MARINE LAYER

Marine Layer Forecasting Report

Green Power Labs Inc.

One Research Drive,
Dartmouth, Nova Scotia,
Canada, B2Y 4M9
Phone: (902) 466-6475
Fax: (902) 466-6889

2014-05-30

This report was prepared by Green Power Labs Inc. (GPLI) in part by a grant provided by the California Public Utilities Commission's (CPUC) California Solar Initiative (CSI) R&D program. The materials in this report reflect GPLI's best judgment based on the information available to the company at the time of report preparation. Any use of this report by a third party, or any reliance on or decisions made based on it, are the responsibility of such third parties. GPLI accepts no responsibility for damages, if any, suffered by any third party as a result of decisions made or actions based on this report.

Contents

Contents 3
1. Executive Summary 4
2. Introduction 4
3. Data Sources 6
   3.1. Geographic data 6
   3.2. Data sources for Live Classification 7
   3.3. Spatial Lookup table 9
   3.4. Historical Data for PR and PM approaches 9
   3.5. Soundings Data Source 9
   3.6. Data for LAPS and WRF 9
   3.7. Geographic masks used in forecast 10
4. Methods 12
   4.1. Marine Layer Classification Methods 12
   4.2. Inversion Detection Algorithm 14
   4.3. Local Analysis and Prediction System / Weather Research and Forecasting Model Marine Layer and Total Cloud Forecasting Methods 15
   4.4. Pattern Recognition Method 17
   4.5. Probabilistic Model 18
5. Results – ML forecast Performance 20
6. Conclusions 30

1. Executive Summary

A number of approaches were tested in developing Marine Layer (ML) forecasts for 36 hours ahead. In particular - Pattern Recognition (PR), Probabilistic Model (PM), and Local Analysis and Prediction System/Weather Research and Forecasting (LAPS/WRF) models were developed and tested, and an optimal combination of models was selected based on their performance vs. time horizon.

Because PM, PR and WRF performed with different accuracy depending on the geographic location, time of day, and forecast horizon a combination of Probabilistic and WRF approach were adopted to deliver the ML forecast. Currently the best mix of models is the use of probabilistic forecast for the present day (7am – 7 pm) and the use of WRF for hours beyond that. The PM approach captures attenuation of ML through the day really well but underperforms on the day ahead. In contrast, the current configuration WRF is capable of developing future cloud fields but is relatively weak in capturing them within the short time horizon. The results of our work are encouraging with overall agreement of the developed forecasts with observations of 60% to 80% (as in number of 1km cells in the service area where ML is forecasted correctly).

It is expected that LAPS/WRF, with some improvements, will become a source of both Marine Layer and solar irradiance forecasts for the SolarSatData® (GPLI's solar power prediction platform - SSD) program in the near future. To achieve this, real time ML classification results will have to be integrated into LAPS/WRF "on the fly" to improve the accuracy of short time horizon forecasts. Testing should be carried on integration of ML forecast with the existing SSD software as well.

It is important to carry on further development of forecasting techniques to improve accuracy by adjusting classification techniques, introducing real time forecast validation, and improving the PM, PR and LAPS/WRF models.

2. Introduction

GPLI has previously developed a successful methodology for using visible and infra-red (IR) channel satellite data to estimate the probability of ML in Southern California. This project is building upon the

previous success. The objective of this project was to develop and implement Marine Layer (ML) forecast for a 1x1 km grid for San Diego County.

San Diego County is a region heavily affected by ML events, having the highest extent and probability of ML occurrence. The influence of ML is the strongest in coastal areas up to 50 km away from coast. Accordingly, previously developed ML classification methodology could be used successfully in the spatial buffer along the coastline. The extent of ML inland is largest in May and recedes towards coast gradually through the summer and fall. We have previously established that an understanding of the climatological drivers for ML events is required to improve the forecasting of ML probabilities. We have addressed this need by implementing state of the art LAPS/WRF weather forecasting augmented by probabilistic modeling and artificial intelligence techniques. The detection of ML in this area is vital for the proper integration of solar resource and power forecasting. Due to the nature of the ML it is particularly difficult to forecast; publically available numerical weather forecast models, with their coarse spatial grids, routinely miss the ML. Marine layer has a similar effect to solar irradiance as clouds, namely that the available solar resource is diminished as a result, ultimately leading to a drop in available PV power.

This report is addressing work carried out by GPLI in 2013 to develop a forecasting system for ML for up to 36 hours ahead. The report addresses data sources and classification methodology and describes in detail the following work: Development and implementation of Pattern Recognition (PR), Probabilistic Model (PM), and LAPS/WRF models; testing different versions of these models; and finally, selection of optimal combinations of models based on their performance vs. time horizon.

3. Data Sources

This section describes sources of information used in classification, modeling and forecasting ML, (with hyperlinks to actual data where available).

3.1. Geographic data

Geographical information used in this project included a digital elevation model (DEM) for the area, as well as the vector data for coastline and SDGE service area boundary. The Digital Elevation Model was obtained from National Elevation Dataset, United States Geological Survey (USGS, Source: http://ned.usgs.gov/). The dataset comes in a resolution of 1/3 Arc Second (approximately 10x10m) and contains a total of 644 files with a total volume of 2.57 GB. Files corresponding to the area of interest were processed into 267 MB ArcGIS grids (ArcGRID) through following steps:

- Zipped map sheet files were downloaded and un-compressed on local drive.
- All map sheets were added to ArcGIS 9.3.1.

US Medium Resolution Coastline (NOAA) was obtained from:

http://shoreline.noaa.gov/data/index.html as shapefile at spatial scale: 1:70,000 with original coverage for Continental United States.

3.2. Data sources for Live Classification

The following data sources were used in ML forecasting:

a) McIDAS (Figure 1)
b) SSEC Data Flow (Figure 2)
c) NOAA FTP (Figure 3)

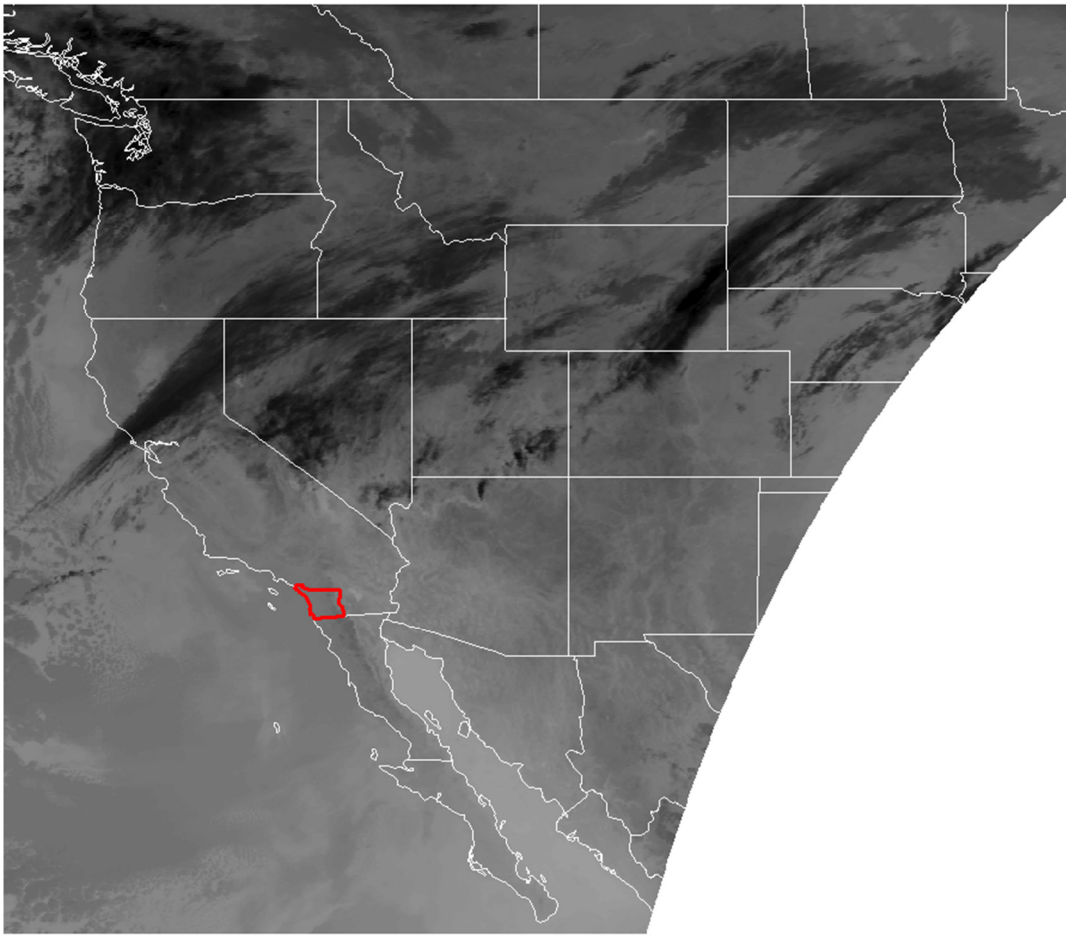


Figure 1. Satellite data obtained through McIDAS. SDGE service area is shown in red.

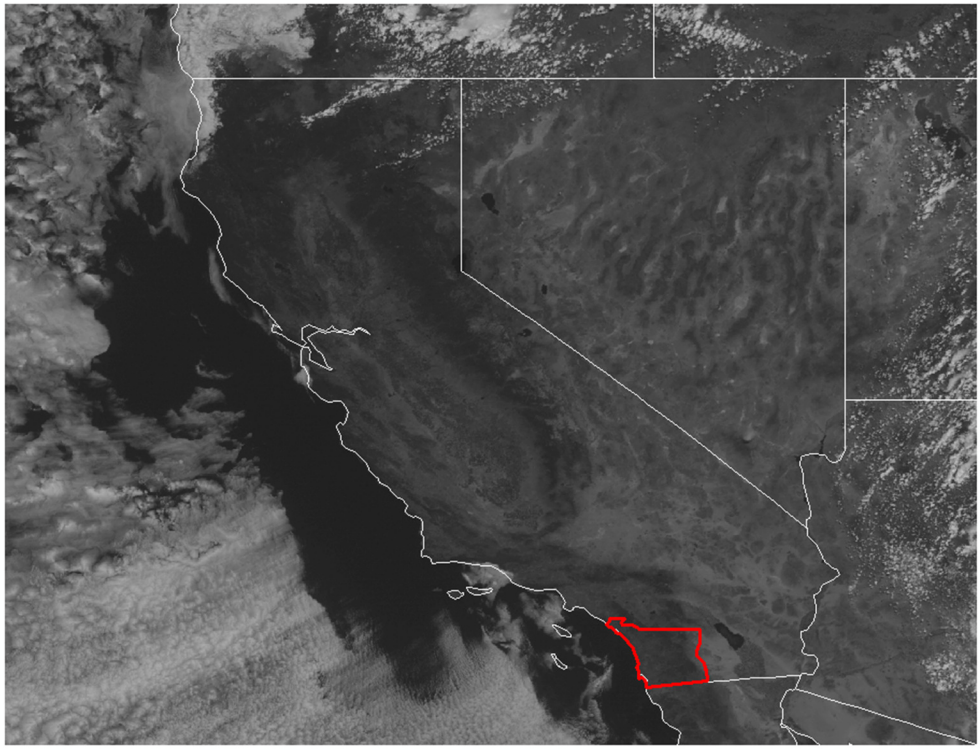


Figure 2. Satellite data obtained through SSEC.

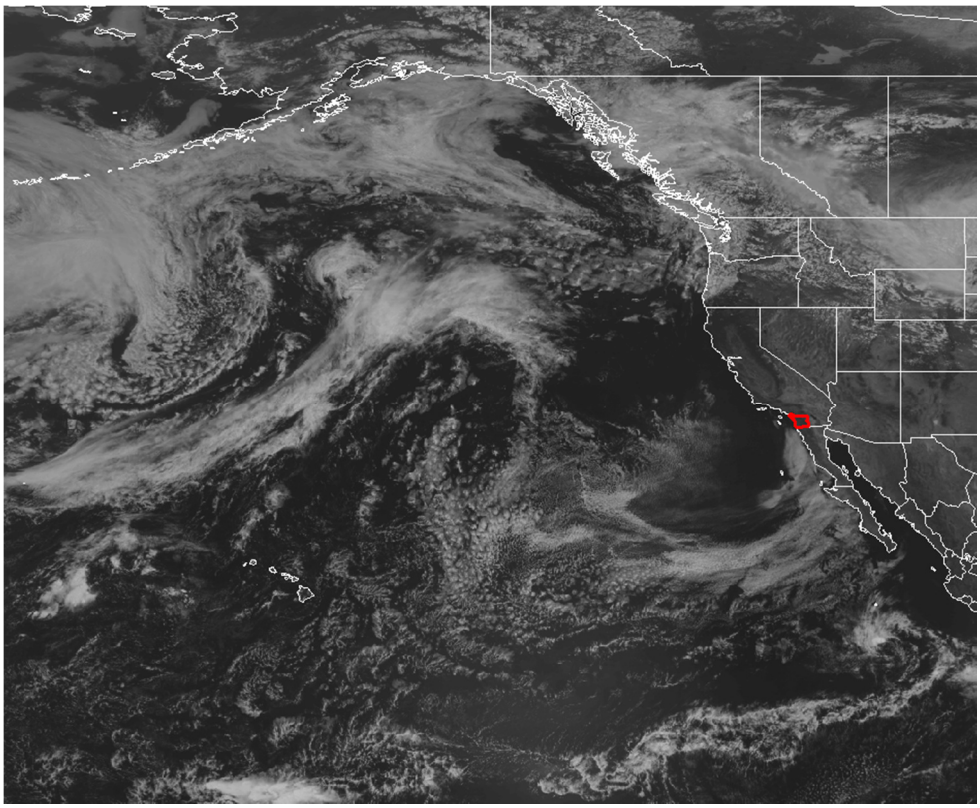


Figure 3. GOES West data coverage.

3.3. Spatial Lookup table

The forecast and analysis were carried on 1x1km grid encompassing San Diego County, see Figure 4 below.

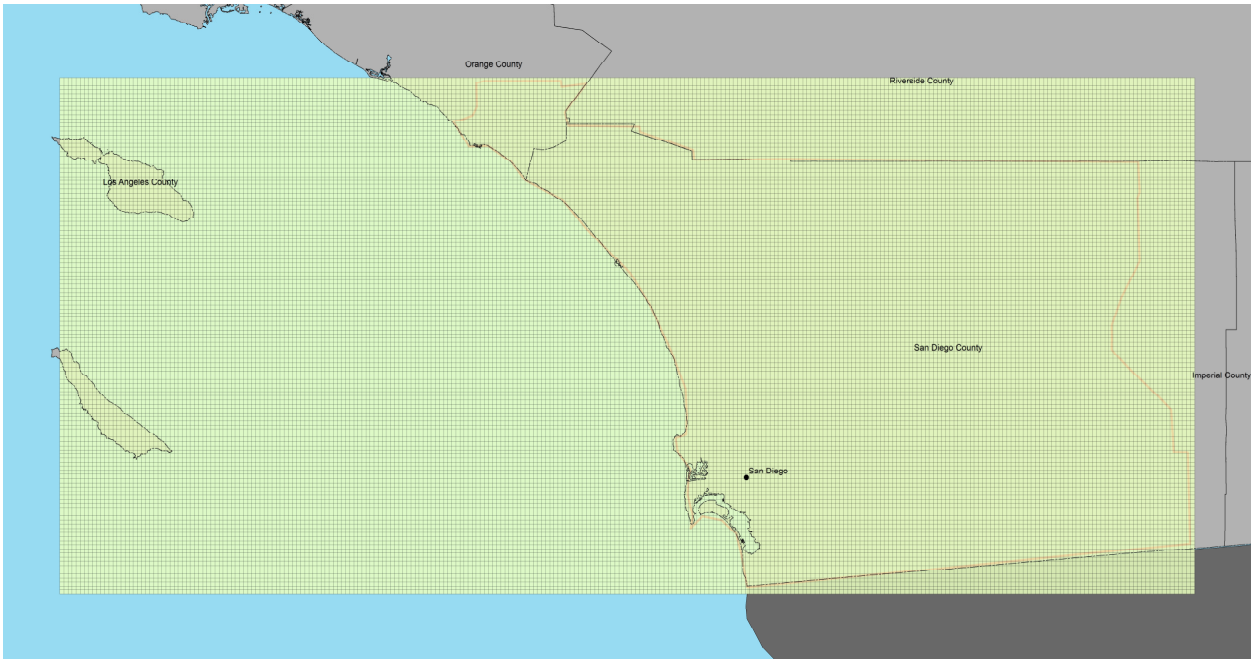


Figure 4. Forecast area with spatial lookup table overlaid. Forecast is provided for each 1x1km cell.

3.4. Historical Data for PR and PM approaches

The following data sources were used in the analysis of historical data used in the PR and PROB approaches.

a) 2000-2010 SSEC provided (historical processing)
b) 2011-2012- SSEC IR and NOAA VIS (historical processing)
c) The historical data is currently being processed with updated classification methods

3.5. Soundings Data Source

All sounding data can be downloaded via the University of Wyoming's atmospheric science website.

3.6. Data for LAPS and WRF

There are a variety of meteorological data that are ingested into the LAPS program to seed the WRF forecast model with the most up to date weather conditions.

For our purpose, we chose to use the North American Mesoscale (NAM) model which covers all of the continental US and has a forecast that extends up to 84 hours (Released every 6 hours).

Our observational data comes primarily from the Meteorological Assimilation Data Ingest System (MADIS). These data provide LAPS with the most recent observations which are then integrated into the model variables.

3.7. Geographic masks used in forecast

To avoid using misclassifications of ML in the arid inland areas we have used spatial masking to the results of classification which constrains the ML to the area where it is most likely. Initially the mask was drawn along the approximate 5% probability level for ML, and later it was substituted with a more accurate mask, see Figure 5 for details.

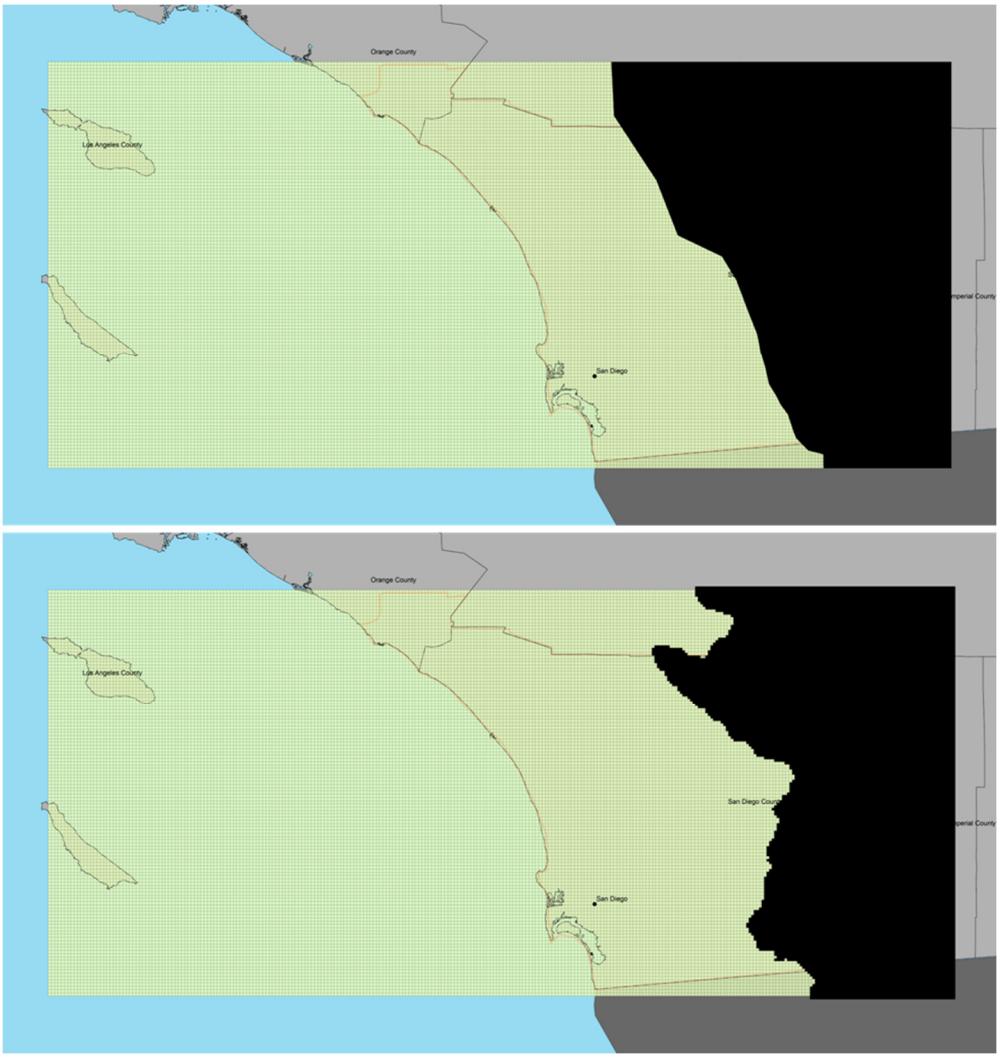


Figure 5. Spatial masks used in forecasting ML - original (top) and advanced (bottom).

4. Methods

This section includes a brief description of the methodology used to classify live ML presence, to establish inversion parameters, to model and forecast ML and total cloud cover using LAPS/WRF, and to forecast ML using Pattern Recognition (PR) and Probabilistic Model (PM) approaches.

4.1. Marine Layer Classification Methods

An automatic ML classification algorithm was implemented for near-real-time ML identification at every square kilometre of the total forecast area. The reclassification process of IR and VIS satellite images were used to detect marine layer along with high/low clouds. Figures 6-8 below outline the process. In Figure 6 we identify all cloud pixels in the image, as well as bare ground pixels. Figure 7 we cross Figure 6 with our ML mask from Figure 5 (above) to mask out clouds not contained within the ML boundary. Lastly, Figure 8, shows the final images used in the ML forecast.

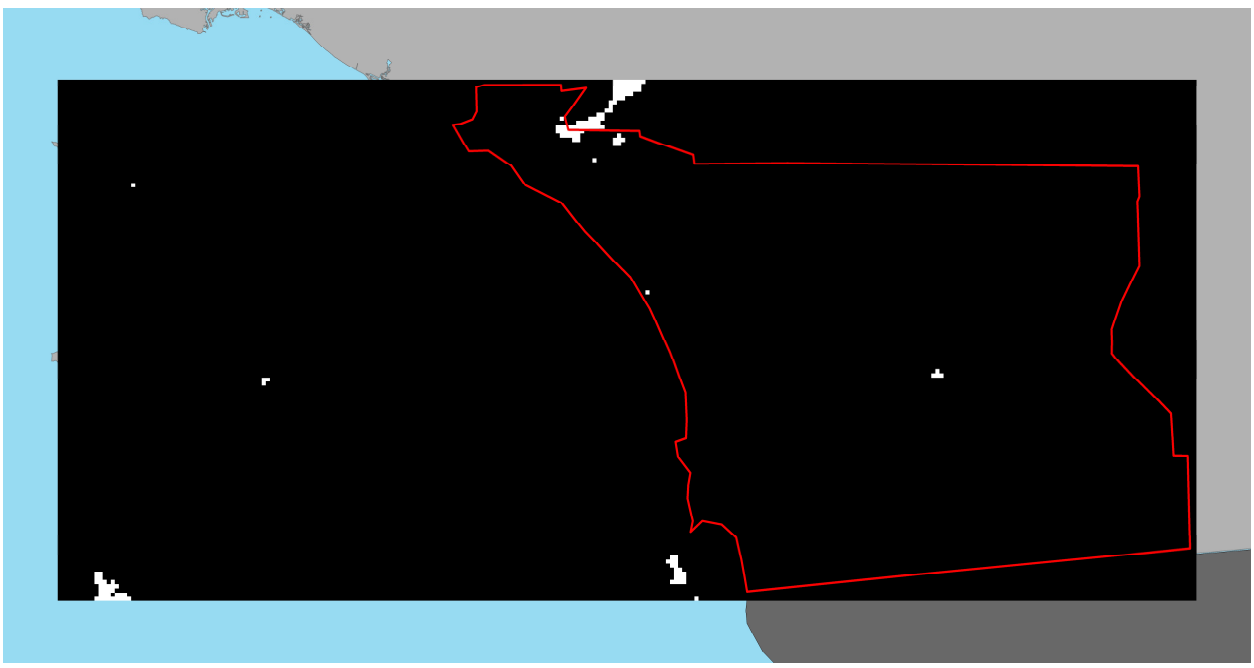


Figure 6 Classification of 2 classes (Cloud or bare ground)

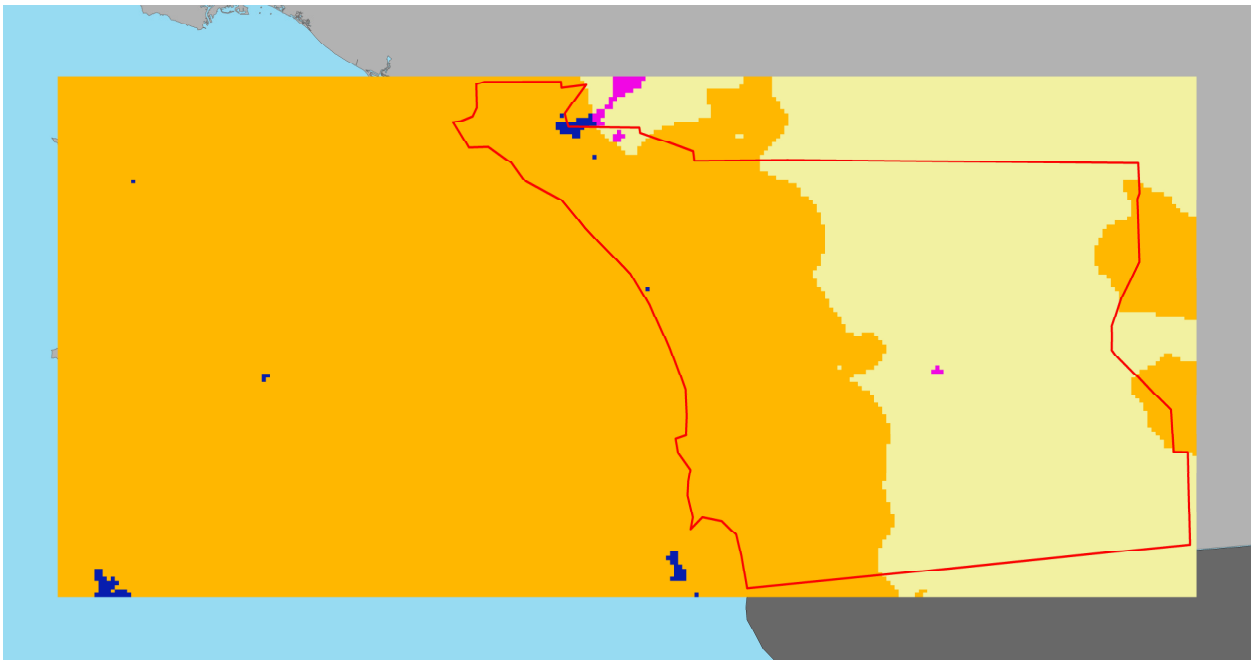


Figure 7 Cross classification with 15% presence of yes/no

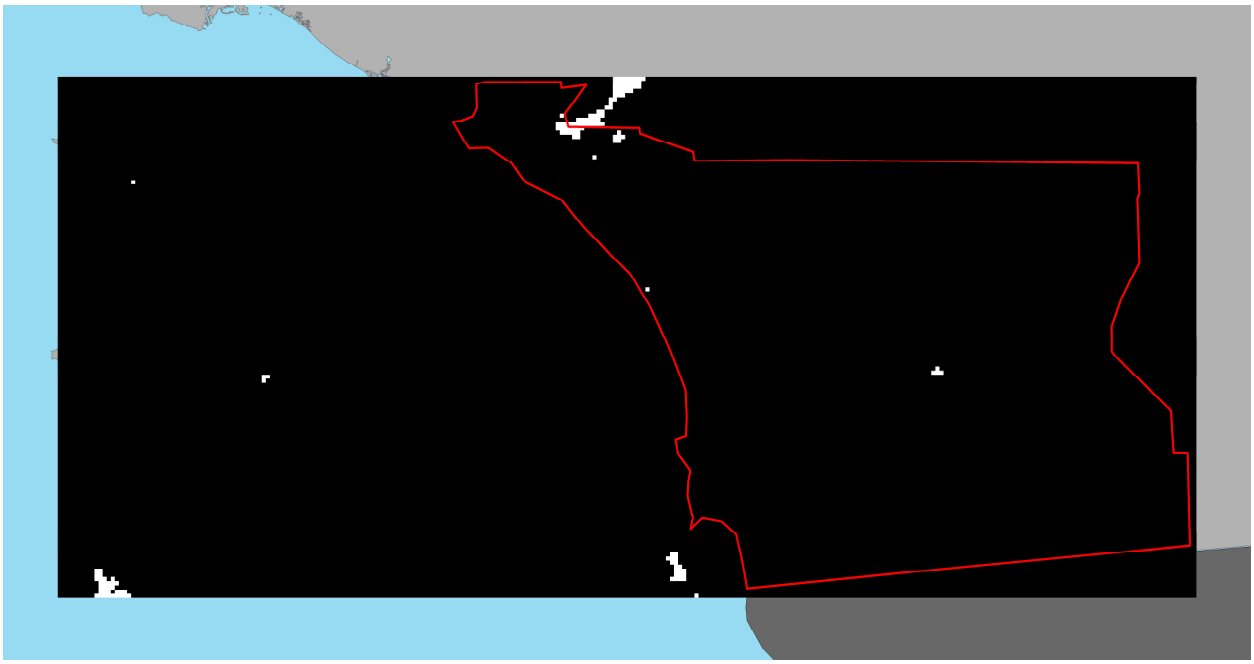


Figure 8 Output to yes/no

4.2. Inversion Detection Algorithm

Atmospheric sounding data that can be used for detecting temperature inversions typically come in a format resembling the following:

Table 1. Example of sounding data.

PRES	HGHT	TEMP	DWPT	RELH	MIXR	DRCT	SKNT
hPa	*m*	*C*	*C*	*%*	*g/kg*	*deg*	*knot*
998	128	14.6	14.1	97	10.23	10	5
995	154	15	14.6	97	10.6	11	4
978	303	15	15	100	11.08	15	1
977.8	305	15.1	15.1	100	11.12	15	1
973	347	16	16	100	11.89	38	1
…							

It is typically sampled at various air pressure levels, each with corresponding heights and other parameters (temperatures, wind speeds, etc.).

Once the algorithm has finished running for a given set of sounding data, the altitude, temperature, wind speed, and wind direction are recorded at the heights identified as the base and top of the temperature inversion, see Figure 10 for two examples of inversion detection.

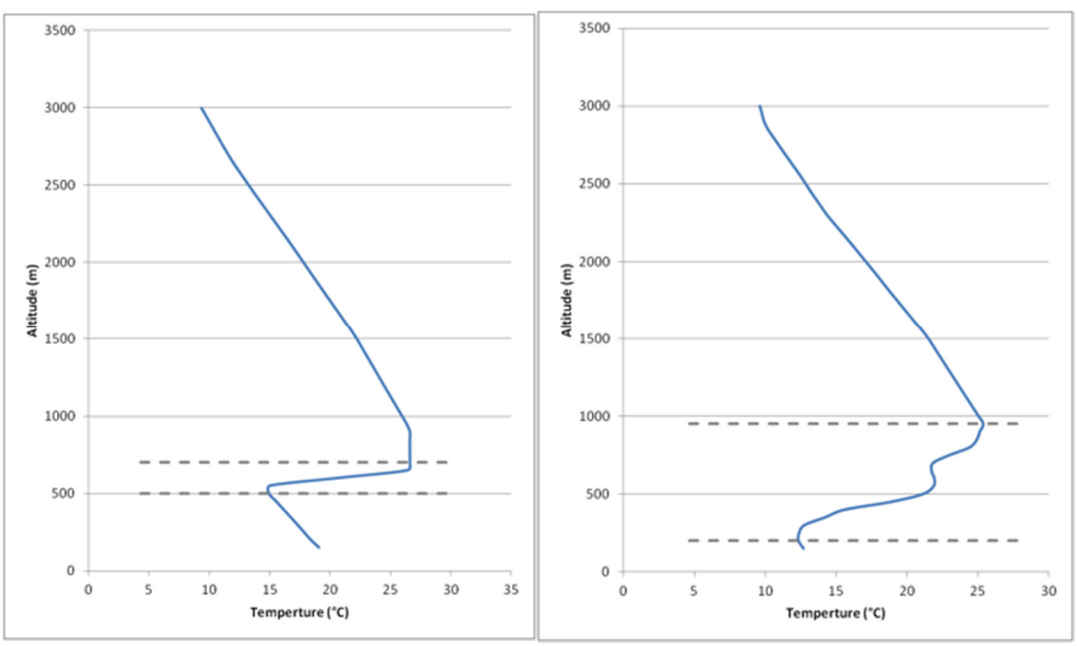


Figure 9. Examples of automatic detection of inversion base and top from a simple (left) and more complex (right) temperature profiles.

4.3. Local Analysis and Prediction System / Weather Research and Forecasting Model Marine Layer and Total Cloud Forecasting Methods

LAPS / WRF was set up to run on the San Diego County area for the purpose of marine layer detection. Our original WRF forecast utilized the Rapid Refresh (RAP) NWP to populate the model seed where observational data was not available. The RAP NWP is released once per hour and has a forecast window of 18 hours and a native resolution of 12 km.

The original LAPS/WRF set up was a grid with a 3km exactly matching the forecast area (above). From the first forecast runs it was evident that boundary errors were appearing in the WRF forecasts along the border, occasionally reporting GHI value in excess of 1500 w/m^2 or values systematically different from surrounding areas . The spatial boundaries were then extended to give a buffer of at least 10km surrounding the original area of interest on each side. The buffer zone ensured that no erroneous forecast values from boundary conditions were found in the final forecast area. Figure 10 below is an example of the extended forecast area. Note that the strip along the top and right hand sides displays these boundary errors.

It was also found that unless the NEXRAD WSR-88D radar data were normalized to include nothing above 200 decibels (dBZs), there were errors showing up in the cloudiness index and GHI variables in the LAPS analysis, see Figures 11,12 , and 13 respectively.

Following the extension of the forecasts spatial extent and the radar upgrade, the forecast length was then increased from 18 hours up to the required 36 hours.

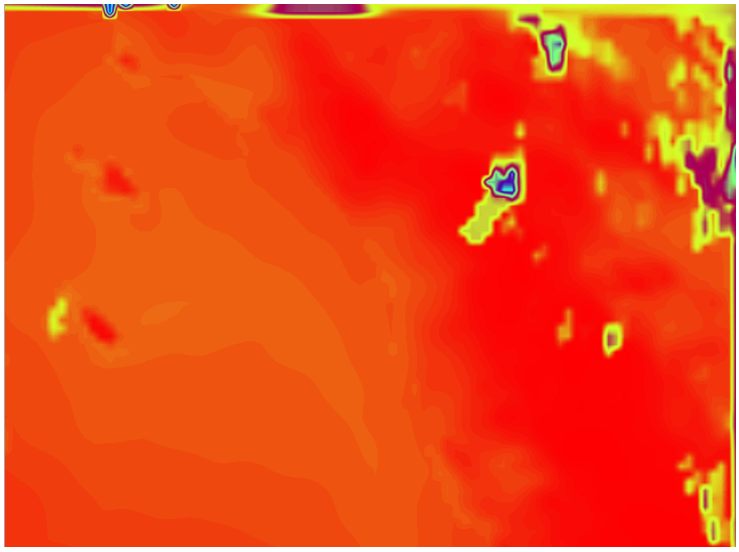


Figure 10. Example of the extended forecast area with boundary errors. Note the strip along the top and right-hand sides that report erroneous GHI values. By extending the forecast boundary, and clipping for the area of interest, the boundary condition errors do not affect the analysis.

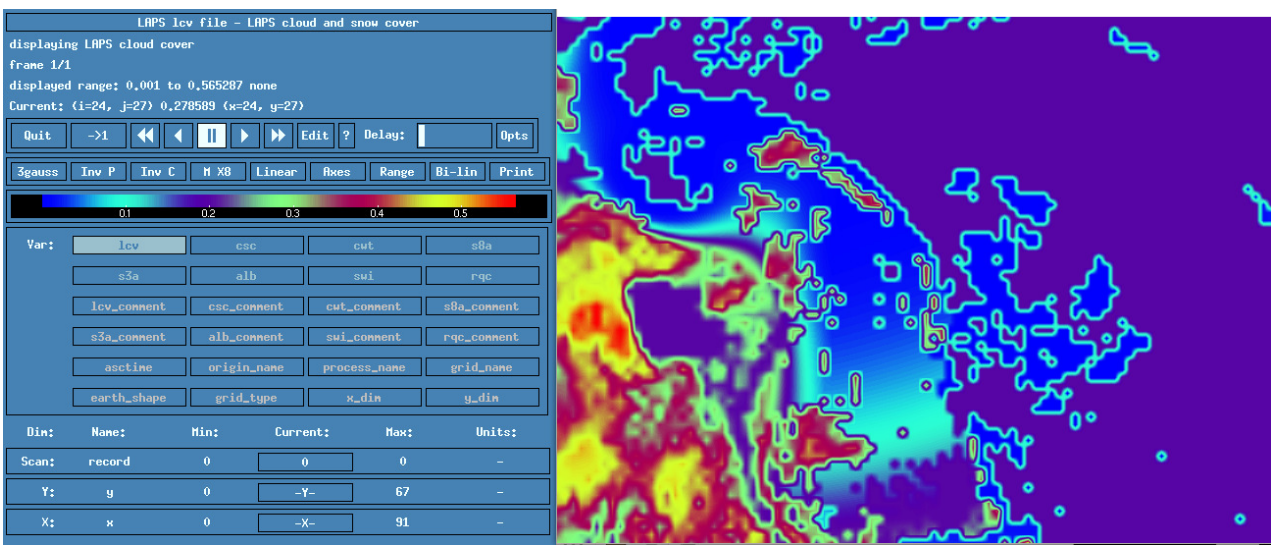


Figure 11. The LAPS analysed cloudiness index for the same region with non-normalized radar

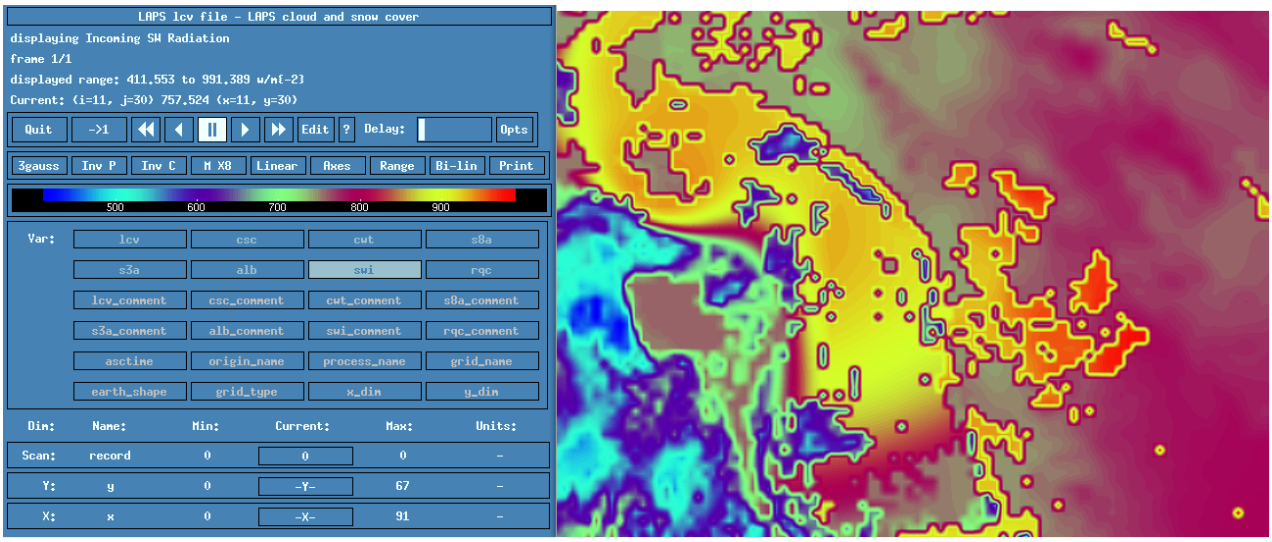


Figure 12. The GHI values for the forecast area with non-normalized radar

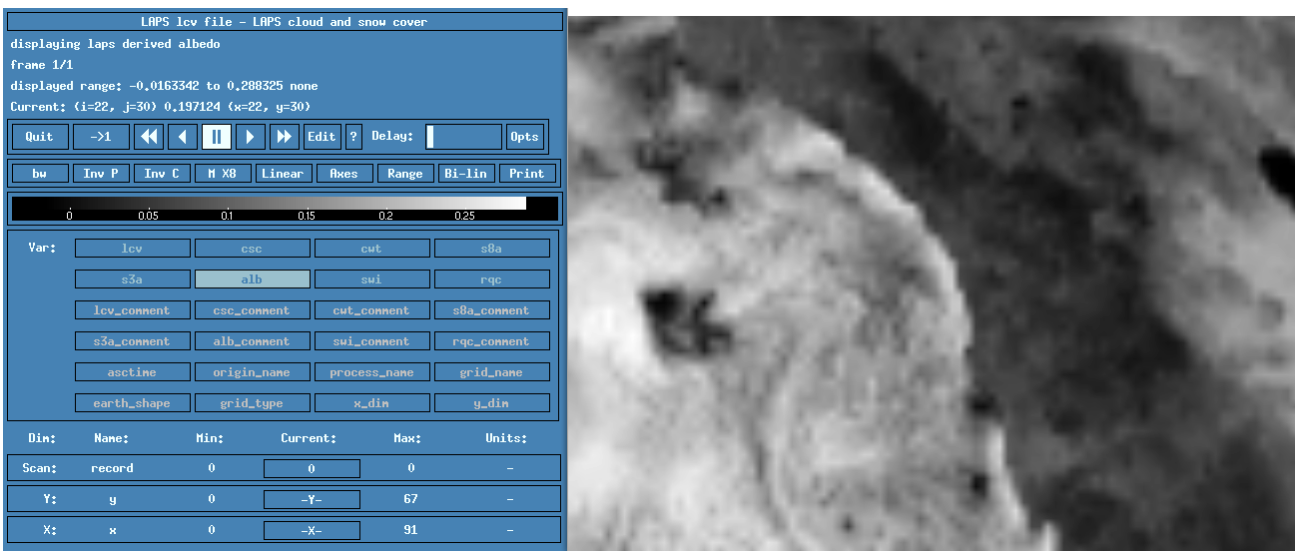


Figure 13. The LAPS analysed albedo values from the satellite images with non-normalized radar.

4.4. Pattern Recognition Method

The Pattern Recognition (PR) model uses historical patterns in the occurrence of the ML to forecast future ML events. This is done by taking the most recent classified ML and looking through a database of historical ML events (2000-present). To create a forecast, the PR stat is calculated for every ML event in the historical database with reference to the current ML to be forecast. This forecasting can, in theory, be applied to any desired time horizon, but the current model was only tested and optimized for a time horizon of 36 hours.

4.5. Probabilistic Model

The probabilistic model (PM) uses the historical rates of the disappearance/reappearance of ML to forecast future ML based on the current classification. Using the historical dataset of ML events, a set of probability grids were developed that use a Markov chain based approach to determine the probability of the ML existing in each of the grid cells in future forecast periods. The resultant grids of probability values are then passed through a filter to turn them into binary ML present/absent grids.

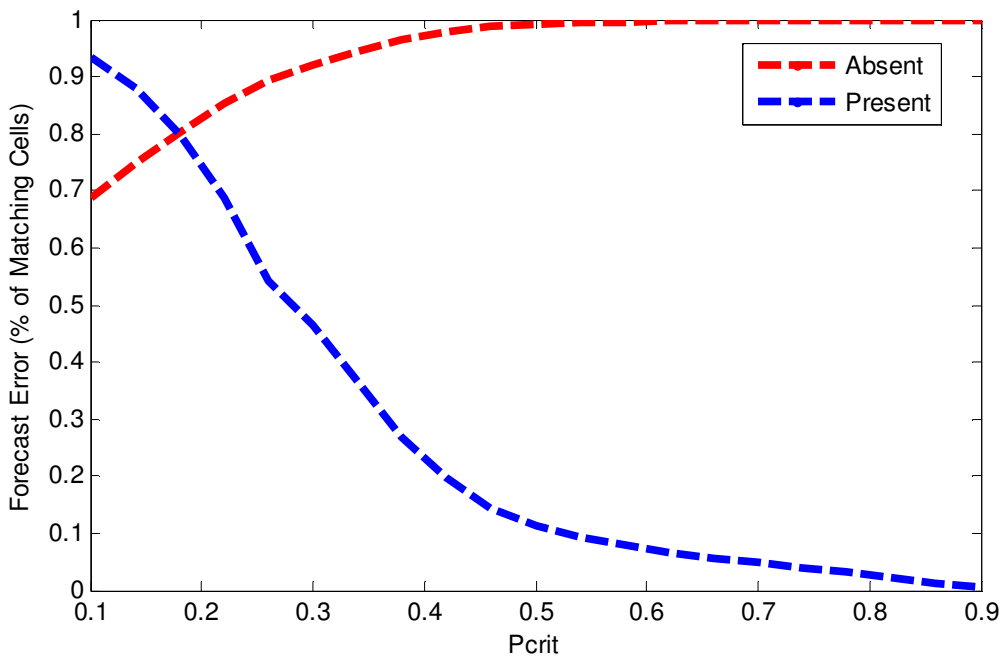


Figure 14. Performance of PM algorithm for clear sky vs. ML conditions.

As is readily apparent (Figure 14), the higher the *Pcrit*, the better the performance on absence and the worse the performance on presence. While this may make the case for a lower Pcrit, It has to be noted the ML is absent roughly 85% of the time, hence it would make sense to bias the model towards predicting absence. Figure 15 below, shows the error when the presence and absence are weighted equally (i.e. assuming they each occur 50% of the time) and weighted based on the weights of actual occurrence:

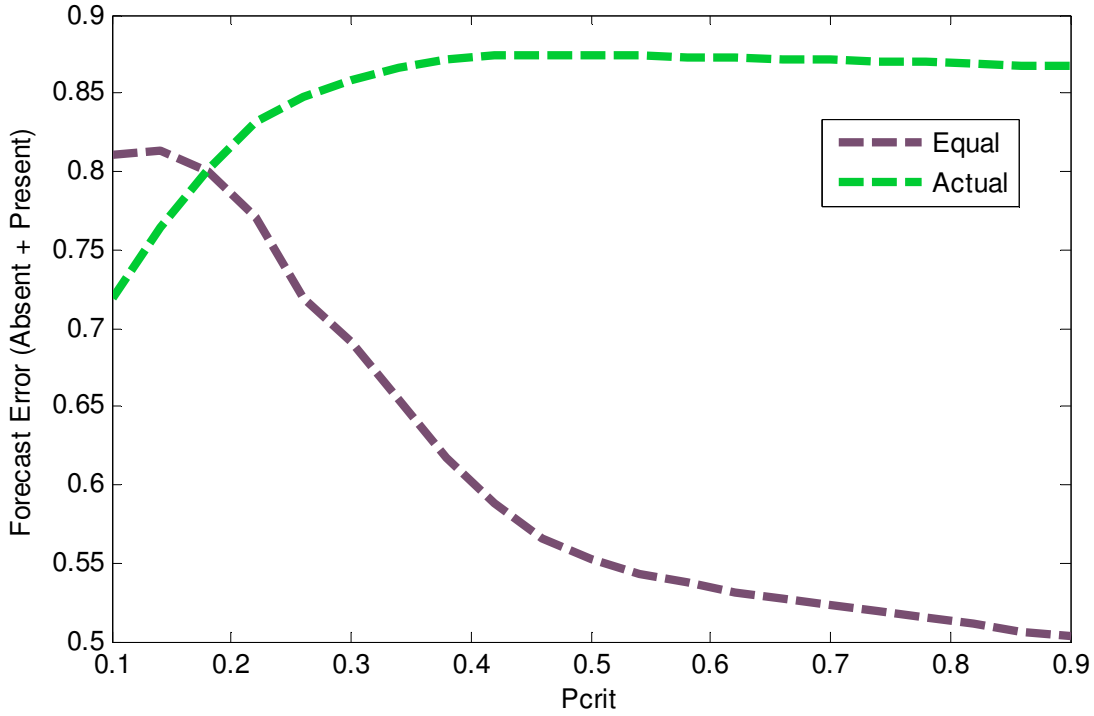


Figure 15. Performance of PM algorithm using two accuracy estimates.

5. Results – ML forecast Performance

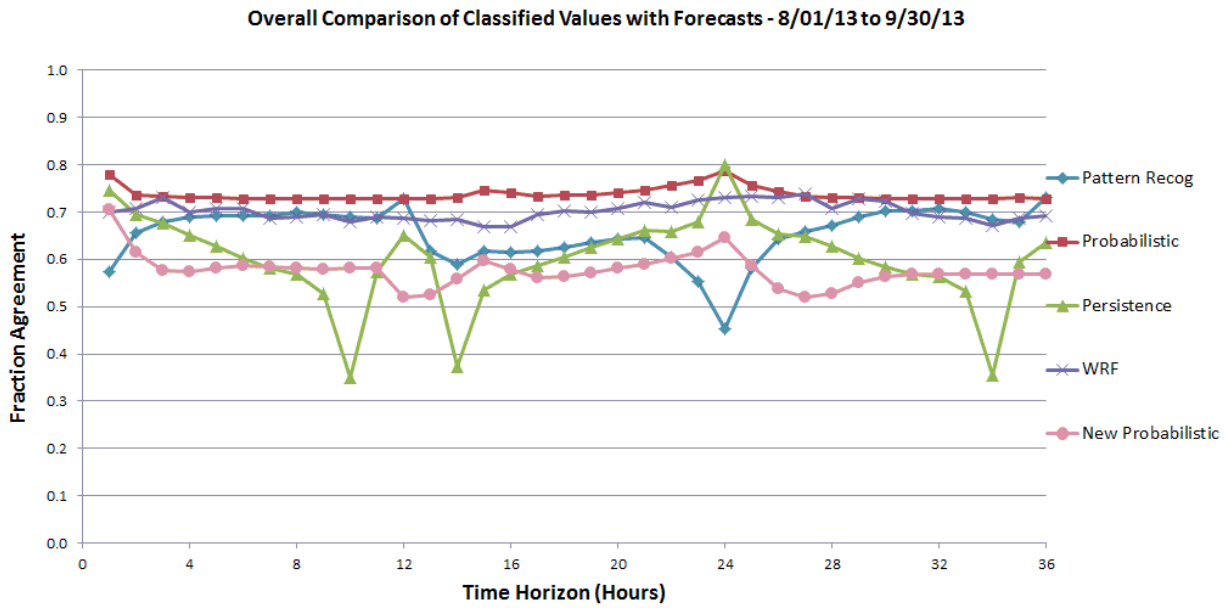
This section addresses performance of the developed system, with reference to different development stages and using several accuracy estimates. Please note that unless otherwise stated, all values are calculated using only daylight hours.

The "Probabilistic" algorithm is the version using a 50% detection threshold, while the "New Probabilistic" method uses a 15% threshold. Unless otherwise noted, the analysis was done for the period of August 1, 2013 to September 30, 2013. The entire 34750 point grid for the SDGE service area is not used in the analysis. Instead a study area consisting of 8049 points that lie in a band between 20 km offshore to 20 km inland of the coast is used. For each set of forecasts, the forecasted marine layer from one to 36 hours out is compared to the classified values for the corresponding time. For example, for the classification created on August 31 at 12:00 PT, the hour-ahead forecast values made at 11:00 were compared to the 12:00 classified values, the two-hour-ahead forecast values made at 10:00 were compared to the 12:00 classification, and so on. All forecasts in the two-month period for all study area locations were binned for each forecast hour, and then the average agreement between forecast and classified was computed for each hour in the time horizon. Only daytime classified values are used in the calculations (daytime defined as 7am to 7pm PDT) though all applicable forecasted values were included. For the Persistence method the previous classified values were used instead of forecasted values.

Two methods are used to compute the agreement between forecasted and classified values. The "actual" method uses no explicit weighting and is a simple average of the agreement at each forecast hour, and thus reflects the actual weather conditions present. The "equal" method is based on the formula used by UCSD:

Agreement = [(# forecast clear & observed clear) / # observed clear + (# forecast cloudy & observed cloudy) / # observed cloudy] / 2.

It should be noted that there is a bias built directly into the equal weighting method that skews the analysis in favour of cloudy days. Both clear and cloudy days are weighted equally here; where it is known in the San Diego area that marine layer is absent roughly 85% of the time. The charts below show the overall comparison between forecasts and classified values.



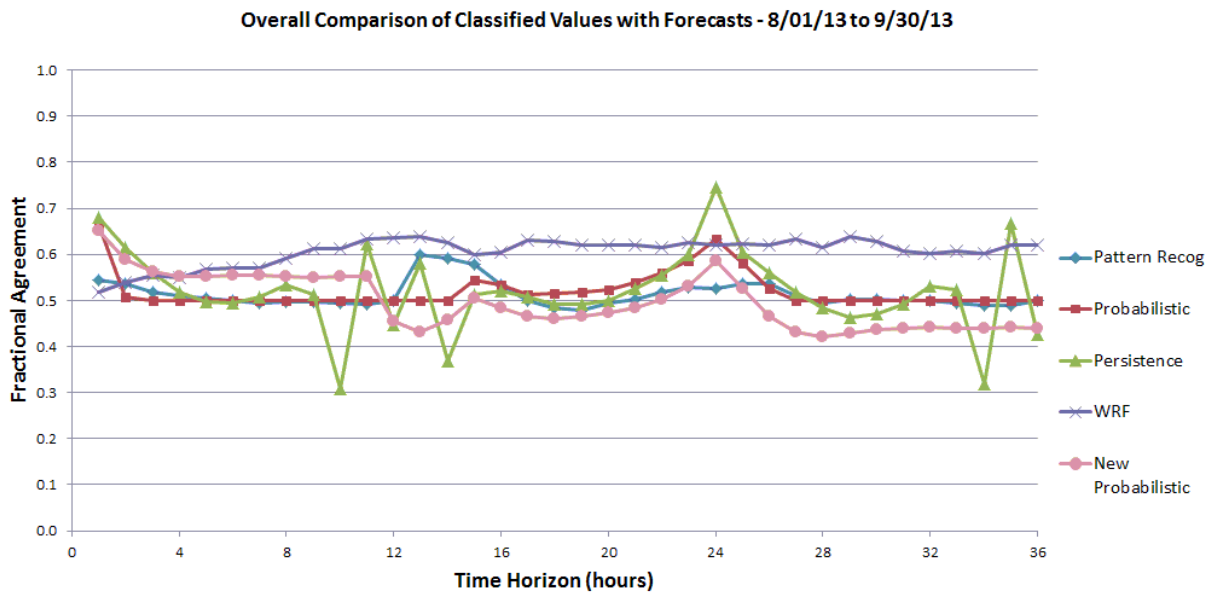


Figure 16. Overall agreement between classified values and forecasted values, averaged over the study area, by forecast hour. The top figure shows the fractional agreement using the "actual" agreement, as defined above, whereas the bottom figure shows the same using the "equal" agreement scheme noted above.

In Figure 16, we present the fractional agreement between the classified ML points and their corresponding forecast points. The actual comparison methodology indicates the most effective method is the Probabilistic algorithm over almost all the entire 36-hour horizon. The equal comparison shows that

the Persistence method is best (though only slightly) for the first two hours, the New Probabilistic method is best for the next two hours, and after that the WRF model is generally the best performer.

Figure 17, below, shows the performance of the algorithms depending on the actual presence or absence of marine layer (calculations were done separately for absence or presence as indicated by the classification). The Probabilistic method is very good at detecting absence, but poor in detecting presence.

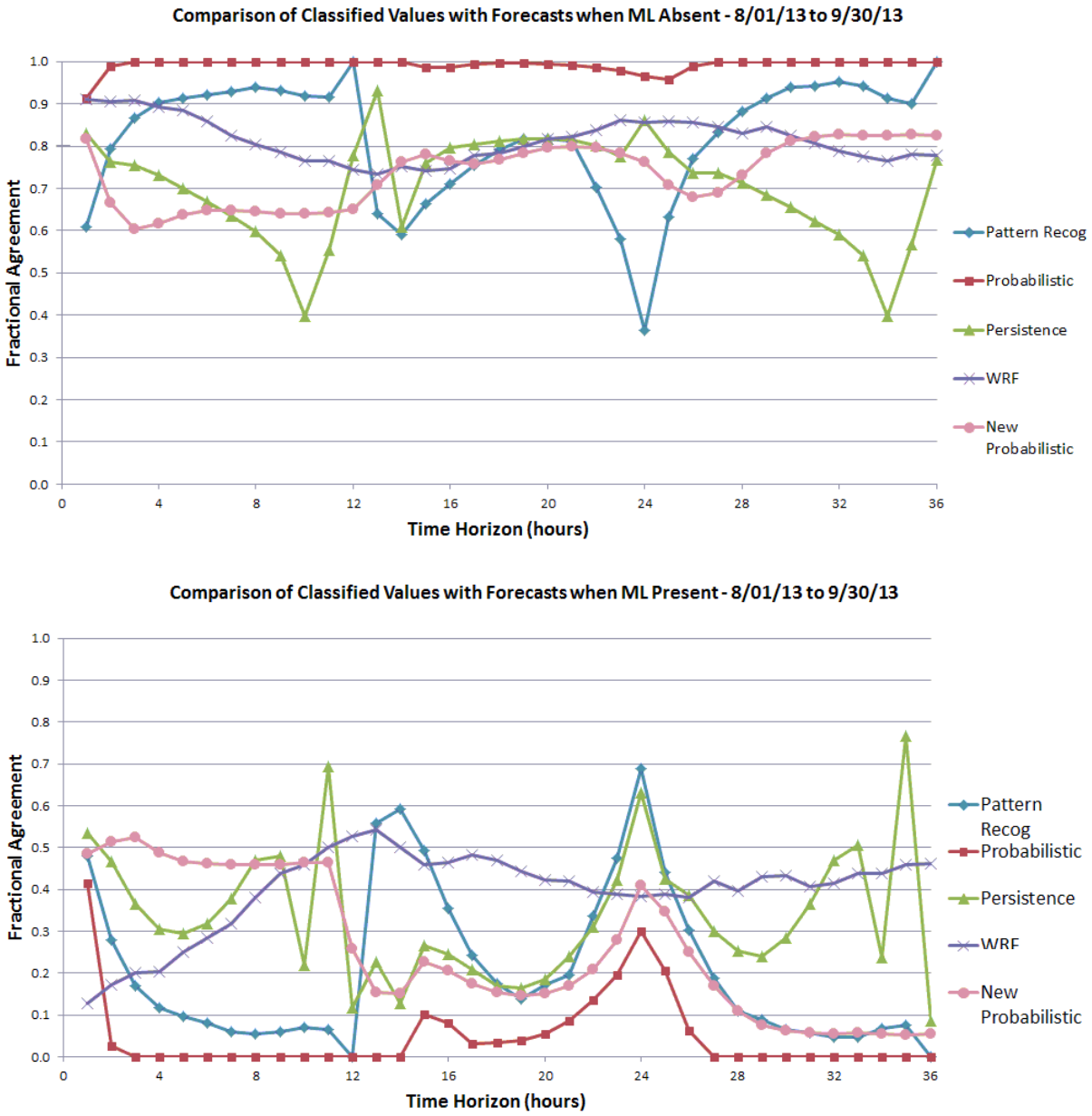


Figure 17. Performance of the algorithms for clear and cloudy cases. Performance is defined as the average agreement between the forecast and classified satellite image.

Figures 18 and 19 below show the algorithm that performed best for each of the locations in the study area for the indicated forecast hour.

The “equal” method puts more emphasis on the presence of clouds, so it is important to consider its implications to the forecasts in the coastal zone. Consideration of the performance based on both methodologies was used to make a decision on the best mix of the forecasting algorithms overall.

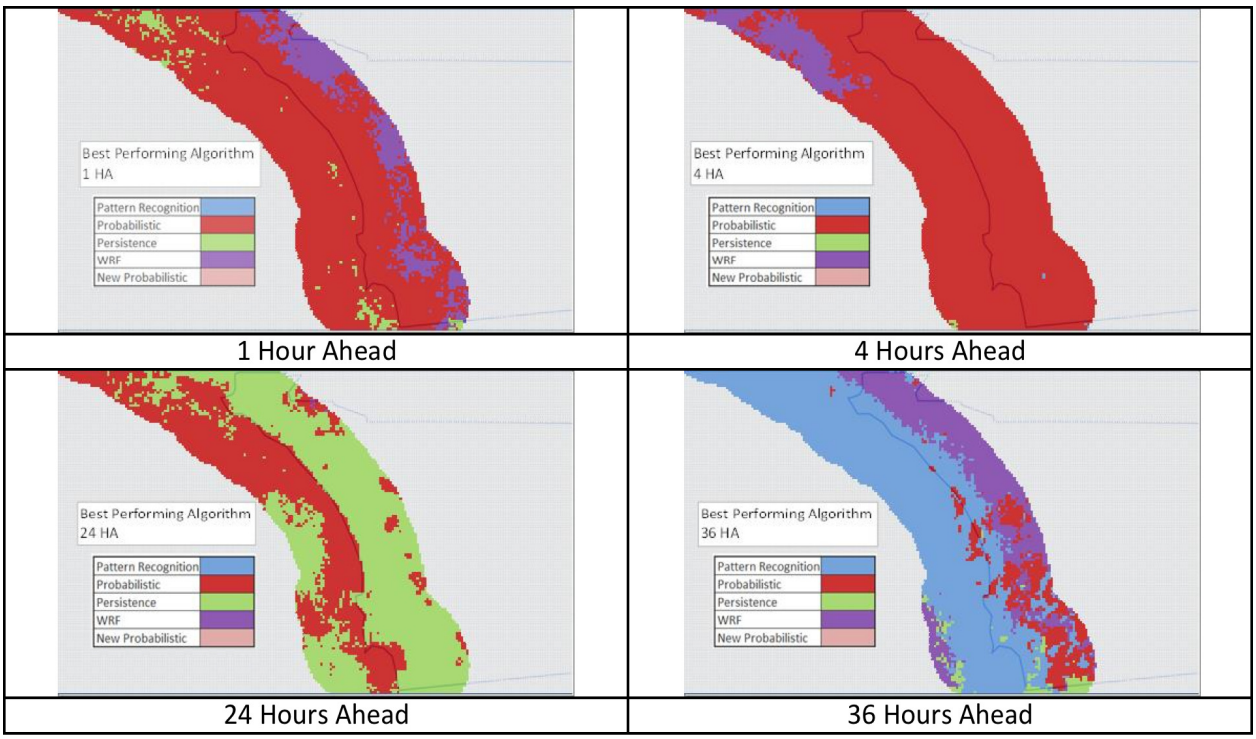


Figure 18. Best performing algorithm by location and forecast hour, actual weighting. See Figures 16 and 17 for colour references.

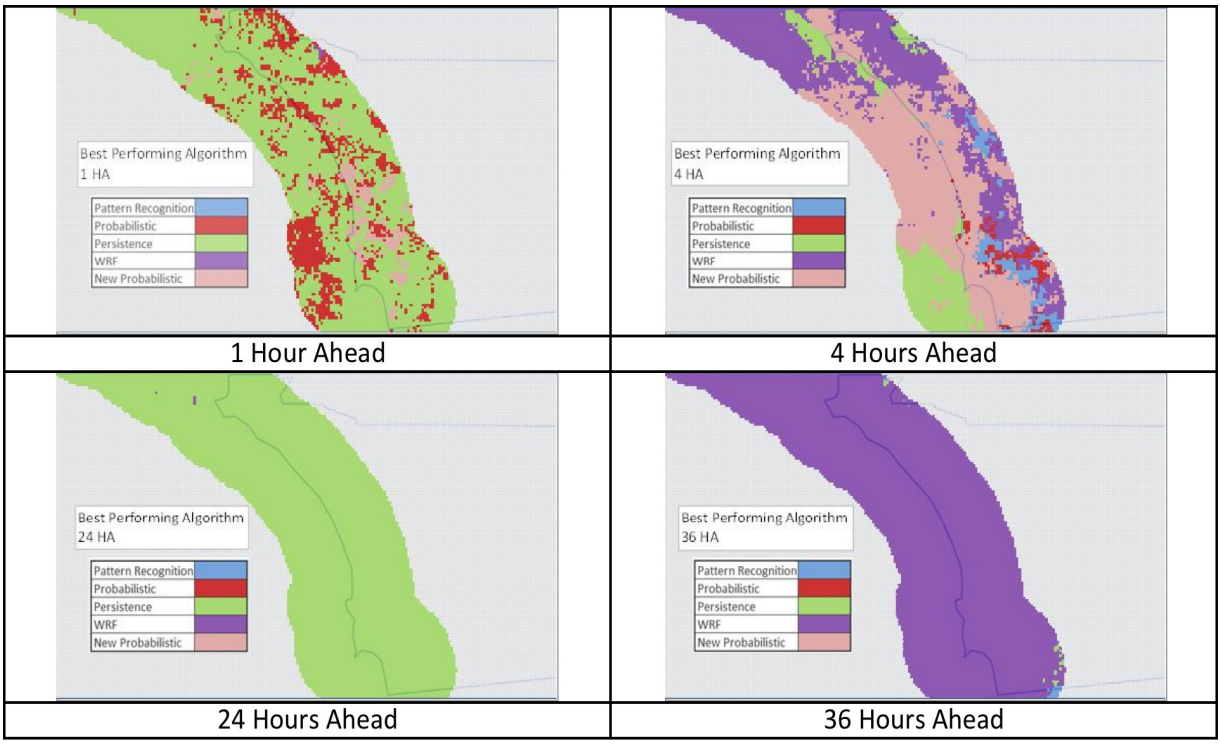


Figure 19. Figure: Best performing algorithm by location and forecast hour, equal weighting.

Figure 20 tracks performance on a weekly basis. A week runs from a Monday to the following Sunday, and the dates on the x-axis denote the start of the week. Each point is the average agreement over the week over all points in the study area over the entire forecast time horizon.

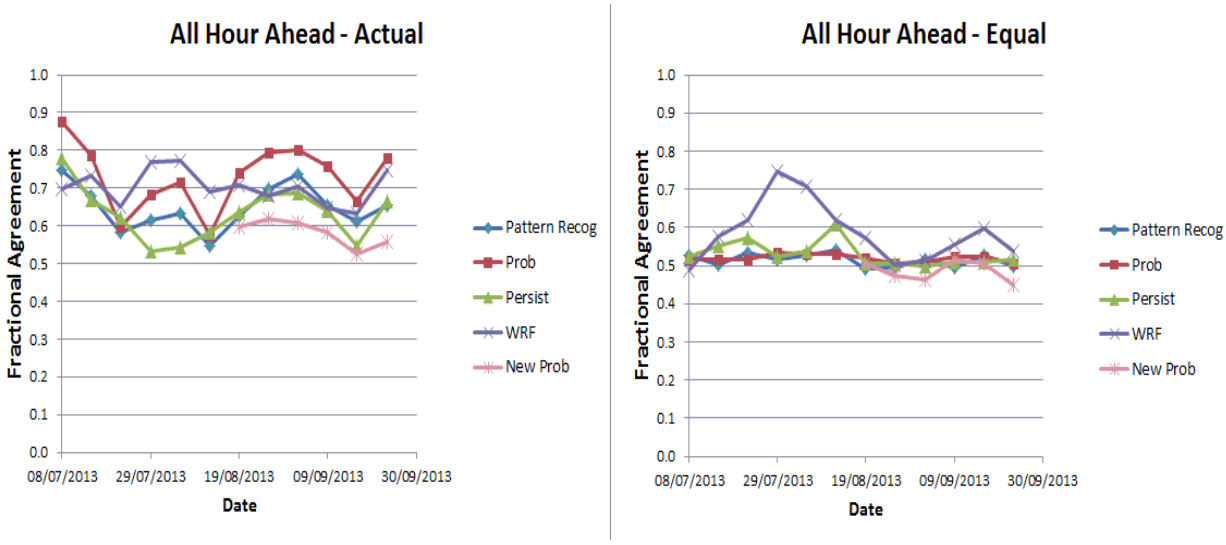


Figure 20. Weekly performance averaged over all points in the study area and over all forecast hours. Points are plotted at the start of the Monday-to-Sunday week.

The next set of charts, Figure 21, show the same type of analysis, but here comparisons are only done if marine layer is present or absent as indicated by the classified values. In addition, the additional orange curve in each chart shows the average fraction of the time that marine layer was absent or present, as determined by the classification.

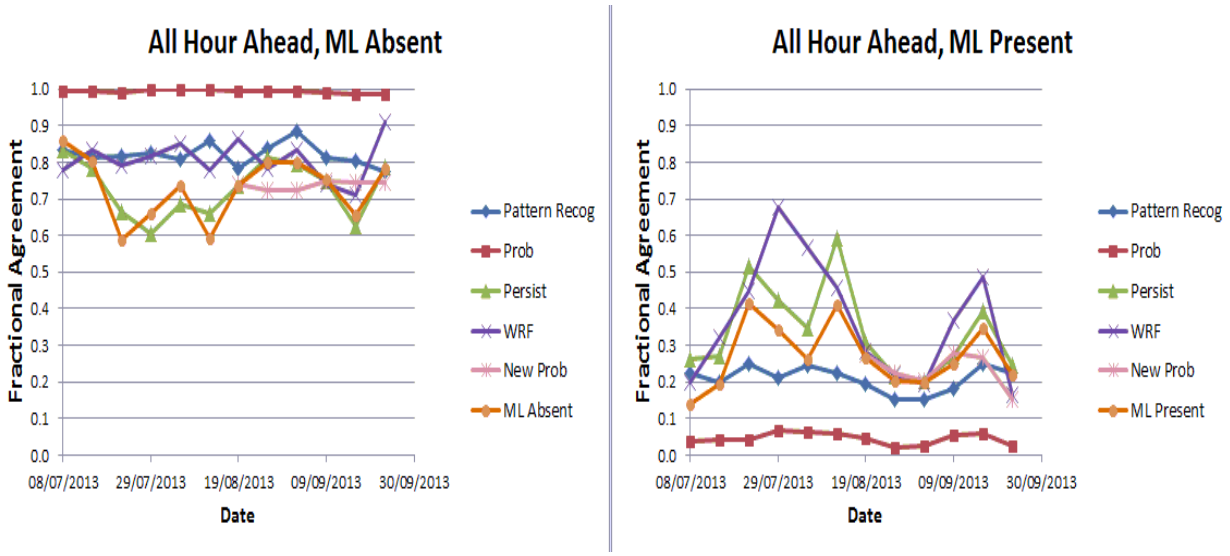


Figure 21. Weekly performance averaged over all points in the study area and over all forecast hours, when marine layer was absent and present. Points are plotted at the start of the Monday-to-Sunday week.

Nighttime classification was implemented on September 11. Figure 22, 23, 24, and 25 below show the performance of the forecasts made at 7 and 9 PDT before and after the implementation date, respectively. The results generally seem better after the implementation of nighttime classification, the notable exception being the first hour of the 7 am forecasts where the algorithms performed more poorly after implementation. Interpretation is difficult because the marine layer situation itself will have differed between the two time periods (the figure above indicates a greater presence of marine layer in the latter period).

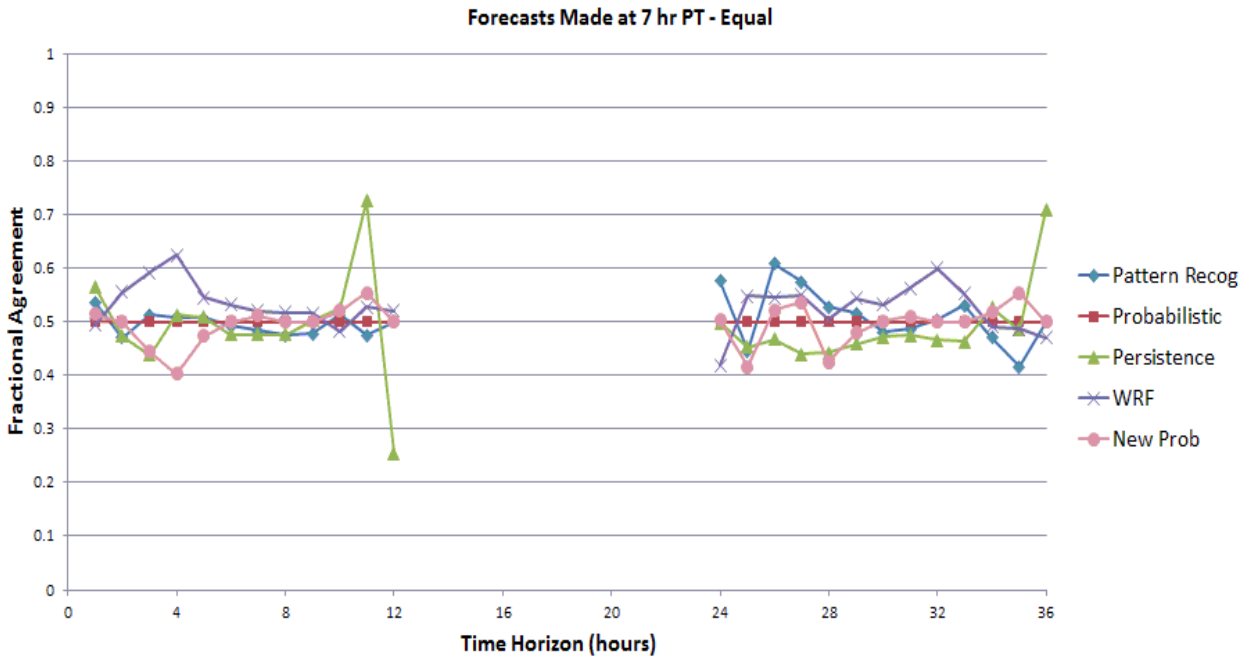


Figure 22. Figure: Performance of forecasts made at 07 hours PDT from August 23 to September 10 (19-day period before nighttime classification was implemented)

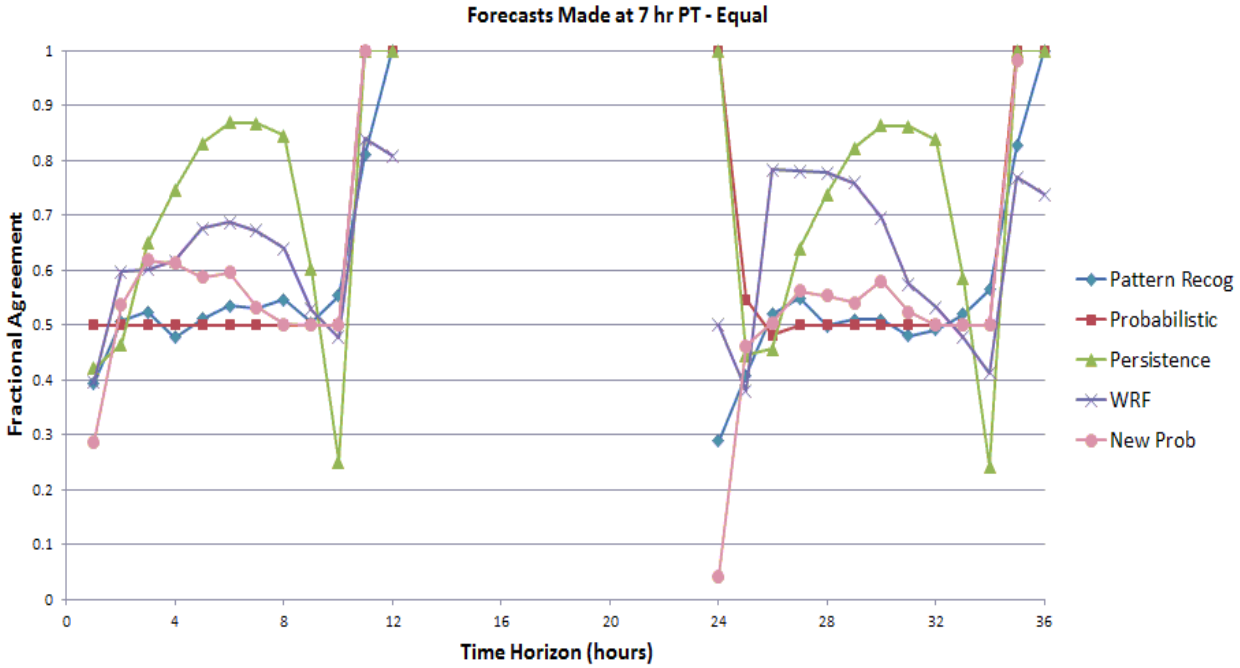


Figure 23. Performance of forecasts made at 07 hours PDT from September 12 to 30 (19-day period after nighttime classification was implemented).

For 7am forecast introduction of nighttime classification resulted in a better performance of persistence for the night hours, current forecast day as well as for the next mid day. PM approach also performed better at evening time.

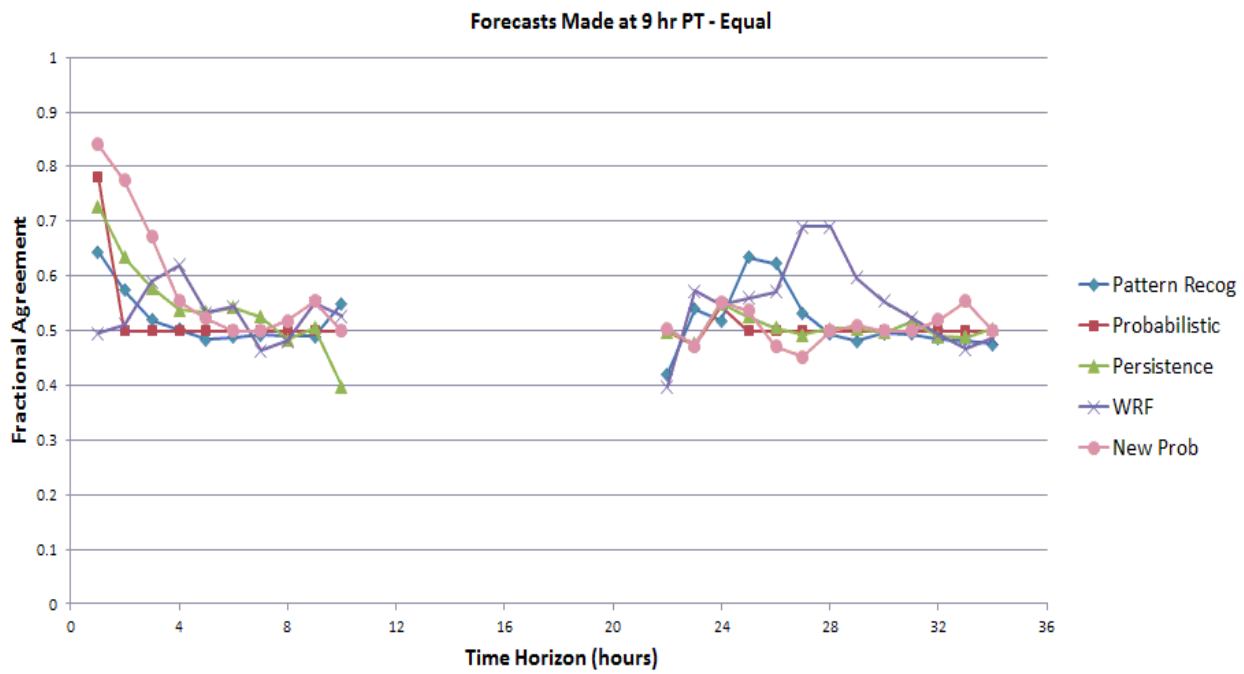


Figure 24. Performance of forecasts made at 09 hours PDT from August 23 to September 10 (19-day period before nighttime classification was implemented).

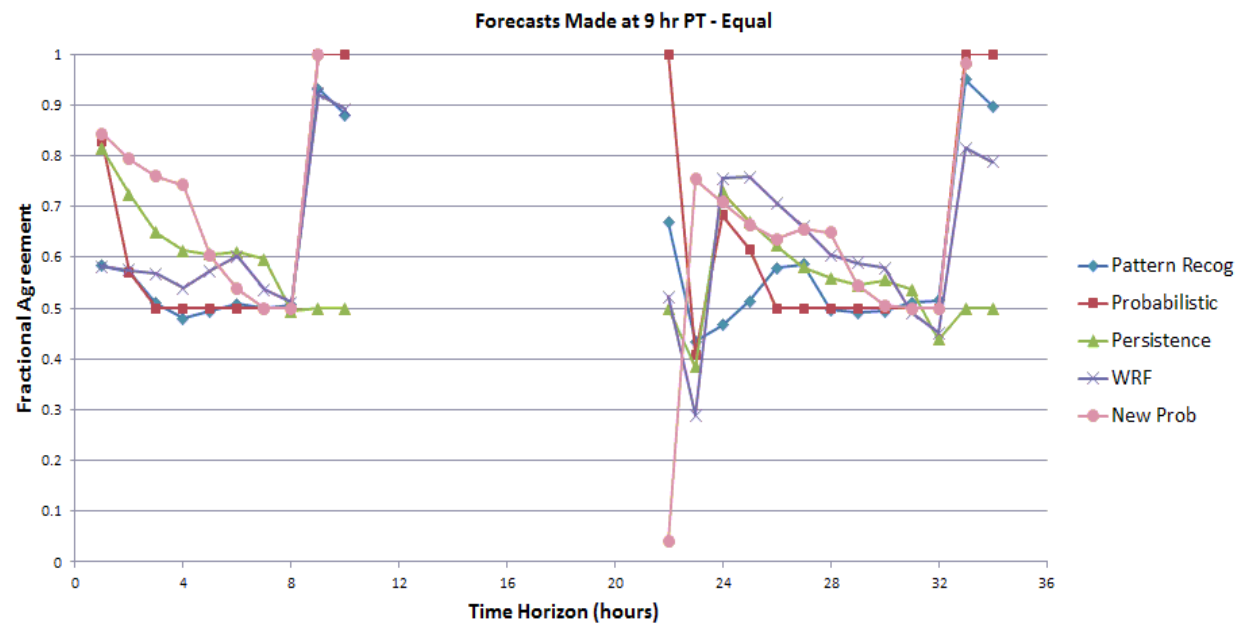


Figure 25. Performance of forecasts made at 09 hours PDT from September 12 to 30 (19-day period after nighttime classification was implemented).

For 9 am forecasts and later the introduction of night ML classification improved both PM and WRF based forecasts.

Table 2 below compares the classified values to observed cloudiness values obtained from 12 METAR stations (Figure 26). Marine layer was defined as the presence of fog, mist, or broken or overcast clouds with heights less than 1500 feet. Classified values from the grid points closest to the weather stations were used. Note that the last two stations (KSEE and KRNM) are outside the coastal-band study area.

Table 2. Agreement between classified and observed marine layer for 12 weather stations for the period May 20 to September 30, between the hours of 7am and 7pm PDT.

METAR Station	Distance from Coast (km)	Agreement (Actual)	Agreement (Equal)
KNXF	0.2	72.4%	66.7%
KNZY	0.2	75.5%	71.0%
KSAN	0.7	75.3%	69.4%
KNRS	1.9	73.5%	69.1%
KOKB	4.3	73.7%	67.5%
KCRQ	5.3	71.7%	67.0%
KNFG	9.6	76.8%	71.6%
KNKX	10.2	79.6%	72.2%
KMYF	10.2	79.6%	71.9%
KSDM	10.4	76.5%	67.2%
KSEE	22.0	78.0%	67.8%
KRNM	33.7	82.0%	63.1%

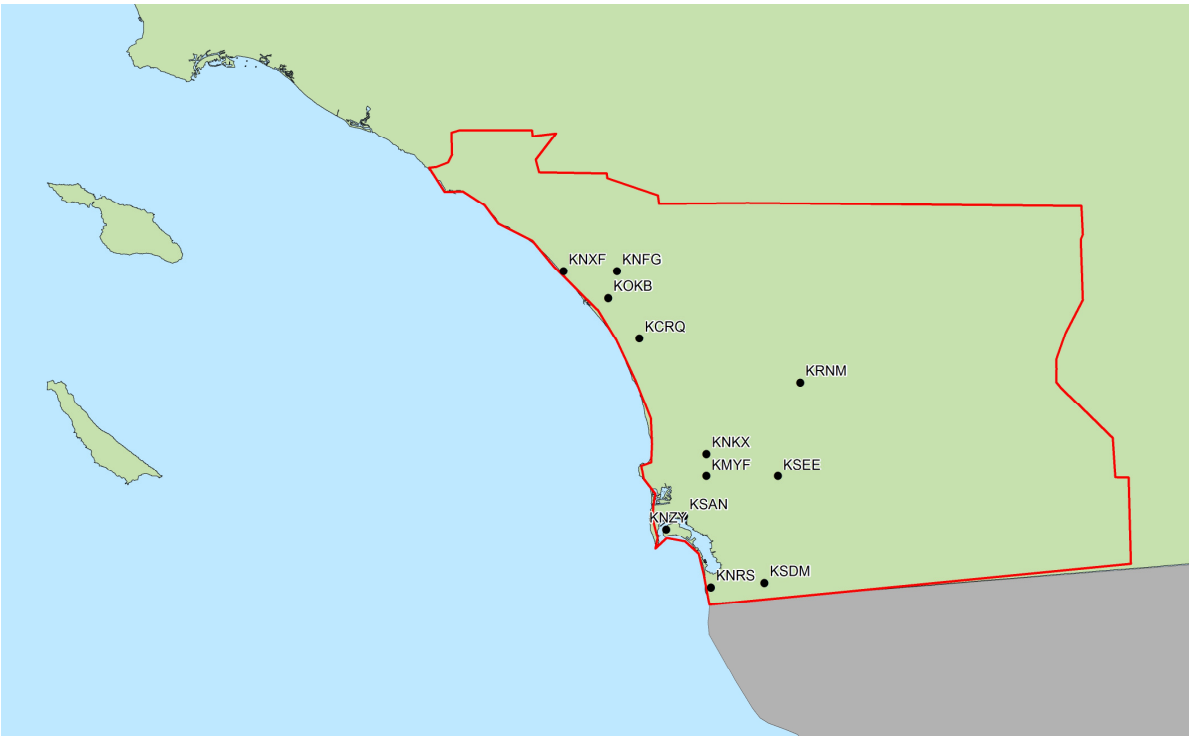


Figure 26. Locations of METAR stations used to compare observations with classified values

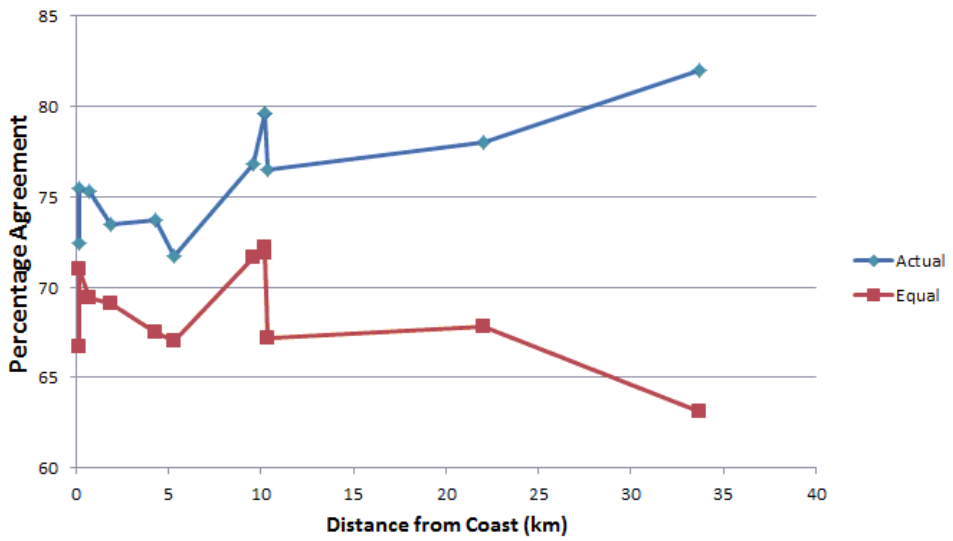


Figure 27. Agreement for Actual and Equal scheme as a function of distance from coast. It can be seen that the actual scheme forecasts best inland, but is always outperforming the forecasts using the equal scheme,

The results of testing forecasts vs METAR observations are very similar in magnitude to the comparison of marine layer forecasts to the classified grids with forecast accuracy of 60 – 80 %.Figure 27 shows the information from Table 2 visualized. It can be seen that for the actual agreement values, there is a trend seen where the forecasts are better as you go further inland. For the equal weighting scheme the values are relatively flat.

6. Conclusions

A number of approaches were tested in developing ML forecasts for 36 hours ahead. They performed with different accuracy depending on the geographic location, time of day, and forecast horizon. As a result a combination of Probabilistic and WRF approach were adopted as best suited for marine layer forecasting in the San Diego County area.

The best (currently) mix of models is the use of probabilistic forecast for the current day (7 am – 7 pm) and the use of LAPS/WRF for hours beyond that. Depending on the critical value of probabilistic threshold PM approach captures well attenuation of ML through the day. In the current configuration LAPS/WRF is capable of developing future cloud fields while being relatively weak in capturing them within the short time horizon.

Although the persistence model appears to do well on the equal weighting scheme, when one looks at the actually occurrence of ML, that is not the case. The reason that persistence does well on the equal weighting scheme is based on the nature of the weighting scheme itself, putting much more emphasis on clouds being present. As discussed above, in general, ML is not present a large majority of the time, thus the equal weighting scheme is heavily biased. Persistence also performs well in the hour ahead diagnostics for several reasons. Compared to WRF, this is due to the fact that numerical models need several hours of spin up time to ensure that clouds are well established and for probabilistic (50%), the ML retreats faster than it occurs naturally. Persistence compared to probabilistic (15%), it can be seen that persistence is not outperforming in the first few hours (Figures 24 25). The above speaks to the need to combine different forecasting techniques to ensure the best possible, accurate, forecast is delivered.

Testing ML forecasts operationally is a challenge because of the seasonal variability in ML presence. Despite these constraints we have had several months of continuous testing in the summer 2013. The results of our work are encouraging with overall agreement of forecasts with observations of 60 to 80% (as in number of 1km cells in the service area where ML is forecasted correctly).

It is expected that LAPS/WRF, with implemented improvements, will become a source of GHI forecast for SSD. Therefore real time ML classification results will have to be integrated into LAPS/WRF "on the fly" to improve short time horizon forecast accuracy. Testing should be carried on integration of ML forecast with the existing SSD software as well.

It is important to carry on further development of forecasting techniques to improve accuracy by adjusting classification techniques, introducing real time forecast validation, improving PM, PR and LAPS/WRF models.