SKY COVER: SHINING LIGHT ON A GLOOMY PROBLEM

Jordan J. Gerth *
University of Wisconsin, Madison, Wisconsin

1. INTRODUCTION

This paper summarizes a methodology to improve the analysis and short-term forecasts of sky cover across the United States, and adjacent coastal areas, using geostationary satellite and in-situ surface station observations of cloud in the celestial dome. The methodology involves the use of linear and mixed integer optimization processes to minimize the mean absolute difference between multi-source sky cover observations and short-term numerical weather prediction model forecasts of cloud and moisture variables.

There are two predominant objectives in this study. The first is to produce a sky cover analysis that is representative of current conditions and suitable for use as validation. The second is to quantify the relationship between sky cover as purported by the analysis and related atmospheric quantities in a cloud-resolving weather forecast model.

2. BLENDED SKY COVER ANALYSIS

The first effort in this study is to create an hourly sky cover analysis assuming several properties about the sky cover quantity. The first proposed property is that sky cover is a temporal average over an entire hour. Second, satellite and in-situ surface observations of cloud are complementary in defining sky cover. The quantity is not exclusive to a single observing platform. Third, the range of the sky cover quantity is always between 0% and 100%. Given these, the goal of this step is to produce an operations-grade sky cover product for the field.

Effective cloud amount (ECA), the product of fractional cloud cover within the field of view (FOV) and cloud emissivity, is the most common method to assess sky cover from satellite observations. In comparison, the United States Office of the Federal Coordinator for Meteorology (OFCM) defines sky cover as "the amount of the celestial dome hidden by clouds and/or obscurations" ("Federal Meteorological Handbook No. 1, Surface Weather Observations and Reports" 2005).

2.1 *Sources of sky observations*

There are three primary sources of sky observations: space-based imagers (i.e., radiometers onboard low earth-orbiting and geostationary satellites); stationary, surface-based instrumentation (e.g., ceilometers); and trained human observers, typically with aid of instrumentation.

---

* *Corresponding author address:* Jordan J. Gerth, Cooperative Institute for Meteorological Satellite Studies, University of Wisconsin, 1225 W. Dayton St., Madison, WI 53706.

There are geometry and/or quality issues with each observation type. Satellites observe the atmosphere from the top, such that high clouds obscure low clouds. Near surface clouds and clouds smaller than the FOV may not be properly represented. In addition, satellite observations are instantaneous, and do not meet the aforementioned temporal average property of sky cover.

In comparison, ceilometers at surface stations fail to detect/report high cloud (over 12 kft) and do not observe the celestial dome. The human observations require estimation and are not as precise. Sky conditions are reported as one of five coverage modes in coded surface observation reports (clear, few, scattered, broken, and overcast).

The satellite sky cover product is based on ECA obtained from the operational GOES imagers (East and West). It is corrected when high (low effective emissivity) cloud obscures underlying low cloud. As part of the product, high effective emissivity cloud is enhanced. Every scan is spatially averaged to produce an ad hoc celestial dome (pixel-centered 11 by 11 box), and all scans within a one-hour window are temporally averaged to produce the final product.

2.2 *Creating the blended analysis*

The blended sky cover analysis leverages the advantageous properties of satellite and in-situ observations in a decision structure that, for a given point, selects one of the two observations, or averages them, based on the scenario. If the surface station observation reports clear (less than 5% celestial dome coverage), the satellite sky cover product value is used. If the surface station observation reports some cloud (5% or better coverage of the celestial dome), the surface observation is used when the value is greater than that from the satellite. In other situations where both observations are available, a weighted average is performed.

The advantages of the blended analysis creation process are that it: evaluates all available data and leverages strengths of multiple observational sources; preserves cloud gradients; adequately resolves diurnal cumulus fields (not missing, not bimodal); and is a temporally continuous and spatially contiguous field (available hourly over the contiguous United States).

3. OPTIMIZING SKY COVER FORECASTS

The second objective is to create an optimal sky cover forecast based on two assumptions. First, the relationship between sky cover and numerical weather prediction cloud and moisture variables is assumed to be approximately linear. The use of a Maclaurin series supports this assumption. Second, the model variables are assumed to adequately represent the atmosphere at the initial time and at times in the future. The goal is to

produce an operations-grade short-term sky cover forecast for the field.

There are two optimizations performed in this project. Each optimization model developed here uses points from gridded input fields, matching points from a gridded truth field, and an objective function. In both optimizations, the objective function minimizes the mean absolute error between the affine expression of the adjustable input fields and the truth field.

Optimization models are designed to contain an objective using a free variable, subject to constraints; terms, matching variables and components; and constraints involving terms. The role of the constraints is to enforce physical relationships. Finally, an optimizer executes the formulated model in order to find a solution—the best objective value without violating any constraints. There are both open-source and commercial optimizers available to solve the defined model, though the commercial optimizers generally arrive at a solution more quickly.

3.1 *Optimizing blended sky cover analysis toward NDFD sky cover one-hour forecast*

The first optimization model compares the NDFD one-hour forecast, as the input field, with the blended sky cover analysis, as the truth field. The National Weather Service (NWS) National Digital Forecast Database (NDFD) contains an operational sky cover forecast field (Glahn and Ruth 2003; "National Weather Service Instruction 10-201, National Digital Forecast Database and Local Database Description and Specifications" 2012) that is one of the only verifying fields of sky cover routinely available for the United States. However, there are notable issues with the national one-hour forecast that include: clear areas with a non-zero cloud cover forecast, vastly different cloud classifications for similar cloud scenes, the lack of spatial continuity between some NWS forecast office areas of responsibility, and temporal trends of sky cover that do not match observations. While a NWS meteorologist maintains the NDFD sky cover grid, the initial input for the grid typically originates from numerical weather prediction model guidance fields. In some cases, a formulation based on numerical model relative humidity is used.

This initial optimization model calculates a unique coefficient for blended sky cover analysis values that fall within prescribed ranges. The optimal coefficient within each value range is then applied to the proper blended sky cover analysis value to produce the adjusted sky cover analysis. Maxima and minima remain in the same location, but the magnitude of gradients can change depending on the magnitude of the coefficient for adjacent sky cover value ranges. Coefficients are constrained to be positive.

3.2 *Optimizing affine expression of HRRR fields toward adjusted sky cover analysis*

In the second optimization model, the truth field is the adjusted sky cover analysis acquired following the results from the previous optimization model. This second optimization model uses a host of input fields that are available on pressure levels from the High-Resolution Rapid Refresh (HRRR) cloud-resolving forecast model (Benjamin 2013): relative humidity, cloud water mixing ratio, cloud ice mixing ratio, rain water mixing ratio, snow mixing ratio, and absolute vorticity, partitioned into positive and negative components. The pressure levels incorporated into the model for the aforementioned fields, except absolute vorticity (200 hPa only), are: 200 hPa, 300 hPa, 500 hPa, 700 hPa, 800 hPa, 850 hPa, 900 hPa, 950 hPa, and 1000 hPa.

The variables incorporated into the optimization model include: one coefficient each for 200 hPa positive and negative absolute vorticity ($m_{200}AV_{200}$), a coefficient for relative humidity quantities ($m_x RH_x$), a threshold for applying a coefficient to the 1000 hPa relative humidity field ($m_{1000}RH_{1000}$ if $RH_{1000} > RH_T$), and a coefficient and scalar for points with non-zero mixing ratio quantities ($m_y MR_y + b_y$ if $MR_y > 0$, otherwise $m_y MR_y$).

Constraints in this optimization model maintain model variables within a range of acceptable values (generally between 0% and 100%). Constraints are implemented to limit the extent of relative humidity and absolute vorticity correlating to cloud, enforce thresholds in mixed integer implementations, and maintain mean values and the approximate value distribution of the output field to similarly match the truth field. Constraints also guide the optimizer to produce a meaningful result. For example, constraints can help to adjust the solution away from scalar adjustments and toward coefficient adjustments in order to maintain spatial gradients in the input fields, where the numerical prediction model is assumed to have skill.

4. RESULTS

Results for this study were obtained through running the aforementioned optimization process hourly from 21 September 2013 through 1 November 2013 over the contiguous United States and adjacent areas. Over this period, the most frequently selected field in the solved affine relationship within the second optimization model was 950 hPa cloud water mixing ratio. In general, cloud water mixing ratio from one or more levels in the lower troposphere was frequently correlated with sky cover. Higher in the troposphere, there was less reliance on cloud water mixing ratio and more reliance on relative humidity. Snow mixing ratio and rain mixing ratio were not commonly included in solutions. This indicates limited model skill on the placement of precipitation processes in the HRRR, at least during the test period.

5. SUMMARY

As a result of this work, an hourly blended sky cover analysis is now produced routinely using multiple sources of sky cover observations. The adjusted sky cover analysis is used to build better numerical weather

prediction model output of sky cover, using a mixed integer optimization methodology. The optimized numerical model output compared to the NDFD one-hour forecast consistently has less mean absolute error than the original/current output. Future work will focus on interacting with the NWS to produce a sky cover analysis of record and validate short-term numerical model forecasts of sky cover.

6. ADDITIONAL READING

A comprehensive review of the approach outlined in this extended abstract is available in the form of a Ph.D. dissertation. A copy of the dissertation, "Sky Cover", is available online at http://go.wisc.edu/85a0k4.

7. REFERENCES

Benjamin, Stan. 2013. "Data Assimilation and Model Updates in the 2013 Rapid Refresh (RAP) and High-Resolution Rapid Refresh (HRRR) Analysis and Forecast Systems."
http://ruc.noaa.gov/pdf/NCEP_HRRR_RAPv2_6jun2013 -Benj-noglob.pdf.

"Federal Meteorological Handbook No. 1, Surface Weather Observations and Reports." 2005. Washington, D.C. http://www.ofcm.gov/fmh-1/fmh1.htm.

Glahn, Harry R, and David P Ruth. 2003. "The New Digital Forecast Database of the National Weather Service." *Bulletin of the American Meteorological Society* 84 (2) (February 1): 195–201. doi:10.1175/BAMS-84-2-195. http://dx.doi.org/10.1175/BAMS-84-2-195.

"National Weather Service Instruction 10-201, National Digital Forecast Database and Local Database Description and Specifications." 2012. http://www.nws.noaa.gov/directives/sym/pd01002001cur r.pdf.