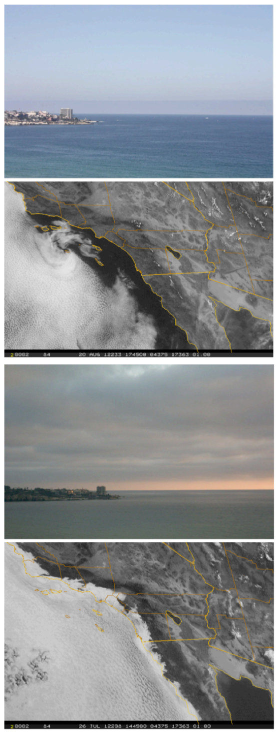
Forecasting the Marine Layer Front through Physical Inputs and Machine Learning Techniques

Lukas Nonnenmacher[1], Rachel E. Schwartz[2], Vladimir Kostylev[3], Carlos F.M. Coimbra[1]

1. Mechanical and Aerospace Engineering, University of California, San Diego, La Jolla, CA, United States; 2. Scripps Institute of Oceanography, University of California, San Diego, La Jolla, CA, United States; 3. Green Power Labs, Dartmouth, NS, Canada

Background

Southern California features one of the highest rates of renewable energy electric grid penetration in the US (about 14.6%). As proposed in the Renewable Portfolio Standard for the State of California, 33% of energy consumed must be supplied by renewable sources by 2020. Solar energy will play a significant role in reaching this goal. The weather in the coastal area of Southern California is highly influenced by the occurrence of low Marine Layer Clouds (MLC) during the summer season. An extensive marine layer system forms over the Pacific Ocean and is frequently propelled inland, causing overcast meteorological conditions on distributed and utility scale solar farms, which in turn cause ramping events and fluctuations in power amplitude and frequency. The formation, movement and behavior of marine layer clouds is modeled as a multivariable system that is governed by factors such as local topography, pressure gradients, coastal winds, solar radiation, etc. To support the decision making process of power suppliers, this work focuses on a pattern recognition based kNN (k-Nearest Neighbor) forecasting system that covers intra-day forecasts reaching from 30 minutes to 4 hours ahead. To achieve higher accuracy, this basic system is combined with decision variables to identify and utilize similar previously recorded conditions.



Dataset

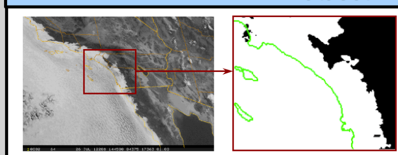


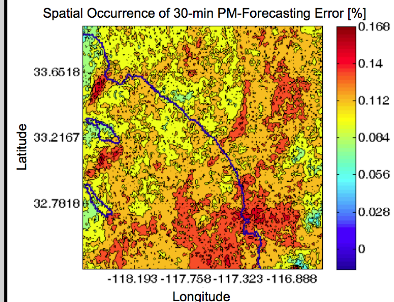
Figure: Example of a image form GOES-West satellite and its translation into a binary Marine Layer Cloud image for the San Diego area.

Procedure: A pair of corresponding VIS and IR images are used to automatically detect MLC. VIS images cannot differentiate low stratus clouds and fog from high altitude clouds. IR images (i.e. temperature maps) can not detect fog (temperature very close of ground temperature), but clearly detects high altitude clouds. Cross-classification of IR and VIS images allows a clear identification of MLC.

Database: 19520 binary images form 2000-2007 + 3 years of test data.

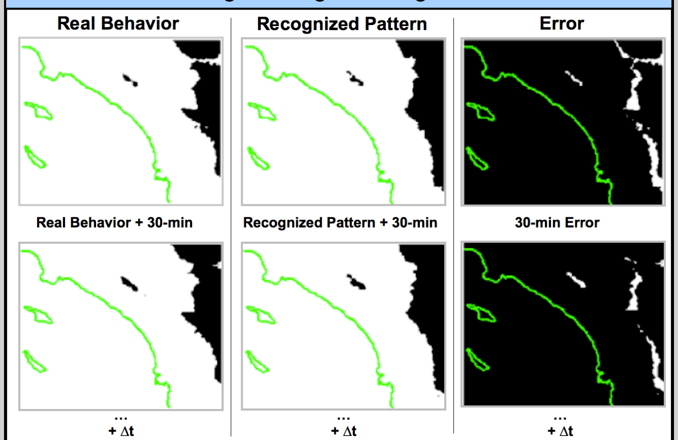
Marine Layer Behavior

Highly Variable Regions

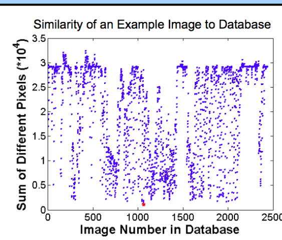


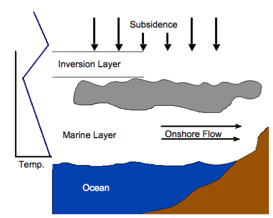
Spatial Distribution of 30-min Persistence Model Error - The figure shows the error in percent of the persistence model applied to binary satellite derived cloud images for 8am - 4pm for May-August. The highest MLC variability occurs in San Diego's Coastal area and coincides with high Photovoltaic (PV) electric grid penetration.

Image Recognition Algorithm



Combination with Decision Variables (Inversion Base Height)

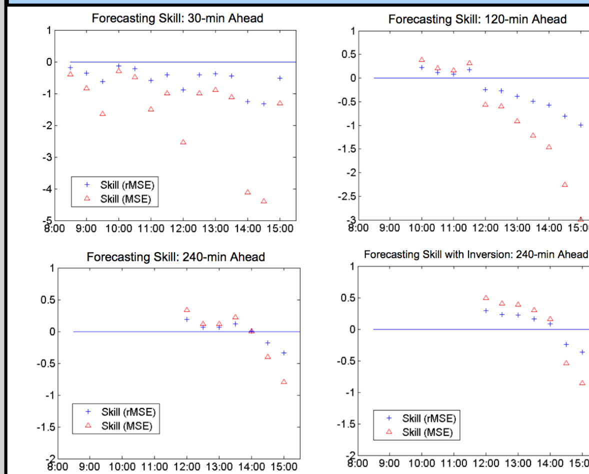




Forecasting Skill

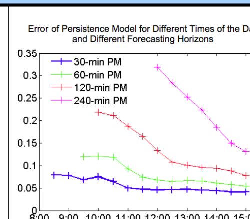
$$s = 1 - \frac{MSE_{PattRec}}{MSE_{Persistence}} \qquad s = 1 - \frac{RMSE_{PattRec}}{RMSE_{Persistence}}$$

Preliminary Results



Discussion

- Positive forecasting skill for long-term, intra-day forecasts (+2h)
- Forecasting skill improves due to poor performance of persistence model for longer time horizons
- Method works for a very specific intra-day time horizons (2h - 4h ahead beginning 8am)



Future Work

- Algorithm optimization
- Combination with additional forecasting methods (intra-day)
- Day ahead forecasting system

Acknowledgements

This work was made possible by the support of SDG&E and Green Power Labs, Inc during a 2012 summer internship. The authors gratefully acknowledge the support of Dallas Cormier and Hugo T.C Pedro