Koninklijk Meteorologisch Instituut

Institut Royal Météorologique

Königliches Meteorologisches Institut

Royal Meteorological Institute

Citizen observations via smartphone in Belgium: data collection and applications

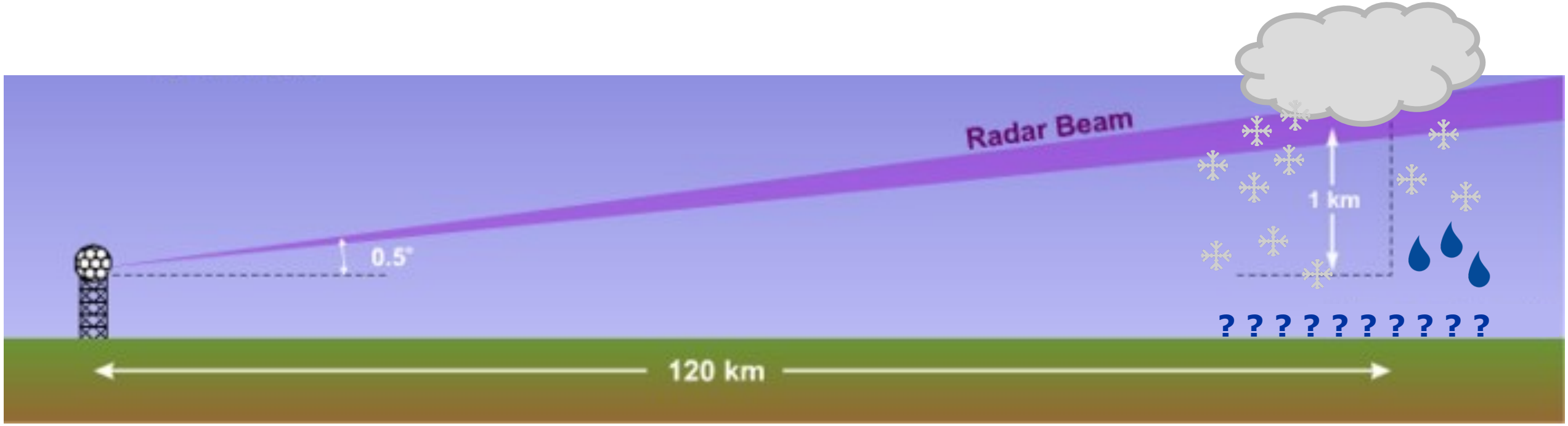
Maarten Reyniers, Laurent Delobbe and Sylvain Watelet

ECSS2023 09.05.2023

Motivation: why citizen reports?

Provide very local meteorological information in addition to classical instruments

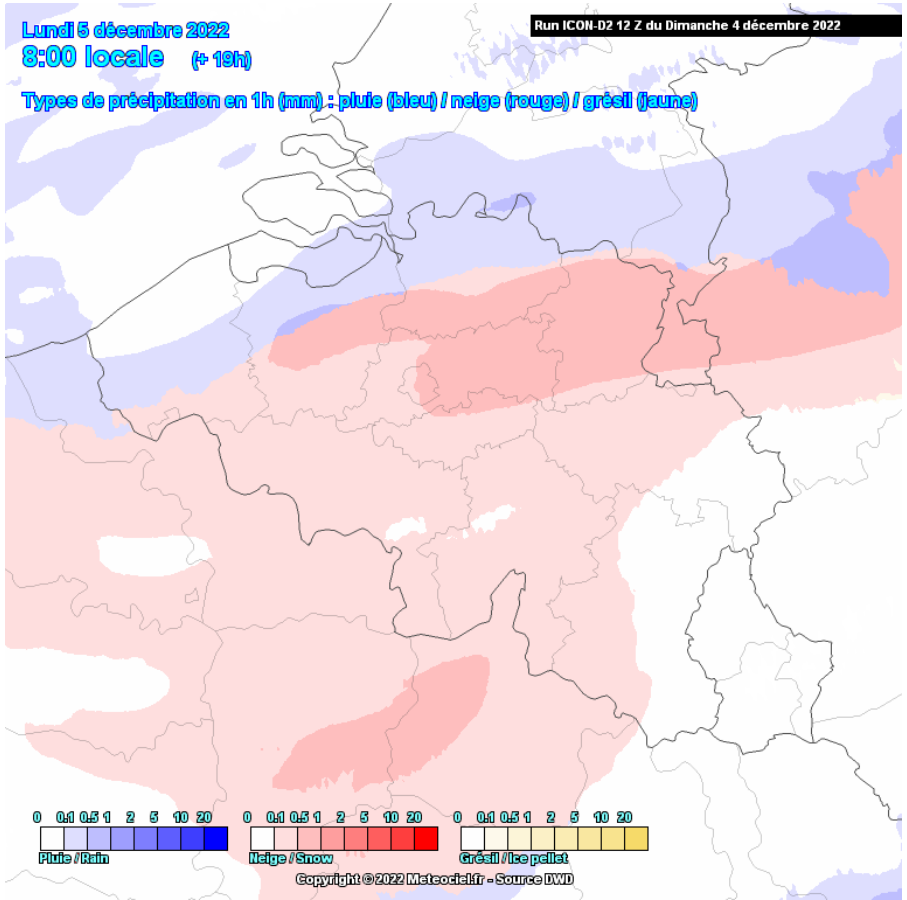
- Verification and improvement of existing nowcasting/forecasting schemes
- Use in daily operations (e.g. weather office)

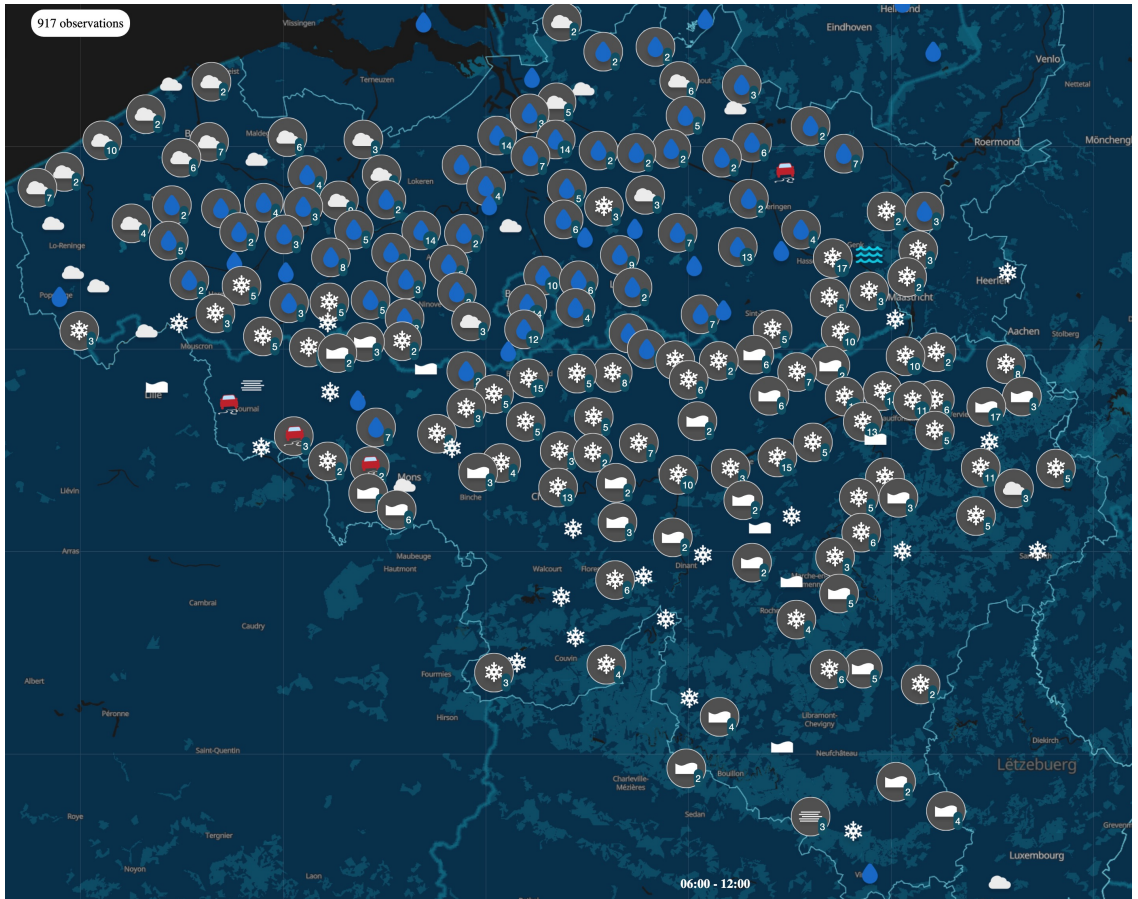


Motivation: why citizen reports?

Provide very local meteorological information in addition to classical instruments

- Verification and improvement of existing nowcasting/forecasting schemes
- Use in daily operations (e.g. weather office)

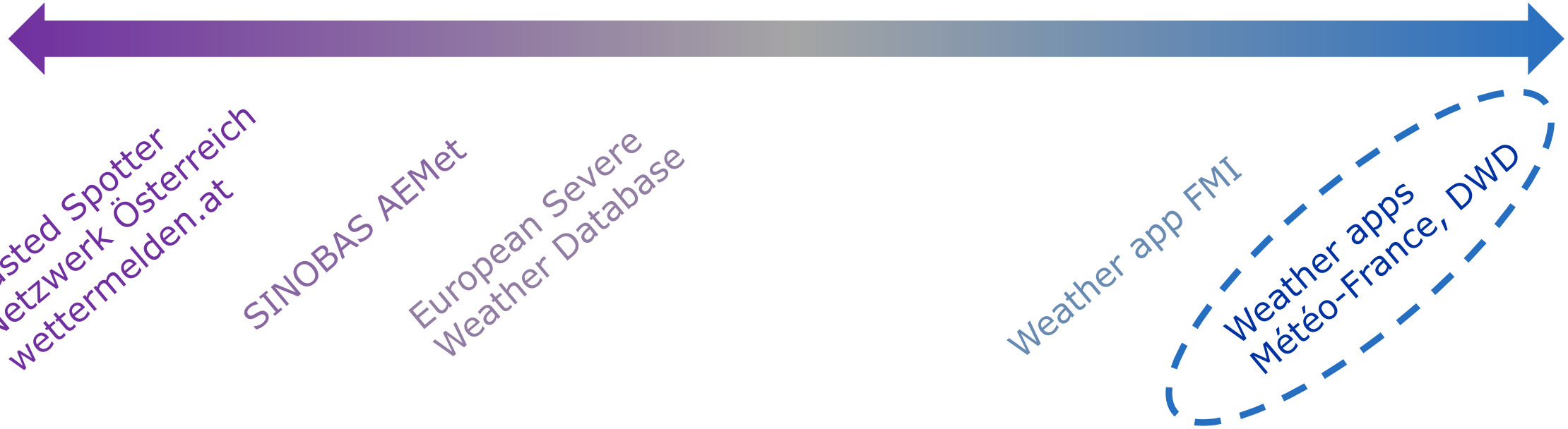




Citizen reports: two possible approaches

Network of trusted/trained observers
Detailed reports
Manual QC, centralised or by peers
Concentrate on high-impact events

Anonymous, open to everyone
Concise reports: only basic info
Automatic QC
All possible observations



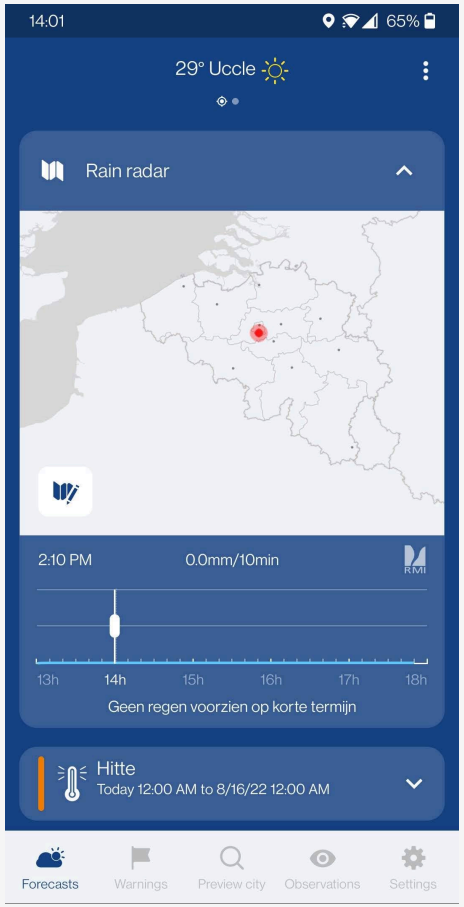
Overview

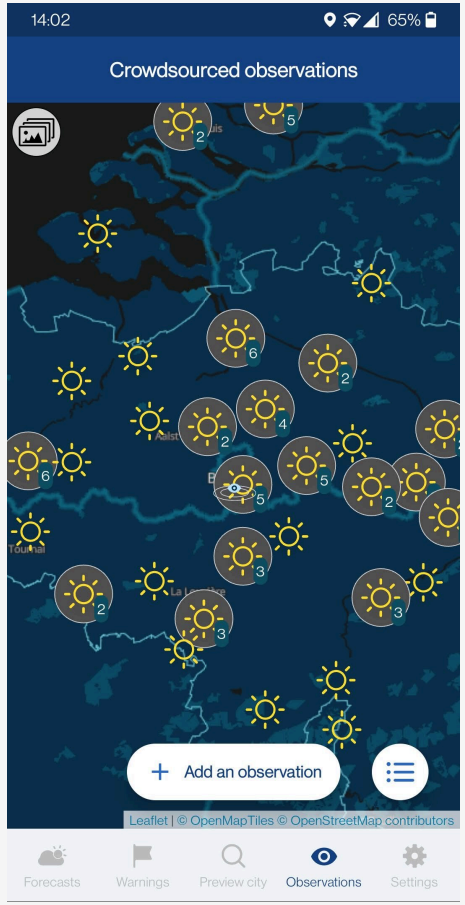
1. General concept
2. Quality control
3. Human biases
4. Use case: hail

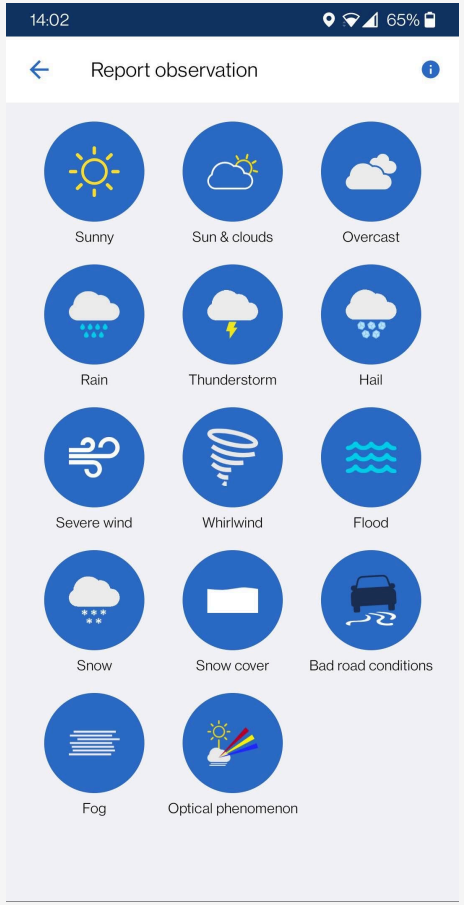
General concept

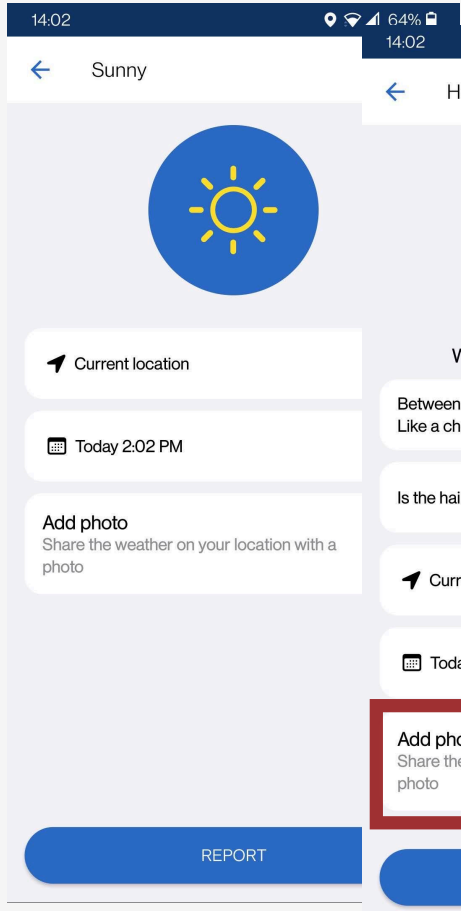
Smartphone app RMIB

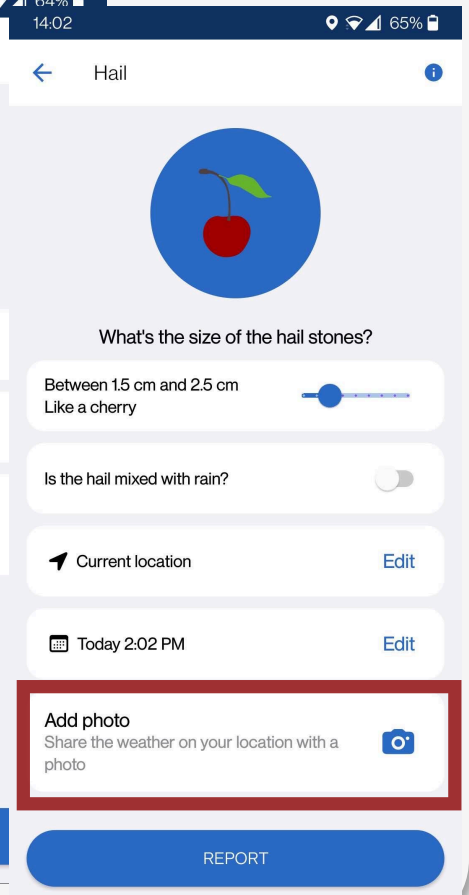
Submit a report in four clicks



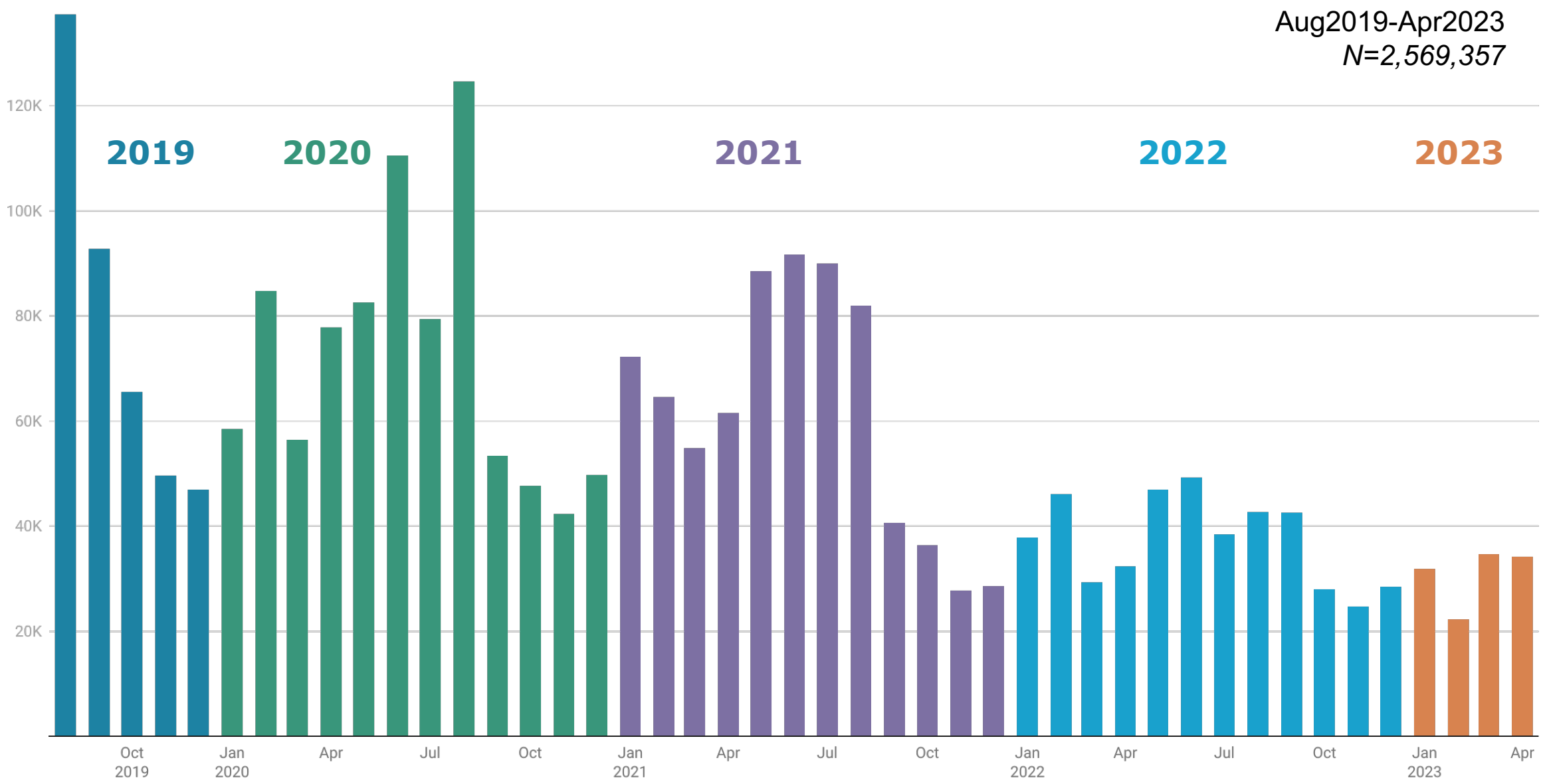




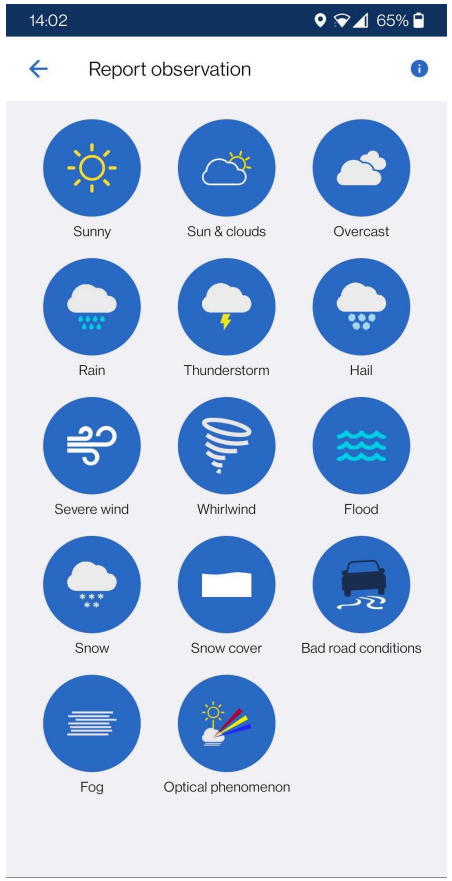




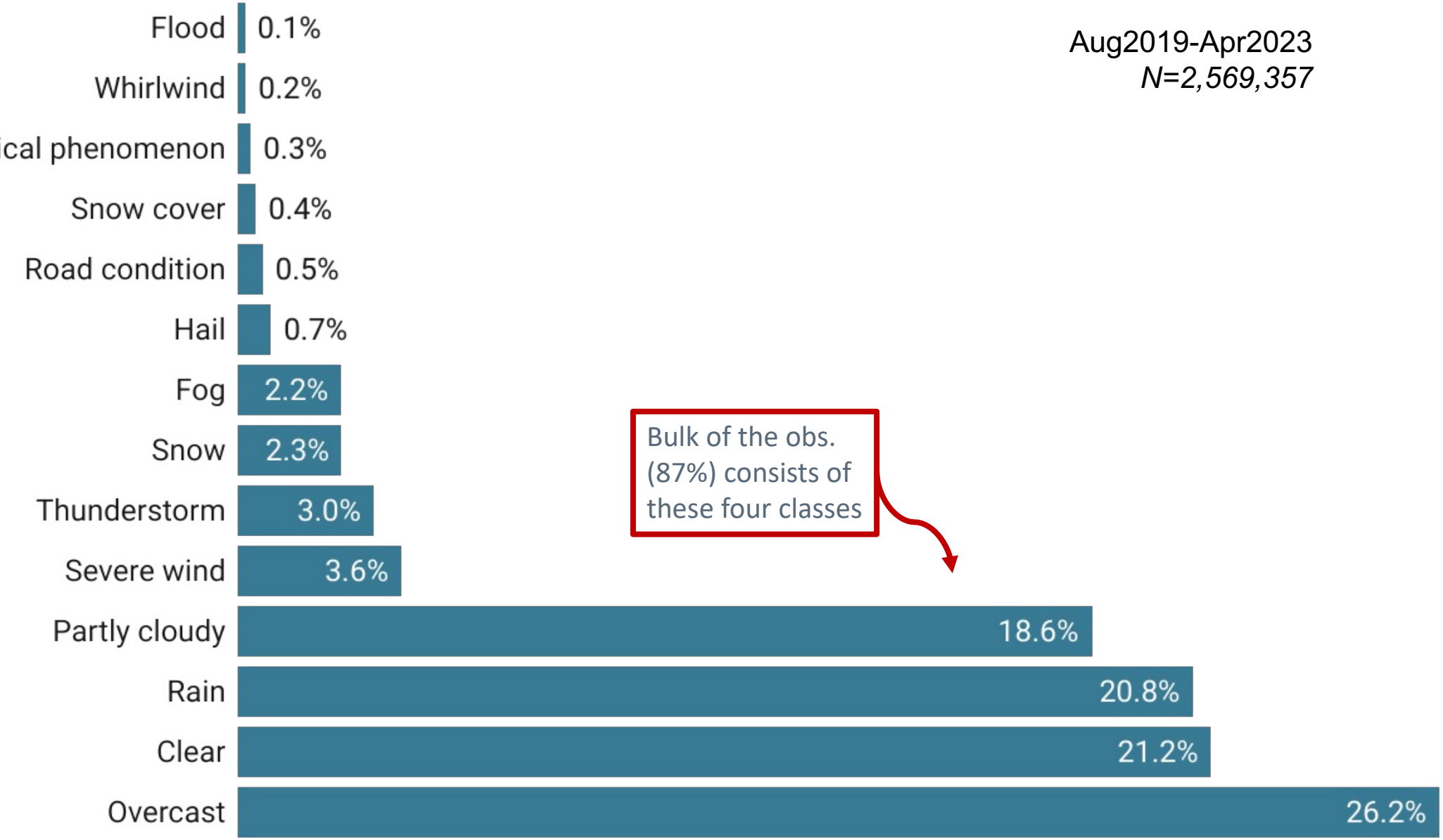
Total number of observations per month



Distribution over different phenomena



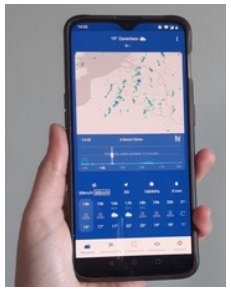
Aug2019-Apr2023
N=2,569,357



Quality control

QC: plausibility check

Every incoming observations is labeled with a quality flag (“sanity check”).



- Timestamp
- Location
- Observation

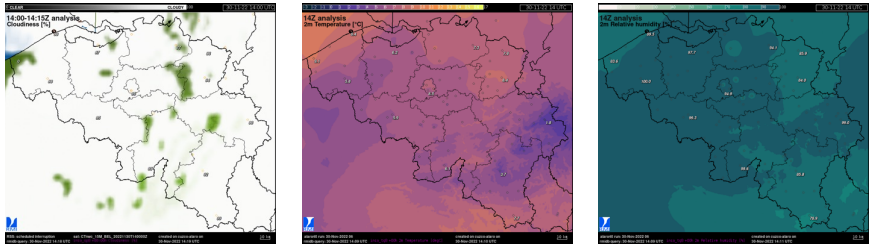
Observations

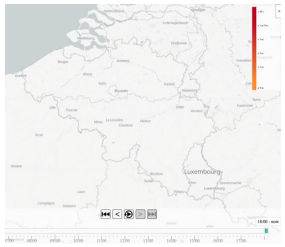
plausibility check

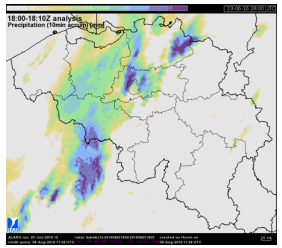
Discrete quantity; only three values:

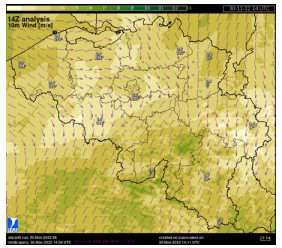
- **Plausible** (score=100)
- **Doubtful** (score=50)
- **False** (score=0)

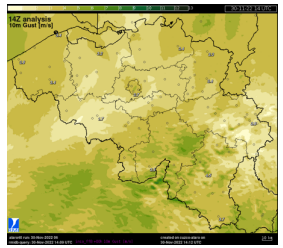
Nowcasting (INCA-BE) and **lightning**

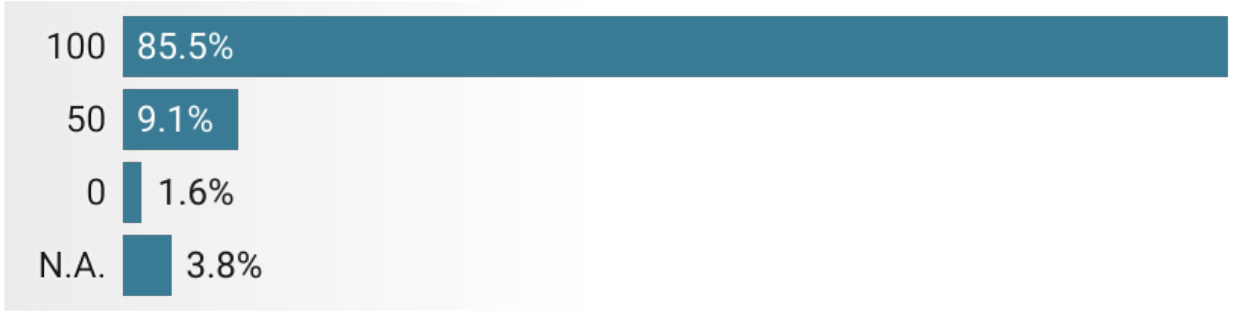








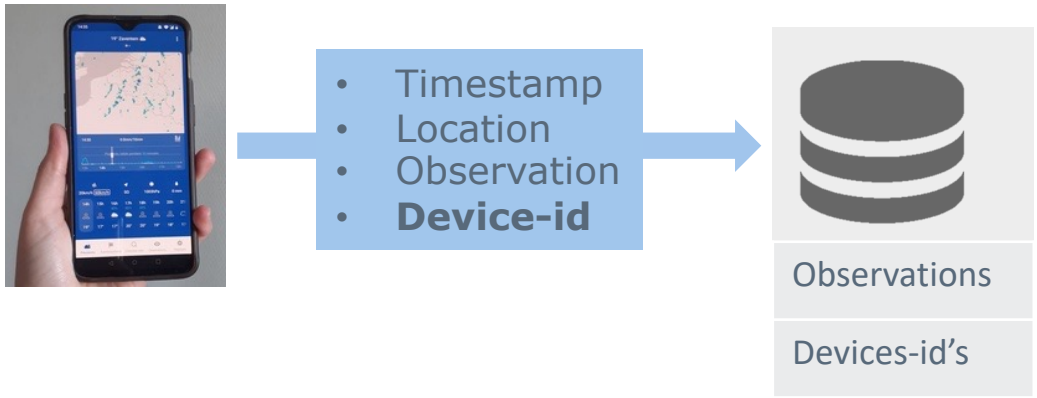




QC: user reputation

(~~User~~) Device reputation ≡

mean plausibility score of all observations received from that *particular device*



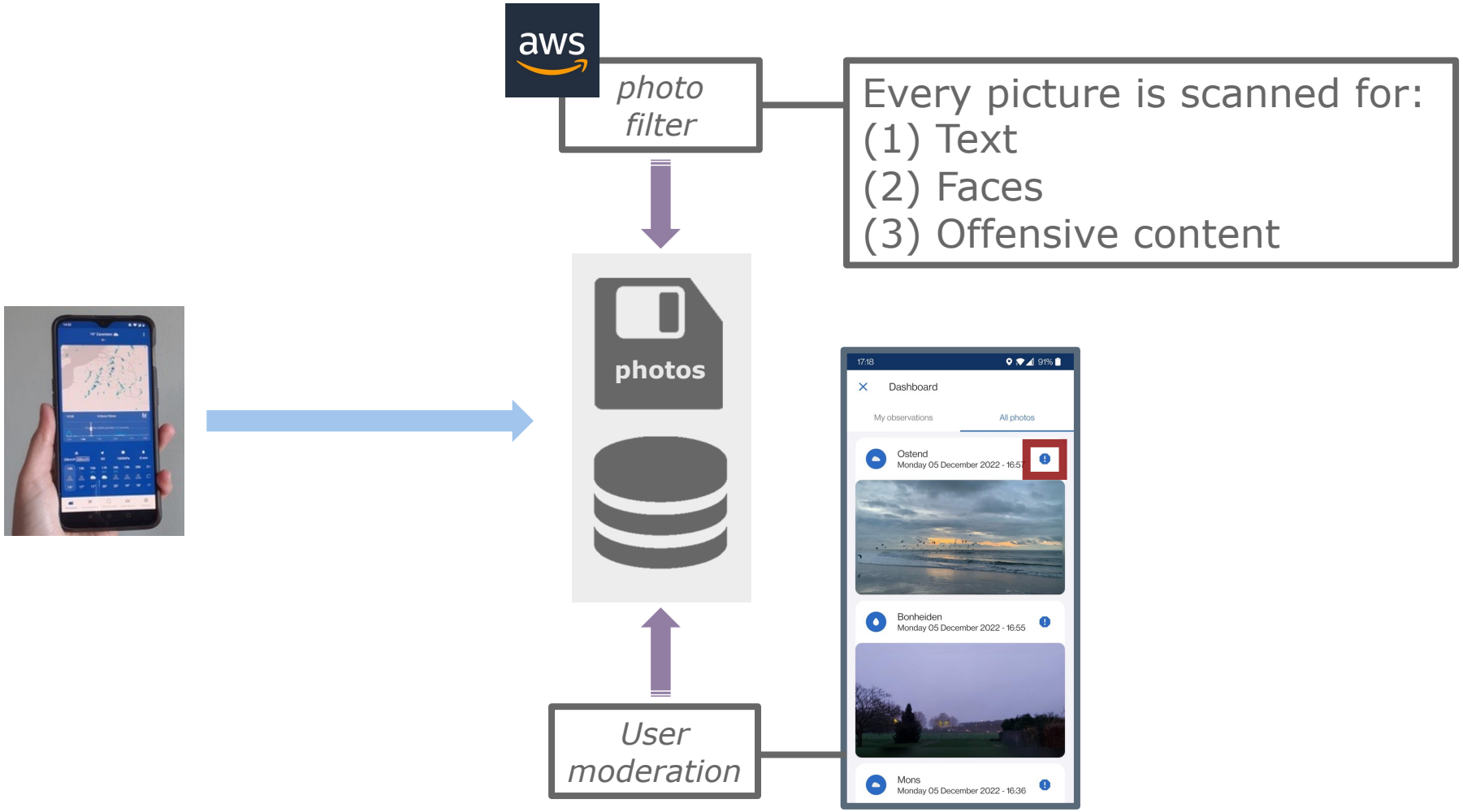
⇒ 0 ≤ user_reliability ≤ 100

78% of the users have a reputation score ≥ 90.
2.6% are trolls (score 0).

QC: photo filter

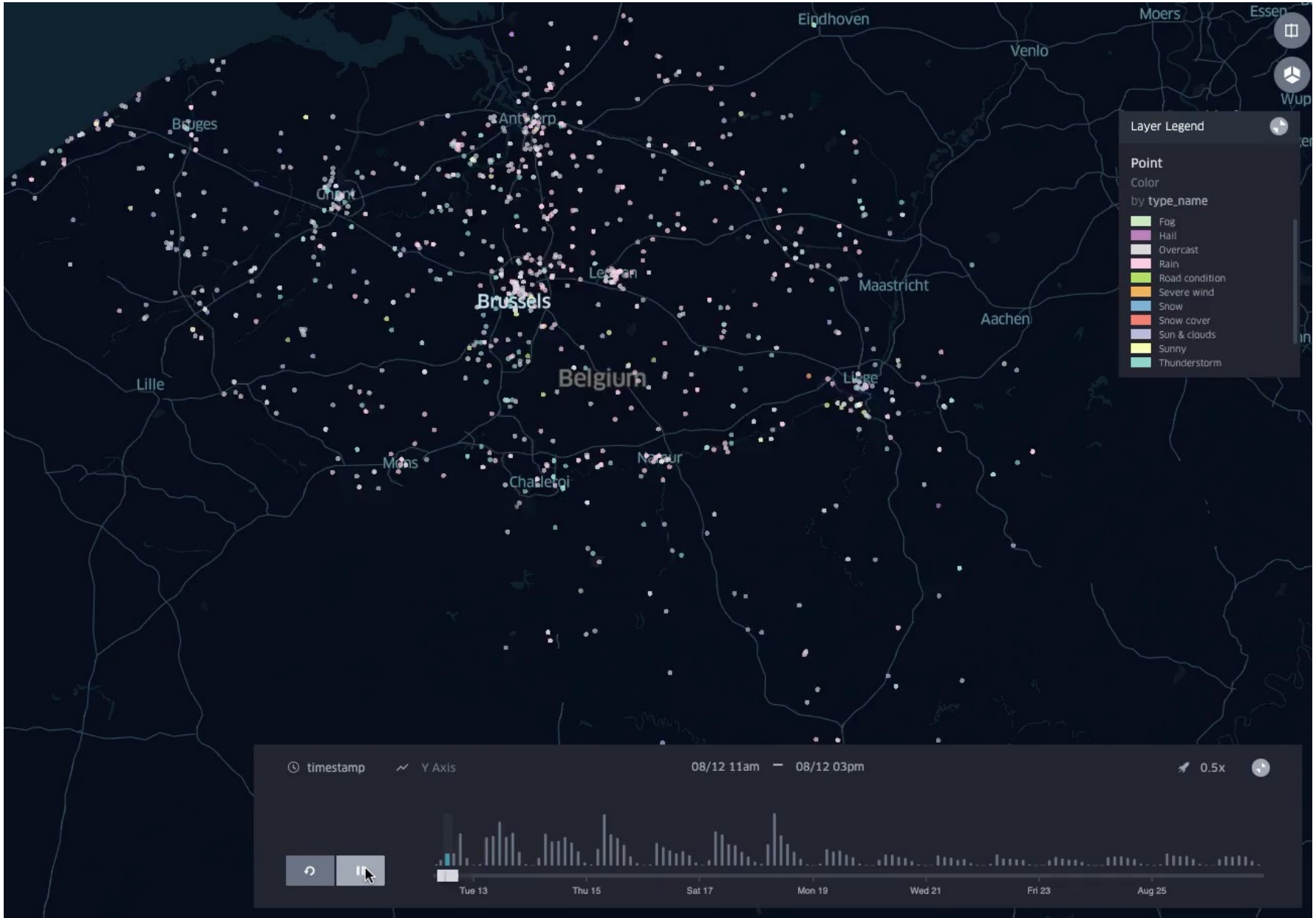
Goal: - To keep the RMIB app **free** from **non-weather content**

- To **protect** other people's **privacy** (faces, licence plates, …)



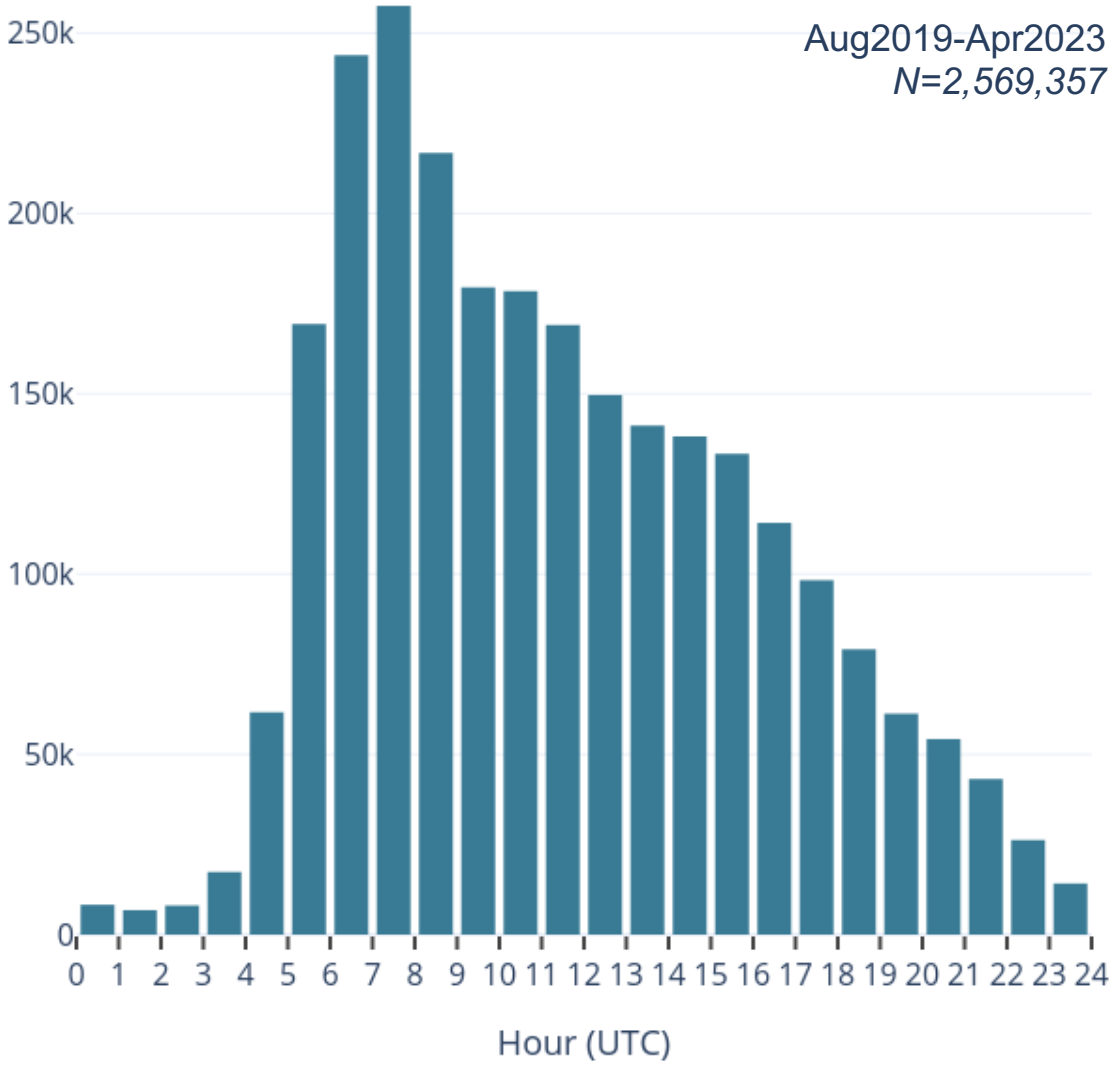
Human biases

(1) Human diurnal cycle



(1) Human diurnal cycle

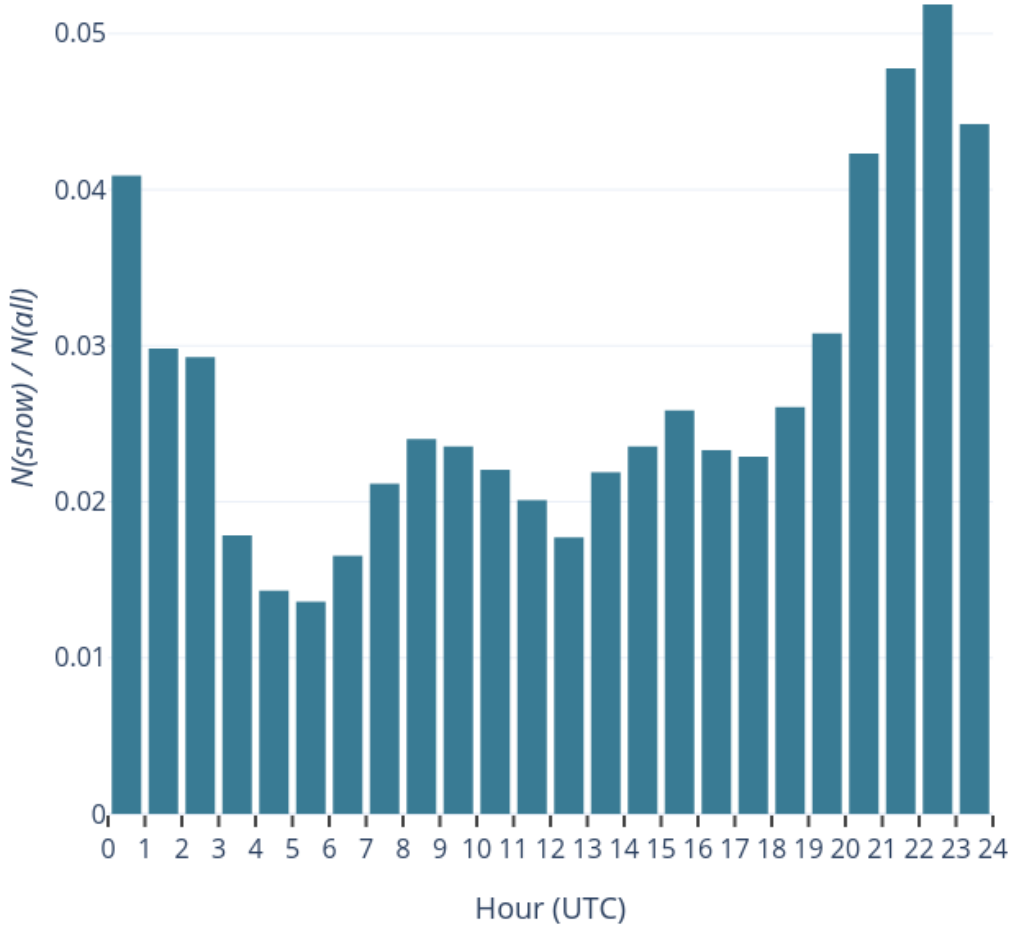
Number of observations aggregated to hourly bins



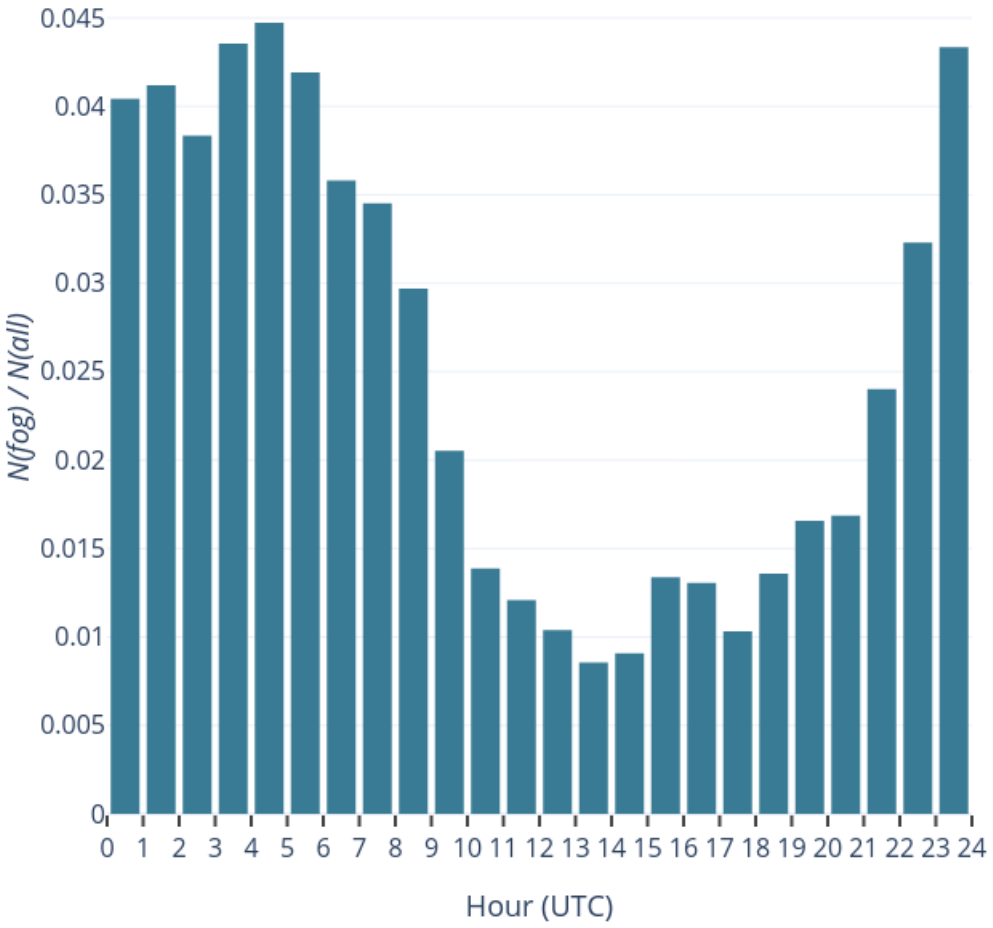
(1) Human diurnal cycle

Intrinsic diurnal cycle of specific **meteorological phenomenon** obtained by dividing by all observations

***snow**/all* *N(snow)=58,164*



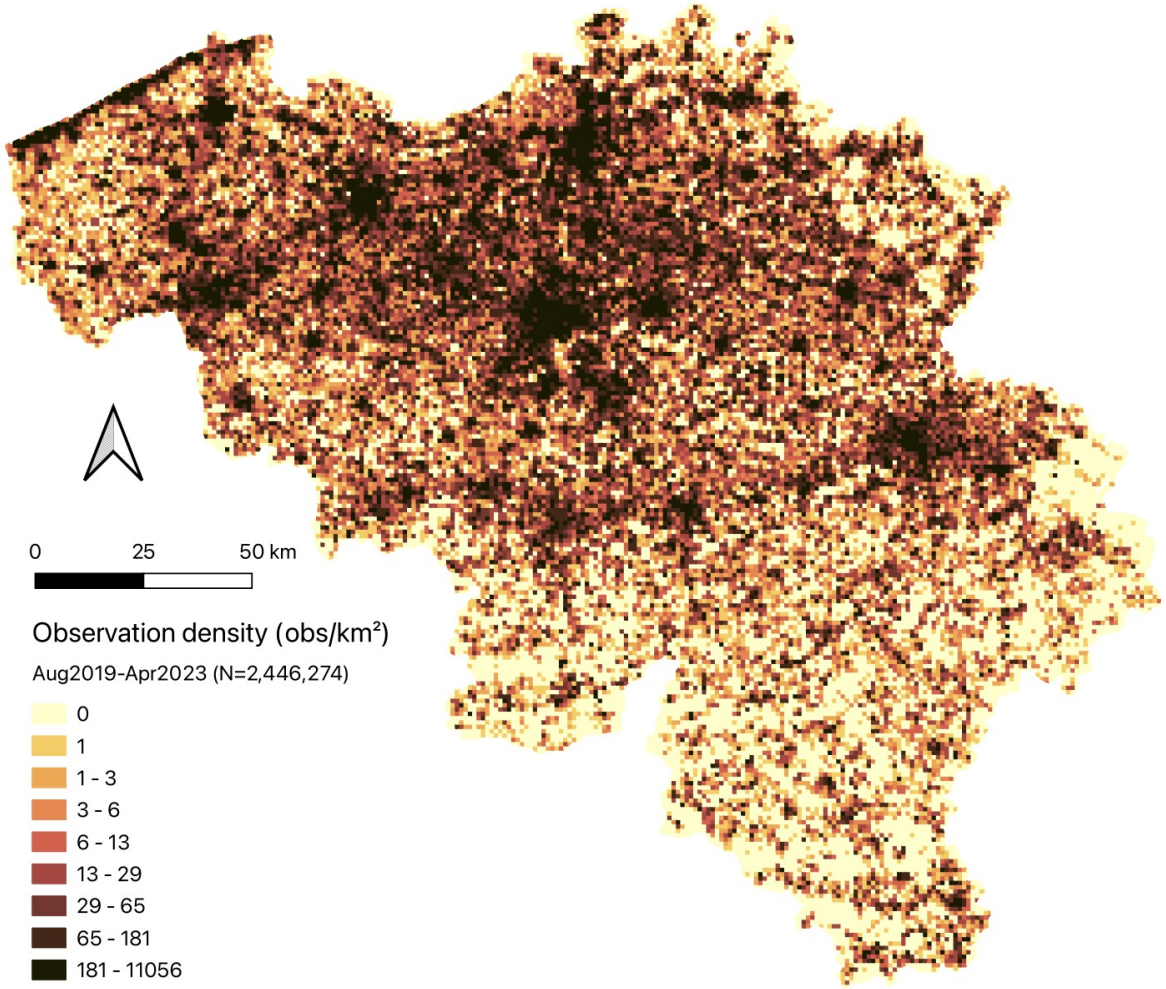
***fog**/all* *N(fog)=57,616*



(2) Population density bias

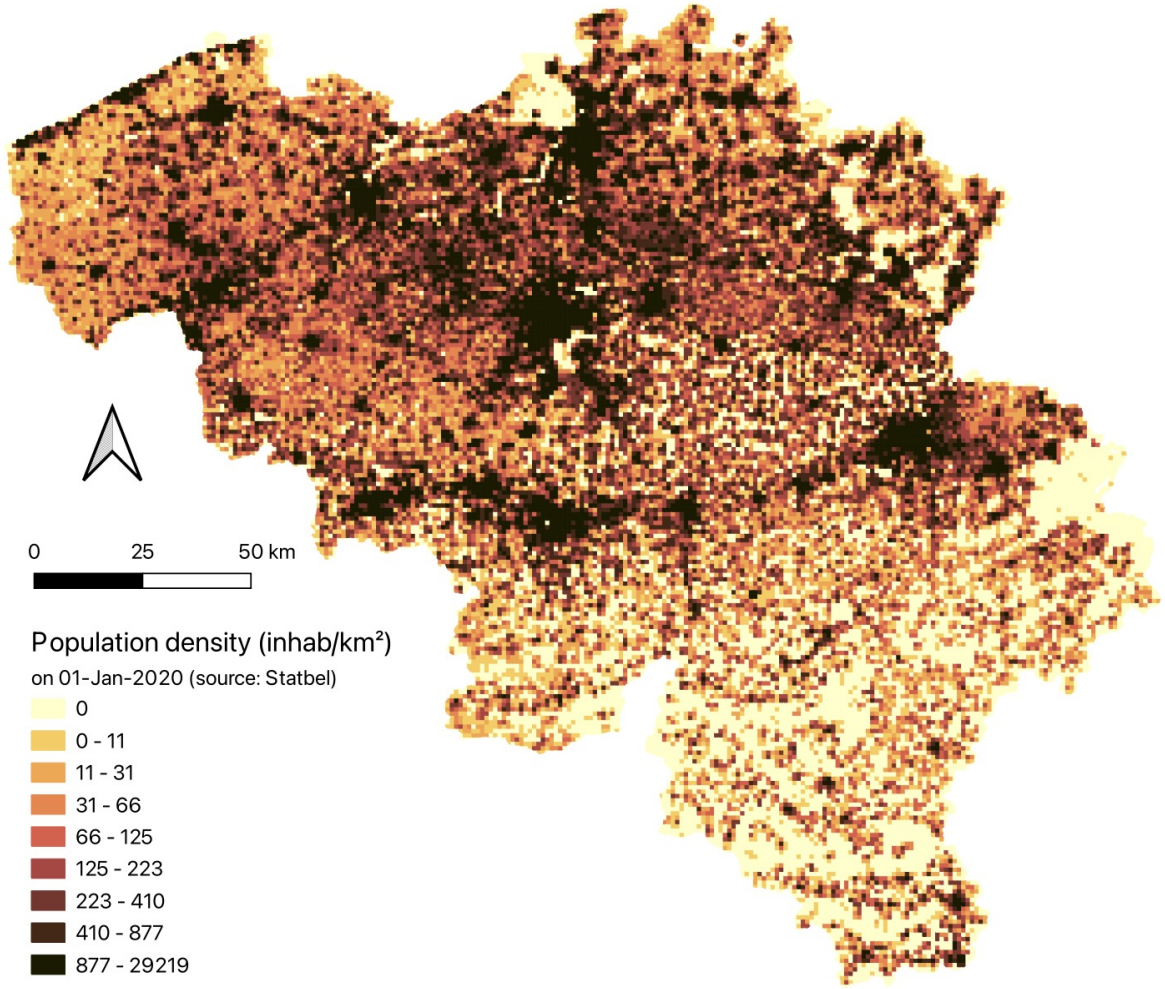
Observations density

Observations per km^2



Population density

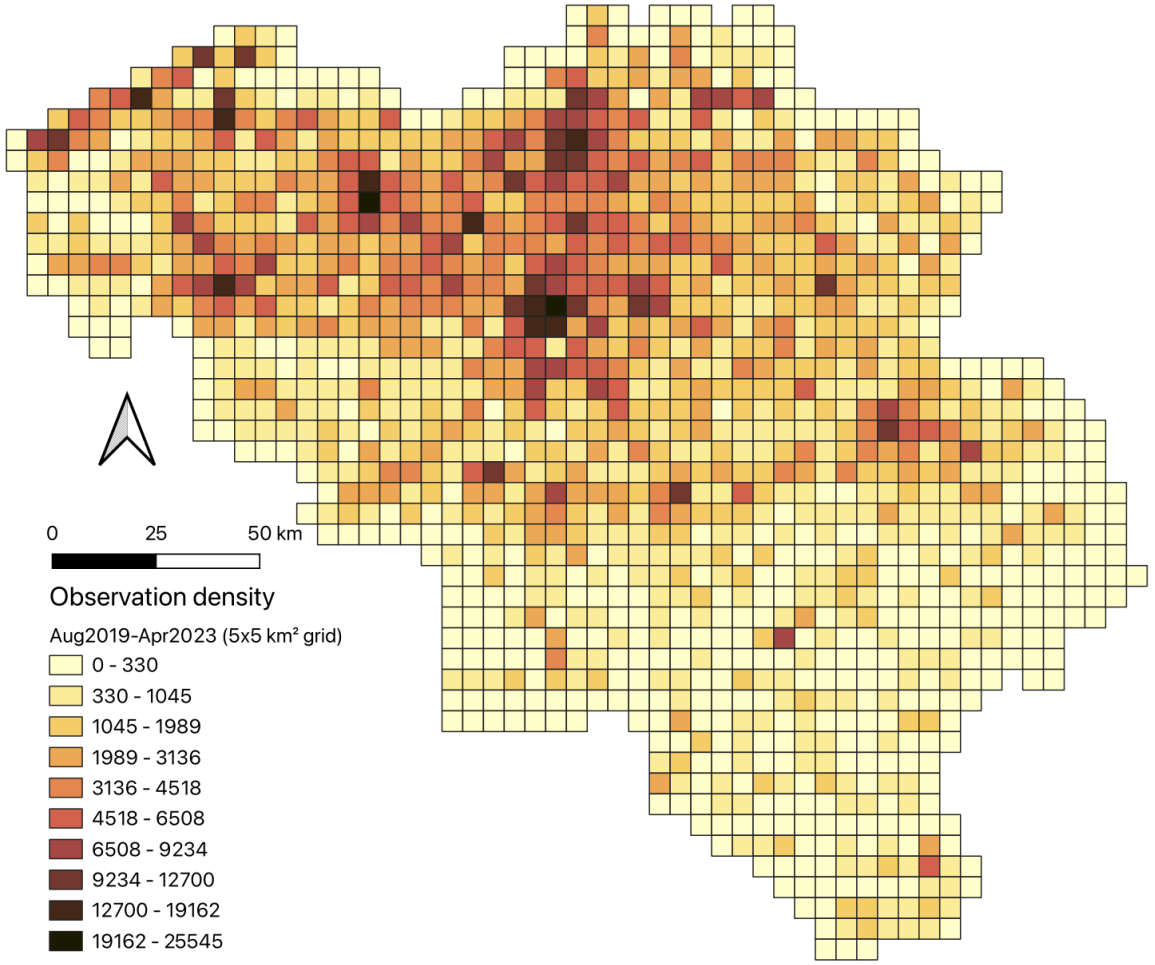
Inhabitants per km^2



(2) Population density bias

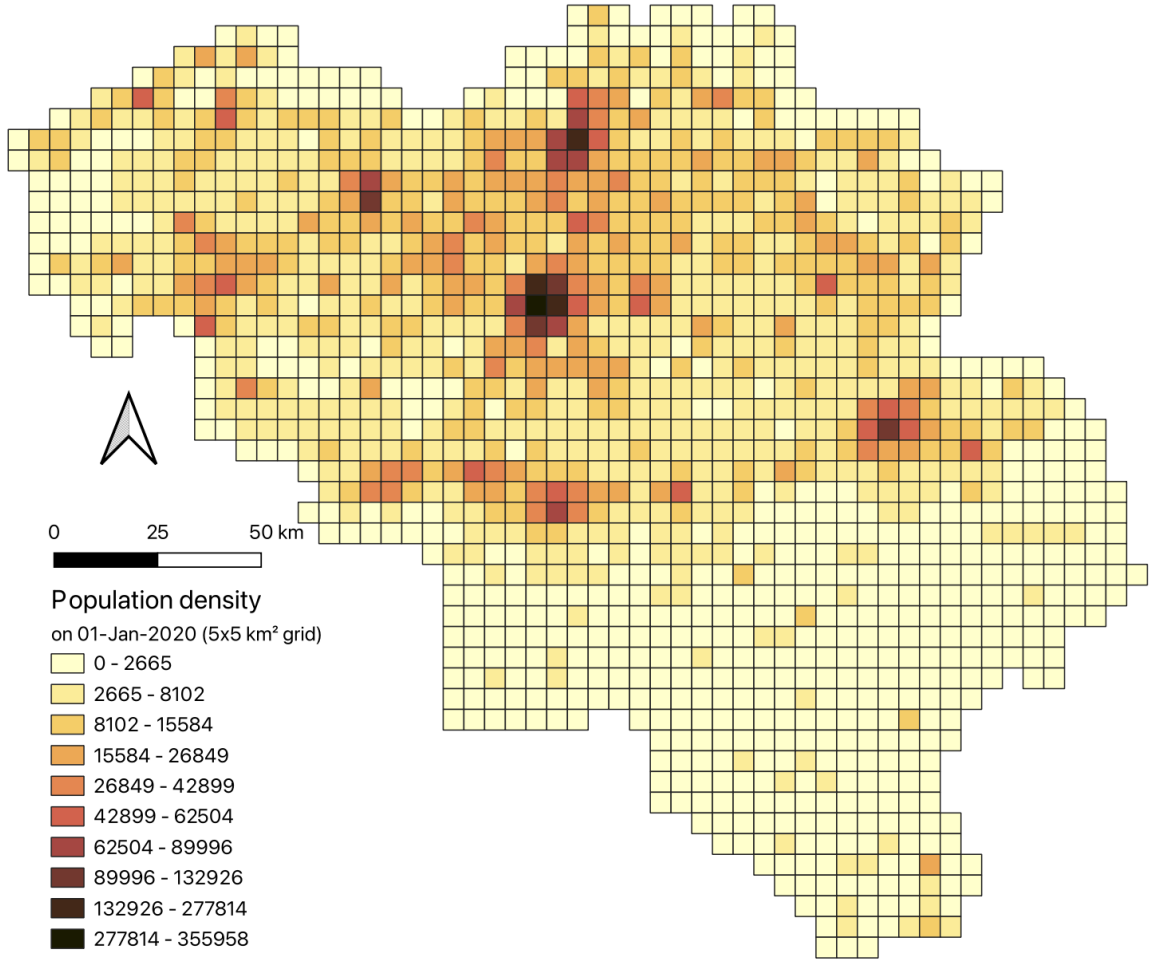
Observation density (5x5 km^2 grid)

Aug2019-Apr2023



Population density (5x5 km^2 grid)

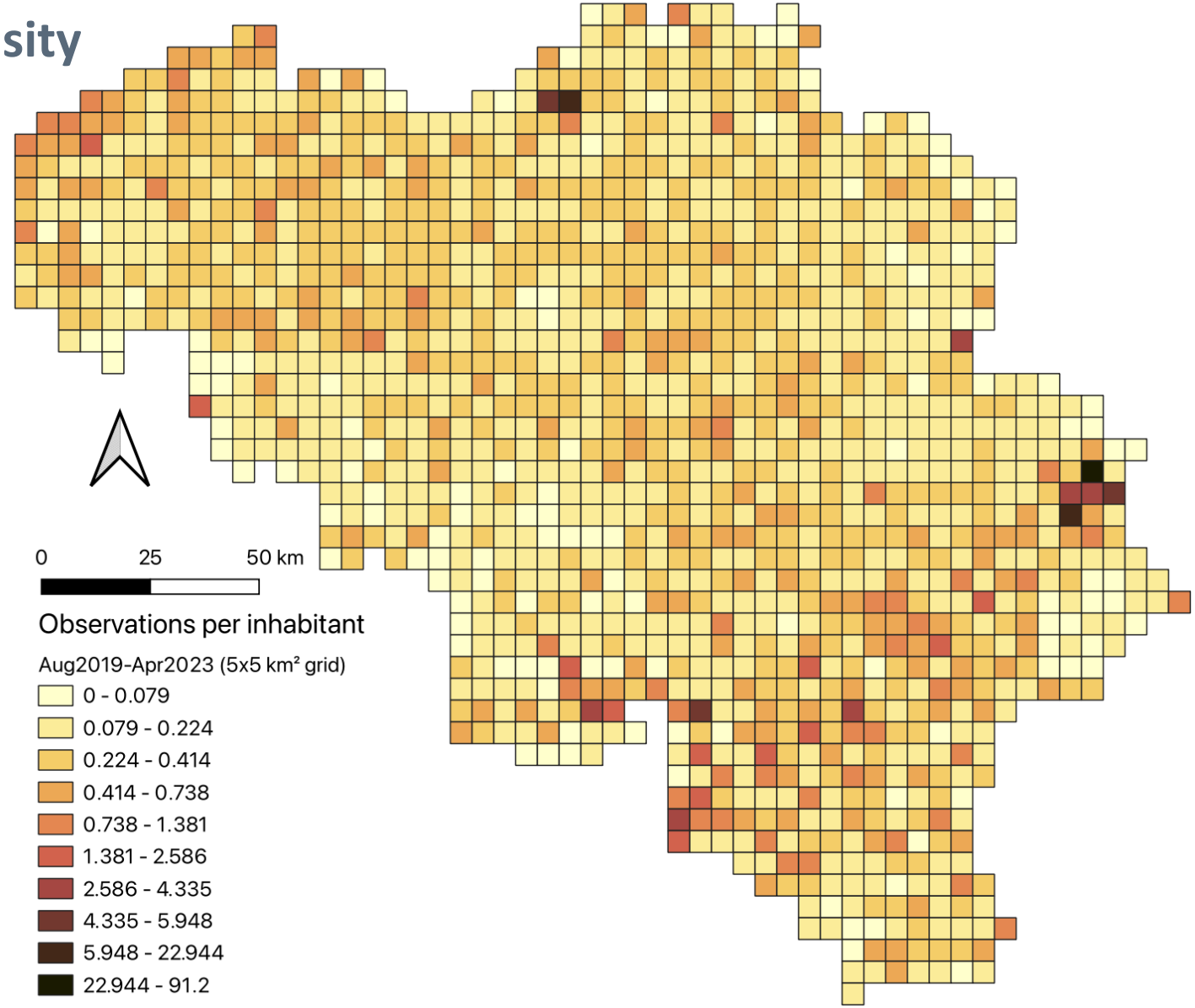
On 01-Jan-2020 (source: Statbel)



(2) Population density bias

“Normalised” observation density

Number of observations per inhabitant (5x5 km^2 grid)



(3) “Excitement” bias

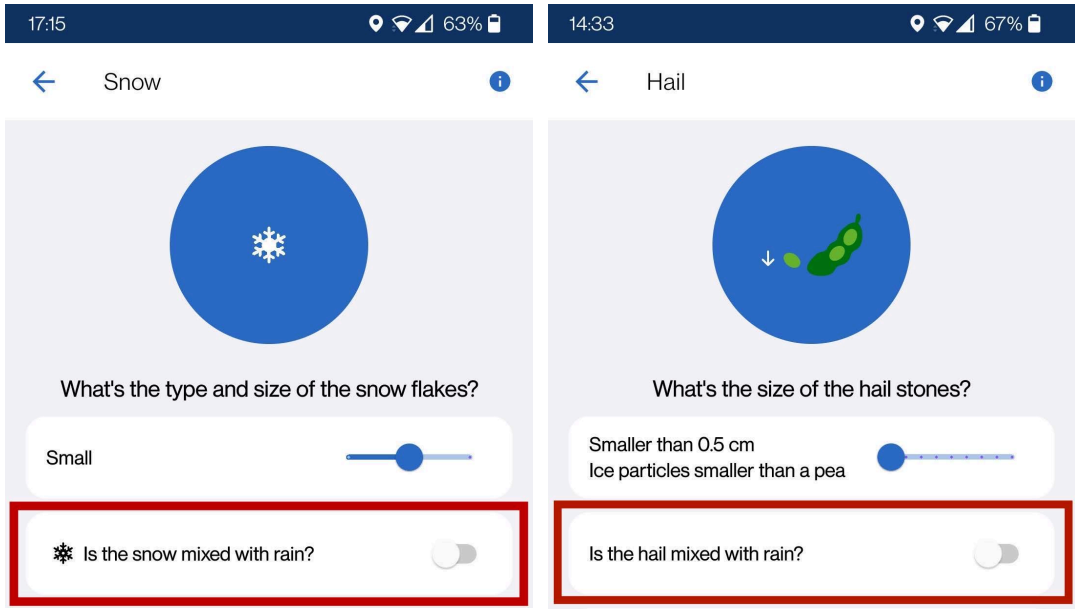
We assume that some types will be **proportionally more reported than others**, for example hail and snow more than rain.

E.g. 23% of the reports are precipitation-related, but it doesn’t rain 23% of the time (which is between 5% and 10%).

(4) In-app user behaviour bias

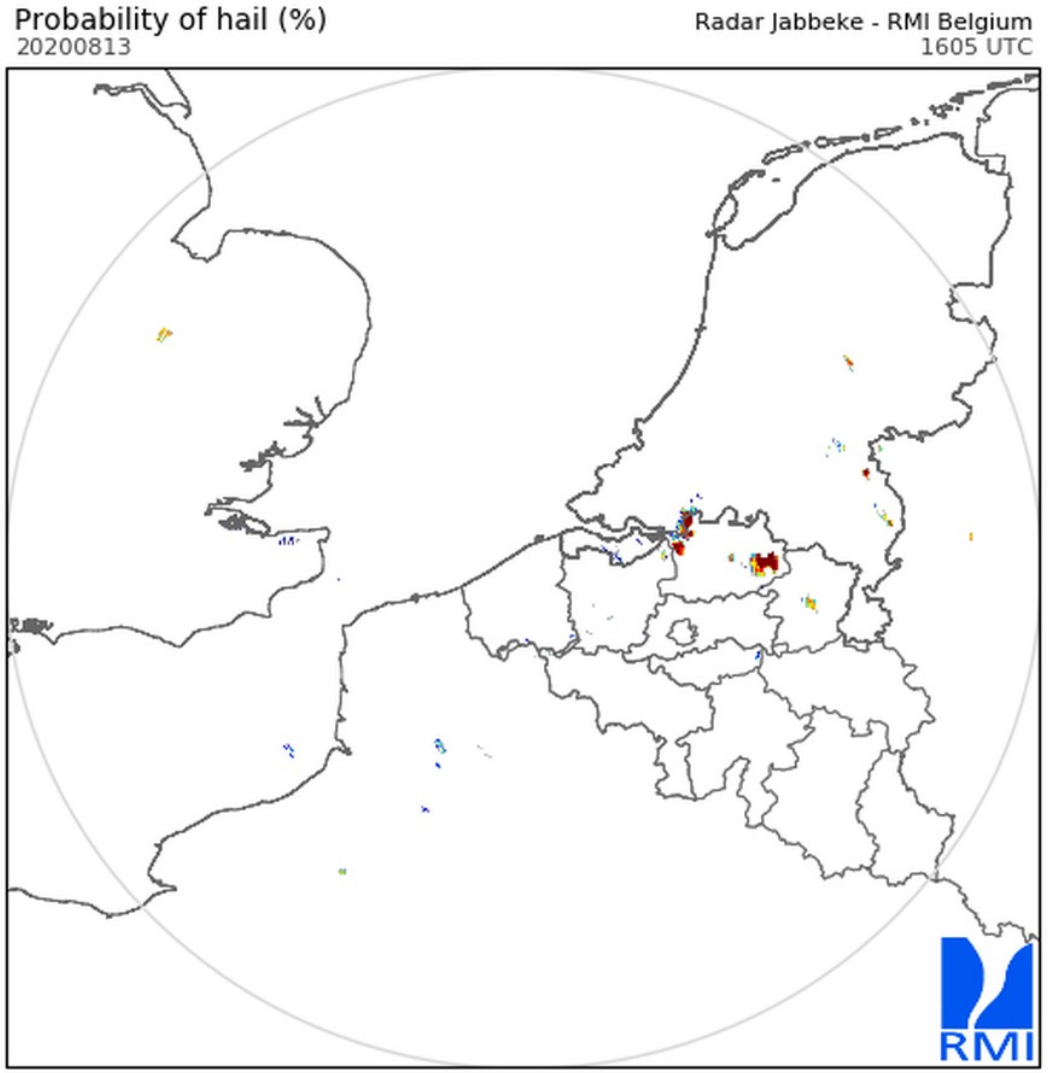
No information on **how users use the app**.

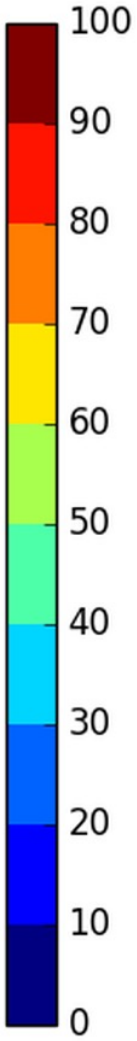
How much effort do users take for a correct observation? ⇒ overrepresentation of app default values.



Use case: hail

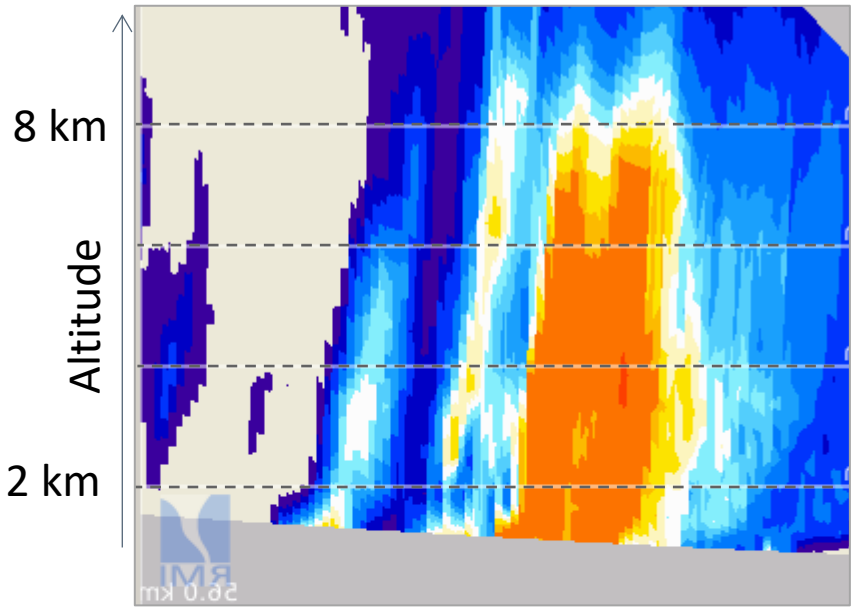
Hail product – Probability Of Hail





Waldvogel's hail detection algorithm

- Derived from difference (in km) between freezing level and Echotop-45dBZ
- Freezing level from RMI ALARO model



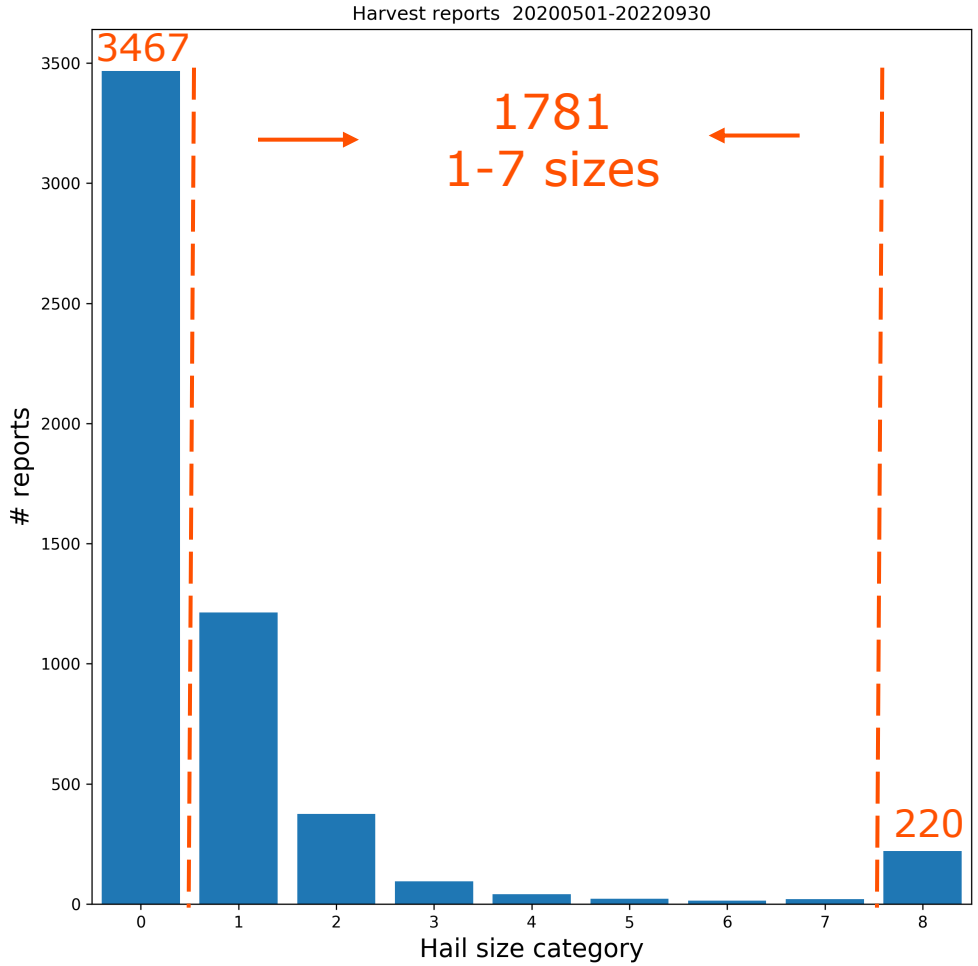
Hail reports May-Sept 2020+2021+2022

May-Sept 2020+2021+2022

5468 hail reports

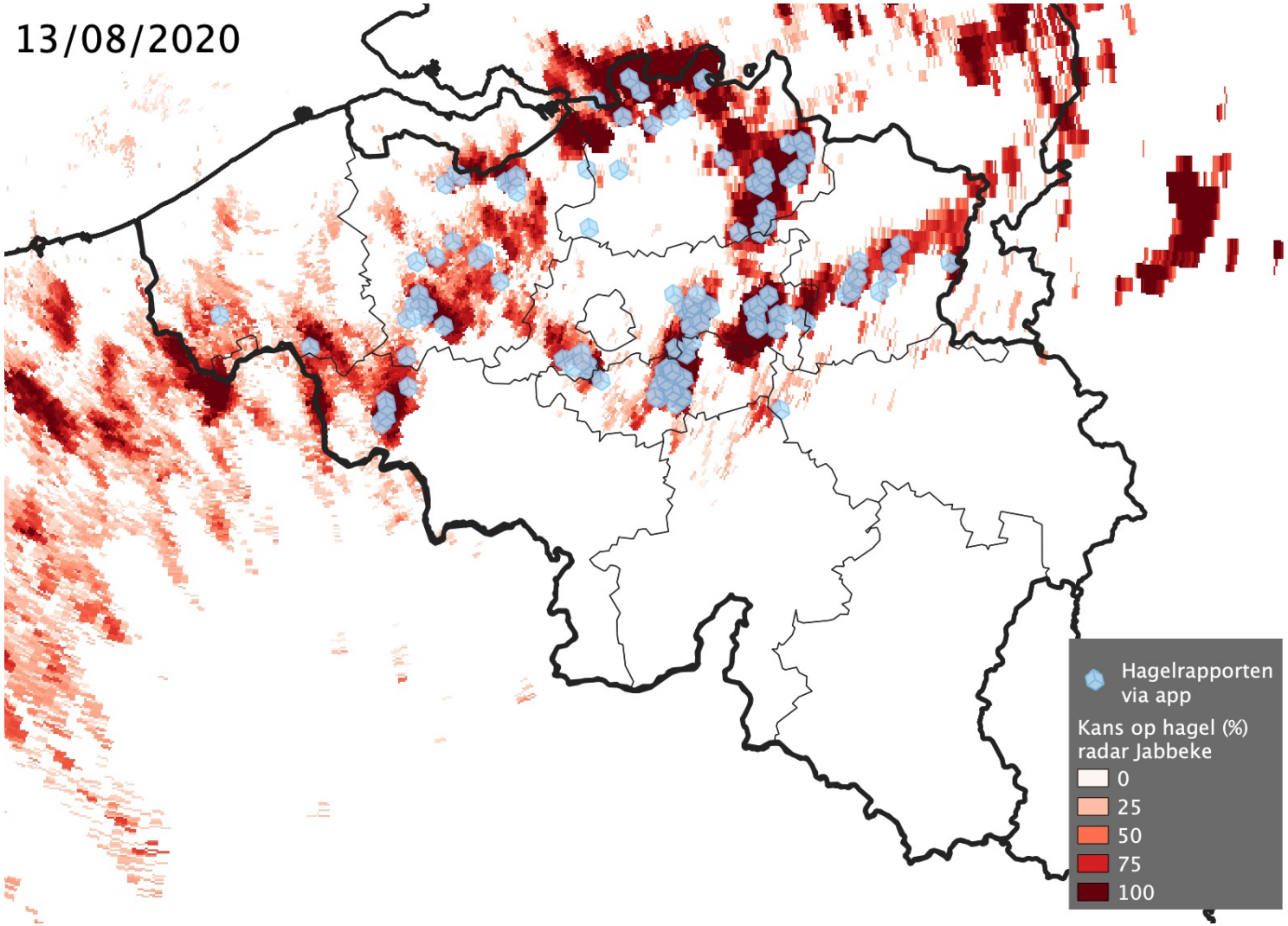
Size categories :

0	Hail stones smaller than a pea
1	Hail stones like a pea
2	Hail stones like a cherry
3	Hail stones like a walnut
4	Hail stones like a ping pong ball
5	Hail stones like a small egg
6	Hail stones like a snooker ball
7	Hail stones like a tennis ball
8	Hail stones like a big apple



Comparison POH versus citizen reports

24h overview 13/08/2020



Verification of hail detection product

- Conversion of POH to YES/NO information:
  hail detected (YES or NO) if POH > POH_threshold

- Only *plausible* reports with exact *GPS coordinates*

- *Distance* to radar < 150 km

- *Tolerance* on *time* and *location*:
  - Max distance between radar obs. and report : 2.5 km
  - Max delay between radar obs. and report : 10 min

Verification May-Sept 2020+2021+2022

Probability of detection (POD) Jabbeke radar

795

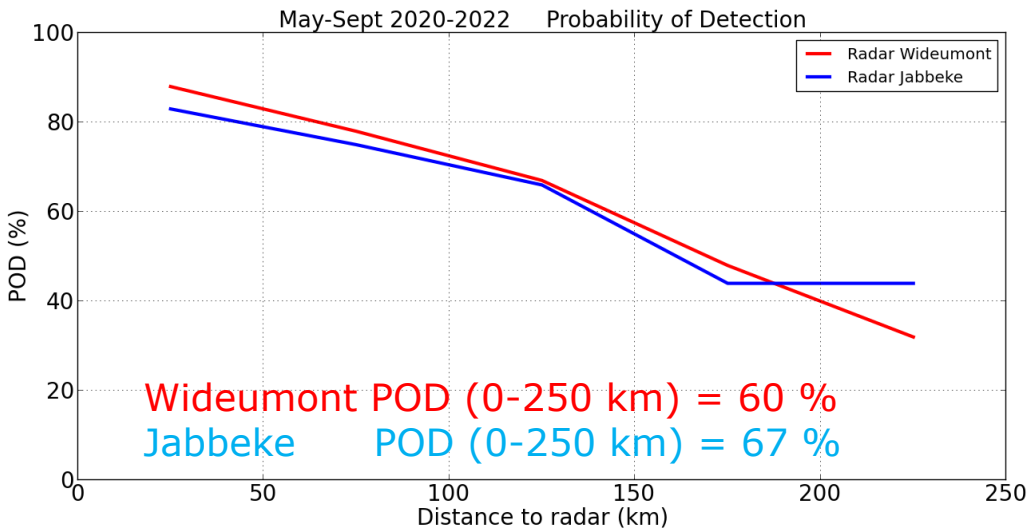
Valid hail reports at less than 150 km from radar

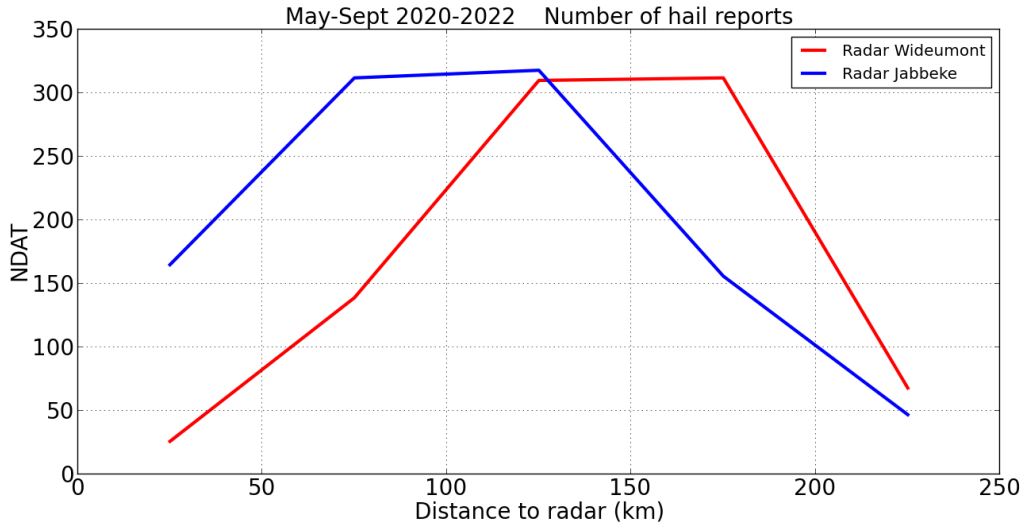
	Detection if probability > 0 % (POH_threshold)	Detection if probability > 50 % (POH_threshold)
	584 hits	449 hits
	211 missed	346 missed
	POD = 0.73	POD = 0.56
	POD = 0.72	POD = 0.62

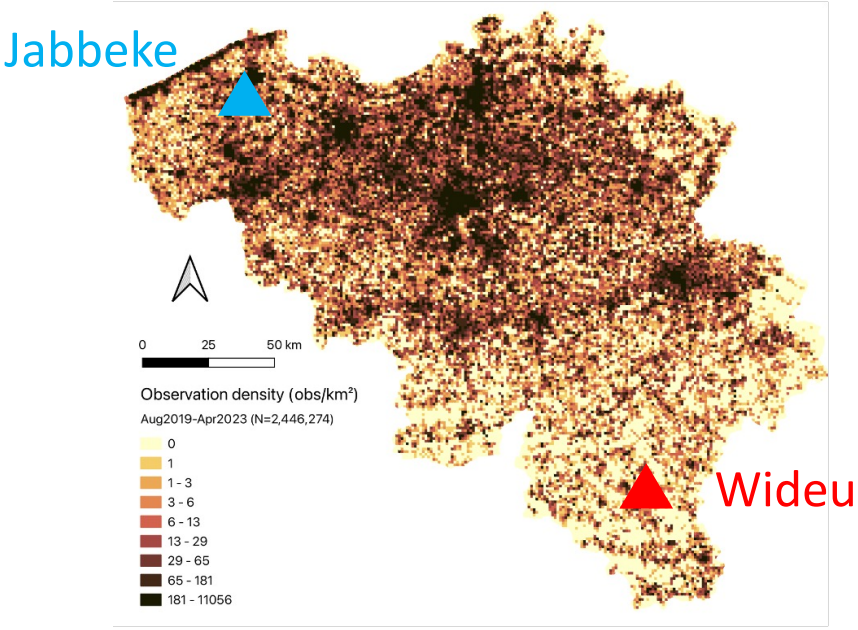
Jabbeke

Wideumont

Impact of the distance to the radar







- Probability of Detection POD decreases with the distance to the radar
- More reports where population density is high
  → more reports at short range for Jabbeke
- Higher POD (0-250 km) for Jabbeke is related to population density
- Benefit of international exchange

False detections May-Sept 2020

No "NO HAIL" reports.
Sunny, partly cloudy, overcast, rain and snow reports are considered as "NO HAIL"

269000
Jabbeke

"NO HAIL" reports at less than 150 km from radar

Detection if probability > 0 % (POH_threshold)

9718 hail detected (false detections)

259282 no hail (correct negatives)

POFD = 3.6 %
Jabbeke

POFD = 3.2 %
Wideumont

Conclusion

- Citizen reporting through smartphone app **new data source**
- **Unstructured** but **numerous** data
- Application :
  - Evaluation of the performance of the radar-based hail detection
  - Development of new method (*talk Sylvain Watelet, Thu*)
  - Verification of official weather warnings and forecasts per commune
- Large potential for the detection/verification of **local phenomena** (fog, snow, hail, …)
- Need to characterize/correct inherent **human-related biases**

THANK YOU

The Royal Meteorological Institute

Het Koninklijk Meteorologisch Instituut

L’Institut Royal Météorologique

Das Königliche Meteorologische Institut

The RMI provides reliable public service realized by empowered staff and based on research, innovation and continuity.

Het KMI verleent een betrouwbare dienstverlening aan het publiek en de overheid gebaseerd op onderzoek, innovatie en continuïteit.

L'IRM fournit un service fiable basé sur la recherche, l'innovation et la continuité au public et aux autorités.

Vertrauenswürdige Dienstleistungen für Öffentlichkeit und Behörden begründet auf Forschung, Innovation und Kontinuität.

RMI