Towards Improving Visibility Forecasts: A Statistical Approach

Kuei-Pao Lu Lin, John Chun-Han
Weather Center Weather Wing, CAF ROC

ABSTRACT

Visibility is a crucial variable influencing airport operations and general aviation safety. In this study, a statistical regression analysis was used to investigate the relationship between the CAF airport-based observations of visibility and the relevant atmospheric physical variables.

The preliminary results reveal that four factors can affect visibility: relative humidity, PM concentration, month of year, and presence of rainfall. The relationship between visibility and these four factors more or less depend on station location, but in general, RH and PM are the variables most related to visibility. With increases in PM and RH, visibility decreases. Moreover, visibility is more sensitive to changes in PM2.5 concentrations than to PM10.

1. Motivation

The objective of this research project is to improve forecasts of visibility at airports over the Taiwan area.

Visibility is a crucial variable influencing airport operations and general aviation safety; hence the need for improving the Air Force's capability in its forecast. While visibility prediction has been recognized as a difficult research task, this work is motivated by the perspective that the current approach to visibility forecasts is ad-hoc and lacking in physical grounding, leaving room for considerable improvements.

2. Methodology

This study makes use of the large observational database gathered by the Air Force at over 10 airports from 2002 to 2007. Every hour, the Air Force measures visibility as well as meteorological variables such as temperature, dew point, rainfall, windspeed, and wind direction. Visibility observations were reported at discrete levels, with a maximum value set to 9999 m. This dataset is combined with hourly measurements of particulate matter (PM) concentrations (in g/m^3) carried out by Taiwan's Environmental Protection Agency (環保署). Measurements of both PM2.5 and PM10 were available, where PM2.5 and PM10 refer to concentrations of PM below diameters of 2.5 μm and 10 μm, respectively. Since PM measurements were not coincident with those from the Air Force, we had to merge the two datasets by simply choosing the PM measurement site closest to each airport. Distances between airport and PM measurement sites ranged from 1 km to 36 km, resulting in inevitable uncertainties due to spatial displacement in visibility and PM.function).

Relationships between visibility and variables were determined through a linear, multiple regression method. The initial statistical model is as follows:

$$Visib = M_i + B_1*[100-RH] + B_2*[PM]+ B_3*[Rain] \quad \text{(Eq. 1)}$$

where Visib indicates visibility, M_i denotes a "month" factor with i ranging from 1~12 indicating the 12 months of the year, RH is relative humidity, PM is the particulate matter concentration, and Rain is a factor indicating presence of rainfall. (100 – RH) can be regarded as a measure of the deviation from saturation—i.e., a measure of the atmosphere's "dryness". Windspeed and wind direction were initially included within the regression model, but due to the fact that these variables were found not to be statistically significant at the 5% level, they were dropped from the model. We also omitted data from the months of May~September due to the significant influence of typhoons that may be hard to incorporate within a statistical model.

Values of M_i , B_1, B_2, and B_3 were established through a least-squares method that minimizes the squared differences between the observed and calculated visibilities. We carried out the multiple regression calculation using the function "lm" within "R", an open source, data analysis software (downloadable from http://www.r-project.org).

3. Results

Results from the multiple regression are shown in Table 1. A rough assessment of the statistical model's performance can be arrived by using the R^2 and residual standard error. Regressions using PM2.5 rather than PM10 exhibit higher R^2 and lower residual standard error, implying that PM2.5 is more closely related to visibility. Hence all of the regression coefficients and subsequent discussions derive from adopting PM2.5 as the PM ncentration.

R^2 range from 0.43 to 0.62, suggesting that 43% to 62% of the observed variance in visibility can be accounted for by the regression model. The residual standard error is on the order of ~1000 m to ~2000 m, indicating that the visibility predicted by the regression deviates from measurements by this amount. We expect part of this deviation to be due to the significant uncertainties in visibility observations.

Airport	R2 (PM10)	R2 (PM2.5)	Residerr (PM10)	Residerr (PM2.5)	month01	month02	month03	month04	month10	month11	month12	RH.unsat [m/%]	PM2.5 [m/(micro-g/m³)]	rainTRUE
RCPO	0.56	0.62	1758	1623	7618	7093	7759	7933	8329	8356	8128	108.3	-63.26	-1905
RCGM	0.54	0.59	1791	1705	7826	7624	7993	8336	8838	8815	8488	94.5	-59.17	-1866
RCMQ	0.44	0.48	2264	2174	6383	5696	6388	6340	7620	7639	7165	124	-55.32	-1525
RCKU	0.47	0.56	1980	1786	7531	7034	7388	8073	8142	8225	8156	111.9	-52.44	-1173
RCNN	0.44	0.57	1817	1592	7157	7299	7439	7748	8033	7888	7632	121.1	-59.15	-1280
RCAY	0.4	0.54	1922	1698	7035	7200	6961	7350	7467	7263	7540	114.4	-47.26	-1156
RCSQ	0.42	0.53	1735	1550	5904	6302	6225	6298	6338	6171	6216	104.9	-39.93	-1348
RCDC	0.42	0.53	1733	1552	5855	6245	6177	6308	6340	6127	6204	104.8	-39.94	-1363
RCQC	0.4	0.49	1701	1591	6734	7191	7624	7971	7691	7740	7518	116.9	-66.01	-787.2
RCYU	0.41	0.47	1420	1222	9454	9593	9411	9237	9418	9479	9503	43.32	-42.38	-1999
RCCS	0.41	0.47	1461	1218	9475	9627	9435	9193	9363	9440	9508	45.24	-41.14	-1965
RCZN	0.36	0.43	1106	1060	9372	9302	9214	9123	9158	9379	9458	38.63	-54.51	-1771

Table 1: Multiple regression results for different airports, based on the statistical model in Eq. 1. Values of M_i for June, July, and August are omitted due to the significant influence of typhoons. For comparison the R^2 and residual standard error for PM10 are also shown.

The values of regression coefficients shown in Table 1 provide the "leverage" different physical variables have on visibility. For instance, the presence of rainfall can decrease visibility by as much as 2000 m. Given a maximum visibility of 9999 m, this translates into a large, 20% decrease in visibility, on average, when rainfall occurs.

The coefficient B_2 in front of PM is negative: in other words, visibility declines with increasing PM concentrations. This is because aerosols comprising particulate matter scatter + absorb radiation as well as serving as cloud condensation nuclei necessary for fog formation.

The coefficients to (100 – RH) are positive, suggesting that as the atmosphere becomes drier, visibility increases. This is likely because fog formation requires saturation conditions for water vapor to condense to the liquid phase.

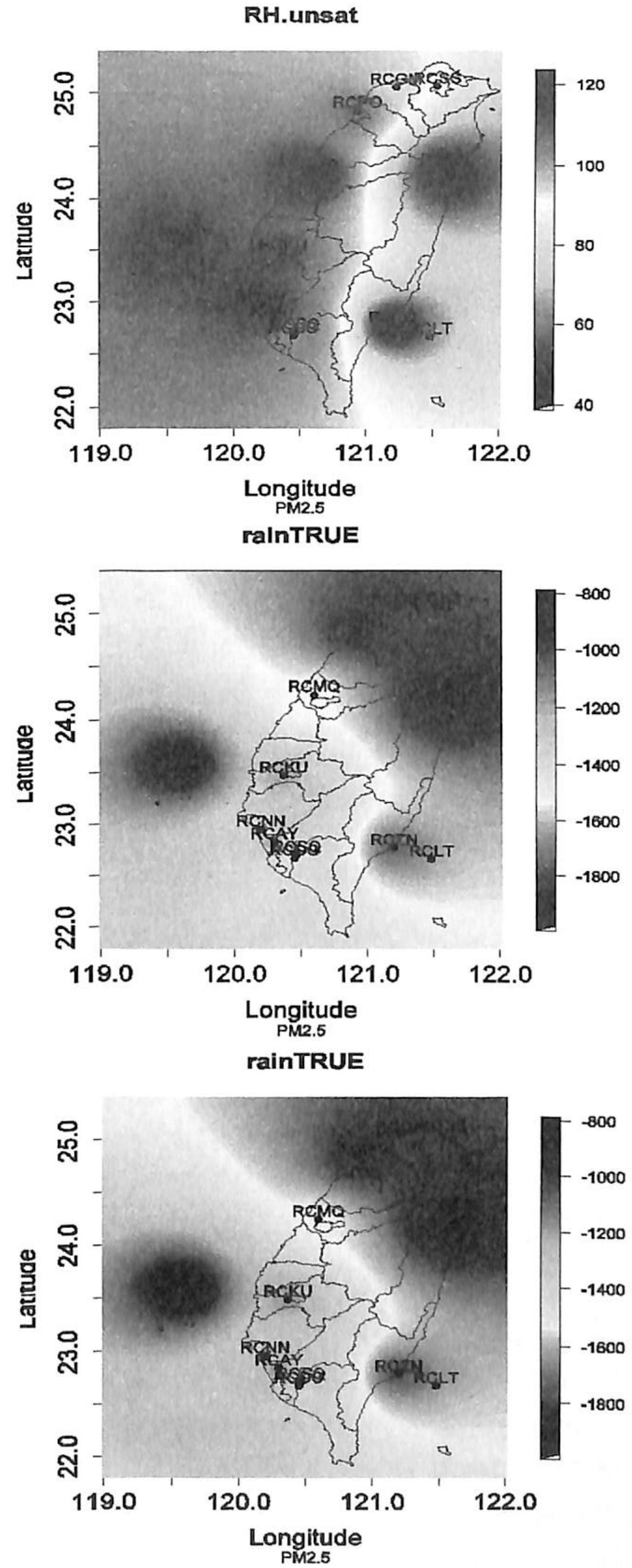

Fig. 1. Spatial patterns of regression coefficients B_1, B_2, and B_3 (from top to bottom). The values at the airports are interpolated to construct a continuous map.

Spatial maps of the regression coefficients are shown in Fig. 1. Distinct spatial patterns can be seen in B_1, B_2, and B_3 . The impacts of RH is clearly more pronounced at airports along the west coast of Taiwan. The response to PM exhibits a more complicated pattern, showing lower sentivities in southern Taiwan and at Hualien, reflecting the fact that other physical mechanisms play more important roles in controlling visibility. The impact of rain exhibits a north-south rather than a east-west gradient, with the rainfall-induced decrease in visibility being larger in northern Taiwan.

4. Summary/Conclusions

In summary, a statistical, regression analysis of airport-based observations of visibility and relevant atmospheric physical variables reveals the following:

- Wind direction and wind speed were found to be poor predictors of visibility.
- Visibility is related to RH, PM concentration, month of year, and presence of rainfall.
- The relationships between the aforementioned variables depend on station location, but in general, RH and PM are the variables most related to visibility.
- With increases in PM and RH, visibility decreases.
- Visibility is more sensitive to changes in PM2.5 concentrations than to PM10.
- Sensitivity to RH is higher in western Taiwan.
- Sensitivity to PM2.5 exhibits a complicated spatial pattern.
- The impact of rainfall in decreasing visibility is larger in northern Taiwan.

5. Future Directions

One future development in the

proposed work is to assess the statistical model by applying it in an actual operational, forecasting setting. This would allow one to examine the regression model's performance in the actual application it was envisioned for. Larger uncertainties are expected, since errors are expected in the predictor variables. For instance, RH and rainfall need to come from NWP model output, rather than actual measured values as is currently the case. The statistical, regression-based approach can also be compared against previous attempts based more on a conceptual approach.

Secondly, another key future direction of this work is to incorporate information provided by atmospheric models. Recent developments in Lagrangian modeling enable the transport of air parcels to be simulated backward in time while incorporating turbulent dispersion. Therefore, by choosing target airports as starting locations and simulating trajectories of air parcels backward in time, air parcel locations at different times prior to the starting time can be known. The Stochastic Time-Inverted Lagrangian Transport (STILT) model carries out such simulations. STILT is an "off-line" tool, which means that it does not solve the equations of motion to derive windfields but ingest windfields produced from another (Eulerian) atmospheric model. High-resolution windfields will be used to drive STILT; suitable products will be derived from mesoscale atmospheric models such as MM5 or WRF.

The STILT atmospheric model yields air parcel trajectories arriving at each airport. This information elucidates the different flow regimes affecting airport weather. Because it is highly probable that specific flow regimes are associated with low visibility events, by identifying the association between flow regimes (as elucidated by STILT) this opens that possibility of improved visibility forecasts based on identifying these regimes by running STILT in a forecast mode.