Intro for Thesisring

Hi everyone, a late submmission to the thesisring but I just finished my Data section and saw that there was still a free spot on the submission list ;)

I tried to write my data section as concise as possible: writing it in such a way that there is the minimal amount of required information, but without missing parts. The thing that I would be interested in most is:

1. Do you agree with this concise writing style?
2. Is there any important information missing from the mentioned datasets?
3. Do the figures help to give a bit of insight into the data and if not, do you have any suggestions on what could possibly be added in terms of figures?

It is not a very long datasection so it should not take too much time, but I'd still like some feedback wherever possible :)

Thank you!

Data

In this study, two types of data are used: CML signal data which is used as input to the model, and precipitation observations as reference to evaluate the model. Both types of data will be elaborated upon in this section.

Commercial Microwave Link data

The CML data used in this research is the same as used in previous research (Overeem et al., 2016). The CML data is required from NOKIA microwave links, operated by T-Mobile. The dataset spans from January 14 2011 until June 30 2013. The dataset consists of 3101 links, but as shown before (Overeem et al., 2016), not all of these links are useful or continuously available. The total number of actual links used will therefore be lower, depending on the quality. Received Signal Level (RSL) in decibel-milliwatts (dBm) is stored at 15-minute intervals, at which the minimum and maximum signal level over the time period are stored. The power resolution of this data is 1 dB.

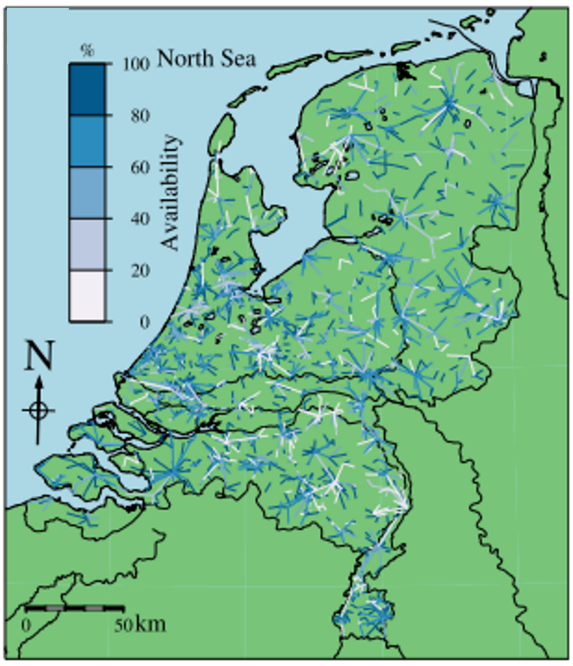

Figure 1: Used CML network in the Netherlands. Retrieved from Overeem et al. (2016)

Apart from the dynamic time signal data, static metadata is included as well. This includes coordinates of the start and end point of the link, distance between these points and frequency per link. Figure 1 shows the distribution of the links throughout the Netherlands. As the frequency at which the links operates and the distance the link covers determine the baseline for the link's signal, their respective distributions can be found in Figure 2. The spikes in the frequency distributions indicate that the links roughly operate at 4 different frequencies. The distance distributions shows that although most links operate at quite short distances (<2 km), there are still links included in the data set that cover over 5 or even 10 km.

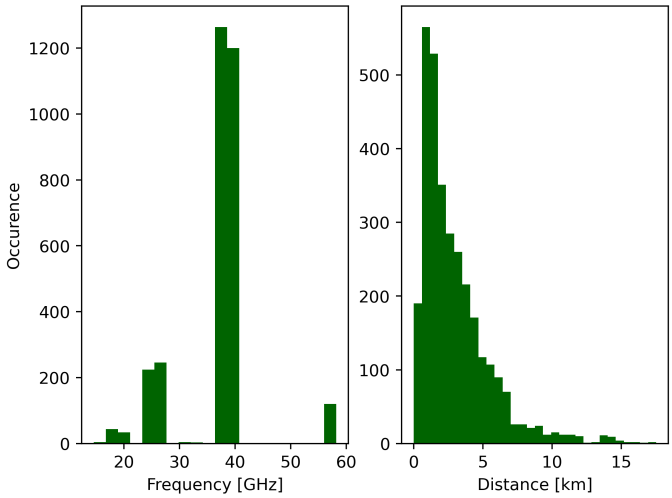

Figure 2: Frequency and distance distributions for the CML data. Figures apply to the final independent test set covering half of 2013

Reference precipitation data

The precipitation observations are taken from the gauge-adjusted radar rainfall product (freely available as "Radar precipitation climatology" via http://climte4impact.eu) with a resolution of 0.9 km^2. The radarmeasurements are adjusted with the use of both automated and manual rain gauge networks operated by the Royal Netherlands Meteorological Institute (KNMI). The gauge-adjusted radar observations are at a 5-min temporal resolution. To match the 15-min intervals from the CML data, the precipitation data is summed to 15-min intervals.

Bibliography

Overeem, A., Leijnse, H., & Uijlenhoet, R. (2016). Retrieval algorithm for rainfall mapping from microwave links in a cellular communication network. *Atmos. Meas. Tech.*, *9*(5), 2425-2444. (AMT) doi: 10.5194/amt-9-2425-2016