

Response to Reviewer 2 Comments

Thank you for taking the time to review our manuscript entitled “**Imitation Learning-based energy management algorithm: Lille Catholic University Smart Grid demonstrator case study**”. We appreciate the effort you have dedicated to provide your insightful feedback and comments. Please find the detailed responses below and the corresponding corrections highlighted within the re-submitted manuscript.

1. Point-by-point response to Comments and Suggestions for Authors

Comments 1: *“Line 353. Incorporating a more detailed discussion about stability and efficiency would enhance this portion.”*

Response 1: Thank you for bringing this to my attention. The mentioned part has been improved by more details (Line 412).

Comments 2: *“In Table 2. The number of iterations is only about 1/1000 (= 100000/100), while the simulation time is about 1/87 (=83.96 min/58 sec). Why they differs such a lot? Are there other factors that have a greater impact on simulation time than the learning aspect?”*

Response 2: Yes, the learning processes of the two algorithms are different. The first algorithm based on Reinforcement Learning (Q-learning) learns through experience. In other words, it learns to make decisions while interacting with the MG simulation environment. Accordingly, at the beginning of the training process it does not has any information about the MG states or the task it is supposed to learn. For this reason, the Q-learning needs more iterations than our proposed approach to explore, understand the MG environment before converging to an optimal control policy. In contrast, our proposed approach based on Imitation Learning and Reinforcement Learning (Imitation-Q-learning) imitates an expert demonstration to learn how to manage the studied microgrid energy. Those demonstrations, as mentioned in our manuscript, are derived from solving a set of energy management problems (several days) using Linear Programming. Since Linear Programming gives the exact theoretical solutions of the energy management problem for several days, the proposed algorithm, Imitation-Q-learning has the perfect demonstrations. Consequently, it does not need a lot of iterations to understand the MG environment and figure out an optimal policy. It simply learns to imitate Linear programming. This makes the proposed approach faster, easier and more effective than the conventional Q-learning algorithm.

2. Response to Comments on the Quality of English Language

Point 1: “The writing can be improved.”

Response 1: Thank you for pointing this out. Accordingly, the manuscript writing has been checked.

3. Additional clarifications

Point 1: “higher resolution figures”

Agree. Accordingly, Prism software has been used to plot all the results with higher resolution. Further, the proposed approach diagram has been replaced.