstrate sea fog density and occurrence patterns to the public and the various agencies and industries that rely on these sea fog forecasts.

6. Conclusion

CART (decision tree) analyses lack a connection to the underlying principles but are capable of forecasting localized sea fog events and are simple to interpret, assuming an archive of historical sea fog, meteorological, and oceanic data are available. For the coastal region between Altamaha Sound and the South Santee River, a sea fog decision tree has been developed that uses eight predictors that correlated with the occurrence of sea fog over a sixteen-year period. These sea fog predictors included wind direction, wind speed, dewpoint depression, SST, air temperature minus SST, dewpoint temperature minus SST, inversion height, and inversion strength. The decision tree featured three branches for forecasting whether sea fog is highly likely, likely, or not likely, based upon favorable ranges in the eight predictors.

The October 2016–April 2017 sea fog season was used to beta test the decision tree. This was an exceptionally inactive season for sea fog with only 18 dates on which sea fog occurred. The sea fog decision tree successfully forecasted sea fog on 17 of those dates and successfully forecasted no fog on 189 of the 194 dates when none occurred. The decision tree also forecasted sea fog on 5 dates that no fog occurred. Thus, the decision tree was much better at predicting sea fog would not form (over 99% success) than it was at predicting sea fog (77%). This indicates that the conditions used in the decision tree err on the side of caution. It is worth noting that 2 of the 6 incorrect forecasts had extenuating circumstances that may explain some of the inaccuracy.

Acknowledgments: Thanks to the staff at the Plymouth State University Weather Center, the National Centers for Environmental Information, and the National Data Buoy Center for creating access to their publicly available data. Chad Gravelle is thanked for providing Fig. 1. The reviewers, the editor, Steven Rowley, and Brian Miretzky are thanked for thoughtful comments that resulted in an improved paper.

REFERENCES

Almuallim, H., S. Kanada, and Y. Akiba, 2001: "Development and applications of decision trees". In *Expert Systems: The Technology of Knowledge Management and Decision Making for the 21st Century, Volume 1*, edited by C. T. Leondes. Academic Press, 53–57.

Alsheimer, F., and B. L. Lindner, 2011: Synoptic-scale precursors to high-impact weather events in the Georgia and South Carolina coastal region. *J. Coastal Res.*, **27** (2), 263–275, Crossref.

Benz, R. F., 2003: Data mining atmospheric/oceanic parameters in the design of a long-range nephelometric forecast tool. Master's thesis, Air Force Institute of Technology, 103 pp.

Breiman, L., J. H. Friedman, R. A. Olshen, and C. J. Stone, 1984: *Classification and Regression Trees*. Chapman & Hall/CRC, 358 pp.

Cho, Y-K., M-O. Kim, and B-C. Kim, 2000: Sea fog around the Korean peninsula. *J. Appl. Meteorol.*, **39**, 2473–2479, Crossref.

Croft, P. J., R. L. Pfost, J. M. Medlin, and G. A. Johnson, 1997: Fog forecasting for the southern region: A conceptual model approach. *Wea. Forecasting*, **12**, 545–556, Crossref.

Fabbian, D., R. de Dear, and S. Lellyett, 2007: Application of artificial neural network forecasts to predict fog at Canberra International Airport. *Wea. Forecasting*, **22**, 372–381, Crossref.

Frysinger, J. R., B. L. Lindner, and S. L. Brueske, 2003: A statistical sea-breeze prediction model for Charleston, South Carolina. Wea. Forecasting, 18, 614–625, CrossRef.

Garmon, J., D. Darbe, and P. J. Croft, 1996: Forecasting significant fog on the Alabama coast: Impact, climatology, and forecast checklist development. NWS Tech. Memo. NWS SR-176. Scientific Services Division, Southern Region, Fort Worth, TX,16 pp. [Available online at repository.library.noaa.gov/view/noaa/6349.]

Hansen, B., 2007: A fuzzy logic-based analog forecasting system for ceiling and visibility. *Wea. Forecasting*, **22**, 1319–1330, Crossref.

Herman, G. R., and R. S. Schumacher, 2016: Using reforecasts to improve forecasting of fog and visibility for aviation. *Wea. Forecasting*, **31**, 467–482, Crossref.

King, J. M., 2007: A detailed study of advection sea fog formation to reduce the operational impacts along the Northern Gulf of Mexico. M. S. thesis, Naval Postgraduate School, 96 pp. [Available online at calhoun.nps.edu/bitstream/handle/10945/3579/07Mar_King_Jason.pdf?sequence=1&isAllowed=y.]

Koračin, D., 2017: Modeling and forecasting marine fog. In *Marine Fog: Challenges and Advancements in Observations, Modeling, and Forecasting*, Edited by D. Koračin and C. E. Dorman. Springer, 425–475.

Koračin, D., J. A. Businger, C. E. Dorman, and J. M. Lewis, 2005: Formation, evolution, and dissipation of coastal sea fog. Boundary-Layer Meteorol., 117, 447–478, Crossref.

Koračin, D., C. E. Dorman, J. M. Lewis, J. G. Hudson, E. M. Wilcox, and A. Torregrosa, 2014: Marine fog: A review. *Atmos. Res*., **143**, 142–175, Crossref.

Lewis, D. M., 2004: Forecasting advective sea fog with the use of classification and regression tree analyses for Kunsan Air Base. M. S. thesis, Air Force Institute of Technology, 90 pp. [Available online at www.dtic.mil/dtic/tr/fulltext/u2/a422963.pdf.]

Lewis, J., D. Koračin, R. Rabin, and J. Businger, 2003: Sea fog off the California coast: Viewed in the context of transient weather systems. *J. Geophys. Res*., **108**, 6-1–6-17, Crossref.

Lewis, J. M., D. Koračin, and K.T. Redmond, 2004: Sea fog research in the United Kingdom and United States: A historical essay including outlook. *Bull. Amer. Meteor. Soc*., **85**, 395–408, Crossref.

Lewis, R. J., 2000: An introduction to classification and regression tree (CART) analysis. *2000 Annual Meeting of the Society for Academic Emergency Medicine*, San Francisco, CA, 14 pp.

Li, P., G. Fu, and C. Lu, 2012: Large-scale environmental influences on the onset, maintenance, and dissipation of six sea fog cases over the Yellow Sea. *Pure Appl. Geophys*., **169**, 983–1000, Crossref.

Lindner, B. L., and J. R. Frysinger, 2007: Bulk atmospheric deposition in the Charleston Harbor watershed. *J. Coastal Res*., **23**, 1452–1461, Crossref.

Lindner, B. L., and C. Cockcroft, 2013: Public perception of hurricane-related hazards. In Coastal Hazards, edited by C. W. Finkl. Springer, pp. 185–210.

Lindner, B. L., and A. Neuhauser, 2018: Climatology and variability of tropical cyclones impacting Charleston, South Carolina. *J. Coastal Res*., in press.

Lindner, B. L., F. Alsheimer, and J. Johnson, 2018a: Assessing improvement in the public's understanding of hurricane storm tides through interactive visualization models. *J. Coastal Res*., in press.

Lindner, B. L., J. Johnson, F. Alsheimer, S. Duke, G. D. Miller, and R. Evsich, 2018b: Increasing risk perception and understanding of hurricane storm tides using an interactive, web-based, visualization approach. *J. Coastal Res*., in press.

Miao, Y., R. Potts, X. Huang, G. Elliot, and R. Rivett, 2012: A fuzzy logic fog forecasting model for Perth airport. *Pure Appl. Geophys*., **169**, 1107–1119, Crossref.

Neumann, J., 1989: Forecasts of fine weather in the literature of classical antiquity. *Bull. Amer. Meteor. Soc*., **70**, 46–48, Crossref.

Roach, W. T., 1995: Back to basics: Fog: Part 3 - The formation and dissipation of sea fog. *Weather*, **50**, 80–84, Crossref.

Song, Y-Y., and Y. Lu, 2015: Decision tree methods: Applications for classification and prediction. *Shanghai Arch. Psychiatry*, **27**, 130–135. [Available online at www.ncbi.nlm.nih.gov/pmc/articles/PMC4466856/.]

Tang, Y., 2012: The effect of variable sea surface temperature on forecasting sea fog and sea breezes: A case study. *J. Appl. Meteor. Climatol*., **51**, 986–1440, Crossref.

Tardif, R., and R. M. Rasmussen, 2007: Event-based climatology and typology of fog in the New York City region. *J. Appl. Meteor. Climatol*., **46**, 1141–1168, Crossref.

Taylor, G. I., 1917: The formation of fog and mist. Quart. *J. Roy. Meteor. Soc*., **43**, 241–268, Crossref.

Trémant, M., 1987. La prévision du brouillard en mer. Météorologie Maritime et Activities Océanographique Connexes, Rapport No. 20. TD no. 211, World Meteorological Organization, Geneva, Switzerland, 34 pp.

Van Schalkwyk, L., and L. L. Dyson, 2013: Climatological characteristics of fog at Cape Town International Airport. *Wea. Forecasting*, **28**, 631–646, Crossref.

Wantuch, F., 2001: Visibility and fog forecasting based on decision tree method. *IDOJARAS*, **105**, 29–38.