Fog Prediction Errors Evaluated for Multiple Physical Parameterization Schemes in the AFWA Mesoscale Ensemble
AFWA Mesoscale Ensemble (MEPS)

- 10-member ARW-WRF ensemble with 3 nests; inner-most has 4-km horizontal resolution, 42 Eta levels, no cumulus parameterization
- Each member gets ICs, BCs from different member of NCEP’s Global Ensemble Forecast System (GEFS)
- Water vapor field is initialized, other water phases are not
- 20-h runs initialized at 00Z every 3-4 days from Nov 2008 to Feb 2009 ➔ 29 total runs
- Configuration based on work by Hacker et al (2011) to obtain “most skillful ensemble with least degree of complexity”

Verification sites (elevation in m)
AFWA Mesoscale Ensemble (MEPS)

- Model perturbations obtained via unique physics suite, in addition to unique lower boundary properties

<table>
<thead>
<tr>
<th>Member</th>
<th>Microphysics</th>
<th>PBL</th>
<th>Shortwave</th>
<th>Longwave</th>
<th>Land Surface</th>
<th>Cumulus (none on inner-most nest)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Kessler</td>
<td>YSU</td>
<td>Dudhia</td>
<td>RRTM</td>
<td>Thermal</td>
<td>KF</td>
</tr>
<tr>
<td>5</td>
<td>WSM6</td>
<td>MYJ</td>
<td>CAM</td>
<td>RRTM</td>
<td>Thermal</td>
<td>KF</td>
</tr>
<tr>
<td>7</td>
<td>Kessler</td>
<td>MYJ</td>
<td>Dudhia</td>
<td>CAM</td>
<td>Noah</td>
<td>BM</td>
</tr>
<tr>
<td>8</td>
<td>Lin</td>
<td>MYJ</td>
<td>CAM</td>
<td>CAM</td>
<td>Noah</td>
<td>Grell</td>
</tr>
<tr>
<td>10</td>
<td>WSM5</td>
<td>YSU</td>
<td>Dudhia</td>
<td>RRTM</td>
<td>Noah</td>
<td>KF</td>
</tr>
<tr>
<td>11</td>
<td>WSM5</td>
<td>MYJ</td>
<td>Dudhia</td>
<td>RRTM</td>
<td>Noah</td>
<td>Grell</td>
</tr>
<tr>
<td>15</td>
<td>Lin</td>
<td>YSU</td>
<td>Dudhia</td>
<td>CAM</td>
<td>RUC</td>
<td>BM</td>
</tr>
<tr>
<td>16</td>
<td>Eta</td>
<td>MYJ</td>
<td>Dudhia</td>
<td>RRTM</td>
<td>RUC</td>
<td>KF</td>
</tr>
<tr>
<td>17</td>
<td>Eta</td>
<td>YSU</td>
<td>CAM</td>
<td>RRTM</td>
<td>RUC</td>
<td>BM</td>
</tr>
<tr>
<td>19</td>
<td>Thompson</td>
<td>MYJ</td>
<td>CAM</td>
<td>CAM</td>
<td>RUC</td>
<td>Grell</td>
</tr>
</tbody>
</table>

Physics suite used by each member
Extracting visibility from WRF

- Relationship between model output ($q_c$, RH, etc.) and visibility cannot be explicitly modeled $\Rightarrow$ need a visibility parameterization
- Desirable to use only critical variables (as determined by first principles) rather than a customized, highly-statistical approach

Stoelinga and Warner, 1999

$$\beta_e = 144.7(q_c)^{0.88}$$

Gultepe, 2006

$$\beta_e = 178.6(q_c)^{0.96}$$

(Vis_{day} related to extinction coef ($\beta_e$) as

$$Vis_{day} = \frac{3.0}{\beta_e}$$

Vis_{night} typically 2-3 times higher)

- Droplet number concentration ($N$) not predicted by microphysics schemes
Layer 1 cloud water RPSS (thresholds of 7, 5, 3, 1 mi)

- After period of spin up, predictions in coastal and mountain regions demonstrate skill relative to persistence.
- Valley region predictions generally not skillful, temporarily drop after sunrise (17-19 h).
- Skill generally increases with forecast hour.
- Parametric visibility parameterization adds no skill ➔ primary source of error is from NWP predictions.
Member climatologies of layer 1 cloud water

- Observations
- Predictions

“light fog” (1~7 mi)

- Predictions highly bimodal in every member
- Excessive zero-$q_c$ predictions, deficit in light fog predictions

Incidence of light fog

Observations: 0.196
Predictions: 0.005
Member climatologies of layer 1 RH

- Large negative bias in layer 1 RH in every member
  - average member bias
    - coastal: -0.182
    - valley: -0.069
    - mountain: -0.014

- Additional $q_c$ error from members restricting fog to very high RH range compared to obs

- Stochastic predictions negatively biased and underdispersive

Distribution of Predictions and Observations
Verification Rank Histograms
Layer 1 temperature

- Warm biases highest overnight, and in coastal region
- Coastal predictions have little diurnal variation, high error variances overnight ➔ seemingly lower predictability
- Post-sunrise warming inadequate in both regions, with larger error variances in valley (observed warming is less consistent)
Layer 1 water vapor

- Near-neutral overnight biases
- $q_v$ error variances lower than temperature error variances in coastal region, comparable in valley region
- Diurnal changes well-predicted
- Insufficient post-sunrise moistening has minor impact on RH compared to temperature biases
Layer 1 and 2-m temperature
Layer 1 and 2-m water vapor

Layer 1

Bias

Rank Histogram

2 Meters

Bias

Rank Histogram

Coastal

Valley

Mountain
Layer 1 and 2-m RH

Bias

Rank Histogram

Bias

Rank Histogram

Coastal

Valley

Mountain
Valley fog
dissipation timing

Valley region post-sunrise count of fog predictions for cases when fog correctly forecast at 14 h (0600 LT)

- Number of cases in verification ranges from 2 (mbr 10) to 16 (mbr 15)
- Individual members exhibit biases in dissipation rate, but no clear systematic bias in this conditional sample
- Poor post-sunrise skill due to cases not shown: high false alarm rate (all members >0.75), low probabilities of detection (<0.30), and improving skill of persistence forecast
Summary and discussion

- Layer 1, $q_c$ predictions are highly bimodal, with virtually no values corresponding to light fog
  - Except in mountains, warm biases minimize $q_c$ production
  - Visibility parameterization error is inconsequential unless incidence of light fog predictions is increased
  - Due to positive resolution, ensemble fog predictions still outperform persistence in coastal and mountain regions after 9 h
  - Post-sunrise skill generally worse, but conditional results in valley region suggest promise
- At 2 meters, less RH bias in coastal and Central Valley regions, but large dry bias in mountains due to cold bias of up to 6 K
  - Error variances at least as good as layer 1 predictions, with better ensemble dispersion
### Member-specific behavior

<table>
<thead>
<tr>
<th>Member</th>
<th>Microphysics</th>
<th>PBL</th>
<th>Shortwave</th>
<th>Longwave</th>
<th>Land Surface</th>
<th>Cumulus (none on inner-most nest)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Kessler</td>
<td>YSU</td>
<td>Dudhia</td>
<td>RRTM</td>
<td>Thermal</td>
<td>KF</td>
</tr>
<tr>
<td>5</td>
<td>WSM6</td>
<td>MYJ</td>
<td>CAM</td>
<td>RRTM</td>
<td>Thermal</td>
<td>KF</td>
</tr>
<tr>
<td>7</td>
<td>Kessler</td>
<td>MYJ</td>
<td>Dudhia</td>
<td>CAM</td>
<td>Noah</td>
<td>BM</td>
</tr>
<tr>
<td>8</td>
<td>Lin</td>
<td>MYJ</td>
<td>CAM</td>
<td>CAM</td>
<td>Noah</td>
<td>Grell</td>
</tr>
<tr>
<td>10</td>
<td>WSM5</td>
<td>YSU</td>
<td>Dudhia</td>
<td>RRTM</td>
<td>Noah</td>
<td>KF</td>
</tr>
<tr>
<td>11</td>
<td>WSM5</td>
<td>MYJ</td>
<td>Dudhia</td>
<td>RRTM</td>
<td>Noah</td>
<td>Grell</td>
</tr>
<tr>
<td>15</td>
<td>Lin</td>
<td>YSU</td>
<td>Dudhia</td>
<td>CAM</td>
<td>RUC</td>
<td>BM</td>
</tr>
<tr>
<td>16</td>
<td>Eta</td>
<td>MYJ</td>
<td>Dudhia</td>
<td>RRTM</td>
<td>RUC</td>
<td>KF</td>
</tr>
<tr>
<td>17</td>
<td>Eta</td>
<td>YSU</td>
<td>CAM</td>
<td>RRTM</td>
<td>RUC</td>
<td>BM</td>
</tr>
<tr>
<td>19</td>
<td>Thompson</td>
<td>MYJ</td>
<td>CAM</td>
<td>CAM</td>
<td>RUC</td>
<td>Grell</td>
</tr>
</tbody>
</table>

#### Accuracy Skill Score

- **Coastal**
- **Valley**
- **Mountain**

![Graphs showing accuracy skill score over forecast hours for different members.](image)
Way forward

- Introduce *gentle* statistical element to make upward adjustments to zero and near-zero $q_c$ predictions
  
  - Layer 1 virtual temperature predictions in valley/mountain region
  
  - Layer 1 absolute moisture and d/dt virtual temperature predictions in coastal region
Backup Slides
**Statistical vs physical techniques in fog prediction**

- Statistical approaches to visibility-in-fog (VIF) prediction include formation of predictors based on observations, NWP output (statistically calibrated), or combination of both
  - Inherently calibrated, often outperform NWP data alone
    - Require long observational record
    - Require stable inputs (i.e., sensitive to NWP platform changes)

- Purely physical techniques place full confidence in NWP output, and convert to desired parameter using only first principles
  - No reliance on observations (only needs model data)
  - First principles valid everywhere
    - No calibration; at mercy of model error
    - First principles can be complex, entail many unknown quantities
Statistical vs physical techniques in fog prediction

- Many military operations are conducted far away from nearest airfield, where statistical calibration or climatological tools don’t exist (targeting, reconnaissance, search and rescue, etc.)

- Aim of this research is to strike appropriate balance between statistical and physical approaches for VIF prediction suitable for remote locations:
  - Use physical approach as baseline, introduce statistical components judiciously only where necessary
  - Gain insight into error characteristics, physical processes, future research needs
Defining visibility

- True visibility determined by complex process involving contrast between object and its background, contrast threshold of observer, and (during day) scattering of ambient sunlight.

- With automated instrumentation, visibility estimated using measured scattering coefficient ($\sigma_s$) within 1.5 ft$^3$ sample
  - During day, based on distance at which brightness difference between object and its background is 5% of the background brightness:
    \[
    \left| \frac{B_0' - B_b'}{B_b} \right| = 0.05 = \exp \left[ - \int_0^{x_{vis}} \sigma_s(x) \, dx \right]
    \]
  - …and if we assume homogeneity within the observing area:
    \[
    x_{vis} = Vis_{day} = \frac{-\ln(0.05)}{\sigma_s}
    \]

- Different algorithm used for nighttime visibility
  - Verification performed against $\sigma_s$ since it is the measured parameter.
**Individual member climatologies of** $q_c$

- In general, NWP correctly models bimodal nature of VIF
- All members have excess zero or near-zero $q_c$ forecasts at expense of intermediate $q_c$ forecasts (less so in mountain region)
- Error suggestive of deficiency in NWP model, not initial conditions
  - Climatology of NWP members avoids intermediate values, despite them being common in observed climatology
NWP forecast error vs visibility parameterization error

Scatter plot of observed $\sigma_s$ vs ensemble mean $\sigma_s$ using SW99 and G06 visibility parameterizations

Correct forecast of <1 mile
False alarms likely to benefit from ensemble spread, whereas most missed opportunities have $q_c \approx 0$, meaning ensemble spread is small.

When all members forecast $q_c \approx 0$, there can be virtually no visibility parameterization dispersion.

Many false alarms are close to verifying in intermediate range, whereas missed opportunities are not.
Figure 1. Nighttime visibility versus daytime visibility for the same extinction coefficient using standard assumptions used by ASOS regarding contrast of an object against its background (during day), luminous intensity (during night), and visual contrast threshold (Rasmussen et al, 1999).
\[
\text{visibility} = \frac{1.8615}{\sigma_s}
\]

\[
\frac{-5.7 + \ln(\text{visibility})}{(1.609 \cdot \text{visibility})} = \sigma_s
\]

| Day     | \[
\text{visibility} = \frac{1.8615}{\sigma_s}
\] |
|---------|--------------------------------------------------|
| Night   | \[
\frac{-5.7 + \ln(\text{visibility})}{(1.609 \cdot \text{visibility})} = \sigma_s
\] |

Table 1. Algorithms used for conversion of \( \sigma_s \) (in km\(^{-1}\)) to visibility (in miles) in all FAA ASOS systems. Given \( \sigma_s \), the nighttime algorithm requires an iterative process to solve for visibility (Belfort Instrument, 2005).
Figure 4. Distribution of $q_t$ forecasts (dark blue bars, top x-axis labels) for each member, and the distribution of observed $q_t$ (light green bars, middle x-axis labels). The first six hours of each case are excluded. The mean and variance of the $q_t$ distribution is indicated (in gm$^{-3}$).
Figure 5. Histogram of $q_x$ forecasts from members 16 (top) and 17 (bottom) for very small but non-zero values. The domain for this plot was created by partitioning bin 2 from Figure 4 into 12 equal sub-bins to allow closer examination. The first 25 non-zero values are also listed for each member.
Figure 6. Same as in Figure 4, but for coastal sites only. Note there is no discontinuity in the y-axis.
Figure 7. Same as in Figure 4, but for Central Valley sites only.
Figure 8. Same as in Figure 4, but for mountain sites only.
<table>
<thead>
<tr>
<th>Metric</th>
<th>Formula</th>
<th>Description</th>
<th>Best Score</th>
<th>Worst Score</th>
</tr>
</thead>
<tbody>
<tr>
<td>Percent Correct</td>
<td>( \frac{\text{correct forecasts}}{\text{total forecasts}} )</td>
<td>Summarizes overall performance with no bias or skill information.</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>Skill Score (relative to persistence)</td>
<td>( \frac{\text{correct forecasts} - \text{persistencecorrect forecasts}}{\text{total forecasts} - \text{persistencecorrect forecasts}} )</td>
<td>Measures overall skill. Value of 0 indicates forecast is no better or worse than persistence forecast.</td>
<td>1</td>
<td>-(\infty)</td>
</tr>
<tr>
<td>Bias</td>
<td>( \frac{\text{total&quot;yes&quot; forecasts}}{\text{total&quot;yes&quot; observations}} )</td>
<td>Reveals whether predictions, on average, are too ambitious or too conservative in forecasting event.</td>
<td>1</td>
<td>Overforecast: +(\infty) Underforecast: 0</td>
</tr>
<tr>
<td>False Alarm Ratio</td>
<td>( \frac{\text{incorrect&quot;yes&quot; forecasts}}{\text{total&quot;yes&quot; forecasts}} )</td>
<td>Answers question “when event is forecast, at what rate does it occur?”</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>Probability of Detection (each member)</td>
<td>( \frac{\text{correct&quot;yes&quot; forecasts}}{\text{total&quot;yes&quot; observations}} )</td>
<td>Answers question “when event occurs, at what rate was it forecast?”</td>
<td>1</td>
<td>0</td>
</tr>
</tbody>
</table>

Table 2. Description of metrics used for binary verification of each member. An "event" refers to an observed \( \sigma_k > 0.29 \text{ km}^{-1} \) due to fog. A “yes” forecast refers to a NWP forecast of \( q_r \geq .00085 \text{ gm}^{-3} \) in the lowest NWP model layer.
Figure 9. Binary verification of inferred presence of cloud water for all sites: a) percent correct, b) skill score relative to persistence, c) bias, d) false alarm ratio, e) probability of detection.
Figure 10. Same as in Figure 9, but for coastal sites only.
Figure 11. Same as in Figure 9, but for Central Valley sites only.
Figure 12. Same as in Figure 9, but for mountain sites only.
<table>
<thead>
<tr>
<th>Metric</th>
<th>Formula</th>
<th>Description</th>
<th>Best Score</th>
<th>Worst Score</th>
</tr>
</thead>
<tbody>
<tr>
<td>Reliability (ensemble suite)</td>
<td>$\frac{1}{M} \sum_{i=1}^{I} N_i \left( \left( p_{i}^* \right) - \bar{q}_i \right)^2$</td>
<td>Measures how well a given forecast probability matches the observed frequency of occurrence</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>Resolution</td>
<td>$\frac{1}{M} \sum_{i=1}^{I} N_i \left( \bar{q}_i - \bar{q} \right)^2$</td>
<td>Measures degree to which ensemble, through its probability forecasts, can parse data into subsamples having frequency of occurrence different from overall climatological frequency</td>
<td>Uncertainty score</td>
<td>0</td>
</tr>
<tr>
<td>Uncertainty</td>
<td>$\bar{q} \left( 1 - \bar{q} \right)$</td>
<td>Does not depend on forecast, only on climatological frequency; indicates level of difficulty in obtaining resolution</td>
<td>N/A – but scores may range from 0 (event occurs 0% or 100% of time, so no resolution possible) to 0.25 (event occurs 50% of time, maximizing potential resolution score)</td>
<td></td>
</tr>
<tr>
<td>Brier Score</td>
<td>reliability – resolution + uncertainty</td>
<td>Combines reliability and resolution to summarize overall ensemble accuracy</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>Brier Skill Score (relative to persistence)</td>
<td>$1 - \frac{\text{Brier Score}}{\text{Persistence}}$</td>
<td>Measures overall stochastic skill of ensemble. Value of 0 indicates forecast is no better or worse than persistence forecast.</td>
<td>1</td>
<td>$-\infty$</td>
</tr>
</tbody>
</table>

$M$ = number of forecast/observation pairs  
$I$ = number of probability bins (11)  
$N$ = number of data pairs in bin $i$  
$p_{i}^*$ = binned forecast probability $(0.0, 0.1, \ldots, 1.0)$ for bin $i$  
$\bar{q}_i$ = observed relative frequency for bin $i$  
$\bar{q}$ = climatological frequency (total occurrences / total forecasts)  

Table 3. Description of metrics used for binary verification of stochastic forecast from ensemble suite.
Figure 13. Binary verification of ensemble suite for inferred presence of cloud water. The left column shows reliability, and the right column shows the resolution and uncertainty for a) all sites, b) coastal sites, c) Central Valley sites, and d) mountainous sites.
Figure 14. Brier Skill Score for the inferred presence of cloud water, using persistence as the reference forecast: a) all sites, b) coastal sites, c) Central Valley sites, and d) mountainous sites.
Visibility vs LWC for various droplet number concentrations (from Gultepe et al 2006)