1. As you’ll find out, this is going to be an iterative process where we are changing your assignment as we go. Sorry for this… but I think this is also what makes it interesting! :)

Reply: I guess we actually can do this from overleaf. I can share links for the document and then we can add comment and change together.

1. Good work in also including a motivation. I’d add some further motivation coming from that we need to make more dynamic use of the resources, closer to the physical limits vs over-dimensioning as done in the past.

Reply: Dear Mathijs, I didn’t understand very well. What do you mean for dynamic use (data-driven)? What do you mean for over-dimensioning?

1. I then see two main use cases:

- research to be able to develop new data-based methods that allow this dynamnic use

- (later) use in operation

Reply: Dear Mathijs, I didn’t fully understand the second part (use in operation).

1. I think it would be interesting and very helpful to provide some concrete use-case examples. E.g.

- a researcher designing algorithms for dealing with congestion

- a researcher working on a method to make a grid reinforcement schedule

=> can you find these from the papers describing research that uses the models and data?

From these use cases we can then hopefully better argue:

- what data and models would they need?

- how would they look for this? (=> which meta-data is important to them?)

For example, about the models I think we need things like:

- if it is a method/model for finding a solution, to which problem? what are the input parameters, what is the output (decision variables)? what is the objective?

- if it is a simulator: also what is the input? what is the level of detail? is the output only about feasibility or also quality of the input? which criteria/measurements are available for a run?

Reply: Dear Mathijs, Thanks for your suggestions for this. It is also what I want to do. For future platform, I believe the data and model should be integrated, that means they are together to be used for a task, like congestion and schedule problem. One thing is that how many and how to categorize the representative tasks in AI and electricity system. We can start from above-mentioned 2 tasks. But also we need to think the holistic map for this.

1. About available guidance & documentation => could you indicate some level (none / basic / …)?

About matching data with models => also here I think the use cases could help to identify which parameters to include so possible matches can be identified automatically.

About the meta-data of the data:

- the date (range) for which this is available

- some indication of the size of a data set: e.g. number of data points?

Reply: Dear Mathijs, thanks for your suggestions, I will revise these part in next 2 weeks .

1. Some detailed comments about the document itself:

What is the aim of this section 2? To critically reflect on the collected/available data?

I don’t completely understand the “directly use” part.

Reply: Dear Mathijs. Section 2 is the part to explain what is the electricity data and model, and summarize the current data and model drawbacks.

Sorry I didn’t explain “directly use” very well. In my perspective, directly use is that: For example, you find a research, named data-driven based eliminating congestion approach. Now you have read the paper, and want to dive into the detail. They open sourced code of model and data. Just download and run successfully. That’s the directly use. Usually, in the code, there must be a data preprocessing, to make the originally data suit to the model training. Currently, Existing platforms cannot meet this demand.

1. Also please indicate what is new next time. + list your questions / doubts and plans for the next two weeks. Can you share a git repo with your code & metadata with us?

Reply: Dear Mathijs. Thanks for your suggestions. Next time reporting I will list my questions doubts and plans for next weeks. I also created a git repo for code and meta-data and the link is: <https://github.com/ShengrenHou/Energy-Data-and-Model-Platform>

For Next 2 weeks, My plan is:

1. Revise and improve the current meta-data issues. Adding the date (range) for which this is available supplement some indication of the size of a data set: e.g. number of data points?
2. Reading papers for congestion and schedule in power market, collecting and summarize what data and model they need. The detail input and output, parameters inside model. How the data is preprocessed.
3. Summarize and write the draft for the report. I will submit the new report at 3, Oct.