REPORT FOR PUBLIC CONSULTATION

Improvement of the quality of input data for congestion management in the frame-work of CREG decision (B)658E/73

June 10, 2022

Content

1. Introduction

On December 9, 2021, CREG issued decision (B)658E/73 on the targets to be achieved by Elia in 2022 in the framework of the system balance as referred to in Article 27 of the tariff methodology. One of the incentives fixed in this decision is “Improvement of the quality of input data for congestion management”. The description indicates that the incentive primarily includes a report on the analysis of the most significant deviations between predictions and reality and an examination of possible short- and long-term solutions and secondly, includes a recommendation and proposal for implementation of concrete solutions in the form of a roadmap for the future.

The objective of the report is to give an overview of the current modelling practices (Individual and Common Grid Model) and to show transparency on the actual quality of the input data for congestion management. The report also covers a root-cause analysis of deviations in the input data (different forecasts f.e. wind, solar, load,…) and possible solutions (possible on short-term or long-term) for improvement.

2. Scope and approach

As proposed in execution of the incentive on "Improvement of transparency with regards to the detection and management of congestion" defined in CREG decision (B)658E/52 of 28 June 2018, Elia publishes, since the beginning of 2020, a quarterly report[1] on congestion management covering a period of three months. This report includes:

- Information on the quality of forecasts used as operational input data for the creation of the Individual Grid Models (IGMs)
- Information on the quality of output data
- Information about the timing, power, location, and purpose for activations of Costly Remedial Actions by Elia.

The incentive defined in decision (B)658E/73 for 2022 builds further on the incentive realized in 2019. Based on the reports mentioned above, deviations can be detected between the input data used for the creation of the IGMs and the reality. On top, based on all the data stored in order to publish the reports for more than 2 years, it is possible to perform a more in-depth analysis. The scope of the incentive is to set up this in-depth analysis of the causes of the deviations in the different forecasts (wind, solar, load,...) and to study possible solutions to improve these forecasts. These solutions can be found in the infeed data, the forecasting model or the resulting forecast.

The different forecasts are analyzed on an individual basis and for each forecast the following aspects are studied and structured in the report (if relevant):

[1] Link: Congestion management (elia.be)

- Infeed data, forecasting model, resulting forecast (into IGM input): see figure below to make the link with the KPIs in the current quarterly reporting
- AS IS versus possible TO BE situation

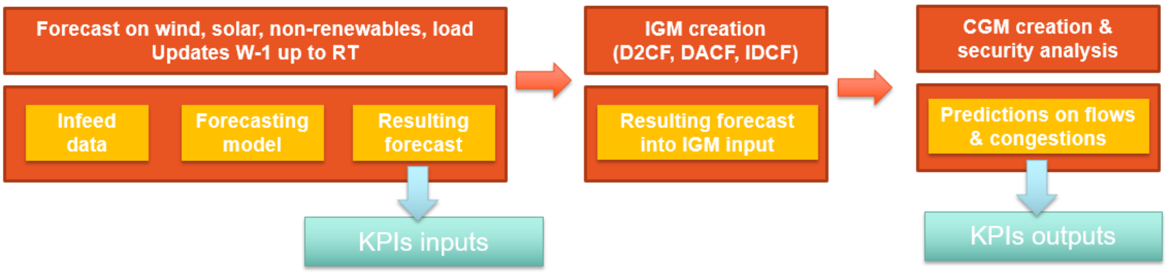


Figure 1: overview of the scope of the analysis performed in the report

It is part of the incentive to start an implementation plan for short-term solutions and to establish a roadmap for long-term solutions. For the long-term solutions, a link with innovation projects and challenges towards the future will be made if relevant.

The goal of the incentive is not as such to improve all the individual forecasts, but to focus on improvements and needed developments in order to keep a good level of congestion management decisions and avoid unnecessary costs.

The report is structured into the following sections, in line with the concrete deliverables as foreseen in the decision of CREG:

- Chapter 4: Transparency on congestion forecast today and current modelling practices. This chapter also includes a mapping of the opportunities for improvements and a benchmarking with other TSOs.
- Chapter 5: Transparency on forecast quality, root-cause analysis on deviations in forecast compared to real time and possible solutions (short- and long-term) to improve the forecasts.
- Chapter 6: Challenges and opportunities for the future in order to prepare the roadmap to be included in the final report.

The content of the report is discussed with several internal and external experts. A workshop is organized to improve the transparency on the content of the report and to introduce the public consultation.

3. Congestion forecast, current modeling practices and performances

Deviations in flow forecasts in day-ahead and intraday (DACFs and IDCFs) have a direct impact on congestion management, since the DACFs and IDCFs are the unique source for congestion detection and congestion solving. Of course deviations on uncongested elements are not relevant (e.g. on strong radial elements typically). Flow forecast deviations in D2CFs are not directly in scope of this document because it does not impact the congestion management (identification and solving of congestions). Of course, there is a strong link between DACFs/IDCFs and D2CFs and any error in D2CF flow forecast may lead to unrealistic cross-border capacities given to the market. The congestions will only materialize if the market goes in this unrealistic position and if the Flow Reliability Margins (FRMs) are not large enough to cover those errors. The focus of the analysis in the report is therefore on the day-ahead and intraday timeframes but most of the proposed solutions will have a positive impact on the D2CFs quality as well. In the description of the congestion forecasting current practices, the D-2 timeframe is also covered.

Note that the network calculations realized in week-ahead are not using forecasts. For this time horizon Elia calculates and solve several scenarios representing "realistic worst cases" ensuring that the outage plan and associated grid topology for the next week is acceptable. An accurate model at each node is of great importance in order to calculate realistic scenarios. However, improving week-ahead forecast would have no influence on the congestion management.

3.1 IGMs building process

IGMs are always built on the "best estimate" principle (i.e. most probable situation is the reference). The figures below give an overview of the D2CF compared to the DACF/IDCF building process. More detailed are provided in the paragraphs below.

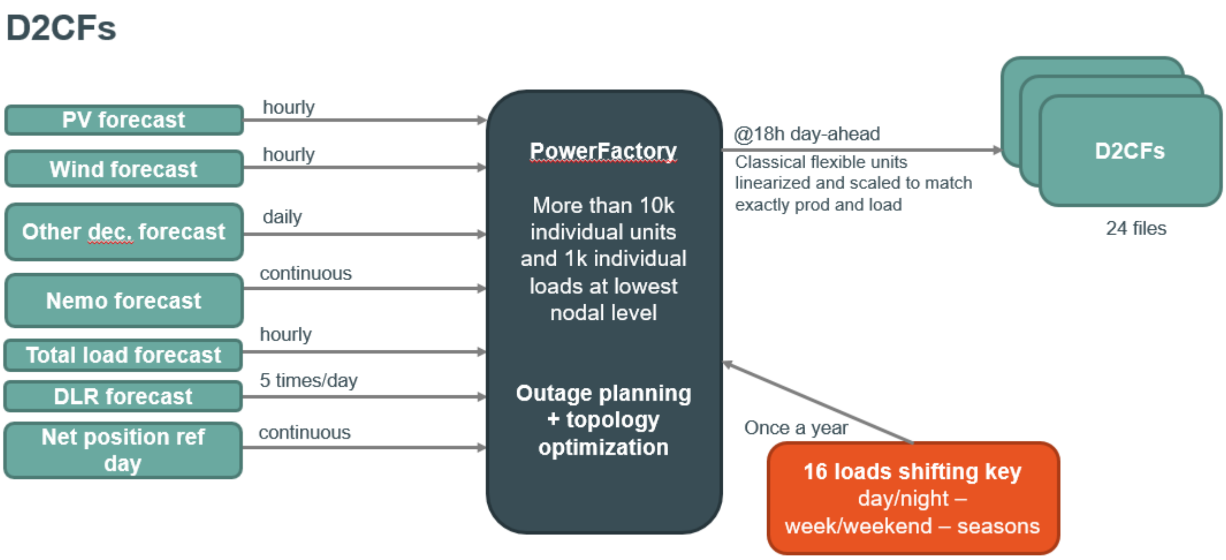


Figure 2: overview of D2CFs building process

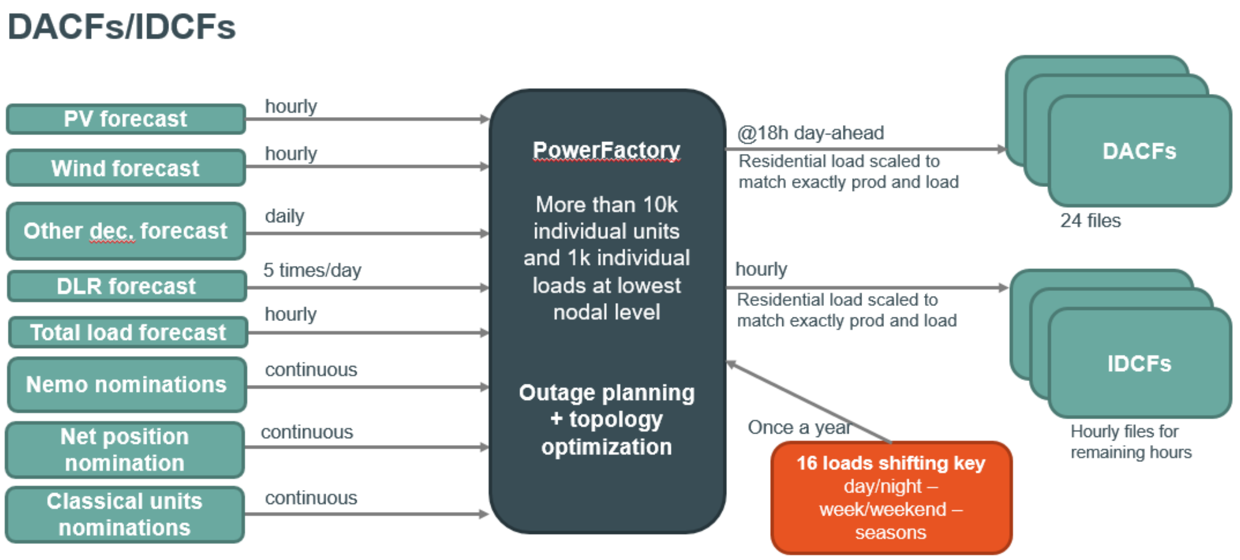


Figure 3: overview of DACFs/IDCFs building process

3.1.1 D2CF

Before 7 p.m., D2CF files for each hour of the day + 2 are created with the following contents and hypotheses:

- Latest (hourly updates mostly) **decentralized production forecasts** at power plant level (i.e. most accurate electrical localization as possible, > 10 000 individual units).
- Latest (hourly updates) **total load forecast** which is then spread over about 1000 individual loads by applying one of the 16 **repartition keys** depending on season, day/night and week/weekend categories.
- All **outages** down to the lowest voltage level and their preventive topology changes when necessary. Usually PSTs taps positions are neutral (will be optimized by the process itself, only north border often anticipated at 15/15/15/15 as a better best estimate position). With GO-live of FB CORE DA early June 2022, initial PSTs taps are pre-optimized (before first capacity calculation) in order to minimize loopflows.
- **Net position** of the referential day imposed by the CGM process in order to have coherent IGMs from all countries (prerequisite for merging purposes). With GO-live FB CORE DA, net positions for CORE countries are based on a centralized forecast.
- Latest **NEMO** flow forecast (forecast bought by Elia from an external provider). With GO-live FB CORE DA, the net position forecast also includes the forecast on NEMO flows improving further its accuracy.
- **Alegro** = Reference Day Schedule (but removed during capacity calculation, in order to be fully optimized by the FB process). With GO-live FB CORE DA, the net position forecast will also predict the initial set point of Alegro.

- **Element ratings**[2] based on forecasted temperature + dynamic line ratings for concerned lines (capped to 5% in order to ensure high reliability).
- **“Must run” conventional units** (essentially nuclear units).

  ➔ The only free parameter is the production of flexible conventional units. Total production of such units is imposed by the other fixed parameters and is then linearly (same % value w.r.t. units maximal production level) ventilated to all available flexible conventional units. In case there are not enough available flexible volumes to reach the imposed target, a scaling on residual load is done. This may happen when the referential day has a strong difference in renewable production level compared to the target day. The resulting imposed net position might be unrealistic in this case.

  ➔ With FB CORE DA go-live the net position CORE forecast ensures that the flexible conventional units are not saturated and no scaling of the residual load will be necessary.

3.1.2 DACF

Before 18h, DACF files for each hour of the day + 1 are created with the following contents and hypotheses:

- Latest (hourly updates mostly) **decentralized production forecasts** at power plant level (i.e. most accurate electrical localization as possible, > 10 000 individual units).
- Latest (hourly updates) **total load forecast** which is then spread over about 1000 individual loads by applying one of the 16 **repartition keys** depending on season, day/night and week/weekend categories. Note that all the industrial loads are fixed but residential loads are the free parameter (see hereunder).
- All **outages** down to the lowest voltage level and their preventive topology changes when necessary. PSTs taps positions based on foreseen market flows, outages and recent taps positions. Topology will at any case evolve throughout the process based on identified congestions.
- **Net position** resulting from the international trades (nominations).
- **NEMO** nominations.
- **Alegro** set points defined by FB MC.
- **Element ratings** based on forecasted temperature + dynamic line ratings for concerned lines (capped to 5%).
- **Conventional units** nominations.

  ➔ The only free parameter is the residential loads. Any error in nominations or in decentralized production forecast will then propagate itself to the residential loads.

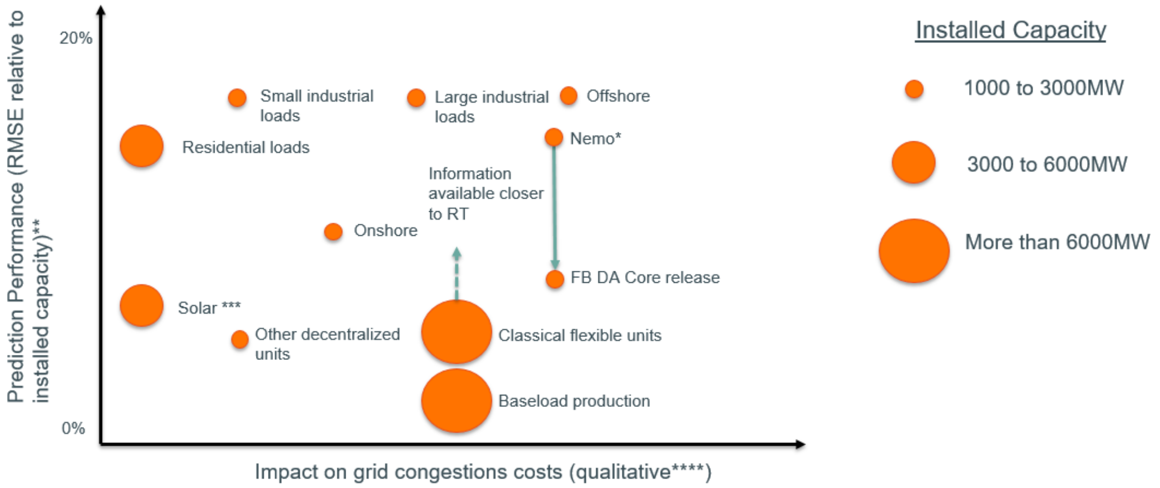
[2] **Dynamic Line Rating (elia.be)**

3.1.3 IDCF

IDCF files are regenerated automatically every hour for the remaining hours of the day with the latest available information following the same principles as DACF files.

3.2 Mapping of opportunities

Before starting to reflect on possible solutions for forecast improvements, it is important to evaluate for which forecasts there is room for improvement, but also to evaluate which improvements have the largest impact on congestions.



*D2CF error
**Lower is better, RMSE day-ahead forecast 2021.
***For daylight hours only
****looking at PTDFs on grid elements might help to quantify the axis (work ongoing)

Figure 4: accuracy vs. impact on congestion for different forecasts

The position on the horizontal axis is strongly related to the concentration of the production (or load). If the units are concentrated in a restricted geographical area (such as the offshore wind parks), the impact on the congestions of a forecast error will be much stronger than a multitude of smaller units equally distributed on the territory (such as residential loads). The electrical localization and the associated costs of available remedial actions may also play a role in the x-axis determination. In order to quantify (not only qualitative indication) the impact on congestions for each of the forecasts the flow deviation on critical grid elements produced by a given deviation on each of the forecasts has to be calculated (PTDF analysis). For the final report, Elia will investigate whether this analysis should be performed.

From the figure 5 it can be concluded that the segmentation of elements leads to interesting results: the total load would appear as a large circle on the bottom-left of the graph (low RMSE and low impact on the congestions). By splitting the loads into different categories and by computing the average of all the individual RMSEs of each category instead of the RMSE of the aggregated time-series, one can see that the large industrial loads have a stronger impact on the congestions (they are concentrated) and their forecasts errors are high.

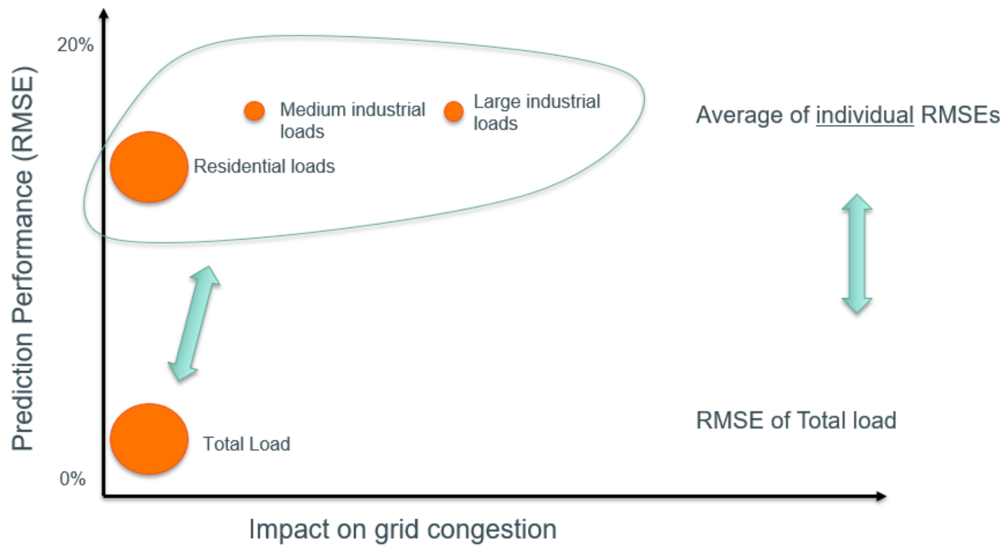


Figure 5: individual loads nodal accuracy comparison with aggregate forecast accuracy

Segmenting the “other decentralized units” leads to the same conclusion: the individual RMSE is higher and the influence on the grid is stronger for the large units (larger than 5MW).

A distinction must also be made regarding the RMSE of an aggregated category and the average of the individual RMSEs of the individuals from this category. In the congestion reduction paradigm, it is important to consider the individual RMSEs for the bigger elements since they are likely to influence the congestions.

3.3 Benchmarking

For some of the main forecasts (solar and wind), a benchmarking compared to different TSOs in Europe is made based on data available in the ENTSO-E database, see figure below.

However, comparing performances of forecasts among different TSOs is not as straightforward. Hereunder are some elements that can explain why a direct comparison might be misleading:

- The number of measuring points: not a single TSO has the exact historical measures of the total PV production of its bidding zone. At Elia this service is purchased from an external provider which uses about 80k measurement points to extrapolate the unmeasured ones. This approach is probably the most accurate way to compile such time-series but many TSOs are estimating it only based on realized weather data. Comparing RMSEs of two TSOs will certainly not determine which of them is the closest to the realized productions (the latter being unknown to both).

- Nodal versus total forecast: since the voltage levels operated by Elia are much lower than for most of the other TSOs (down to 36 kV) the information that Elia has is much more precise. It means that Elia must predict and model its grid in a much more detailed fashion.

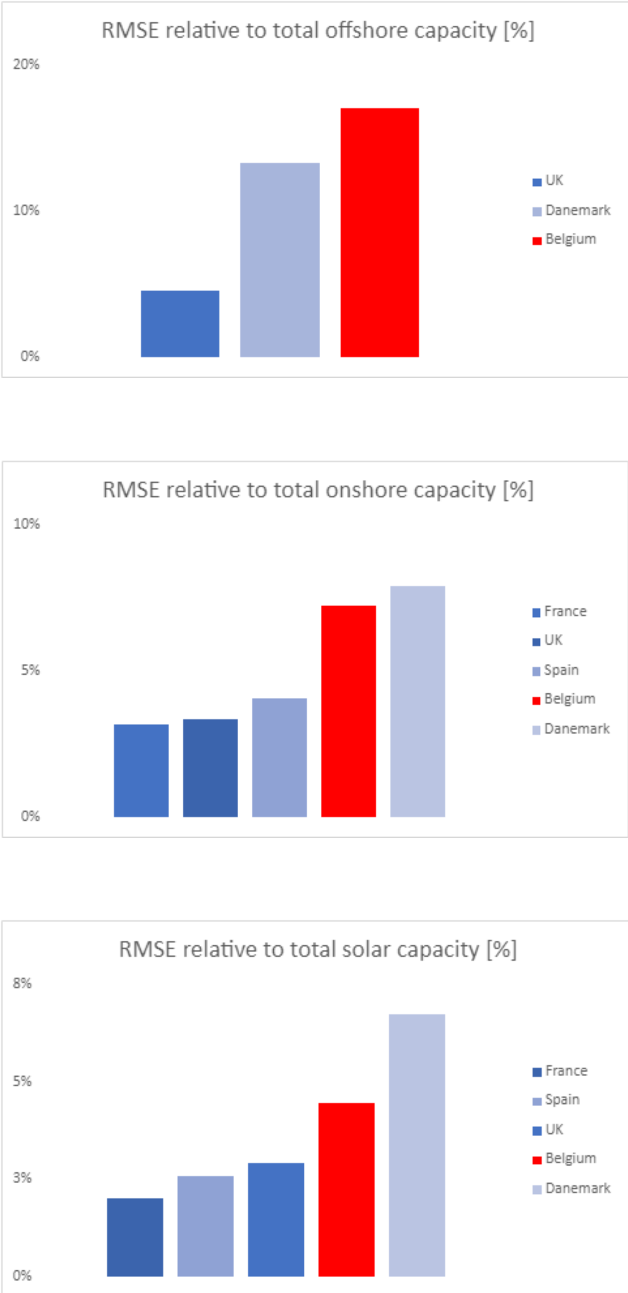


Figure 6: RMSE comparison for some forecast compared to other TSOs

However, the general trend seems to show that there is room for improvement for all renewable forecasts. The reader must keep in mind that the RMSE is just a partial indicator of the performance of a given forecast but for the sake of clarity this is the KPI that was compared for this benchmarking between several selected TSOs.

In order to improve the benchmarking, it would be needed to contact the individual TSOs in order to better understand their approaches and potentially learn from their methodology and evaluate whether it is compatible with Elia's forecasts.

4. Root-cause analysis and possible solutions

The different forecasts are analyzed on an individual basis, focusing on the following aspects:

- Infeed data, forecasting model, resulting forecast
- AS IS versus possible TO BE situation

For the possible TO BE situations, a first indication is given on the effort (low, medium, high) and impact on congestions (low, medium, high). This information will be taken into account later on in the short-term implementation plan and long-term roadmap.

4.1 Wind Forecast

- About 500 individual onshore parks amounting for 2700 MW listed in PISA database
- 85% of those parks are lower than 10MW
- Each park is modelled and forecasted individually by an external provider
- Each park is geographically and electrically precisely located
- RMSEs for onshore forecast at individual level is ~10% (~7% at aggregated level) and ~18% for offshore (same at aggregated level because all at a similar location).
- Impact on congestions is also high (also for some onshore wind parks because they are sometimes concentrated on the same area, saturating the lower voltage grid).

No game-changers are found yet to significantly improve this forecast but a structural data quality check could be realized on the onshore cadaster file. Fine-tuning GPS locations, installed capacities, measurement quality etc. could already improve the forecast accuracy at nodal level.

4.1.1 Infeed data

AS IS

- **Cadaster**: Installed capacities by GPS locations are necessary. More detailed data (number of turbines, turbine type/brand, hub height) were used in the past but with the evolution of forecasting algorithms (ML, AI, auto-tune etc.) those are not helping anymore and allowed to simplify the cadaster and concentrate on the inputs actually improving the prediction.

- **Historical data**:
  - **Offshore**: For every offshore park, the power output is directly measured. The power output is available both in real time and historical value.
  - **Onshore**: For onshore parks only 25% of the total installed capacity is measured in real time, the rest being up-scaled based on the average power factor of the measured ones. A couple of days later the tool receives accurate metering data for about 80% of the onshore parks. Those metering data have recently been added as an ex-post correction of the published estimation of the final produced energy as soon as they are available.
- **Wind curtailment ordered by Elia**: are shared with the external provider so that those event can be discarded from the training period for its predictive algorithm.

POSSIBLE TO BE

- **Compare the power factor:** The idea is to compare the power factor of the parks that are geographically close, the main benefit of this analysis would be to detect any error in the installed capacity or a structural error of the provider. This can be used to detect long-term maintenance. Moreover, Elia could improve the upscaling rule by applying a different upscaling factor based on the locations by using the power factor of measured parks in the vicinity of unmeasured parks or based on technology etc.)
  - **Impact low:** the other checks proposed below should probably solve potential issues. Applying a more elaborate upscaling might make sense but the trouble remains the lack of real time data.
  - **Effort low:** The work should not be too hard to do and can be done with easily available data.
- **Using VITEC tool to spot inconsistencies** in the cadaster data (fine-tune GPS coordinates & installed capacities, issues on measurement data, etc.). This tool displays the RMSEs of each individual parks and can identify some inconsistencies.
  - **Impact medium:** Elia has already spotted such issues in the past so a structural check for all wind parks would probably bring some benefit.
  - **Effort low/medium:** Elia could ask a student or starter to deep-dive into this relatively elaborate tool.
- **Compare real time measurements with ex-post metering:** This data is only recently available but Elia should check the coherency on both data sources.
  - **Impact medium:** Elia has already spotted such issues in the past so a structural check for all wind parks would probably bring some benefit.
  - **Effort low:** Both data series will be available in an internal tool at park level. This is not yet the case but will be available by end 2022.
- **Increasing the real time measurement ratio**: all TSO-connected units do have such measurement but Elia still struggles to receive it for DSO-connected units because they need to install an RTU (remote terminal unit) and meet complex requirements.

  - **Impact low:** it will greatly increase production estimation for the real time and the last hours in the past but as soon as Elia has metering data the added-value is gone. Considering the fact that forecast is mostly based on historical data (last 90 days) already including metering data for most of it, this improvement will only help for the short-term corrections of the forecast (the coming 2-3 hours). Increasing the "nowcasting" is interesting but still a bit late for being taken into congestion management.
  - **Effort low:** with iCAROS, a new protocol of live data exchange will be set up in order to receive type-B units. Elia expects to receive a lot more live data. Thanks to the Load and Generation project the mapping will be straightforward.
- **Including maintenance information**, certainly for offshore parks. Today Elia is using the forecast as input for the IGMs assuming that our forecast is in general better than those of the owners. Actually some owners might do a great job while others may not have an as elaborate forecast as Elia. Today Elia always supposes that the full park is available, the idea would be to scale the forecast based on the maintenance information.
  - **Impact low:** offshore parks are usually available when the wind blows and they usually stop a single turbine at a time but sometimes they also have big forced or planned outages of several turbines.
  - **Effort medium**: Using an internal tool to read and compare offshore nominations with offshore predictions at wind park level as well as the max available capacity provided by the BRP would allow the operators adapt the final forecast accordingly.
- **Using measurement data from other offshore parks in the surroundings:** in order to better forecast ramping events. When offshore wind changes rapidly it is a challenge for the prediction to be time-accurate.
  - **Impact low:** such event does not happen too often and IGMs granularity is hourly while the challenge is more on the 15min accuracy. Moreover, the errors only happen during the ramping events.
  - **Effort high:** need international collaboration and set up.

4.1.2 Forecasting model

AS IS

- Currently, **two providers** are supplying data for the wind power forecast.
  - The first supplier (IRM) focuses only on the offshore wind parks and still uses "classical" method (i.e. non Machine Learning) to provide the forecast. This is also a forecast optimized for the power predictions during storms.
  - The second supplier (VITEC) is used for the onshore production and uses machine learning and data from several weather forecasters in order to provide an accurate forecast.

➔ IRM offshore forecast is supposedly better during storms but Elia realized that VITEC outperforms during normal situations. The performance gap between the two providers is around 3% (18% RMSE for IRM and 15% for VITEC).

- VITEC: hourly updates. Buys weather data from many global model on specific weather stations near wind parks. Model retrains itself automatically every day based on the 90 last days. It uses very recent data to correct the forecast of the next hours to come. It removes outliers from the training period.
- IRM: 5 updates/day. Model is not retrained except on demand and it takes a lot of effort. No outliers detection, no machine learning.

POSSIBLE TO BE

- **Periodically challenge the suppliers:** The idea is to use existing in-house machine learning software and feed them with the data provided by the suppliers. Elia can then detect if there is room for improvement (or new strange behavior) for the forecast quality.
    - **Impact Low:** Elia organizes usually a call for tender every 3 years and selects the best candidates. Tracking performances on a shorter period remains valid to ensure that any issue is detected and solved as quickly as possible.
    - **Effort Low:** some basic machine learning tools available on the market and easy to use so the check should be quite fast.

  Same idea could be pushed further with a live monitoring dashboard, notifying users in case of strange behavior in order to avoid any issue to propagate for a long period of time.
    - **Impact medium**: Elia is focusing on forecast quality since a long time and each year the processes and tools are getting more robust, as such Elia can anticipate lower bug rates in the future.
    - **Effort high**: A live tool performing many automated checks and actions should be designed. On the other hand, such tool would benefit to all existing forecasts potentially.

4.1.3 Resulting forecast

AS IS

- **RPN imports latest forecasted data** from EFTool and send them to PF at the exact electrical node each time an IGM is created
- **RPN/PISA/PF sync** -> see below on other decentralized productions (for AS IS and possible TO BE)

POSSIBLE TO BE

- **Choose supplier in function of wind speed:** in order to keep IRM for storm situations (when wind speed is > 20m/s for example) and use VITEC otherwise.
    - **Impact High:** The expected outcome is high, Elia already tried to push IRM to use ML and other possible improvements but their tool is not easily configurable/editable.
    - **Effort Low:** VITEC could do the combination in their tool upfront.

4.2 PV Forecast

- In January 2022, 8800 individual[3] units amounting for 4690 MWp[4] were listed in PISA tool while the estimated power was more around 5900MWp. 500 virtual power plants spread over the grid are filling the missing installed power.
- All those units are connected on lowest voltage levels available in our model (low influence on flow forecast on Elia's grid) and the DA RMSE for 2021 is ~4% (if computed only for daylight hours) and ~1% (if computed for all hours).
- Elia buys state of the art estimated measures and forecasts from an external provider.

There seems to be little room for improvement but considering the upcoming increasing of PV units Elia should ensure the same level of accuracy in the coming years.

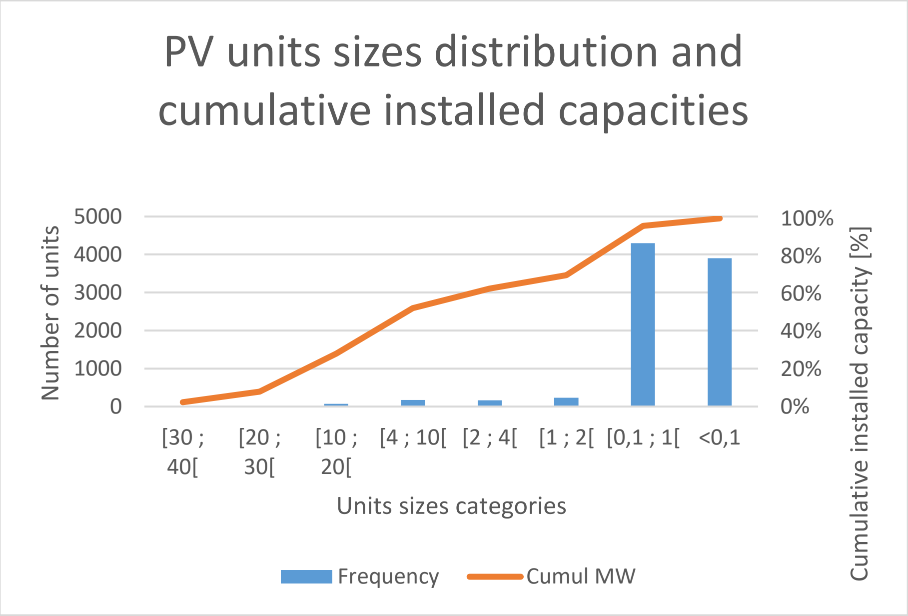


Figure 7: Installed capacities of PV units in Elia grid model

[3] Those units are often an aggregation of even smaller PV installations.
[4] Elia uses MWpeak (MWp) for the time being because there are no indications for the time being that inverters would limit power outputs in Belgium.

4.2.1 Infeed data

AS IS

- **Cadaster management**: Installed capacities at communal-level (~400 communes) + filter TSO/DSO-connected[5] are required[6]. Cadaster is generated by PISA on demand. Based on information of authorities[7] Elia can estimate missing capacities and scale existing units to match with a regional total target. Estimation of monthly capacity increases by region is then applied to anticipate monthly evolution of the cadaster. This is predefined in the cadaster file sent to the external provider and EFTool meaning that updates will be applied automatically at the predefined dates (no need for cumbersome coordination with external provider and our own tool). Twice a year Elia updates the cadaster for the 12 months in the future based on an extract from PISA + reevaluation of best estimates from the authorities. Hereunder some numbers for January 2022 are listed as an example (**gap of 1200MW = 20%** between PISA and best estimate).

	PISA (JAN - 2022)	ASIS JAN 2021	Best estimate JAN 2022	Best estimate JAN 2023
Flanders	3598	3470	4300	4600
Wallonia	980	1263	1500	1700
Brussels	110	55	130	140
TOTAL	4689	4788	5930	6440

- **Historical values**: are themselves a forecast based on a subset of measured installations. This is provided by an external provider[8]. The upscaling is done based on 80k measurement points amounting for more than 1GW of installed capacities. They are covering all regions in Belgium and all installations' sizes. On paper this is the best Elia can hope for as methodology to recreate the estimated produced energy. The only downside is the black-box effect and our incapacity to challenge the received values.

[5] TSO/DSO information does not improve in any ways the quality of the flows forecast, relevant for publication & settlement purposes.

[6] In 2021 it was still necessary to distinguish <10kVA units from the rest (need of Synergrid). This is finished and make it easier for Elia to upgrade cadaster. We can now use PISA as main source.

[7] **Carte dynamique (solaire et éolien) de la Wallonie - Site énergie du Service public de Wallonie** and **Cijfers energiekaart - Energiesparen**

[8] Energy Meteo and Services

POSSIBLE TO BE

- **Reducing the gap from PISA** by identifying the troublemakers and simplifying the data exchange procedure for the DSOs.
  - **Impact low:** current upscaling method is not shocking for such types of units but it would be valuable to have at least no medium/big units missing in our models because those generate clear local disturbances. Elia can anticipate that results will not be perfect considering the fact that this level of information is probably not always known by the DSOs themselves.
  - **Effort low:** Advocating for DSOs to increase installed capacity coverage is not really on Elia's hand but it will not be a big effort for Elia to identify the troublemakers if any and improve the collaboration and data exchange with them.
- **Smart meter data:** could increase coverage of measured installations to improve estimation of realized production.
  - **Impact low:** current coverage is already great.
  - **Effort high:** smart meter data have usually no information on the PV production isolated. Often this is combined with consumption data making the exercise worthless.

4.2.2 Forecasting model

AS IS

- **External provider**[9], selected among 6 pre-selected candidates on a live trial run of 3 months. RMSE is typically used to evaluate forecasts performances in such selection. Elia shares the real time data (in this case from same provider) and the cadaster file as inputs. The external provider buys weather data near the installed units (in this case for all units in Belgium). Usually such provider buys those weather data from several providers/weather models and launch many forecasting methods in parallel with different approaches. Then the provider combines all those forecasts in order to minimize RMSE (with a special attention to DA forecast).

POSSIBLE TO BE

- **Combining several providers**: is something Elia already tried without success. Nowadays external providers are already doing this upfront. They buy many independent weather data from different weather models and they use many forecasting algorithms in parallel. They combine all their different predictions with ML and AI in order to optimize RMSE (or other parameters on demand).
  - **Impact very low:** give a backup in case of issue but providers have high reliability standards (by selection).

---

[9] Energy Meteo and Services

- **Effort high**: at least doubling the costs for buying the data + IT/Process effort for combination.

4.2.3 Resulting forecast

AS IS

- Same as for other decentralized productions (see hereunder) except that Elia can easily upscale the installed capacities of some 500 virtual units in RPN and PF in order to match with total target of installed capacities. Those virtual units are spread in function of the potential of future installations (Elia can assume a good hypothesis). This way IGMs are following correctly the total capacities with a relatively good repartition even if a full synchronization of PISA/RPN/PF is not performed.

POSSIBLE TO BE

- See other decentralized productions

4.3 Other decentralized productions (non-wind, non-PV)

- 2152MW installed capacity by 2022
- 900 individual units (each of them modeled in PowerFactory) and 1750MW of CHPs.
- This category is slowly but continuously increasing, so it becomes more and more interesting to invest efforts to improve this forecast. This forecasting model has not been changed since many years.
- Battery storages are progressively entering into this category, Elia needs to anticipate how to best forecast them.
- RMSE ~5% for the aggregated category, for the units with a power output larger than 5 MW (the one with the strongest impact on grid congestion) the average individual RMSE is 15%. Considering the size and numbers of such units it seems that building individual models for the most problematic units would make sense. In parallel, Elia could improve the current methodology for the other units.

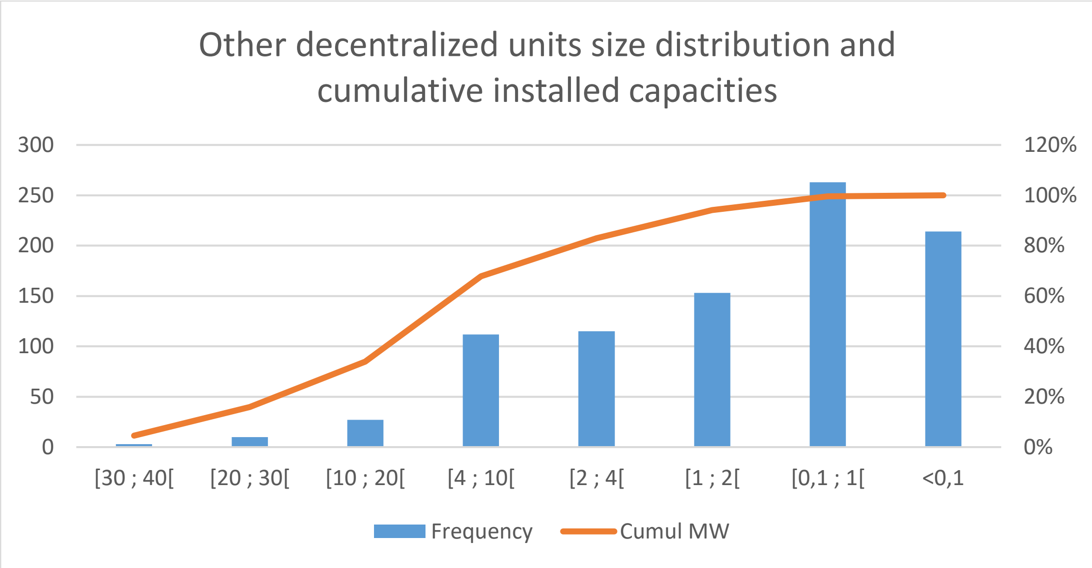


Figure 8: Installed capacities of other decentralized (non-wind, non-PV) units in Elia grid model

4.3.1 Infeed data

AS IS

- **Cadaster:** installed capacities + production profile + electrical location + Mnemonic TIC (metering point) are necessary and are updated typically once a year based on PISA reference. PISA contains a very good reference for those type of units; Elia assume it is almost complete.

➔ **NEW:** this has been improved since early 2022 with updates possible at any frequency thanks to IT projects and process optimization. One button to generate cadaster file by filtering “InService”, not PV, not Wind, type A or B. The idea is to trigger an update each time there is a delta of 50 MW to avoid unnecessary workload.

- **Historical values for each individual units:** ~65% of total installed capacity is effectively metered. Reminder: metering is not a real time measurement; the data is accessible in our systems with at least 1-day lag, sometimes longer.
- **Production profile**: the historical measurement estimations of unmetered units are an upscaling based on the power factor of all metered units from the same production type. Here the production profile currently in use: batteries, CHP large, CHP small, Run of river, incinerators.

$$UM_{unit} = \frac{\sum_{Profile} P_{meas}}{\sum_{Profile} P_{installed}} * P_{inst,unit}$$

POSSIBLE TO BE

- **Directly link ("1-to-1") between TIC and PISA** units to avoid double counting (also true for onshore winds). This is under discussion with metering services to evaluate feasibility.
    - **Impact medium:** not so many units concerned and most of them are manually corrected.
    - **Effort low:** depending on metering services possibilities, alternative is to add and maintain a new field in PISA with weighting factors.
- **Add missing metering** for the biggest units (e.g. add TIC for 30 biggest missing units -> coverage goes > 90% having a TIC) + ensure that existing TIC are functioning correctly + ensure that metering services are systematically requesting metering data for all units big enough.
    - **Impact medium:** already 65% coverage but biggest units have sensible effect on the grid because concentrated in one electrical location).
    - **Effort low/medium:** depending were we put the threshold.
- **Using real time data** (measured or estimated by the state estimator) instead of metering data to make the forecast.
    - **Impact medium**: will improve the short-term forecast but also allows hourly updates of all time-horizons forecasts.
    - **Effort medium**: thanks to Load and Generation project, PF and EMS granularity will be aligned, data will be accessible and thanks to iCAROS project, simplified communication protocol will allow to receive more real time data.
- **Improve production type categories:** CHP baseload vs CHP stopping during weekends and nights, units usually not producing (typically used as manual Frequency Restoration Reserve, mFRR), …
    - **Impact low**: only valuable for unmetered units (35% today but will hopefully reduce in the future).
    - **Effort medium**: Need to adapt PISA accordingly (no IT development) and verify that all users are fine with this change or apply the categorization-logic into EFTool (IT development)
- **ON/OFF nomination of B type units:** this will be available by the implementation of iCAROS phase 2.
    - **Impact low**: quality of those nominations still to be demonstrated but it is certainly a valuable input data because OFF status are practically impossible to forecast in most cases.
    - **Effort medium**: no IT structure today to feed model with this input.
- **Add market data:** typically the spot price of DA/ID market could be relevant for a better prediction of such units' behavior. If prices are very high, some might produce more than usual and vice versa. A quick view on the graph below shows the influence of several factors on the production of such units. One can see that the price is a strong driver but there is a multitude of other parameters. A more detailed analysis is required in order to be able to improve the forecast.
    - **Impact medium:** quite marginal effect today but probably progressively bigger in the future.
    - **Effort low**: quite a basic data easily accessible.

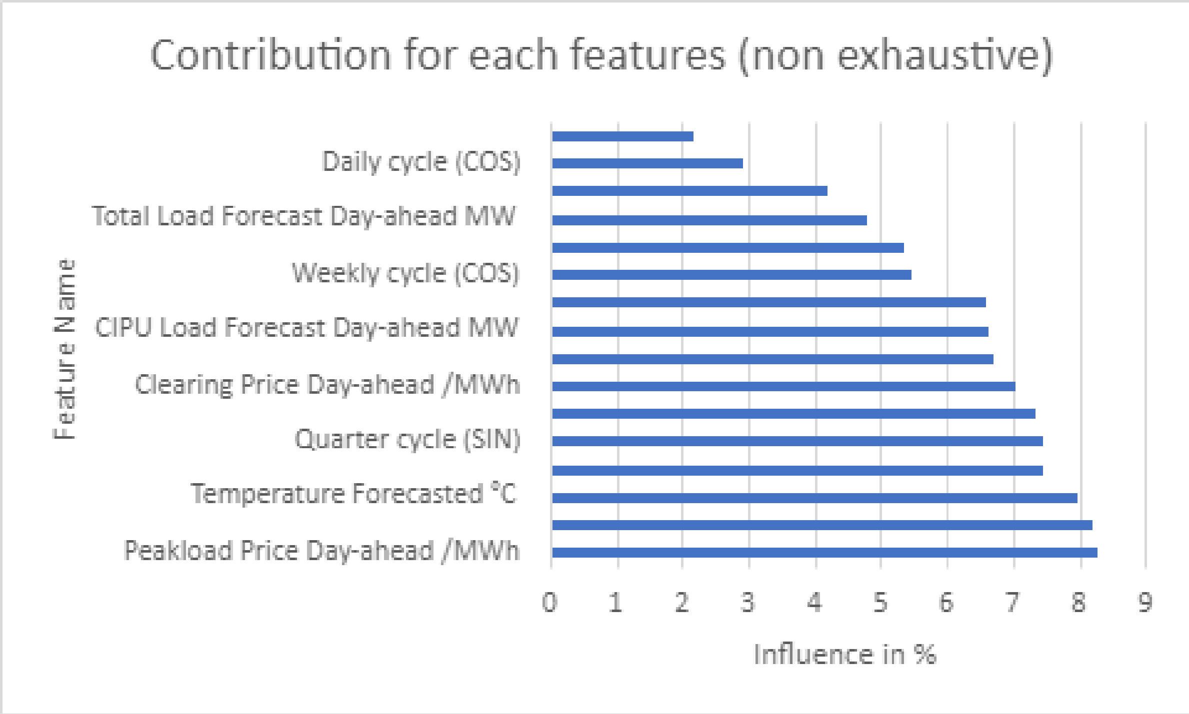


Figure 9: Identified drivers for forecasting other decentralized (non-wind, non-PV) units production levels

4.3.2 Forecasting model

AS IS

- **Internal tool** with basic algorithm. The tool looks at the **3 most recent days** with metered data available from the same calendar, category and for each quarter of an hour to be forecasted it computes the average of the power factor for each profile and then apply it to each unit.

$$Pforecast_{unit,QH} = \frac{\sum_{For\ each\ refday} \sum_{Machines\ with\ same\ profile} Pmeas_{QH}}{\sum_{For\ each\ refday} \sum_{Machines\ with\ same\ profile} Pinst_{QH}} * \ Pinst_{unit,QH}$$

  ○

- **Update** is done once at 6h24 for values extending from “Now” to day + 7.
- **Timers** for data imports of the most recent metered data (metering data are not real time, they are mostly available 1 or 2 days ex-post).
- **Calendar days** (categories used: Weekdays, Saturdays and Sundays).

POSSIBLE TO BE

- **Switch from a forecast per profile to an individual forecast.** The tool linearizes the individual forecasts based on their profile, which makes little sense. The tool could simply apply the same logic at an individual unit level :
    - <u>Pforecast_unit,QH = sum_for each ref day (unit Pmeas_QH)/number of ref days</u>
  - **Impact high**: this would greatly improve the forecast at nodal level because for sure all units within any given profile do not behave the same.
  - **Effort low:** keeping the same tools and main logic, just adapting some formulas.
- Look for **external** partners
  - **Impact high:** for sure external providers will be able to meet our need and improve significantly current model.
  - **Effort high:** high costs w.r.t. todays' + high workload to organize tender and selection + lot of changes in current processes. Considering the fact that Elia has all the main drivers in-house it would make little sense to look for external providers like Elia typically does when complex weather data are required.
- Adding **ML/AI** modules to existing logic
  - **Impact high:** this will most probably correct many issues notably from the inputs perspectives.
  - **Effort high**: not feasible in current tool.
- Add **market data** to existing logic: correction of forecast based on spot prices in the past and next day.
  - **Impact low**: probably not a good way to apply the same correction per production profile, this is more specific to each unit.
  - **Effort high**: current tool is basic and has no built-in forecasting modules.
- Add historical values and nominations of **pump-turbine units and big batteries (scheduled ones)**. They could be good drivers/proxies for small batteries production forecast but this is for the time being impossible to study by lack of data. A mitigation measure could be to foresee the possibility in the tools and processes to feed on those inputs when batteries penetration will be higher.
- Building **individual forecast** with off-the-shelf forecasting modules and some basic inputs (market prices, total load, historical data) and stop with current referential days basic approach. Hybrid situation could be also imagined were only biggest units are individually modelled (e.g. 40 biggest units = 35% of installed capacity) while the rest would still be forecasted with current method (with still other identified improvements).
  - **Impact high:** this will most probably correct many issues notably from the inputs perspectives. Concrete numbers are not yet known, certainly not if other identified possible improvement of the current logic are also taken into account. However, our findings on load forecasts showed already good results while such units are strongly linked to load behaviors.
  - **Effort high**: Internal tool is more a data handling tool than a forecasting tool and it would be probably necessary to make this module on a separate tool. Of course such a tool would be beneficial for other forecasts as well.

- **Increase updates rate:** ideally hourly updates but only if this makes sense (in this case also needed to have input data on hourly frequency).
    - **Impact very low**: TIC values are not updated more than once a day. Only relevant if model uses other inputs with smaller time granularity (real time data, market data, etc.).
    - **Effort low:** just a tool configuration.
- **Improve profiles definitions** (see in input section)
- **Change the 3 last reference days rule** (weighted average with more weights on more recent days, using only reference day, …)
    - **Impact low**: still to be assessed, but impact expected to be low.
    - **Effort low:** minor changes of the code.
- **Fine-tune timers** in order to have more recent data available (today 2 days delay) even if not all metered data are yet available.
    - **Impact low:** current times = 2-days delay between availability of metered data vs. real time while lot of data available in day+1.
    - **Effort low:** just changing configuration of current tool (no IT development) but some analyses to be done to find optimum solution.
- **Improve calendar days categories**: bridge days = Saturdays, Bank holidays = Sundays, 1st of January = special day, Christmas holidays = special holidays, …
    - **Impact medium**: current logic is very basic, e.g. 1st of Jan = normal day while very special day in practice. In this case improvement only valid for those special days (~10% of the years).
    - **Effort medium**: small analysis to be done but probably very similar to current load categorization (CHPs are strongly linked to industries/loads). Then some IT developments (minor code changes) to be foreseen.

4.3.3 Resulting forecast

AS IS

- **Availability/centralization**: forecasted values, per quarter of an hour, from real time to D+7, per individual units are made available for all operational tools.
- **RPN imports latest forecasted data** from EFTool and sends them to PF at the exact electrical node each time an IGM is created.
- RPN & PF production units[10] cadasters are **synchronized with PISA** once every year. There are many IT issues for this process, this is well-known and efforts are being done to smooth this process in the future. In

[10] All units are concerned (PV, Wind, other RES, centralized)

the meantime, when updating the cadaster in EFTool, new units are not pushed to IGMs until next synchronization with PISA/RPN/PF. Nonetheless, it remains pertinent to regularly update EFTool cadaster for other processes.

POSSIBLE TO BE

- Smooth process of **synchronizing PISA with RPN and PF** and increase frequency of synchronizations.
    - **Impact high:** structural errors due to this lack of synchronism.
    - **Effort high**: IT challenge is not to be underestimated, notably because the long-term planning department is using the same tool and models but for all future horizons and such update should be validated for all the future years too.

4.4 Load Forecast

There are two important aspects here:

- the **aggregated** total load of Belgium and
- how to ventilate this total load into about **1000 individual loads** at lowest possible nodal level.

10GW total load on Christmas Eve is most probably different from 10GW on Monday morning in April. The contributions of each individual load are not the same. Currently Elia is using 16 different load repartition keys to capture the main behaviors at play. A recent study with an external data scientist confirmed that those 16 vectors are capturing most of the behaviors and constitute a good compromise between accuracy and operability. The 16 vectors are the result of the following categorization: day/night, weekday/weekend and the 4 seasons (2 x 2 x 4 = 16 vectors).

Today the total load is relatively easy to predict (RMSE ~3% in 2021) based on some simple inputs and the impact on congestions of the total load forecast errors is low due to the diffusion of this error into all individual residential loads (~6GW residential load on yearly average). Indeed industrial loads are fixed by the selected repartition key while the residential loads are scaled to meet the total load target during IGMs creation.
However, in the upcoming years, Elia expects important changes in the load contributors: electric vehicles penetration, the increase in heat pumps, the load flexibility development and adaptation of consumption based on market prices, ... All those factors will probably deteriorate the RMSE of the total load forecast if the same methodology is kept while at the same time the total load will only increase in the future, increasing its impact on congestions altogether. Nonetheless, by the lack of data, those future evolutions are hard to anticipate accurately. Elia preconize to keep an eye on those evolutions in upcoming tenders for external forecast providers and to make sure that our tools and processes are ready when needed.

Regarding the individual load forecast Elia calculated that RMSEs are on average around 15% for residential loads and around 17% for big industries (>40MW). Applying individual forecasts for the highest loads seems promising w.r.t. influence on congestion management. One of the findings is that for industries it would only make sense if Elia has access to very recent past data (day-ahead typically) to feed the model (ML 2 in the table below). For residential loads

Elia sees that this prerequisite is less important. Results also indicate that schedules received for big industries still outperform the ML 2. For these industries, it would make perfect sense to receive and use the schedules, having a demonstrated impact on congestions.

RMSE [%]	Large ind	Residential
AS IS	16,9	15,1
ML 1	16,4	12,6
ML 2	13,4	12,1
Offtakes	12,1	/

Figure 10: RMSEs for several forecasting methods

4.4.1 Infeed data

AS IS

- **Historical data** of total load are computed by Elia and shared with the external provider via a live stream.
- **Equivalent temperature** is computed by the external provider based on some simple weather data.
- **Calendar days**: some special events are highlighted and categorized by the external provider in order to best cope with special days (bridge days, holidays etc.).
- **16 repartition keys** calculated once a year based on last year metering data. The average values of a given individual load for the selected timestamp in the past is computed and used in the repartition key. For some special industries a worst case load is chosen to best represent and anticipate the potential congestions (typically for electric oven when the average would be much lower than their peaks while they typically reach their peaks on a daily basis).

POSSIBLE TO BE

- **Market prices** will certainly become a crucial driver for total load prediction even if today no such correlation exists.
  - **Impact unknown** (lack of data)**.**
  - **Effort low:** during next selection of external provider Elia will impose that such input is taken into consideration.
- **Recent past data** can clearly help individual load models when unpredictable behavior on loads occurs. This could be demonstrated with ML2 model on industrial forecast accuracy.
  - **Impact medium/high:** depends if Elia has access to latest QH data or if Elia has only access to the data with some time lag.
  - **Effort medium/high:** to have data from the latest QH requires a perfect alignment between PF and EMS models. This is foreseen in the Load and Generation project. Having the data with some time lag would still require many adaptations in our tools and processes.

- **Historical data of different categories** as electric vehicles, heat pumps, residential batteries etc. It could improve the total load forecast. Specific models could separately forecast each of those categories and then combined by a total load forecast model.
  - **Impact unknown** (lack of data).
  - **Effort high:** this represent many data while the latter are not directly connected to Elia grid.
- **Schedules of biggest industries** are certainly helpful when some unpredictable behaviors are at stake as typically revision periods, change of typical consumption patterns due to supply chain issues or other externalities. The ML/AI models Elia has tried, could not beat most of the schedules received for industries with a load superior to 40MW. This indicates that for the biggest industries, those schedules could be either used as final forecasts or eventually used as input parameter for a ML/AI model. Request
  - **Impact high**: Some industries have already load levels of several hundreds of MWs, with the progressive electrification of more industrial processes such loads will only increase in number.
  - **Effort high**: requesting such schedules represents a lot of work for the asset owners and should be set in place only in case if a clear link exists between congested elements and the set point of those industries. Such selection would make sure we only ask schedules for the biggest industries (the bigger it is the greater its influence on flows) located in congested areas.

4.4.2 Forecasting model

AS IS

- Elia buys the total load forecast from an external provider, the trial phase allows to make a good selection based on performances and prices while the 16 repartition keys are well performing for a top-down approach (i.e. forecast of total load then ventilation among individual loads). Consequently, there is little room for improvement if we do not change the paradigm altogether.

POSSIBLE TO BE

- Use **machine learning to forecast the loads individually**. With such approach, RMSE at nodal level can be reduced for both industrial and residential loads as shown in figure 10.
  - **Impact medium/high:** For the biggest loads, the impact will be tangible already today. More importantly, the new types of loads (e.g.electric vehicles) will introduce more exotic behaviors in the future while those are typically well detected by machine learning (there is no need for complex modelling if this is confirmed). With the electrification of industries, mitigating the error on such forecast might be more and more valuable.
  - **Effort Medium:** The main challenge is to integrate this in the existing IT infrastructure. For industrial loads, it seems that live stream of recent data is key to really improve the forecast quality. This requires much more efforts in terms of model alignments and data acquisitions.

- **Launch tender** to find a new provider and explore the possibility to do an **in-house** load forecast with AI Center of Excellence and 50Hz.
    - **Impact Medium/High:** Elia might see an improvement in the performances but Elia could also ask the provider to offer new services such as categorization of the loads in order to prepare the future (increase in electric car, heat pumps, load elasticity with respect to prices, etc.)
    - **Effort Medium:** Enlarging the scope of the forecast requires more data while those are not directly available at Elia (see 4.4.1).
- **Use categorization** in order to improve the forecast (Proof of Concept with an external data scientist). This approach is another way to tackle the problem of different types of loads compared to the machine learning approach. Forecasting each load individually with the additional information that they belong to the same cluster (i.e. they show similar patterns), may sometimes further improve the final individual forecast.
    - **Impact medium:** It is hard to quantify the gain for such a method before doing it but with the development of new usage, this might be more important in the future.
    - **Effort medium/high:** The effort seems harder than for ML because it requires a more in-depth analysis in order to categorize the usages (less automatic than ML).

4.4.3 Resulting forecast

AS IS

- Industrial loads are fixed by the selected repartition key while the residential loads are scaled to meet the total load forecast.

POSSIBLE TO BE

- **Combining total load forecast** with the total load forecast reconstructed with all the day-ahead & intraday market positions. A machine learning algorithm could be used, in addition with other infeed parameters to best combine both time-series.
    - **Impact low**: if total load forecast is well optimized upfront, there is a fair chance that the total load based on the market positions will not contain much additional information.
    - **Effort low**: an off-the-shelf ML/AI model could set this up.

4.5 Conventional units

Nominations are directly used for DACF and IDCF files (no forecasting model). The quality of those nominations has been high in the past but this should be closely monitored in the framework of the progressive balancing obligation relaxation and higher uncertainty on DA/ID horizon for BRPs (intermittent prod etc.) meaning that decision making are pushed closer and closer to real time.

4.5.1 Forecasting model

POSSIBLE TO BE

- **Forecast of intraday market moves** based on open position of Belgium, DA market prices and offer & demand curves, nominated Belgian redispatching bids (volumes and prices). A proof of concept has been set-up in 2021 to evaluate feasibility and performances.
    - **Impact Low:** Nowadays nominated volumes remain most of the time very accurate. For some extreme situations the proof of concept could anticipate some moves correctly but changing the official nominated power cannot easily happen in an official IGM (net position is imposed by the merging process).
    - **Effort High:** many new inputs and scripts are required to make it work in daily operations but feasibility has been proven. If impact is starting to grow due to increasing uncertainty in day-ahead horizon, Elia can always reopen this initiative.

4.6 Net position

Net positions resulting from nominated power exchanges between European countries is used as a fixed parameter for DACF/IDCF files. For D2CF files the net position from a reference day is applied but this has evolved since the FB DA CORE go-live early June. A centralized forecast is predicting the net positions for all CORE countries. Elia applies the resulting net positions of Belgium directly into its D2CFs. RSCs are managing this forecast and it is not part of this exercise because it only concerns the D2CF files.

4.6.1 Forecasting model

POSSIBLE TO BE

- **Net Position correction** based on open positions and on latest market prices curves from all European countries could be done at RSC level. Some most probable exchanges could be anticipated in order to reduce the open positions and as a result improve the net positions imposed by the CGMs process.
    - **Impact medium:** nowadays open positions of Belgium and surrounding countries remains acceptable. A proof of concept could be done to assess feasibility and gains.
    - **Effort high**: this should be done at RSC level with a close collaboration of as many as possible TSOs.

4.7 Grid topology

Outages (planned and unplanned) are all included in the IGMs of Elia as well as their necessary preventive topological actions. PSTs taps are mainly at neutral taps for D2CF (actually 15/15/15/15 to anticipate structural loopflows) but this is of course only an initial state and optimized PSTs taps position will be taken to increase Flow-based domain in estimated market direction.

For the initial version of DACF, PSTs are already optimized by looking at the market flows ("inputs") resulting from FB MC. Of course this is a rough estimation ("model") considering the fact that the loopflows are not yet known at this stage. During the DACF process iterations, PSTs will be adapted to relieve eventual congestions on Elia's grid or even abroad.

4.8 Nemo Forecast

The forecast on the Nemo flow is only relevant for D-2 time horizon because later on there are deterministic nominations. Previous model for the Nemo forecast was bought from an external provider but since early June, with the go live of FB CORE DA, this model has been replaced. The new model is outperforming the previous one mostly because it is based on a broader range of infeed data. This forecast is not anymore into the hands of Elia and it seems that there is not much room for improvement.

4.9 Dynamic Line Rating

Dynamic Line Rating (DLR) allows to adapt the rating of equipped lines based on the local weather conditions (i.e. cooling of the line). On average this means an increase of capacity w.r.t. the more conservative static line rating. Of course such technology never ensures that the capacity will be increased when needed. Using DLR is actually not improving the congestion management since it adds a new uncertainty into the system. The only positive effect on congestion management is when the available capacities are progressively increasing closer to real time. This is often the case because the uncertainty tends to reduce when closer to real time allowing the system to grant higher values while keeping the same risk level.

4.9.1 Infeed data

AS IS

- **Real time measurements:** About 30 lines are currently equipped with Ampacimon modules and a dozen underground cables have the real time thermal rating installed. This represents all the typically congested lines for which the bus bay elements are not limiting the DLR.
- **Lines/bottlenecks selection**: selection of element to be equipped, is based on a techno-economic study. Estimated avoided costs and installation costs are the main drivers but Elia also looks at the benefit of such investments in terms of risk mitigation. Typical candidates are lines monitored by the market coupling. Long infrastructure works, typically for High Temperature Low Sag (HTLS) conductors reinforcements, are also good candidates during the outages period. The long-term planning department also identifies future candidates many years in advance based on their security analyses results.
- **Temperature forecasts**: are used to anticipate which reference rating should be applied for each IGM. Indeed Elia changes the reference ratings of its grid based on the measured temperature increasing in average the

available ratings w.r.t. the previous static rating approach (same reference rating applied by fixed predefined periods).

POSSIBLE TO BE

- **Improving the long-term detection criteria's:** DLR never ensures that capacity will be increased when needed except when the congestion is clearly linked to high wind production. It is then rarely possible to consider such technology to release an identified bottleneck, it is only a good (but costly) mitigation measure for temporary situations. Consequently, DLR installation adds costs without ensuring that the congestions will be released meaning that most of the installations of DLR are validated quite late with the consequence that sometimes bus bays must first be reinforced delaying the installation by many years. It could be more optimal to set up a process identifying possible candidates far in advance and making sure that their bays will not be limiting when it is estimated that there is a risk of congestion on this element. By doing this Elia will make sure that if the bottleneck materialize itself, Elia will be able to quickly install DLR (could be done in 6 months).
    - **Impact high**: this could be very helpful to anticipate the Ampacimon installation, certainly in the case of limiting bus bays. Changing bus bays elements may take years and should be planned coherently in the infrastructure 4 years-ahead plan.
    - **Effort medium**: ideally long-term planning could generate a list of most loaded expected elements for each year in the future. Those would be natural candidates for DLR and at least evaluation of need and feasibility could be launched well ahead of time.
- **Irradiation forecast:** could be used to further fine-tune the applied rating of each element of the grid. The simplest way to implement it, could be a day / night forecast as input. This would typically increase the rating every night while keeping the current rating during daylight. More advanced models could be envisaged based on a real irradiation forecast during days.
    - **Impact medium**: such approach has the advantage to increase ratings of the entire grid at once and we know for sure that this would only mean higher or equivalent ratings
    - **Effort medium**: a simple rule could be applied on the short run by using the same approach as the one done with temperature-based ratings.
- **Temperature and irradiation behavior by conductor type:** instead of by element type and voltage. Today Elia clusters elements by their types (cables, lines, transformers etc.) and by voltage levels. For each of these categories Elia takes the less performing element w.r.t. temperature and irradiation behavior as the reference for the entire category. By doing this Elia underrates most of the grid. To tap all the potential Elia should enrich the models (Power Factory for planning and EMS for real time grid operation) with all those necessary information.
    - **Impact high**: again this will boost the capacities of almost all elements of the grid.
    - **Effort high**: this represent a huge work in terms of data handling and tools improvements.

4.9.2 Forecasting model

AS IS

- **Ampacimon forecast:** uses the past measured ratings, the past local weather measures and the past global weather predictions to build a predictive model. The provider trains the model on a fixed 3 months period manually selected. Only 14 elements, those being monitored in the Flow-based, have such forecast due to economic optimization. For the remaining lines without this forecast, it is up to the operators to do it by hand based on recent data and weather predictions. Consequently, a margin needs to be applied in order to ensure that the used rating will not be too optimistic w.r.t. the rating effectively available in real time.

POSSIBLE TO BE

- **Revamping the Ampacimon forecast**: Ampacimon is presently working on an all-new forecasting model with:
  - machine learning and artificial intelligence embedded,
  - improved learning periods (either the last 3 years of available data or maybe a rolling window of 90 days in order to better follow the seasonality)
  - merging their 2 days-ahead forecast with their short-term forecast (now up to H+6).

  - **Impact high**: this new license is expected to provide decent gains, certainly for DA and ID timeframes.
  - **Effort high**: this is a complete rework of their product but of course the work for Elia is very limited.
- **Installation of new short-term forecast (H+6 and less, ongoing):** Operators might use the forecast 6h as a reliable source for their decisions making in terms of congestions management. Present version stops at 4h-ahead and is not reliable enough to be used.
  - **Impact medium**: this manual use of a reliable forecast 6h can be an interesting decision support, but cannot be easily implemented automatically into our IGMs without IT developments. Elia will probably wait for the 2 days-ahead forecast improvement for the automatic inclusions of the forecasts into the IGMs.
  - **Effort medium**: new version of this forecast is still to be deployed, and then a statistical study is needed to determine its reliability.

4.9.3 Resulting forecast

AS IS

- **Capping rules:** in order to ensure a 99% reliability of the forecasted values in the IGMs, forecast results provided by Ampacimon are capped to 105% of the applied reference rating, before putting it automatically in the IGMs. This limits the potential of the DLR, but is mandatory due to the weak reliability of the current results.

POSSIBLE TO BE

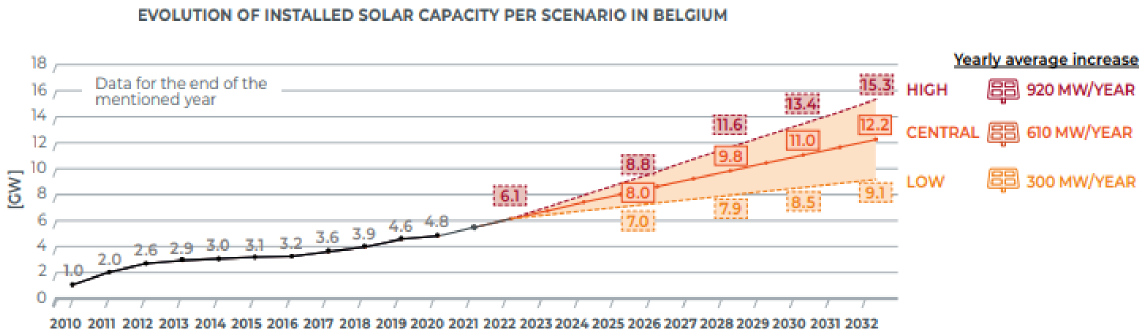
- **Increase of the capping factors**: this improvement is linked to the Horizon forecast improvement mentioned above but Elia can already anticipate that the ID caps will be greatly released based on some internal studies.
    - Impact high: this improvement will allow to better align the real time ratings and the forecasted ones used in IGMs, facilitating a lot the decision making when preparing costly remedial actions.
    - Effort medium: once the new version of the 2 days-ahead forecast is available, a statistical study is needed to determine the new capping factors while conserving the same reliability level.

5. Challenges and opportunities for the future

In order to build a short-term implementation plan and long-term roadmap it is important to keep in mind future challenges and opportunities and to make the link with ongoing projects. Some challenges trigger improvement of forecasts on short-term, while others indicate that it is better to wait before starting any implementation.

5.1 Increase of installed capacity of decentralized production

The expected growth of each forecasted data can help to anticipate the future troublemakers in the flow forecasts from IGMs. It is well-known that all PV, onshore and offshore wind productions will increase dramatically in the coming years (see figures below from the adequacy and flexibility study for Belgium[11]).



[11] Link: https://www.elia.be/-/media/project/elia/shared/documents/elia-group/publications/studies-and-reports/20210701_adequacy-flexibility-study-2021_en_v2.pdf

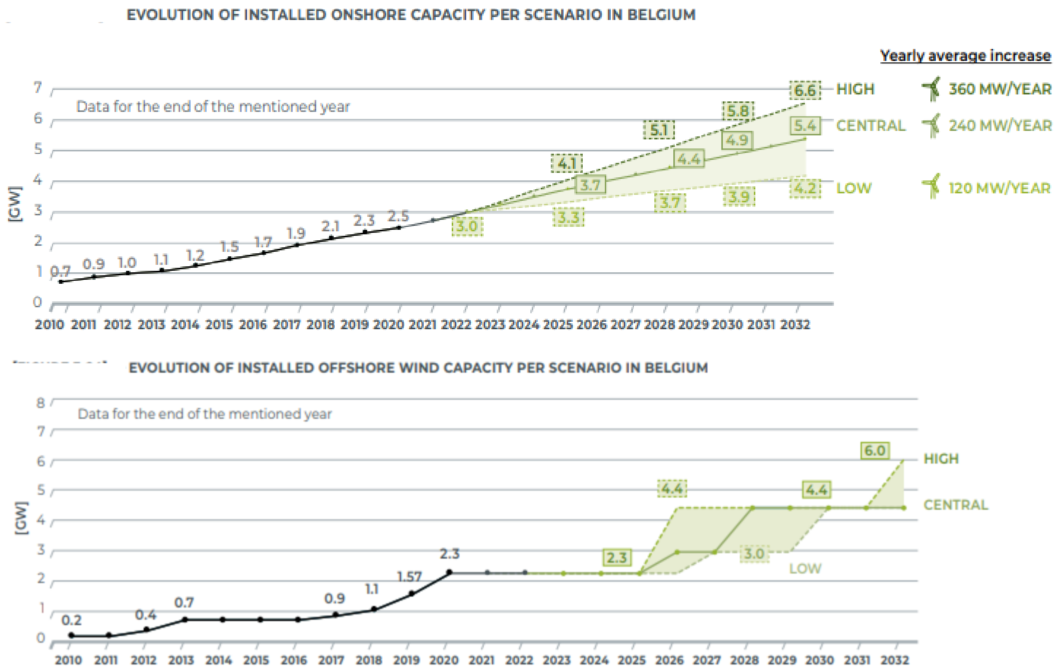


Figure 11: Evolution of solar and wind capacity per scenario in Belgium

5.2 Regional Operational Security Coordination (ROSC)

By mid-2024, the ROSC v1 will be live for the CORE region. A centralized and partly automatized security analysis will be realized notably for the costly remedial action optimization. Any error in the IGMs flow predictions on the monitored elements will directly impact the volumes and prices of redispatching. There will be less room for human assessment in case of an error in the IGMs files, the importance of a qualitative IGMs will consequently increase.

5.3 Update of information closer to real time

Updates on production information might become available closer to real time. BRPs are progressively authorized to nominate unbalanced positions to cope with the intrinsic growing uncertainty of the intermittent productions notably. Up to now, Elia did not detect an impact on congestion management and Elia does not predict a negative impact towards the near future. Elia should keep monitoring that the current practices (i.e. load scaling for DACFs and IDCFs) is a good approach.

5.4 Consumer centricity

The energy landscape is changing fast and in a deep manner, the consumer will be more active and aware about his consumption. Elia is already working with medium/large sized consumers in order to develop interactive tools that are able to forecast the individual consumption of the consumer. The idea is to have a platform that allows our customer to know what is their forecasted individual load and be able to improve this forecast for their use.

This is important because the behavior of the load will change soon, from a relatively inelastic load (relative to pricing) to a much more volatile load. A big chunk of this volatility will come from electric cars charging, heat pumps and batteries (charging, discharging, or holding) and the price will, likely, be a strong driver.
Elia must be agile in the way it forecast the load, more interactions with the customers means that the needs are better understood and anticipated. These profound changes will also occurs for the smaller customers so the knowledge of consumption dynamic profile will have to be even stronger.

5.5 Innovation projects and incentive

The innovation department of Elia is deeply involved in the improvement of the forecasting in general (not only for congestion management). The department is contributing to the task on several axes. The first is that the resources of the AI CoE (Artificial Intelligence Center of Excellence) are available to help the business to improve the forecasts. The data scientists of the center have a deep knowledge in machine learning and data handling, which is very valuable for the work done now regarding the reduction of the congestion costs through improving forecast. Moreover, the center is collaborating across business units inside the Elia Group (in Belgium and in Germany). This means that a broad range of experience and competence are put together increasing the general quality of the center.
The first concrete pilot of this AI CoE was the grid losses forecast, which was already developed in-house in Germany and then made in Belgium using the expertise from the German colleagues. The next project will be working together, across both Germany and Belgium with the AI CoE, on the forecast for Deterministic Frequency Deviations (DfDs).
Some other projects and incentives are ongoing (also linked with congestion maganement), for which it is important to align.

5.6 iCAROS

New regulation, a changing energy landscape and an evolution in operational needs call for a major evolution for the coordination of assets and congestion management. iCAROS project, “Integrated Coordination of Assets for Redispatching and Operational Security”, aims to redesign all process of operational data exchange between Market Parties and Elia for outage planning, scheduling and congestion management. All type B-C-D units (>1MW) connected either to TSO and DSO will have to provide these data. The implementation of iCAROS is distributed in three phases:

- Phase 1: implementation clarification of target design for system relevant assets ≥ 25 MW
- Phase 2: extension of implementation clarification of target design to all system relevant assets ≥ 1 MW (only availability plans for DSO-connected) & demand facilities (only TSO-connected)
- Phase 3: full extension of implementation clarification of target design to all system relevant assets ≥ 1 MW & demand facilities (only TSO-connected)

Simplified communication protocol and process for real time data of type B units will be set up. This will allows DSOs to share real time data easily without investing in expensive remote terminal units (RTUs). Consequently it can be expected that Elia will have far more real time data at disposal which is one of the enabler of better forecasts.

5.7 Load and generation project

There is a project in the starting blocks for aiming at better alignment of Power Factory (PF) models (offline studies) and EMS models (live data acquisition). The approach is shown in the graph below. Any steps in that direction may only help to increase the quality of the IGMs allowing a better comparison at nodal level of measured data from EMS and forecasted data from PF. It will also allow to feed forecasting models with real time data (estimation or measurements). The project will also foresee automatic monitoring and sanity checks on measured and estimated data improving the general quality of measured data as well as model data.

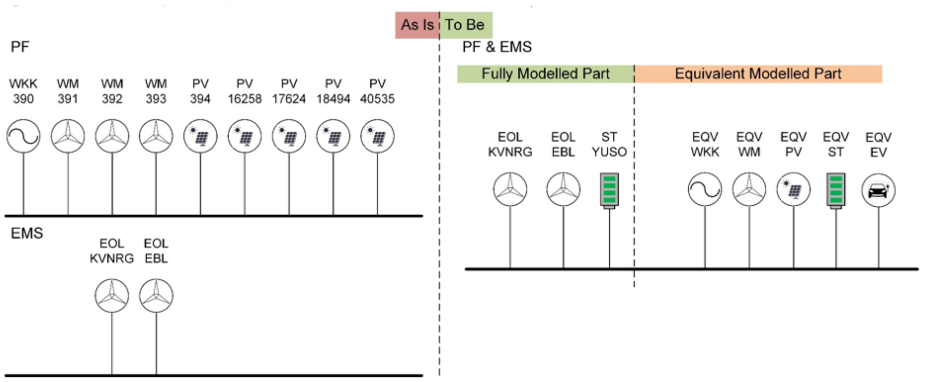


Figure 12: Evolution of the modelling of the real time tool for grid operation at Elia

5.8 Smart IGMs

The current approach in the IGMs is to forecast everything separately and then combine all data in the IGMs in order to have a forecast on the flows on all lines (needed for detection of congestions). Therefore, up to now all forecasts (Wind, PV, CHP, Nemo, load) are tackled separately. An idea for the future in the framework of congestion management would be to explore if dynamic and/or weighted combinations (90% percentile, average value, …) between all available forecasts input would lead to better IGM and so to a better detection of potential congestions.

To achieve that initiative, Elia is working in 3 phases, in collaboration with an external data scientist:

1. Understanding market drivers:

   Work up the various available drivers (price, renewable production, ….) to define periods with similar market conditions. The aim here is to better understand the influence of market drivers on the physical flows and by extension on the congestion. The results of this phase are clusters of data, sorted according to market conditions and underlying dynamics. Those clusters will lead to better modeling and more accurate predictions.
2. Optimization of periods and clusters

   The clusters from phase 1 depend on a few parameters: the timeframe of the reviewed period, the number of chosen clusters and some internal parameters needed for the calculation. The result of this second phase is an algorithm to optimize those parameters by using a simple prediction model.
3. Prediction

   The result of this third phase is a prediction algorithm that can be used to predict flows and so support congestion management.

6. Next steps

After the consultation period, Elia will collect all comments and feedback from the market parties. Elia will analyze these comments and integrate them into a consultation report together with Elia's responses. In addition to the consultation report, Elia will publish the reactions of the market parties (including names) on the website, unless it is explicitly stated that the contribution is to be considered confidential.

The comments will also be taken as much as possible into account in the final report that will be published by the end of the year. These comments can include any further questions on the aspects described in the report, additional needs or ideas for further analyses or any feedback on the identified possible solutions.

The final report will also cover a more elaborated roadmap towards the future including a short-term implementation plan indicating the impact based on some representative cases with congestion management costs and a long-term roadmap making the link with future challenges and projects (chapter 5). The implementation plan and roadmap will take into account the impact versus effort analysis made (chapter 4) and the evaluation of room for improvement (chapter 3).

Annex 1: Glossary

- **AI:** Artificial Intelligence
- **BRP:** Balance Responsible Party
- **CORE:** capacity calculation region in which Elia participates (16 TSOs)
- **DLR:** Dynamic Line Rating
- **EDW:** Electronic Data Warehouse (tool storing and distributing many operational data for ex-post analyses and reporting)
- **EFTool:** Elia Forecasting Tool, an internal tool acting as intermediary between different data sources and used for visualizing different forecasts
- **EMS:** Energy Management System, the tool used for real time grid operation at Elia.
- **FB:** Flow Based
- **FBMC:** Flow Based Market Coupling
- **iCAROS:** Integrated Coordination of Assets for Redispatching and Operational Security, the project that ….
- **IGM**: Individual Grid Model. This is an extended description of the grid state for a specific target hour for a given TSO. Such hourly files are created for 3 times horizons:
  - **D2CF**: Day +2 Congestion Forecast
  - **DACF**: Day-Ahead Congestion Forecast
  - **IDCF**: IntraDay Congestion Forecast
- **CGM**: Common Grid Model is the product of the combination of IGMs of many TSOs for a given hour.
- **L&G:** Load & Generation is an ongoing internal project aiming at aligning the real time model (EMS) with the planning model (Power Factory).
- **ML:** Machine Learning
- **PF:** Power Factory, commercial load flow tool used at Elia for operational planning and other grid calculations
- **PISA:** Elia database listing all known production units in Belgium (past, present and future units)
- **PST:** Phase Shifting Transformer
- **PTDF:** Power Transfer Distribution Factors
- **PV:** Photo Voltaic
- **QH:** Quarter hour
- **RMSE:** Root Mean Square Error
- **RPN**: Referential Production Netcalc, internal centralized netcalc production database
- **RSC:** Regional Security Coordinator
- **TIC (Mnemonic):** Traitement Intégré des Comptages, used to manage metering data