GNSS data quality check in the EPN network

A. FABIAN, C. BRUYNINX, J. LEGRAND, A. MIGLIO

Royal Observatory of Belgium

Online | 4–8 May 2020

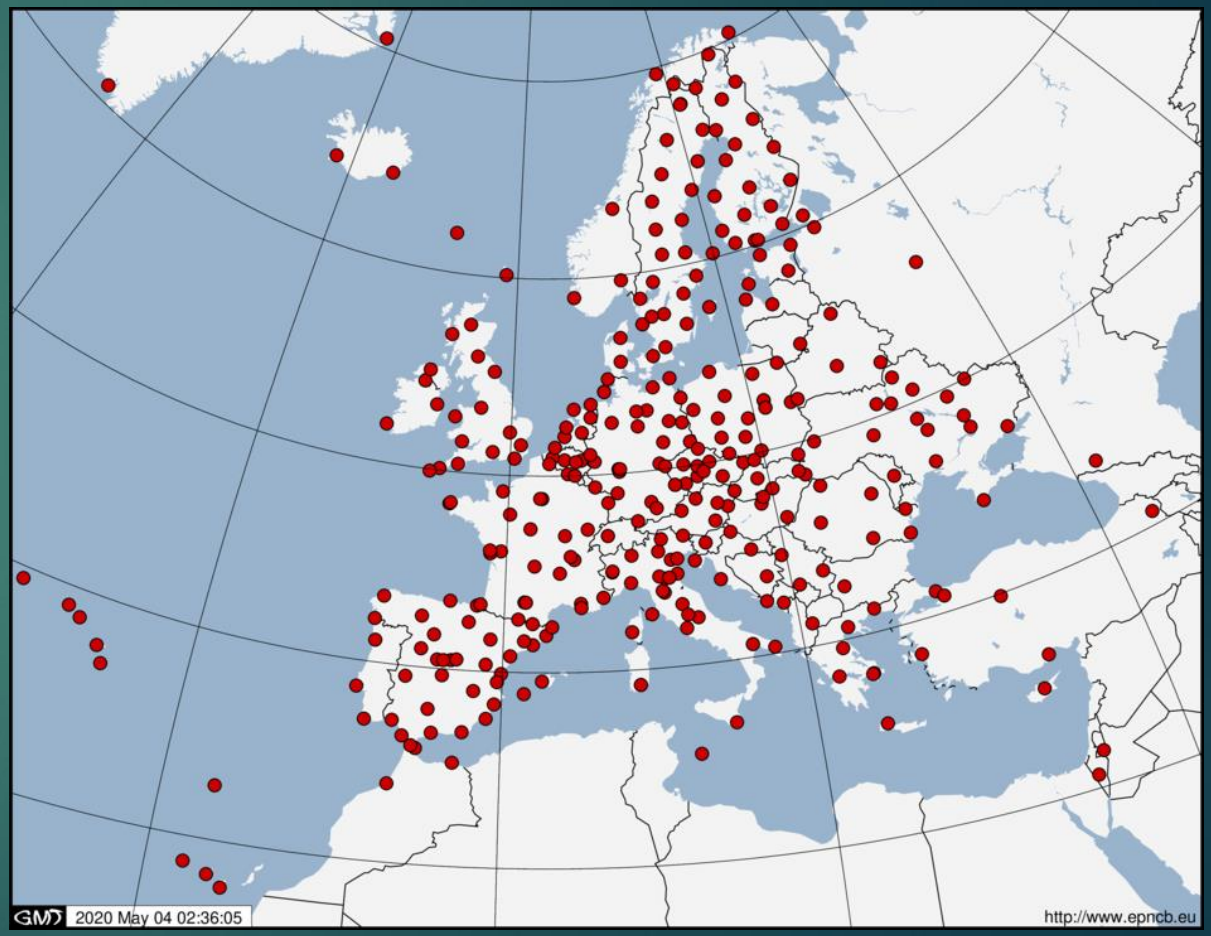
Introduction

Global Navigation Satellite Systems (GNSS) is a widely spread and most effective technique for geodetic applications and monitoring the Earth's atmosphere. Therefore, the density of the GNSS networks have grown considerable since the last decade. Each of the networks collects huge amounts of data from permanently operating GNSS stations. The quality of the data is variable, depending on the evaluated time period and satellite system. Conventionally, the quality information is extracted from daily estimates of different types of GNSS parameters such as number of data gaps, multipath level, number of cycle slips, number of dual frequency observations with respect to the expected number, and from their combinations.

In the following slides I would like to focusing on the receiver-dependent errors and the daily quality monitoring of the EUREF Permanent GNSS Network (EPN). Then the requirements and design of a new GNSS data quality monitoring system which is capable of handling more than 3000 stations will be presented. Followed by the ongoing improvements taking advantage of artificial intelligence techniques.

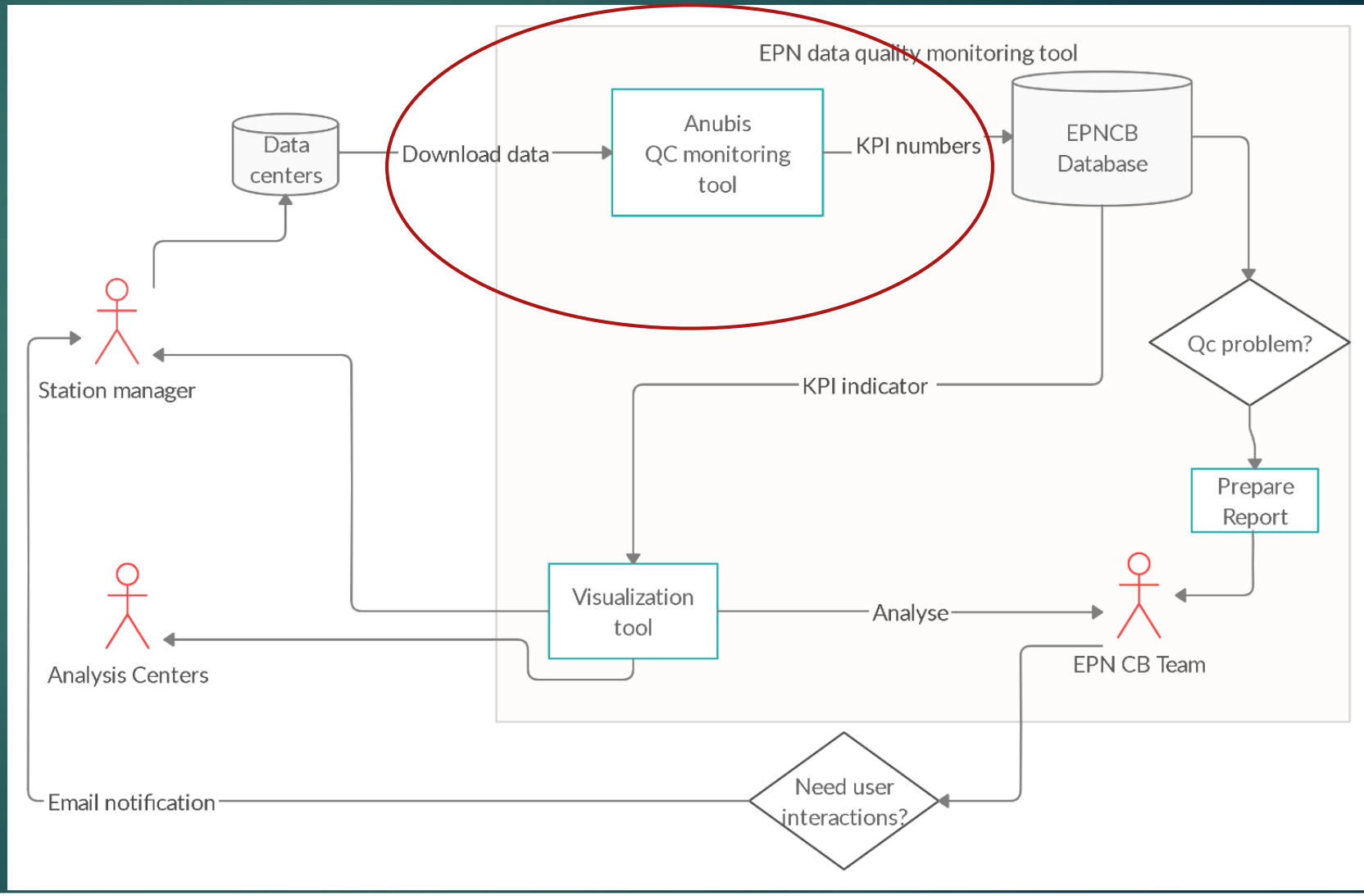
Background - EPN network

- EUREF Permanent GNSS Network Consists of
  - 300+ continuously operating GNSS reference stations
  - Data centers which provide access to more than 20 years of observations
  - Analysis centers that analyze GNSS data
  - The EPN central bureau (EPN CB, Bruyninx et al., 2019) which is responsible for the daily monitoring of the stations and management of the EPN.



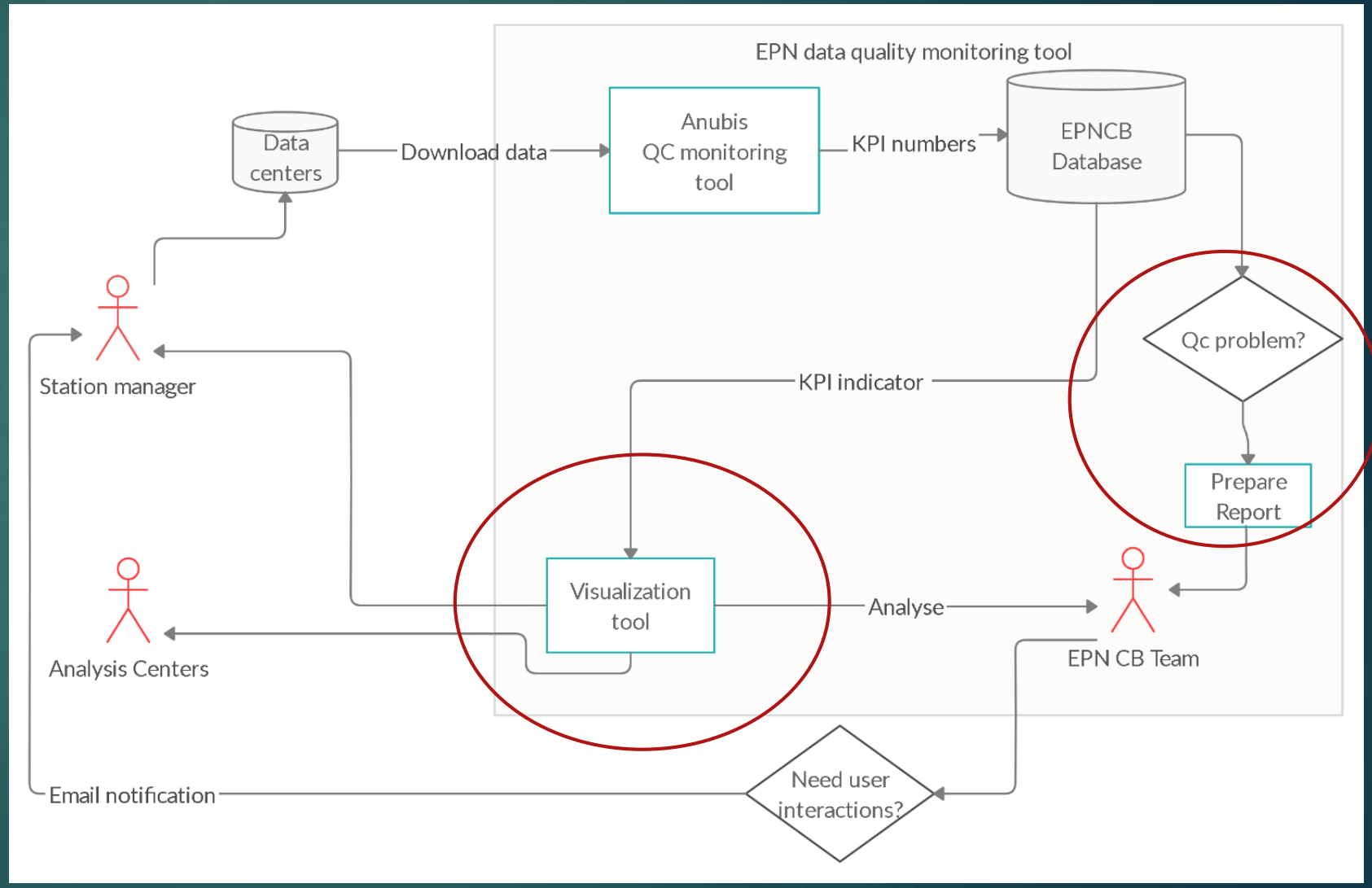
Daily data monitoring tool in EPN CB

- Downloads the data from data centers
- G-nut/Anubis software is used to extract the daily performance of the EPN stations and stores the constellation specific daily Key performance indicator (KPI) numbers (e.g. multipath value, cycle slips, observed/expected observation, missing epochs, number of observed satellites) in the EPN CB database.



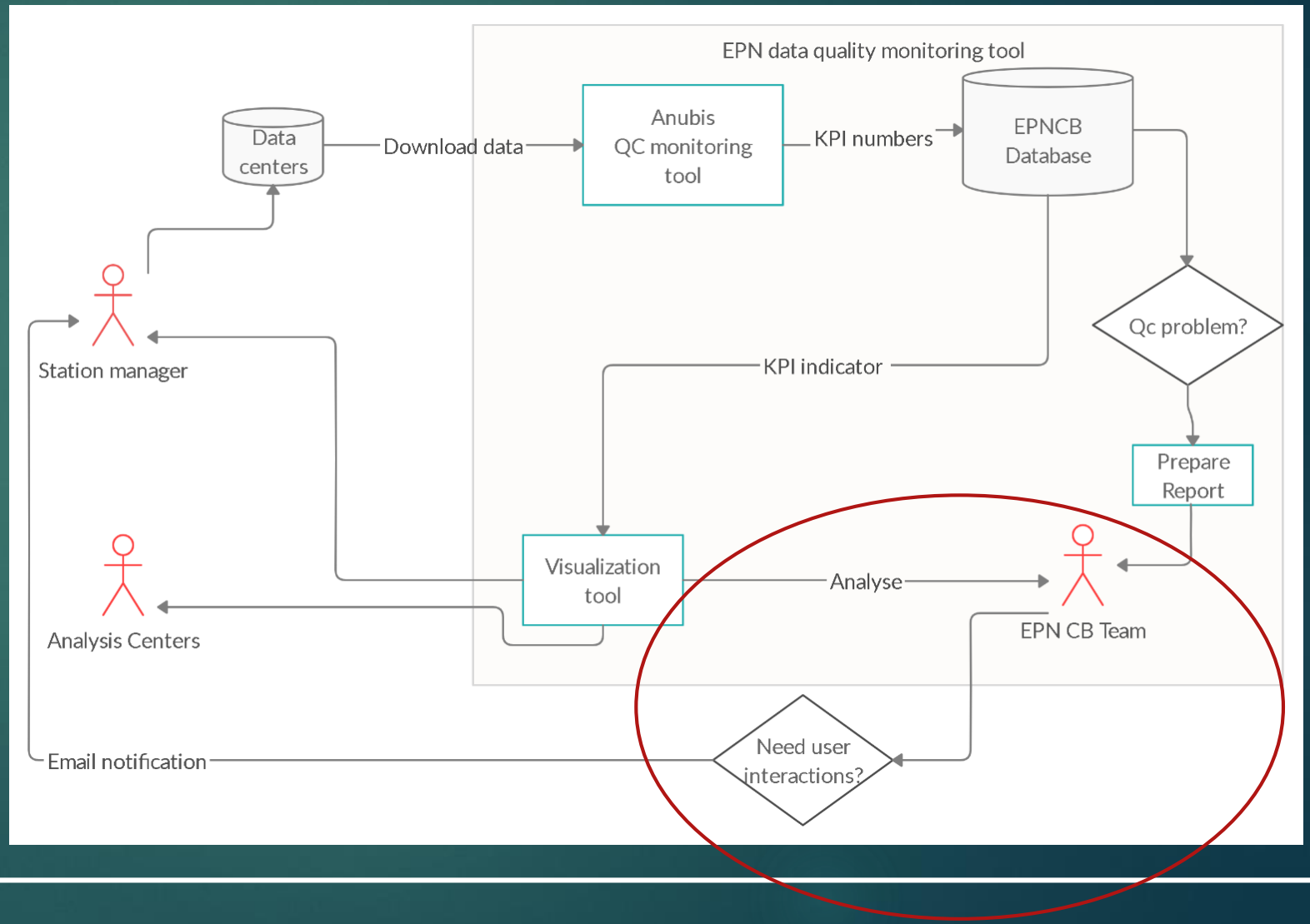
Daily data monitoring tool in EPN CB

- A user interface allows to visualize and compare the performance of the stations.
- The other extraction of the database is a report which contains statistical information. We use
  - preprocessing to remove the incomplete data files.
  - and compute the sliding averages for 80 and 14 days sliding windows for selected KPI.
  - and trigger alarm based if the average exceeds the predefined threshold.

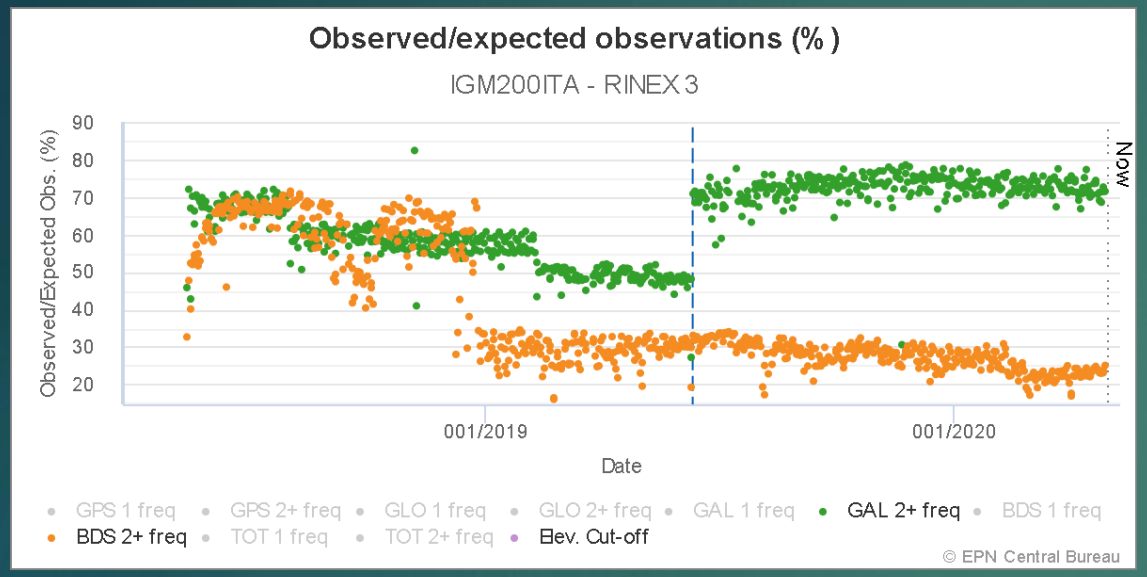


Daily data monitoring tool in EPN CB

- The EPN CB team gets the report and eliminates manually the False Positive (FP) values and investigates the origin of the quality degradation.
- The degradation can be caused by
  - receiver/firmware problem
  - receiver set up (elevation cut off) or local interference
  - satellite-dependent degradation
- The EPN CB team has to contact the station manager/analysis centers/manufacturer depending on the problem.



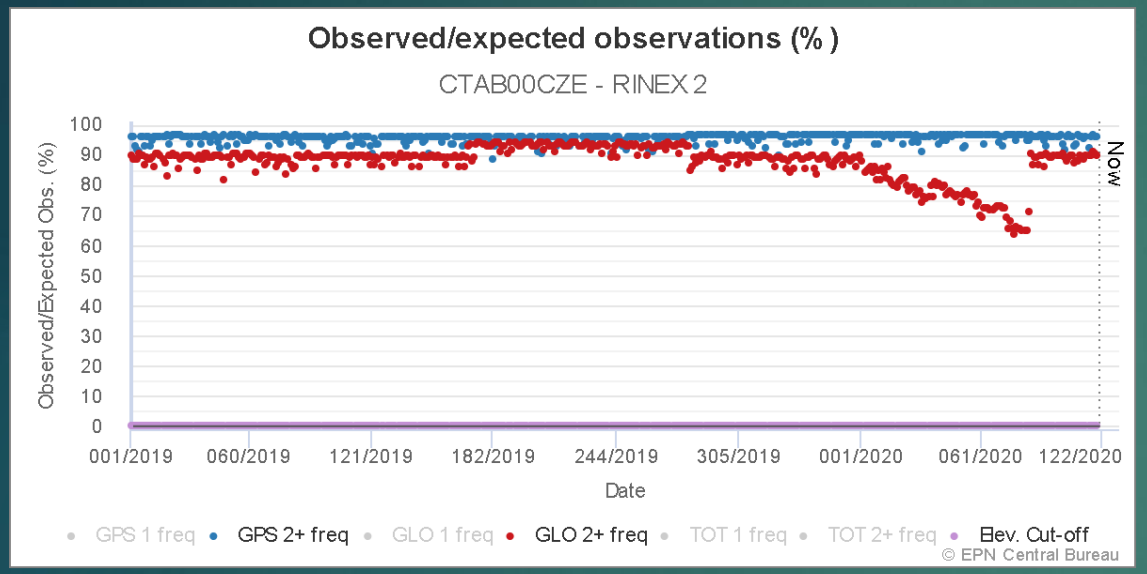
Typical results of the data quality checks I.



Problem with the Galileo observations:

- The system detected that the Obs/Exp (%) of the Galileo observations was falling down
- The EPN CB team had to check all of the stations with the same receiver and firmware
- All of the receivers with the same firmware had the same issue
- The issue was solved after the firmware was upgraded

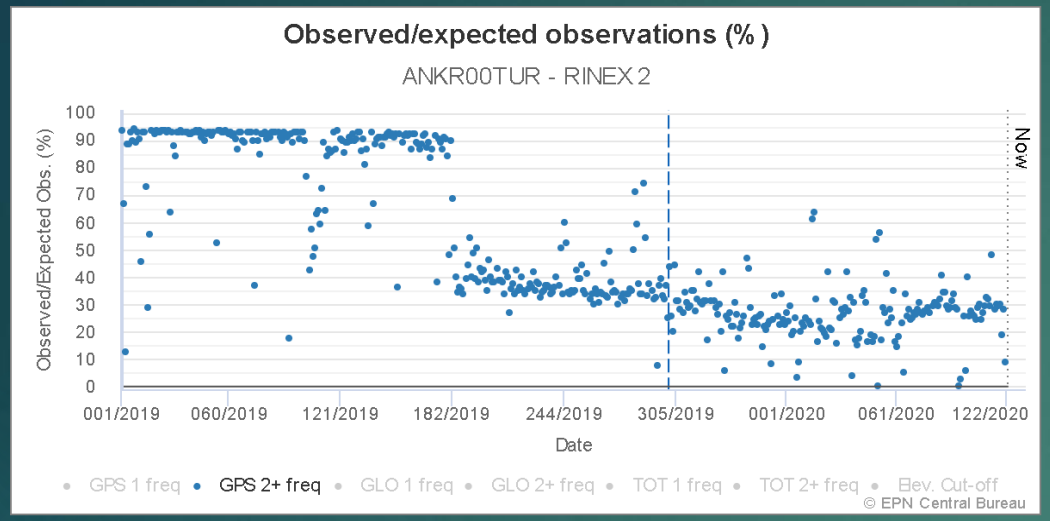
Typical results of the data quality checks II.

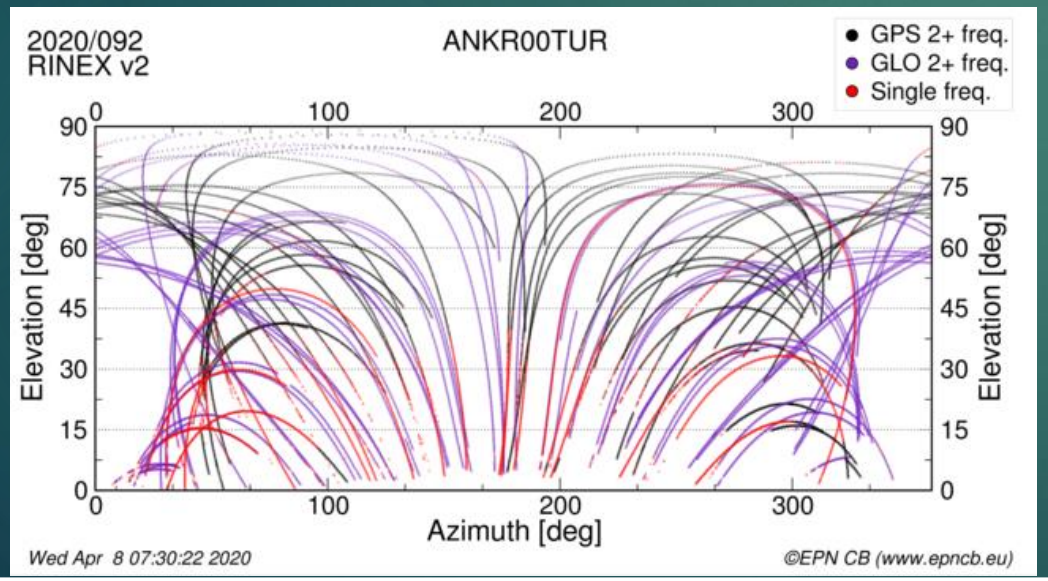


Problem with the GLONASS observations :

- The system detected that the Obs/Exp (%) of the GLONASS observations was falling down
- The EPN CB team had to check all of the stations with the same receiver and firmware
- All of the receivers with the same firmware had the same issue
- The EPN CB team contacted the station manager
- Measured engine reset solved this issue

Typical results of the data quality checks III.





The Obs/Exp (%) was falling down suddenly:

- The system detected the degradation.
- Several satellites were not tracked below 60° elevation.
- Effect on the multipath and on the cycle slips.
- Experienced effect on the position time series.
- The quality degradation only affects this station.
- The EPN CB team contacted the station manager and the analysis centers and inactivated the station due to the on-site degradation.

Motivation to improve the system

- The EPN CB is operationally collecting and analysing the quality of more than 300 GNSS stations.
- In the upcoming years, this data-monitoring tool will be used to also monitor the GNSS component of the European Plate Observing System (EPOS) expected to include more than 3000 GNSS stations.
- This anticipated inflation of GNSS stations to be monitored will make it increasingly challenging to select the high quality GNSS data.
- The current system is limited
  - does not exploit correlations between the daily data quality, coordinate time series and the GNSS station metadata (such as equipment type and receiver firmware) often common to many stations.
  - It is using predefined threshold to triggers alarms.
  - It is a time-consuming task to detect and eliminate the False Positive (FP) values.
  - It is a lot of manual work to investigate the source of the data quality degradation.
- The current semi-automatic method is not designed to handle the larger amounts of data.

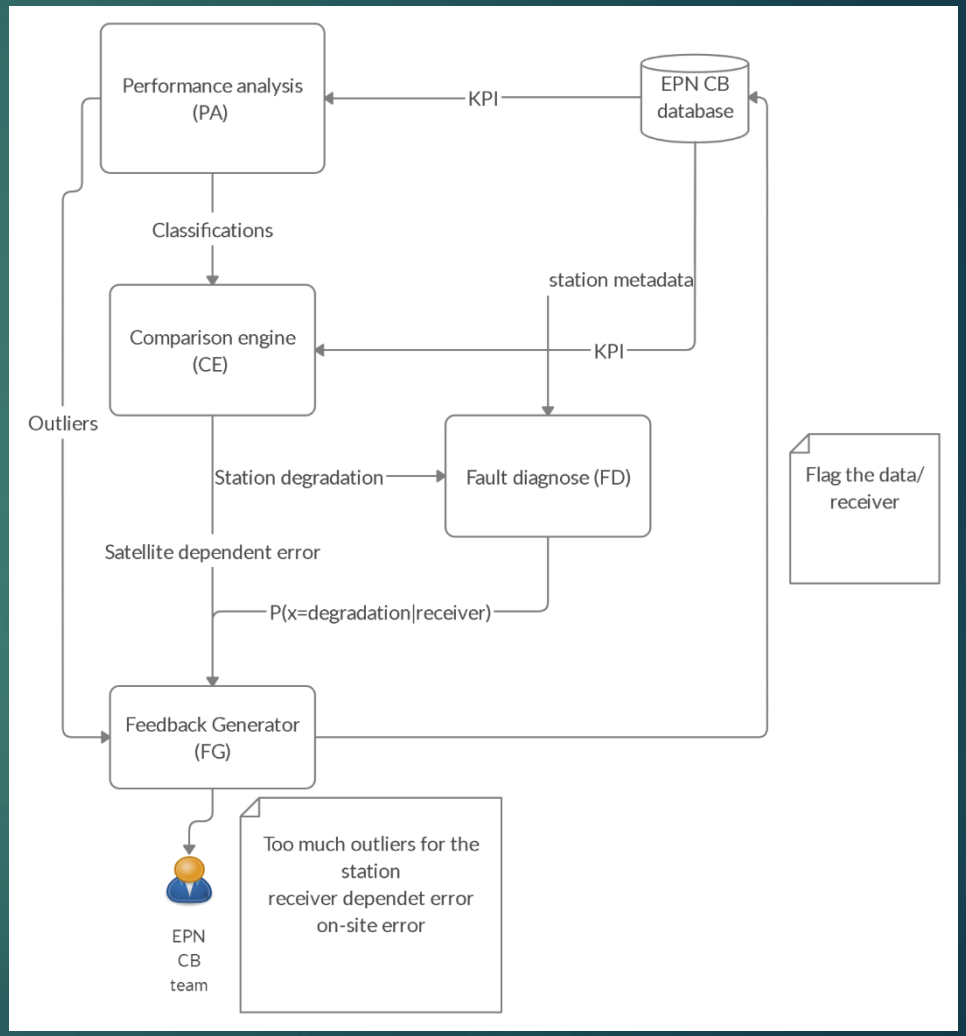
Requirements for the new system

The new monitoring system has to

- be an expert AI system which acts as human expert and applies the expertise rules and knowledge
- analyse most of the KPI indicators (e.g. multipath, cycle slips) and classified their behavior. (e.g. degradation, jump)
- minimalize the number of the FP values
- find the correlations between the daily data quality, coordinate time series and the GNSS station metadata (such as equipment type and receiver firmware)
- find the root cause of the problem and report to the EPN CB team
- make a final decision on the quality degradation and save it to the database

Possible design of the new system

- contains the following components:
  - **Performance analysis (PA)** which eliminates the outliers and normalizes the indicators and determines the classification for each of the KPI.
  - **Comparison engine (CE)** which compares the KPI of each station at a specific day.
  - **Fault diagnose (FD)** which determines a possible area of fault and a consequence effect of the fault. Analyze which KPIs are affected the most.
  - **Feedback Generator (FG)** which delivers report on the performance of station.

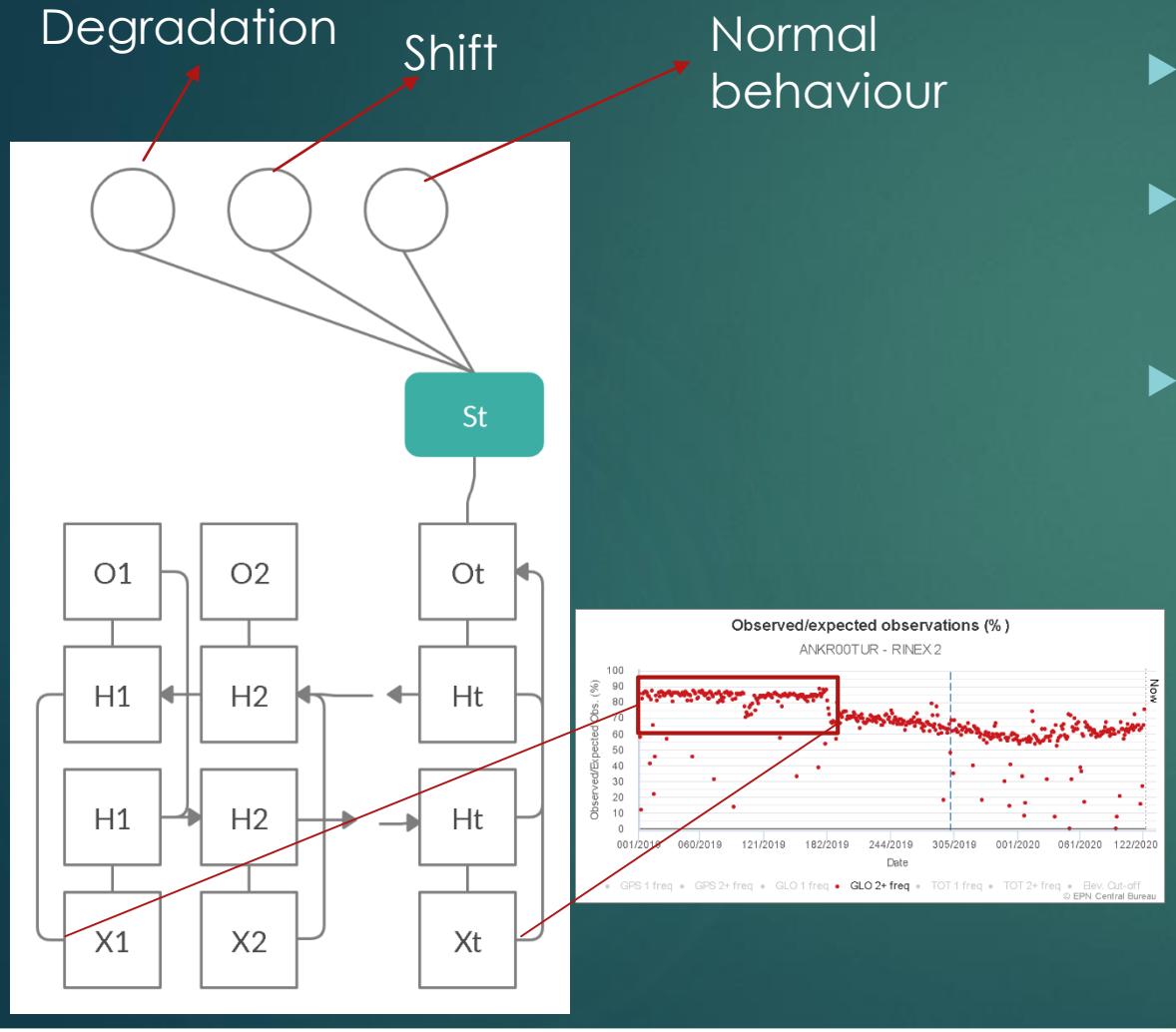


Performance analysis (PA)

The aim of this component is to check each of the performance indicators, eliminate the outliers and determine the classification for each of the KPIs.

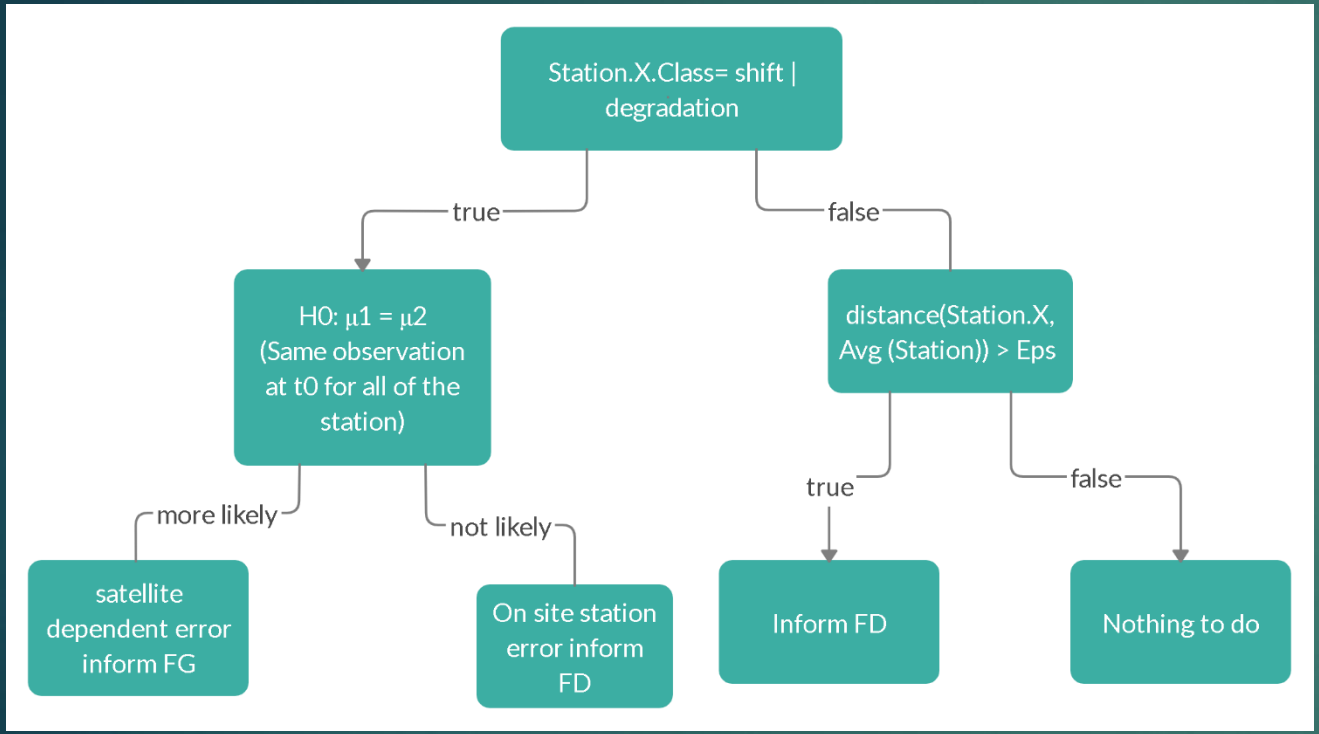
- Step 1: Eliminate outliers
  - Use thresholds (e.g. not a full day of observation) or a support vector machine depending on the type of the indicator.
- Step 2: Classification of the KPIs
  - Use a recurrent neural network (RNN) for the deeper understanding of the behaviour of each KPI.

Performance analysis (PA) - Step 2: Classification of the KPIs



- RNN is designed to recognize patterns in sequences of data.
- At each step in time the RNN uses results from the previous sequences and its output is used as an input for the next sequence.
- As a result, the RNN gives the classification vector as an output and provides this information to CE.
  - Possible classification for the KPI is
    - Degradation
    - Shift
    - Normal behavior

Comparison engine (CE)



The aim of this component is

- to decide whether the root cause of the degradation is day dependent or station dependent
- to provide the relative size of the degradation

- Step 1: If classification of the station is degradation or shift then the system makes a statistical hypothesis test if all of the stations are affected by the same degradation or not.
- Step 2: This component compares the performance of each of the stations on a specific day and computes the relative size of the degradation.
- Step 3: If the component explores any abnormal behavior then informs the FD or FG components.

Fault diagnose (FD)

The aim of this component to determine a possible area of faults and approximates the final decision.

The component

- is triggered when a quality degradation is affected by one of the stations for a specific day.
- finds the linear combination of the KPI parameters which is responsible for the greatest variation in performance.
- is a classifier to compare the current value of the indicator and all of the stations where the same receiver is installed.
- returns back the root cause of the problem.

Feedback generator (FG)

- Delivers report on the performance of station. The aim of this component is to update the flag for the specific daily data and for the specific receiver or for a station for a given period.
- The FG notifies the EPN CB team if a new station or a receiver problem popped up.

Summary

These slides presented the currently used method of GNSS data quality checking and its limitations. Based on more than 20 years of GNSS observations collected in the EPN, we showed typical cases of correlations between the different data quality metrics and GNSS station metadata. Then, we set up the requirements and design of the new GNSS data quality monitoring system capable of handling more than 3000 stations. Based on the collected EPN samples and the typical cases, we introduced the ongoing improvements taking advantage of artificial intelligence techniques and showed a possible design of the system.

Based on this component model we are starting the development of the new GNSS data quality monitoring tool.