

OPERATIONAL CONCEPT FOR OBSERVATION-BASED FORECASTING OF CLEAR SKY CONDITION

Timothy J. Hall, Carl N. Mutchler, Greg J. Bloy, Rachel N. Thessin, Stephanie K. Gaffney and J. J. Lareau
The Aerospace Corporation
15049 Conference Center Drive
Chantilly, VA 20151

1. INTRODUCTION

A range of military, civil and commercial activities benefit from cloud-free sky conditions. Passive optical or thermal systems, such as those on unmanned aerial systems, need a cloud-free-line-of-sight in order to sense their targets (Norquist 1999). Solar energy available to fuel photovoltaic power plants is strongly modulated by clouds (Girodo et al. 2006) due to their ability to reflect incoming shortwave radiation. Additionally, electricity demand on power grids is correlated to the amount of solar irradiance. Very short-range (i.e., up to 6-hr) sky condition forecasts are useful to decision makers for these applications.

In this paper, we present a concept for global, obs-based very short-range sky condition (i.e., total cloud or cloud-free amount) forecasting. This system would utilize meteorological satellite (METSAT) data and data-assimilation derived analyses to build feature databases for training advanced predictive learning algorithms. Such a system would be potentially valuable for global, point-specific forecasts – particularly in data-sparse regions.

The 0 to 6-hr time frame is a sweet-spot for obs-based (i.e., empirical) weather forecasting techniques (Bankert and Hadjimichael 2007, Hansen 2007, Vislocky and Fritsch 1997). The most primitive, obs-based forecast is basic persistence (BP) (i.e., the future weather condition will be the same as the current weather condition). The climatological-expectancy-of-persistence (CEP), also called the persistence probability, combines BP with climatology (Enger et al. 1962). CEP was developed as an objective tool for operational forecasters to help them predict future cloud

ceiling height and visibility by matching an initial condition with historical conditions and categorizing the initial condition in terms of stratified climatological data. Common stratification variables include month, time of day, wind direction, ceiling height, and visibility. Climatological forecasting is an empirical method based on the statistics of average historical weather conditions.

Satellite cloud climatologies (SCC) can be used to make climatological forecasts based on historical averages. Kelly (1988) introduced the concept of using digital satellite cloud imagery to generate wind-stratified, CEP-based cloud forecasts. CEP forecast techniques have been applied (or at least suggested) by a number of authors including Combs et al. (2004), Connell et al. (2001), Hall et al. (1998), and Reinke et al. (2003).

Advanced, obs-based techniques, driven by the analysis of current and historical data, have emerged from a number of technical disciplines including statistics, applied mathematics, artificial intelligence, cognitive psychology, engineering, knowledge discovery in databases, and meteorology. Non-parametric, predictive learning algorithms, trained on past surface observations have been applied to cloud forecasting by numerous investigators (e.g., Wiener et al. 2004, Bankert and Hadjimichael 2007). Hansen (2007) used a statistical method based on a k-nn algorithm to predict ceiling height and visibility.

In the literature, no advanced, obs-based sky condition forecast techniques have been developed based on meteorological satellite information. In two companion conference papers (Hall et al. 2010a, Hall et al. 2010b), research results are presented describing the development and demonstration of a number of obs-based prediction algorithms, based primarily on meteorological satellite data, that performed very well against several benchmark algorithms. A small subset of research results are presented in this paper as evidence of the viability of the

Corresponding author address: Tim Hall, The Aerospace Corporation, 8455 Colesville Rd., Silver Spring, MD 20910; email: Timothy.J.Hall@aero.org

operational concept. Our findings suggest that forecast centers with the responsibility to forecast sky condition should consider integrating predictions from one or more advanced, obs-based techniques into their production process.

2. FORECAST SYSTEM CONCEPT

The forecast system concept is divided into multiple components. First, global (i.e., merged) METSAT composites and cloud/no cloud (CNC) maps are created and archived, along with corresponding data assimilation (DA)-based, gridded analyses. Second, these data are used to build a feature database for the point of interest (i.e., the local collection target). Third, this feature database is used to train an empirical prediction algorithm. Fourth, the prediction algorithm is initialized with features extracted or derived from current satellite imagery and gridded, numerical model analysis or forecast fields. Finally, the prediction algorithm is run to generate a very short-range, probabilistic forecast.

A key aspect of this concept is the notion that the weather patterns relevant to the sky condition at any point are unique to that location. While it may be reasonable (to a first-order approximation) to consider the weather patterns at two adjacent points (or small areas) to be the same, it is not necessary to do so given the temporal and spatial resolution of available METSAT data. Therefore, we propose a system which trains, in near real-time, a prediction model(s) for any local target of interest.

a. Hypothetical decision maker scenario

To provide context for our concept, a hypothetical scenario involving a military decision maker is now presented. In the scenario, a military commander manages a unique (i.e., high-demand, low availability), high-flying, high-cost unmanned aerial system (UAS) that provides surveillance capabilities. The UAS's sensors are ineffective in the presence of a cloud ceiling between the platform and any intended surveillance target.

Due to the nature of the mission and the operational environment, the commander is faced with frequent decisions to send the UAS (or not) to surveil some *local* target of interest (i.e., a small area $\leq 16 \text{ km}^2$) in a data-denied region. To support the tasking process, the

forecast system provides the probability of a clear sky condition over the target at 1-hr increments for a window of 1-5 hr from the time a decision is made. The commander is assumed to act or not in order to maximize the net benefit (i.e., to minimize cost and operational risk), on average, over many decisions. Support from a human forecaster is not available to assist in this scenario.

b. Satellite imagery and gridded analyses

The first component of the forecast system concept involves two databases including a METSAT archive and an archive of corresponding, DA-based analyses. For a global application, these data would probably amount to dozens to hundreds of terabytes. Just a few years ago this would have been a daunting prospect; however, recent advances in data storage and management capabilities have now made this a tractable problem.

Our research suggests that a CHANCES-class (Reinke et al. 2003, Vonder Haar et al. 1995) METSAT database would provide the foundation for the conceived forecast system. The Air Force sponsored CHANCES project pioneered the creation of a global, research-quality METSAT database at 1-hr temporal and 5-km horizontal spatial resolution (Fig. 1). Each CHANCES "image" was created by stitching together all available low-earth orbiting (LEO) and geostationary (GEO) weather satellite imagery into a global mosaic. Challenges in creating these mosaics included registration and navigation issues (to ensure geographic alignment of pixels from image to image), and continuity at the seams where the imagery from multiple satellites had to be joined together. CHANCES also included a binary CNC map for every hour (Fig. 1). A significant challenge resolved by the CHANCES team to create the CNC maps was the cross-calibration of similar information (i.e., longwave infrared radiance) collected by different satellites.

A barrier to development of a real-time application using CHANCES has been the storage media. During the time of the CHANCES project, the storage solution (given budget constraints and project objectives) was a high-capacity digital tape archive with file compression. In this format, data retrieval and subsequent processing was a time-consuming, labor intensive process. To support the present investigation, several terabytes of data were stored on portable hard drives connected to a

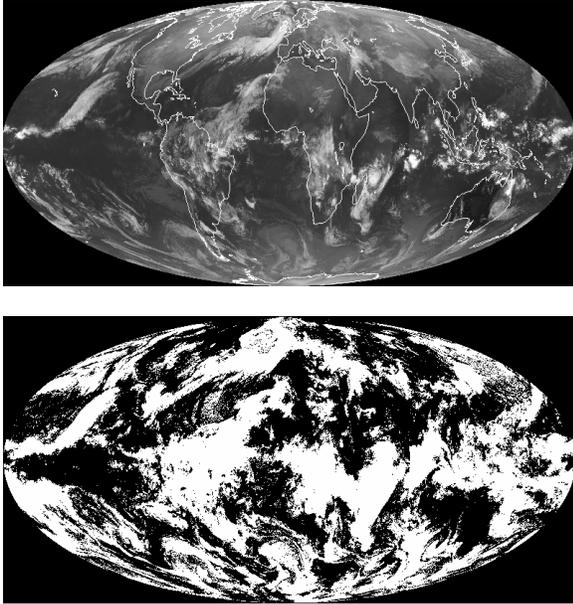


Figure 1: Example of a global CHANCES infrared radiance image mosaic (top) and the corresponding CNC image (bottom) from Reinke et al. (2003). Cloud is white and no cloud is black in the CNC image. (Reprinted with permission of John Forsythe, CIRA/CSU)

PC that were accessed by the research team over a local area network. This type of architecture, on a slightly larger scale, would enable near real-time forecasting for any, world-wide local target.

To complement the METSAT data, our concept includes an archive of DA-based, gridded meteorological analyses. From these data, many useful variables can be extracted or derived, and combined with cloud structural features from the METSAT data for a more complete characterization of the atmosphere. There are numerous, state-of-the-art global or regional-scale numerical modeling systems that can provide these data. Applying temporal and spatial interpolation, these analyses can be treated as observations for training of data-driven prediction algorithms.

c. Training of forecast algorithm(s)

The second component of this concept involves training a forecast algorithm. Training requires a feature database. In our concept, the feature database for a local target is built by extracting cloud structural features from a METSAT archive, and meteorological parameters from DA-based analyses. This

process results in a multi-year time-series of feature vectors with each vector representing the atmospheric state at a specific point in time.

Once the feature database is built, a forecast algorithm is trained. There are numerous, obs-based prediction algorithms which are viable for implementation in this concept ranging from machine learning methods to statistical techniques. Approaches tested in this investigation (Hall et al. 2010a) included artificial neural networks (ANN) (Zhang et al. 1999), Random Forests (Breiman 2001), decision tree induction (Quinlan 1986), k-nearest neighbor (k-nn) (Fukunaga 1990), Bayes classification (Fukunaga 1990), satellite cloud climatology (SCC) (Reinke et al. 1992), and the climatological-expectancy-of-persistence (CEP) (Enger et al. 1962, Kelly 1988, Hall et al. 1998, Reinke et al. 2003). Note that the notion of training does not strictly apply to the k-nn, SCC, and CEP methods.

d. Generating very short-range forecasts

The third component of the concept starts with characterization of the “current” state of the atmosphere in the region of interest, and concludes with production of a probabilistic forecast using one of the forecast methods. This characterization is done by extracting features from very recent obs to build a feature vector representing the “current” state of the atmosphere. The data used include current satellite imagery, and the most current DA-based analysis. In lieu of an analysis, the most recent, gridded model forecast (i.e., a 3-hr or 6-hr forecast) may be used as a surrogate source of meteorological parameters.

Finally, the “current” feature vector is plugged into the forecast algorithm to generate a very short-range probabilistic forecast. Depending on the method, not all of the features may be used. For instance, the CEP technique (as implemented in this investigation) only required the initial sky condition at the local target point along with the time of day and day of year. An ANN model, on the other hand, may make effective use of a hundred or more features.

3. PROTOTYPE SYSTEM CONCEPT PREDICTIONS USING A RANDOM FOREST ALGORITHM

For this investigation, numerous prediction algorithms demonstrated the viability of this

forecast system concept (Hall et al. 2010a, 2010b). In this section, the development of the feature database used for algorithm training and testing is discussed. Finally, results from implementation of a predictive learning algorithm called Random Forest are presented for Ft Hood which was one of six local targets used in this research (Hall et al. 2010a).

a. Data

The research database for this project, which spans from 1 May 2003 to 29 June 2008, consists of features (Table 1) extracted from satellite imagery and meteorological parameters derived or extracted from analysis fields generated by the NCEP's Eta model data assimilation system (EDAS) (Black 1994). The Eta analyses used in this investigation were extracted from a North American sector archived at 3-hr temporal resolution, 40-km horizontal spatial resolution, and 25 vertical levels. These data are maintained for research in the NCAR Computational & Information Systems Laboratory (CISL) Research Data Archive (RDA).

Cloud structural features were extracted from digital weather satellite imagery collected by NOAA Geostationary Operational Environmental Satellites (GOES-10 and GOES-12). Half-hourly GOES-12 imagery from the National Climate Data Center (NCDC) archive comprised the primary source of satellite observations. GOES-10 imagery was used to fill a 13-d GOES-12 data gap during Dec 2007.

Processing these data involved multiple steps. First, we read and saved the raw GOES variable (GVAR) visible and IR (10.7- μm and 3.9- μm channels) images for each half hour at their native resolution. Next, these data were registered, quality controlled, and projected to a common, equal-area projection with a uniform 4-km pixel spacing.

Once the GOES data were processed, we applied a cloud mask based on the bispectral composite threshold (BCT) technique presented by Jedlovec et al. 2008. The BCT algorithm uses multi-day composites of the difference in the 10.7- and 3.9- μm channel brightness temperatures ($BT_{10.7}$ and $BT_{3.9}$) to represent spatially and temporally varying cloud-free sky condition thresholds for cloud tests. An attractive feature of the BCT method is that it provides relatively consistent day-and-night cloud detection. The result of applying the BCT algorithm in this investigation was a five-year,

half-hourly time series of cloud-no cloud (CNC) image composites, each of which represented a map of the clouds at a specific observation time. These satellite data served as a regional, CHANCES-class database.

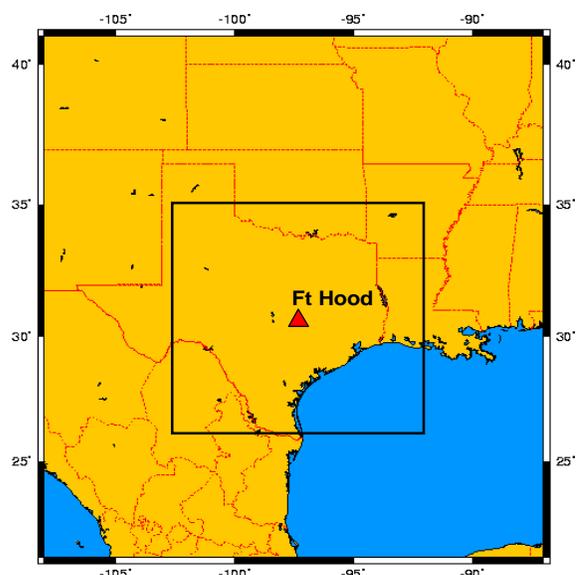


Figure 2: Map showing the Ft Hood local target (indicated by the triangle) at 31.14°N 97.72°W. The black rectangle surrounding Ft Hood represents the outer extent of the 1000 x 1000-km region in which METSAT data were used as a source of features. (Reprinted courtesy of NOAA)

b. Feature database development

A 5-yr database comprised of 105 features (Table 1) was built for the Ft Hood region (Fig. 2). The features were a mix of real and categorical variables extracted or derived from the EDAS model analyses, satellite-based CNC composites and astronomical calculations (e.g., solar zenith angle).

As mentioned above, an objective of this research was to develop and demonstrate an approach with global applicability. Due to the potential low quality of data assimilation-based moisture analyses outside of heavily observed regions such as the Continental United States (CONUS) and Europe (Andersson et al. 2004), the meteorological variables extracted from the Eta analyses were restricted to potential temperature (θ), pressure, wind and geopotential height, and parameters derived from them such as the mean layer vector wind (MLVW) (Blanchard and Lopez 1985) and dry static stability ($\Delta\theta/\Delta z$).

Table 1: List and description of features selected as predictors. Features refer to the sky condition of the pixel at Ft Hood unless specified other wise. N, W, E, and S refer to points 100-km north, west, east and south of Ft Hood, respectively. Ft Hood is referenced as the target (T) in the table.

		TABLE 1. (Continued)	
1	Ordinal date (integer from 1-366 used to define day of yr)	79	% area cloudy in upwind 400-km cone with vertex at T and 10.7 μ m IR BT < 233.15K
2	Time of day (GMT)		
3	Sunrise (time of sunrise to nearest 15 min.)	80	% area cloudy in upwind 500-km cone with vertex at T and 10.7 μ m IR BT < 233.15K
4	Sunset (time of sunset to nearest 15 min.)		
5	Maximum Solar Elevation	81	% area cloudy in upwind 100-km cone with vertex at T and 10.7 μ m IR BT \geq 233.15K
6	Solar Zenith Angle		
7	500-mb geopotential hgt.	82	% area cloudy in upwind 200-km cone with vertex at T and 10.7 μ m IR BT \geq 233.15K
8	700-mb geopotential hgt.		
9	500-mb geopotential hgt. anomaly for time of yr (\pm 15 d)	83	% area cloudy in upwind 300-km cone with vertex at T and 10.7 μ m IR BT \geq 233.15K
10	700-mb geopotential hgt. anomaly for time of yr (\pm 15 d)		
11	500-mb geopotential hgt. anomaly 6-hr change	84	% area cloudy in upwind 400-km cone with vertex at T and 10.7 μ m IR BT \geq 233.15K
12	700-mb geopotential hgt. anomaly 6-hr change		
13	500-mb geopotential hgt. anomaly and 6-hr change regime	85	% area cloudy in upwind 500-km cone with vertex at T and 10.7 μ m IR BT \geq 233.15K
14	700-mb geopotential hgt. anomaly and 6-hr change regime		
15-19	700-mb pressure vertical velocity at T, N, S, E, W	86	% area with identical sky condition (clear or cloudy) 24-hr prior in 52 x 52-km region
20-24	Best (4-layer) lifted index at T, N, S, E, W		
25-29	1000 to 700-mb MLVW speed at T, N, S, E, W	87	6-hr change in % area cloudy in 500x500-km region with 10.7 μ m IR BT < 233.15K
30-34	1000 to 700-mb MLVW direction at T, N, S, E, W		
35	300-mb maximum wind speed in a 1000x1000-km region	88	Avg. half-hourly rate of change in 10.7 μ m IR BT in 52x52-km region over 3-hr
36-40	300-150-mb MLVW speed at T, N, E, S, W		
41-45	300-150-mb MLVW direction at T, N, E, S, W	89	# hr out of previous 8-hr >80% clear in 52x52-km region
46	MS LP difference from N to S	90	# hr out of previous 6-hr >60% cloudy in 52x52-km region
47	MS LP difference from W to E	91	# hr out of previous 3-hr >60% cloudy in 52x52-km region
48	MS LP difference from T to: E, W, N, and S	92	# hr sky condition at T clear out of previous 6-hr
49	MS LP 3-hr change	93	# hr sky condition at T cloudy out of previous 6-hr
50	d(θ)/d(z) from 1000 to 700-mb	94	# hr sky condition at T clear out of previous 3-hr
51	d(θ)/d(z) from 850 to 500-mb	95	# hr sky condition at T cloudy out of previous 3-hr
52	d(θ)/d(z) from 900 to 850-mb	96	Avg. half-hourly rate of change in % area clear in 100x100-km region over 3-hr
53	1000-500-mb thickness difference from N to S		
54	1000-500-mb thickness difference from W to E	97	Avg. half-hourly rate of change in % area clear in 500x500-km region over preceding 3-hr
55	1000-500-mb thickness 6-hr change		
56	Sky condition (cloudy or clear)	98	Avg. half-hourly rate of change in % area clear in NW quadrant of 500x500-km region over preceding 3-hr
57	% area cloudy in 1000x1000-km region		
58	% area cloudy in 500x500-km region	99	Avg. half-hourly rate of change in % area clear in SW quadrant of 500x500-km region over preceding 3-hr
59	% area cloudy pixels in 100x100-km region		
60	% area cloudy in 52x52-km region	100	Avg. half-hourly rate of change in % area clear in SE quadrant of 500x500-km region over preceding 3-hr
61	% area cloudy in 12x12-km region		
62	% area cloudy in N sextant of 500x500-km region	101	Avg. half-hourly rate of change in % area clear in NE quadrant of 500x500-km region over preceding 3-hr
63	% area cloudy in NE sextant of 500x500-km region		
64	% area cloudy in E sextant of 500x500-km region	102	Mean of distribution of 10.7 μ m IR BT
65	% area cloudy in SE sextant of 500x500-km region	103	Variance of distribution of 10.7 μ m IR BT
66	% area cloudy in S sextant of 500x500-km region	104	Skew of distribution of 10.7 μ m IR BT
67	% area cloudy in SW sextant of 500x500-km region	105	Kurtosis of distribution of 10.7 μ m IR BT
68	% area cloudy in W sextant of 500x500-km region		
69	% area cloudy in NW sextant of 500x500-km region		
70	% area cloudy in 52x52-km region when \geq 10% of 500x500-km region cloudy		
71	% area cloudy in 500x500-km region with 10.7 μ m IR BT < 233.15K		
72	% area cloudy N quadrant of 300x300-km region with 10.7 μ m IR BT < 233.15K		
73	% area cloudy E quadrant of 300x300-km region with 10.7 μ m IR BT < 233.15K		
74	% area cloudy S quadrant of 300x300-km region with 10.7 μ m IR BT < 233.15K		
75	% area cloudy W quadrant of 300x300-km region with 10.7 μ m IR BT < 233.15K		
76	% area cloudy in upwind 100-km cone with vertex at T and 10.7 μ m IR BT < 233.15K		
77	% area cloudy in upwind 200-km cone with vertex at T and 10.7 μ m IR BT < 233.15K		
78	% area cloudy in upwind 300-km cone with vertex at T and 10.7 μ m IR BT < 233.15K		

of the target. Since the EDAS analyses are three-hourly, these data were interpolated temporally in order to populate the database at the half-hourly frequency of the CNC composites. Fifty cloud structural features were extracted from the CNC maps and BT_{10.7} imagery. The cloud features (Table 1) fall into five categories:

(1) Static sky condition features (56-75) that represent the percent coverage of cloudy or clear pixels at the current observation time in some region near or around the target.

(2) Static sky condition features stratified by MLVW (76-85).

(3) Dynamic sky condition features (87, 88, 96-101) created by analyzing the change (or trend) in percent area coverage of cloudy or clear conditions over an interval of time (e.g., 6 hr).

(4) Sky condition persistence features (86, 89-95) that capture the persistency of a particular sky condition over an interval of time.

(5) IR image statistical features (102-105) derived from the distribution of BT_{10.7} values in each 10.7- μ m image including mean, variance, skew, and kurtosis.

For Random Forest feature selection, and testing, these data were divided into two subsets. The first three years were designated as training data. The last two years of data were reserved for testing. All performance metrics, discussed below, are based on validation against the 2-yr test dataset.

c. Forecast performance benchmarks

Three benchmark forecast methods were computed as detailed below based on information in the feature databases.

1) Basic persistence

BP of clear for any given forecast interval was taken to be a 0 or 100% probability of clear at forecast time based on the initial sky condition. This translates to a 100% forecast probability of clear if the initial sky condition was clear in the CNC composite and 0% forecast probability of clear if the initial sky condition was cloudy.

2) Climatological-expectancy-of-persistence

CEP was derived using the 3-yr training dataset and was calculated as the probability

(i.e., the frequency of occurrence) of clear at each forecast interval (1, 2, 3, 4, and 5 hr) given the current (i.e., initial) sky condition (i.e., cloudy or clear) based on training data events within ± 1 hr and ± 30 d of that time of day and day of year. There was no differentiation between true persistence and recurrence in these calculations.

3) Satellite Cloud Climatology (SCC)

The SCC forecasts were based on the unconditional, prior probability of cloud-free conditions at the target calculated using the 3-yr training dataset for given (time of day, day of year) combinations. For all observations within ± 1 hr and ± 30 d, the percentage of occurrences with clear conditions were used as the a priori (i.e., climatological) sky condition probability for that time of day and time of year.

d. Random Forest forecast algorithm

The Random Forest method was developed by Leo Breiman (Breiman 2001) to improve the performance of his decision tree algorithm (Breiman et al. 1998). Random Forest creates an ensemble of decision trees by training on a random redistribution of the training set. Each distribution is generated by randomly drawing N samples (with replacement), where N is the size of the training set. A tree is grown on a fixed-size subset of features randomly drawn on each round. The algorithm outputs the class that is the mode of the output by the individual trees. In this investigation, the output was interpreted as a probability. Various analytical techniques were used to prune the feature space to 78 features (Hall et al. 2010a) for the final implementation.

e. Proof-of-concept, Random Forest performance

No single measure of performance can completely and unambiguously describe the quality of a forecast system. Therefore, our approach to assess the RF, CEP, SCC and BP forecasts was multifaceted. Overall performance potential was assessed using relative operating characteristic (ROC) analysis and its associated ROC score (Harvey et al. 1992, Mason and Graham 2002). Additional insights were gleaned from an ensemble of metrics including sharpness, accuracy, expected best cost (EBC), and reliability. Hall et al. (2010b) contains a detailed description of these metrics.

Following Mason (1982) and Fawcett (2006), ROC curves are plotted on a two-dimensional graph with hit rate (i.e., positives correctly forecasted divided by total positives) on the y-axis and false alarm rate (i.e., negatives incorrectly forecasted divided by total negatives) on the x-axis. Given a set of probabilistic forecasts, the information needed to create a set of hit rate, false alarm rate pairs (to create the ROC curve) can be generated from categorical forecasts by stepping a decision probability p_c through a range of values (i.e., decision thresholds). The ROC score is the area under the ROC curve and is considered a single-number-summary of forecast algorithm performance (Mason and Graham 2002). BP has one hit rate, false alarm rate pair associated with it and, hence, results in only a single point in ROC space.

ROC curves for Random Forest, CEP, and BP are shown in Fig. 3. Perfect performance (i.e., a ROC score = 1.0) is represented on a ROC graph by the upper left-hand corner. A ROC curve oriented along the major diagonal line from the origin (0,0) to the point (1,1) represents random forecasting (i.e., a ROC score of 0.5). In Fig. 3, it is clear that the performance potential of the Random Forest forecasts exceeds the other methods for all possible decision thresholds for the 1-, 3-, and 5-hr forecast intervals. The same holds for the 2- and 4-hr intervals (not shown). Based on the ROC curves, the performance potential for SCC exceeded random forecasting but was well under all other methods. For the 5-hr forecast, the ROC scores of Random Forest, CEP, and SCC were 0.8532, 0.7714, and 0.6073, respectively.

Sharpness (Fig. 4) indicates the tendency of a probabilistic forecast method to correctly assign extreme probability values (i.e., the tendency toward correct categorical forecasts). Forecast performance (in terms of *sharpness*) is dependent on the amount of separation between the probability values output by the algorithm when the true class is clear and when the true class is cloudy. Therefore, a histogram of the distribution of probabilities plotted against the data of how each instance turned out was used to graphically assess *sharpness*. The sharper the performance, the more the graph assumes a shape of a bathtub's cross-section. The probability distributions shown in Fig. 4

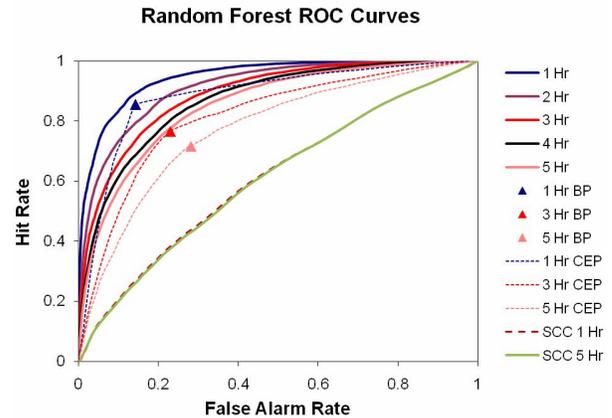


Figure 3: ROC graph for Random Forest, CEP, SCC and BP for the Ft Hood, local target.

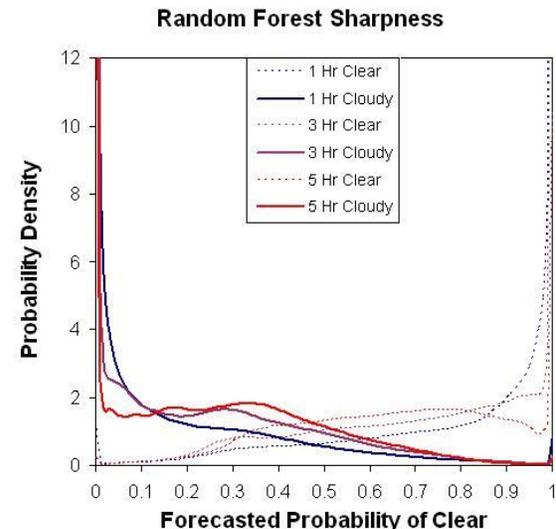


Figure 4: Histogram of the distribution of probabilities of how each instance turned out that graphically depicts the *sharpness* of the Random Forest algorithm.

reveal that the Random Forest forecasts for Ft Hood were very sharp. For comparisons to other algorithms refer to Hall et al. (2010b).

Accuracy for this investigation (Fig. 5) was taken as the percent correct match (PCM) which was defined as the percent of all forecasts that turned out to be correct (i.e., either a hit or a correct negative). Determination of PCM requires choosing a probabilistic decision threshold at which the forecast is made. Given a forecast sky condition probability provided by the Random Forest algorithm in this investigation, a threshold of 0.5 was used to transform each probabilistic forecast into a categorical forecast of cloudy or clear. This threshold minimizes the

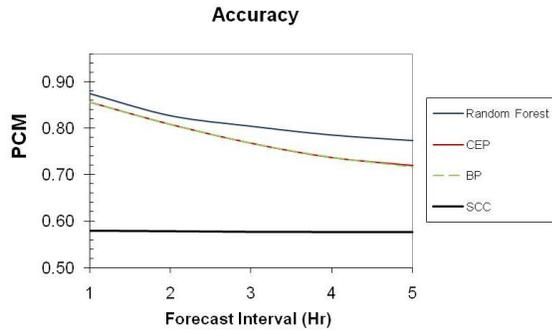


Figure 5: Accuracy of Random Forest, CEP, BP, and SCC forecast methods for forecast intervals of 1-5 hr for Ft Hood local target.

probability of error, $P(\text{error})$ which is equal to $1 - P(\text{correct forecast})$. As shown in Fig. 5, the accuracy of the Random Forest algorithm decreases with increasing forecast interval, but clearly exceeds the accuracy of the benchmark methods including CEP, BP, and SCC. CEP and BP are competitive, however, at the 1- and 2-hr forecast intervals.

EBC is a value metric that was used in this investigation to assess performance in the context of the decision maker's average, net benefit from the obs-based forecasts. The basic premise of the cost-loss problem is that a decision maker is faced with the uncertain prospect of a weather event (E). As discussed by Murphy and Ehrendorfer (1987), the prototype cost-loss scenario is a problem involving two possible decisions to act or not and the two weather events (i.e., clear or cloudy). $E = 0$ corresponds to clear conditions and $E = 1$ corresponds to cloudy conditions. Let $f = 0$ represent a categorical forecast of clear sky condition and $f = 1$ a cloudy sky condition forecast. The decision maker incurs a cost c (> 0) if action is taken and ($E = 1$), a cost equivalent to $(c - v)$ if action is taken and $E = 0$, and a cost equivalent to $(v - c)$ if no action is taken and ($E = 0$). Here, v (≥ 0) is the additional value of taking the action when ($E = 0$) (not including the cost of the action). Note that if ($v > c$), the cost would turn out to be negative, meaning a "profit" is realized. For this investigation, this problem was considered in terms of an expected best cost expressed in terms of the value-cost ratio ($\alpha = v/c$). If ($E = 1$), then ($v = 0$). The decision maker is assumed to take action or not in order to maximize "profit" (i.e., minimize cost) such that $\alpha > 1$.

The use of a specific probabilistic threshold to transform the probabilistic forecast output by

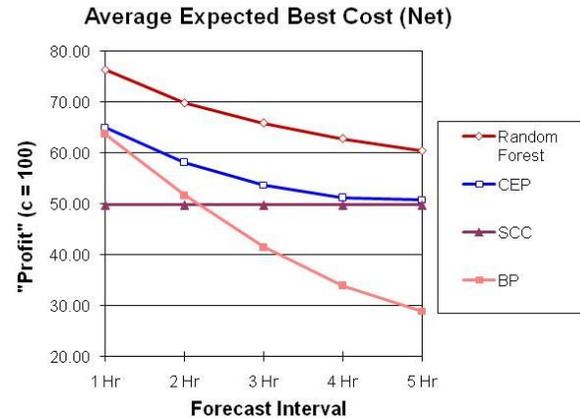


Figure 6: EBC for with $\alpha = 3$ for Random Forest, CEP, SCC, and BP forecast methods across all forecast intervals assuming a cost (c) = 100 monetary units.

the Random Forest algorithm into a categorical forecast will generate the following probabilities: $P_{00} = P(f = 0 | E = 0)$, $P_{10} = P(f = 1 | E = 0)$, $P_{01} = P(f = 0 | E = 1)$, $P_{11} = P(f = 1 | E = 1)$. The test data can also be used to generate the a priori probabilities $P_0 = P(E = 0)$, and $P_1 = P(E = 1)$. It can be shown that the *EBC*, or max net "profit" per action taken, on average, is equivalent to:

$$EBC = c \{ [(1 - \alpha) + 2(\alpha - 1)P_{10}]P_0 + P_{01}P_1 \}$$

Assuming a cost of 100 monetary units, the variation in *EBC* (converted to "profit" by plotting $-EBC$) for Random Forest, CEP, SCC, and BP with increasing forecast interval, for a value-cost ratio $\alpha = 3$ is shown in Fig. 6. The average profit per action of the Random Forest forecasts exceeds CEP by about 15% across all forecast intervals. Random Forest forecasts are about 15% more profitable than BP for 1-hr. and over 100% more profitable for 5-hr forecasts.

Reliability is equivalent to bias and answers the question of how well the predicted probabilities of an event correspond to their observed frequencies. It complements ROC analysis and EBC. Reliability for the Random Forest algorithm at each forecast interval is shown in Fig. 7 by plotting the observed frequency of clear versus the forecast probability of clear (to the nearest tenth). Theoretical, perfect reliability is shown by the emboldened line that extends from the origin (0,0) to the point (1,1) in the figure. The Random Forest algorithm

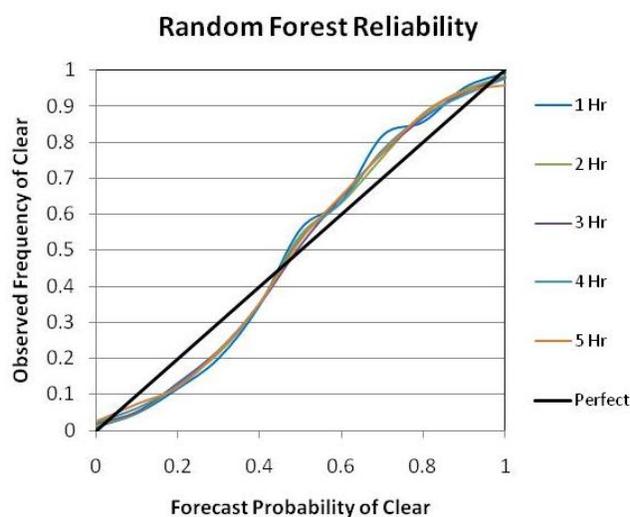


Figure 7: Reliability graph for Random Forest.

had slight tendencies to under-forecast when the observed frequency of clear is above 0.5, and over-forecast when the observed frequency is below 0.5. The average root mean square reliability error (RMSRE) over all forecast intervals was 0.056. This was slightly higher than the average RMSRE for CEP (0.0255) and SCC (0.0484).

4. CONCLUSION

In this paper, a new concept for production of obs-based sky condition forecasts was presented. Our research demonstrated an approach to automated, very short-range forecasting that could potentially be implemented globally. The key enablers are a multi-year, CHANCES-class METSAT database, and advanced prediction algorithms that can be trained to make probabilistic forecasts when presented with a “current” weather pattern as defined by a set of cloud structural features. Future implementation of a CHANCES-class database would rely on NPOESS imagery (along with data from other low-earth orbiting weather satellites) for information north (south) and of 60° N lat (60° S lat), along with geostationary weather imagery for the region spanning from the tropics through the mid-latitudes. Data assimilation-based analyses were found to improve the performance of the prediction algorithms when used in conjunction with the METSAT data. Based on our work, data storage and processing would appear not be significant constraints for an operational system.

During this investigation, several algorithms that we tested were trained on multiple years of data and produced forecasts in less than 10 min on standard desktop computers. This does not include the time for initial construction of the feature database. The Random Forest algorithm, discussed in this paper, was among the top performing methods. Other obs-based methods that we tested also performed well. The results of all algorithms across six target areas (representing different weather regimes) are presented in two companion conference papers (Hall et al. 2010a, 2010b).

Our findings indicate that the performance of obs-based sky condition techniques for very short-range forecasting may exceed the performance of state-of-the-art algorithms currently in use such as those based on advection (Kopp et al. 1997) and cloud-motion vectors (Banniehr et al. 1996, Hammer et al. 2003, Hamill and Nehr Korn 1993). Additionally, advanced obs-based forecasts would provide a useful baseline for on-going work to develop very short-range forecast techniques based on numerical modeling. The next step in demonstrating this concept would be to conduct trials to compare performance of the obs-based forecasts we have demonstrated with current, operational techniques. We encourage other research groups to take up this challenge.

5. ACKNOWLEDGEMENTS

This work was sponsored by the NPOESS program under contract number DG133E07CQ0005.

9. REFERENCES

- Andersson, E., C. Cardinali, M. Fisher, E. Holm, L. Isaksen, Y. Tremolet and A. Hollinsworth, 2004: Developments in ECMWF’s 4D-Var System. *16th Conference on Numerical Weather Prediction*, Seattle, WA, Jan 12-15, 2004.
- Bankert, R. L., and M. Hadjimichael, 2007: Data mining numerical model output for single-station cloud-ceiling forecast algorithms. *Wea. Forecasting*, **22**, 1123-1131.
- Banniehr, L., M. Rohn, and G. Warniecke, 1996: A functional analytic method to derive displacement vector fields from satellite image

- sequences. *Int. J. Remote Sensing*, **17**, 383-392.
- Breiman, L., 2001: Random forests. *Machine Learning*, **45**, 5-32.
- Breiman, L., J. H. Friedman, R. A. Olshen, and C. J. Stone, 1998: *Classification and regression trees*. Chapman & Hall, 368 pp.
- Black, T. L., 1994: The new NMC mesoscale Eta Model: Description and forecast examples. *Wea. Forecasting*, **9**, 265-278.
- Blanchard, D. O., and R. E. Lopez, 1985: Spatial patterns of convection in south Florida. *Mon. Wea. Rev.*, **113**, 1282-1299.
- Combs, C.L., W. Blier, W. Strach, M., DeMaria, 2004: Exploring the timing of fog formation and dissipation over San Francisco Bay area using satellite cloud composites. Preprints, *13th Conf. Satellite Meteorology and Oceanography*, Norfolk, VA, Amer. Meteor. Soc.
- Connell, B. H., K. Gould, and J. F. W. Purdom, 2001: High-resolution GOES-8 visible and infrared cloud frequency composites over northern Florida during the summers 1996-1999. *Wea. Forecasting*, **16**, 713-724.
- Enger, I., L. J. Reed, and J. E. MacMonegle, 1962: An evaluation of 2-7-hr aviation terminal-forecasting techniques. Interim Report Technical Publication 20 prepared for the Federal Aviation Agency, Systems Research and Development Service by the Travelers Research Center Inc., 43 pp.
- Fawcett, T., 2006: An introduction to ROC analysis. *Pattern Recognition Letters*, **27**, 861-874.
- Fukunaga, K., 1990: *Introduction to Statistical Pattern Recognition*. Academic Press, 592 pp.
- Girodo, M., R. W. Mueller, and D. Heinemann, 2006: Influence of three-dimensional cloud effects on satellite derived solar irradiance estimation—first approaches to improve the Heliosat methods. *Solar Energy*, **80**, 1145-1159.
- Hall, T. J., C. N. Mutchler, G. J. Bloy, R. N. Thessin, S. K. Gaffney, and J. J. Lareau, 2010a: Comparison of artificial intelligence and statistical techniques for probabilistic forecasting of sky condition. *20th Conf. on Probability and Statistics*, Atlanta, GA, Amer. Meteor. Soc.
- Hall, T. J., R. N. Thessin, G. J. Bloy, and C. N. Mutchler, 2010b: Comparison of artificial intelligence and statistical techniques for probabilistic forecasting of sky condition. *14th Conf. on Aviation, Range, and Aerospace Meteorology*, Atlanta, GA, Amer. Meteor. Soc.
- Hall, T. J., D. L. Reinke, and T. H. Vonder Haar, 1998: Forecasting applications of high-resolution satellite cloud composite climatologies. *Wea. Forecasting*, **13**, 16-23.
- Hammer, A., D. Heinemann, C. Hoyer, R. Kuhlemann, E. Lorenz, R. Muller, and H. G. Beyer, 2003: Solar energy assessment using remote sensing technologies, *Remote Sensing of Environment*, **86**, 423-432.
- Hamill, T. M., and T. Nehr Korn, 1993: A short-term cloud forecast scheme using cross correlations. *Wea. Forecasting*, **8**, 401-411.
- Hansen, B., 2007: A fuzzy logic-based analog forecasting system for ceiling and visibility. *Wea. Forecasting*, **22**, 1319-1330.
- Harvey, L. O. Jr., K. R. Hammond, C. M. Lusk, and E. F. Mross, 1992: The application of signal detection theory to weather forecasting behavior. *Mon. Wea. Rev.*, **120**, 863-883.
- Jedlovec, G. J., S. L. Haines, and F. J. LaFontaine, 2008: Spatial and temporal varying thresholds for cloud detection in GOES imagery. *IEEE Trans. Geosci. Rem. Sens.*, **46**, 1-13.
- Kelly, F. P., 1988: Spatial and temporal short range total cloud cover estimation by metric analysis of composite imagery, Ph.D. dissertation, Department of Atmospheric Science, Colorado State University, Fort Collins, CO, 152 pp.
- Mason, S. J., and N. E. Graham, 2002: Areas beneath the relative operating characteristic (ROC) and relative operating levels (ROL) curves: Statistical significance and interpretation. *Q. J. R. Meteorol. Soc.*, **128**, 2145-2166.
- Kopp, T. J., M. M. Wonsick, L. E. Cantrell, and F. D. Bieker, 1997: ADVCLD: Air Force Global Weather Center's updated cloud forecast model.

Preprints, *Cloud Impacts on DOD Operations and Systems 1997 Conference*, Newport, RI, 107-110. PL-TR-97-2112, Phillips Laboratory (AFMC), Hanscom AFB, MA, 278 pp.

Mason, I., 1982: A model for assessment of weather forecasts. *Aus. Met. Mag.*, **30**, 291-303.

Murphy, A. H., and M. Ehrendorfer, 1987: On the relationship between the accuracy and value of forecasts in the cost-loss ratio situation. *Wea. Forecasting*, **2**, 243-251.

Norquist, D. C., 1999: Cloud predictions diagnosed from global weather model forecasts. *Mon. Wea. Rev.*, **128**, 3528-3555.

Reinke, D. L., C. L. Combs, S. Q. Kidder and T. H. Vonder Haar, 1992: Satellite cloud composite climatologies: A new high-resolution tool in atmospheric research and forecasting. *Bull. Amer. Meteor. Soc.*, **73**, 278-285.

Reinke, D. L., J. M. Forsythe, J. A. Kankiewicz, K. R. Dean, C. L. Combs, and T. H. Vonder Haar, 2003: Development and applications of regional cloud products from the CHANCES global cloud database. *12th Conf. on Satellite Meteorology and Oceanography*, Long Beach, CA, Amer. Meteor. Soc.

Vislocky, R. L., and J. M. Fritsch, 1997: An automated, observation-based system for short-term prediction of ceiling and visibility. *Wea. Forecasting*, **12**, 31-43.

Vonder Haar, T.H., D.L. Reinke, K.E. Eis, J.L. Behunek, C.R. Chaapel, C.L. Combs, J.M. Forsythe, and M.A. Ringerud, 1995: Climatological and Historical Analysis of Clouds for Environmental Simulations (CHANCES) Database. Final Report prepared for Phillips Laboratory under Contract No. F-19628-93-C-0197, July 1995.

Wiener, G., P. H. Herzegh, R. Bateman, and B. Chorbajian, 2004: Utilizing site-based data mining in national ceiling and visibility forecasting. *11th Conf. on Aviation, Range, and Aerospace Meteor.*, Hyannis, MA, Amer. Meteor. Soc.