Drink. Water Eng. Sci. Discuss.,
https://doi.org/10.5194/dwes-2017-19-AC2, 2017

Open Access
Drinking Water
Engineering and Science
Discussions

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

Nicolas Cheifetz et al.

nicolas.cheifetz@veolia.com

Received and published: 2 June 2017

Dear Referee,

First, we would like to thank you for taking time to review this paper and providing us constructive comments and suggestions. The major comments concern the application of the proposed approach and the consideration of periodicities.

Identifying the major usage profiles from water consumption is an interesting topic to water utilities. Indeed, the resulting segmentation helps the water companies to gain a better knowledge about users consuming the distributed water. The user is having a better experience with the tools developed by their water utility. For instance, users at

Veolia Eau d'Ile de France (Paris area in France) can already monitor their water index / consumption on a dedicated website for free. Based on our clustering results, people could compare with similar patterns and adapt their consumptions according to their needs. In addition, customer services might alert the user if a leakage is detected. Concerning the grid management, each prototype can be used to represent the water behavior of users belonging to the same cluster. An erratic water pattern (like in cluster 2) can be the sign of a leakage and might initiate a corrective action. Sampling a large amount of water meters is useful for several topics (e.g. tracking the meter metrology, estimating the global consumption modes based on a limited number of meters) ; such sampling analysis is straightforward using our meter segmentation.

In this article, a double seasonality is taken into account: daily and weekly. This prior knowledge is included both in the modeling of the Fourier-based decomposition and the prototype functions of the Fourier Regression mixture (FReMix) model. To clarify the definition of the time series used for clustering, please read the last paragraph of section 3: *From each time series y_i, the model parameters defined by Equation (1) are thus identified, and the periodic seasonal pattern defined by $x_i = (x_{i1}, \ldots, x_{im})$, with $m = 168$, is extracted. Due to the periodicity of the series $(x_{it}, \ldots, x_{iT})$ defined by Equation (2), it should be noted that the first terms $m = 168$ are sufficient to characterize the time series. [...] The set of normalized seasonal patterns is used as input data for the clustering step [...].* In other words, the time series x_i have a length of 168 due to the trigonometric modeling of the chosen Fourier basis decomposition. The Fourier coefficients are identified performing a multiple linear regression on the global time series y_i detrended (by subtracting a moving average of order 168) which limits the effect of a long seasonality. Finally, our sample has a duration of 15 months (with 5 seasons at most) and we are more interested about identifying the major mode of consumption than estimating water demand profiles with local changes and a fine granularity.

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