On the Verification of Thunderstorm MetObjects During the 2015 Toronto Games

Dominique Brunet, David Sills, Norbert Driedger, Emma Hung and Janti Reid
Cloud Physics and Severe Weather Research Section, ECCC
July 28, 2016

WMO WWRP 4th International Symposium on Nowcasting and Very-short-range Forecasting (WSN16) , Hong Kong, China, 25-29 July 2016
Toronto 2015 Pan Am Games

- Large sporting event with over 6000 athletes from 41 Pan American countries.
- Main weather concern for organizers was lightning occurrence in outdoor events.
- Environment Canada provided site specific operational weather alerts.
- Experimental thunderstorm nowcasts were generated at a ‘Research Support Desk’ (RSD) operated by four science staff.
1. The Verification Experiment
2. Relaxation: Classical Scores
3. NWP-based MetObjects
4. MetObject-based Verification
The Verification Experiment

Goals:

1) Intercomparison of different NWP-based thunderstorm nowcasts/forecasts and of a human-generated nowcast/forecast

2) Development of verification methodology to improve such intercomparison

Previous work:

Verification Domain

- Pan Am: July 9\textsuperscript{th}-26\textsuperscript{th} - ParaPan Am: August 8\textsuperscript{th}-15\textsuperscript{th}
- Plus few extra days in July-August 2015
- Total: 31 days
Observations

• CLDN: Canadian Lightning Detection Network (Vaisala + ECCC)
• Thunderstorm occurrence: “If you can hear the thunder…”

July 25, 2015, 18-21Z (2-5 pm)

• For verification purposes, we define thunderstorm occurrence as within 25km radius from lightning flashes
NWP-based Thunderstorm Forecasts

All forecasts are valid for July 25, 2015, 18-21Z. Runtime: 12Z. Issue time: 15Z.

- **RDPS-Stat**
  - Statistical post-processing of regional deterministic forecast (RDPS).
  - Source: Bill Burrows (ECCC-Edmonton)

- **RDPS-Sci**
  - Calibrated post-processing of regional deterministic forecast (RDPS) based on latest thunderstorm initiation science.
  - Source: Neil Taylor (ECCC-Montréal)

- **REPS-TI**
  - Calibrated regional ensemble thunderstorm forecast.
  - Source: Ron Frénette (ECCC-Montréal)

Forecaster Generated MetObjects

- RSD: human forecasters use prototype software for generation of MetObjects nowcasts
- MetObjects were based on combination of NWP nowcasts, observations including climatology, and conceptual models.


MO-RSD: TS prob for non-severe
Issued at 15Z, valid 18-21Z

MetObjects: Vector representation (points, lines, areas) of meteorological concepts with attributes
Thunderstorm MetObjects: areas with severity (severe/non-severe) and probability attributes (chance/likely/certain).
1. The Verification Experiment
2. Relaxation: Classical Scores
3. NWP-based MetObjects
4. MetObject-based Verification
Classical Verification

Contingency table

<table>
<thead>
<tr>
<th></th>
<th>Observed</th>
<th>Not observed</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Forecasted</strong></td>
<td>H = #hits</td>
<td>F = #false alarms</td>
</tr>
<tr>
<td><strong>Not forecasted</strong></td>
<td>M = #misses</td>
<td>C = #correct negatives</td>
</tr>
</tbody>
</table>

Toronto Games 2015 Cases for Day 1 @ 18Z

CSI for chance/20% (non-severe)

BSS
Reliability Plot and ROC Curve

4-categories:
None (0-20%)
Chance (20-50%)
Likely (50-80%)
Certain (80-100%)

• MO-RSD and RDPS-Sci are over-forecasting
• MO-RSD and REPS-TI have best ROC scores
Problem: Forecasts are sensitive to the relaxation parameter.
Solution: Bring all forecasts to the same scale (as MetObjects) so that they are robust to the choice of relaxation parameter.
1. The Verification Experiment
2. Relaxation: Classical Scores
3. NWP-based MetObjects
4. MetObject-based Verification
NWP-based MetObjects

Goal: Find optimal smoothing and thresholds to transform NWP into MetObjects.

We use MO-RSD (human-generated) as the reference, so that all data have the same level of smoothness.

Decomposition of Brier’s score into calibration and refinement term:

$$\frac{1}{K} \sum_k (f_k - o_k)^2 = \sum_i w_i (y_i - p_i)^2 + \sum_i w_i p_i (1 - p_i)$$

1) Find optimal smoothing for Chance category: minimize refinement term.
2) Find optimal thresholding for Chance/Likely/Certain categories: minimize calibration term.
MetObject Extraction Example

Statistical post-processing (RDPS-Stat)

Thunderstorm initiation (RDPS-Sci)

Calibrated ensemble (REPS-TI)

MO-RSD (human-generated)
The relative order of the scores does not depend strongly on the relaxation parameter anymore. The scores are also more similar.
Results for Optimal Relaxation

Same optimal relaxation parameter to minimize bias: 40 km

CSI: chance category

BSS
1. The Verification Experiment
2. Relaxation: Classical Scores
3. NWP-based MetObjects
4. MetObject-based Verification
Double Penalty Problem

Which verification method better fits with our intuition of what is a good forecast?

They have all the same classical scores!
Distance Between Two MetObjects

Point-to-point distance for \( x, y \in X \):
\[
d(x, y) = \sqrt{\sum_i (x_i - y_i)^2}
\]

Point-to-set distance for \( x \in X, A \subset X \):
\[
D(x, A) = \min_{a \in A} d(x, a)
\]

Set-to-set distance (Hausdorff) for \( A, B \subset X \):
\[
H(A, B) = \max \left( \max_{a \in A} D(a, B), \max_{b \in B} D(b, A) \right)
\]

Set-to-set distance (generalized Hausdorff):
Replace min and/or max by power sums.

The geometry type of A and B can be areas, lines or cloud of points.

Distance-based Verification Example

CASE: July 25, 18-21Z

MO-RDPS-Sci
CHC: 559km
LKY: 829km
CRT: $\infty$

MO-RDPS-Stat
CHC: 511km
LKY: 296km
CRT: $\infty$

MO-REPS-TI
CHC: 440km
LKY: 208km
CRT: 298km

MO-RSD
CHC: 364km
LKY: 419km
CRT: 548km
Conclusions

• Compared three NWP-based forecasts with human-generated forecast (MetObject) for thunderstorm nowcast with 3 hours lead time.
• REPS-TI forecast does best according to the classical scores, but the results are sensitive to the choice of relaxation parameter.
• Extracting MetObjects from NWP-based forecasts allows a comparison at the same scale and with the same calibration: REPS-TI and human-generated forecast (MO-RSD) come out on top with no statistically significant difference.
• The MetObject approach fits naturally an object-based verification paradigm. As an example, we demonstrate the computation of distance between MetObjects.
Questions or Comments?

Thank You!

Dominique Brunet
Meteorological Research Division
Science & Technology Branch
Environment and Climate Change Canada
Dominique.Brunet@canada.ca
Caveats and Outlook

Caveats:
- The NWP-based MetObjects were computed in hindcast.
- The forecasters did not have access to the NWP-based MetObjects.
- Limited number of cases and of human forecasters.
- Trying to compare a 4-category forecast to continuous probabilistic forecasts.

Outlook:
- Hard to settle the human VS machine forecaster controversy because of all the caveats, but getting closer to a rigorous methodology.
- Verification methods could also be extended for nowcasts based on persistence and extrapolation.
- Distance between MetObjects only one step into a full object-based framework, but distance can also be used for grouping and matching.
- The Toronto Panam 2015 Dataset will be shared Open Access, so an intercomparison of lightning nowcast can be possible.

QUESTIONS/COMMENTS?
- Contact Dominique Brunet at Dominique.Brunet@canada.ca
Object-based verification

Group/match objects.
For each matched objects:
Features extraction:
  Centroid, angle, aspect ratio, size
Features/object comparison:
  Difference, log-ratio, distance
Summary or aggregation of pairwise scores