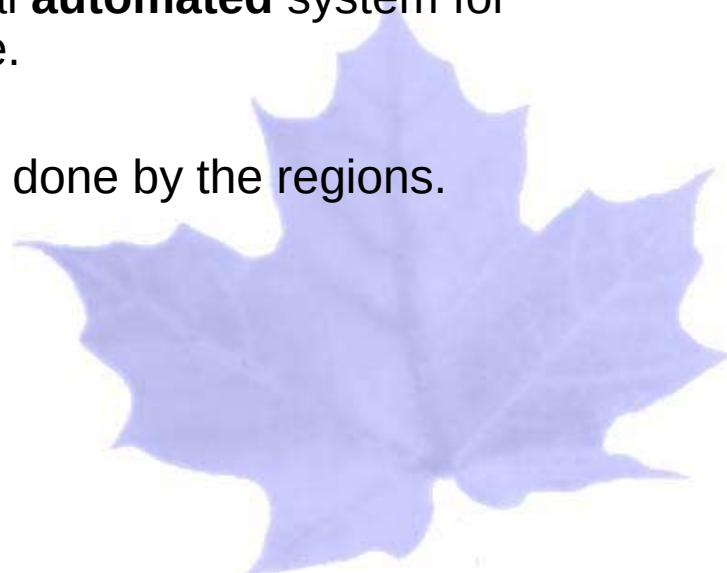
Verification of Canadian Public Weather Forecasts by an Automated System

International Verification Methods Workshop

Sep 15-17, 2004

Background

- In 2001, the Meteorological Service of Canada (MSC) embarked on the exercise of "**Living within Our Means – Focus for the Future."**
- The exercise identified the need for a national **automated** system for measuring MSC's public forecast programme.
- Verification of the public forecast programme done by the regions.



Canadian Public Forecasts

FPCN11 CWUL 090844
FORECASTS FOR WESTERN QUEBEC ISSUED BY ENVIRONMENT CANADA AT
05:00 AM
EDT THURSDAY 9 SEPTEMBER 2004 FOR TODAY AND FRIDAY. NEXT
FORECASTS
ISSUED AT 11:30 AM.

MONTREAL METROPOLITAIN-LAVAL.
HEAVY RAINFALL WARNING IN EFFECT.
TODAY..RAIN AT TIMES HEAVY. RISK OF A THUNDERSTORM THIS
AFTERNOON.
AMOUNTS EXPECTED OF 30 TO 40 MM. WINDS NORTHEAST 20 KM/H
INCREASING
TO 40 WITH GUSTS TO 60 THIS MORNING. HIGH 18.
TONIGHT..RAIN. AMOUNTS EXPECTED OF 15 TO 25 MM. WINDS NORTHEAST
40
KM/H WITH GUSTS TO 60 BECOMING LIGHT THIS EVENING. LOW 15.
FRIDAY..RAIN ENDING IN THE MORNING. CLEARING LATER ON. AMOUNTS
EXPECTED OF 5 MM. WINDS NORTHWEST 20 KM/H BECOMING LIGHT IN THE
AFTERNOON. HIGH 22.

.. .. ..

FPCN54 CWUL 090801
EXTENDED FORECASTS FOR SATURDAY SUNDAY AND MONDAY FOR WESTERN QUEBEC
ISSUED BY ENVIRONMENT CANADA AT 5.00 AM EDT THURSDAY 9 SEPTEMBER
2004.
THE NEXT SCHEDULED FORECAST WILL BE ISSUED AT 11.30 AM.

MONTREAL METROPOLITAIN-LAVAL.
SATURDAY..SUNNY. LOW 11. HIGH 25.
SUNDAY..A MIX OF SUN AND CLOUD. LOW 17. HIGH 25.
MONDAY..CLOUDY. 30 PERCENT CHANCE OF SHOWERS. LOW 13. HIGH 20.
NORMALS FOR THE PERIOD..LOW 10. HIGH 21.

.. .. ..

System Features

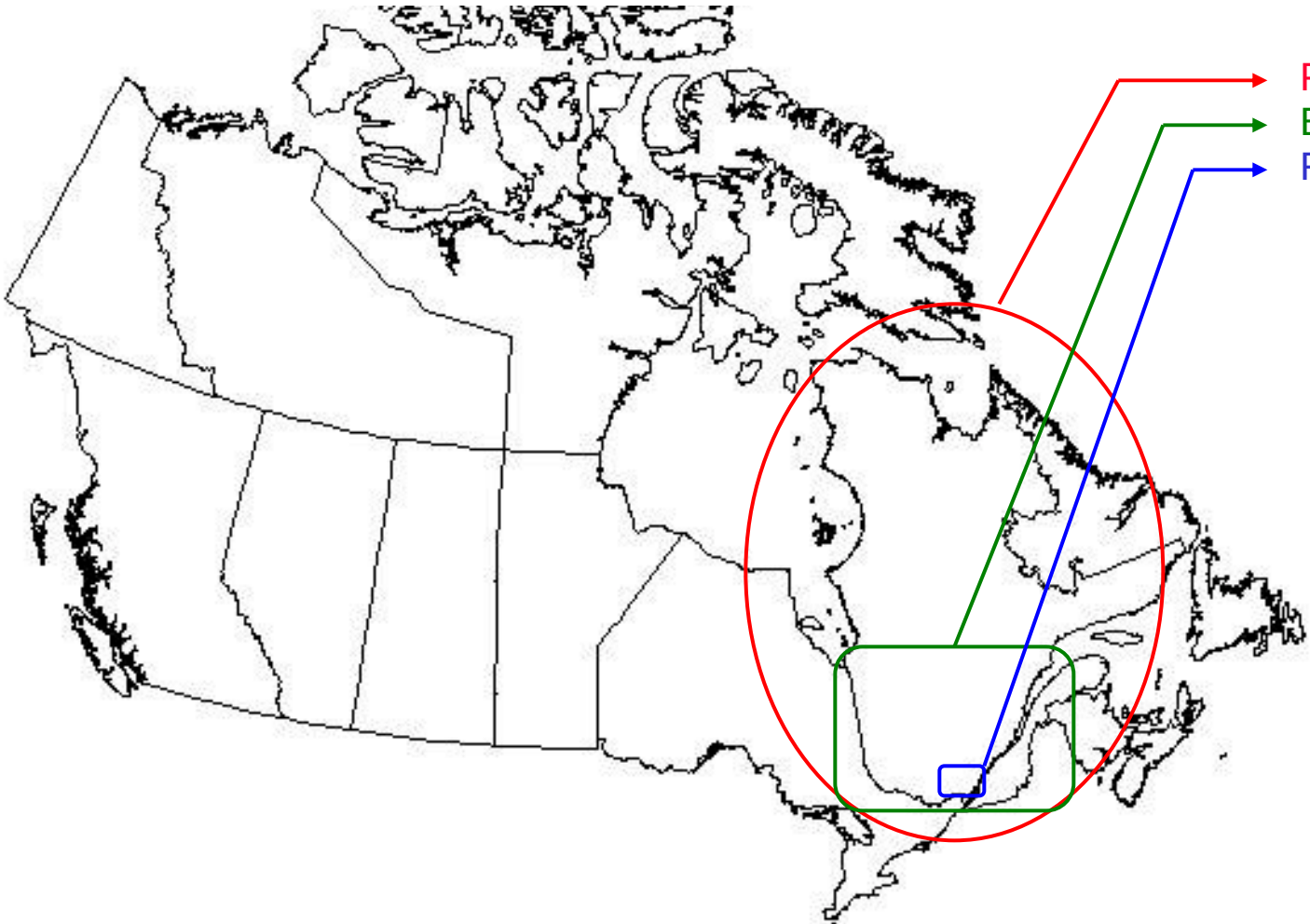
- Design of the system begin in late 2001
- System became operational in April 2004
- Web interface
- Monthly performance reports
- On demand reports for a selected set of elements
- Scores for internal MSC users only

Performance Scores

- Currently tracking scores for nine weather elements
  - Daily maximum temperature
  - Daily minimum temperature
  - 3-hourly temperature
  - 3-hourly probability of precipitation
  - 24-hour (06-06Z) precipitation amounts
  - 3-hourly wind direction
  - 3-hourly wind speed
  - 3-hourly sky opacity

- Scores
  - Bias
  - Root Mean Square Error
  - Percent Correct within Certain Threshold
  - Probability of Detection
  - False Alarm Ratio
  - Critical Success Index
  - Brier Score

Performance Scores



Forecast Data

- MSC forecasters use a tool called **Scribe** to generate text public forecasts.
- **“Scribe”** is similar to National Weather Service’s Interactive Forecast Preparation System IFPS.
- Forecast data for verification are based on Meteocode files from Scribe, not text version of public forecasts
- Using Meteocode, no parsing of text forecasts needed in our system.
- One to one relationship between Meteocode and text public forecast is assumed.

Forecast Data

```
version:     ( 3.00 );
entete: (FPCN71 CWUL EST5EDT regulier  2004 09 09 0900 00
         prochaine_prevision 2004 09 09 1530 30);

regions: (r71.1);
accum:  (8 37 pluie totale pres_de 50 70);
avert:  (7 34 avertissement nil pluie_abondante);
ciel:  (0 37 10 10 2) (37 42 7 7 4) (42 48 7 1 5)
       (48 53 0 0) (53 78 0 0) (78 93 4 4 8)
       (93 120 7 7 3) (120 126 7 1 4) (126 144 0 0);
indice_uv:  (17 18 6.5) (41 42 5.8);
temp:  (-3 0 point_intermediaire 17) (0 6 min 14)
       (6 12 point_intermediaire 16) (12 18 max 18)
       (18 24 point_intermediaire 15) (24 30 point_intermediaire 17)
       (30 36 min 15) (36 42 point_intermediaire 20)
       (42 45 max 22) (45 51 point_intermediaire 17)
       (51 54 min 11) (54 60 point_intermediaire 15)
       (60 66 point_intermediaire 23) (66 69 max 25)
       (69 75 point_intermediaire 19) (75 81 point_intermediaire 18)
       (81 84 min 17) (84 90 max 25)
       (90 96 point_intermediaire 19) (96 102 point_intermediaire 15)
       (102 108 min 13) (108 114 point_intermediaire 19)
       (114 117 max 20) (117 123 point_intermediaire 13)
       (123 129 min 11) (129 135 point_intermediaire 17)
       (135 141 max 21) (141 240 point_intermediaire 17);
.. .. ..
.. .. ..
```

```
FPCN11 CWUL 090844
FORECASTS FOR WESTERN QUEBEC ISSUED BY ENVIRONMENT CANADA AT 05:00 AM
EDT THURSDAY 9 SEPTEMBER 2004 FOR TODAY AND FRIDAY. NEXT FORECASTS
ISSUED AT 11:30 AM.

MONTREAL METROPOLITAIN-LAVAL.
HEAVY RAINFALL WARNING IN EFFECT.
TODAY..RAIN AT TIMES HEAVY. RISK OF A THUNDERSTORM THIS AFTERNOON.
AMOUNTS OF 30 TO 40 MM EXPECTED . WINDS NORTHEAST 20 KM/H INCREASING
TO 40 WITH GUSTS TO 60 THIS MORNING. HIGH 18.
TONIGHT..RAIN. AMOUNTS OF 15 TO 25 MM EXPECTED. WINDS NORTHEAST 40
KM/H WITH GUSTS TO 60 BECOMING LIGHT THIS EVENING. LOW 15.
FRIDAY..RAIN ENDING IN THE MORNING. CLEARING LATER ON. AMOUNTS
EXPECTED OF 5 MM. WINDS NORTHWEST 20 KM/H BECOMING LIGHT IN THE
AFTERNOON. HIGH 22.
```

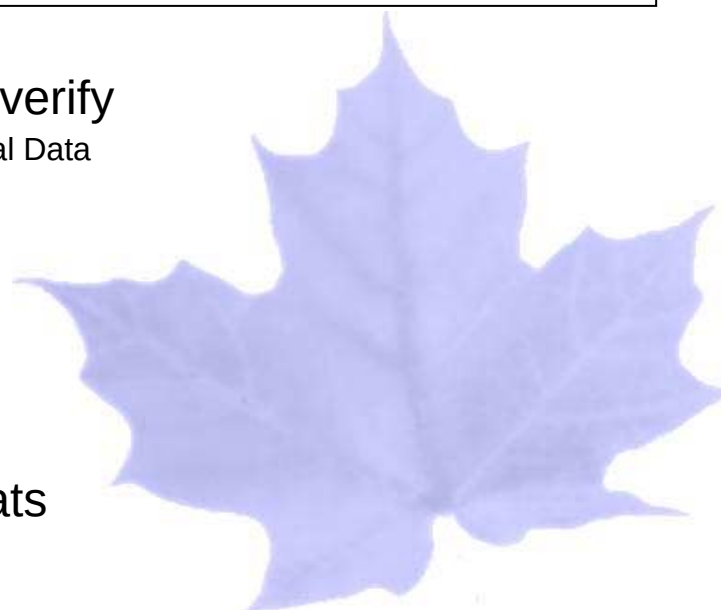
Observation Data

- 06Z synoptic reports (SM's) used to verify
  - daily maximum temperature forecast
  - daily minimum temperature
  - 24-hour precipitation amounts (06 – 06Z)

SM 160600
71624 32969 01807 10191 20170 39950 40155 56014 333 10268
20184 70000 555 31618 40190=

- Hourly observations in BUFR format used to verify
  - **B**inary **U**niversal **F**orm of the **R**epresentation of Meteorological Data
  - 3-hourly temperature
  - 3-hourly probability of precipitation
  - 3-hourly wind direction
  - 3-hourly wind speed
  - 3-hourly sky condition

- SM's and BUFR are very well structure formats



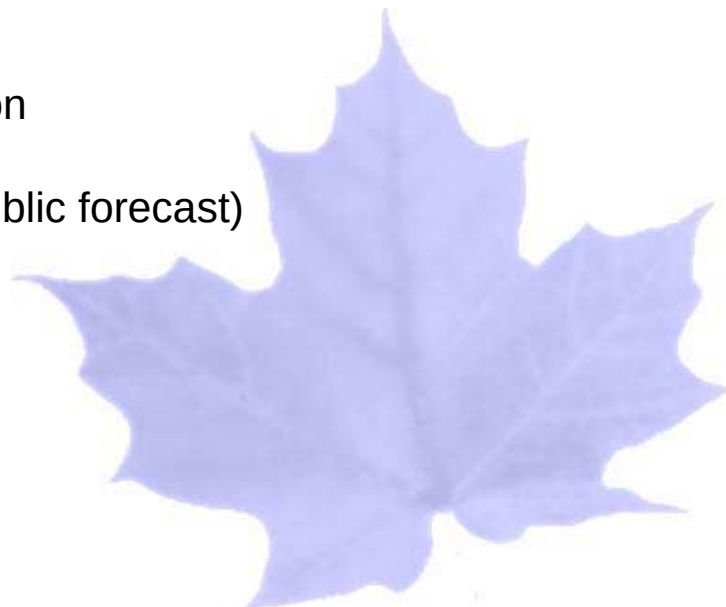
Hardware

- Linux cluster
- PostgreSQL database
- Web interface

- Hourly observations and synoptic reports

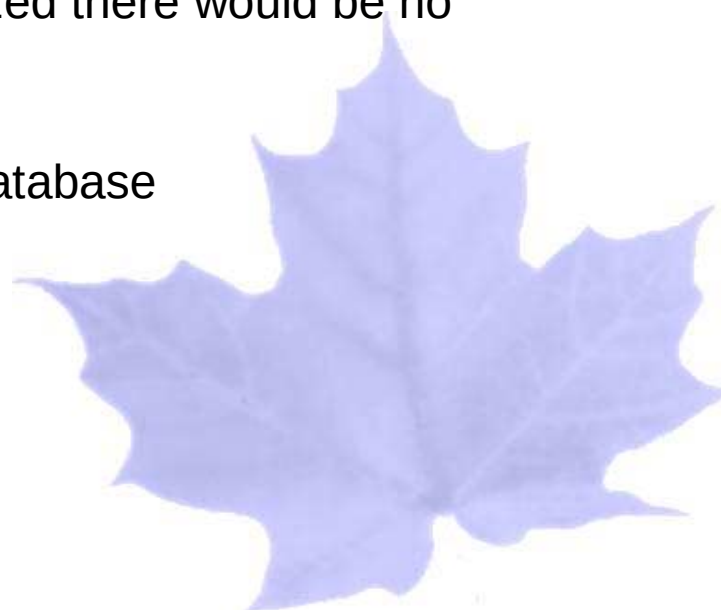
Challenges / Issues

- Large volume of data
- Quality control of data
- Lack of observations
- Which stations to use to verify forecasts for a region
- Which is the official forecast (Meteocode or text public forecast)
- Creditability in our scores



Large volume of data

- Original plan in 2001: a database with observations and forecast to be provided to us before implementation (Production Database Project)
- Six months before implementation, we realized there would be no database unless we built our own
- Much complications with building our own database
  - Hardware infrastructure
  - Storage required 1000 GB+
  - Relatively new technology
  - On-going IT support/resources
- Maintenance of database
  - Backing up large volume of data
  - Disaster recovery



Quality Control of Data

- Quality control originally planned to be part of the Production Database Project
- Very difficult to quality control every piece of data ingested into the system
- Some errors relatively easy to spot:
  - -60°C for minimum temperature near Toronto
  - automated QC possible
- Some are not straight-forward:

```
SPECI CYTZ 091144Z AUTO 36013G20KT 3 1/2SM RA BKN014 OVC066 16/15 A2976 RMK SKYX/=
METAR CYTZ 091200Z AUTO 00011G19KT 3SM RA FEW009 BKN014 OVC066 16/15 A2976 RMK PCPN 3.5MM PAST HR SLP080
    56014 SKYX/=
SPECI CYTZ 091223Z AUTO 36010G18KT 3 1/2SM RA OVC016 16/15 A2976 RMK SKYX/=
```

- Shy away from using precipitation reported by automatic observing stations

Lack of Observations

- Goal: provide performance statistics for all of Canada
- Precipitation forecasting considered very important by the general public
- Precipitation (occurrence/amount) verification challenging
  - Do not have reliable precipitation observations in many forecast regions
  - Network too sparse to represent precipitation pattern
- Complications with inferring precipitation from radar echoes
  - Not all of Canada has radar coverage
  - Need algorithms to deal with ground echoes, virga, phase change bright banding, etc
  - Decide which ZR relationship to use
  - Event more information to be ingested into database
- Free formatted observations like volunteer reports cannot be used

Which Observing Stations to Use?

- We verify region forecasts, NOT point forecasts
- Need to match forecast for a region with representative stations
- Complications:
  - No representative observations
  - what weighting scheme to use if there are more than one
- Even point forecast not always straight forward:

  e.g. Today: Sunny. High 25 **except 18 near the shore**.

Forecast Performance for FPCN11 CWTO / City of Toronto
Verified against CYYZ CYTZ

Statistics based on forecaster-adjusted Scribe data are shown in bold. Corresponding values for unadjusted Scribe data in brackets							
	Sample Size	Bias	Root Mean Square Error	Absolute difference between forecast and observed values			
				% at or within 1 deg	% at or within 3 deg	% at or within 5 deg	% at or within 10 deg
Day 1	62 (62)	**0.35** (1.35)	**1.89** (2.11)	**47** (39)	**92** (84)	**97** (97)	**100** (100)
Day 2	62 (62)	**0.64** (1.64)	**2.25** (2.53)	**48** (34)	**82** (74)	**95** (94)	**100** (100)
Day 3	62 (62)	**0.48** (1.42)	**2.07** (2.41)	**48** (40)	**82** (79)	**97** (95)	**100** (100)
Day 4	62 (62)	**0.48** (1.29)	**2.82** (2.59)	**37** (31)	**82** (74)	**90** (97)	**100** (100)
Day 5	62 (62)	**0.35** (0.87)	**2.18** (2.38)	**35** (47)	**87** (89)	**97** (94)	**100** (100)

Official Forecast

- Which is official: Meteocode or text forecast?
- Currently text forecasts are “official”
  - need disclaimer to indicate our scores may not reflect official forecasts

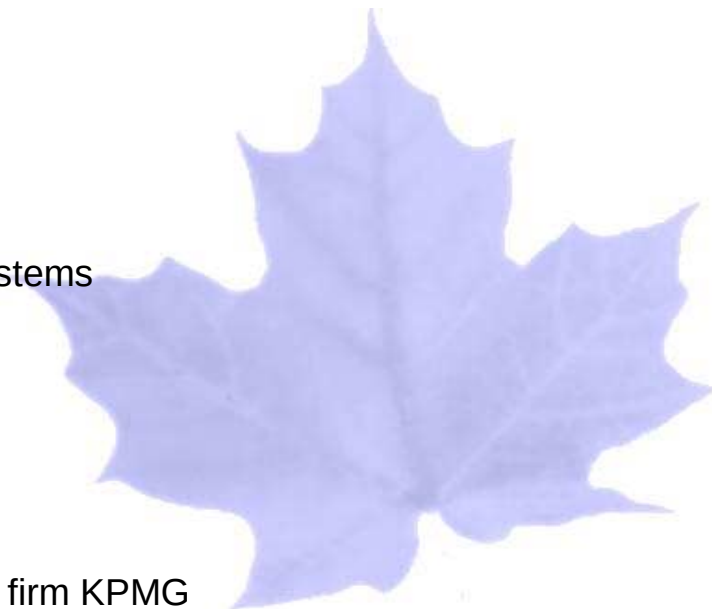
Monthly Maximum Temperature Verification for Ontario
Forecast Data Extracted from Early Morning Regular Issues*
Verification Period 2004-08-01 to 2004-08-31

* Forecast data are extracted from Scribe Meteocode files. Due to difference in level of Scribe implementation in the regions, official forecast data may not necessarily reflect contents of the corresponding forecast bulletins transmitted.

- Use of Meteocode not completely standardized in MSC regions
  - regional practices may differ
- Meteocode files offer more elements to be verified

Credibility

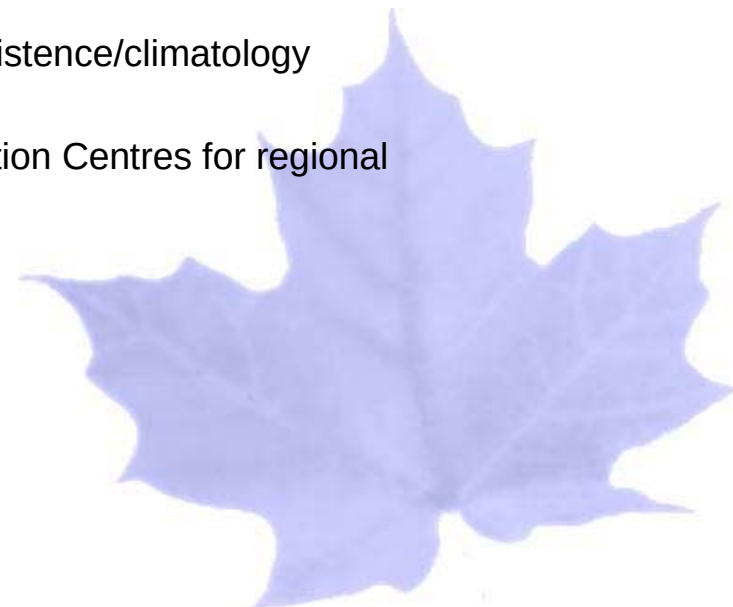
- Scores may be used for many purposes and by many users
- Question from users: Why should we trust your scores?
- Issue have yet to be addressed
- What we can do now
  - Provide raw data for transparency
  - Compare our scores with those from existing MSC systems
- What we need to do in the future
  - Establish standards in operation
    - ISO 9000/9001 standard registration
  - Hire independent party to review operation
    - our aviation system was reviewed by accounting firm KPMG



High Impact Weather

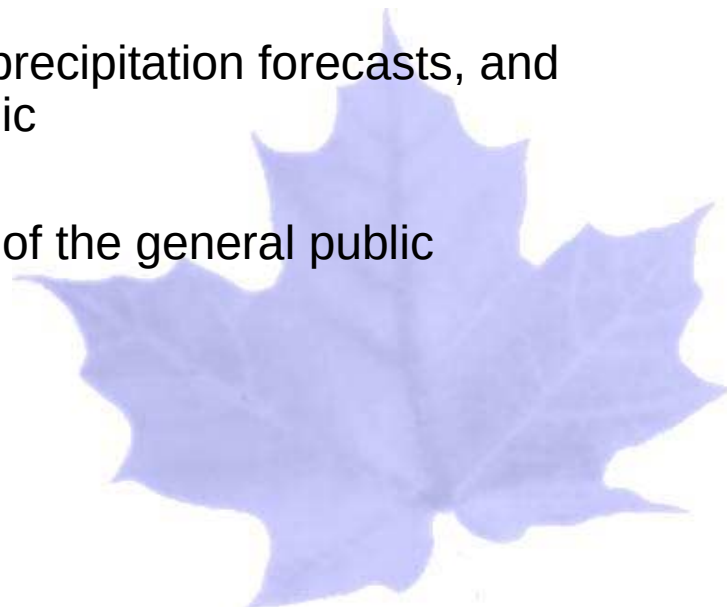
- Abrupt changes in temperature/wind velocity, near freezing temperatures, etc

- What we can do in the immediate future
  - stratify temperatures based on departures from persistence/climatology
    - Brooks and Doswell (1996)
  - partner with national labs and regional Storm Prediction Centres for regional verification efforts



Scores for the General Public in Canada

- Our division has been tasked with building a plan for disseminating performance scores to the general public
- Survey in 2002 indicates weather warnings, precipitation forecasts, and temperature forecasts most important to public
- Scores should be easily understood by most of the general public
- Production of scores should be automated

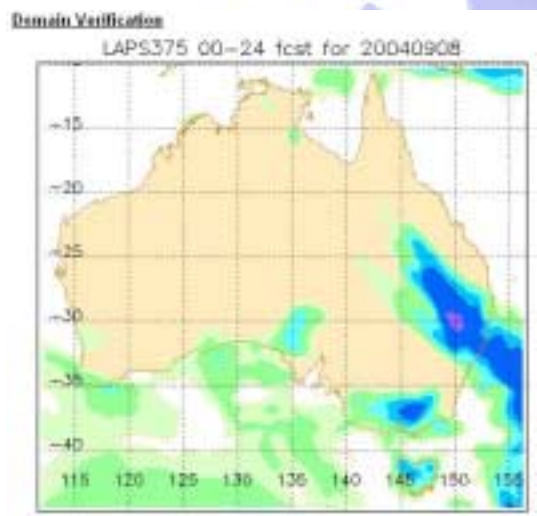


Verification Scores for the General Public – World Examples

- American
  - Point verification of the National Digital Forecast Database
  - Maximum/minimum/dew point temperatures, prob. of precip., etc
  - MAE, bias, Brier Score, etc

- British
  - Point verification for 5 cities
  - Maximum/minimum/wind
  - daily forecast error

- Australian
  - Grid point verification
  - Quantitative Precip. Forecast
  - MAE, bias, POD, FAR, etc



Proposed Scores for the General Public

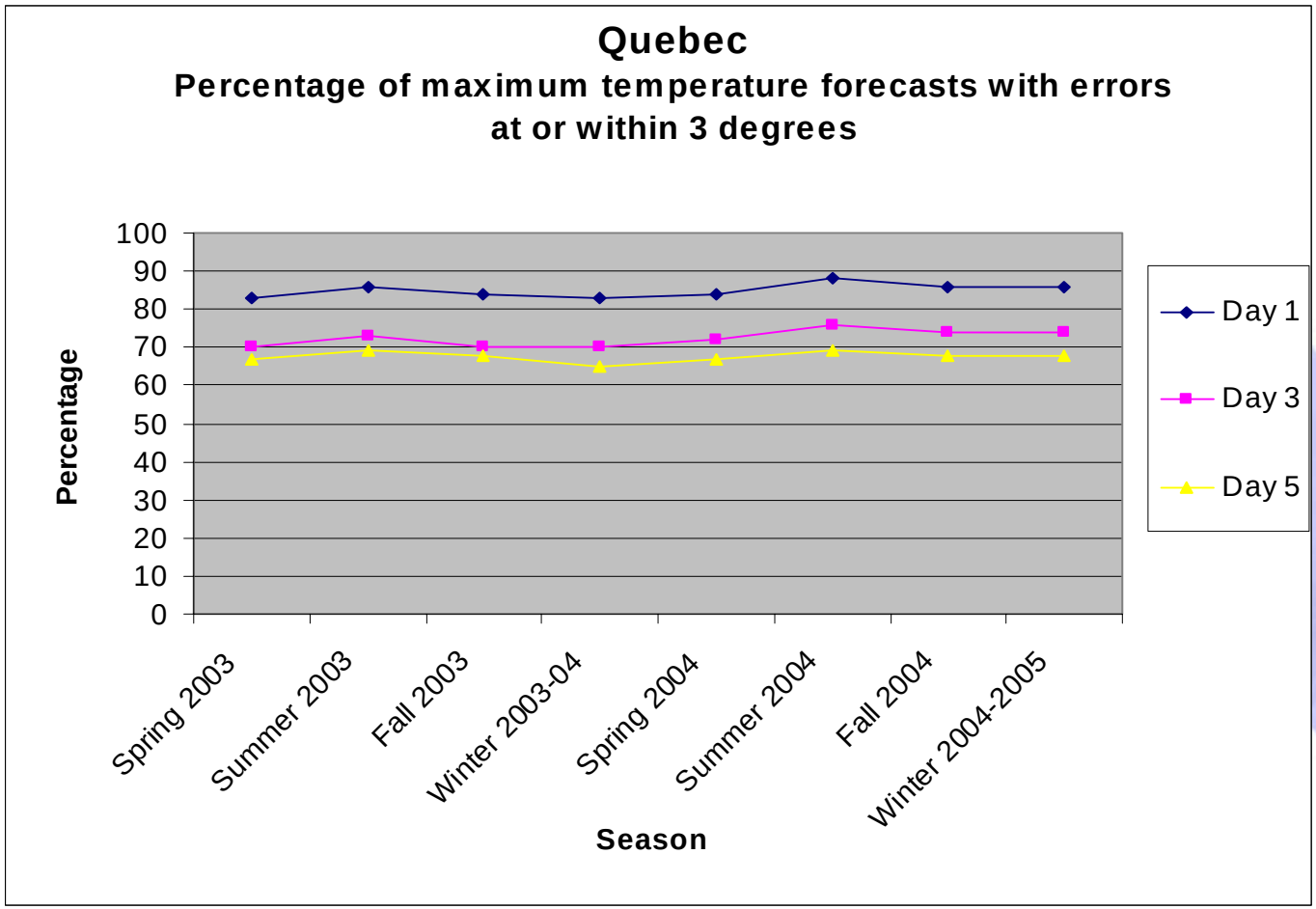
- Weather elements verified
  - Daily maximum temperature
  - Daily minimum temperature
  - Precipitation occurrence
- Daily Maximum/Minimum temperature
  - Percent correct within 3°C of observed value (0 to 100%)
  - 3°C threshold corresponds to results of 2002 survey
- Precipitation Occurrence
  - Positively oriented Brier Score (0 to 100%)
- Scores relatively intuitive to the public
  - 0% to 100% → bad to excellent

$$BS = 1 - \left(\frac{1}{N}\sum_{i=1}^{N}(p_i - o_i)^2\right)$$

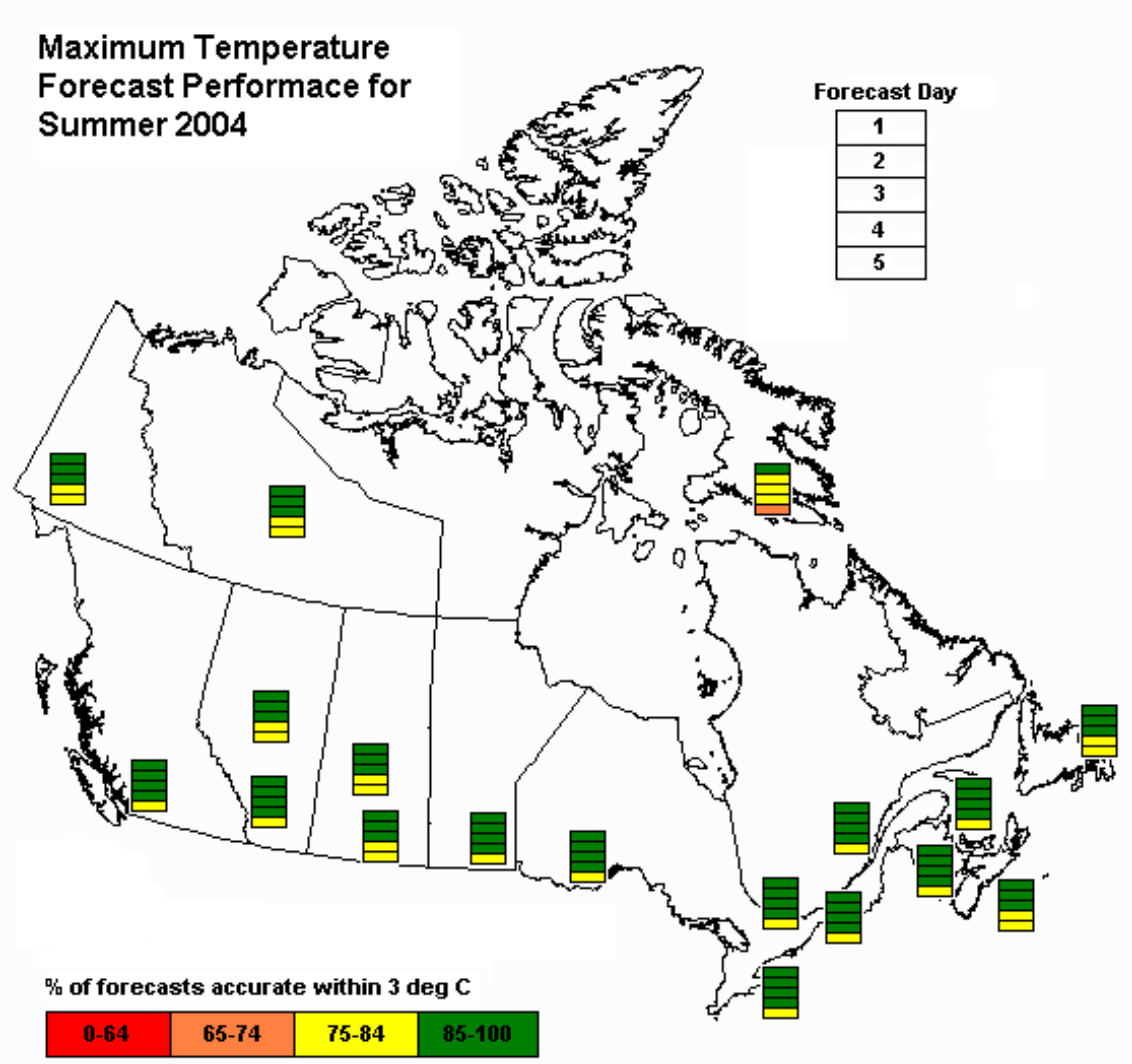
Proposed Scores for the General Public

- Scores calculated for major centres in Canada
- Scores reports to be published every three months
- Hosted on main public web site for official Environment Canada forecasts
- Both graphical and tabular formats provided

Proposed Display of Scores



Proposed Display of Scores



Other Proposed Ideas

- Creation of a composite score
  - e.g. one score that captures the performance of forecast precipitation, temperature, etc
  - assign arbitrary weights to forecast elements
  - Phoenix Score model
  - New Zealand weather service model
- Creation of aggregate scores
  - e.g. one score that captures the performance of forecast temperature for the whole country