OneNet Open Call Webinar| 11/01/2022

Advanced Data Quality Analysis of Data Exchange Platforms

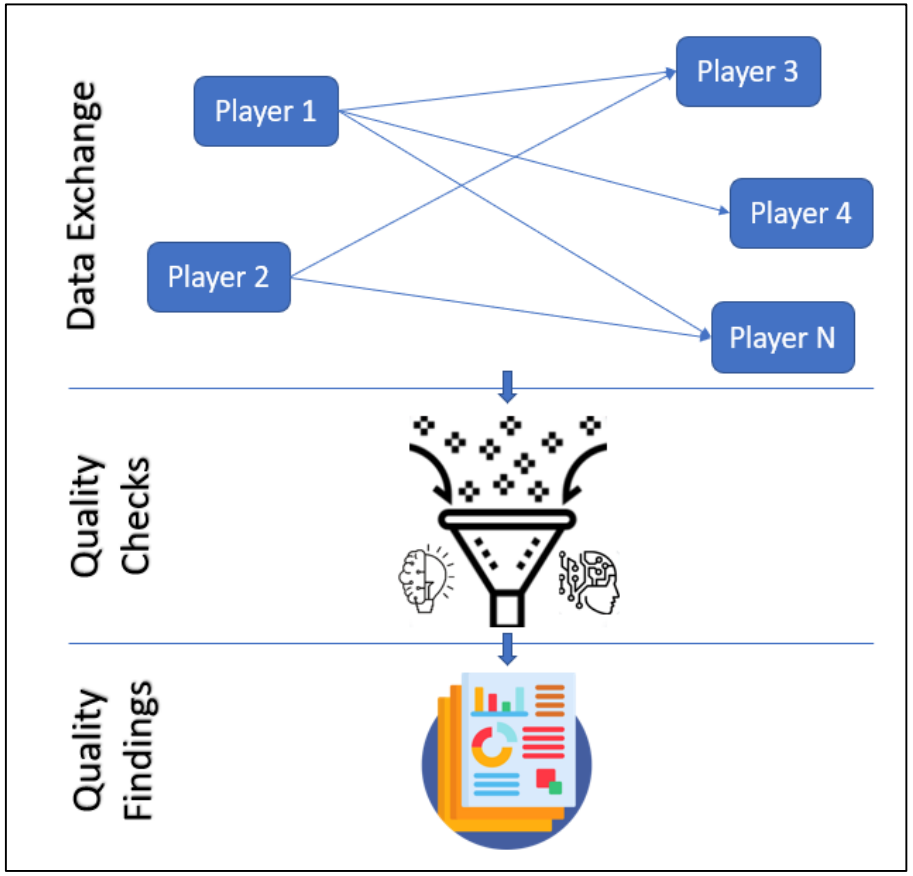
Ali Fahri & Dmitry Belichenko
ENTSO-E

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 957739

Outline

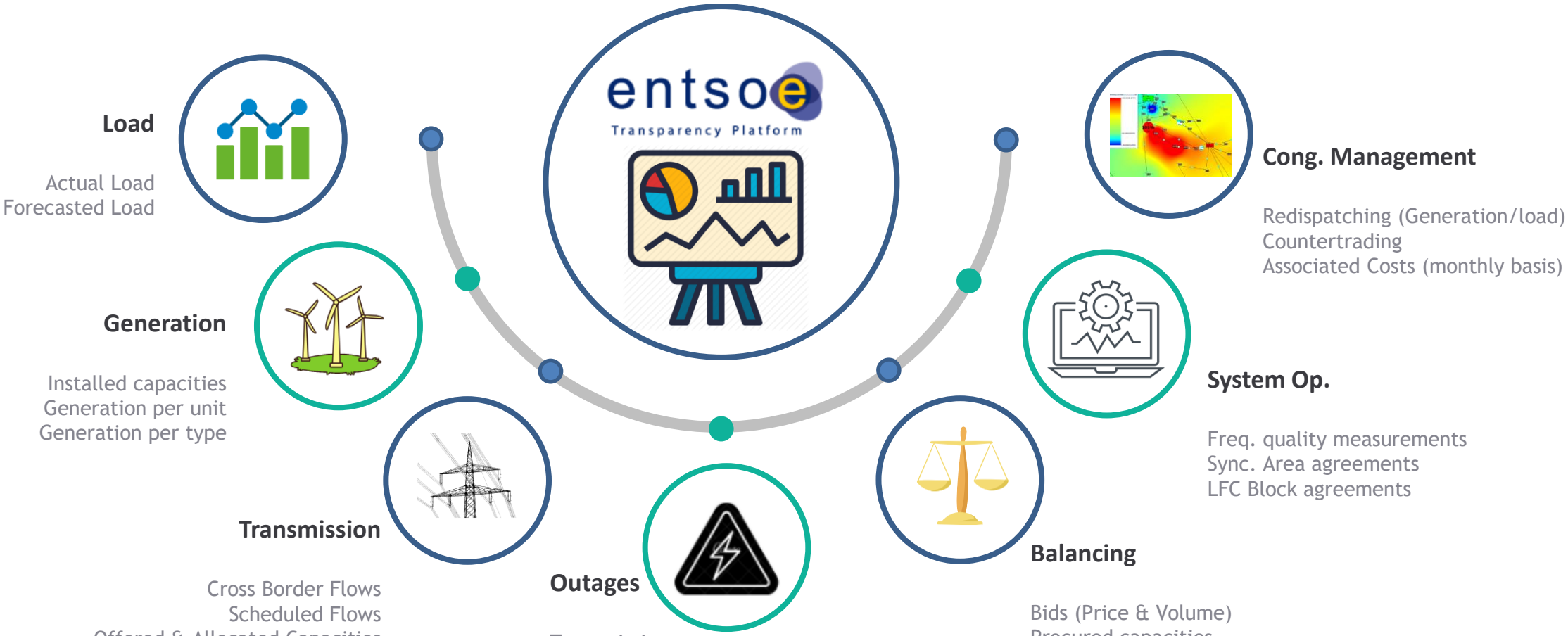
- Scope & Objectives of the Scenario
- Overview of the TP data
- Data Download Options from TP
- Existing Quality Analysis Methods
- Q&A

Scope & Objectives of the Scenario



- Apply advanced **machine learning** algorithms on the data exchanged between different players in the TSO-DSO-Consumer value chain
- Detect **outliers** from timeseries where standard methodologies are not sufficient
- Enhance the **quality of the data** by highlighting the abnormalities
- The developed tool is expected to be part of project's middleware for continuous quality checks
- The output will be the result of the data quality analysis performed regularly
- Transparency Platform Data can be used to test the developed tools

Overview of TP Data



Data Download Options

Web GUI

Limited download
Different formats

Data Repository

Asynchronous process
Up to 50 MB
XML format (zipped)

SFTP

Bulk download
CSV format
Monthly grouped
Hourly updates

Subscriptions

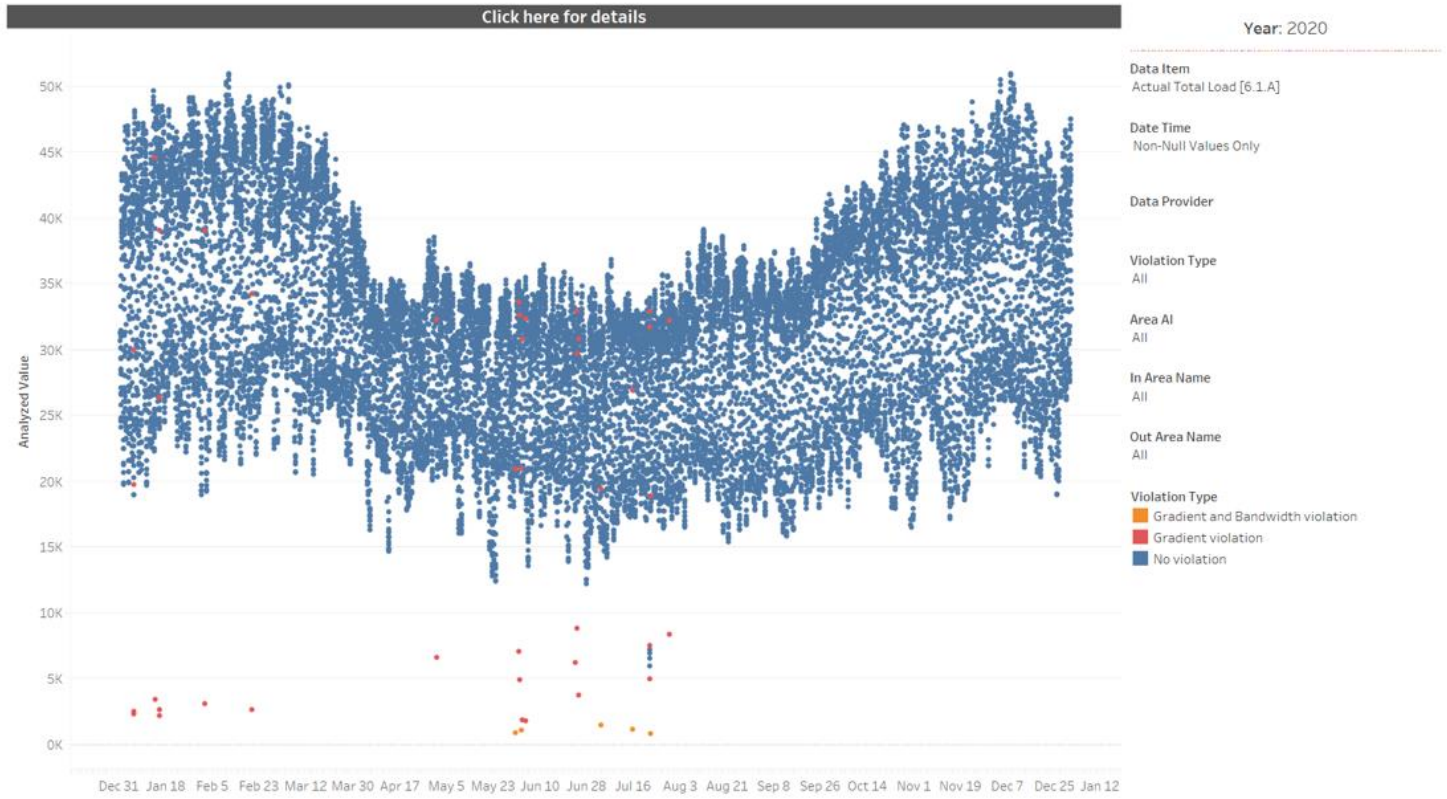
Data feed
XML format
Web service/ECP

REST API

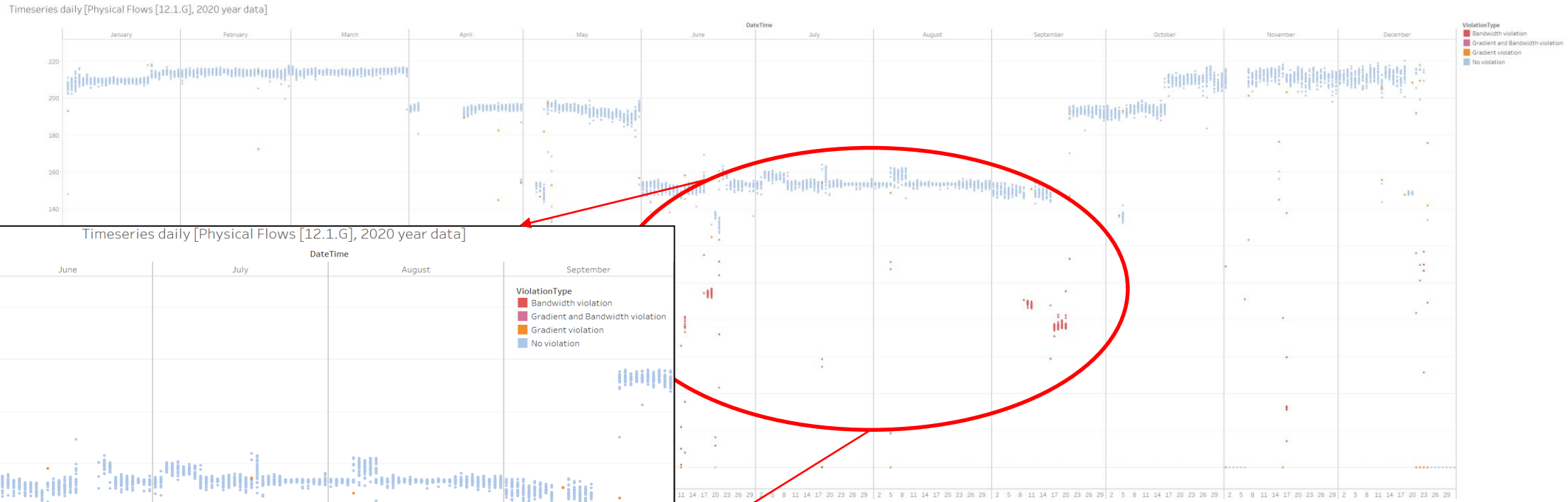
Almost real time
Flexibility on requests
Request Limits Apply
XML format

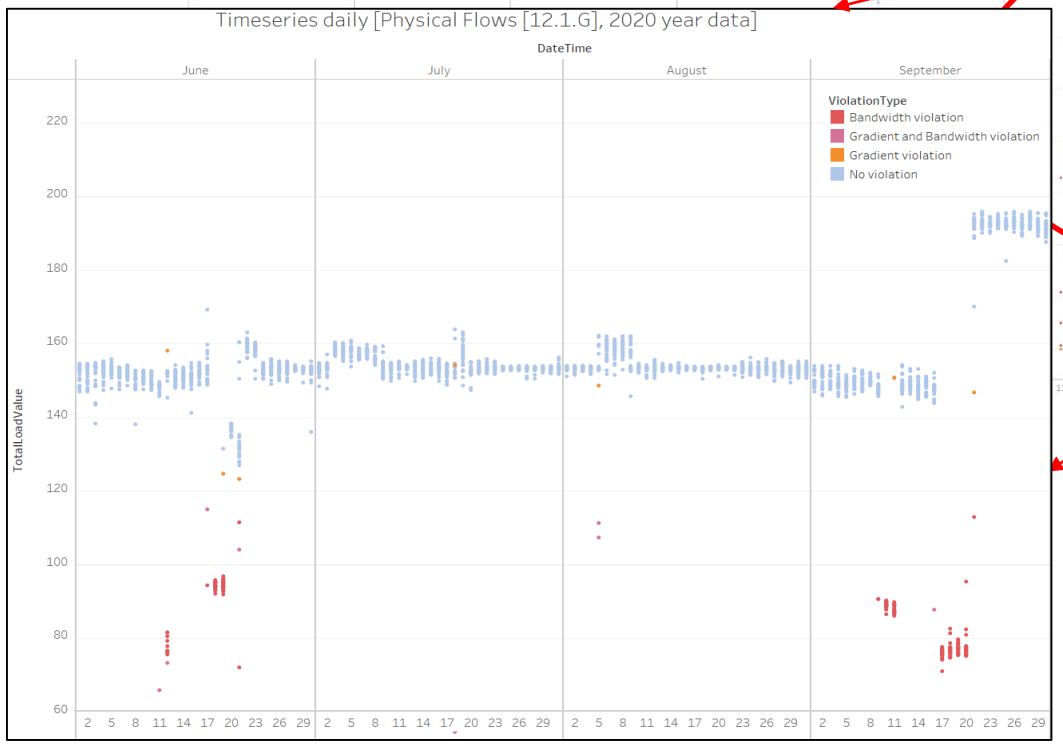
Existing Quality Analysis Methods

The initial proof of concept (PoC) done for Actual Total Load [6.1.A] Data Item has resulted in acceptable quality analysis results based on the Median Absolute Deviation technique (https://www.academia.edu/5324493/Detecting_outliers_Do_not_use_standard_deviation_around_the_mean_use_absolute_deviation_around_the_median).



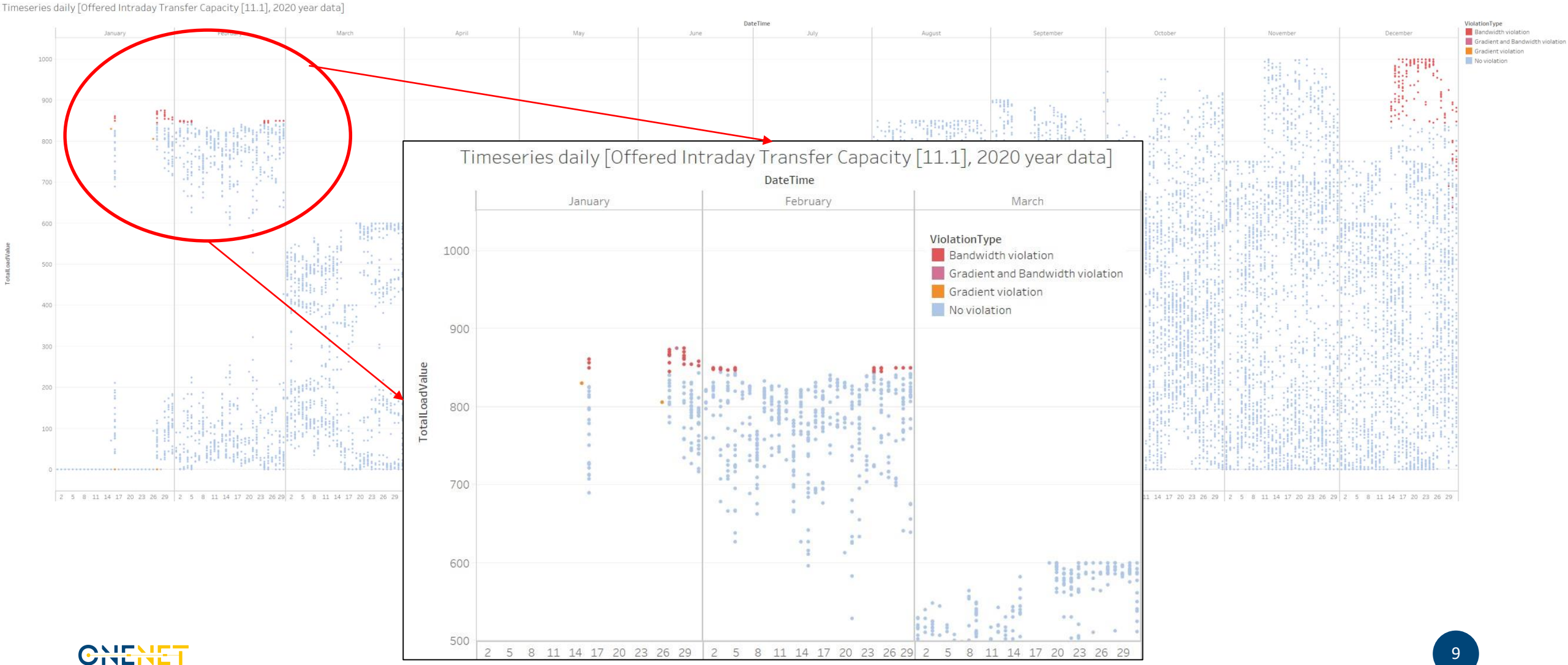
More Complex Cases - 1





More Complex Cases - 2

More Complex Cases – 3



Thank You

Ali Fahri & Dmitry Belichenko

Contact Information

ali.fahri@entsoe.eu

dmitry.belichenko@entsoe.eu

transparency@entsoe.eu