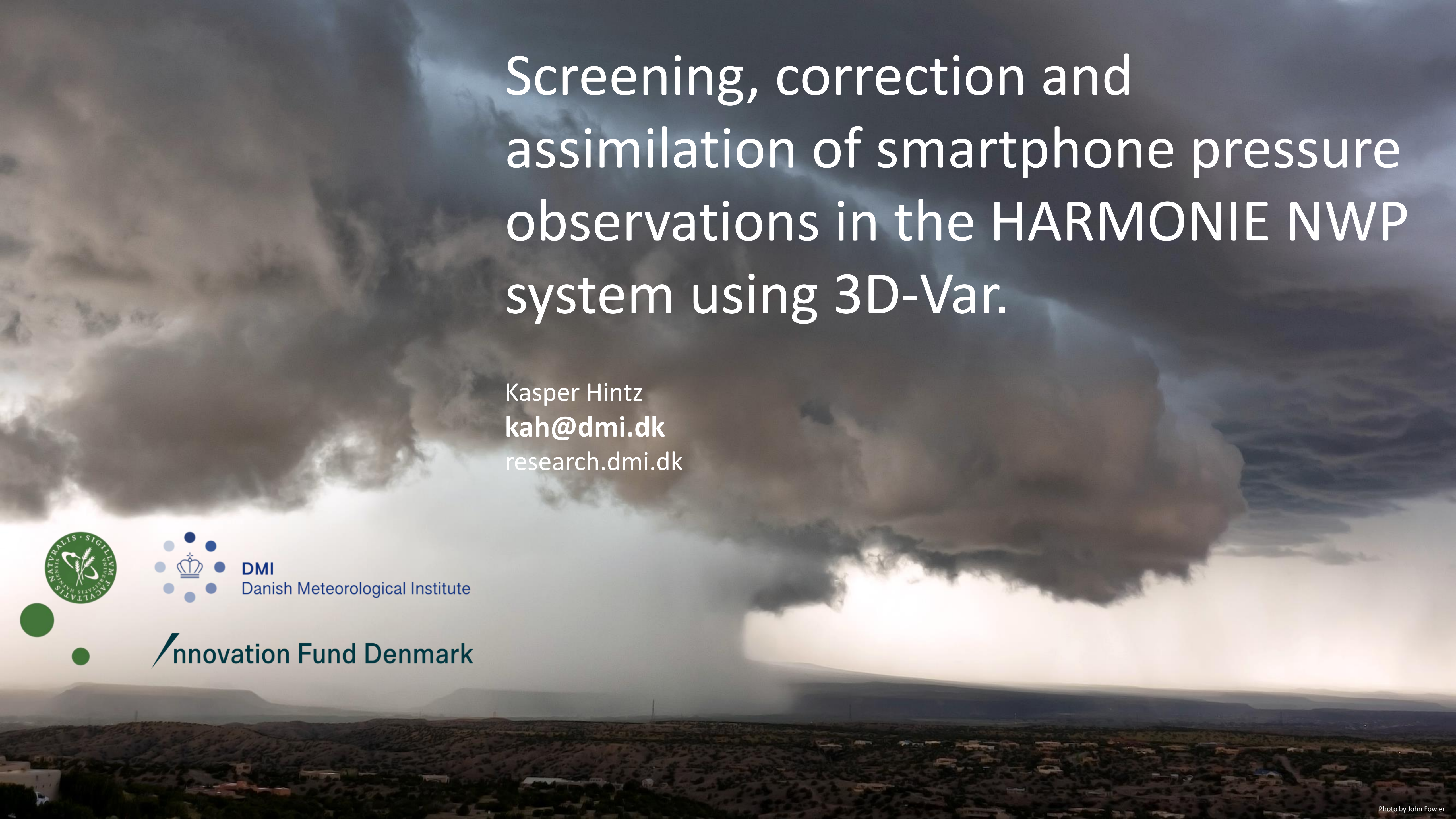


Screening, correction and assimilation of smartphone pressure observations in the HARMONIE NWP system using 3D-Var.

Kasper Hintz
kah@dmi.dk
research.dmi.dk

DMI
Danish Meteorological Institute

Innovation Fund Denmark

Photo by John Fowler

20 August 2007
Photo: Casper H Petersen

2 July 2011
Photo: Unknown

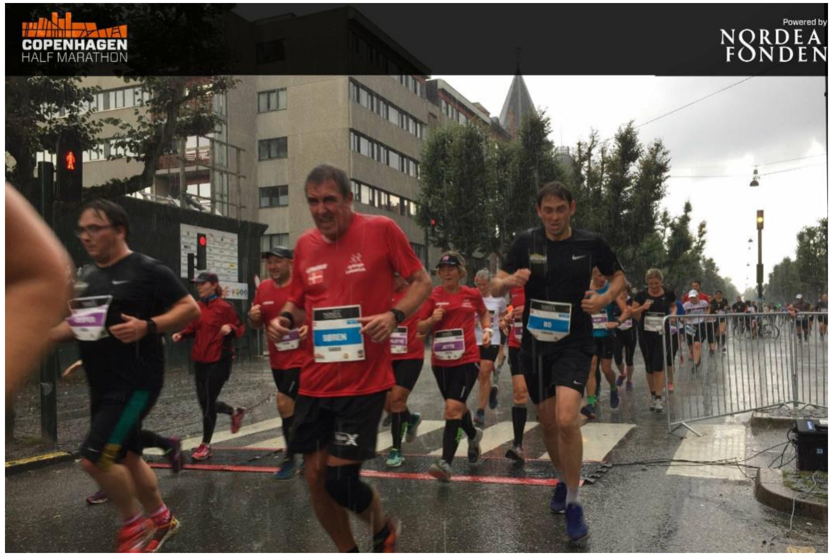
4 September 2015
Photo: Jens Larsen

14 August 2010
Photo: Jens Dresling


POLITIKEN

DANMARK | KULTUR | DEBAT | INTERN. | MENU

Løbet blev aflyst:
Københavns halvmaraton ramt af voldsomt uvejr



Motivation

The primary goal is to make better 'nowcasting' forecasts and products for our meteorologists as we need to provide a forecast capability and service to warn rapidly developing, extreme weather events.

This means rapid update cycles and better use of current observations and exploits new high-resolution observation types. (more on this later)

DMI
Danish Meteorological Institute

Innovation Fund Denmark

Smartphone Pressure Observations

What is our motivation?

Why are there barometers in smartphones?

Link with Applications

How to get the observations

Data policy and privacy constraints

Screening and Correction

Standard checks of SPOs

Predicting errors using machine learning models

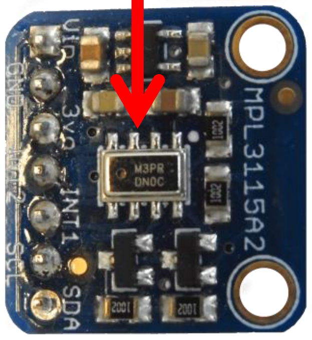
Assimilation in the HARMONIE system

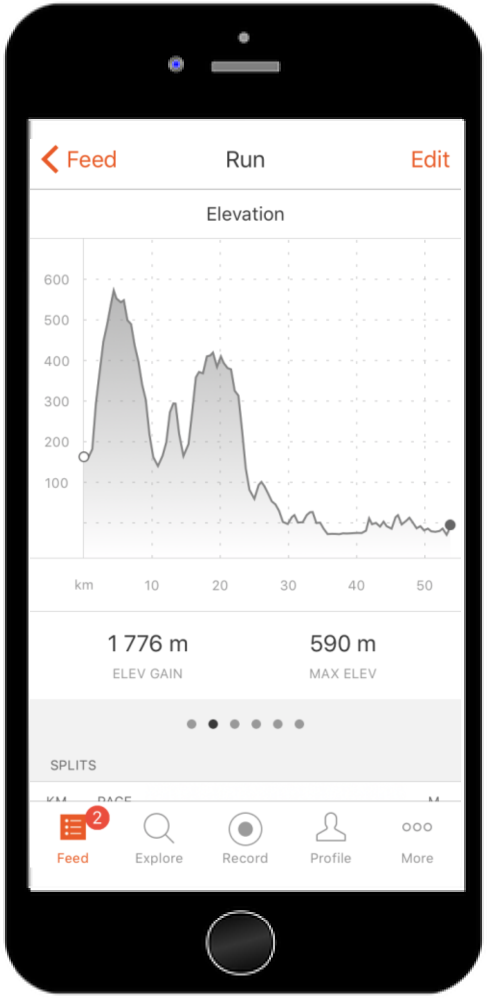
Results and future aspects and experiments

Smartphone Pressure Observations

Most smartphones measures the atmospheric pressure

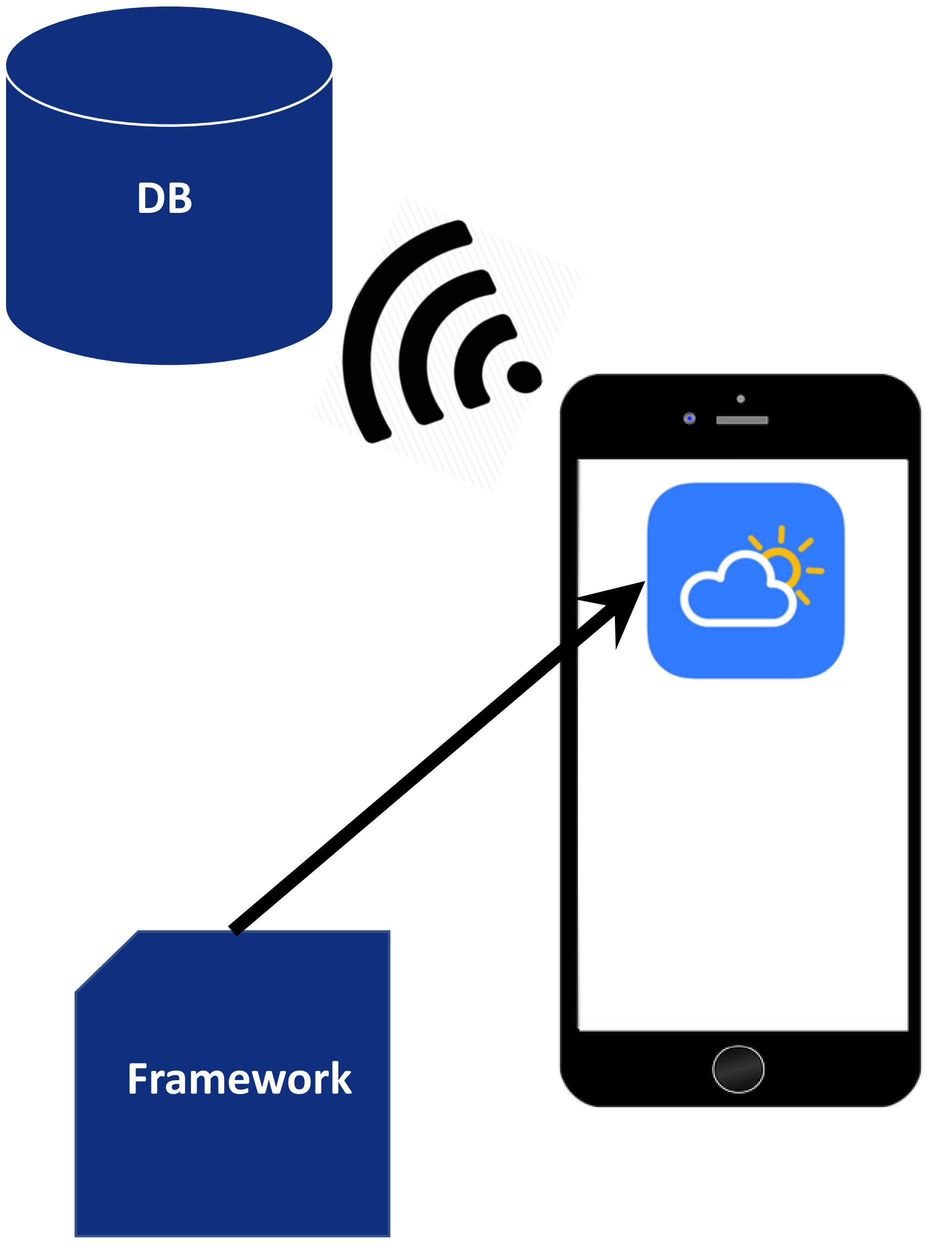
- Used to monitor changes in altitude and acquire a fix of the location of the device faster.





So what?

- Pressure is an essential variable in NWP and are being assimilated from conventional sources today.
- Can potentially also be used for verification and/or nowcasting purposes on convective scales.

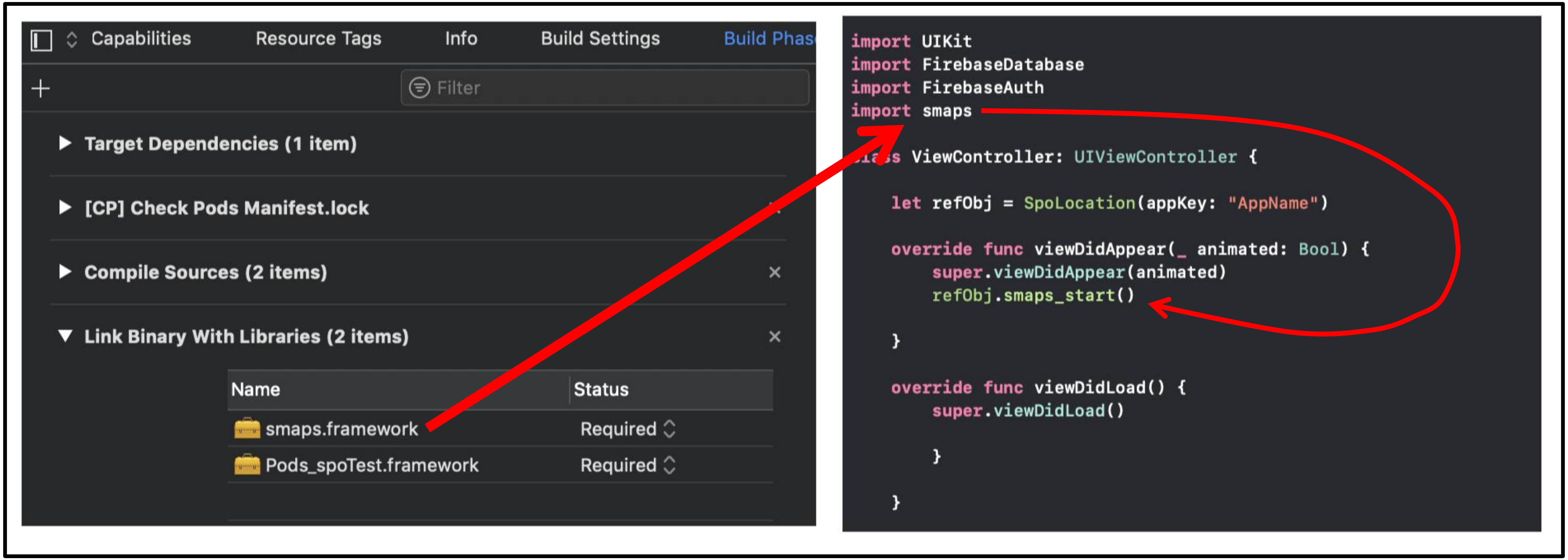


We **do *not*** develop and maintain a full-stack app with a main focus on collecting observations.

We **do** develop and maintain a framework with the sole purpose of collecting observations.

Objective: Keep data processing in the meteorological community and collaborate on data collection between meteorological services.

Basic example of starting observation collection



Screendump from Xcode (MacOS IDE)

There is still room for improvement and we welcome all collaboration.

kah@dmi.dk | hev@dmi.dk

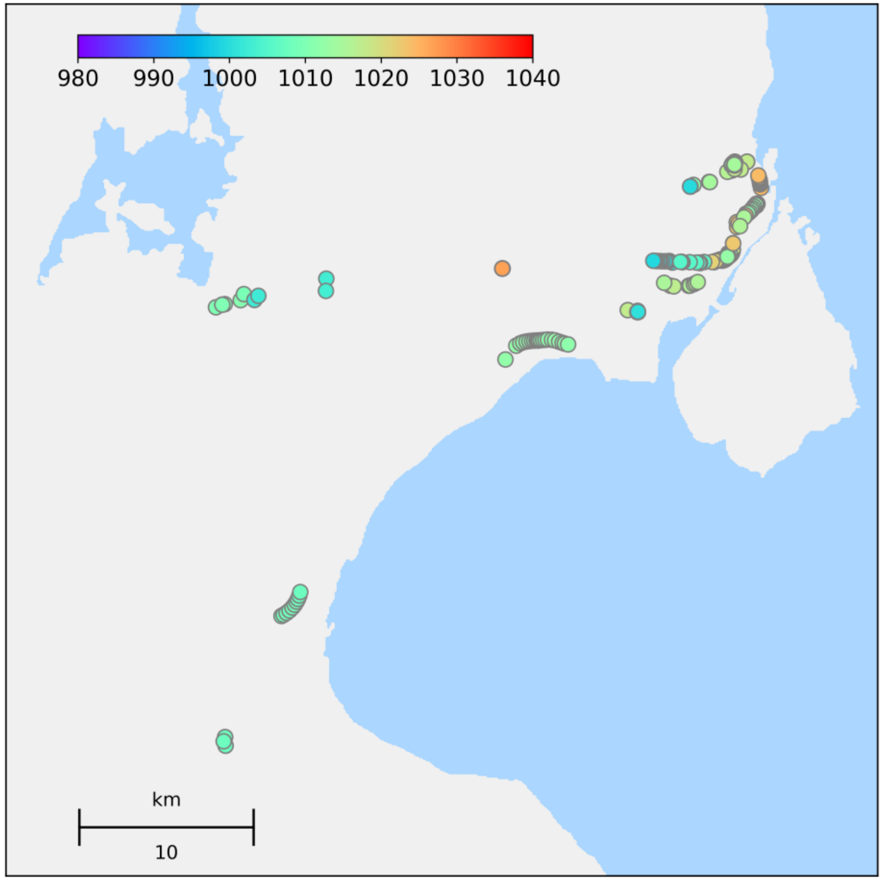
It is still unknown if extra (meta)data from the smartphones can be used for quality control and/or correction; **McNicholas and Mass (2018)** showed excellent results indicating, that this is the case.

Therefore the following is collected if available:

- **Pressure**
- **Latitude**
- **Longitude**
- **Altitude**
- **Timestamp**
- **User ID**
- **Acceleration** in three dimensions
- **Speed** of device
- Horizontal **Accuracy**
- Vertical **Accuracy**
- STD of relevant parameters, calculated on the phone directly

Furthermore, the following is appended:

- **Residual** (observation-background)
- **DHM** (Danish Terrain Model) height at location



Makes data personal and hence processing must comply with the GDPR act from EU.

How will "Brexit" affect data collection in the UK?

Some comments on privacy issues

Currently users are asked for permission to fetch data.

This can introduce problems later:

- Permissions must be collected, registered and can be cancelled.
- If the intended data usage changes you have to ask for permission again.
- Can give bias in data (*does properly not apply for NWP*)

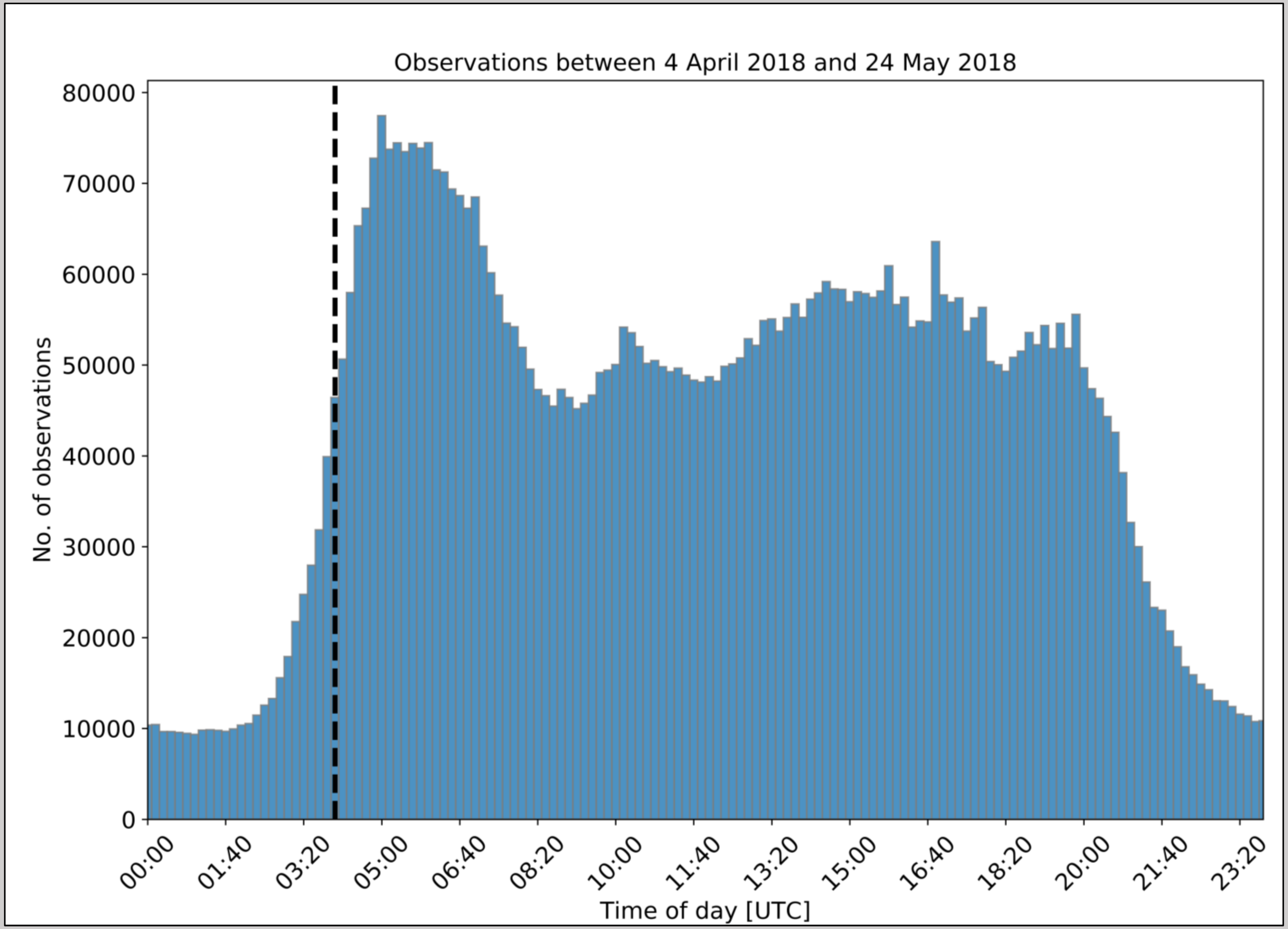
Fully anonymized data are not governed by GDPR. (A way around GDPR?)

- Requires data minimization and generalization that often ‘destroys’ the value of the data.

Article 6: ***Consent***, Contract, Compliance with a legal obligation, Vital Interests, ***Public Interest***, Legitimate Interests.

- Consent is only one among multiple reasons for “Lawfulness of processing”.

Daily reports from smartphones between 4 April - 24 May 2018



During the first year **61,728,672** observations was collected from **149,782** unique devices.

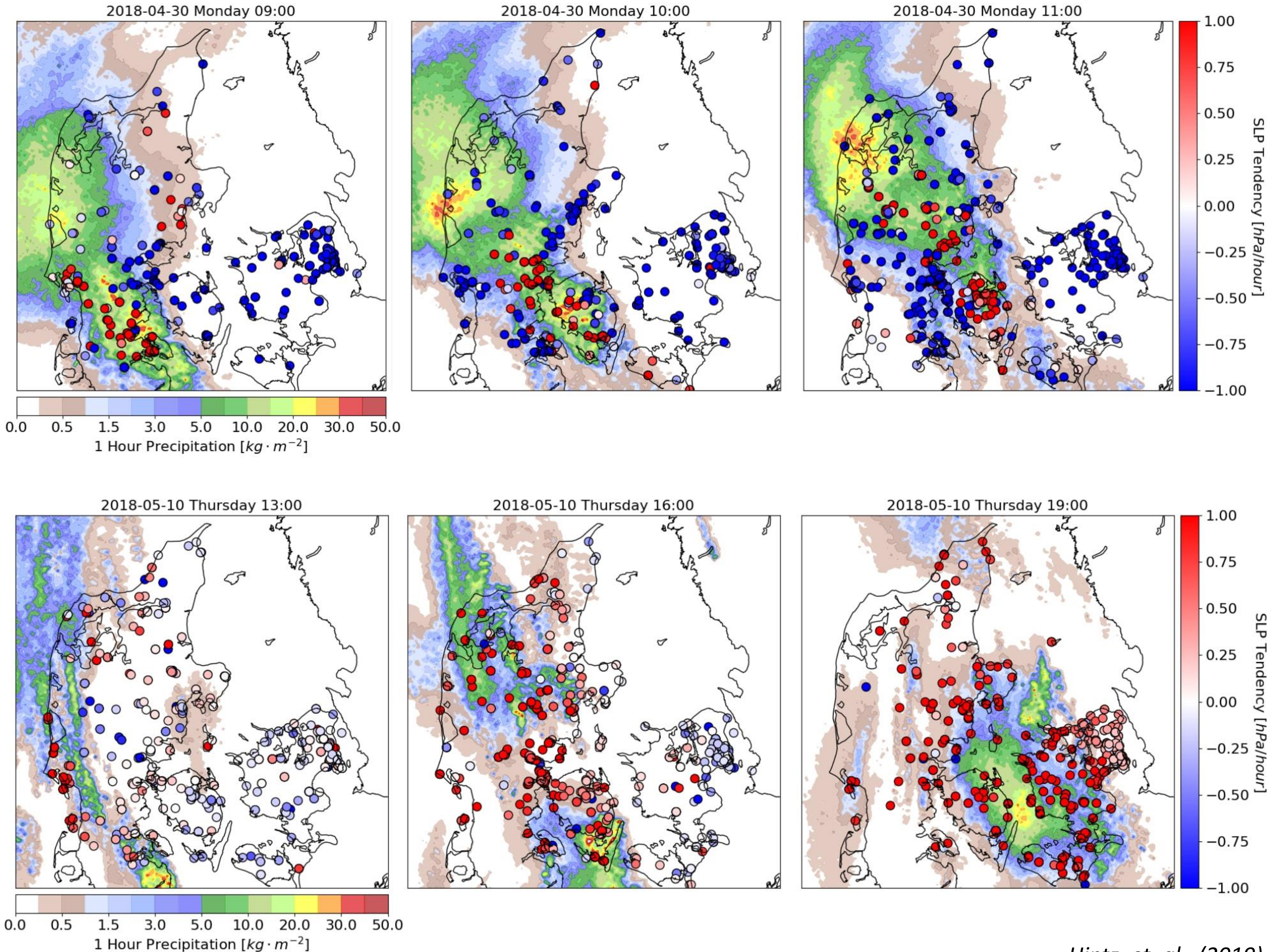
Assuming this is scalable to the rest of Europe, gives 15 million smartphones (some of which reports more than once per day)

10^{-6} € per observation in maintaining costs.

Pressure tendencies (raw data) plotted with radar-estimated precipitation.

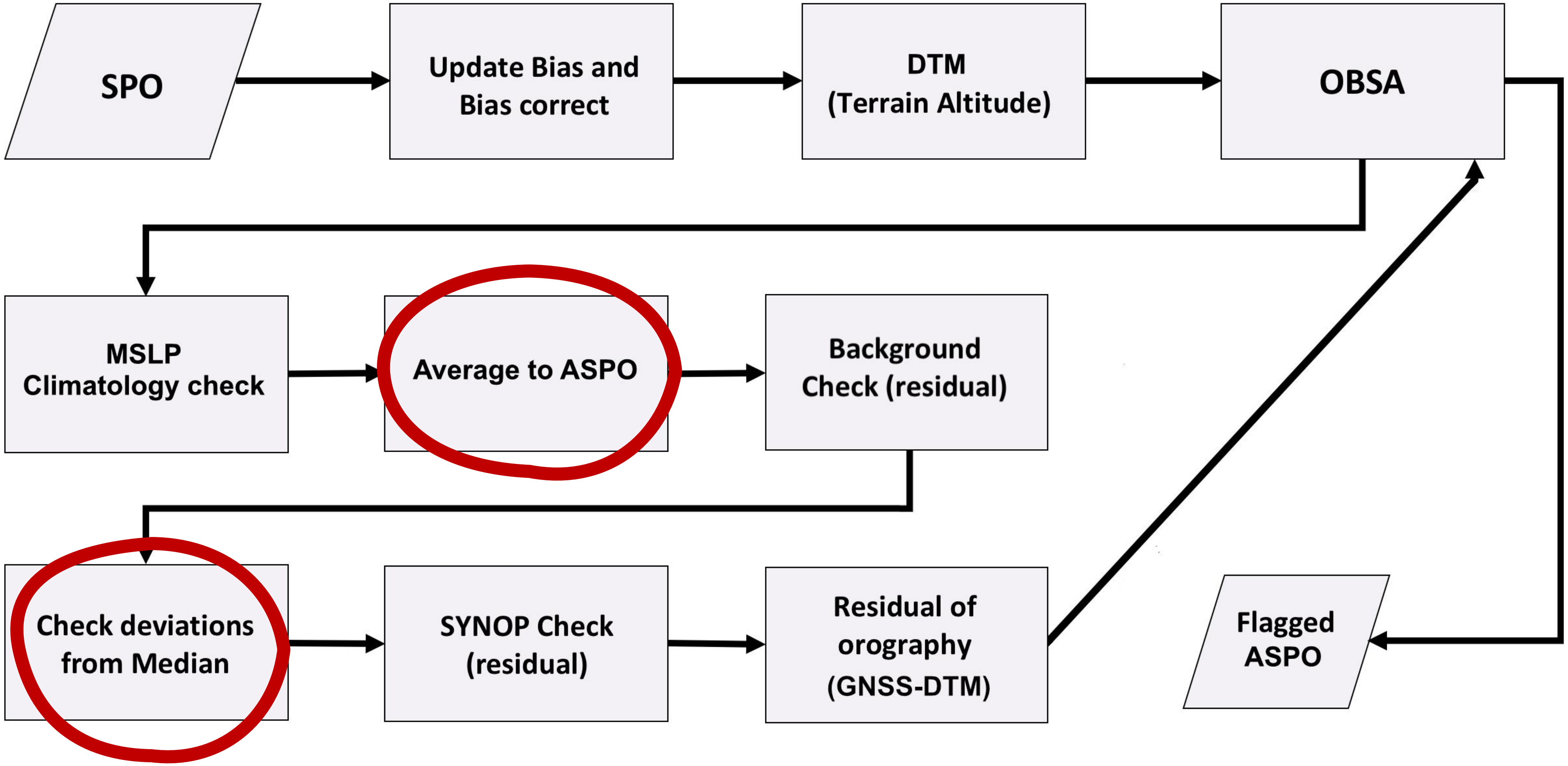
Frontal zones can be identified directly.

Necessary to ensure only good observations are assimilated without throwing too much away.



Hintz, et. al., (2019)

Screening of SPOs (Smartphone Pressure Observations)

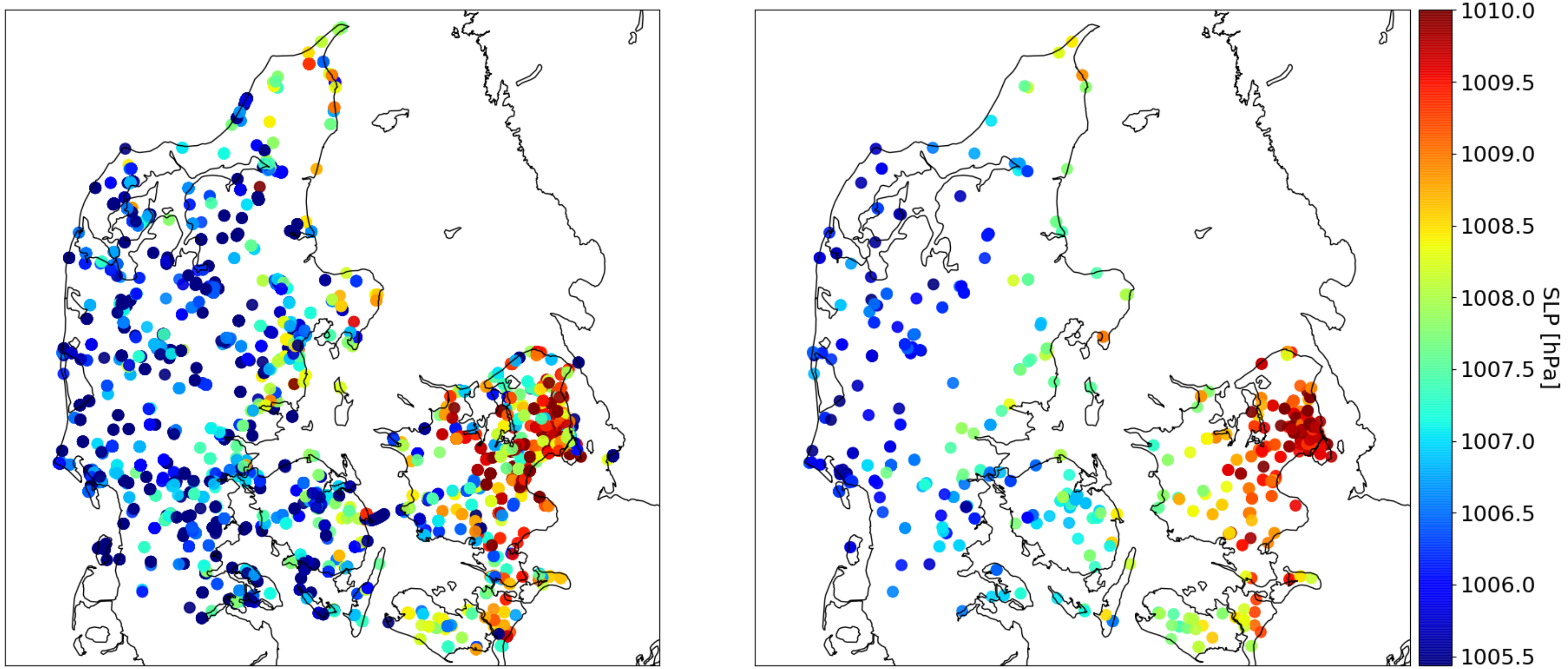


Each observation is allocated with a **Flag** value of zero. Each routine adds a penalty to the **flag** value if it is found suspicious.

Different settings for the median check has been applied to filter more or less observations.

EXP_MEDV1 (**'loose' median check**): SPOs deviating more than 1.0 hPa from median within 20 km is flagged.

EXP_MEDV2 (**'strict' median check**): SPOs deviating more than 0.2 hPa from median within 50 km is flagged.

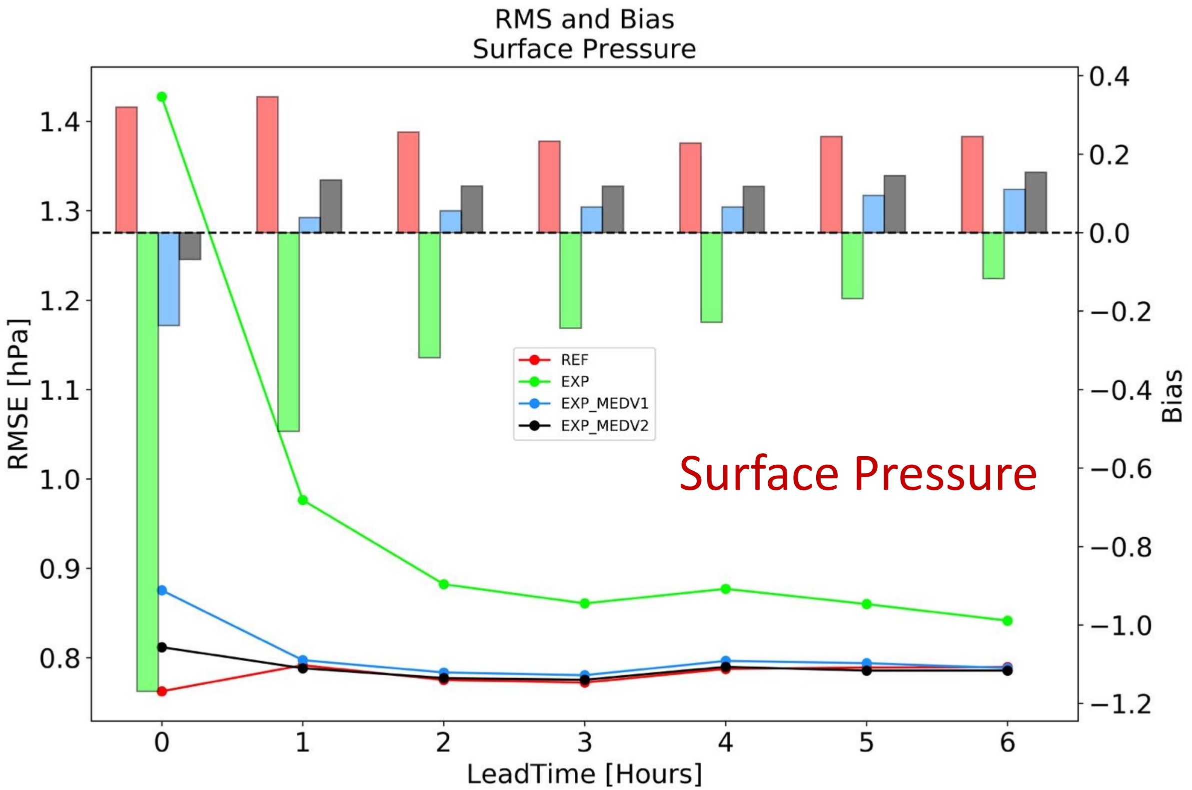


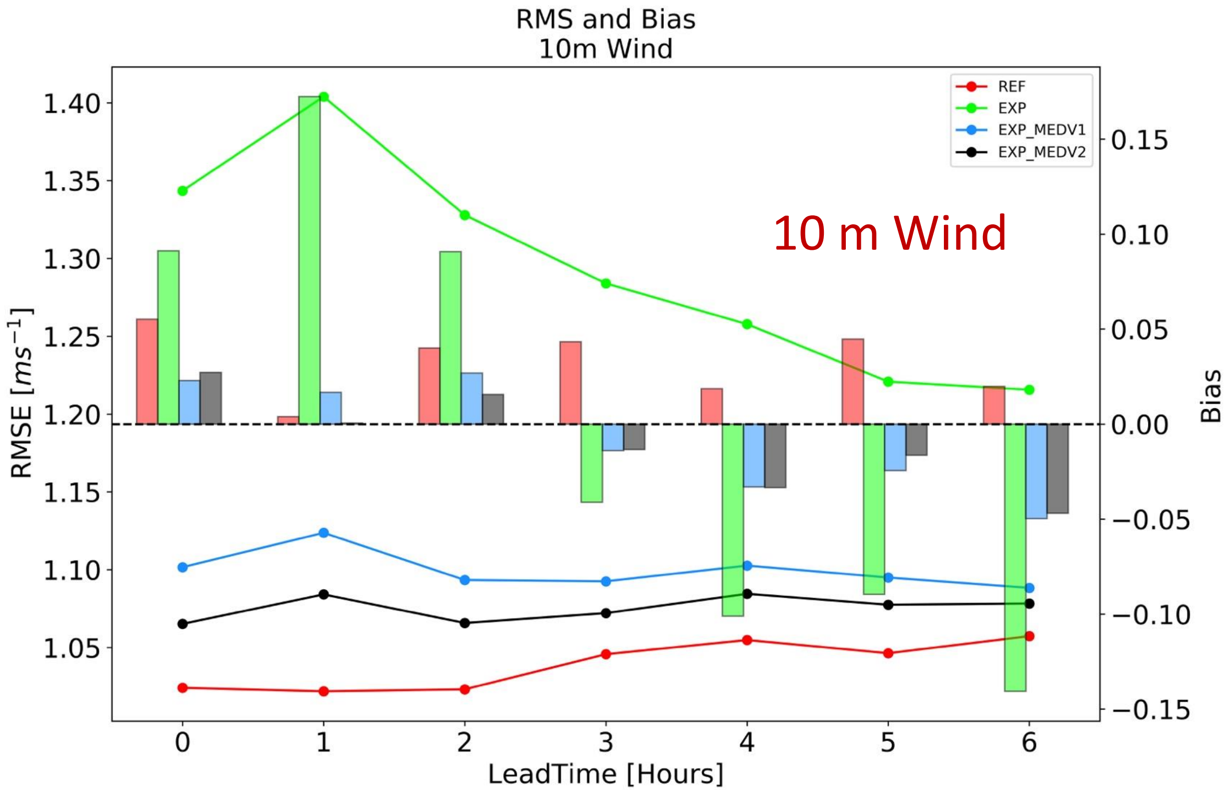
10th of May 2018, **9-10 UTC.** **'Loose' Median Check**

In average about 90 % of all observations are flagged ('strict' check)

NWP Experiments with 3D-Var (Harmonie c40h1)

Date range: 5th May 00 UTC – 10th May 12 UTC. **DA Cycle:** 3 Hours.





Ref: No pressure observations from Denmark

EXP: No filtering of SPO

EXP_MEDV1: Filtering with ‘loose’ median check

EXP_MEDV2: Filtering with ‘strict’ median check

In another experiment, bias decreased from 0.35 hPa to -0.15 hPa over two months using SPOs

To facilitate more observations experiments correcting SPOs before screening using different machine learning algorithms are tested.

Training data: April 2018 – November 2018

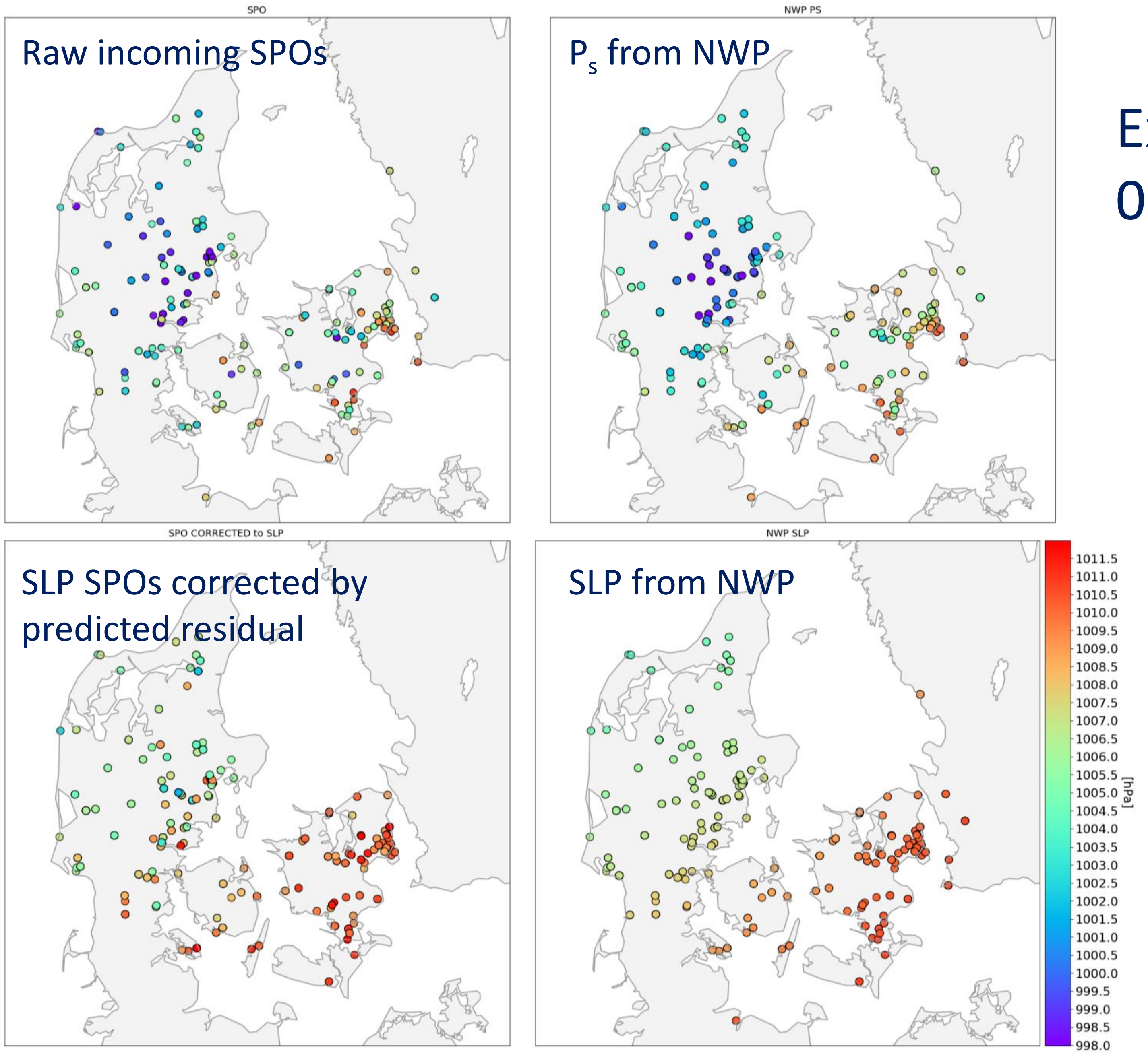
Validation data: December 2018

Y: Target of prediction: Residual between NWP and observation

X: Features: Ps, latitude, longitude, σ_p, hour of the day (still experimenting)

Y = F(X) (Added to observation)

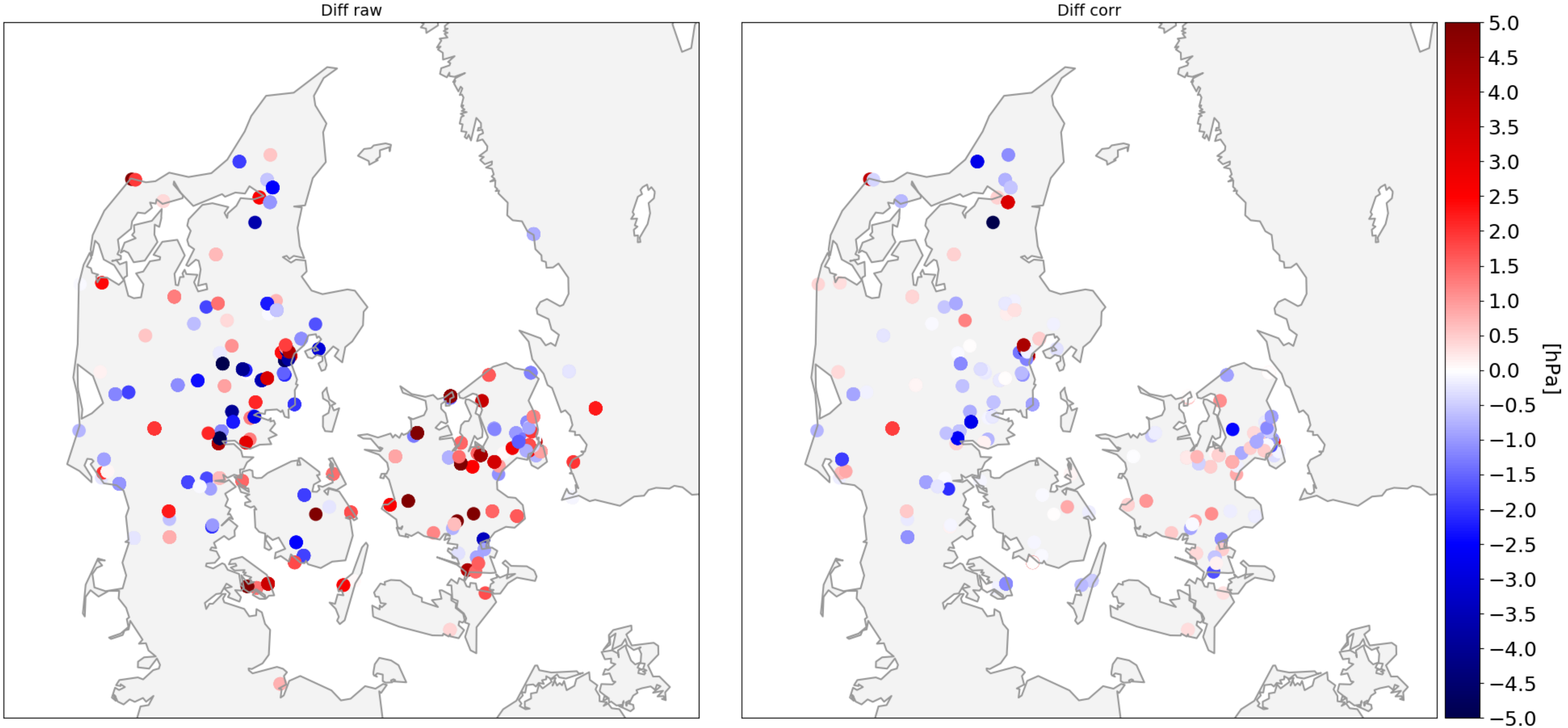
An individual model is trained for each device. Three simple models, a k-Nearest-Neighbour, a random forest and a Gradient Booster have be tested with different features.



Example for 2018 1st December, 0630 - 0640 UTC.

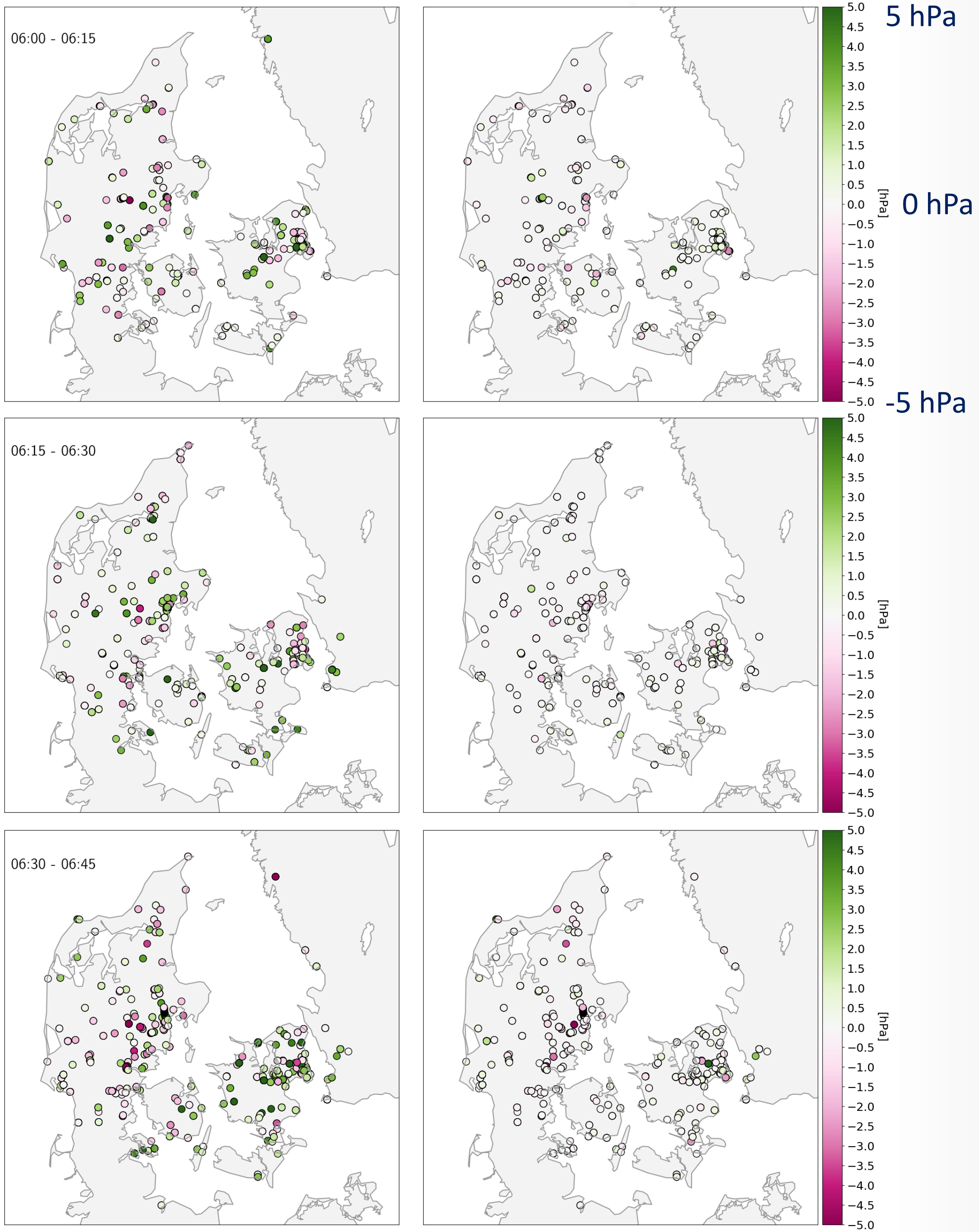
Models trained with a Random forest using only:
Observed Pressure,
Latitude,
Longitude

Example for 2018 1st December, 0630 - 0640 UTC. (Zooming in)



SPO - NWP

ML Corrected SPO - NWP



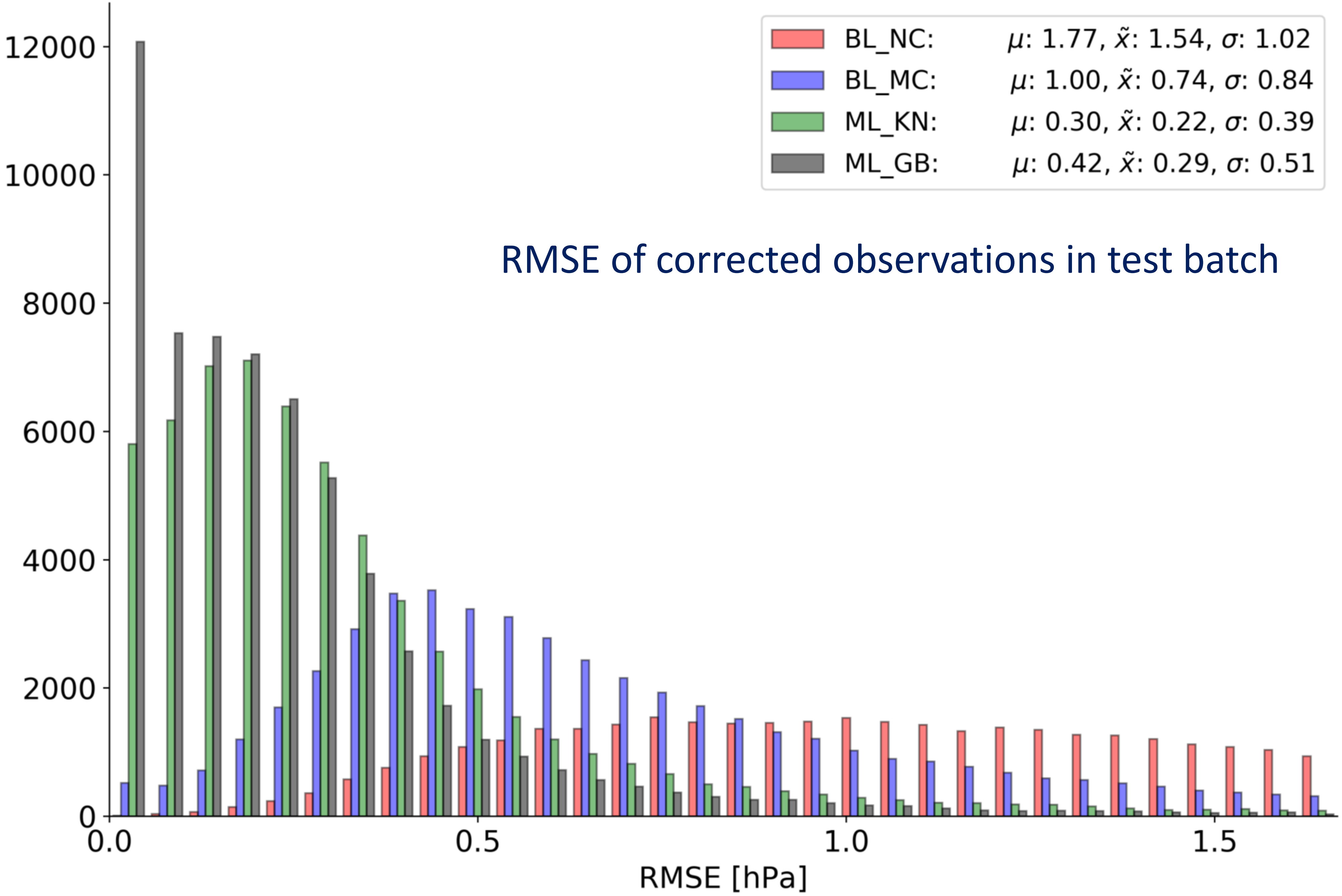
Residual of SPOs using a short-term NWP forecast as reference.

Left Column: Uncorrected observations

Right Column: Corrected observations

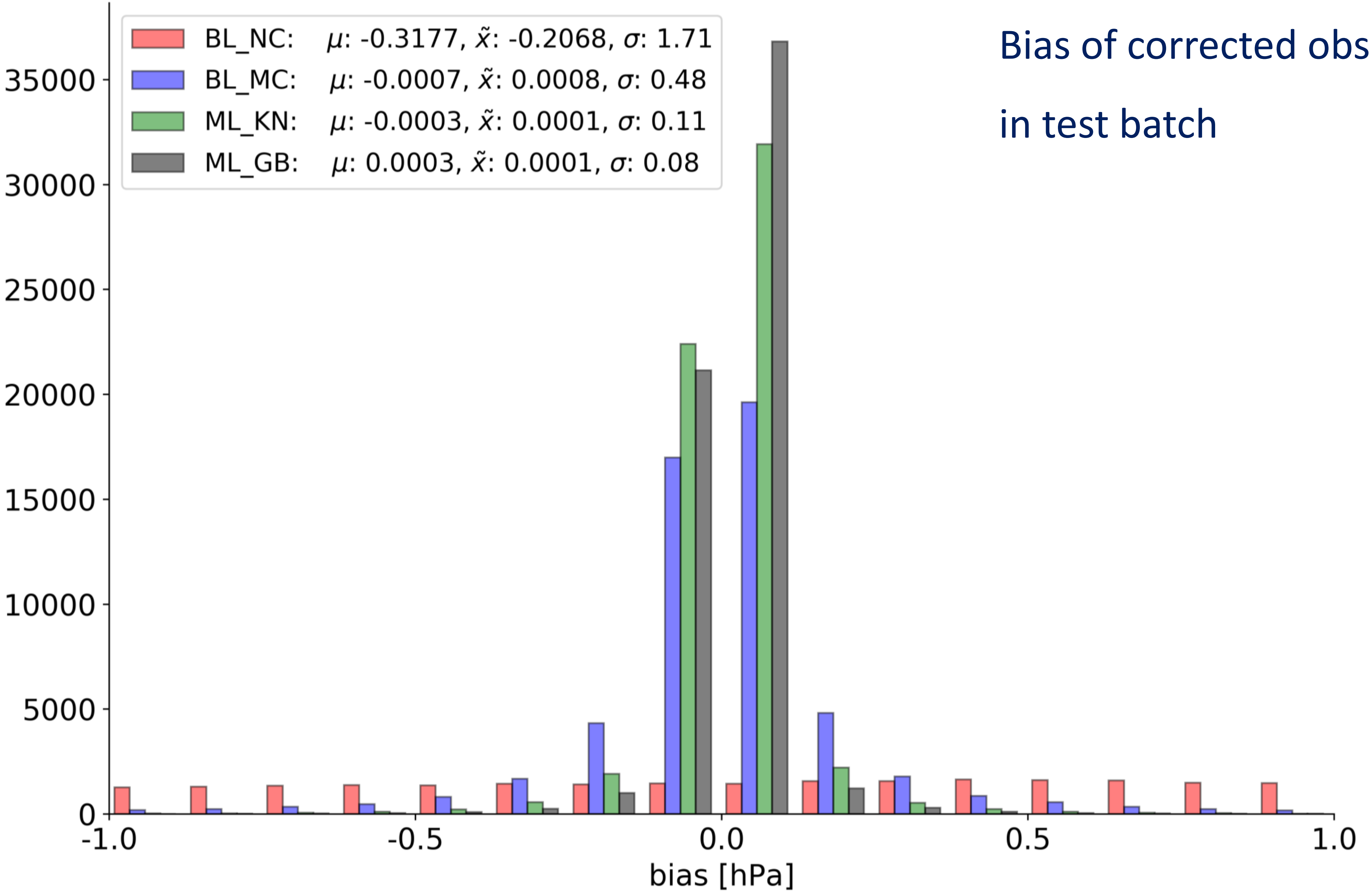
(using a Gradient Boosting algorithm with Ps, latitude, longitude, σ_p and hour of the day)

Rows: Each row is 15 minutes



RMSE of corrected observations in test batch

Bias of corrected observations in test batch



Data Assimilation in the DMI HARMONIE NWP system

Dr Xiaohua Yang, DMI, sets up an objective of the DMI HARMONIE NWP system.

Data assimilation for very high resolution is a necessary capability for NWS to forecast local, small scale weather for extremes.

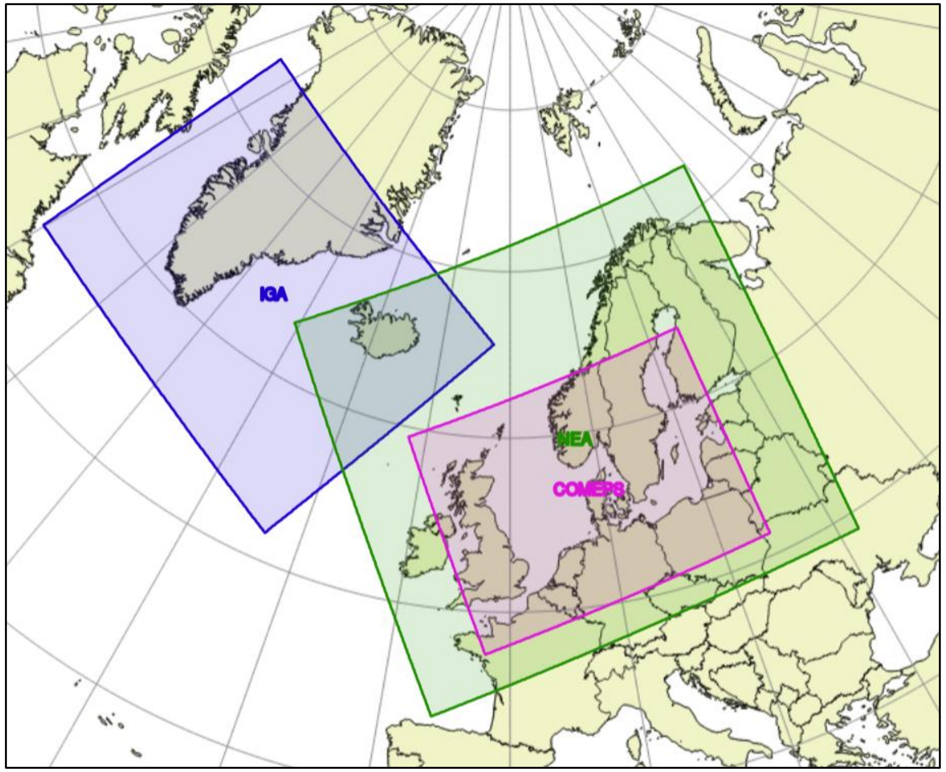
- Most nowcasting/RUC applications are targeted for here and now weather (Aviation, airport, special/local events, automatic/mobile weather
- Many extreme events are associated with small scales and limited predictability (Small spatial and well as temporal scale)

For some of the high impact weather (summer convection...), the time between first observed phenomena and finish of it are within a few hours.

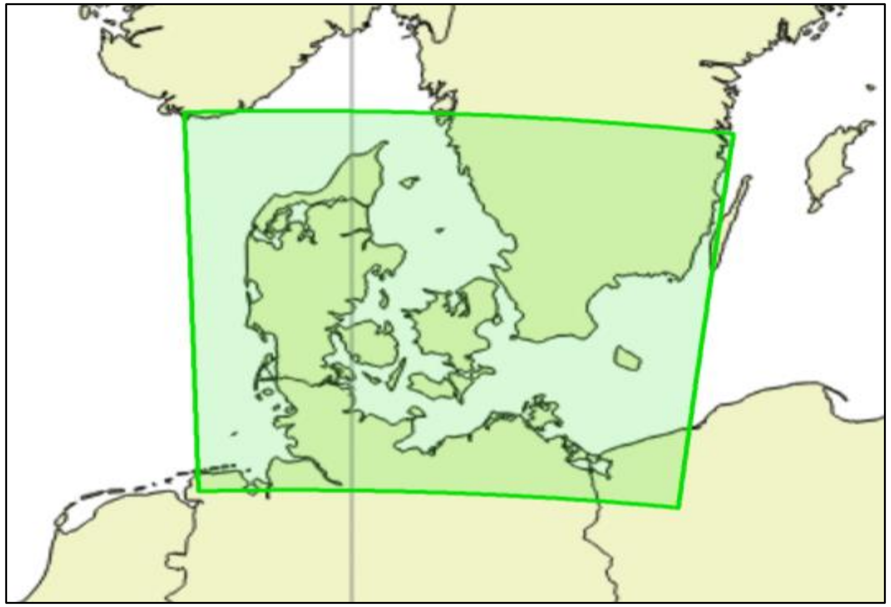
For NWP end-users, <u>timeliness</u> and <u>consistency with observations</u> are part of quality indicators.

*Harmonie-nowcasting shall go **sub-km grid scale.***

“HARMONIE-lite”



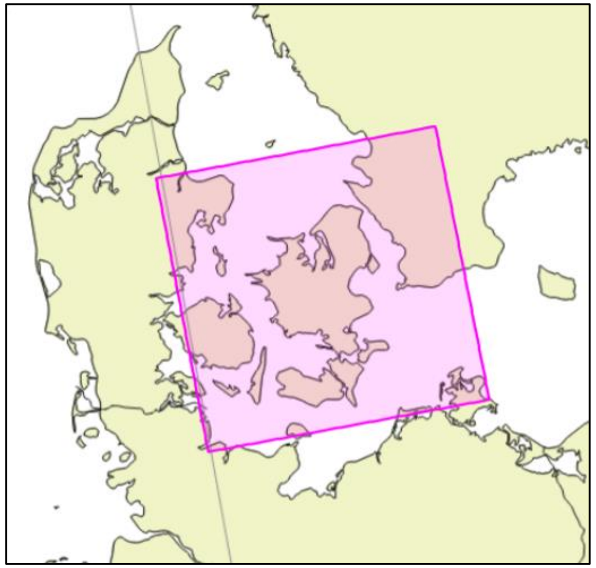
NEA, IGA and COMEPS operational domains.

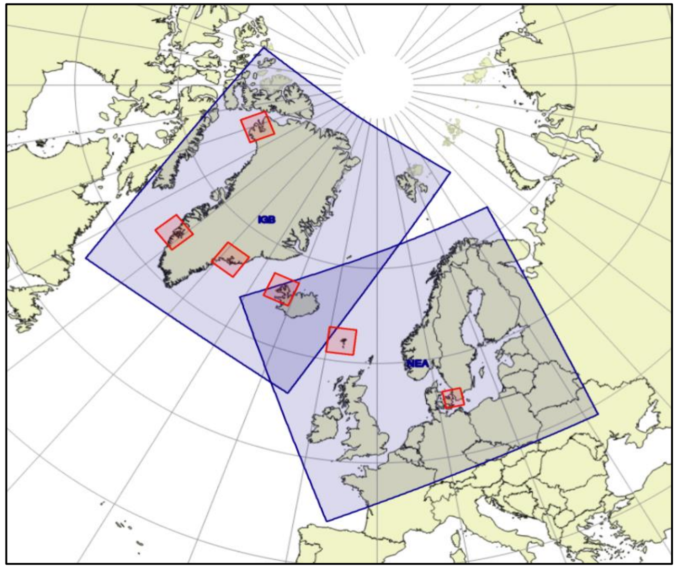


DK750 Model domain
Future Nowcasting Model

DA in same resolution as the model.

DK500 Model domain
(Test setup)



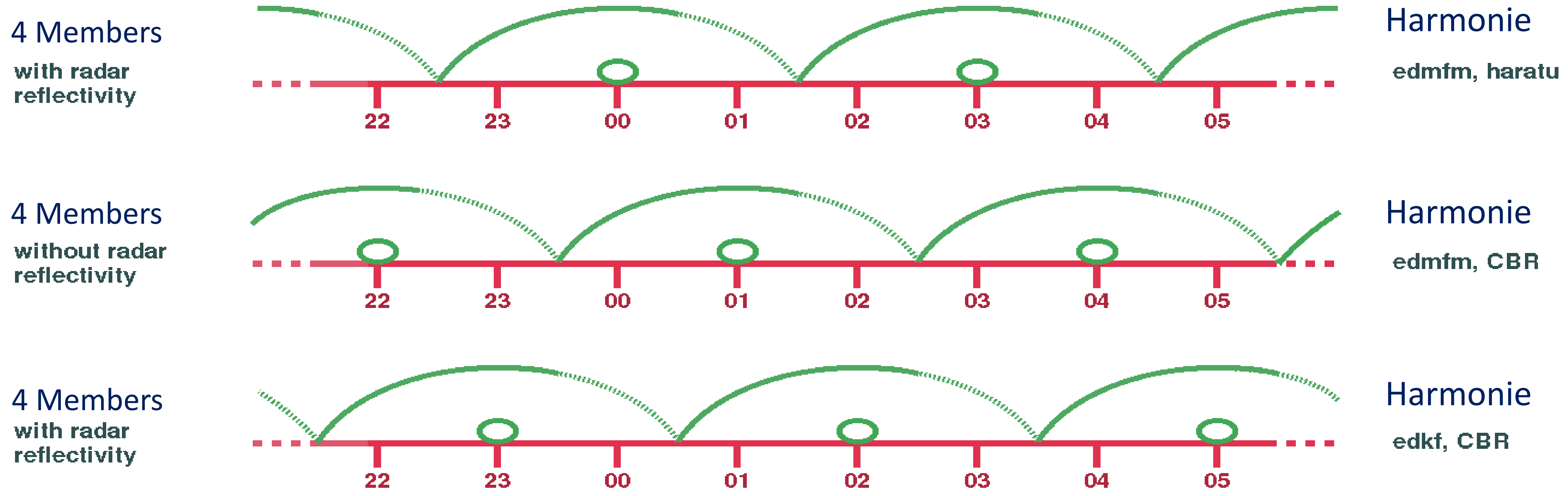


Observation data used for short range forecasts at DMI

	Production frequency	Delay time between last observations and forecast at DMI
Radar advection:	every 10 min	25 min
Nowcasting, Harmonie	every 1 h	30 min
COMEPS (EPS)	every 1h	2h 15 min
COMEPS Nowcasting, targeted	every 10 min	35 min
NEA (Operational Harmonie)	every 3h	2h30m - 4h30 min
IFS (ECMWF)	every 6h	3h45m - 8h:45 min

Slide credits: **Xiaohua Yang, DMI**

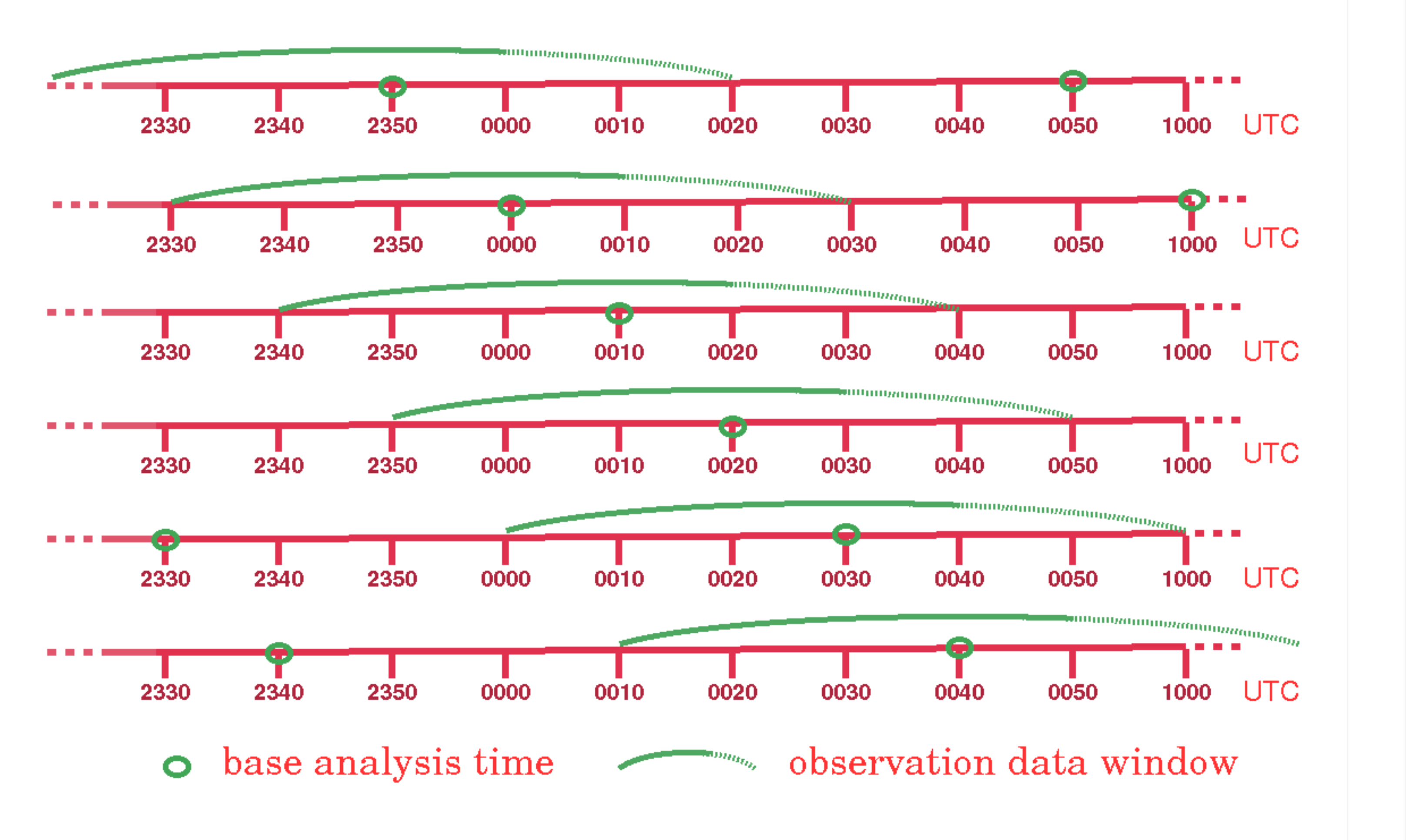
DMI-COMEPS: Frequent Analysis with Overlapping Windows



COMEPS = (3DVAR control on 3h window each hour: 4 perturbed members each hour) 6 hours = 24 perturbed members assembled each hour.

DMI COMEPS: Xiaohua Yang and Henrik Feddersen, DMI

DMI COMEPS Nowcasting product: Frequent Analysis with Overlapping Windows



Not yet operational

Harmonie-Lite (750m): new forecast every 10 minutes.

Each suite (rows) uses different observation batches in 3DVar.

SPOs (and other crowdsourced data) can enter one or more of these observation bathes.

Conclusions and recap

Motivation

- Observations are essential for NWP and nowcasting capability.
- Crowdsourced data has both advantages and disadvantages over conventional data (e.g., low-cost observations, but poor quality).
- There exist many sources of data, only a few have been mentioned here.

Pressure from smartphones

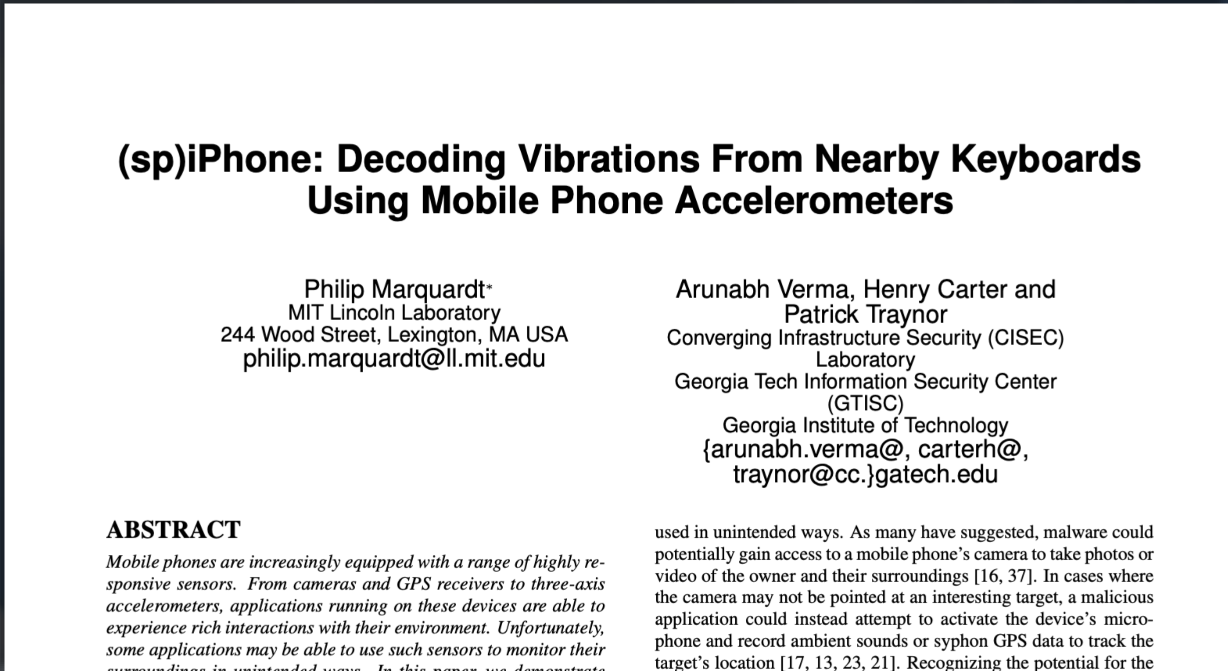
- Data collection has been very successful. Screening works but can be improved, for example, with ML algorithms as shown.
- SPOs have been successfully integrated into the DMI HARMONIE setup.

Remarks

- Third party data can be problematic because it is not always known how data have been processed before delivery.
- More experimentation with data assimilation is needed, to fully utilize these observations (ongoing).

Conclusions on privacy issues

- A unique user ID is used to bias-correct observations from each sensor. There are possible methods to solve this but not yet implemented.
- Legal issues must be considered, such as the GDPR act from the European Union. User consent is not the only way forward.
- These issues make it difficult to share data between scientists (e.g. Price et al. 2018).



(sp)iPhone: Decoding Vibrations From Nearby Keyboards Using Mobile Phone Accelerometers

Philip Marquardt*
MIT Lincoln Laboratory
244 Wood Street, Lexington, MA USA
philip.marquardt@ll.mit.edu

Arunabh Verma, Henry Carter and Patrick Traynor
Converging Infrastructure Security (CISEC) Laboratory
Georgia Tech Information Security Center (GTISC)
Georgia Institute of Technology
{arunabh.verma@, carterh@, traynor@cc.}gatech.edu

ABSTRACT

Mobile phones are increasingly equipped with a range of highly responsive sensors. From cameras and GPS receivers to three-axis accelerometers, applications running on these devices are able to experience rich interactions with their environment. Unfortunately, some applications may be able to use such sensors to monitor their

used in unintended ways. As many have suggested, malware could potentially gain access to a mobile phone's camera to take photos or video of the owner and their surroundings [16, 37]. In cases where the camera may not be pointed at an interesting target, a malicious application could instead attempt to activate the device's microphone and record ambient sounds or syphon GPS data to track the target's location [17, 13, 23, 21]. Recognizing the potential for the