Crowdsourced data improves temperature forecasts on Yr.no

Thomas Nipen¹, Ivar Seiderstad¹, Cristian Lussana², Nina E. Larsgård³ & Mareile Wolff³
Norwegian Meteorological Service
¹ Center for development of weather forecast
² Division for climate services
³ Division for observation quality and data processing
small – bigger – biggest …
Unconventional data sources are increasing

- Private weather stations and meteorological sensors
- Smartphones (pressure, temperature)
- Carsensors (temperature, precipitation)
… and the models?

yr.no distributes temperature forecasts which may be more than 3 °C off the actual values, even if some postprocessing during inversion situations is already in place.
Why amateur stations at MET?

- models (NWP/RRA) can still have large errors
- models get higher resolution
- Large number of stations now available
- Need for accuracy can be less in post-processing than for other applications

Diagram:

- New data sources
  - Data assimilation
  - NWP model
  - Post-processing
  - End-users
Netatmo @ MET Norway

- Live data feed (every 10 min)
- Archive (2013-now)
Added value of a denser network
Added value of a denser network

Cold pools

Tryvann

Inversions

Urban effect

Ocean effect

28/03 2018 05:00 UTC
What about the quality?

• Optimal interpolation analysis with netatmo stations introduces a small shift, which can be corrected for.

• Need for accuracy can be less in post-processing than for other applications.
How do we quality check the observations?

Spatial quality check (about 30% are removed)
Quality checks are performed every hour independently.

Station exposed to sun

Station ok later in the day
How do we quality check the observations?

Isolated Netatmo-stations are removed
...but not isolated WMO-stations.
MET Analysis Nordic Temperature 1x1 km

Corrections cover large parts of the country
Nowcasts and forecasts significantly improved

Corrections have greatest influence the first 12 hours
Results are cross-validated against WMO stations.
Application: Automatic weather forecast

Launched on *Yr.no* 19 march 2018
…offers possibility for additional quality checks of Met Norway’s conventional stations

The number of available weather observations is constantly increasing.

We need to be ready to check massive amount of data in a reasonable time.

→ Taking the best out these two techniques:

• individual tests of parameters from one location based on meteorological knowledge and experience
• assessing the quality of a data-”crowd”
Summary

An increasing number and variety of crowdsourced data are available and advanced quality control techniques are under development.

Effective quality control for crowdsourced temperature data in place and amount of large now-cast errors could be significantly decreased for individual locations.

Working on: Combination of “big data” statistical analysis techniques with more traditional quality control based on meteorological knowledge and advanced measurement technique and available metadata.

There are different demands on quality of data for different purposes

Open Data Access - Open Source Software - Open for Collaboration

• Post-processing  https://github.com/metno/gridpp
• Quality Control  https://github.com/metno/TITAN
• Post-processed dataset (and raw forecast)  http://thredds.met.no/thredds/catalog/metpplatest/catalog.html
We are not talking either or...

Crowdsourced data have a high spatial distribution of stations, there where people are.

Crowdsourced data have no global coverage – especially more extreme climates require advances observation methods (quality and reliability).

Crowdsourced parameters do not deliver reliable meteorological measurements during a crisis.

Crowdsourced data are not for free (fast and ever-changing technology, dependencies on data providers, privacy issues).

Crowdsourced data are found, so far, for only a limited amount of parameters.

High quality time series for climate monitoring and continuation of long time series needs the continuation of high quality meteorological/climatological stations.

Crowdsourced data needs to be verified with independent and controlled observations – as not all errors can be identified by the shown quality checks.

Important with independent knowledge and understanding of meteorological measurement technology to make optimal use of crowdsourced data.
Some meteorological philosophy...

If you are performing meteorological observations. do it right!
If not, it’s better NOT to do it!

Andrija Mohorovičić (1857-1936, Kroatian meteorologist and seismologist)
A paradox?

No data is better than bad data

Any data is better than no data *

*from 1st day panel discussion on New Data Sources
Utskifting av bakgrunnbilde:
- Høyreklikk på lysbildet og velg «Formater bakgrunn»
- Under «Fyll», velg «Bilde eller tekstur» og deretter «Fil…»
- Velg ønsket bakgrunnbilde og klikk «Åpne»
- Avslutt med å velge «Lukk»
Quality control of Netatmo temperatures

1. Altitude check
   ± 5 STDEV of neighbouring altitudes

2. Buddy check
   ± 5 STDEV of neighbouring observations

3. Forecast ensemble check
   ± 5 ensemble STDEV of ensemble mean

4. Spatial consistency check*
   Close to the cross-validated field