

Interactive comment on “Modeling and Clustering Water Demand Patterns from Real-World Smart Meter Data” by Nicolas Cheifetz et al.

Anonymous Referee #1

Received and published: 11 April 2017

The authors present in this manuscript an interesting process to modelling and clustering hourly water demand in urban areas. I personally like this approach. However, I have a number of important concerns regarding both the manuscript and the mentioned process.

- The set of methodologies selected to analyse the time series of water demand should be specifically justified. I mean, what are the advantages/disadvantages to handling this data with Fourier and no other options such as wavelets? Why Functional clustering is suitable for this case? On this way, Section 2 should be extended and also tailored to the specific problem of water demand.
- The authors propose a univariate time series analysis, which is suitable to have a good approach of the selected methodologies. However, they don't explain why they

C1

are not considering exogenous variables such as different weather inputs. This is an important point as the 1st sentence of the manuscript specifically mention "climate change" as one of the main reasons why this work can have a proper research impact.

- The data description in 5.1 should clarify the actual data handled by the authors. At this point, this is not easy to figure it out from where each of the time series is coming.
- The authors propose 8 as the ideal number of clusters (BIC criterion). However, these 8 clusters are further classified as member of one over 4 categories. This makes me wonder why don't directly use 4 groups. Still, why the authors use clustering and they don't use classification tools or any semi-supervised approach as they have certain background information which is not taken into account on the clustering process. I also miss a measure of time series distance to approach how well is the clustering configuration.
- Importantly, I miss an explanation of the usefulness of the approach. One of the basic applications should be quantify how the current proposal may support further predictive models. How much accuracy is expected to gain by using this process?

As general assessment, I recommend to do major revisions to the manuscript ahead to attempt to be published. The authors should work in depth aiming to have a higher quality work which current version can't be accepted.

Interactive comment on Drink. Water Eng. Sci. Discuss., doi:10.5194/dwes-2017-19, 2017.

C2