



State Estimation / Data Driven Load Flow

23-015 SNI

27 January 2023

Building 026
Koningstraat 27-1d
6811 DG Arnhem

Postbus 2013
2800 BD Gouda

026 352 37 00
www.phasetophase.nl



Copyright © Phase to Phase BV, Arnhem, the Netherlands. All rights reserved.

The contents of this report may only be transmitted to third parties in its entirety. Application of the copyright notice and disclaimer is compulsory.

Phase to Phase BV disclaims liability for any direct, indirect, consequential or incidental damages that may result from the use of the information or data, or from the inability to use the information or data.



CONTENTS

CONTENTS ii

- 1 Introduction..... 1
- 2 What is different?..... 1
- 3 The workflow..... 2
- 4 Normal Distribution for load modelling 2
 - 4.1 Bottom up Approach 2
- 5 Weighted Least Squares (WLS)..... 3
- 6 Implementation in Vision Network Analysis 3
 - 6.1 State Estimation dialog box 3
 - 6.2 Format for input Excel files..... 4
 - 6.2.1 Assigning data to element (MV/LV transformer and industry users)..... 4
 - 6.2.2 Assigning data to branch 5
 - 6.2.3 Assigning data from bijzonderheden tab 5
 - 6.2.4 Pre-processor 6
- 7 Example 7
 - 7.1 Network description 7
 - 7.2 Input..... 7
 - 7.3 Saving data and time frame selection options 8
 - 7.4 Viewing the results 9
- 8 Potential Applications 10
- 9 References..... 11

1 INTRODUCTION

Due to energy transition, the distribution grid is undergoing a rapid change. This is due to the introduction of more and more components in the grid especially in the distribution network side. It is estimated that by the year 2050, most of the electricity generation will be satisfied by renewable sources of energy such as solar panels and wind turbines. Further, fossil-fueled cars are to be gradually replaced by electric vehicles. These Distributed Energy Resources (DER) are paving way for the revolution in the energy industry. The inclusion of DER in the power system affects the voltage profiles, power quality, protection and in general the operation of the grid creating possible congestion issues.

This warrants a need for ensuring a way to keep the grid in check with constant monitoring. Monitoring can be performed either online in an operation center or offline using recent measurements over certain period of time. Offline network analysis allows to estimate all network variables and to examine how close the network is to potential congestion problems. This is where state estimation comes into picture. Predominantly used in transmission networks in the past, Distribution System State Estimation (DSSE) has been on the rise. One of the features when coming to the distribution network side is the fact that often no data is available at some points of the network. Even though measurement devices are available at certain locations, it is still possible that there can be data gaps making some timestamps unobserved.

For this very reason, a tool for State Estimation coupled with different methods to fill this missing data is introduced in Vision Network Analysis. The estimates from this calculation can be used for future planning or other short term grid applications depending on the requirement.

Guide overview

The report consists of a background and the method that is used for State Estimation/Data Driven Load Flow (DDLDF) that is implemented. The general assumptions taken are mentioned as well. The way to use the module is explained with a small network example towards the end.

2 WHAT IS DIFFERENT?

Before delving deep into this module, the difference between the already existing pseudo-monitor tool and also load flow with respect to state estimation needs to be discussed.

In the traditional load flow method, the powers of all loads and generators have to be defined at the start of the calculation, and no uncertainty is considered. This is perfect in an ideal scenario when everything is known but in reality, that is not the case. Further, in the load flow formulation no powers/currents through the branches can be specified explicitly. State estimation, on the contrary, is flexible in this respect since it allows to use different types of measurements. It takes into account the available measurement data, and with the help of pseudo-measurements the gaps are filled with some variance associated so that after computation the most probable estimate can be found out. In certain sense, this process can be seen as an extension of the load flow method, therefore state estimation is also referred in this report as Data Driven Load Flow.

3 THE WORKFLOW

The DDLF workflow implementation is shown in Figure 1 and is explained in more detail in Chapter 6. The measurements are broadly classified into two categories: real and pseudo measurements. The real measurements are those that are obtained directly from the meter data. Pseudo measurements are those which are generated in order to fill in the gaps of meter data to make the system observable. It can be either done using previously available data or using more advanced methods like taking reference of generalized profiles or using a bottom up approach from the LV side (employing Gaia LV Network Design in the future).

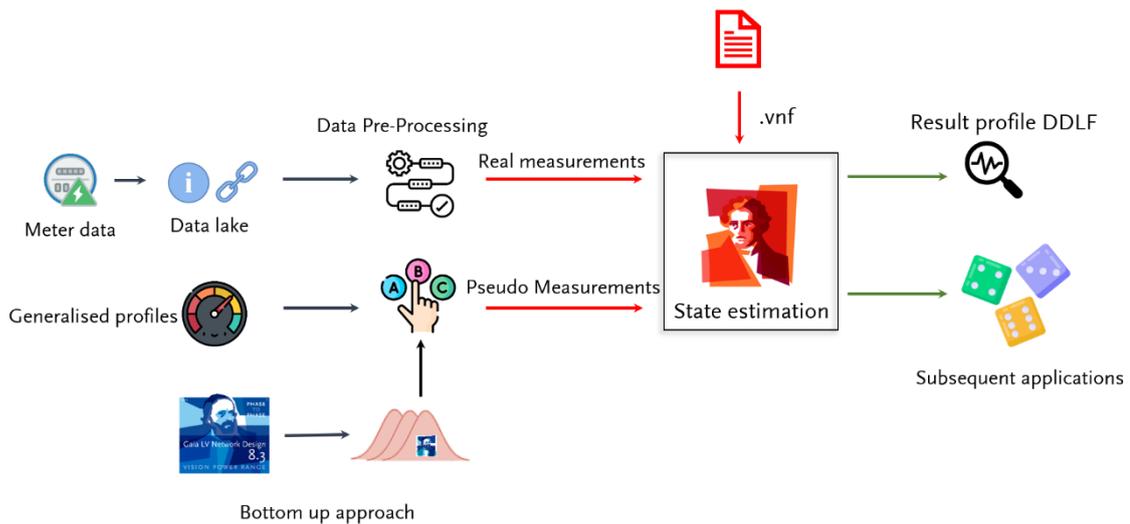


Figure 1: State Estimation/ DDLF workflow

With this measurement data in place, coupled with the .vnf file which contains the network data it is possible to perform state estimation or the data driven load flow. The profiles for all the variables needed like the voltages at all nodes, power flows and so on are obtained. This data can be used for further applications depending on the requirement and therefore opens up a world of possibilities.

4 NORMAL DISTRIBUTION FOR LOAD MODELLING

The considered inputs are assumed to follow a normal distribution pattern and a mean with standard deviation is used as data/measurement input for the module. For real measurements, the variance is considered to be low but for pseudo-measurements it is larger.

An advanced load modelling version in Gaia which takes into consideration different type of loads can be a possibility to generate the profiles and accordingly calculate the mean as well as variance for the pseudo-measurements to get a better estimate on the whole. This is the bottom up approach described briefly below.

4.1 Bottom up Approach

If the LV networks are available corresponding to the transformer loads in the MV network, it is now possible to do a bottom up calculation and compute the mean as well as standard deviations of the active

and reactive power injections at the LV side of the transformer by adding up all the powers from individual apparatus as well as the Gaussian mixture modelled loads.

The bottom up approach employs the strategy of using Monte Carlo sampling on the Gaussian mixture distribution of different types of loads in a household ranging from 200 to 1000 samples in every time instance. The samples are added up for each household first to get the equivalent load for a particular time instance. After this, the total power is summed up for all the households present in the LV network which gives a large set of possible values (depending on the sample number taken) for that time instance.

The mean and the standard deviation is then calculated for this dataset which gives the probable range in which the equivalent load value lies for the LV network on the LV side of the MV/LV transformer.

5 WEIGHTED LEAST SQUARES (WLS)

WLS is the conventional method used for state estimation and its modelling is based on [1]. A brief overview of the working of the algorithm for the distribution network scenario is explained in it. Further statistical tests are performed in [2] to give a relevant justification for the algorithm's selection. This is the method employed in the module designed in Vision Network Analysis and few of the assumptions are:

- **State Variables:** The bus voltages and angles are considered as the state of the system which has to be calculated.
- **Measurements:** The active and reactive power injections as well as the real and reactive power flows along with bus voltage magnitudes and magnitude of current flows can be considered.
- **Bad Data:** It is assumed that no bad data has to be processed considering the large variances present due to the pseudo-measurements.

6 IMPLEMENTATION IN VISION NETWORK ANALYSIS

The State Estimation module is meant to replace the conventional Pseudo-monitor module which is already available in Vision Network Analysis. The designed module is an add-on along with the already existing options in the dropdown of the Pseudo monitor option for now. To use this option, select the **Data Driven Load Flow** option from the **Analyses** section in the **Calculation** tab on the Vision ribbon.

When clicked, a dialog box with further options to customize the inputs as well as the desired way to process data and the outcome appears.

6.1 State Estimation dialog box

The dialog box that appears is as shown in Figure 2. The inputs to be specified are the measurement files (in a particular format which is explained later on) and if the pre-processor option is to be used based on the user's data quality.

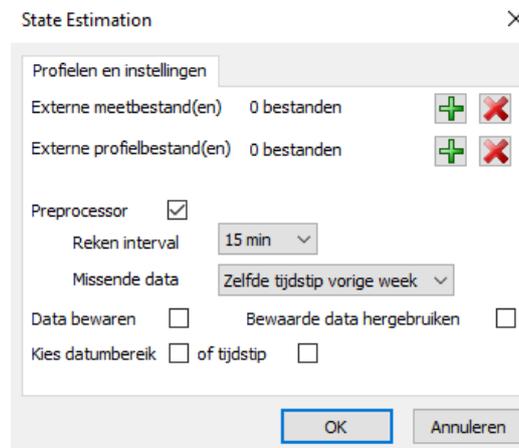


Figure 2: State Estimation Dialog box

In order to run state estimation, it is necessary that there is at least one input of the measurement files corresponding to the elements in the network. By default, the preprocessor feature is enabled. This comes with two further settings that can be customized by the user. The sampling time interval for the output data and the choice of how to fill in the missing gaps. In case the user is confident with the quality of data, it is possible to disable the preprocessor and start the calculation.

To further elaborate on the options for data pre-processing, the calculation interval refers to how frequently a state estimation calculation is performed for the selected time-period. The missing data option refers to the way a gap is filled in with reference to previously available timestamps data from the input measurement file.

6.2 Format for input Excel files

The data can be provided as an input in the form of Excel sheets at this moment of time. Considering the bigger picture and practicality that might involve quite a large number of networks, other solutions will be considered for future implementations of the module.

Keeping this aside for now, the inputs that can be handled by the module are active/reactive power and current injections at the nodes; active/reactive power flows and currents in branches of the network.

6.2.1 ASSIGNING DATA TO ELEMENT (MV/LV TRANSFORMER AND INDUSTRY USERS)

A particular format needs to be followed for the Excel sheet to recognize and differentiate between the inputs. To connect a profile/measurement with the element in Vision, the format of **NodeName.ElementName** needs to be used as shown in Figure 3. Depending on the type of input, the corresponding unit is also entered. For example **W/kW/MW** for active power, **var/kvar/Mvar** for reactive power, **A/kA** for currents and **V/kV** for voltages. Additionally, the standard deviations can also be specified through the input, for example **W_std/kW_std/MW_std** for active power and **var_std/kvar_std/Mvar_std** for reactive powers. This helps in assigning data to the corresponding element. It is advisable to input all parameters in the same class of units. The same process needs to be followed to assign data to industry users too.

Note: Active and reactive power injections at nodes are the most frequently used measurements.

	A	B	C	D	E
1	datum & tijd	NodeName.ElementName	NodeName.ElementName	NodeName.ElementName	NodeName.ElementName
2		MW	Mvar	MW_std	Mvar_std
3	07/02/2022	0,18493071	0,03836006	0,01383925	0,00281018

Figure 3: Data input format for elements

6.2.2 ASSIGNING DATA TO BRANCH

In order to assign a profile to a branch, the name of the measuring field **Name** needs to be specified instead of the previous format used as shown in Figure 4. The rest of the procedure is similar as to the previous case of assigning to an element but here it is for the branch.

Note: Here, the measuring field represented by  helps in identifying the branch and cross-referencing it. It is also side specific which means it indicates the side (direction) of the branch where the measurement is taken.

	A	B	C
1	datum & tijd	Name	Name
2		kW	kvar
3	01/01/2021	108	48

Figure 4: Input format for branch element

From here, the data is then further “pre-processed” to double check the input and ensure consistency depending on if the option to do so is enabled.

Note: The units of variables should be uniform in the excel files throughout the measurement and profile files to avoid computational errors.

6.2.3 ASSIGNING DATA FROM BIJZONDERHEDEN TAB

To input the standard deviations of the pseudo-measurements it is possible also via the Bijzonderheden tab in the branches or the elements as shown in Figure 5. This helps to internally replace the variances value to the equivalent arrived from the value of the standard deviation in this field. For Vision to be able to read these, it is important that the Kenmerk field is either set to **P_std** or **Q_std** depending on what the user wants to input and the standard deviations values have to be specified in **MW** or **MVar**.

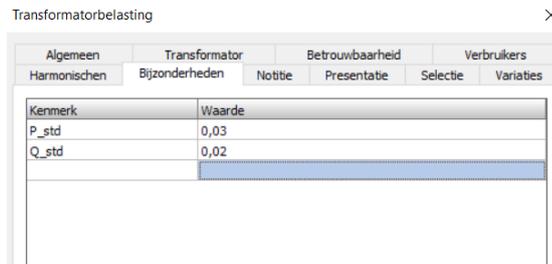


Figure 5: Input via bijzonderheden tab

6.2.4 PRE-PROCESSOR

The time interval between computed estimates can also be adjusted according to requirements. The default setting is 15 min. Also, the period for calculation as well as if needed a particular timestamp can additionally be selected to be estimated.

An option to skip this step is also included if the user is confident with the data at hand and also reduces computation time. But when this checkbox is disabled, it is to be noted that some advanced features can no longer be used. The data is exactly taken as it is given in the input. In this scenario, the only value currently being automatically filled-in is the variance in case the values for it are absent. In the future a user input for standard deviations via Excel sheet will be also supported.

The default setting for interval selection is set as 15 minutes, and it can be changed to other intervals if needed (5min, 10min, 15min, 30min, 1hour, 12 hours, 1 day and 1 month). It is important that when such an adjustment is made, the frequency of measurements should correspond to initial multiples of input frequency taken so as to prevent eventual skipping of values.

The way to fill up the missing gaps includes into five possible methods:

1. **Taking the previous timestamp value:** Using this method is justifiable when the data has very few missing gaps and therefore is reasonable to use the last known measured value.
2. **The same time-stamp value from previous day:** In the event of having the data of a whole day or part of the day missing, it is not advisable anymore to use the first method considering a large chunk of values are filled in with the same value which does not give a good picture of the reality.
3. **The same time-stamp value from previous week:** Similarly if there are large gaps over the week, it is possible to see the trends from the previous week and leverage those to get a good estimate.
4. **Mean value (10 timestamps):** Using the mean value from the previously available data is another approach. The last available timestamps refer to the same timestamp values in the previous days. A large range of one month is considered for this purpose taking into account that all the days might not have available measurement at the required timestamp. In that case, the pre-processor searches for ten available values of the timestamp required in thirty previous days going day by day and determines an average of them. This helps in taking into account a wider set of behavioral range and take the averaged value as the reference. For the first 30 days, method 3 is followed by default since there is no enough historical data available from prior time instances.
5. **Using a reference profile:** This is a unique way compared to the other methods where the measurement file is further supported by a reference profile file which helps to shed more light on what an approximated value can be for a particular instance (determined as an input in another Excel file). For the timestamp where a measurement is unavailable, the value from the reference profile file is used to bridge the gap and continue with the computation. It is important to note that the interval of measurements in the reference profile file should be the same as the selected one in the preprocessor options.

Note: It is important that if the methods 2 and 3 are selected, in case of having missing gaps in the first day or the first week there is not enough data from previous timestamps. Hence, the most suitable method of the three (1, 2 or 3) is chosen. After this first hurdle's crossed, the initially selected method is continued for the rest of the dataset.

In case no standard deviation is provided either from the profiles (for any kind of measurement) or the Bijzonderheden tab(for pseudo-measurements), the selection of variances is done on the basis of type of measurement as a default value explained in detail below.

- For real measurements, the value is assumed to be reliable as it is taken from the measuring device which implies that the uncertainty is low [3].
- For the pseudo-measurements, a larger variance is considered which is 100 times larger than the one considered for real measurements. This is because an extra degree of freedom is provided so that with state estimation, the variable converges to the more accurate value.

Note: The computation time is higher when pre-processor is on due to all of these additional factors considered.

7 EXAMPLE

The usage of the module is explained here with the help of a simple network, and the obtained results are explained.

7.1 Network description

A network with eleven nodes and ten transformer loads is taken for the purpose of demonstration as shown in Figure 6. The nominal voltage level of the network is 10 kV. The transformers are rated 10/0.4 kV.

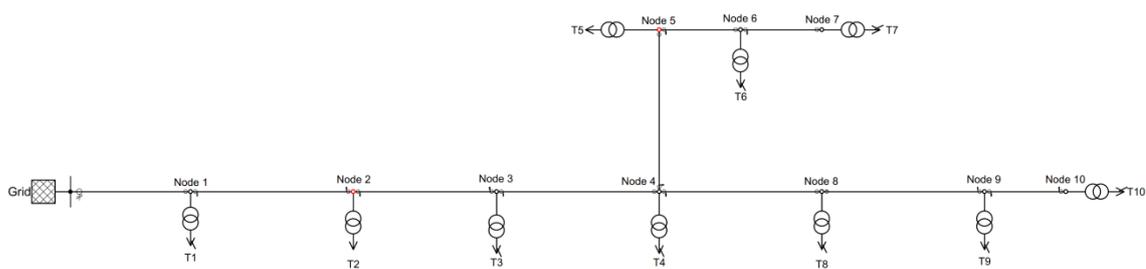


Figure 6: Example of a part of MV Network

7.2 Input

To perform state estimation on the network, use **DDLf** (Data driven load flow) of the Vision ribbon. The dialog box shown in Figure 7 appears. Depending on the status of the preprocessor check box, the other two options for time interval and filling missing data are available as seen in Figure 7.

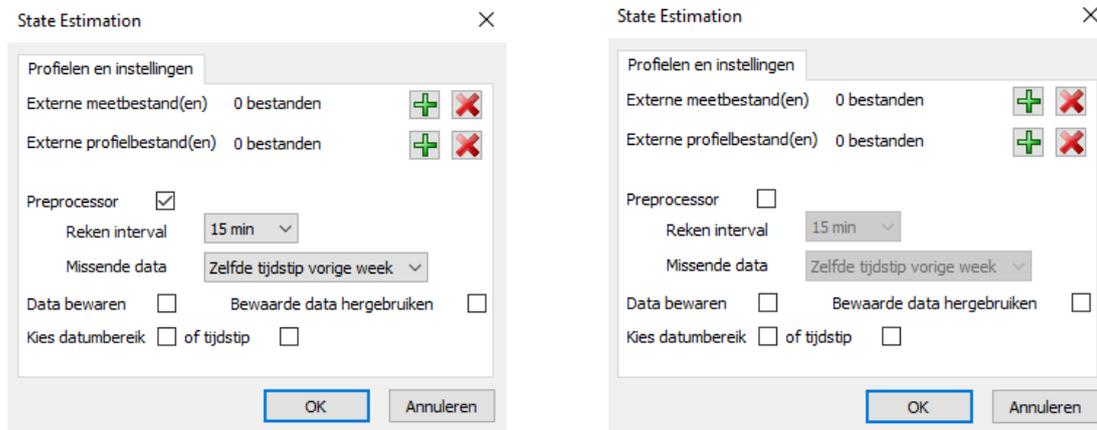


Figure 7: Input with pre-processor on/off

The input can be fed by clicking on the add file button (with green plus icon) beside external measurement file and then a dialog box as shown in Figure 8 opens up. Files can be added by going to the required directory. Multiple files can also be selected with the help of Ctrl and Shift buttons. Otherwise it is also possible to add the files individually by performing the action repeatedly.

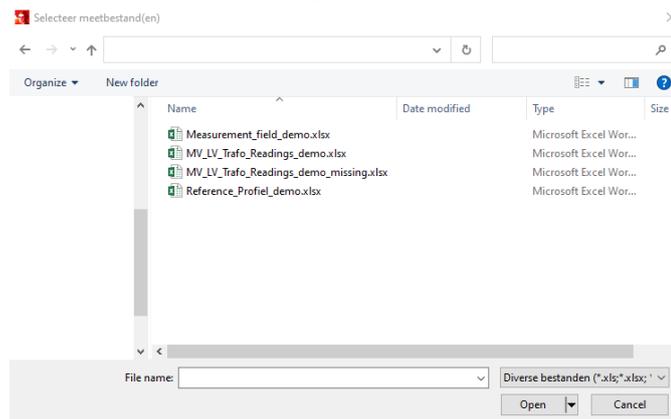


Figure 8: Excel input dialog box

Further, in a similar way profile files (pseudo-measurements) can be added with the other file addition button if in the preprocessor Reference profile (the missing data is filled up using this file as the basis) option is selected. In the other scenarios, even if a profile file is added along with the measurement file it is neglected.

If the preprocessor option is on, further adjustments can be made on the time interval and filling the missing data as explained in Section 6.2.4.

7.3 Saving data and time frame selection options

The other options are to save and reuse data in case multiple computations are needed for the same network. This can be done by checking the box Save Data during the first run. After the files are read, data is not erased after state estimation computation is completed. From the second run, also checking the box which says Re-use data helps to spare significant amount of time by not reading in the datafiles again each and every time.

Further, a particular timestamp or a particular period in which the user needs state estimation to be performed on can also be selected as shown in Figure 9.



Figure 9: Selecting time-frame/time-stamp for computation

7.4 Viewing the results

After computation, the results can be viewed for any object by right clicking it (a dialog box as shown in Figure 10 opens up) and using the Details button to display numerical results and the Graph button for making a plot.

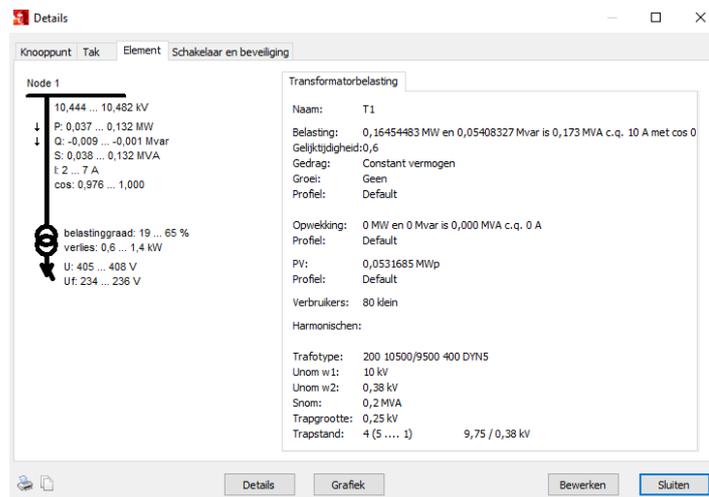


Figure 10: Results in detail form pop-up

As seen, the values shown are the range between maximum and minimum in which the variable varies in a computed time period. It is to be noted that these maximum or minimum values of different variables do not necessarily occur at the same instance (i.e. $P_{max} = 0,132$ MW and $Q_{max} = -0,001$ Mvar shown in the example above might happen at different times).

Data can be viewed by going into the details and then can be exported or perform other operations for further analysis using the toolbar shown in Figure 11. The data can be sorted with the filter option where only particular rows or columns can be selected to view. The data can be exported to an Excel datasheet with the Excel icon in the toolbar. Few specific rows or columns and also individual cells can be selected and copied to the clipboard. Also rows and columns can be deleted if needed by using the last two icons in the toolbar.

Details state estimation

Resultaat van transformatorbelasting T1 aan Node 1

tijd	P		Q		S	I	cos	belastinggraad	U	
	MW	Mvar	MVA	A					%	kV
07/02/2022 00:00	0,049	-0,008	0,050	3	0,989	25	10,476	235		
07/02/2022 00:15	0,043	-0,009	0,043	2	0,980	21	10,477	236		
07/02/2022 00:30	0,049	-0,008	0,049	3	0,986	24	10,478	236		
07/02/2022 00:45	0,051	-0,008	0,052	3	0,988	26	10,478	236		
07/02/2022 01:00	0,045	-0,007	0,046	3	0,988	23	10,479	236		
07/02/2022 01:15	0,045	-0,007	0,045	2	0,987	22	10,479	236		
07/02/2022 01:30	0,043	-0,008	0,044	2	0,983	22	10,480	236		
07/02/2022 01:45	0,040	-0,008	0,041	2	0,981	20	10,480	236		
07/02/2022 02:00	0,042	-0,008	0,042	2	0,983	21	10,480	236		
07/02/2022 02:15	0,041	-0,008	0,042	2	0,981	21	10,481	236		
07/02/2022 02:30	0,040	-0,008	0,041	2	0,979	20	10,480	236		
07/02/2022 02:45	0,047	-0,007	0,047	3	0,987	23	10,481	236		
07/02/2022 03:00	0,045	-0,007	0,045	3	0,987	23	10,481	236		

Figure 11: Results overview (Details)

The different types of variables like loading percentage, active power, reactive power, voltage, current and so on that need to be plotted can be selected in the graph view with the dropdown box at the bottom as shown in Figure 12 (for transformer element T1). The graph can also be exported in multiple formats (metafile and bitmap) using the right mouse button context menu. The plotting of individual points can be eliminated to display only the continuous line by unchecking the Points option. Even multiple objects can be selected at the same time and plotted.

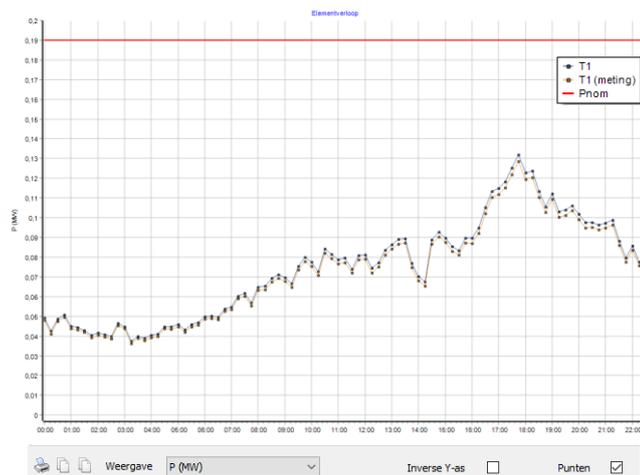


Figure 12: Results overview (Graphical)

The other option to view the graphical results is through the measuring field element. By right clicking the measuring field element connected to an object in Vision Network Analysis (for example the branch, transformer, node or load), then going to the Measurement Files option and clicking on the Graph button.

8 POTENTIAL APPLICATIONS

Few of the potential applications are as follows:

- Calibrated Profiles and Short-Term Network Planning:** The estimated profiles of distribution transformers can be obtained for the network locations where no measurements are currently available. These calibrated profiles can be further used for short-term planning studies involving addition of new components and reconfigurations of the network.

- **Congestion Analysis using Recent Data:** Network states over recent period of time (for example, the last year) can be estimated, and an analysis of how close the network was to a potential congestion at different time instances can be performed. Since state estimation module determines all variables in the network, all cables and transformers (even those not measured) can be evaluated for overloading or for reaching certain loading limit (for example, a cable was loaded more than 80% twice a year for one hour).
- **Optimal Meter Placement:** Using state estimation, meter placement studies can be performed to help narrow down the most economic and accurate scenario to be used.

9 REFERENCES

- [1] A. Abur and A. G. Expósito, Power system state estimation: theory and implementation, New York: NY: Marcel Dekker, 327pp, ISBN: 978-0-8247-5570-6, 2004.
- [2] S. S. Nibhanupudi, "State Estimation in Medium Voltage Distribution Networks," MSc Thesis, TU Delft, 2021.
- [3] N. Blaauwbroek, D. Kuiken, P. Nguyen, H. Vedder, M. Roggenkamp and H. Slootweg, Distribution network monitoring: Interaction between EU legal conditions and state estimation accuracy, Energy Policy, vol. 115, pp. 78–87, April 2018.