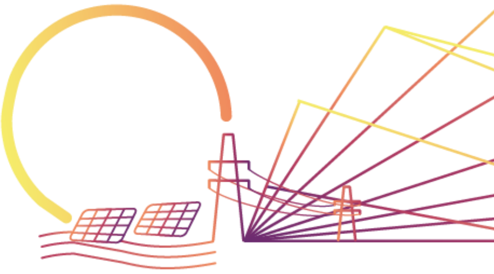
This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement Nº 953016

D2.1 Definition of needs from industry, evaluation of the existing simulation tools and models available on the market

T2.1 Definition of needs from industry, evaluation of the existing simulation tools and models available on the market

Grant Agreement nº:	**953016**
Call:	H2020-LC-SC3-2020-RES-IA-CSA / LC-SC3-RES-33-2020
Project title:	**Smooth, REliable aNd Dispatchable Integration of PV in EU Grids**
Project acronym:	SERENDI-PV
Type of Action:	Innovation Action
Granted by:	Innovation and Networks Executive Agency (INEA)
Project coordinator:	Fundación TECNALIA Research & Innovation
Project website address:	https://serendipv.eu/
Start date of the project:	October 2020
Duration:	48 months
Document Ref.:	SERENDI-PV_D2.1 Needs of Industry and Tools and models evaluation_with track changes.docx
Lead Beneficiary:	CYTHELIA ENERGY
Doc. Dissemination Level:	**P** – Public
Due Date for Deliverable:	31/03/2021 (M6)
Actual Submission date:	12/08/2021 (M11)
Version	V1



Summary

This deliverable is an output of task T2.1, which is divided in two subtasks:

- Task T2.1.1 Needs of the industry
- Task T2.1.2 Evaluation of the available simulation tools and models

The present report is therefore organised in 2 related sections.

To assess the needs of the industry concerning PV systems simulations, a survey was elaborated with the participation of all the partners involved in the task and distributed to the partners' contact and through social media. The **first section** presents in detail the survey and its outcomes.

The **second section** presents the methodology adopted for the evaluation of simulation tools and models, followed by the presentation and the discussion of the results.

Document Information

Title	Definition of needs from industry, evaluation of the existing simulation tools and models available on the market
Lead Beneficiary	Cythelia Energy
Contributors	All partners
Distribution	PU - Public
Report Name	SERENDI-PV_D2.1 Needs of Industry and Tools and models evaluation_with track changes.docx

Document History

Date	Version	Prepared by	Organisation	Approved by	Notes
12/08/2021	V1	Ismaël Lokhat	Cythelia Energy	TECNALIA	Submitted to the EC

Acknowledgements

The work described in this publication has received funding from the European Union’s Horizon 2020 research and innovation programme under grant agreement Nº 953016.

Disclaimer

Contents

Tables

Figures

1 EXECUTIVE SUMMARY

1.1 Description of the deliverable content and purpose

This document contains the results obtained from the work performed in task T2.1.

It is divided in two parts, each one addressing one of the two subtasks of task T2.1:

1. Presentation of the survey conducted and its results
2. Presentation of the evaluation of tools and models

The objective of the work performed in this task was to have a better understanding of the industry needs in concerning the simulation of PV systems, and to compare a selection of the existing tools and models available on the market.

1.2 Reference material

This document has taken information from the following documents previously produced by SERENDI-PV:

- D10.1 Project Management Plan, for the detailed description of the WP2 tasks
- D11.1 H - Requirement No. 1 – Humans, for the requirements to follow for doing the survey
- D11.2 Protection of Personal Data (POPD) - Requirement No. 2, for the requirements to follow for doing the survey

1.3 Relation with other activities in the project

Table 1.1 depicts the main links of this deliverable to other activities (work packages, tasks, deliverables, etc.) within SERENDI-PV project. The table should be considered along with the current document for further understanding of the deliverable contents and purpose.

Table 1.1: Relation between current deliverable and other activities in the project

Project activity	Relation with current deliverable
WP2 T2.1	The outcomes of this deliverable are based on the partners' activity in Task 2.1 (led by CYT) from M1 to M6.
WP2 T2.2-T2.6	The outcomes of this deliverable serve as "reference" for identifying and quantifying the industry needs and limitations (of the state of the art), which will be addressed through the innovations on PV simulations in WP2.
WP1- WP8	Similarly, the "reference" nature of D2.1 can provide useful insights into industry needs and KPIs in relation PV simulations (e.g. Energy Performance Index, EPI). These aspects are jointly addressed (from a different scope) in WP1 and WP8.

1.4 Abbreviation list

Table 1.2: Abbreviation list

Abbreviation	Meaning
BIPV	Building Integrated Photovoltaic
DNI	Direct Normal Irradiation / Irradiance
GHI	Global Horizontal Irradiation / Irradiance
GTI	Global Tilted Irradiation / Irradiance
LeTID	Light and elevated Temperature Induced Degradation
LID	Light Induced Degradation
MBE	Mean Bias Error
MBWE	Mean Bias Weighted Error
MODIS	MODerate-resolution Imaging Spectroradiometer
NMBE	Normalised Mean Bias Error
NMBWE	Normalised Mean Bias Weighted Error
NRMSE	Normalised Root Mean Square Error
NRMSWE	Normalised Root Mean Squared Weighted Error
PV	Photovoltaic
QC	Quality Control
RMSE	Root Mean Square Error
RMSWE	Root Mean Squared Weighted Error
SAM	Solar Advisor Model
SDAT	Solargis Data Analyst
WP	Work Package

2 Needs from industry

The purpose of the task T2.1.1 was to assess the needs of the PV industry through a survey.

This section presents both the survey and its results.

2.1 Survey

2.1.1 Presentation

The survey was elaborated thanks to the participation of the partners involved in the task.

Google form was used to allow an easy page layout and online release.

After an introduction of the survey and its purpose, the survey was organized in 7 sections:

1. Information concerning the profile of the respondent (Figure 2.1). To comply with the General Data Protection Regulation, it has been decided that the survey would be anonymous, as it was described in *D11.1* and *D11.2*.
2. Meteorological data (Figure 2.2). The aim of this short section was to identify the type of data used.
3. The software usage of the respondent (Figure 2.3). This section allowed the respondent to indicate up to 3 software in a list of 35.

   The list of proposed software is given in the table below, ranged in alphabetical order.

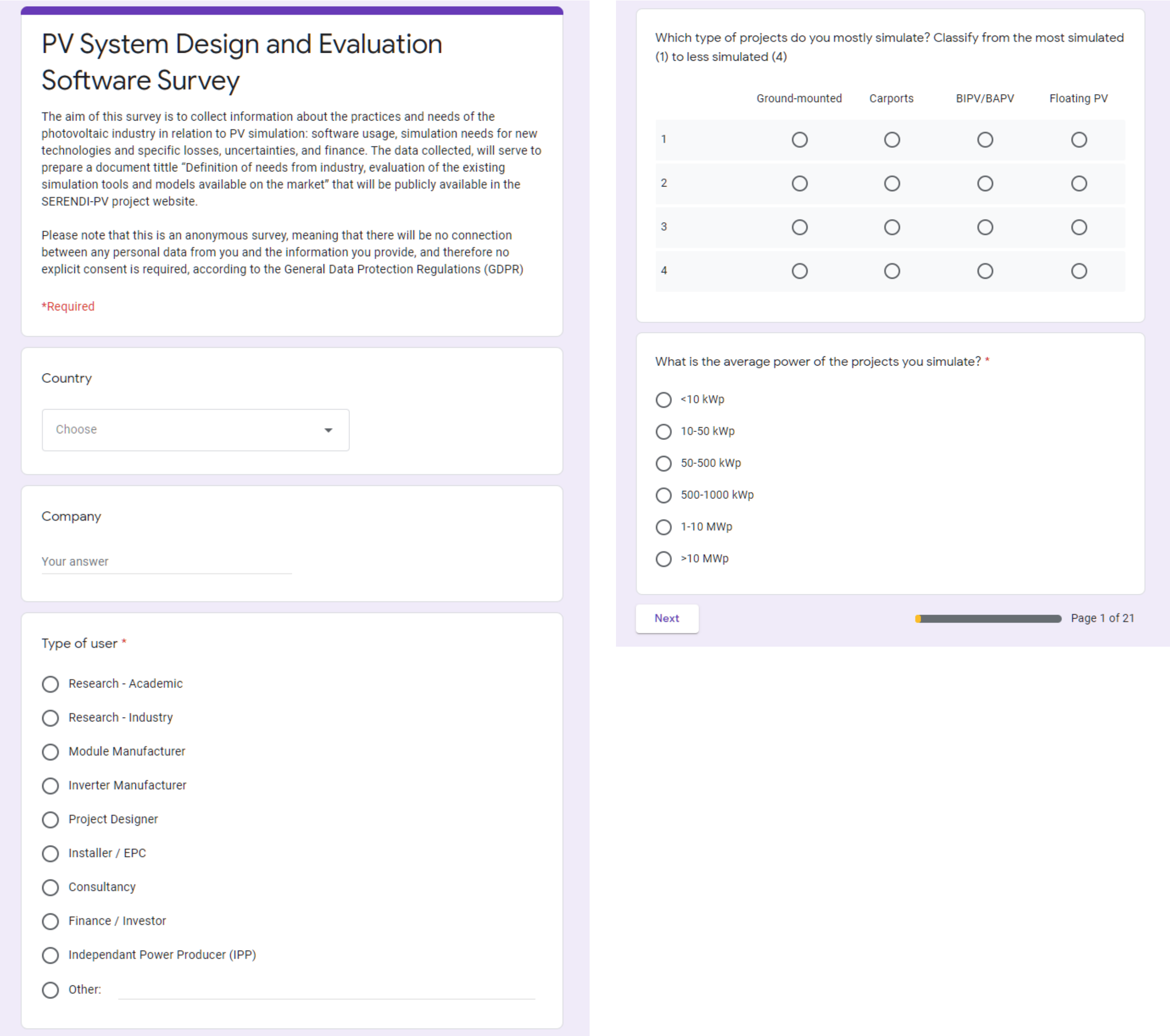
Table 2.1: List of proposed software tools

Archelios PRO	Lusim	PVGIS	SAM (NREL)
BIMSOLAR	PC1D	PV Scout	SMARTS
CASSYS	Plan4Solar PV	PVSites	Solar PRO
CECPV Calculator	Plant Predict	PVSYST	Solargis
Easysolar APP	Polysun	PVWatts	SolarGo
Global Solar Atlas	PVcase	PV*SOL	Solarius PV
Helios 3D	PV Complete	Quokka	Solar Mapper
Homer PRO	pvDesign (Ratedpower)	RETScreen Expert	Zenit
Insel	PV Designer (Solmetric)	Skelion	

   The respondent had the possibility to indicate the name of a software not listed.

   The respondent was then asked to tell which features he would like to have in the software he’s using.

4. Losses evaluation section (Figure 2.4 to Figure 2.6), which relates to the work that will be conducted in the task 2.2 of SERENDI-PV. This section was organised in 3 sub-sections: degradation losses, soiling losses and snow losses. A few general questions preceded these subsections in order to grasp the importance of these losses for the respondents
5. New technologies section (Figure 2.7 to Figure 2.9), which relates to the task 2.3. This section was organised in 3 sub-sections: bifacial PV, floating PV and BIPV
6. Uncertainties section (Figure 2.10) which relates to the task 2.4.
7. Finance section (Figure 2.11), which relates to the task 2.5. Only the Project designers, EPCs, IPPs and investors are invited to answer the questions of this section.



PV System Design and Evaluation Software Survey

The aim of this survey is to collect information about the practices and needs of the photovoltaic industry in relation to PV simulation: software usage, simulation needs for new technologies and specific losses, uncertainties, and finance. The data collected, will serve to prepare a document tittle "Definition of needs from industry, evaluation of the existing simulation tools and models available on the market" that will be publicly available in the SERENDI-PV project website.

Please note that this is an anonymous survey, meaning that there will be no connection between any personal data from you and the information you provide, and therefore no explicit consent is required, according to the General Data Protection Regulations (GDPR)

*Required

Country

Choose

Company

Your answer

Type of user *

○ Research - Academic
○ Research - Industry
○ Module Manufacturer
○ Inverter Manufacturer
○ Project Designer
○ Installer / EPC
○ Consultancy
○ Finance / Investor
○ Independant Power Producer (IPP)
○ Other:

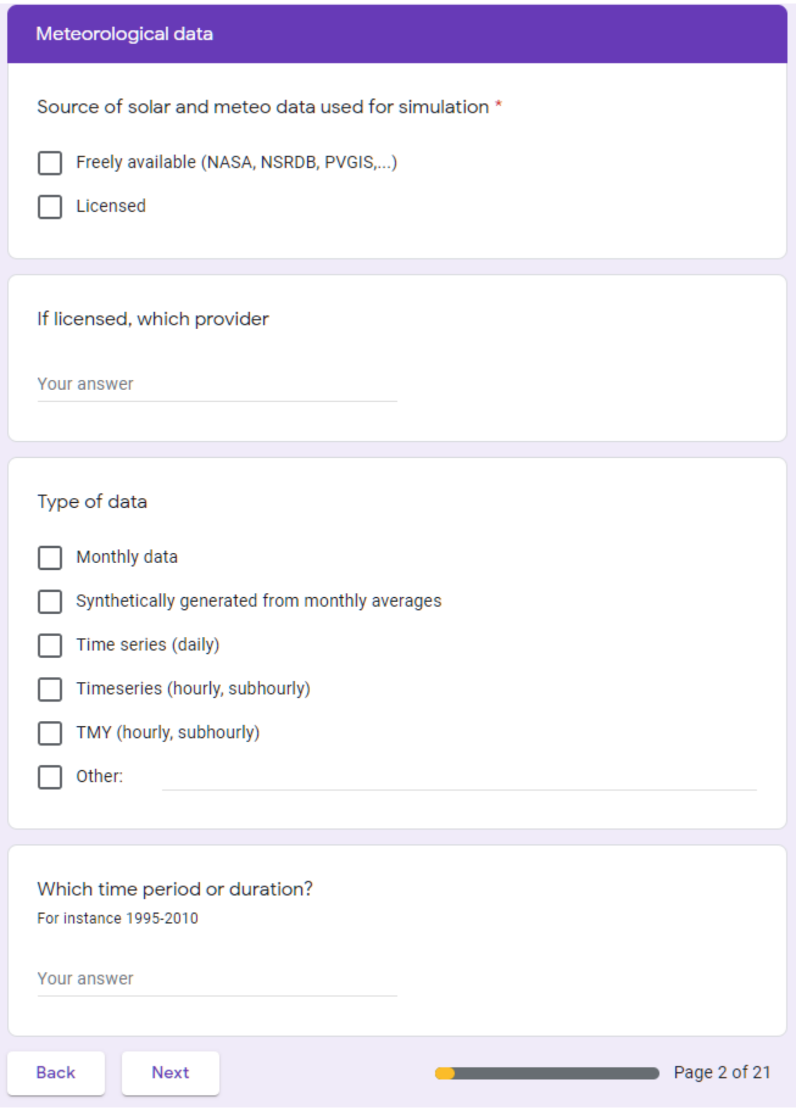
Which type of projects do you mostly simulate? Classify from the most simulated (1) to less simulated (4)

	Ground-mounted	Carports	BIPV/BAPV	Floating PV
1	○	○	○	○
2	○	○	○	○
3	○	○	○	○
4	○	○	○	○

What is the average power of the projects you simulate? *

○ <10 kWp
○ 10-50 kWp
○ 50-500 kWp
○ 500-1000 kWp
○ 1-10 MWp
○ >10 MWp

Figure 2.1: Survey description – section 1: Information concerning the profile of the respondent



Meteorological data

Source of solar and meteo data used for simulation *

☐ Freely available (NASA, NSRDB, PVGIS,...)
☐ Licensed

If licensed, which provider

Your answer

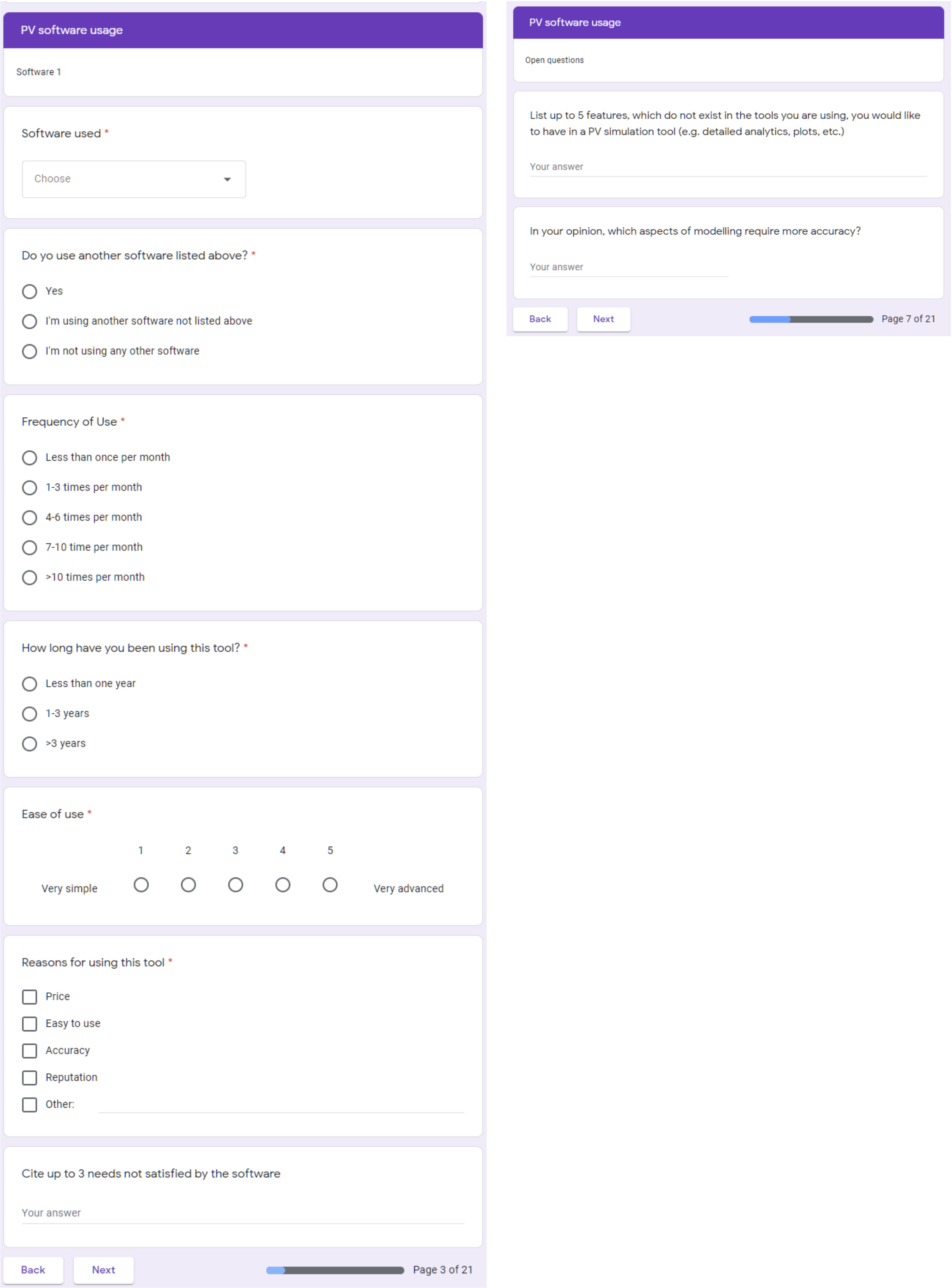
Type of data

☐ Monthly data
☐ Synthetically generated from monthly averages
☐ Time series (daily)
☐ Timeseries (hourly, subhourly)
☐ TMY (hourly, subhourly)
☐ Other:

Which time period or duration?
For instance 1995-2010

Your answer

Figure 2.2: Survey description – section 2: Meteorological data



PV software usage

Software 1

Software used *

Choose ▾

Do yo use another software listed above? *

○ Yes
○ I'm using another software not listed above
○ I'm not using any other software

Frequency of Use *

○ Less than once per month
○ 1-3 times per month
○ 4-6 times per month
○ 7-10 time per month
○ >10 times per month

How long have you been using this tool? *

○ Less than one year
○ 1-3 years
○ >3 years

Ease of use *

	1	2	3	4	5	
Very simple	○	○	○	○	○	Very advanced

Reasons for using this tool *

☐ Price
☐ Easy to use
☐ Accuracy
☐ Reputation
☐ Other: ____

Cite up to 3 needs not satisfied by the software

Your answer

Back Next Page 3 of 21

PV software usage

Open questions

List up to 5 features, which do not exist in the tools you are using, you would like to have in a PV simulation tool (e.g. detailed analytics, plots, etc.)

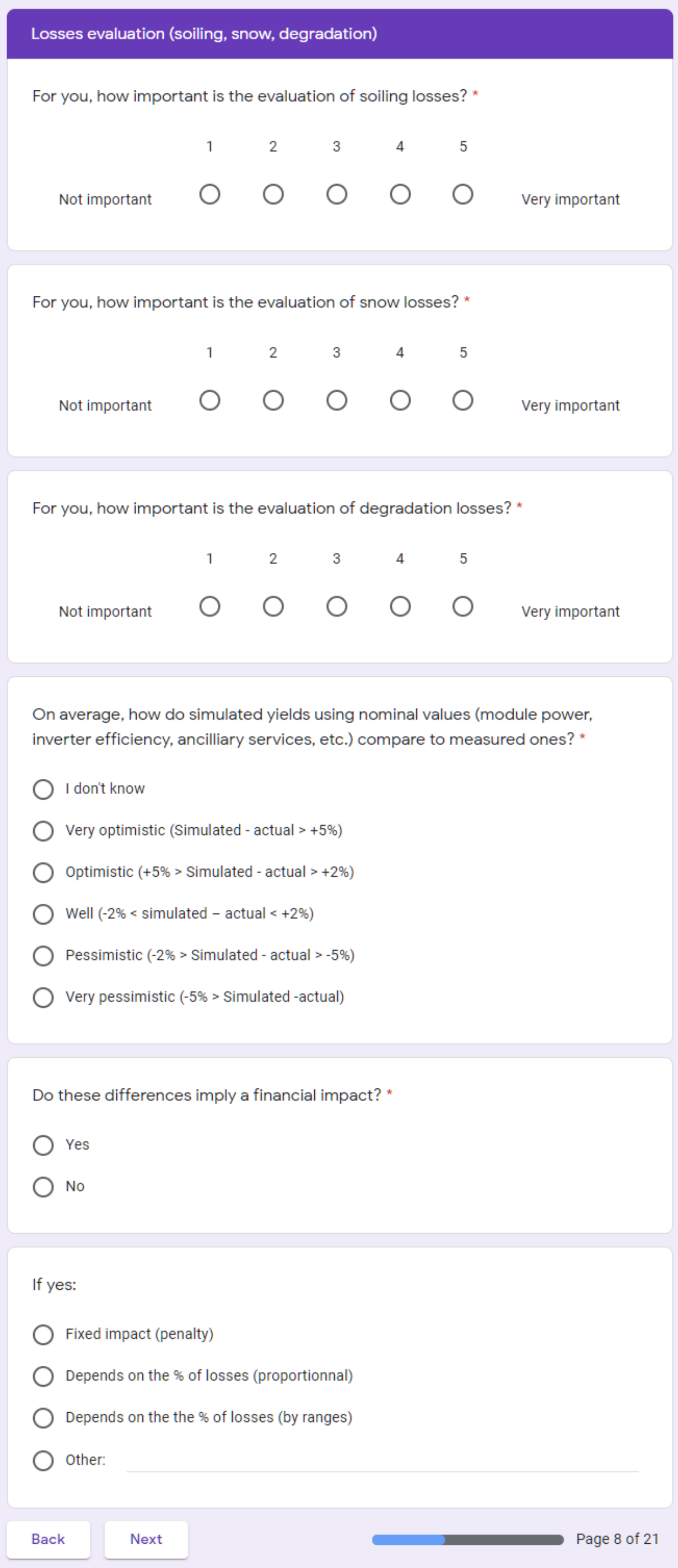
Your answer

In your opinion, which aspects of modelling require more accuracy?

Your answer

Back Next Page 7 of 21

Figure 2.3: Survey description – section 3: The software usage of the respondent



Losses evaluation (soiling, snow, degradation)

For you, how important is the evaluation of soiling losses? *

	1	2	3	4	5	
Not important	○	○	○	○	○	Very important

For you, how important is the evaluation of snow losses? *

	1	2	3	4	5	
Not important	○	○	○	○	○	Very important

For you, how important is the evaluation of degradation losses? *

	1	2	3	4	5	
Not important	○	○	○	○	○	Very important

On average, how do simulated yields using nominal values (module power, inverter efficiency, ancilliary services, etc.) compare to measured ones? *

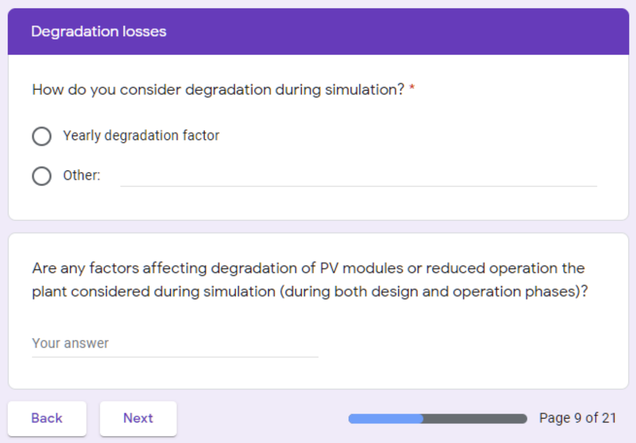
○ I don't know
○ Very optimistic (Simulated - actual > +5%)
○ Optimistic (+5% > Simulated - actual > +2%)
○ Well (-2% < simulated – actual < +2%)
○ Pessimistic (-2% > Simulated - actual > -5%)
○ Very pessimistic (-5% > Simulated -actual)

Do these differences imply a financial impact? *

○ Yes
○ No

If yes:

○ Fixed impact (penalty)
○ Depends on the % of losses (proportionnal)
○ Depends on the the % of losses (by ranges)
○ Other: ____



Degradation losses

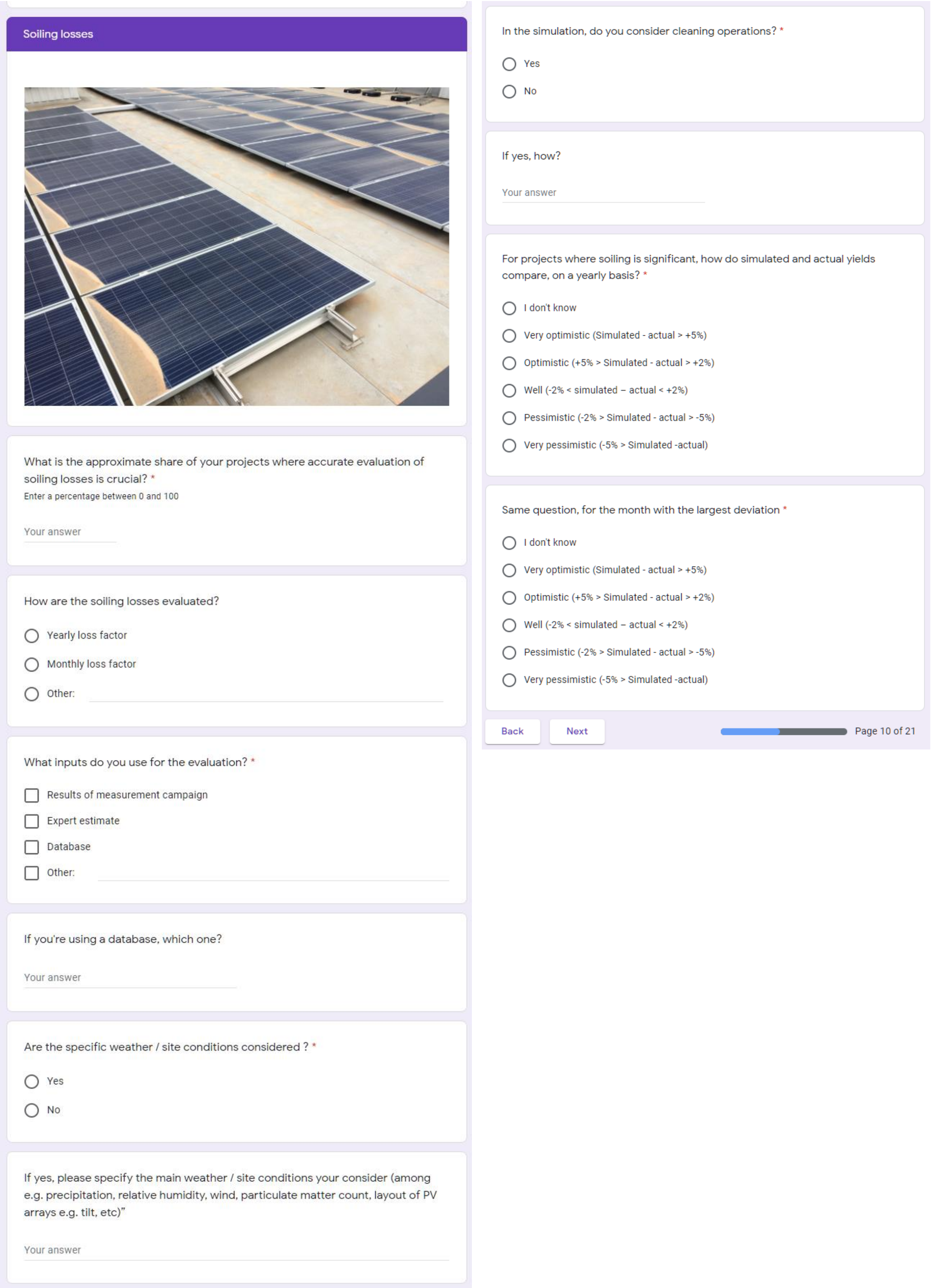
How do you consider degradation during simulation? *

○ Yearly degradation factor
○ Other: ____

Are any factors affecting degradation of PV modules or reduced operation the plant considered during simulation (during both design and operation phases)?

Your answer

Figure 2.4: Survey description – section 4 (1): Losses evaluation



Soiling losses

What is the approximate share of your projects where accurate evaluation of soiling losses is crucial? *
Enter a percentage between 0 and 100

Your answer

How are the soiling losses evaluated?

○ Yearly loss factor
○ Monthly loss factor
○ Other:

What inputs do you use for the evaluation? *

☐ Results of measurement campaign
☐ Expert estimate
☐ Database
☐ Other:

If you're using a database, which one?

Your answer

Are the specific weather / site conditions considered ? *

○ Yes
○ No

If yes, please specify the main weather / site conditions your consider (among e.g. precipitation, relative humidity, wind, particulate matter count, layout of PV arrays e.g. tilt, etc)"

Your answer

In the simulation, do you consider cleaning operations? *

○ Yes
○ No

If yes, how?

Your answer

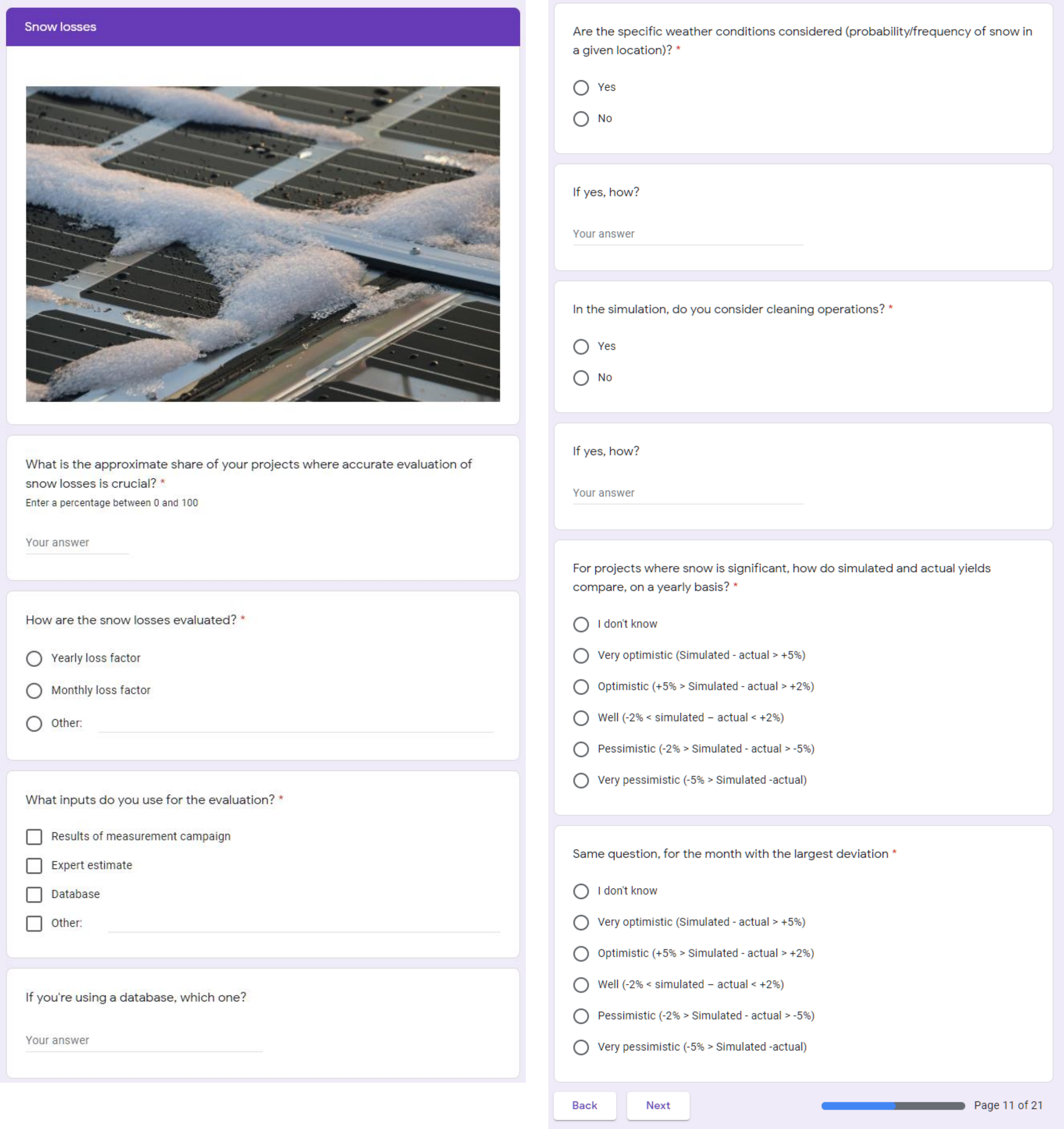
For projects where soiling is significant, how do simulated and actual yields compare, on a yearly basis? *

○ I don't know
○ Very optimistic (Simulated - actual > +5%)
○ Optimistic (+5% > Simulated - actual > +2%)
○ Well (-2% < simulated – actual < +2%)
○ Pessimistic (-2% > Simulated - actual > -5%)
○ Very pessimistic (-5% > Simulated -actual)

Same question, for the month with the largest deviation *

○ I don't know
○ Very optimistic (Simulated - actual > +5%)
○ Optimistic (+5% > Simulated - actual > +2%)
○ Well (-2% < simulated – actual < +2%)
○ Pessimistic (-2% > Simulated - actual > -5%)
○ Very pessimistic (-5% > Simulated -actual)

Figure 2.5: Survey description – section 4 (2) : Losses evaluation



Snow losses

What is the approximate share of your projects where accurate evaluation of snow losses is crucial? *
Enter a percentage between 0 and 100

Your answer

How are the snow losses evaluated? *

○ Yearly loss factor
○ Monthly loss factor
○ Other:

What inputs do you use for the evaluation? *

☐ Results of measurement campaign
☐ Expert estimate
☐ Database
☐ Other:

If you're using a database, which one?

Your answer

Are the specific weather conditions considered (probability/frequency of snow in a given location)? *

○ Yes
○ No

If yes, how?

Your answer

In the simulation, do you consider cleaning operations? *

○ Yes
○ No

If yes, how?

Your answer

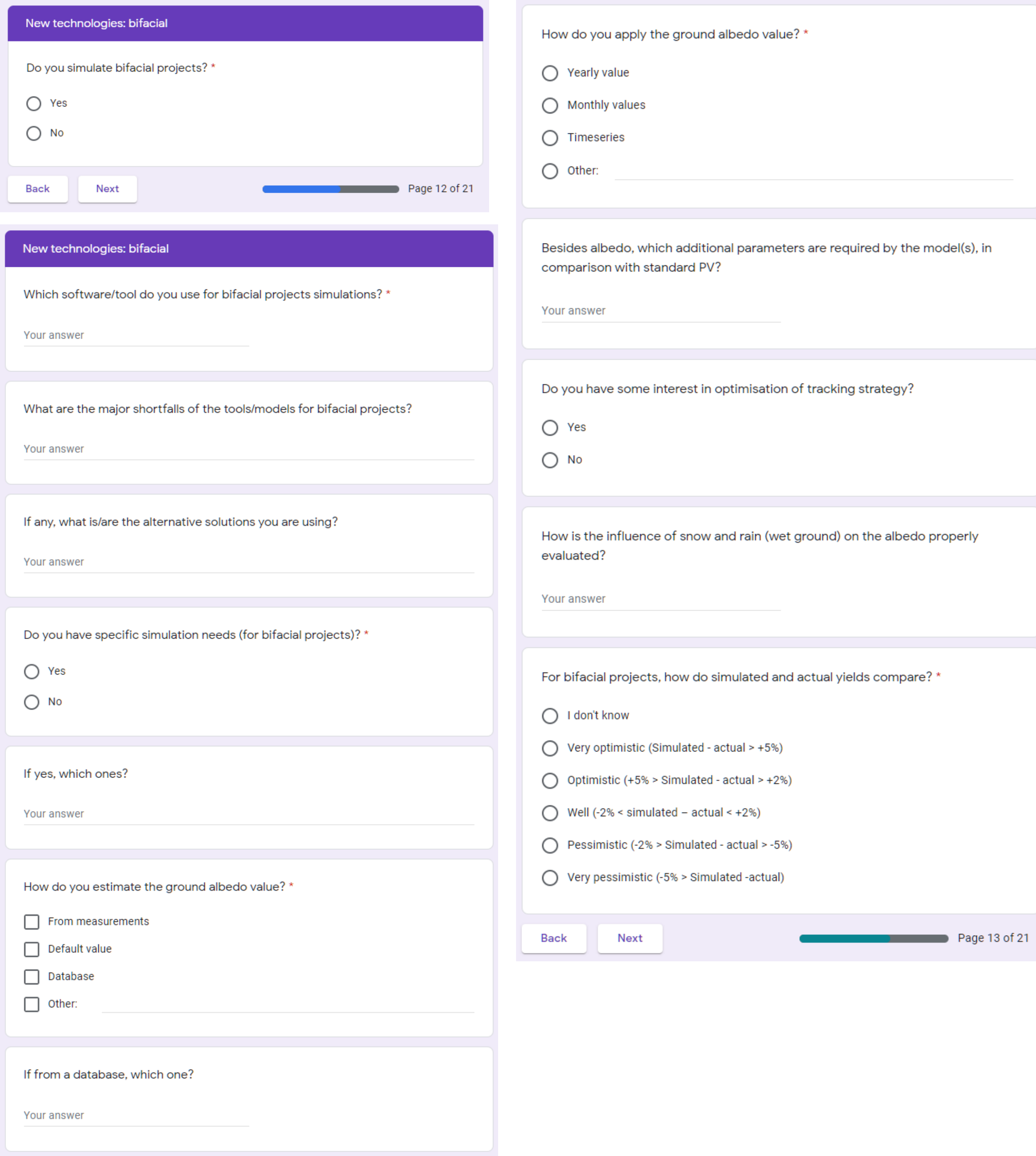
For projects where snow is significant, how do simulated and actual yields compare, on a yearly basis? *

○ I don't know
○ Very optimistic (Simulated - actual > +5%)
○ Optimistic (+5% > Simulated - actual > +2%)
○ Well (-2% < simulated – actual < +2%)
○ Pessimistic (-2% > Simulated - actual > -5%)
○ Very pessimistic (-5% > Simulated -actual)

Same question, for the month with the largest deviation *

○ I don't know
○ Very optimistic (Simulated - actual > +5%)
○ Optimistic (+5% > Simulated - actual > +2%)
○ Well (-2% < simulated – actual < +2%)
○ Pessimistic (-2% > Simulated - actual > -5%)
○ Very pessimistic (-5% > Simulated -actual)

Figure 2.6: Survey description – section 4 (3): Losses evaluation



New technologies: bifacial

Do you simulate bifacial projects? *

○ Yes

○ No

New technologies: bifacial

Which software/tool do you use for bifacial projects simulations? *

Your answer

What are the major shortfalls of the tools/models for bifacial projects?

Your answer

If any, what is/are the alternative solutions you are using?

Your answer

Do you have specific simulation needs (for bifacial projects)? *

○ Yes

○ No

If yes, which ones?

Your answer

How do you estimate the ground albedo value? *

☐ From measurements

☐ Default value

☐ Database

☐ Other:

If from a database, which one?

Your answer

How do you apply the ground albedo value? *

○ Yearly value

○ Monthly values

○ Timeseries

○ Other:

Besides albedo, which additional parameters are required by the model(s), in comparison with standard PV?

Your answer

Do you have some interest in optimisation of tracking strategy?

○ Yes

○ No

How is the influence of snow and rain (wet ground) on the albedo properly evaluated?

Your answer

For bifacial projects, how do simulated and actual yields compare? *

○ I don't know

○ Very optimistic (Simulated - actual > +5%)

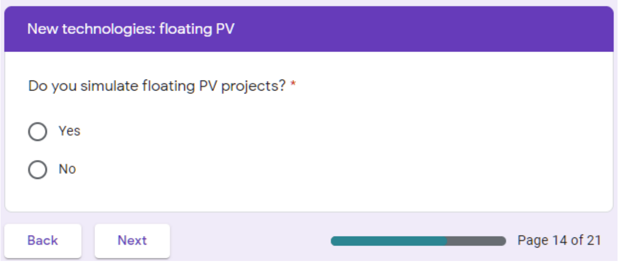
○ Optimistic (+5% > Simulated - actual > +2%)

○ Well (-2% < simulated – actual < +2%)

○ Pessimistic (-2% > Simulated - actual > -5%)

○ Very pessimistic (-5% > Simulated -actual)

Figure 2.7: Survey description – section 5 (1): New technologies

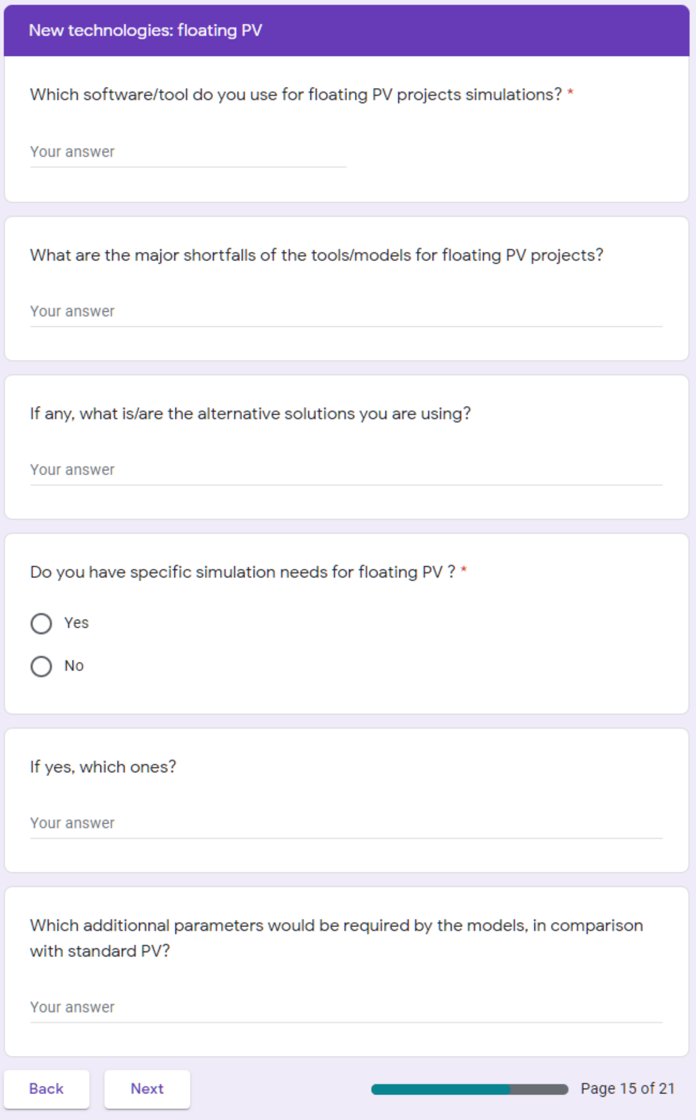


New technologies: floating PV

Do you simulate floating PV projects? *

○ Yes

○ No



New technologies: floating PV

Which software/tool do you use for floating PV projects simulations? *

Your answer

What are the major shortfalls of the tools/models for floating PV projects?

Your answer

If any, what is/are the alternative solutions you are using?

Your answer

Do you have specific simulation needs for floating PV ? *

○ Yes

○ No

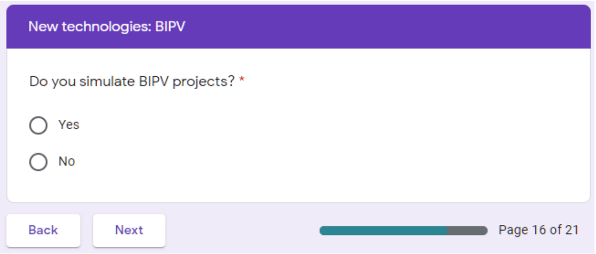
If yes, which ones?

Your answer

Which additionnal parameters would be required by the models, in comparison with standard PV?

Your answer

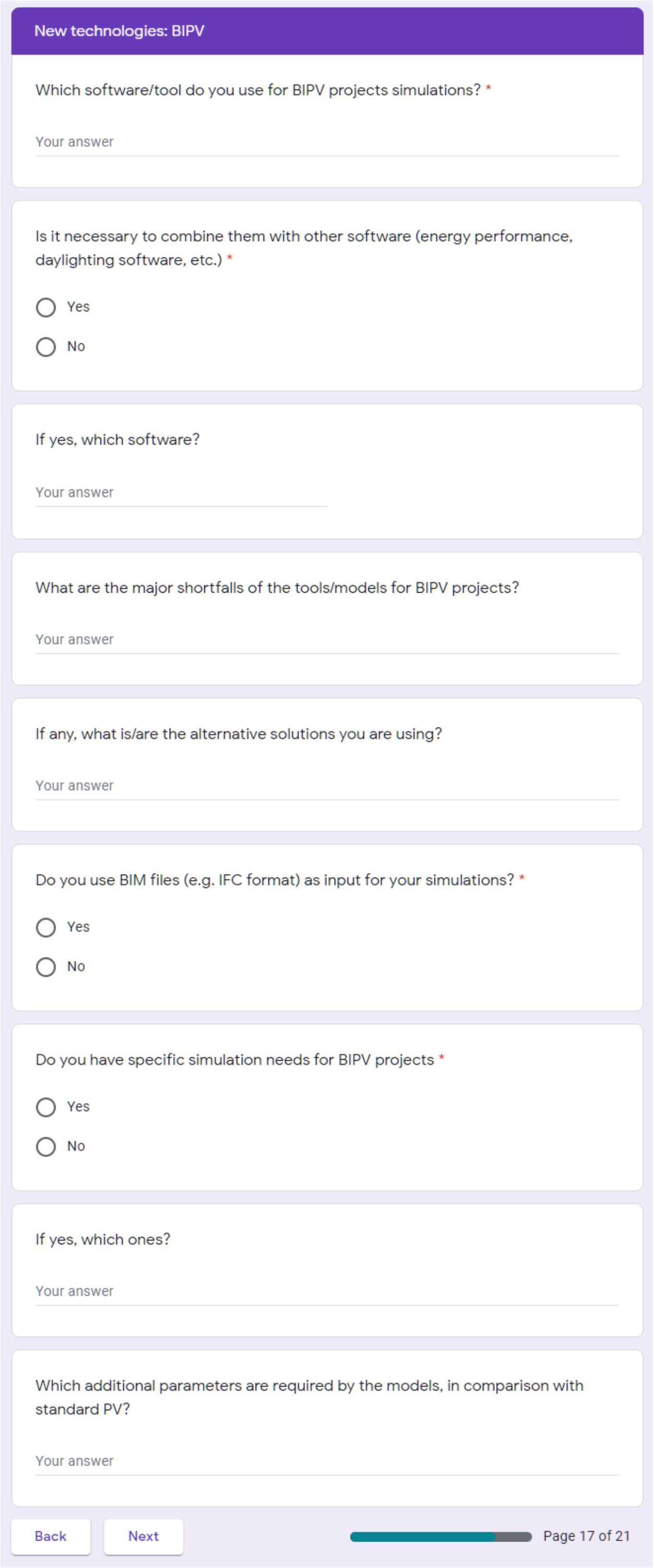
Figure 2.8: Survey description – section 5 (2): New technologies



Do you simulate BIPV projects? *

○ Yes

○ No



Which software/tool do you use for BIPV projects simulations? *

Your answer

Is it necessary to combine them with other software (energy performance, daylighting software, etc.) *

○ Yes

○ No

If yes, which software?

Your answer

What are the major shortfalls of the tools/models for BIPV projects?

Your answer

If any, what is/are the alternative solutions you are using?

Your answer

Do you use BIM files (e.g. IFC format) as input for your simulations? *

○ Yes

○ No

Do you have specific simulation needs for BIPV projects *

○ Yes

○ No

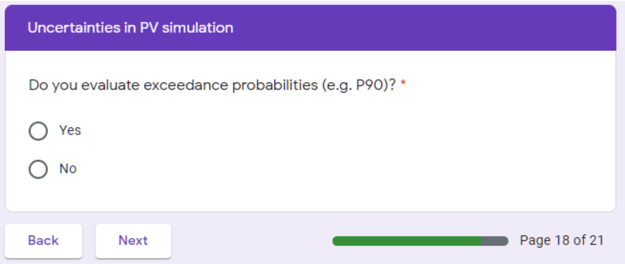
If yes, which ones?

Your answer

Which additional parameters are required by the models, in comparison with standard PV?

Your answer

Figure 2.9: Survey description – section 5 (3): New technologies

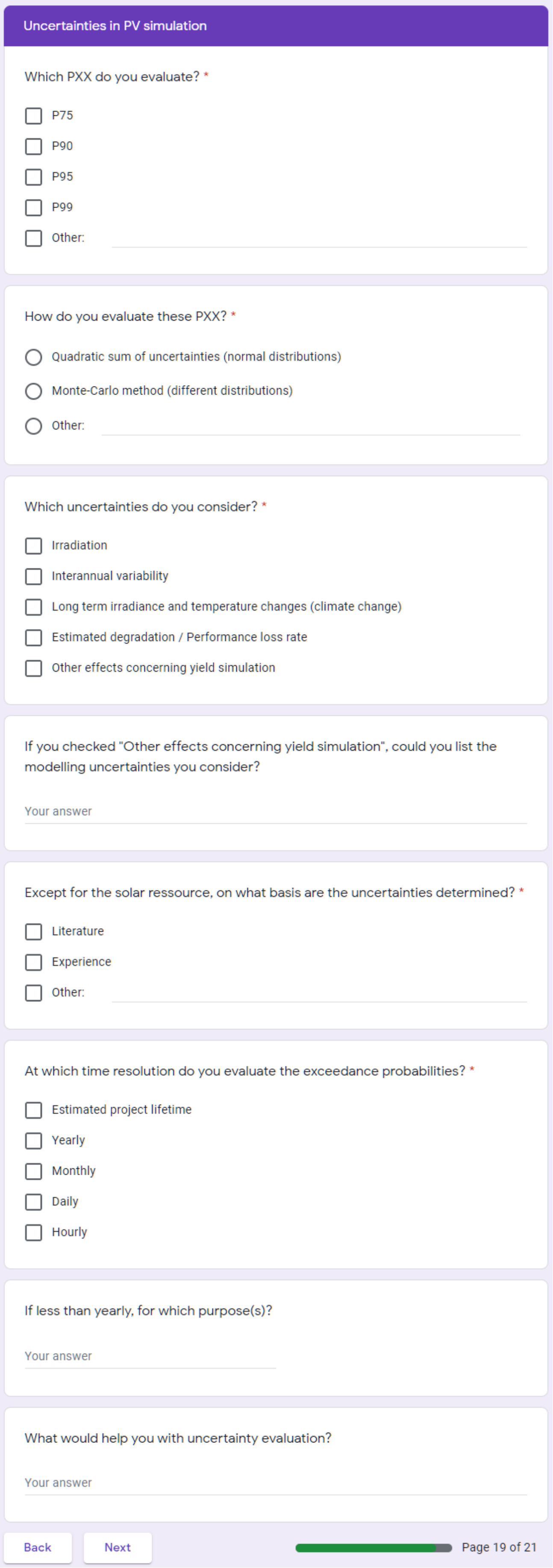


Uncertainties in PV simulation

Do you evaluate exceedance probabilities (e.g. P90)? *

○ Yes

○ No



Uncertainties in PV simulation

Which PXX do you evaluate? *

☐ P75
☐ P90
☐ P95
☐ P99
☐ Other:

How do you evaluate these PXX? *

○ Quadratic sum of uncertainties (normal distributions)
○ Monte-Carlo method (different distributions)
○ Other:

Which uncertainties do you consider? *

☐ Irradiation
☐ Interannual variability
☐ Long term irradiance and temperature changes (climate change)
☐ Estimated degradation / Performance loss rate
☐ Other effects concerning yield simulation

If you checked "Other effects concerning yield simulation", could you list the modelling uncertainties you consider?

Your answer

Except for the solar ressource, on what basis are the uncertainties determined? *

☐ Literature
☐ Experience
☐ Other:

At which time resolution do you evaluate the exceedance probabilities? *

☐ Estimated project lifetime
☐ Yearly
☐ Monthly
☐ Daily
☐ Hourly

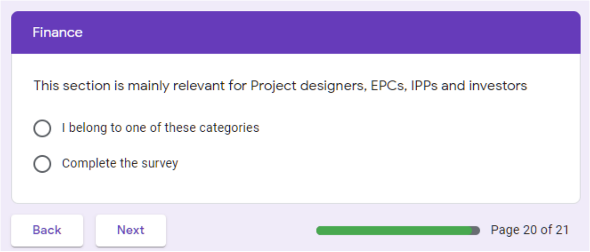
If less than yearly, for which purpose(s)?

Your answer

What would help you with uncertainty evaluation?

Your answer

Figure 2.10: Survey description – section 6: Uncertainties

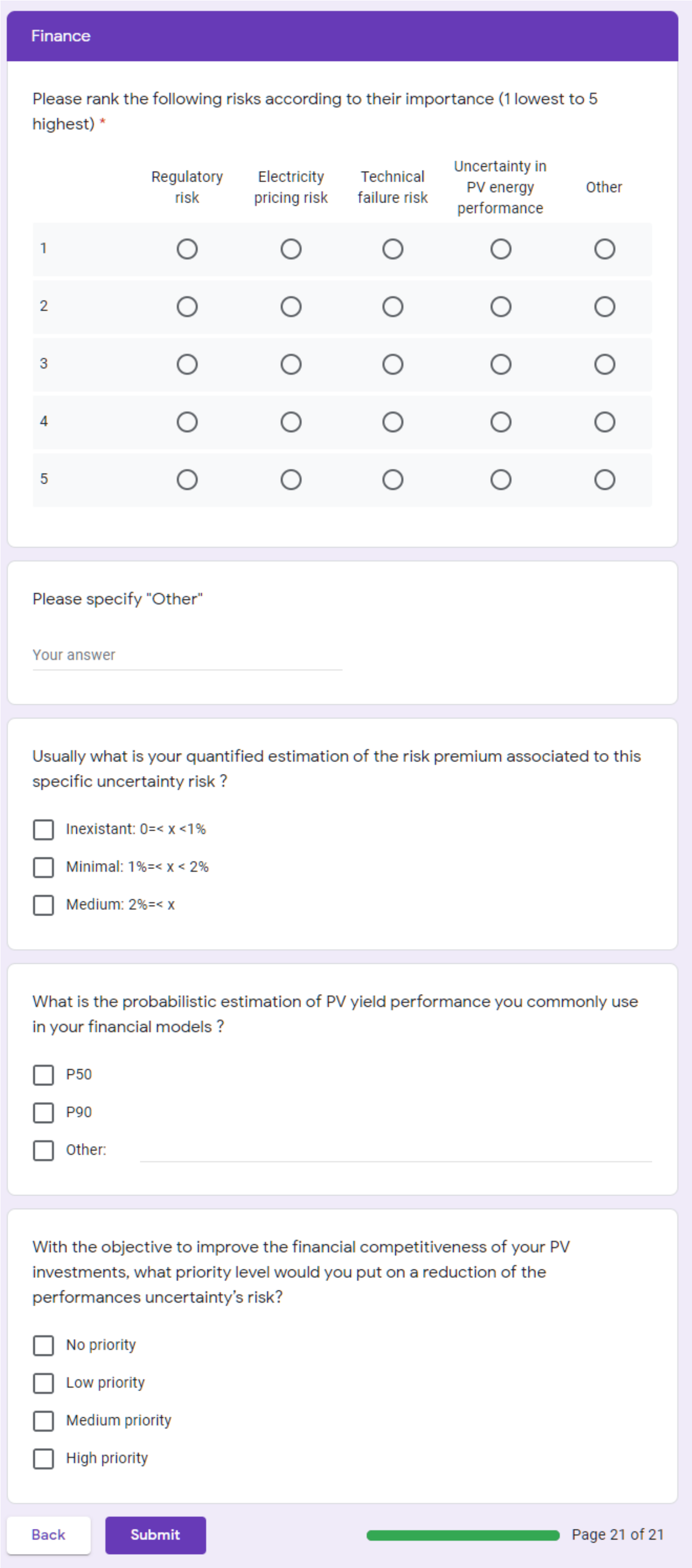


Finance

This section is mainly relevant for Project designers, EPCs, IPPs and investors

○ I belong to one of these categories

○ Complete the survey



Finance

Please rank the following risks according to their importance (1 lowest to 5 highest) *

	Regulatory risk	Electricity pricing risk	Technical failure risk	Uncertainty in PV energy performance	Other
1	○	○	○	○	○
2	○	○	○	○	○
3	○	○	○	○	○
4	○	○	○	○	○
5	○	○	○	○	○

Please specify "Other"

Your answer

Usually what is your quantified estimation of the risk premium associated to this specific uncertainty risk ?

☐ Inexistant: 0=< x <1%

☐ Minimal: 1%=< x < 2%

☐ Medium: 2%=< x

What is the probabilistic estimation of PV yield performance you commonly use in your financial models ?

☐ P50

☐ P90

☐ Other:

With the objective to improve the financial competitiveness of your PV investments, what priority level would you put on a reduction of the performances uncertainty's risk?

☐ No priority

☐ Low priority

☐ Medium priority

☐ High priority

Figure 2.11: Survey description – section 7: Finance

2.1.2 Distribution of the survey

The distribution of the survey was performed through the following ways:

- WIP, the partner in charge of the dissemination & communication related to the project, prepared **LinkedIn and Twitter posts,** which were then shared by the partners of the project.

WIP Renewable Energies
912 followers
3mo •

The new #SERENDIPV project is looking to better understand its stakeholders' needs, to do so we invite #PV industry professionals to fill out the consortium`s new survey on PV System #design and #Software usage + more. Survey link here: https://lnkd.in/dmC3X9F #renewableenergy #solarenergy

YOU ARE INVITED TO PARTICIPATE
"Understanding the PV industry survey: PV simulation and software usage, needs, uncertainties, growth & finance"

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement Nº 953016

Figure 2.12: Survey distribution – LinkedIn post

WIP Renewable Energies @WIPRenewables · 9 mars

The new #SERENDIPV project is looking to better understand its stakeholders' needs, to do so we invite #PV industry professionals to fill out the consortium`s new survey on PV System #design and #Software usage + more. Survey link here: wip-munich.de/projects/proje... #renewableenergy

YOU ARE INVITED TO PARTICIPATE
"Understanding the PV industry survey: PV simulation and software usage, needs, uncertainties, growth & finance"

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement Nº 953016

Figure 2.13: Survey distribution – Twitter post

- **Blogs articles**

  https://www.wip-munich.de/projects/project-serendi-pv/

  https://solargis.com/blog/solargis-news/serendi-pv-solargis-joins-large-european-rd-initiative-to-increase-penetration-and-integration-of-photovoltaics

- **Articles in newsletters and emailing campaigns** dedicated to the survey reached more than 11 000 recipients
- **Direct emailing** to 150 recipients. Emails were sent by the partners, to chosen contacts, i.e., people interested by the topics of the survey, and more likely to answer it. Reminders were also sent.

2.2 Responses and analysis

Despite the efforts to distribute the survey, **only 50 recipients replied** to the questionnaire.

Their responses are presented according to the different sections of the survey.

2.2.1 Section 1: General information

The profiles of the respondents are well balanced between the different categories, except for “Manufacturer” who are less likely users of the models and simulation tools addressed by SERENDI-PV.

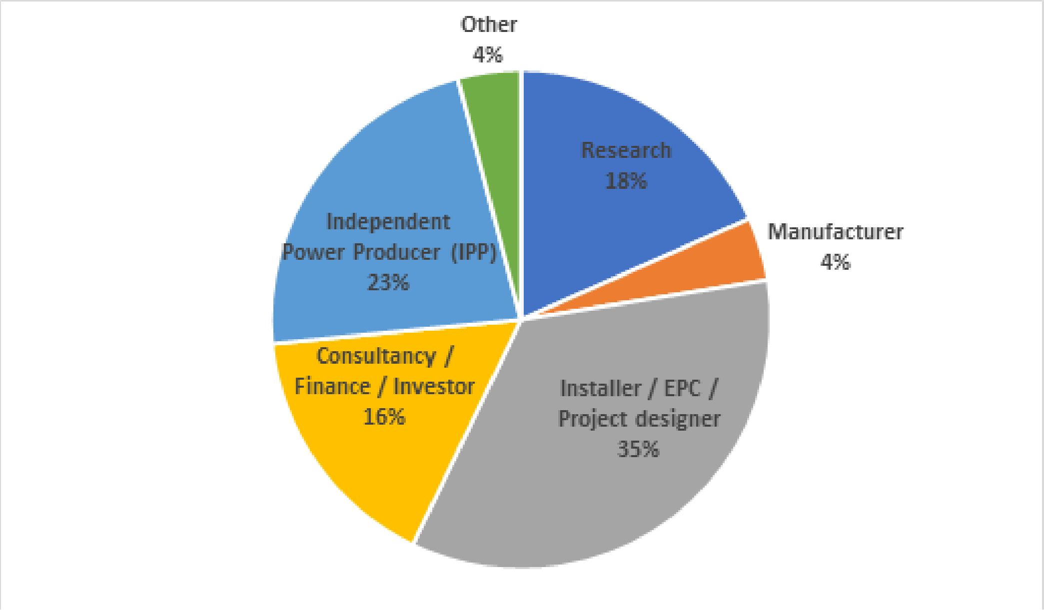


Figure 2.14: Type of users who responded the survey

The respondents’ countries mainly reflect the countries where the project partners are active or have direct contacts. As can be seen in Figure 2.15 most of the respondents are Europeans.

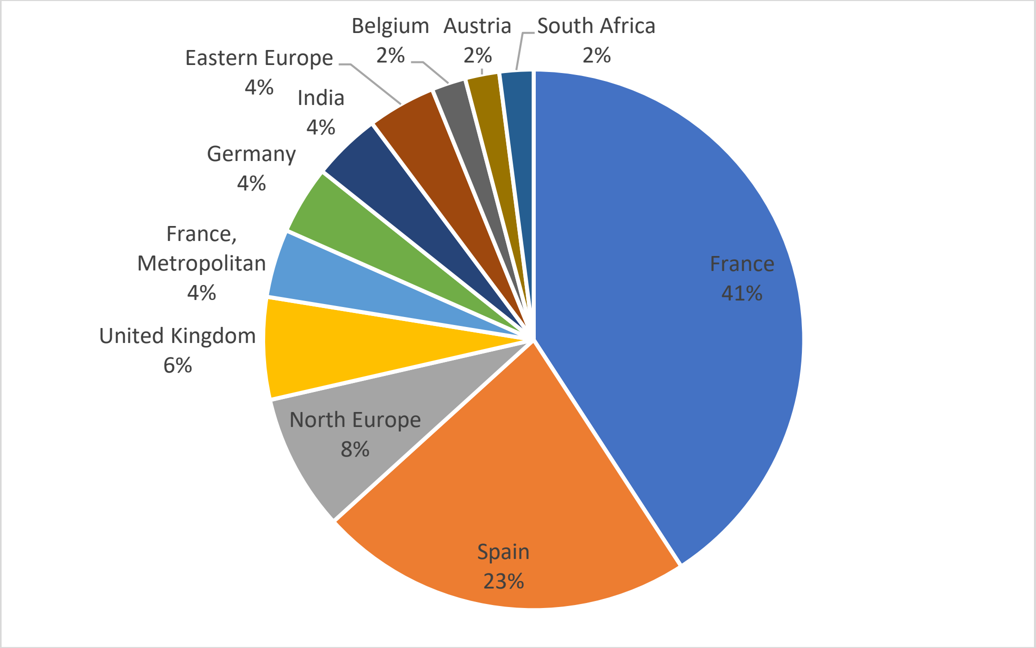


Figure 2.15: Countries of respondents

Large ground-mounted projects are mostly simulated, followed by BIPV/BAPV projects.

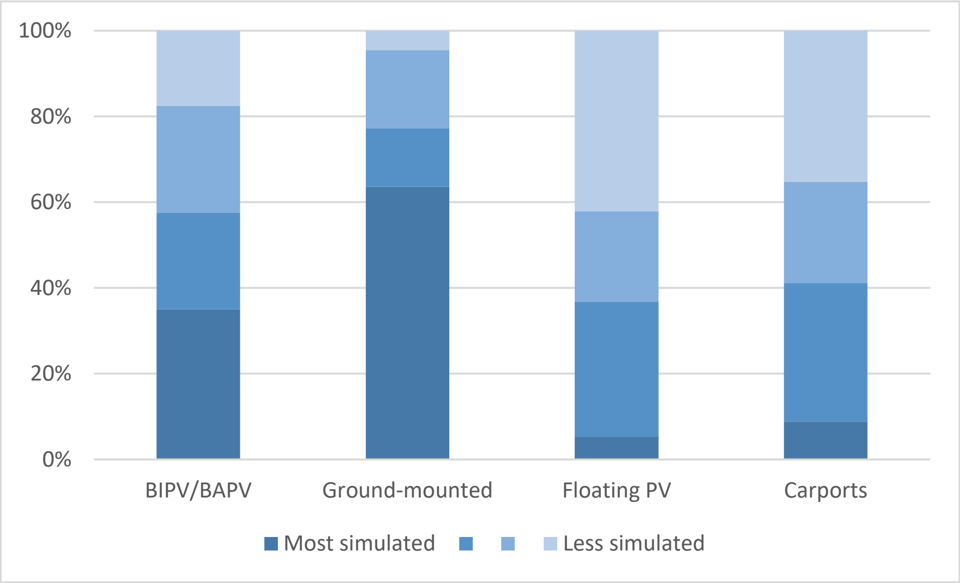


Figure 2.16: Type of projects simulated

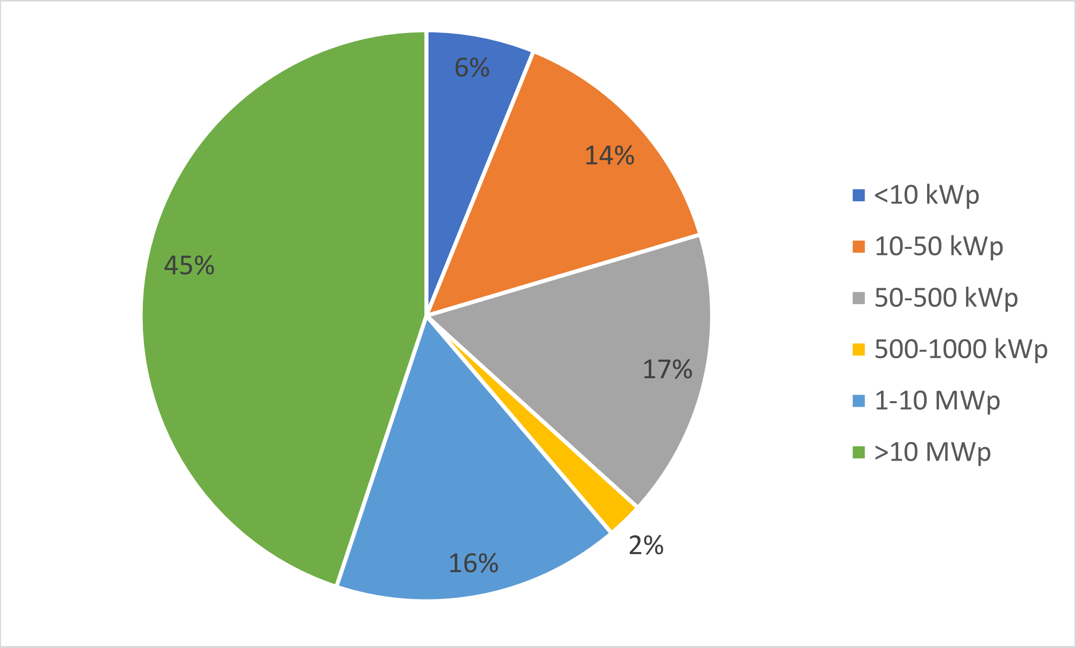


Figure 2.17: Average power of simulated projects

2.2.2 Section 2: Meteorological data

Respondents are using both free and licensed meteorological data.

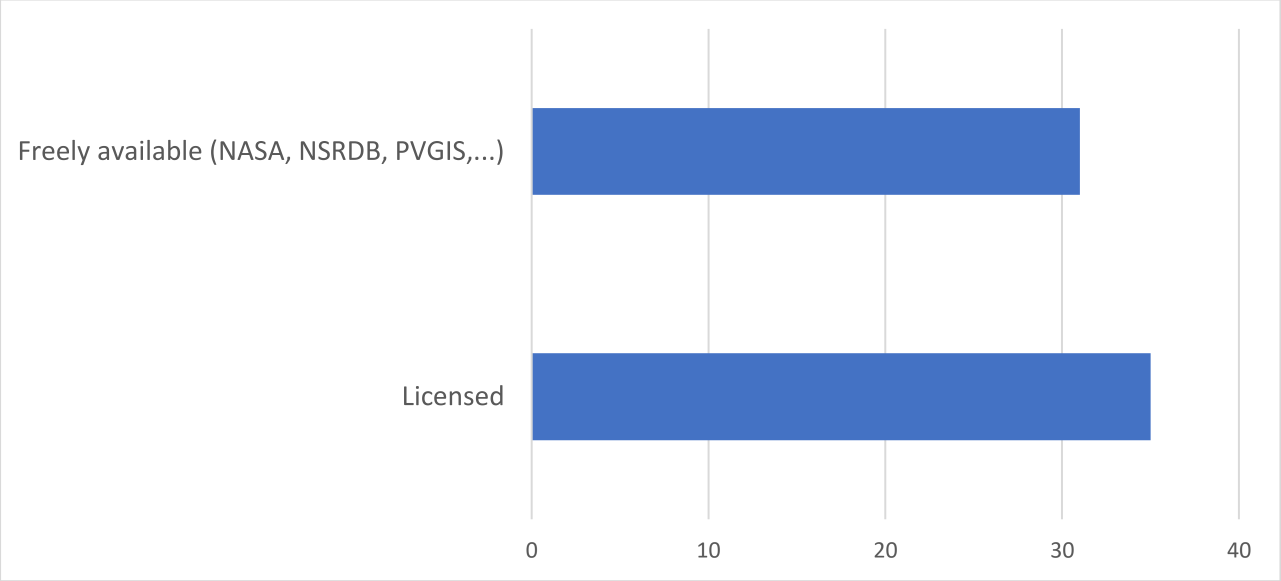


Figure 2.18: Source of solar and meteo data used for simulation

Most of the respondents who specified the provider of licensed data are using Solargis data (74%). The other licensed sources are: Metenorm, PVGIS, Soda-Helioclim, Reuniwatt and solar cadastre (from Cythelia). Historical data of 15 and more years are preferred over less extended time ranges.

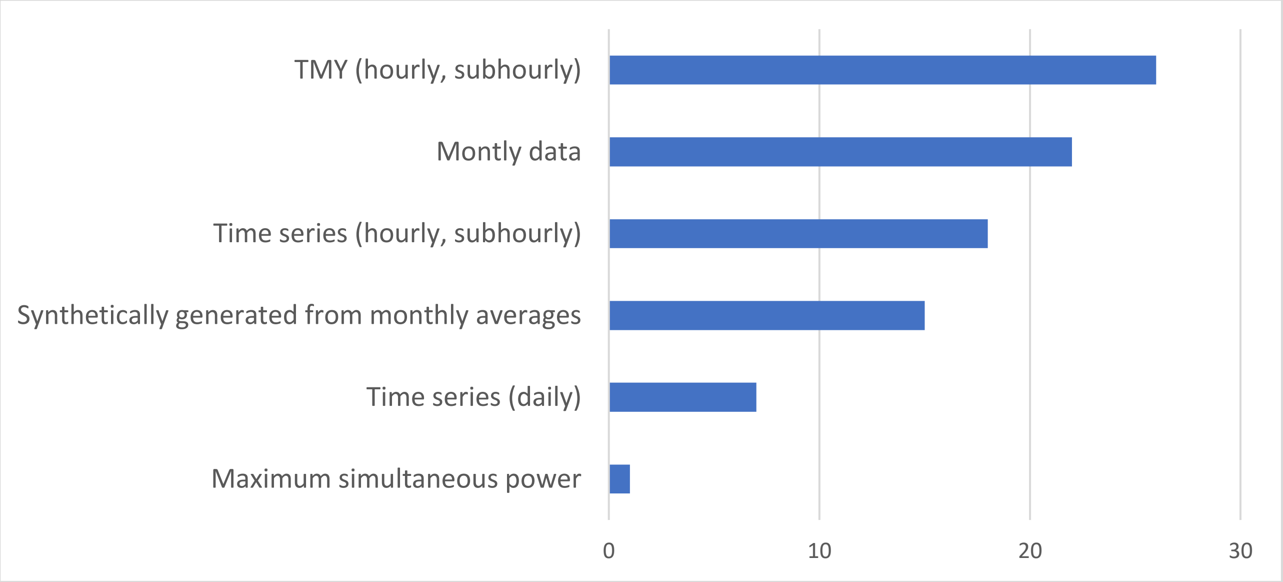


Figure 2.19: Type of meteorological data

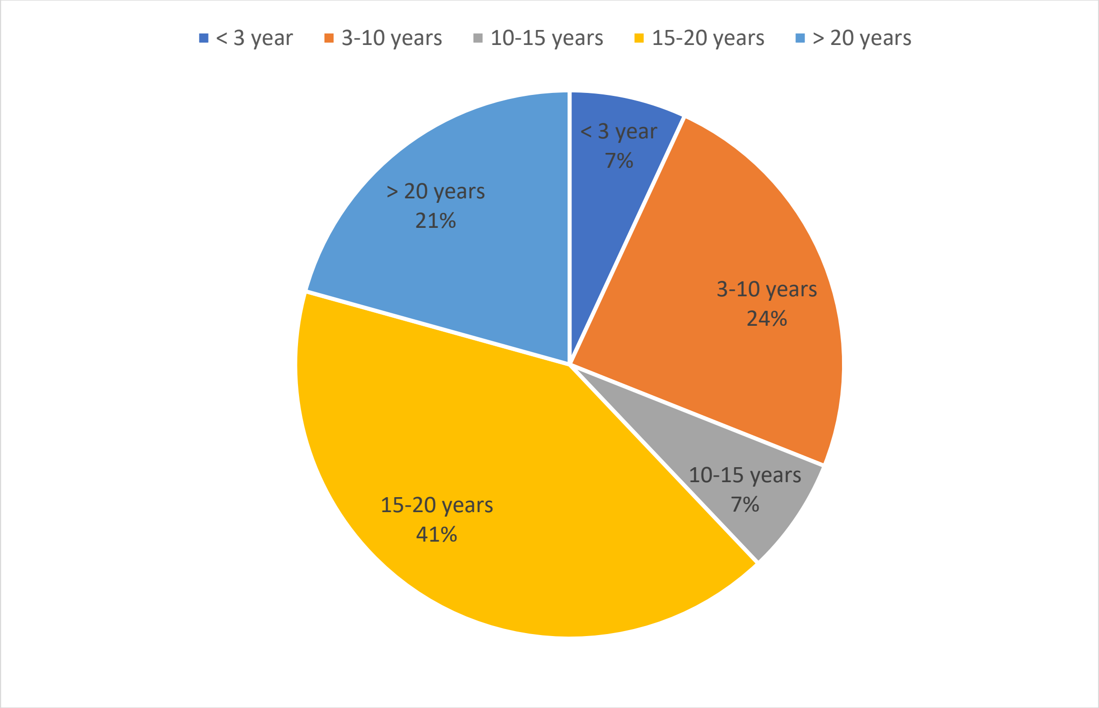


Figure 2.20: Time ranges of meteorological data

2.2.3 Section 3: PV Software usage

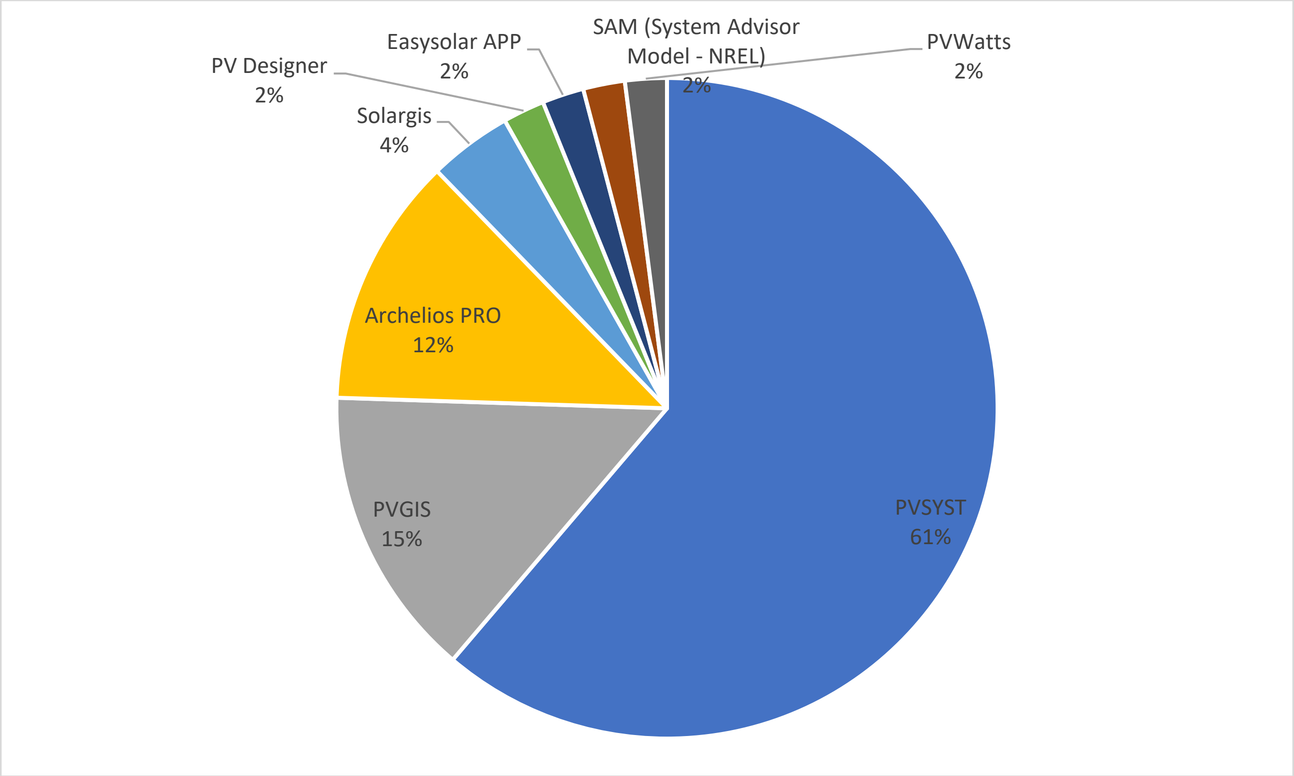


Figure 2.21: First software used – 49 responses

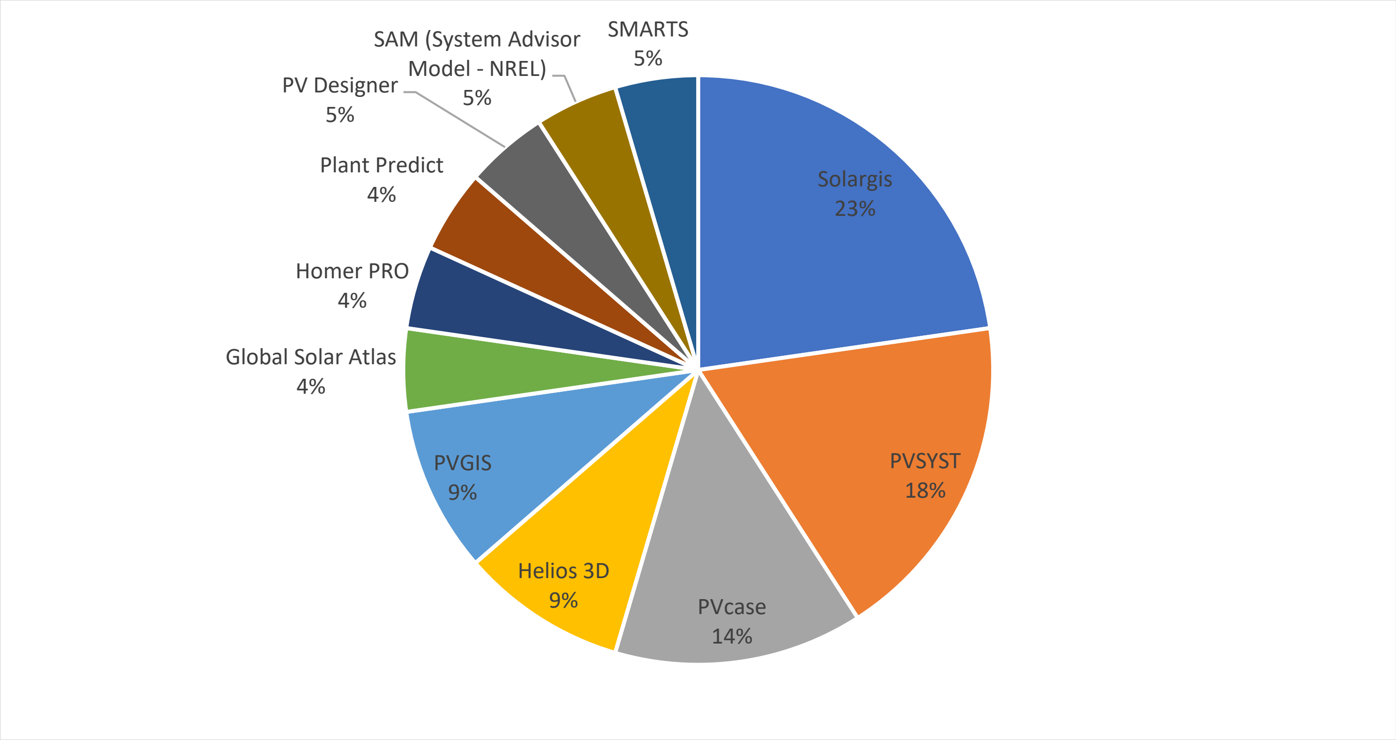


Figure 2.22: Second software used – 22 responses

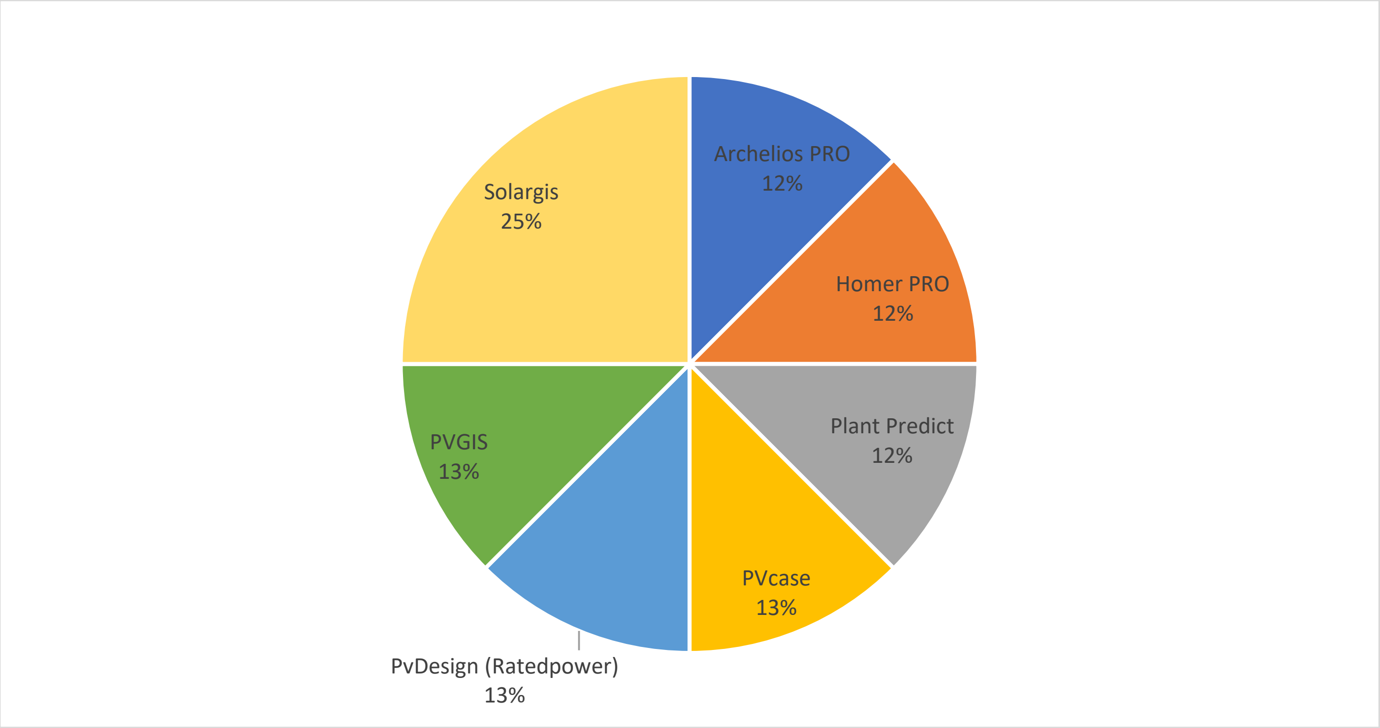


Figure 2.23: Third software used – 8 responses

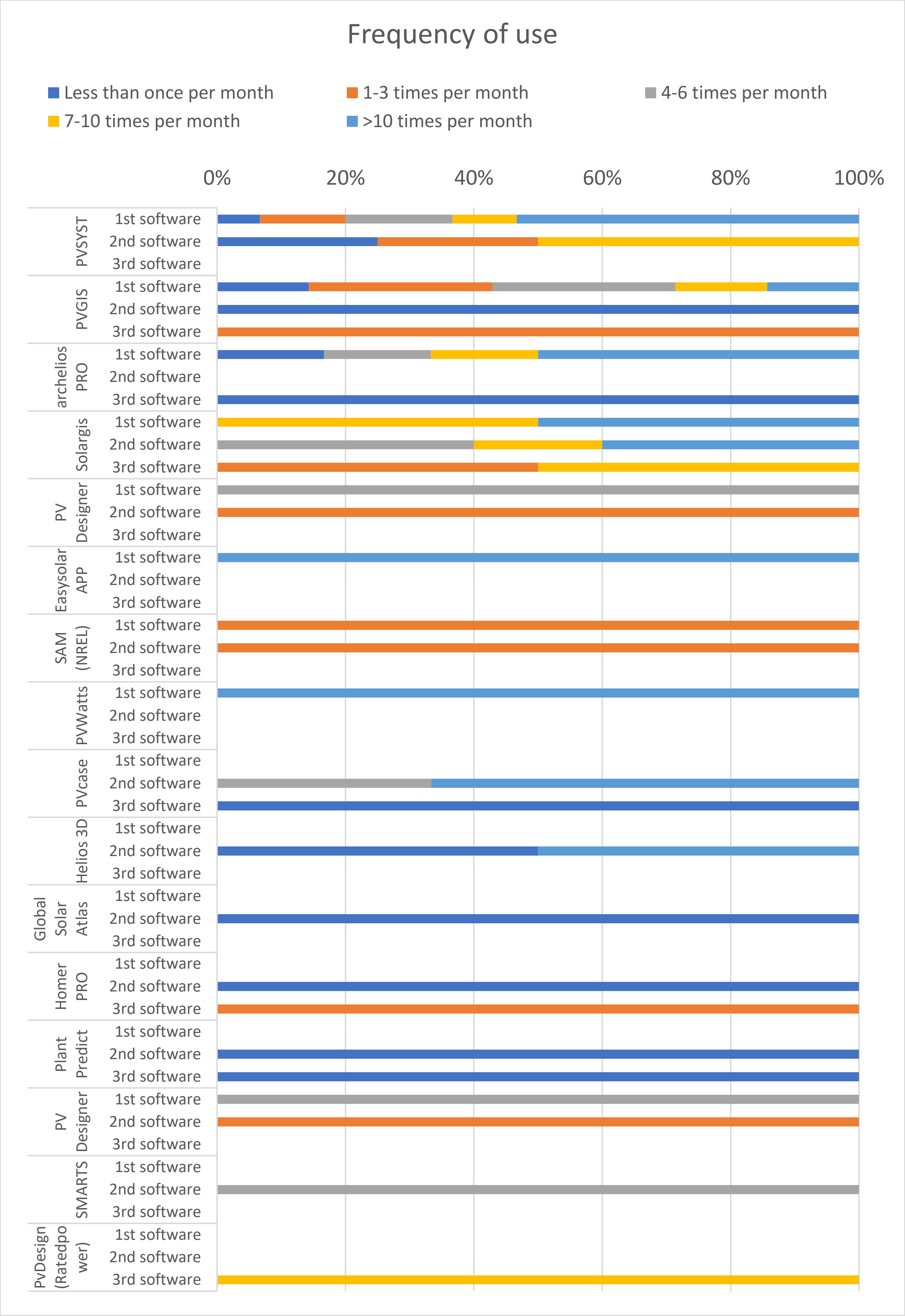


Figure 2.24: Frequency of use

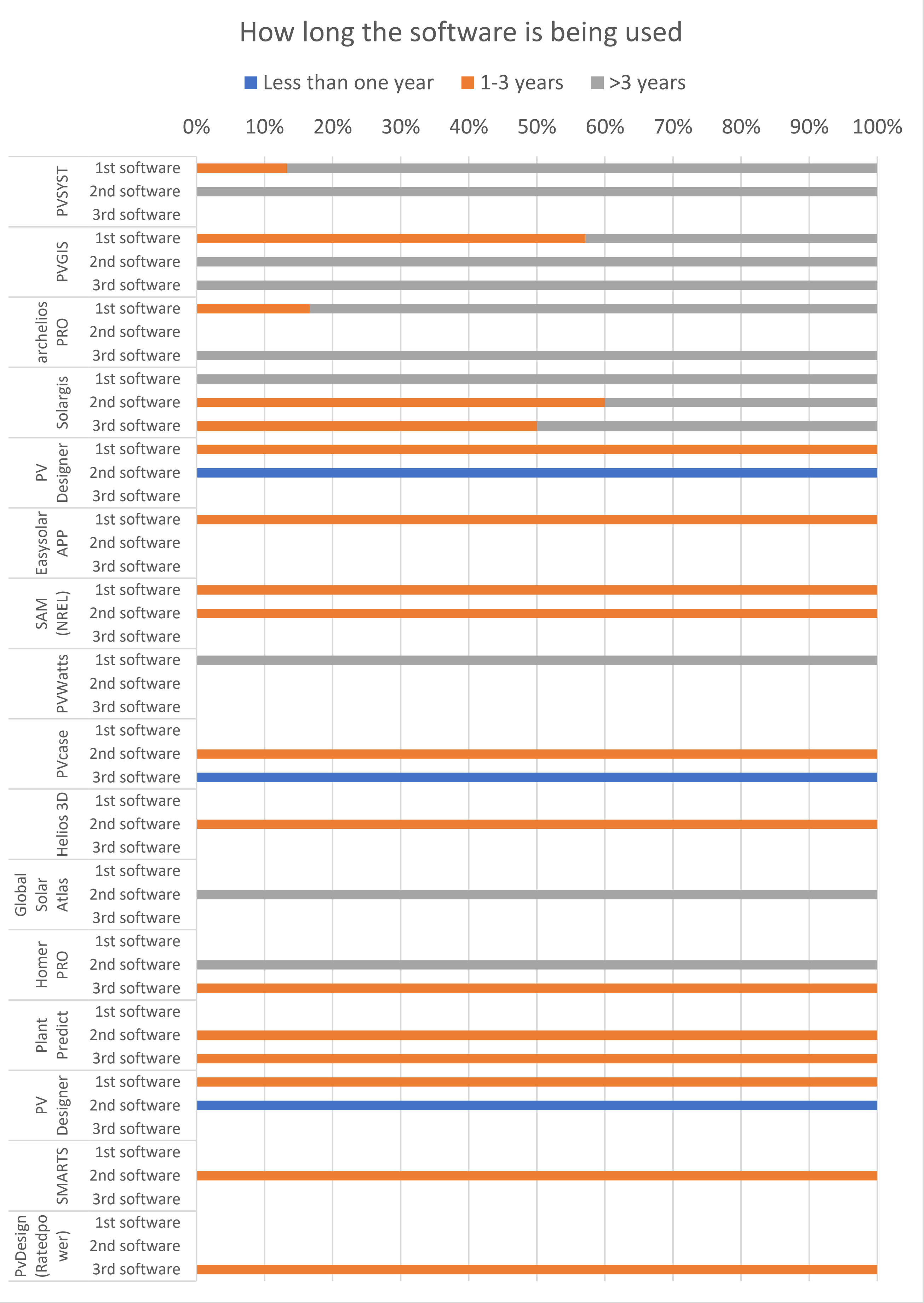


Figure 2.25: How long is it being used

The perceived ease of use of each software is given in the figure below (ordered as a function of their usage).

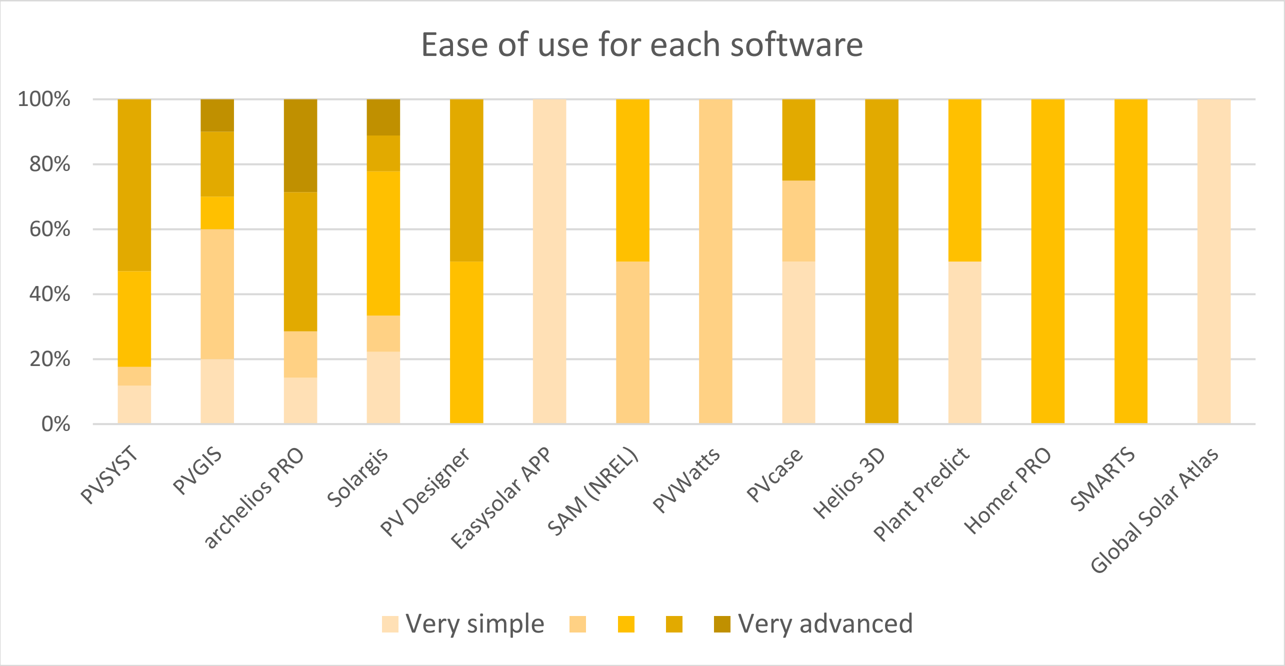


Figure 2.26: Ease of use of each software

The main reasons for using the software are given in the figure below (ordered as a function of their usage:

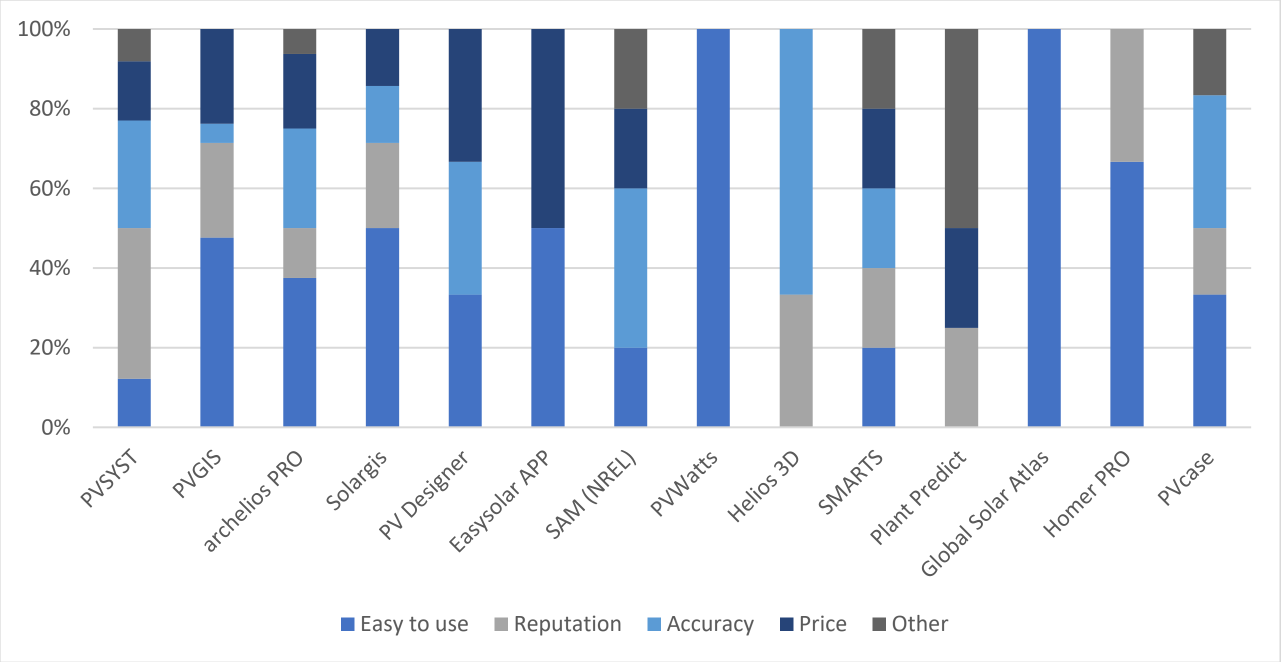


Figure 2.27: Reasons for using each software

The following “Other” reasons were given:

- Archelios PRO: link with existing 3D tool
- PVSYST: bankability, requested by clients, industry standard
- SMARTS: quick simulation of spectral effects
- PVcase: semi-automatic generation of layout, flexibility
- Plant Predict: online interface and API

The respondents were asked to give 5 features they would like to have in a PV simulation tool. The features cited more than three times are (in descending order, in **bold** the topics addressed in WP2 of SERENDI-PV):

- Shadow graphs
- **Bifacial, bifacial on trackers**
- Help to fill **losses factors (soiling, snow**, cable,…)
- Flexibility on layout
- Self-consumption simulation
- **Floating**
- History data simulation and detailed analytics
- **Import/compatibility (DXF, DWG, BIM, CAD, …)**

Among the answers concerning the aspects of modelling which require more accuracy, the most frequent were related to the following topics: losses, irradiance, bifacial PV (including with trackers), shadings, floating PV.

As for the needs not satisfied by the software, the answers are not easy to interpret. Direct interviews would be required for a better understanding and will be considered to be carried out in the context of the upcoming Tasks 2.2 and 2.3 of WP2.

2.2.4 Section 4: Losses evaluation

2.2.4.1 General

Following picture depicts the prominence of the three losses SERENDI-PV addresses:

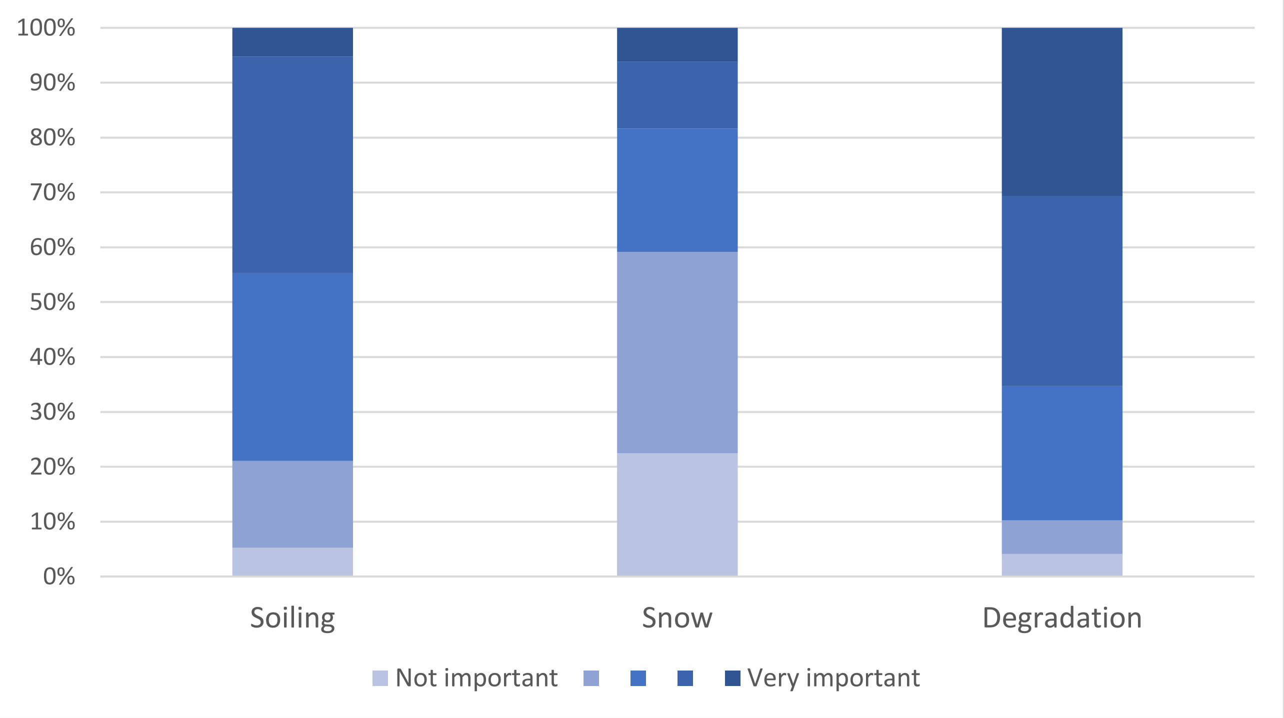


Figure 2.28: Prominence of the 3 losses SERENDI-PV addresses for the respondents

37 respondents gave the estimation of comparison between simulation results and measurements.

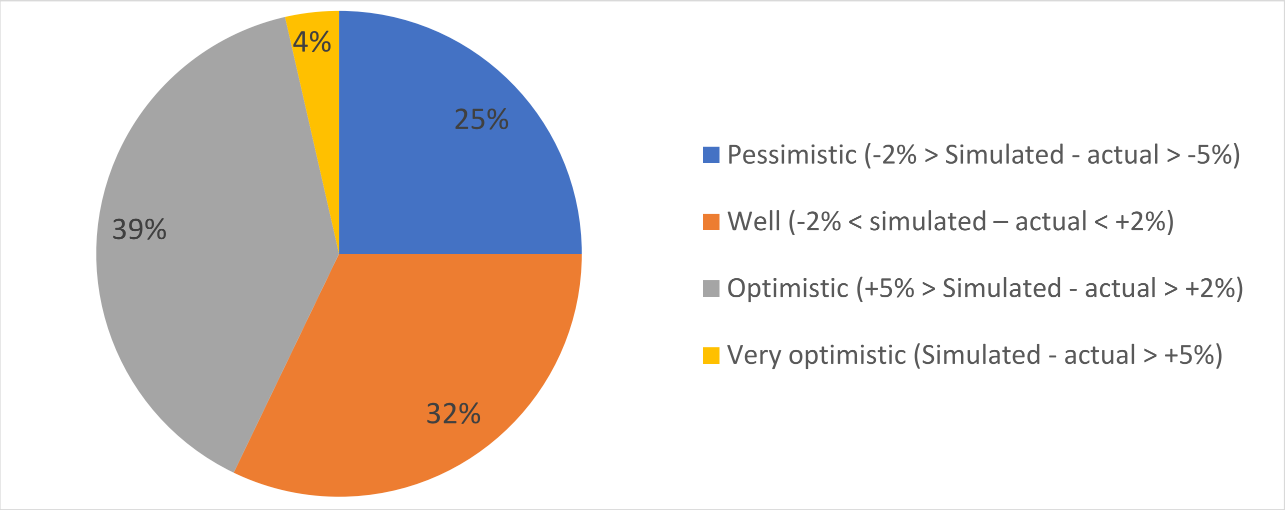


Figure 2.29: Comparison between simulation results and measurements

For 74% of the respondents, the deviation does have a financial impact.

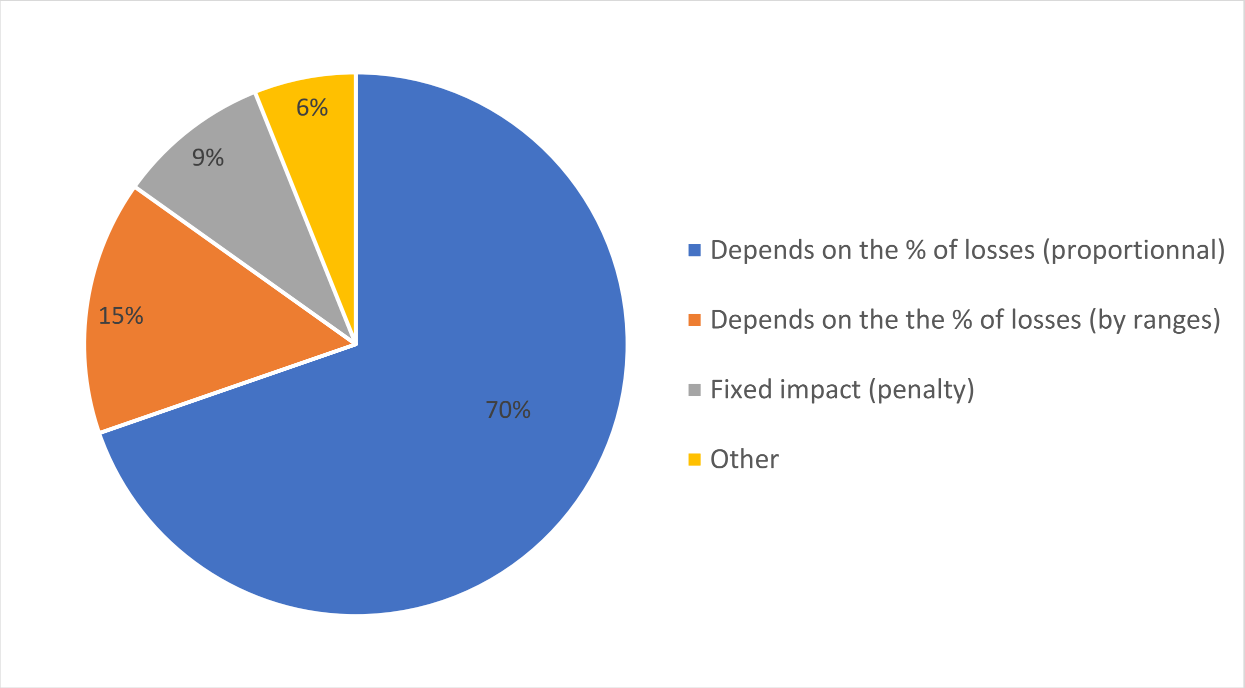


Figure 2.30: Impact of yield deviations

2.2.4.2 Degradation losses

90% of respondent consider degradation in their simulation with a yearly degradation factor. The following details are given:

- Different value for the first year, same value for the other years
- Accounting of the clipping of the inverters

The answers to the question "Are any factors affecting degradation of PV modules or reduced operation the plant considered during simulation (during both design and operation phases)?" were not all relevant, except the following ones:

- Decreasing availability of the plant
- LID, LeTID
- Additional mismatch
- Grid curtailment

2.2.4.3 Soiling losses

The following graph shows the share of projects for which soiling losses are perceived as crucial. For 40% of the respondents, this share is less than 10%. For 55% of them, this share is less 20%, and so on.

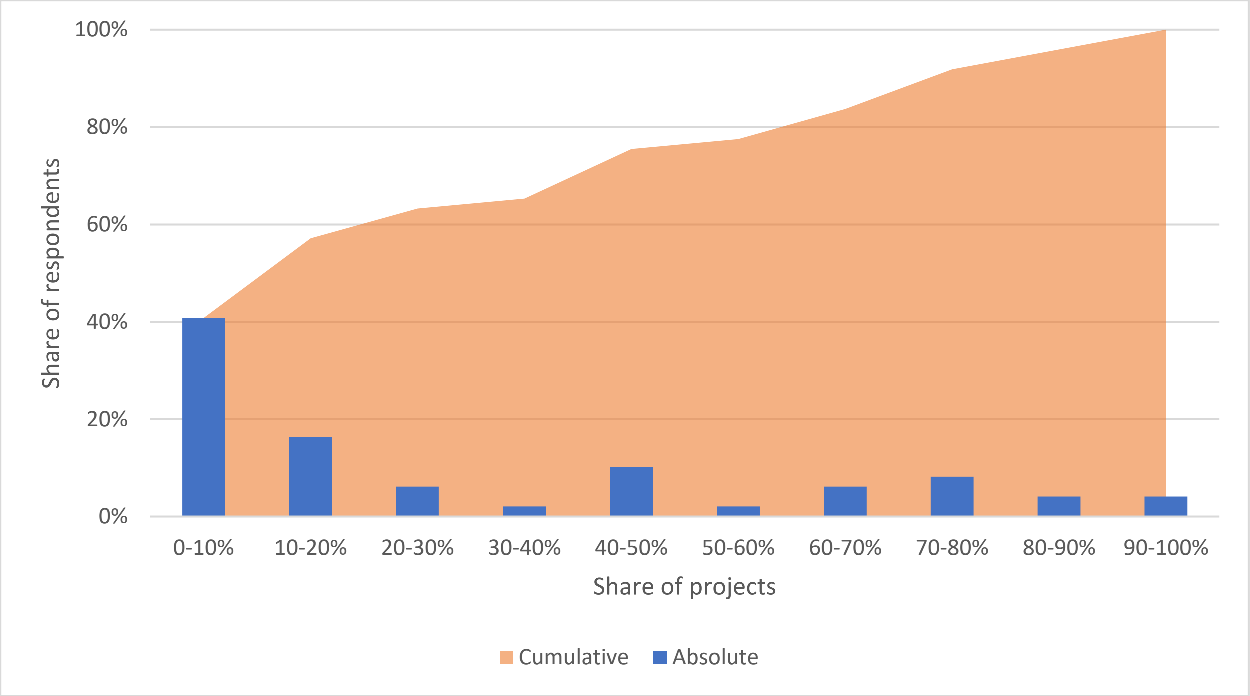


Figure 2.31: Approximate share of projects where soiling losses are crucial

Soiling losses are evaluated mainly with:

- yearly loss factor (65%),
- less frequently using monthly loss factor (35%).

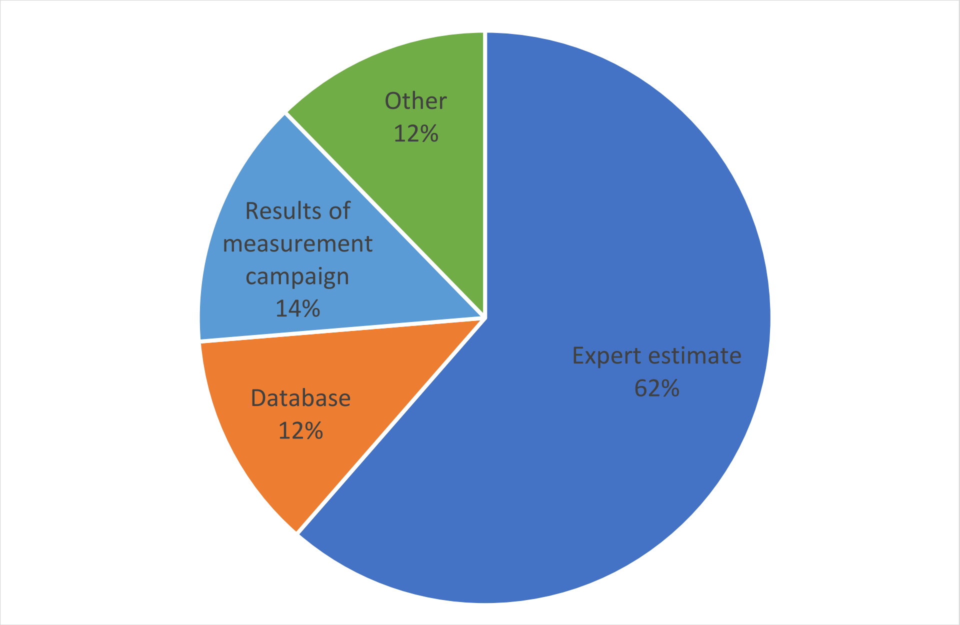


Figure 2.32: Inputs used for the evaluation of soiling losses

If not internal, the databases used are:

- NASA
- Prospect
- SODA MERA
- Solargis
- Météo France (for France)

78% of the respondents do consider specific site/weather conditions for the estimation of soiling losses.

The factors considered are (in descending order of responses):

- Precipitation
- Tilt / layout of PV array
- Sand/dust/pollution
- Relative humidity
- Agricultural season
- Birds

Half of the respondents do consider cleaning operations in the simulation. The cleaning operations are considered in the simulations by adjusting the soiling loss, based on expert guess. The use of experimental data has been cited three times.

32 out of 49 respondents do not know how simulated and actual yields compare for projects where soiling is significant. The remaining ones find their simulation rather optimistic or in alignment with the reality.

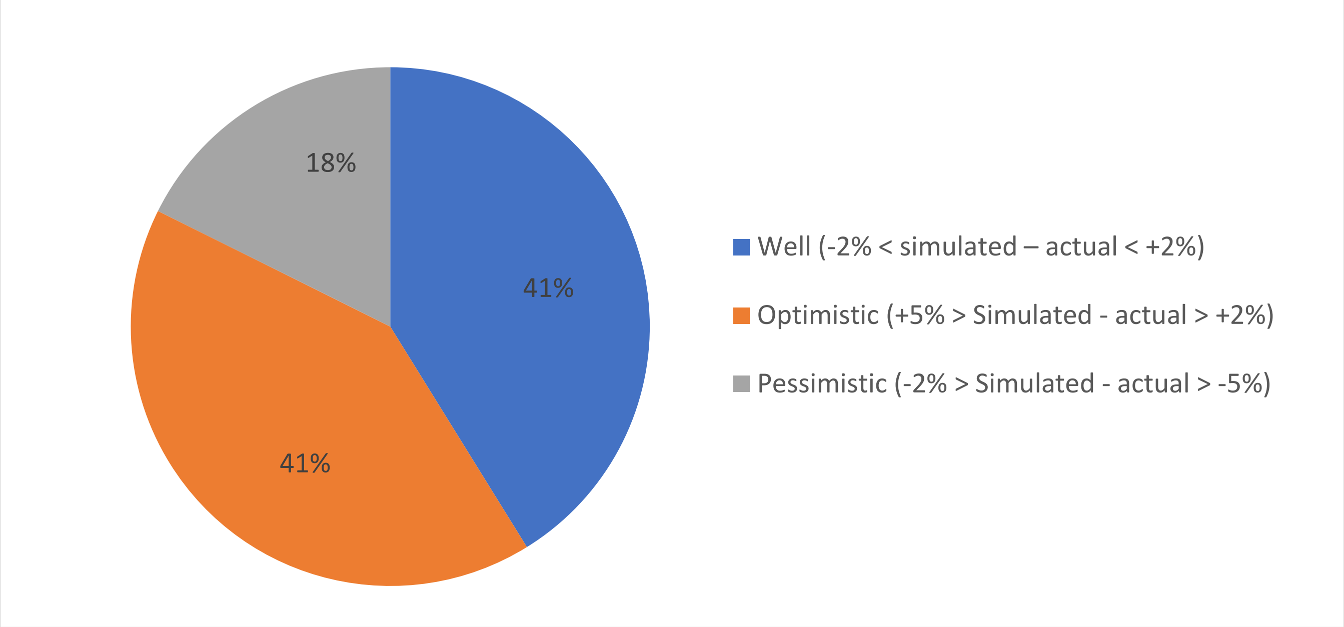


Figure 2.33: Simulated vs actual yields for projects where soiling is significant

For the month(s) with the largest deviation, the simulation tends to overestimate the yield.

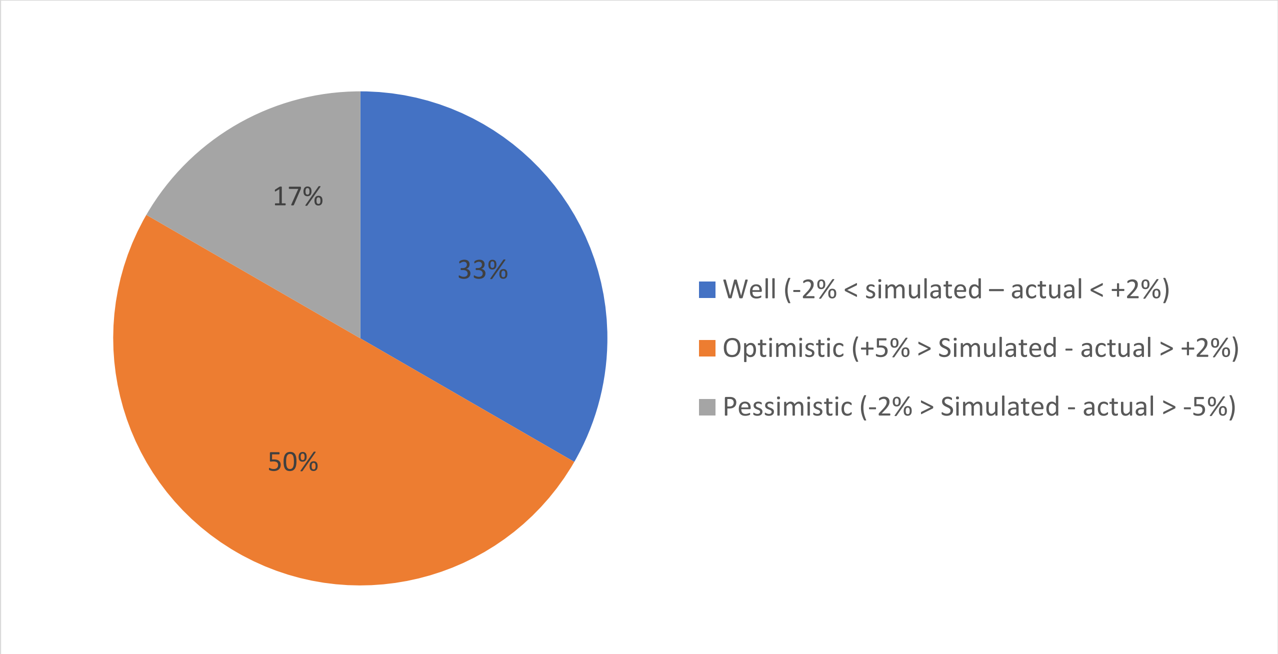


Figure 2.34: Simulated vs actual yields for projects where soiling is significant (for the month with the largest deviation)

2.2.4.4 Snow losses

The following graph shows the share of projects for which snow losses are perceived as crucial. The reading of this graph is similar to the one above concerning soiling losses. For almost 70% of the respondents, the share of projects for which snow losses are important is less than 10%.

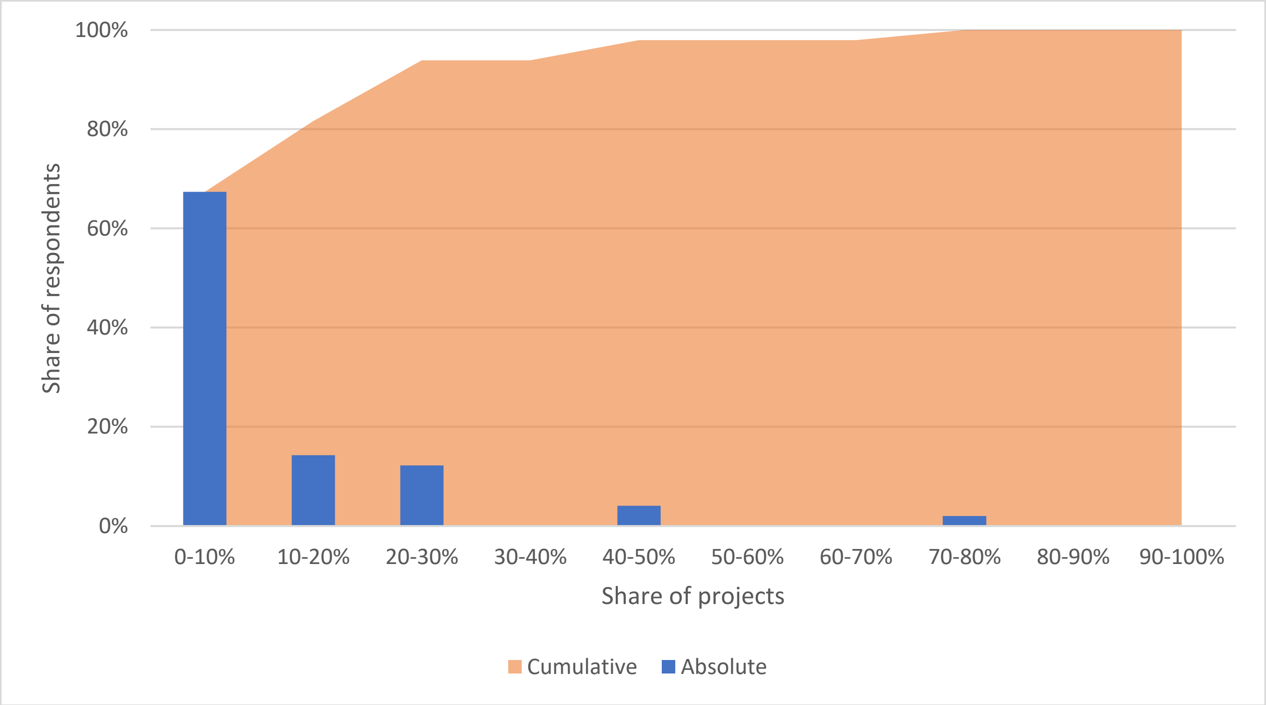


Figure 2.35: Share of projects where snow losses are crucial

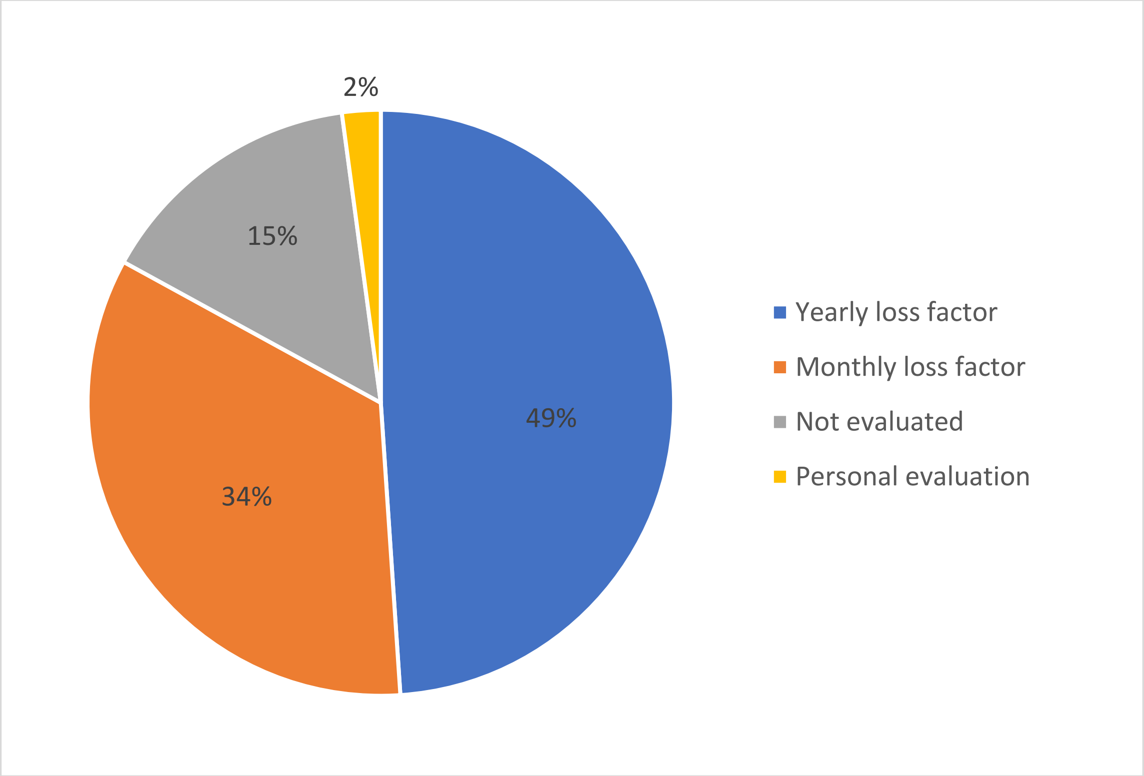


Figure 2.36: How snow losses are considered

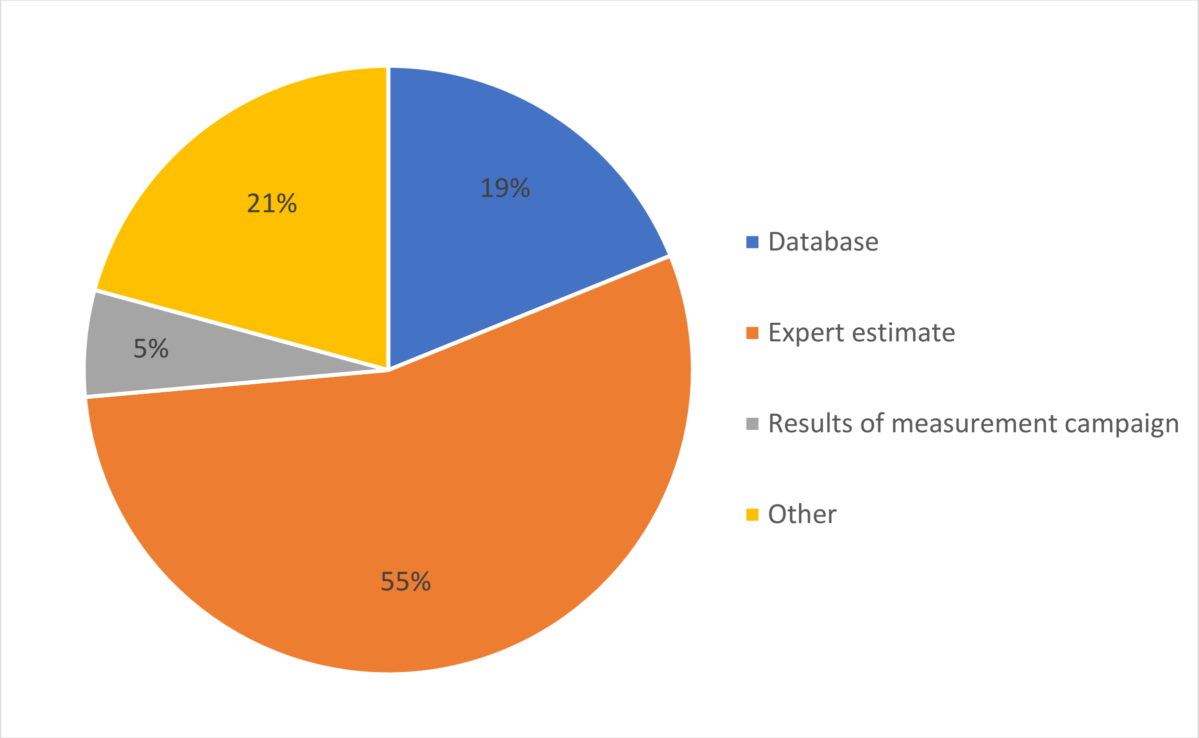


Figure 2.37: How snow losses are evaluated

Other:

- Estimate (5%), which is the same as “expert estimate” category
- Internal model (2%)
- Literature (2%)
- N/A (9%)

The databases used are (in descending order):

- Solargis
- Météo France
- ERA5
- Meteonorm
- SODA MERRA

Half of the respondents do consider local weather conditions, and for those who specified how (15), gave the following details:

- Probability / frequency of snow precipitation
- Days of snow per year
- Use of database and historic data

The characteristics of the installation is also considered by a few of them (x 2): tilt, snow guard.

Only 24% of the respondents do consider snow cleaning operations in the simulation.

Out of 49, 42 respondents do not know how simulated and actual yields compare for projects where snow is significant. The remaining ones find their simulation rather optimistic or in alignment with the reality.

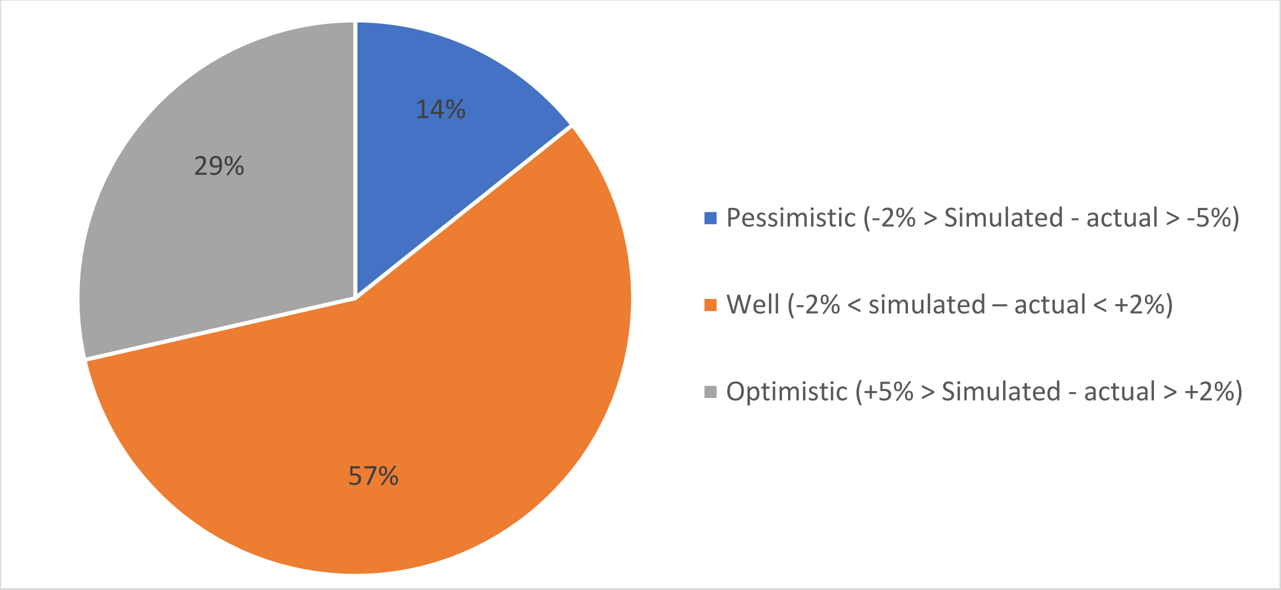


Figure 2.38: Simulated vs actual yields for projects where snow is significant

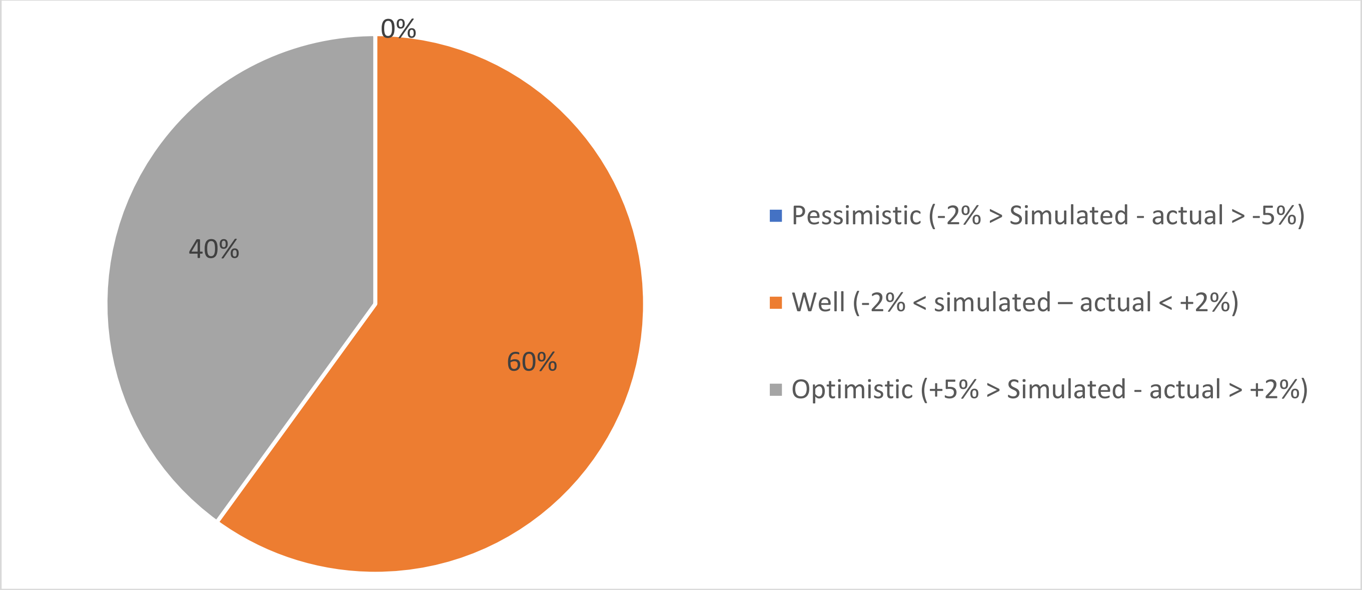


Figure 2.39: Simulated vs actual yields for projects where snow is significant (for the month with the largest deviation)

2.2.5 Section 5: New technologies

2.2.5.1 Bifacial

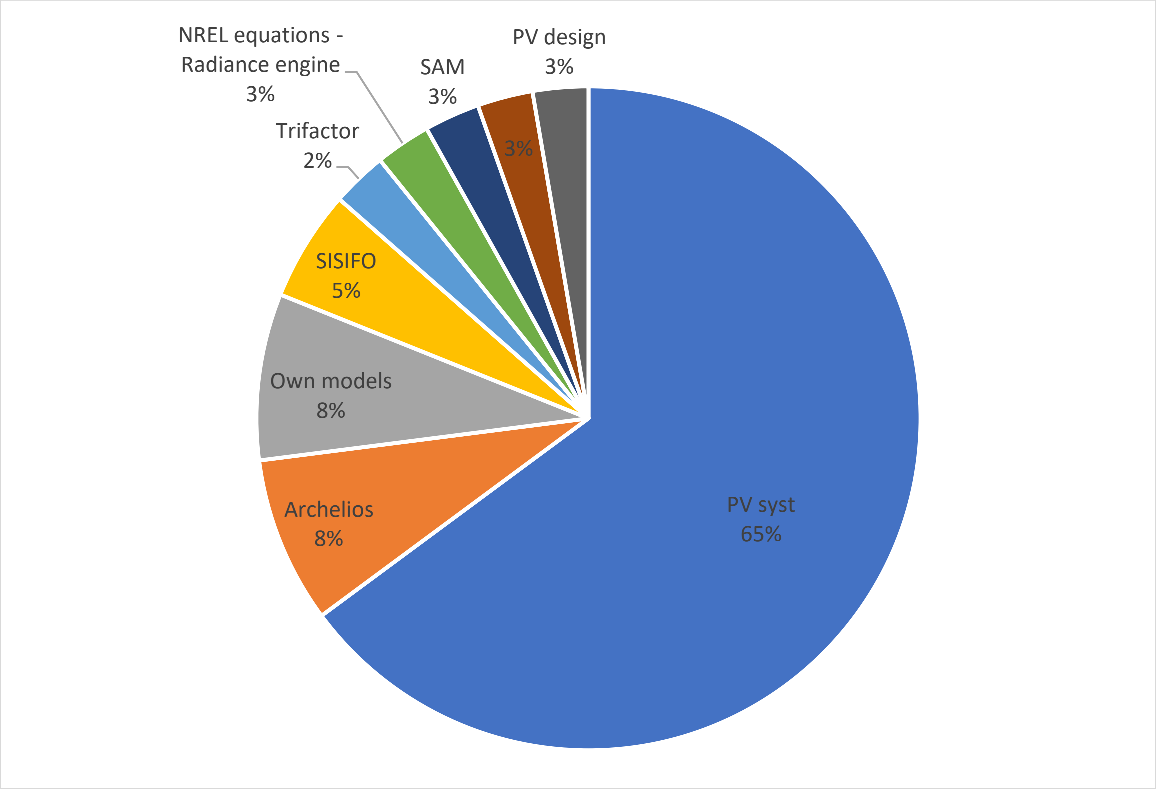


Figure 2.40: Tools used for bifacial simulation – 37 responses

Major shortfalls for PVSYST mentioned by users are:

- The limits of the model. Users are expecting more precise calculation of the rear side irradiance: ‘no edge effect”, “shading on rear side”, “irradiance inhomogeneity”, “no 3D simulation”, “backtracking 3d”.
- The unknown uncertainties on backside irradiance modelling and albedo

20 respondents do have specific needs for bifacial projects simulation, and 19 respondents specified their needs.

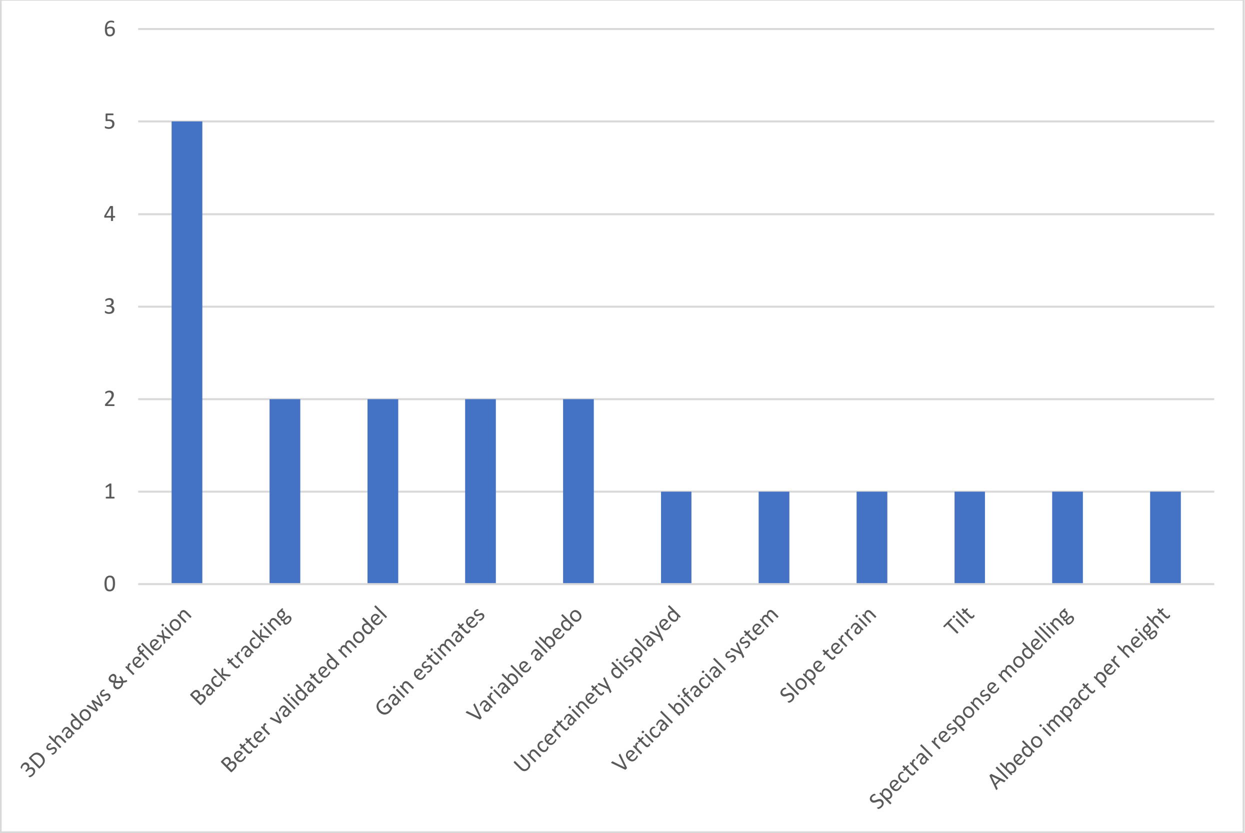


Figure 2.41: Specified needs for bifacial PV simulation

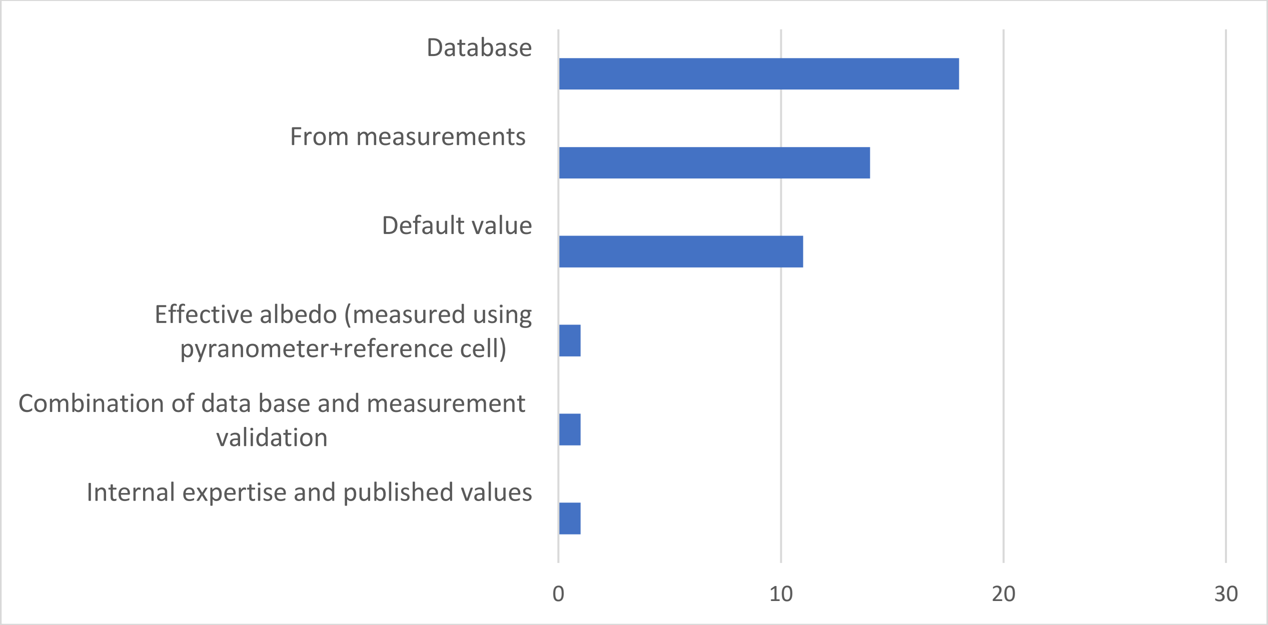


Figure 2.42: Estimation of ground albedo value

15 respondents specified the databases they use for the estimation of albedo:

- Solargis: 13
- NASA/MODIS: 2

75% of 32 respondents apply monthly albedo values, and the remaining 25% are using annual value. No one is using timeseries.

In comparison with standard PV, the respondents identified the following additional parameters.

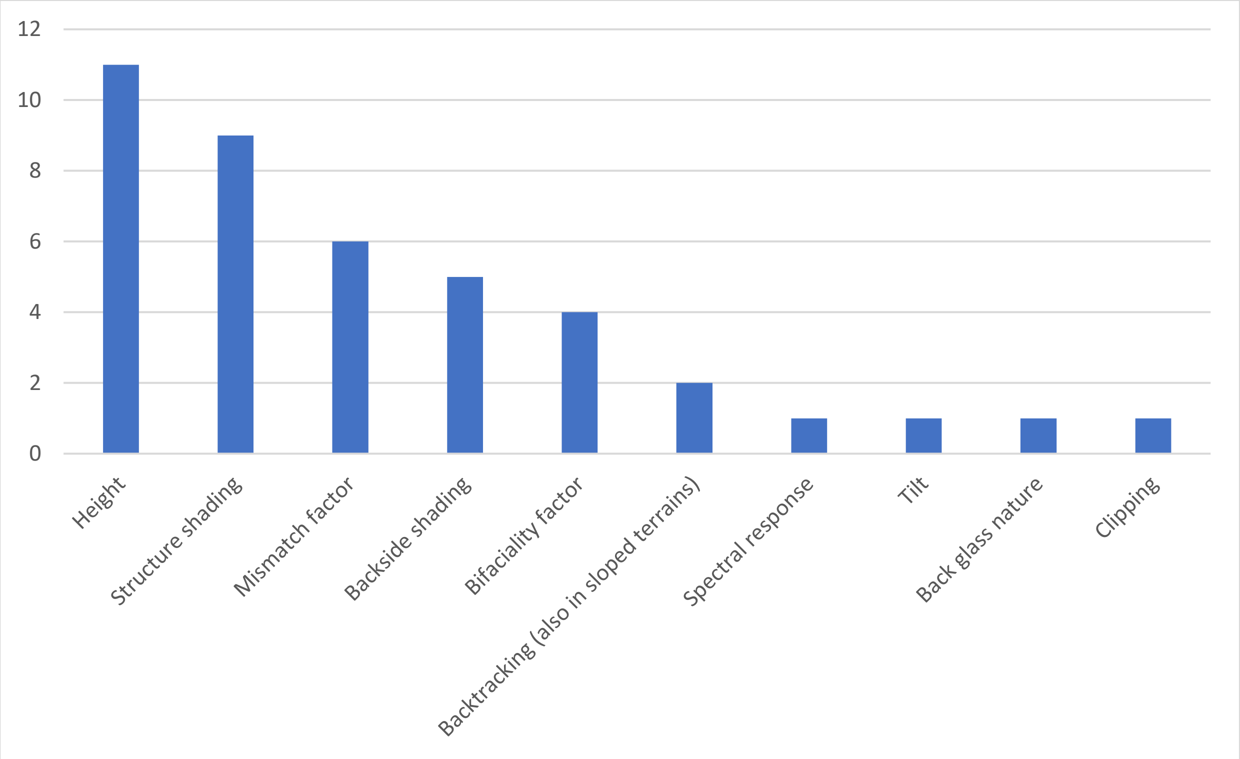


Figure 2.43: Additional parameters required for bifacial in comparison with standard PV

The influence of snow and rain on the albedo is often not considered, and if so, done through monthly averages (empirical estimation).

21 respondents out of 32, do not know how simulated and actual yields compare for bifacial projects. The remaining ones find their simulation rather optimistic or in alignment with the reality.

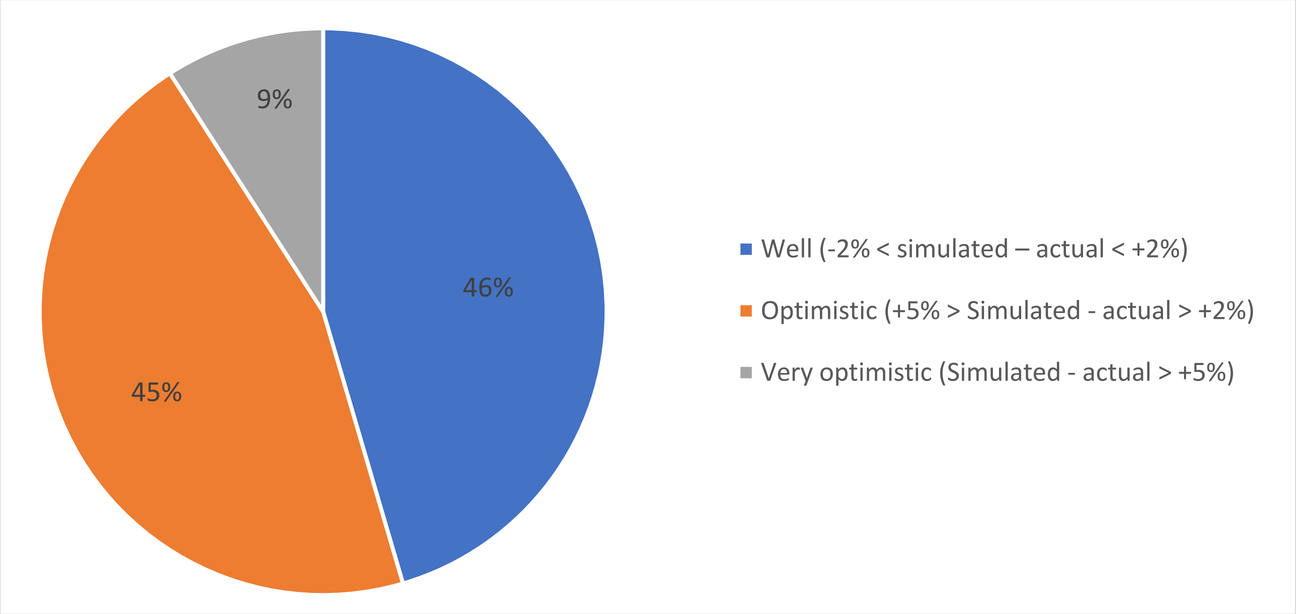


Figure 2.44: Simulated vs actual yields for bifacial projects

2.2.5.2 Floating

32% of the respondents are performing simulations of floating PV systems.

Again, PVSYST is the most used software for this kind of projects. The major identified shortfalls are related to:

- Modelling of module temperature: “water and air temperature not well considered”, “thermal constants”, “mounting technology”,
- Water albedo

The specific needs of the simulation of floating PV, according to 8 answers, are related to:

- the thermal model, as mentioned above among the shortfalls,
- the need for specific values to adjust the parameters of existing models,
- the mismatch due to waves

In comparison with standard PV, the respondents identified the following additional parameters.

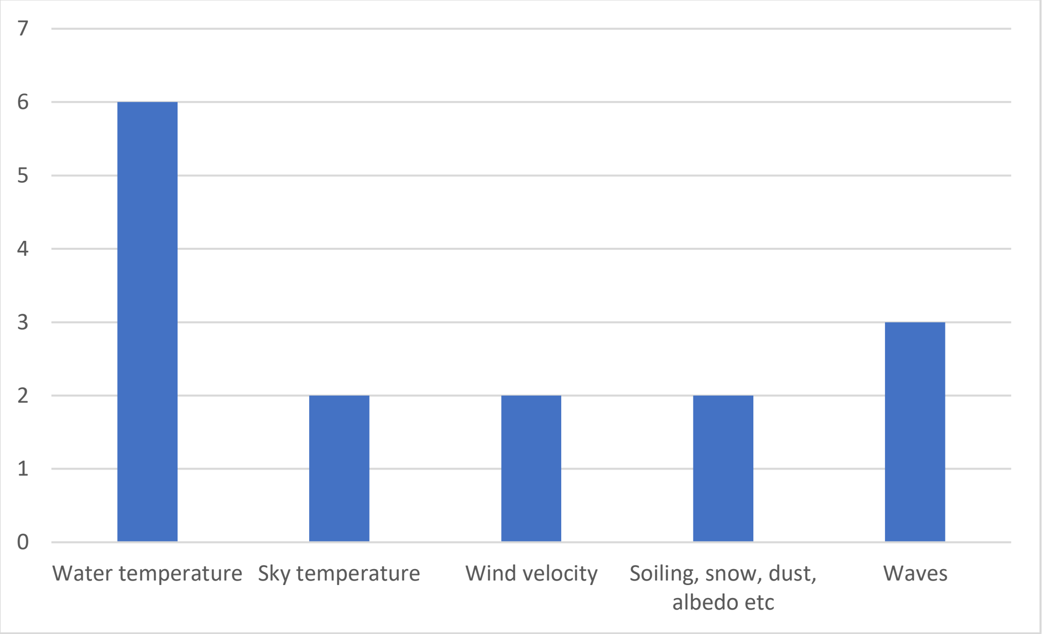


Figure 2.45: Additional parameters required for floating PV in comparison with standard PV

2.2.5.3 BIPV

28% of the respondents are performing simulations of BIPV systems.

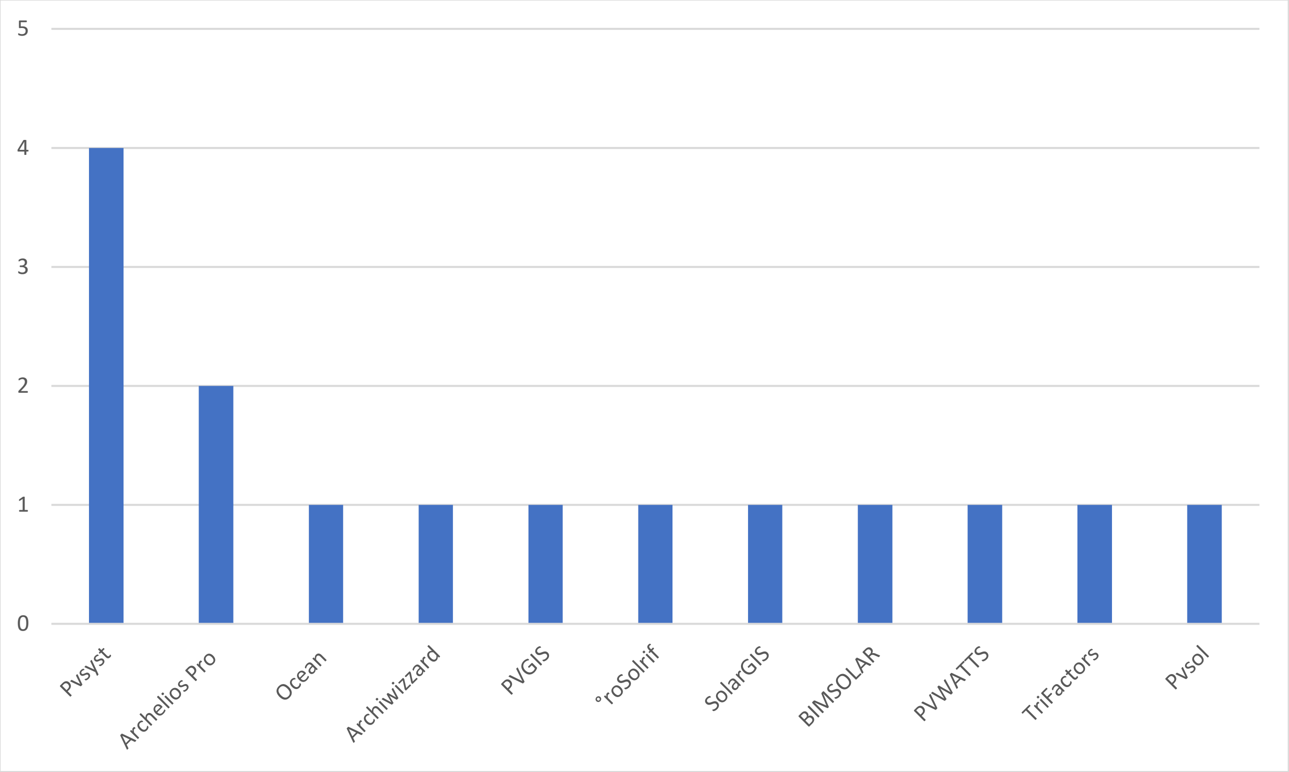


Figure 2.46: Software used for BIPV simulation

3 respondents (over 14) have the necessity to combine the simulation with another software.

4 respondents are using BIM format files for the simulations.

Responding to the question asking for the additional parameters required for the modelling of BIPV, in comparison with standard PV, the respondents (3 answers) gave the issues related to BIPV modelling:

- Need for 3D modelling
- Simulation of shadow and “micro-shadow”
- Thermal aspects
- Pollution
- Reflexion (from nearby buildings)

2.2.6 Section 6: Uncertainties

Most of the respondents (58%) do evaluate exceedance probabilities.

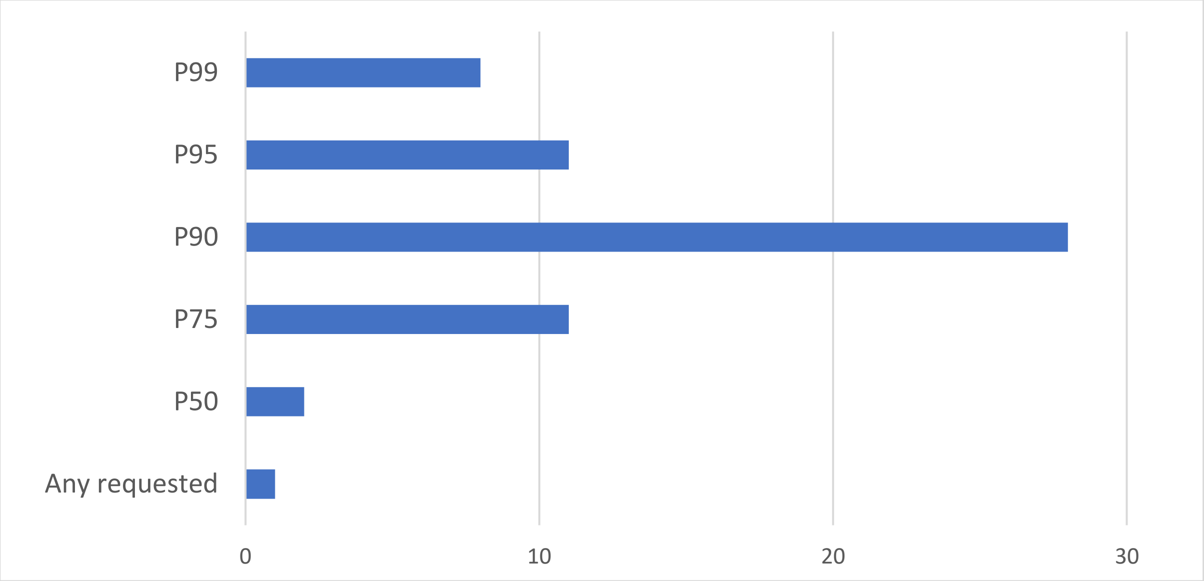


Figure 2.47: PXX evaluated (29 responses)

Only 4% of the respondents use Monte-Carlo methods to evaluate the exceedance probabilities. The vast majority use the quadratic sum of uncertainties and consider only normal distributions.

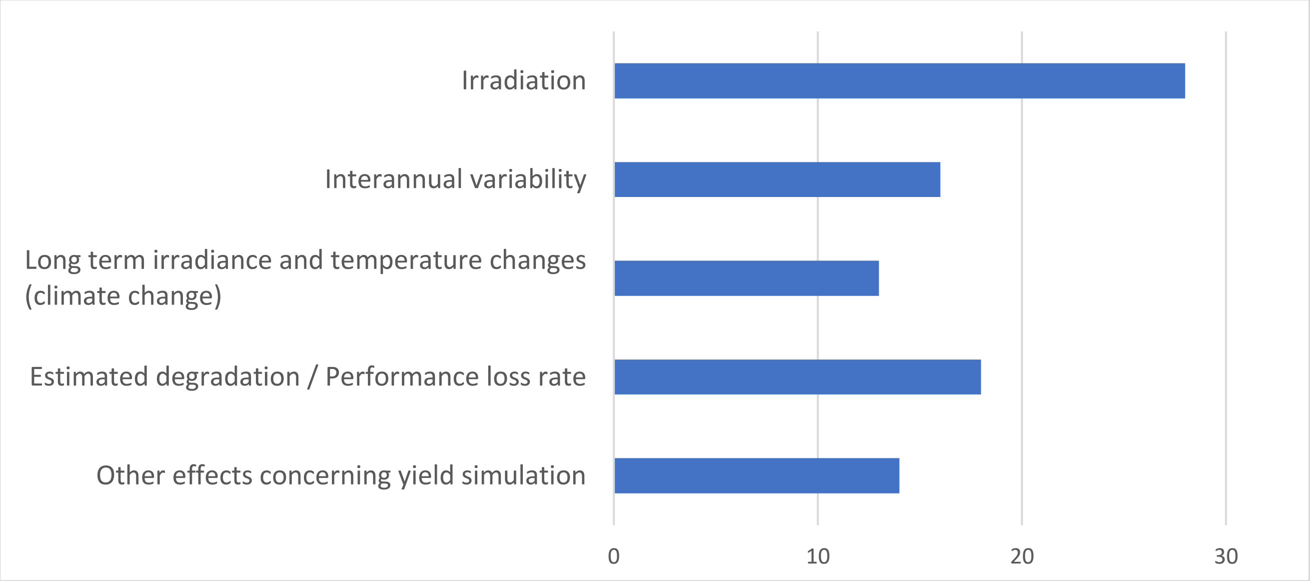


Figure 2.48: Uncertainties considered (29 answers)

The other effects considered are related to

- The equipment (modules, inverters, cables)
- The simulation uncertainties
- The 3D modelling

The uncertainties are mostly determined from literature and experience.

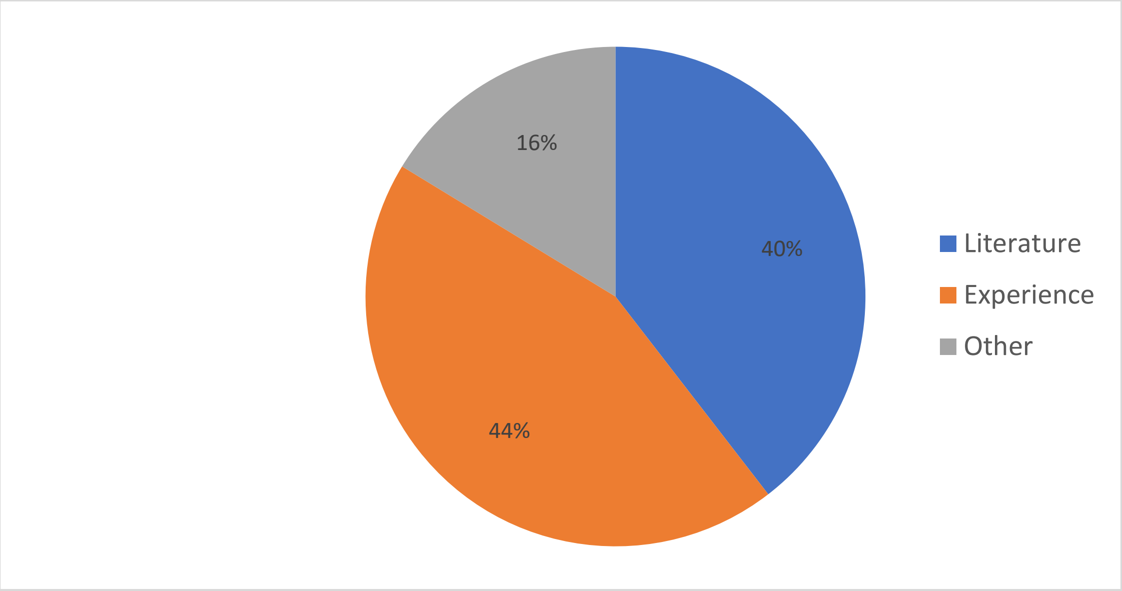


Figure 2.49: Determination of uncertainties (28 responses)

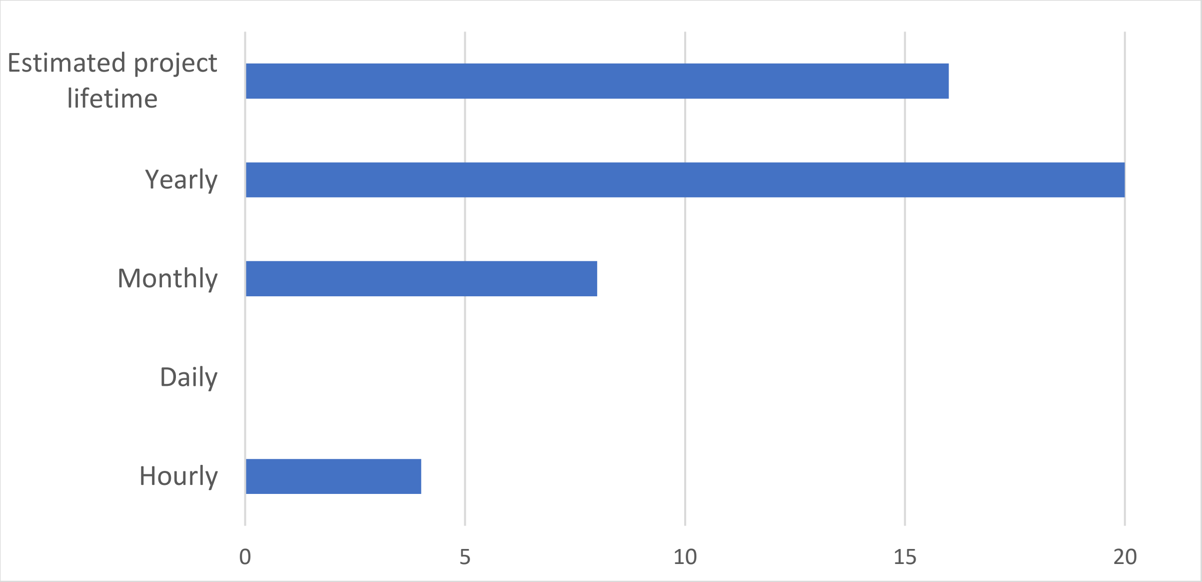


Figure 2.50: Time resolution for the evaluation of exceedance probabilities

The reasons given for the calculation of sub-yearly probabilities are:

- Financial modelling
- Monthly PPA
- Regular assessment
- Analysis of the evolution and reason of the deviation

The expectations concerning this topic are related:

- A better understanding of the models and values for uncertainties: guidelines, explanations
- The uncertainties for bifacial and floating PV

2.2.7 Section 7: Finance

40 respondents answered the questions of this section, meaning that they declared to belong to one of these categories: project designer, EPC, IPP or investor.

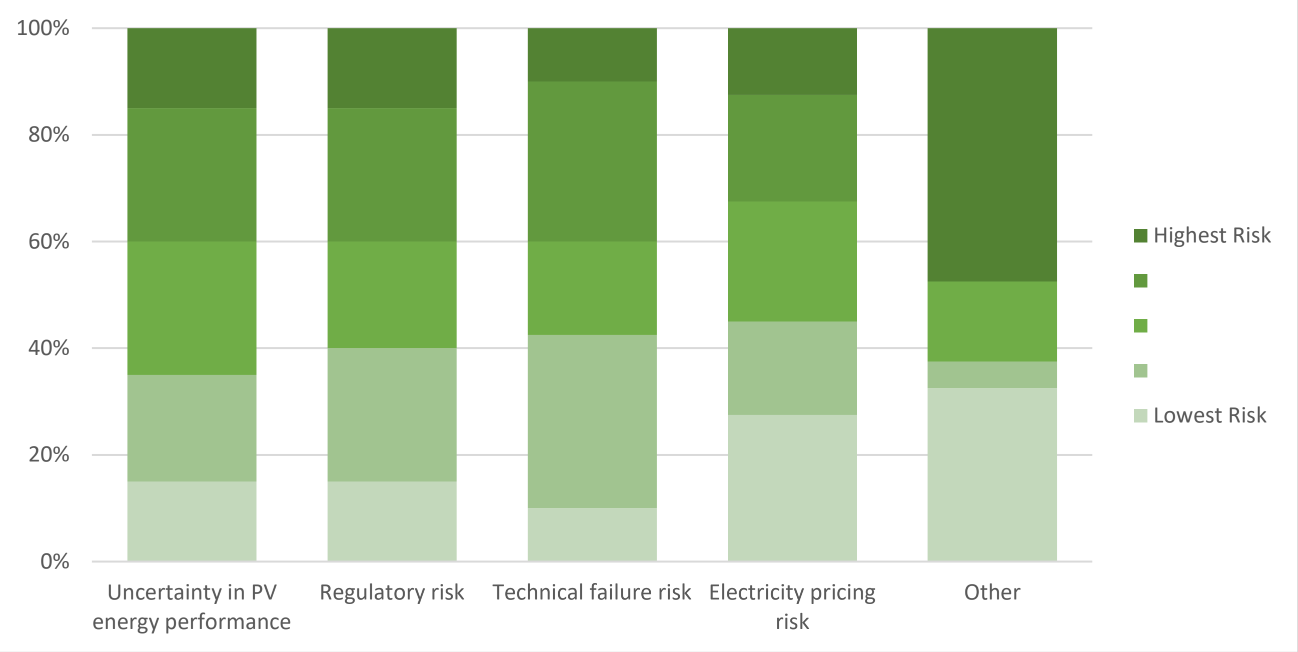


Figure 2.51: Ranking of risks

Detail of “Other” category is found in the table below.

Unfortunately, many of those who ranked “Other” as the highest risk did not give more details.

Table 2.2: Risk ranking for “other”

1 = lowest risk	2	3	4	5 = highest risk
New construction creating shades	Curtailment. Although for merchant plants, market & price risk clearly far higher.	Suppliers bankrupt during the project life Natural disasters	/	Project owner's indecisiveness
Meteorological risks	No response x1	Bankability of new technologies		Environment
Project specific factor		Risk of taking short-cuts in simulation because the software is too complicated		Price of equipment
Supply chain instability		Grid impact		Lack of constitutive material of solar panel
Town planning		No response x1		Grid fees
No response x8				**No response x16**

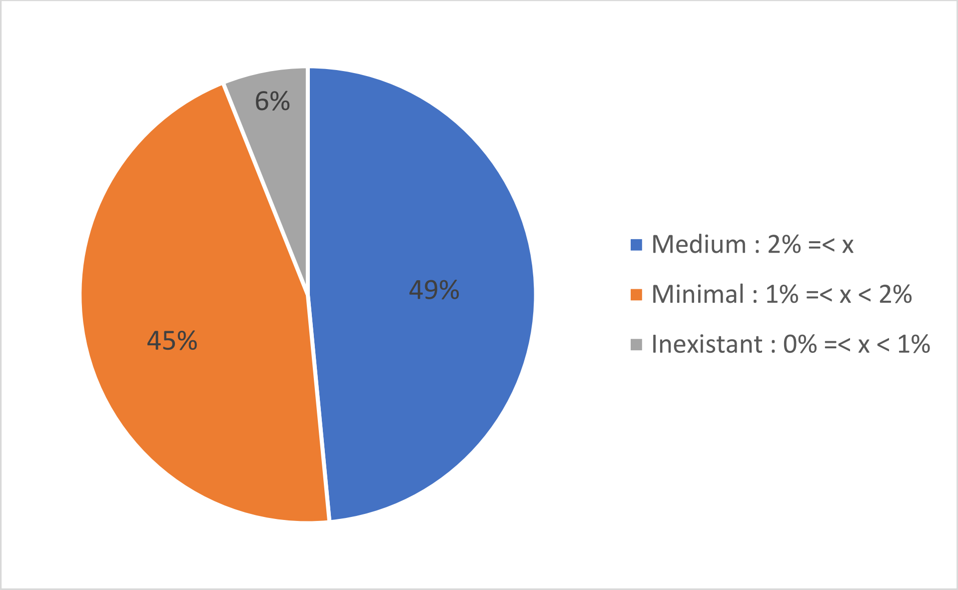


Figure 2.52: Estimation of the risk premium associated to the uncertainty risk

The P50 value is commonly used in the financial models, followed by the P90.

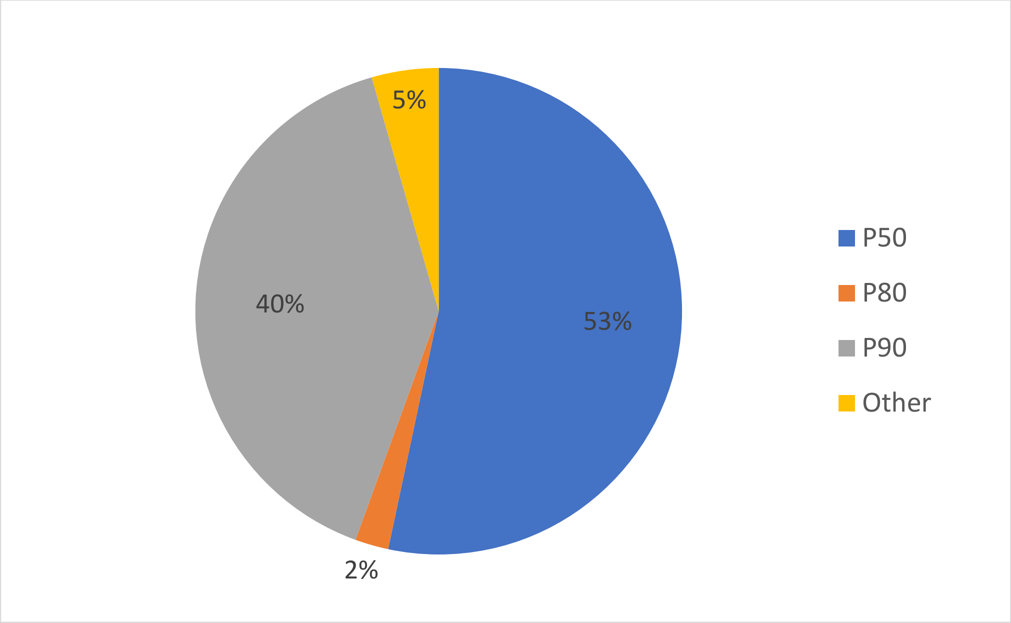


Figure 2.53: Probabilistic estimation of PV yield used in financial models (34 responses)

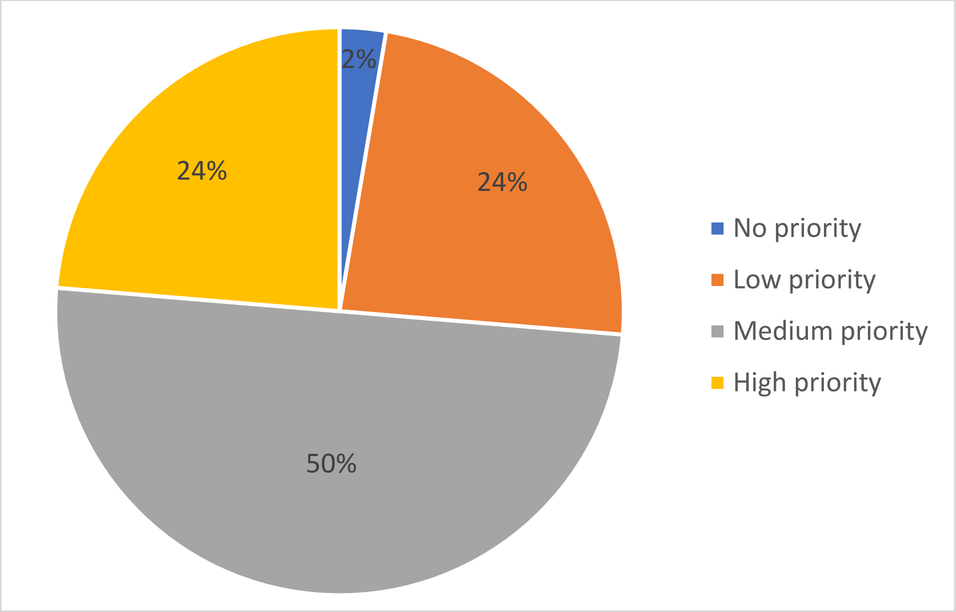


Figure 2.54: Priority level put on the reduction of risk

3 Evaluation of simulation tools and models

3.1 Context – Rationale of the study

The commercial ("black-box") solutions for PV (energy yield) simulations are significantly limited due to exclusion of particularities of new or emerging PV system designs (e.g. bifacial PV, floating PV). Moreover, the influence of certain environmental stressors, as well as different loss mechanisms and/or failure propagation over time is also not taken into consideration. As such, the impact of individual site-, design- and technology- specific parameters on the PV performance and reliability is not possible to neither model/calculate nor assess accurately. Mentioned phenomena limit not only the accuracy of simulation tools but also their resilience towards data-driven PV forecasting and inspections for preventive maintenance. Overall financial yield is therefore lowered.

Recent advances in physics-based models for tailored multi-factor (electrical, optical, thermal) PV simulations are particularly promising to overcome the aforementioned limitations. However, current models are still not compatible for integration into state-of-the-art PV simulation tools. Moreover, most of the models are computationally intensive, thus not suitable for calculating the lifetime performance of PV systems, especially utility-scale ones. Considering, for instance, the case of bifacial PV, to avoid the increasing computational complexity and runtime, most of the existing physics-based models simplify the bifacial PV module's response to ambient conditions by modelling a single, "typical" module within the array and then extrapolate the results to a full-size array. As a result, the impact of mismatch effects caused by spatial variations of bifacial irradiance is not considered, thus inducing significant errors or uncertainties in the estimated (simulated) bifacial PV energy yield.

Addressing described R&D gaps and limitations in energy yield modelling of new PV system technologies, is exactly the principal objective of WP2 in SERENDI-PV. To better position the aimed innovations of WP2, in relation to the state-of-the-art research, industry needs and existing tools for PV simulations, next to the survey of subtask 2.1.2, SERENDI-PV partners performed a "benchmark" study to evaluate state-of-the-art or commercial tools/models for PV energy yield simulations. A similar relevant benchmark study was recently carried out and publicly presented in the context of the IEA PVPS Task 13 collaborative platform, with the participation of some of the SERENDI-PV partners. In that case, the study focused mostly on the climate-dependence and uncertainties of the PV energy yield assessments, as well as their impact on PV LCOE [1,2].

The following section describe the followed methodology, and more specifically:

- The hypothesis and parameters for the simulations, including the meteorological data (§ 3.2.3),
- The performance indicators used to assess the results of simulations (§ 3.2.5), to go a little further than the comparison between measured and simulated yearly yields.

3.2 Methodology

3.2.1 The evaluated PV simulation tools

Through this study, the following 8 PV simulation tools were evaluated: **Archelios PRO (Cythelia Energy), PVSYST, SGIS Evaluate (Solargis), TriFactors (CEA), Zenit (Fraunhofer), LUSim (Lucisun), SISIFO (QPV) and SAM (Solar Advisor Model – NREL)**. These tools comprise either established commercial solutions or (proprietary) in-house software prototypes of the involved partners. In order to minimize uncertainties due to user bias, all partners have defined common parameters and hypotheses/assumptions for their respective simulations. **For confidentiality reasons, software names are anonymised in the chapter** *3.3 Results: Analysis and Discussion***.**

It should be highlighted that the scope of the study is not to compare the aforementioned simulation tools between them only; but also to evaluate (both qualitatively and quantitatively) the "performance" and design limitations of such state-of-the-art tools.

3.2.2 The simulated PV plants

A total number of 7 PV plants were simulated by all partners, for the evaluation of the above PV simulation tools. The selection of these PV plants was elaborated jointly by "simulation tool owners" and "PV plant/data providers" partners in T2.1, to allow:

i) evaluation on diverse PV system sizes, designs, technologies or site characteristics,
ii) evaluation under different climatic profiles,
iii) availability of actual historic data (PV production and meteo data) of one year, selected by the data providers from two criteria:
    i. the a priori availability of the data
    ii. the age of the plant, the youngest possible, to minimise the degradation parameter

Indicatively, the installed capacity of the simulated PV plants ranges from 250 kW_p (for the smaller, commercial-scale one) up to 21 MW_p (for the larger, utility-scale one). Table 3.1 gives an overview of the simulated PV plants.

For reasons of confidentiality regarding certain type of information or data, the **PV plants are anonymized**. Yet, general descriptive information and data of the PV plants' sites, that are relevant to the simulation and evaluation results are still provided.

Table 3.1: Overview of the simulated PV plants

	Type / technology	Climate profile
"PV Plant 1"	Monofacial, Fixed tilt	Warm temperate / Mediterranean
"PV Plant 2"	Bifacial, Fixed tilt	Inter-tropical zone, tropical/oceanic
"PV Plant 3"	Monofacial, 1-axis tracker	Warm temperate / Mediterranean
"PV Plant 4"	Monofacial, 1-axis tracker	Highly arid, warm (desert)
"PV Plant 5"	Monofacial, Roof-mounted	Warm temperate / Mediterranean
"PV Plant 6"	Monofacial, BIPV	Warm temperate / Mediterranean
"PV Plant 7"	Monofacial, Floating PV	Warm temperate / Mediterranean / water reservoir microclimate

3.2.3 Hypotheses / Assumptions

3.2.3.1 Solar and meteorological data, albedo

For all the studied PV plants, ground measurements of meteorological parameters, including solar irradiation, are available. The primary intention was to use these measurements as inputs for the simulation, but this idea was abandoned for the following reasons:

- For some plants, information about sensors were not available, like the type of instrument, the calibration and recalibration dates, cleaning/maintenance logs, etc.
- Quality Control (QC) of the measurements showed many inconsistencies (see § 3.2.4)

Moreover, a specific effort would have been necessary to convert each measurements dataset into a unique format, to allow simple imports in the different software.

For the above reasons, the use of satellite data seemed to be more practicable, therefore solar resource and meteorological data were supplied from Solargis database. This is a high-resolution global database of solar resource and meteorological parameters, operated by the company of the same name. Its geographical extent covers most of the land surface between latitudes 60º North and 55º South.

Solar radiation is calculated by numerical models, which are parameterized by a set of inputs characterizing the cloud transmittance, state of the atmosphere and terrain conditions. In the Solargis approach, solar irradiance is calculated in 5 steps:

1. Calculation of clear-sky irradiance, assuming all atmospheric effects except clouds,
2. Calculation of cloud properties from satellite data,
3. Integration of clear-sky irradiance and cloud effects and calculation of global horizontal irradiance (GHI),
4. Calculation of direct normal irradiance (DNI) from GHI and clear-sky irradiance,
5. Calculation of global tilted irradiance (GTI) from GHI and DNI.

The calculation procedure also included terrain disaggregation model for enhancing spatial representation – from the satellite resolution to the resolution of digital terrain model. As a final result, irradiation parameters are provided in original 15-minute time step, 250 m spatial resolution and time representation depending on a region (starting in years 1994/1999/2007, according to available satellite data history). Comprehensive overview of the Solargis model was made available in several publications [3, 4, 5]. The related uncertainty and requirements for bankability are discussed in [6, 7, 8].

Meteorological parameters are an important part of a solar energy project assessment as they determine the operating conditions and the effectiveness of solar power plant operations. Meteorological data can be collected by two approaches:

1. By measuring at meteorological sites, and
2. Computing by meteorological models.

In Solargis, the **meteorological data** is derived from the meteorological models available for the region. Several models are available, but good option is to use ERA5 European Atmospheric Reanalysis (source ECMWF) [9]. The original spatial resolution of the models (25 km) is enhanced to 1 km for air temperature and air pressure by spatial disaggregation and use of the Digital Elevation Model SRTM-3. The spatial resolution of other parameters is unchanged. Original time resolution is 1 hour.

Surface **albedo**, another important parameter for photovoltaic production modelling, mostly for bifacial installations, is also available in Solargis database. Ground albedo parameter is derived from the MODerate-resolution Imaging Spectroradiometer (MODIS) albedo data product, version 6 (MCD43A3) [10, 11, 12]. Daily value represents temporally weighted average of data from 16 days long window. The original MODIS data is available in 1 to 2-day frequency. The spatial resolution is 0.5 km and the temporal resolution is 1 day.

An overview of the provided data parameters is given in Table 3.2.

Table 3.2: List of the provided solar and meteo data parameters

Parameter type	Parameter name	Time resolution
Solar	GHI – Global horizontal irradiation	• Time series in original 15-minute step • Time series in hourly aggregation
	DNI – Direct normal irradiation	
	DIF – Diffuse irradiation	
Meteo	TEMP – Air temperature	
	WS – Wind speed at 10m	
	WD – Wind direction at 10m	
	RH – Relative humidity	
	AP – Atmospheric pressure	
	PWAT – Precipitable water	
Albedo	ALB – Ground albedo	• Monthly long-term averages

The fundamental difference between a satellite observation and a ground measurement is that signal received by the satellite radiometer integrates an area (a footprint of visible and infrared channels represents an area of several square kilometres) while a ground station represents a pinpoint measurement. This results in a mismatch when comparing instantaneous values from these two observation instruments, mainly during intermittent cloudy weather and changing aerosol load.

A solution is to correlate satellite-derived data with ground measurements to understand the source of discrepancy and subsequently to reduce the uncertainty of the resulting historical time series. After correlation, the site adaptation of the model is applied with an aim to remove general trends of disagreement between the measurements and the model data. Important is to avoid matching of data representing the extreme cases, such as dust storms or volcano ash outbreaks, which do not represent prevailing conditions at a site. This principle also mitigates propagation of short-term issue in the ground measurements into the site adaptation results.

Therefore, the site adaptation focuses on seasonal trends. At the monthly level, some disagreements between the measured and site-adapted data may exist. To achieve reasonable results, high-quality ground measurements should be available for a period of about one year, so that all seasons are included. In case of a tight time schedule, a shorter period may be considered for on-site measurements. However, such data may not be capable to cover all deviations. In optimal case, two years of data provide more robust results and allow decreasing uncertainty of resulting site-adapted data.

Prior to the comparison with satellite-based solar resource data, the ground-measured irradiance has to be quality-controlled. Quality control (QC) is based on methods defined in SERI QC procedures, Younes et al. [13, 14] and implemented in-house by Solargis. The tests are applied in two runs: (i) first, the automatic tests are run to identify the obvious issues; next (ii) by the visual inspection we identify and flag inconsistencies, which are of more complex nature. Visual inspection is an iterative and time-consuming process.

The automatic QC tests include:

- Identification of missing values
- Correction of time shifts
- Evaluation of measurements against sun position (Sun below and above horizon)
- Comparing the data with possible minimum and maximum irradiance limits
- Evaluation of consistency of GHI, DIF and DNI by comparing the redundant measurements (if available).

The visual quality control aims to identify and flag the following erroneous patterns:

- Shading from nearby objects (near shading) or mountains (far shading)
- Regular data error patterns
- Irregular anomalies
- Comparison of measurements from different instruments (if available).

For purposes of the quality control and site adaptation of the satellite model for selected sites, several datasets measured by ground meteorological stations were provided by partners. After quality control, some of the measured datasets were rejected due to issues which degraded the quality of measured data to level, where site-adaptation of the satellite model will not improve accuracy of satellite data. Therefore, only dataset for "PV Plant 4" was used for site adaptation and delivery of adapted time series. For this plant, simulations were performed using both standard and adapted dataset.

3.2.3.2 Simulation parameters

As stated above, to minimize uncertainties due to user bias, all partners have defined common parameters and hypotheses/assumptions for their respective simulations.

The first step was to understand for each software the different modelling steps (or losses) being considered by each software (see Table 3.4 below).

As PVSYST is well-known by the different partners, the simulation parameters for each plant were discussed and chosen for this software, at first. The parameters for the other software were then defined considering the PVSYST parametrization as a reference.

In Table 3.3 below, the rationale behind each parameter is given.

Table 3.3: Simulation parameters rationale

Modelling steps / losses	Info from plant owner	Explanation
Transposition model		As it is the only model available in all software, the Perez model was chosen for all the plants.
Albedo	No	For each plant, monthly albedo values were determined from Solargis database.
Soling losses	No	Measurement of soiling are not available. Annual loss factor was estimated for each plant based on local microclimatic conditions.
Spectral correction		Not considered in simulations.
Module quality / Tolerance	Yes	From module datasheet. Quarter of the difference between min and max values.
LID losses	No	2% default value for all plants (p-type silicon modules).
Module mismatch	For one plant only	If not already estimated by the plant owner, default values of 0,5% or 1% depending on the age of the plant.
Module ventilation	No	Default value except for roof-integrated (less ventilation) and floating (higher ventilation, to reflect the a priori lower ambient temperature) systems.

Modelling steps / losses	Info from plant owner	Explanation
Annual degradation factor	No	Default value of 0,5%/year.
Bifacial: shed transparency	No	Conservative value: 0%.
Bifacial: non uniformity of rear irradiance		Default value: 10%.
Bifacial: shadow from structure	No	Default value: 15%. As the bifacial are installed on greenhouses, this value is high to consider the optical losses between the rows.
DC cables losses	For some plants	If not available, 1% at STC, which is the recommended value in the countries were the plants are located.
AC cables losses	For some plants	If not available, 1% at STC, which is the recommended value in the countries were the plants are located.
Transformer losses	For some plants	Not available. 0,1% for irons losses and 1% for resistive losses.
Auxiliaries	No	Not considered.
Unavailability	No	Not considered. Unavailability is corrected post-simulations based on actual production data.

Table 3.4: Modelling steps / losses consideration

	PVSYST	archelios PRO	Solargis Evaluate	TriFactors (CEA)	Zenit	LUSim	SISIFO	SAM
Transposition model	Perez, Hay	Perez	Perez	Perez, Maxwell, Lam_Li, Louche, Kartveit_Olseth, Reindl, Orgill_Hollands	Perez, Hay, Klucher	Hay, Perez	Perez, Hay	Perez, HDKR, Isotropic
Albedo	Monthly values	Monthly values	Monthly values	Hourly values	Annual or monthly	Monthly	Monthly values	Monthly
Soiling losses	Annual (or monthly) loss factor(s)	Annual loss factor	Monthly values	NA (but can be calculated with loss factor)	Annual or monthly	Monthly	Monthly values	Annual or monthly
Spectral correction	only for a-Si:H optional: First Solar model	NA	First Solar model	NA	Annual or monthly	Hourly	N. Martin	NA
Module quality / Tolerance	loss/gain factor	loss/gain factor	Included in mismatch	Included in mismatch	Included in mismatch	loss/gain factor	loss/gain factor	NA
LID losses	loss factor	loss factor	Included in mismatch	NA (but can be calculated with loss factor)	Not considered as it should be considered in nameplate power according to IEC 61215:2016	loss factor	loss factor	NA
Module mismatch	loss factor *+ detailed calculation of this loss factor*	loss factor	loss factor	NA (but can be calculated with loss factor)	loss factor	Based on I-V curves	loss factor	NA
Module ventilation /	thermal losses factor Uc (+Uv)	back-side of module ventilation factor	thermal coefficient	linear or NOCT model	thermal coefficient	thermal coefficient	thermal coefficient	NOCT / heat transfer

	PVSYST	archelios PRO	Solargis Evaluate	TriFactors (CEA)	Zenit	LUSim	SISIFO	SAM
thermal behaviour								
Annual degradation	annual degradation factor + mismatch	annual degradation factor	annual degradation factor	NA (but can be calculated with degradation factor)	annual degradation factor	annual degradation factor + mismatch	annual degradation factor	annual degradation factor
Bifacial: shed transparency	transmission factor	transmission factor		Hypothesis: no structure		GPU-based 3D view factors		
Bifacial: non uniformity of rear irradiance	mismatch loss factor	mismatch loss factor		Included in calculation of IV curve of modules		Included in I-V curves		
DC and AC cables losses	percentage losses at STC *or detailed calculation from cables sections and length*	percentage losses at STC	percentage losses at STC	NA (but can be calculated with loss factor)	percentage losses at STC	percentage losses at STC *or detailed calculation from cables sections and length*	percentage losses at STC	percentage losses at STC
Auxiliaries	User defined	NA	NA	NA	NA		User defined	NA
Unavailability	Unavailability probability + random (or user-defined) distribution of unavailability periods	loss factor	loss factor	loss factor	loss factor	Unavailability probability	loss factor	loss factor / hourly

3.2.4 Quality Control of production data

Similar to solar or meteorological data, if production datasets are intended for further processing (comparison, performance evaluation, etc.), quality control of provided data is required. Some sequences of quality check may be automatized (detection of missing values, time shifts, consistency, etc.), while others still need visual control by trained operator (finding error patterns, anomalies). To make this process more comfortable, an SDAT (Solargis Data Analyst) software is being developed in Solargis. The software is designed for analysis, importing, exporting, converting data formats, harmonization, visualization, comparison, gap filling and error detection in measured data. Currently it is designed mostly for irradiation data processing, but great part of the functionality is applicable to meteorological measurements as well as to analysis of power production time series.

For purposes of the comparison between real power production and software simulations, datasets with one year of electrical production were provided by partners for evaluated power plants. One exception is Plant 2, where only 4 months of data were available (the plant was commissioned in late 2020). As simulation of partial years is not possible for all evaluated software, it was not found relevant to include this plant in the analysis. However, for the work planned in WP2 on bifacial modelling, the data from this plant will be very useful.

Provided datasets were imported into SDAT software, where time reference, invalid, maximum, minimum or static values were checked together with visual inspection. All non-valid values were flagged and removed from further processing. The identification and removal of these problematic timestamps is important to have the less unbiased comparison between simulated and actual production. If not done, the analysis could be seriously flawed. Finally, sub-hourly data were aggregated and exported to hourly datasets, so that all tested modelling tools were able to work with provided data.

Table 3.5 provides summary results of power production datasets quality control. Figure 3.1 provide example of one year flagged data with indication of issues found during quality control. Figure 3.2 provide several examples of issues found during quality control.

Table 3.5: List of the provided data parameters

Power plant	Data provided (time step, time span)	QC data amount passed [%]
“PV Plant 1”	10-minute, year 2012	87.9
“PV Plant 2”	10-minute, year 2021	Not analysed
“PV Plant 3”	10-minute, year 2016	95.7
“PV Plant 4”	15-minute, year 2018	95.7
“PV Plant 5”	15-minute, year 2019	88.6
“PV Plant 6”	10-minute, year 2014	92.7
“PV Plant 7”	10-minute, year 2020	73.3

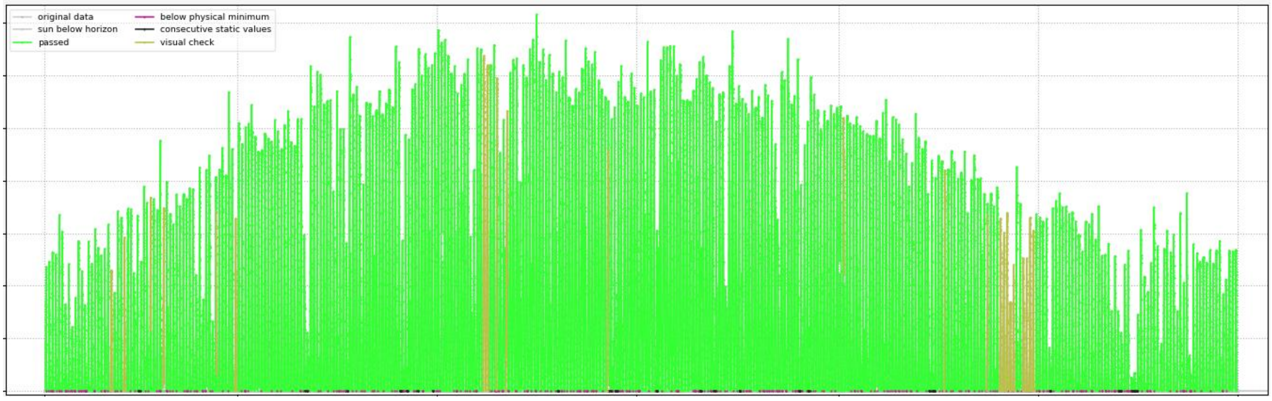


Figure 3.1: Example of Quality Control results
(green flags data passed QC, other colours data with issues)

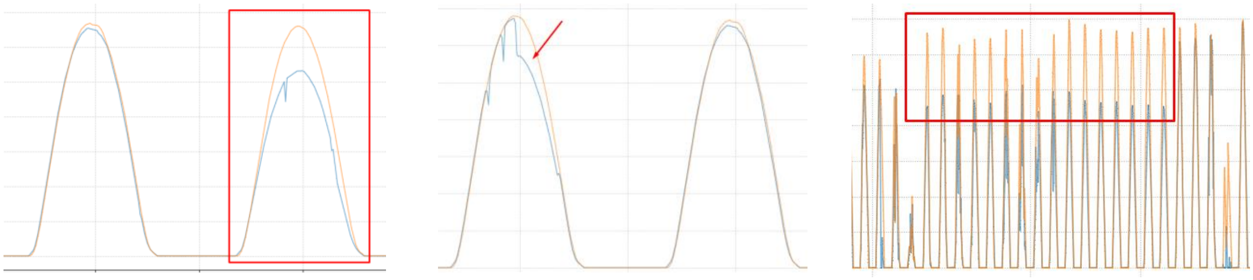


Figure 3.2: Example of Quality Control
Partial production - issues in inverters or string sections
(orange line reference data, blue line real production)

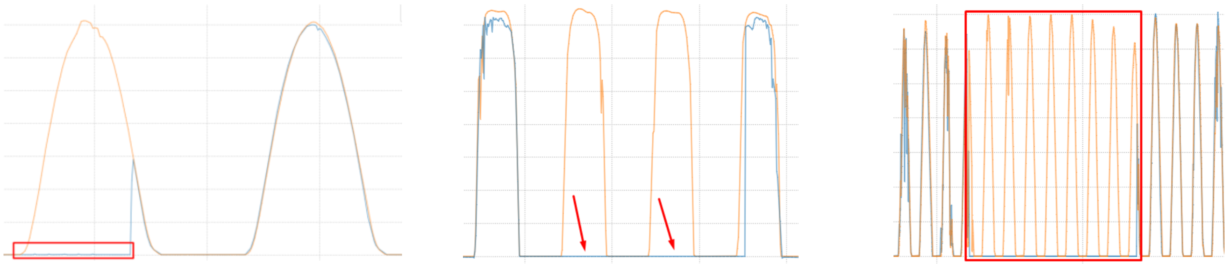


Figure 3.3: Example of Quality Control
Complete outage of production
(orange line reference data, blue line real production)

3.2.5 Key Performance Indicators (KPIs)

The following KPIs are computed, taking into account only non-zero values.

The subscript “actual” refers to actual measured power or energy.

Table 3.6: Key Performance Indicators

KPI	Acronym	Formula
Relative difference		$\frac{Yield_{simulated,year}}{Yield_{actual,year}} - 1$
Mean Bias Error	MBE	$\frac{1}{n}\sum_{i}\left(P_{simulated,i} - P_{actual,i}\right)$
Root Mean Square Error	RMSE	$\sqrt{\frac{1}{n}\sum_{i}\left(P_{simulated,i} - P_{actual,i}\right)^{2}}$
Normalised Mean Bias Error	NMBE	$MBE \Big/ \frac{1}{n}\sum_{i} P_{actual,i}$
Normalised Root Mean Square Error	NRMSE	$RMSE \Big/ \frac{1}{n}\sum_{i} P_{actual,i}$
Mean Bias Weighted Error	MBWE	$\frac{1}{n}\sum_{i}\frac{\left(P_{simulated,i} - P_{actual,i}\right) * P_{actual,i}}{\frac{1}{n}\sum_{i} P_{actual,i}}$
Root Mean Squared Weighted Error	RMSWE	$\sqrt{\frac{1}{n}\sum_{i}\frac{\left(P_{simulated,i} - P_{actual,i}\right)^{2} * P_{actual,i}}{\frac{1}{n}\sum_{i} P_{actual,i}}}$
Normalised Mean Bias Weighted Error	NMBWE	$MBWE \Big/ \frac{1}{n}\sum_{i} P_{actual,i}$
Normalised Root Mean Squared Weighted Error	NRMSWE	$RMSWE \Big/ \frac{1}{n}\sum_{i} P_{actual,i}$

With n = the number of non-zero values

Except for the first one, the indicators are calculated on hourly and daily values.

3.2.6 Limitations

The purpose of this work is to assess the accuracy of the modelling, and to do so, and as stated above, the inputs of the models must reflect the reality as close as possible. Despite the efforts presented above, there are several parameters or hypotheses for which the uncertainty could not be reduced or assessed, starting with the plants ‘characteristics which are not perfectly known. For instance:

- The actual nominal power is based on the modules’ datasheets, and not the actual flash-tests; the degradation rate of the modules is not based on measurements; same for the LID losses,
- Efficiency curve of the inverters was not available for all plants,

- Assumptions have been made concerning the electrical wiring of the modules was not known,
- The DC and AC cables losses at STC,
- Etc.

As for the other simulation parameters:

- Use of satellite data instead of ground measurement. The comparison will encompass the discrepancies of both irradiation/meteo and PV system modelling,
- Soiling losses measurements were not available,
- Same as for albedo, even if its influence is limited.
- Periods of unavailability are not fully known for all plants. Nevertheless, these periods could have been approached thanks to the Quality Control procedure presented below (§ 3.2.4).

3.3 Results: Analysis and Discussion

The raw comparison between actual production and simulated yield is given below. **For confidentiality reasons, the names of the software are not given and are replaced by code names** (S1, S2, …).

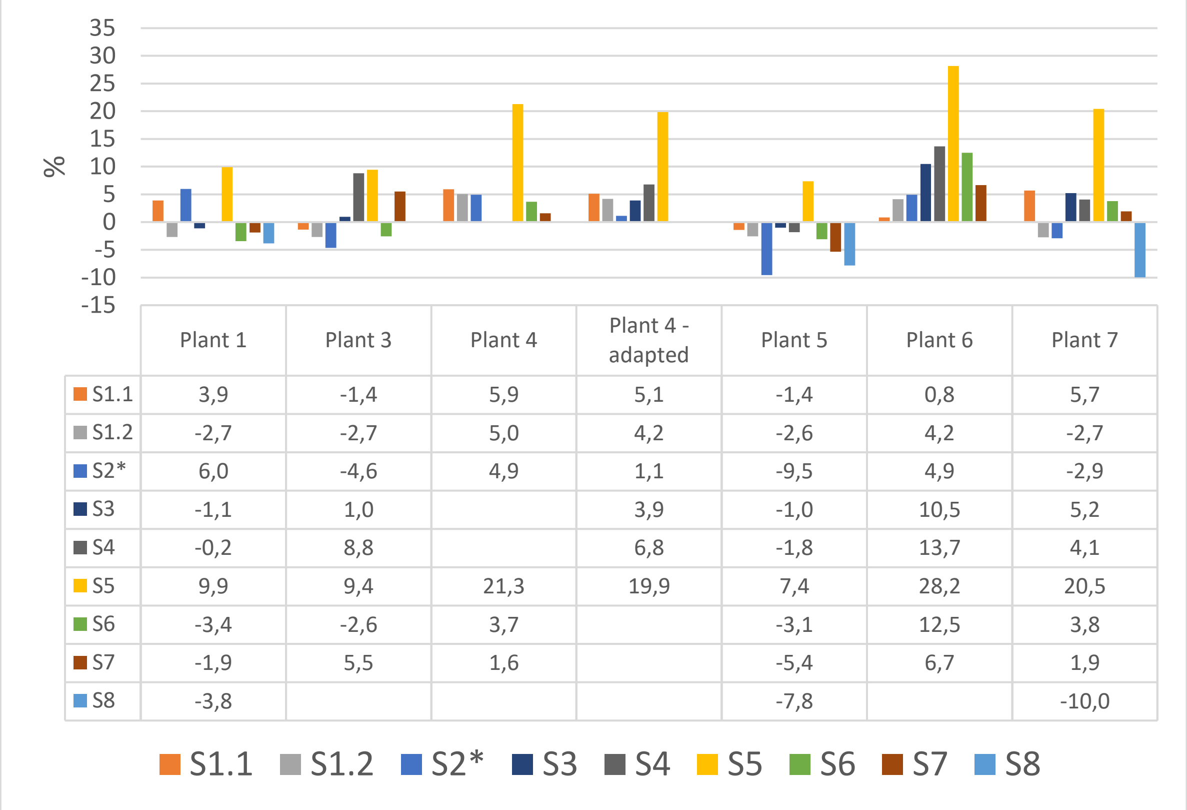
Software 1 has been used by two different partners. Even if the simulation parameters have been set up prior to the simulations, differences are still found. This stresses out the fact that **the "human factor" remains in the simulation process and is a source of uncertainty**. This is an interesting topic which is not addressed in this project, which focuses on the modelling, but may deserve a specific task in a future one.

The commercial version of Software 2 is based on calculation performed on average days for each month. Therefore, the hourly and daily KPIs are not calculated for this software. A pre-alpha version of this software marked "S2*" has been used to compute hourly and daily values.

All simulation tools are used to simulate a "universal" type of PV systems (i.e. both monofacial and bifacial ones), except for the case of Software 5 which is rather optimized for simulating bifacial PV systems, and output of this software is at the moment limited to DC power. This explains the high differences observed on all plants.

Software 8, at the moment when simulations were ran, did not integrate the possibility to automatically execute a full simulation on systems with trackers and bifacial systems. This is why no results are presented for this software for plants 3 and 4. Since then, the GPU part of the tool used for bifacial/tracker irradiation gain evaluations has been matched to the yield estimator part (it is therefore possible at the current moment to fully execute a simulation for all the different systems).

Globally, the yield is overestimated, as unavailability periods are considered in the comparison. After removing these periods and other erroneous data (as mentioned in § 3.2.4), the differences are reduced (Figure 3.4). Figure 5.1, in annex, gives the relative difference between yearly simulated and measured production, considering all the values, before QC.

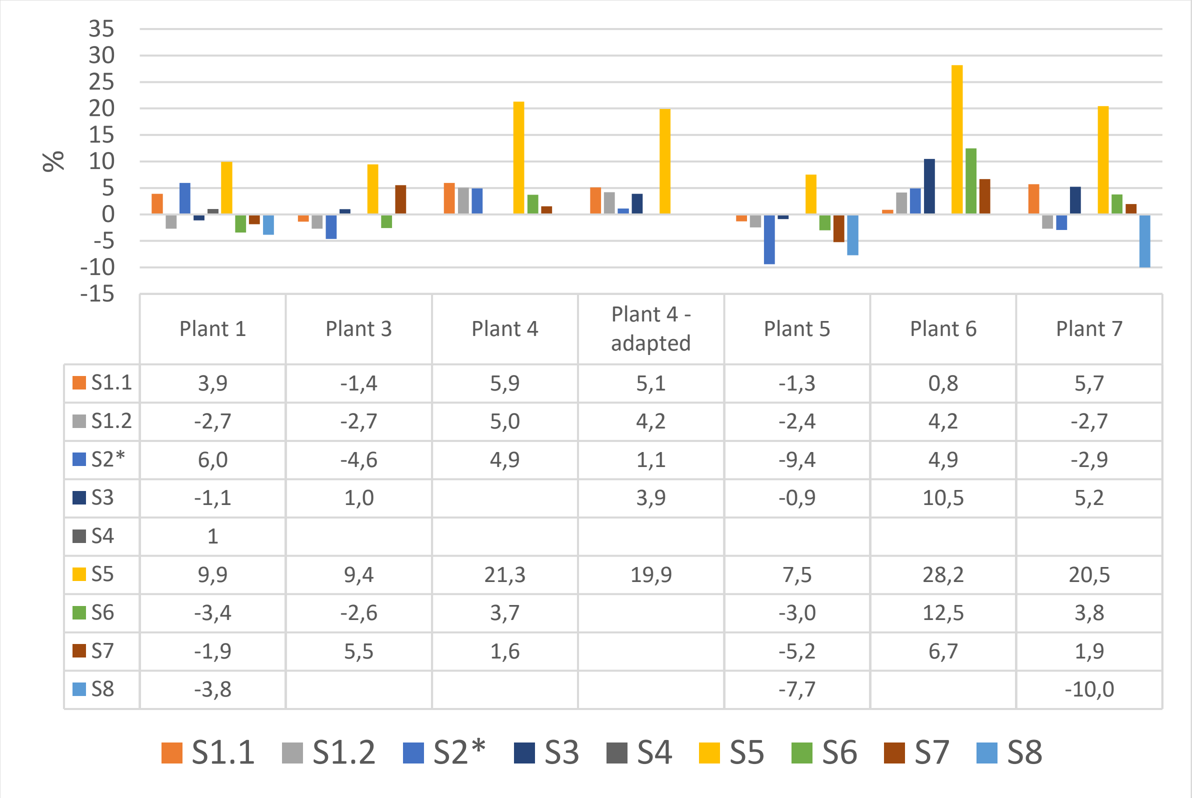


	Plant 1	Plant 3	Plant 4	Plant 4 - adapted	Plant 5	Plant 6	Plant 7
S1.1	3,9	-1,4	5,9	5,1	-1,4	0,8	5,7
S1.2	-2,7	-2,7	5,0	4,2	-2,6	4,2	-2,7
S2*	6,0	-4,6	4,9	1,1	-9,5	4,9	-2,9
S3	-1,1	1,0		3,9	-1,0	10,5	5,2
S4	-0,2	8,8		6,8	-1,8	13,7	4,1
S5	9,9	9,4	21,3	19,9	7,4	28,2	20,5
S6	-3,4	-2,6	3,7		-3,1	12,5	3,8
S7	-1,9	5,5	1,6		-5,4	6,7	1,9
S8	-3,8				-7,8		-10,0

Figure 3.4: Relative difference between simulated and measured production for one year (corrected according to available measured periods)

Plant 4 was simulated with several software using “non-adapted” and adapted satellite data. The use of adapted data always improves the results.

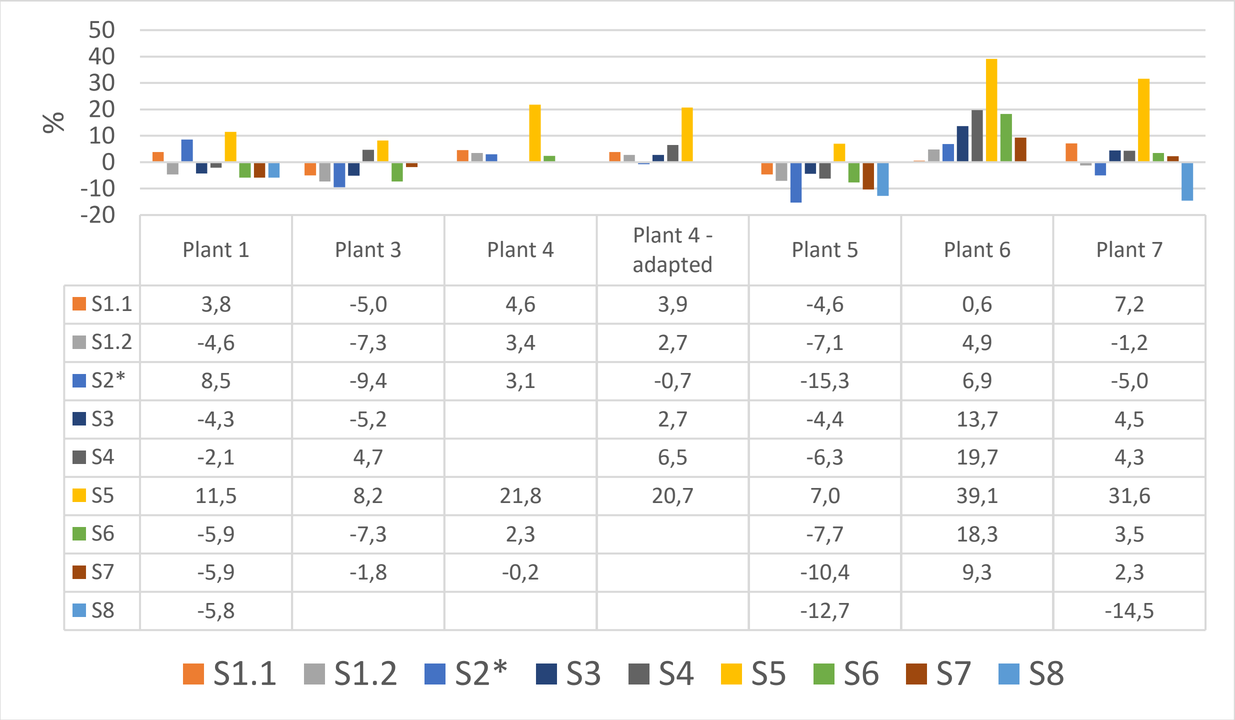
Unsurprisingly, MBE figures calculated with both hourly and daily values are very close to the above relative differences (calculated for one year).



	Plant 1	Plant 3	Plant 4	Plant 4 - adapted	Plant 5	Plant 6	Plant 7
S1.1	3,9	-1,4	5,9	5,1	-1,3	0,8	5,7
S1.2	-2,7	-2,7	5,0	4,2	-2,4	4,2	-2,7
S2*	6,0	-4,6	4,9	1,1	-9,4	4,9	-2,9
S3	-1,1	1,0		3,9	-0,9	10,5	5,2
S4	1						
S5	9,9	9,4	21,3	19,9	7,5	28,2	20,5
S6	-3,4	-2,6	3,7		-3,0	12,5	3,8
S7	-1,9	5,5	1,6		-5,2	6,7	1,9
S8	-3,8				-7,7		-10,0

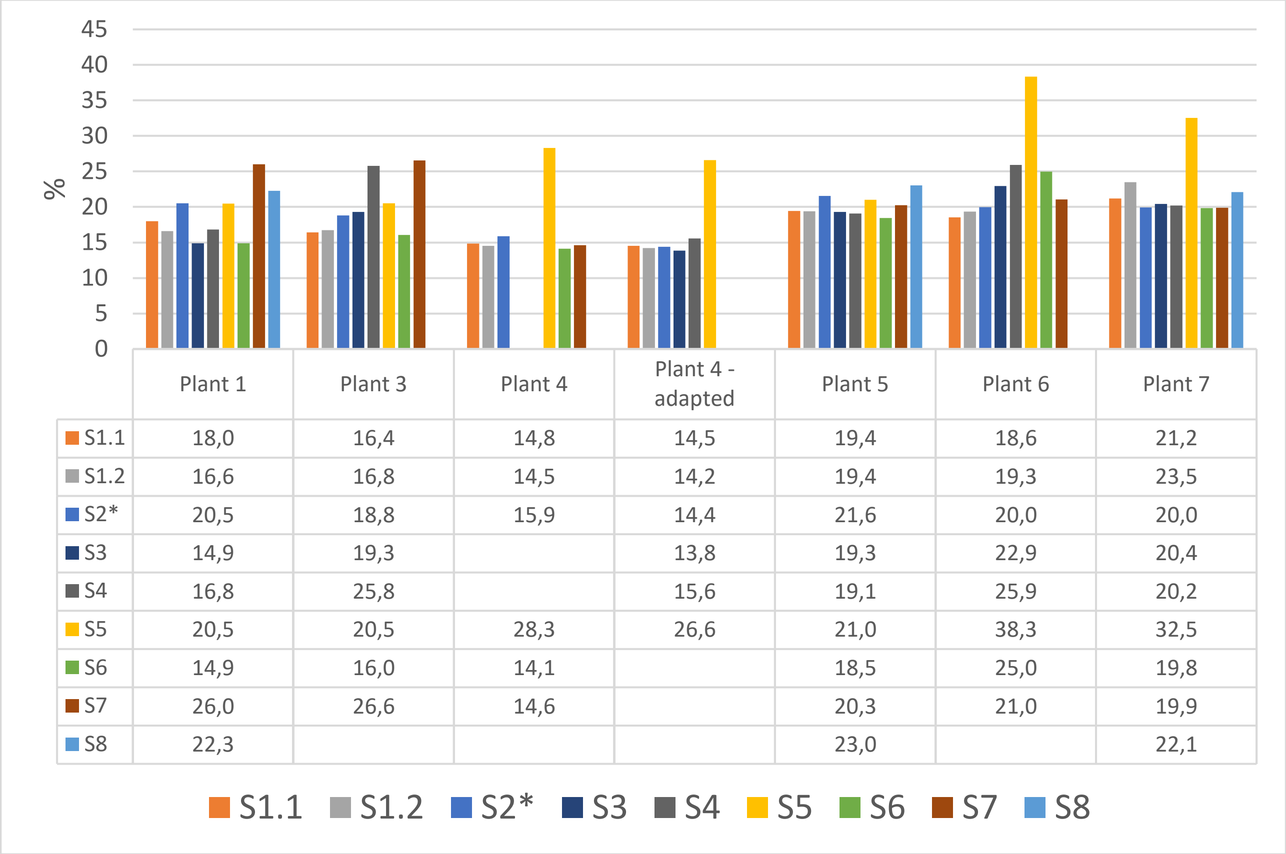
Figure 3.5: Normalised Mean Bias Errors

The following figures present the hourly values of NMBWE, NRMSE and NRMSWE. The daily values are found in annex.



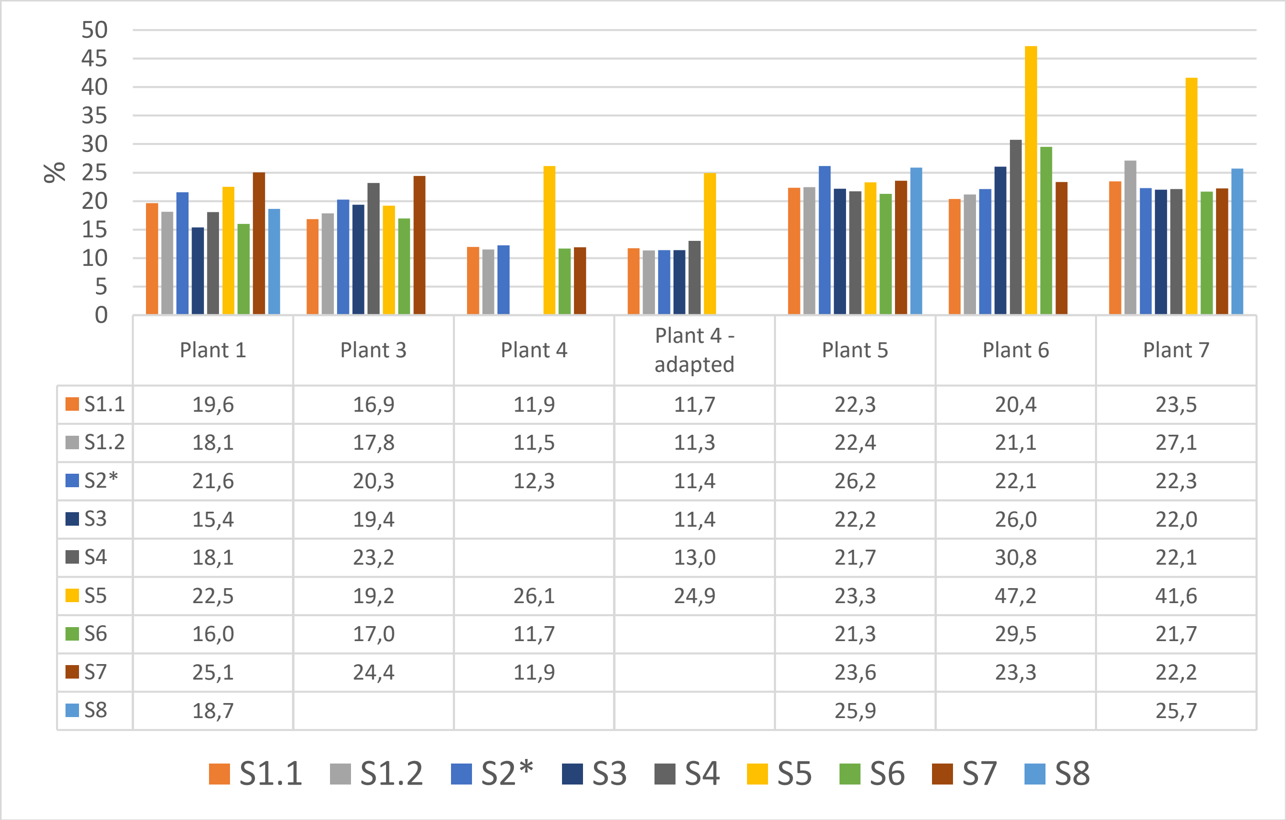
	Plant 1	Plant 3	Plant 4	Plant 4 - adapted	Plant 5	Plant 6	Plant 7
S1.1	3,8	-5,0	4,6	3,9	-4,6	0,6	7,2
S1.2	-4,6	-7,3	3,4	2,7	-7,1	4,9	-1,2
S2*	8,5	-9,4	3,1	-0,7	-15,3	6,9	-5,0
S3	-4,3	-5,2		2,7	-4,4	13,7	4,5
S4	-2,1	4,7		6,5	-6,3	19,7	4,3
S5	11,5	8,2	21,8	20,7	7,0	39,1	31,6
S6	-5,9	-7,3	2,3		-7,7	18,3	3,5
S7	-5,9	-1,8	-0,2		-10,4	9,3	2,3
S8	-5,8				-12,7		-14,5

Figure 3.6: Normalised Mean Bias Weighted Errors (hourly)



	Plant 1	Plant 3	Plant 4	Plant 4 - adapted	Plant 5	Plant 6	Plant 7
S1.1	18,0	16,4	14,8	14,5	19,4	18,6	21,2
S1.2	16,6	16,8	14,5	14,2	19,4	19,3	23,5
S2*	20,5	18,8	15,9	14,4	21,6	20,0	20,0
S3	14,9	19,3		13,8	19,3	22,9	20,4
S4	16,8	25,8		15,6	19,1	25,9	20,2
S5	20,5	20,5	28,3	26,6	21,0	38,3	32,5
S6	14,9	16,0	14,1		18,5	25,0	19,8
S7	26,0	26,6	14,6		20,3	21,0	19,9
S8	22,3				23,0		22,1

Figure 3.7: Normalised Root Mean Square Errors (hourly)



	Plant 1	Plant 3	Plant 4	Plant 4 - adapted	Plant 5	Plant 6	Plant 7
S1.1	19,6	16,9	11,9	11,7	22,3	20,4	23,5
S1.2	18,1	17,8	11,5	11,3	22,4	21,1	27,1
S2*	21,6	20,3	12,3	11,4	26,2	22,1	22,3
S3	15,4	19,4		11,4	22,2	26,0	22,0
S4	18,1	23,2		13,0	21,7	30,8	22,1
S5	22,5	19,2	26,1	24,9	23,3	47,2	41,6
S6	16,0	17,0	11,7		21,3	29,5	21,7
S7	25,1	24,4	11,9		23,6	23,3	22,2
S8	18,7				25,9		25,7

Figure 3.8: Normalised Root Mean Square Weighted Errors (hourly)

The plants which present the less good KPIs, i.e. the highest MB(W)E and RMS(W)E are plants 5, 6 and 7, which belong to the types of technology to be studied in WP2, BIPV and floating PV.

The scatter plots presented below illustrate the hourly simulated vs measured data, for the different simulation tools of the study, indicatively for the case of Plant 5.

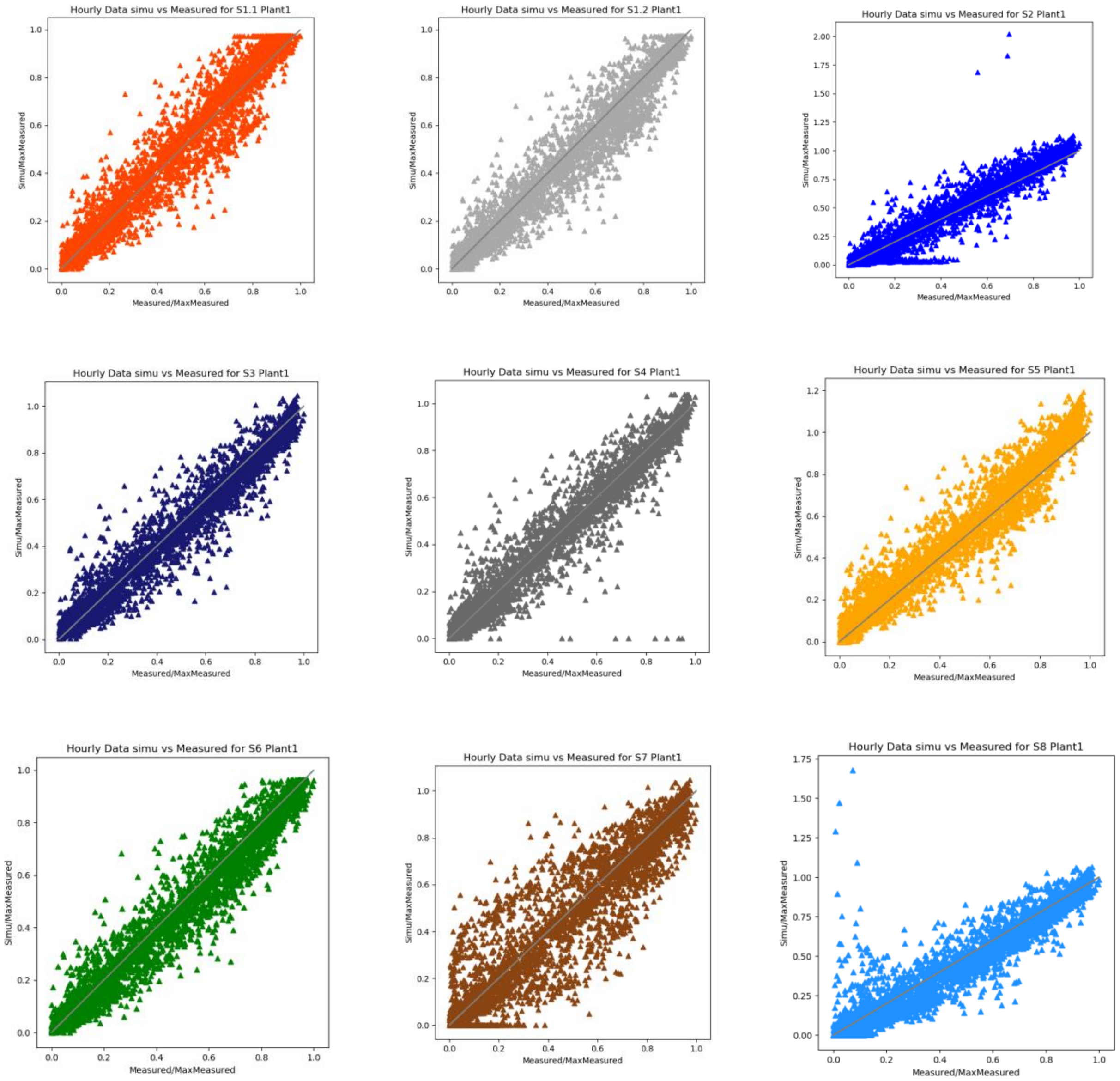


Figure 3.9: Scatter plots (hourly)

The work presented in this reported will be continued in the other tasks of WP2, to monitor the improvements of the models during the project. Besides, some plants' characteristics or simulation parameters which were at this stage of the project not known or not well characterised, should be more ascertained, thanks to the work planned in WP8 (Demonstration).

4 REFERENCES

[1] Moser, D., Lindig, S., Herz, M., Müller, B., Horvath, I., Schils, A., Ramesh, S., Green, M., Vedde, J., Barnard, D., Herteleer, B., Tsanakas, J.A., Ascencio-Vásquez, J., Richter, M. (2020). Benchmarking Yield Assessment Exercise in Different Climates within an International Collaboration Framework. 37th European Photovoltaic Solar Energy Conference and Exhibition. pp. 1317 - 1323. DOI: 10.4229/EUPVSEC20202020-5CO.10.3.

[2] Moser D. et al. (2020). Uncertainties in Yield Assessments and PV LCOE. Report IEA-PVPS T13-18:2020. ISBN 978-3-907281-06-2

[3] Perez R., Cebecauer T., Suri M., 2014. Semi-Empirical Satellite Models. In Kleissl J. (ed.) Solar Energy Forecasting and Resource Assessment. Academic press.

[4] Cebecauer T., Šúri M., Perez R., High performance MSG satellite model for operational solar energy applications. ASES National Solar Conference, Phoenix, USA, 2010.

[5] Šúri M., Cebecauer T., Perez P., Quality procedures of SolarGIS for provision site-specific solar resource information. Conference SolarPACES 2010, September 2010, Perpignan, France.

[6] Šúri M., Cebecauer T., 2014. Satellite-based solar resource data: Model validation statistics versus user's uncertainty. ASES SOLAR 2014 Conference, San Francisco, 7-9 July 2014.

[7] Cebecauer T., Suri M., Gueymard C., Uncertainty sources in satellite-derived Direct Normal Irradiance: How can prediction accuracy be improved globally? Proceedings of the SolarPACES Conference, Granada, Spain, 20-23 Sept 2011.

[8] Šúri M., Cebecauer T., Requirements and standards for bankable DNI data products in CSP projects, Proceedings of the SolarPACES 2011 Conference, September 2011, Granada, Spain.

[9] European Centre for Medium-Range Weather Forecasts. 2017, updated monthly. ERA5 Reanalysis. Research Data Archive at the National Center for Atmospheric Research, Computational and Information Systems Laboratory. https://doi.org/10.5065/D6X34W69. Accessed 30-04-2019.

[10] NASA MODIS website https://modis.gsfc.nasa.gov/data/

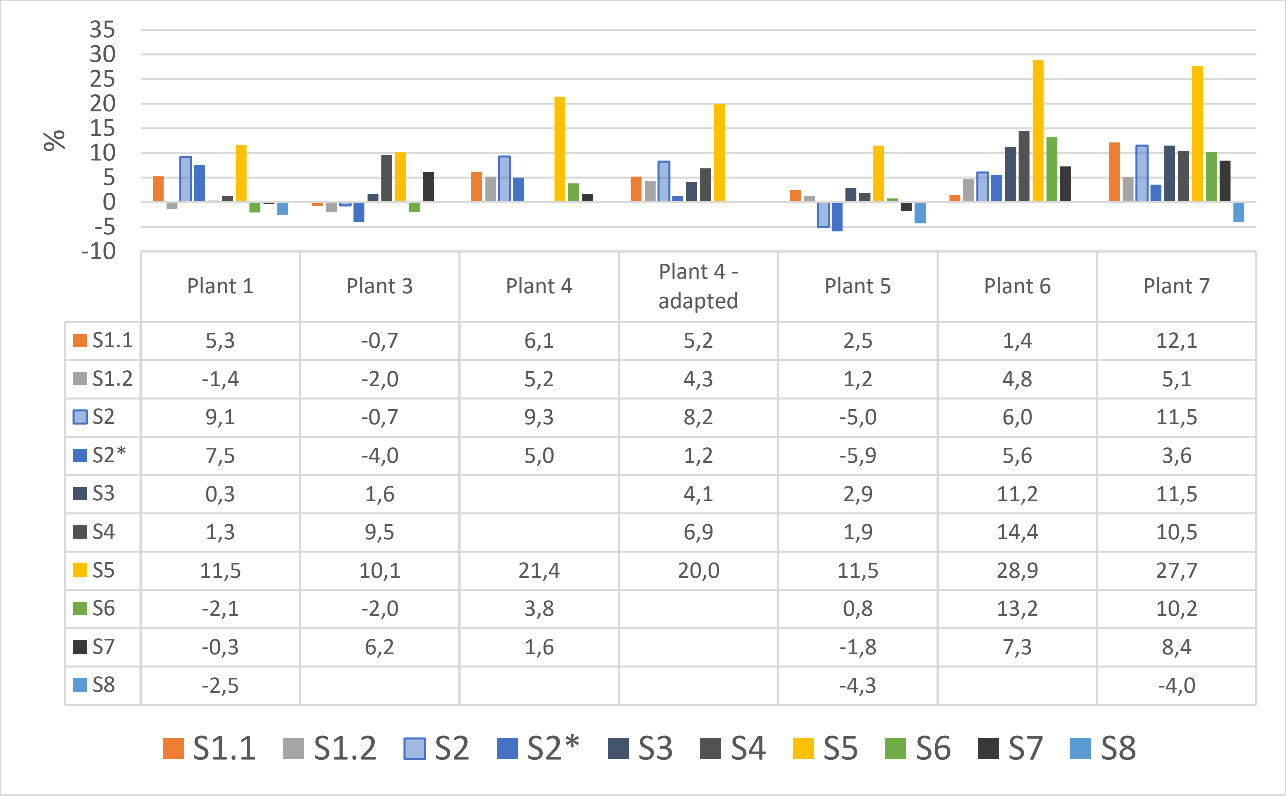
[11] AppEEARS Team. (2019). Application for Extracting and Exploring Analysis Ready Samples (AppEEARS). Ver. 2.24. NASA EOSDIS Land Processes Distributed Active Archive Center (LP DAAC), USGS/Earth Resources Observation and Science (EROS) Center, Sioux Falls, South Dakota, USA. https://lpdaacsvc.cr.usgs.gov/appeears/

[12] Schaaf, C., Wang, Z. (2015). MCD43A3 MODIS/Terra+Aqua BRDF/Albedo Daily L3 Global - 500m V006. NASA EOSDIS Land Processes DAAC. doi: 10.5067/MODIS/MCD43A3.006

[13] NREL, 1993. User's Manual for SERI QC Software-Assessing the Quality of Solar Radiation Data. NREL/TP-463-5608. Golden, CO: National Renewable Energy Laboratory.

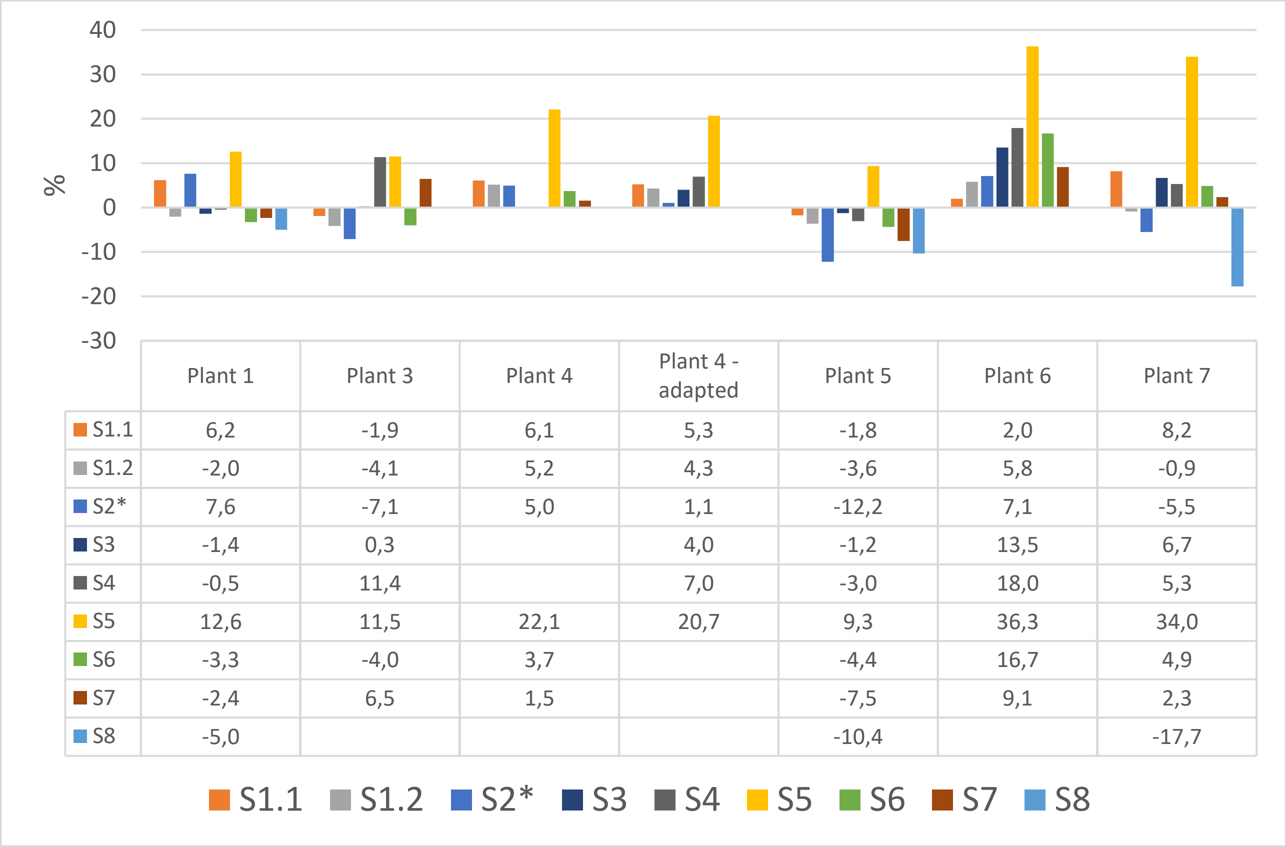
[14] Younes S., Claywell R. and Muneer T, 2005. Quality control of solar radiation data: Present status and proposed new approaches. Solar Energy 30, 1533-1549.

5 ANNEX



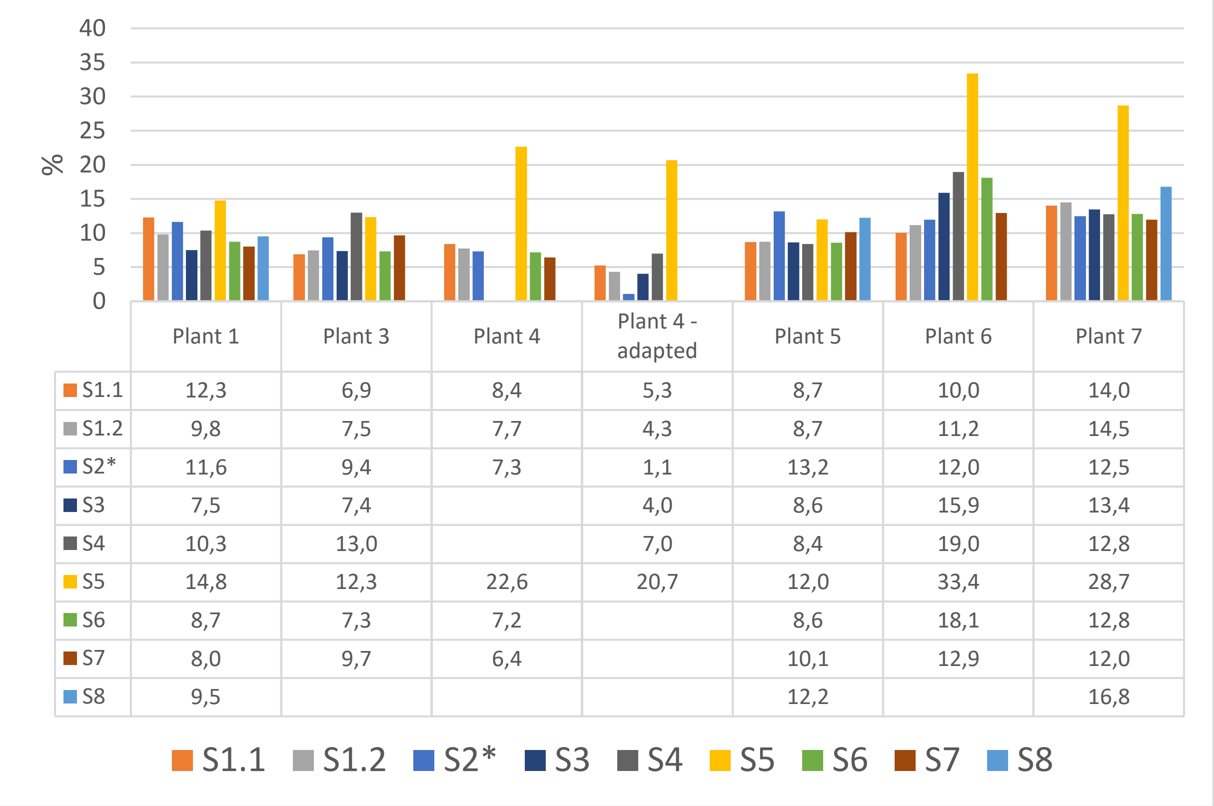
	Plant 1	Plant 3	Plant 4	Plant 4 - adapted	Plant 5	Plant 6	Plant 7
S1.1	5,3	-0,7	6,1	5,2	2,5	1,4	12,1
S1.2	-1,4	-2,0	5,2	4,3	1,2	4,8	5,1
S2	9,1	-0,7	9,3	8,2	-5,0	6,0	11,5
S2*	7,5	-4,0	5,0	1,2	-5,9	5,6	3,6
S3	0,3	1,6		4,1	2,9	11,2	11,5
S4	1,3	9,5		6,9	1,9	14,4	10,5
S5	11,5	10,1	21,4	20,0	11,5	28,9	27,7
S6	-2,1	-2,0	3,8		0,8	13,2	10,2
S7	-0,3	6,2	1,6		-1,8	7,3	8,4
S8	-2,5				-4,3		-4,0

Figure 5.1: Relative difference between simulated and measured production (non-corrected)



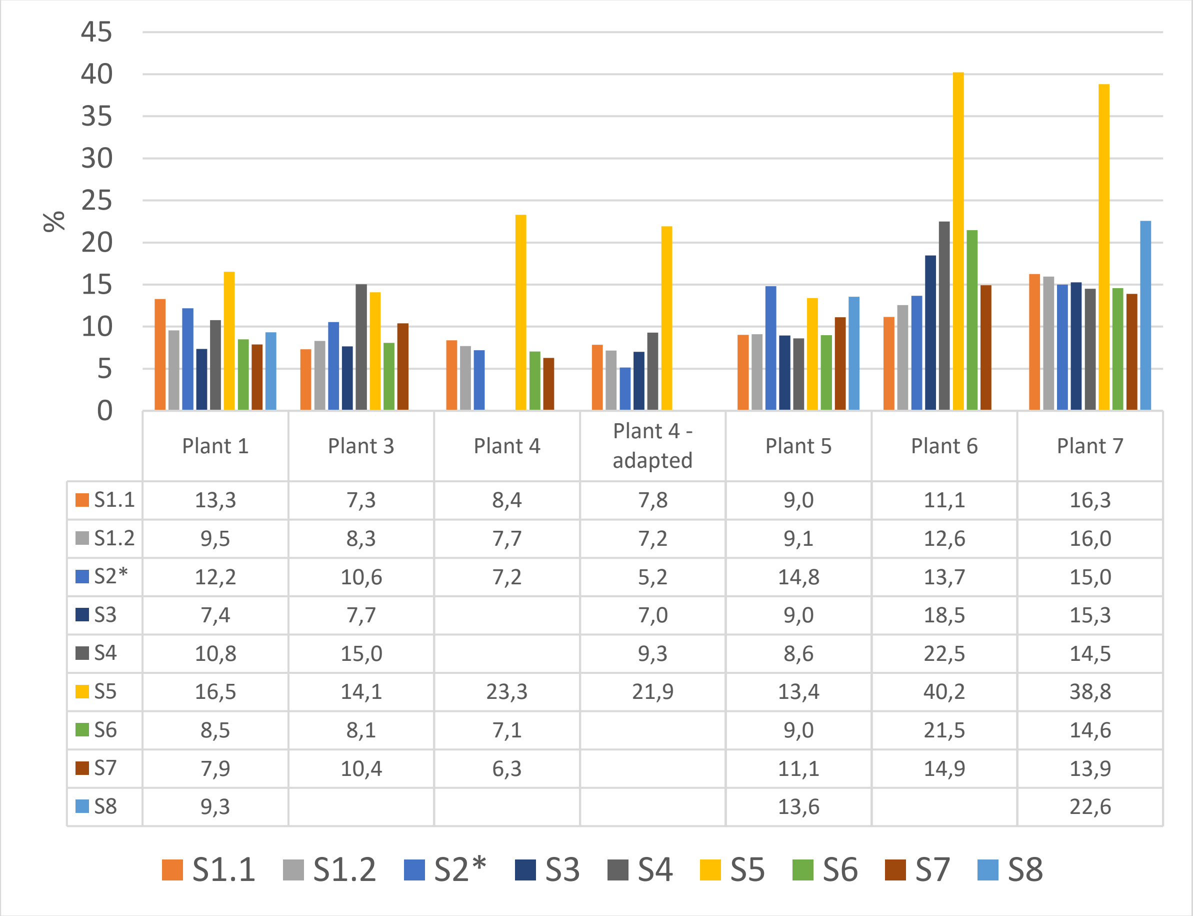
	Plant 1	Plant 3	Plant 4	Plant 4 - adapted	Plant 5	Plant 6	Plant 7
S1.1	6,2	-1,9	6,1	5,3	-1,8	2,0	8,2
S1.2	-2,0	-4,1	5,2	4,3	-3,6	5,8	-0,9
S2*	7,6	-7,1	5,0	1,1	-12,2	7,1	-5,5
S3	-1,4	0,3		4,0	-1,2	13,5	6,7
S4	-0,5	11,4		7,0	-3,0	18,0	5,3
S5	12,6	11,5	22,1	20,7	9,3	36,3	34,0
S6	-3,3	-4,0	3,7		-4,4	16,7	4,9
S7	-2,4	6,5	1,5		-7,5	9,1	2,3
S8	-5,0				-10,4		-17,7

Figure 5.2: Normalised Mean Bias Weighted Errors (daily)



	Plant 1	Plant 3	Plant 4	Plant 4 - adapted	Plant 5	Plant 6	Plant 7
S1.1	12,3	6,9	8,4	5,3	8,7	10,0	14,0
S1.2	9,8	7,5	7,7	4,3	8,7	11,2	14,5
S2*	11,6	9,4	7,3	1,1	13,2	12,0	12,5
S3	7,5	7,4		4,0	8,6	15,9	13,4
S4	10,3	13,0		7,0	8,4	19,0	12,8
S5	14,8	12,3	22,6	20,7	12,0	33,4	28,7
S6	8,7	7,3	7,2		8,6	18,1	12,8
S7	8,0	9,7	6,4		10,1	12,9	12,0
S8	9,5				12,2		16,8

Figure 5.3: Normalised Root Mean Square Errors (daily)



	Plant 1	Plant 3	Plant 4	Plant 4 - adapted	Plant 5	Plant 6	Plant 7
S1.1	13,3	7,3	8,4	7,8	9,0	11,1	16,3
S1.2	9,5	8,3	7,7	7,2	9,1	12,6	16,0
S2*	12,2	10,6	7,2	5,2	14,8	13,7	15,0
S3	7,4	7,7		7,0	9,0	18,5	15,3
S4	10,8	15,0		9,3	8,6	22,5	14,5
S5	16,5	14,1	23,3	21,9	13,4	40,2	38,8
S6	8,5	8,1	7,1		9,0	21,5	14,6
S7	7,9	10,4	6,3		11,1	14,9	13,9
S8	9,3				13,6		22,6

Figure 5.4: Normalised Root Mean Square Weighted Errors (daily)