This is a supplementary material for the paper: “A Dynamic Ridesplitting Method with Potential Pick-up Probability Based on GPS Trajectories”. Here we provide the white noise assumption analysis and the details about the matrix settings associated with the kalmen filter model in the pick-up probability prediction.

For each grid, the proposed Kalman filter model assumes the number of ride requests in the same time interval to be stable between two consecutive days. Also, the number of ride requests is stable between consecutive time intervals in the same day. Therefore, there are two noise assumptions in the model. For a time interval t in day m, first, the ride requests number difference (i.e., the process noise ω_m) from the time interval t of the previous day m-1 is assumed to be white noise; second, the ride requests number difference (i.e., the measurement noise v_m) from previous time interval $t-1$ of day m is assumed to be the white noise. These two assumptions are supported by our analysis of ride requests number distribution on two real-world datasets.

ω_m and v_m are white noise if the sequence of ride requests number differences is identically distributed with a mean of zero and are not autocorrelated. We checked these two conditions for both datasets.

On the San Francisco dataset, the distribution of process noise ω_m is shown as the following figure 1, the mean is 0.01 and the standard deviation is 1.34:

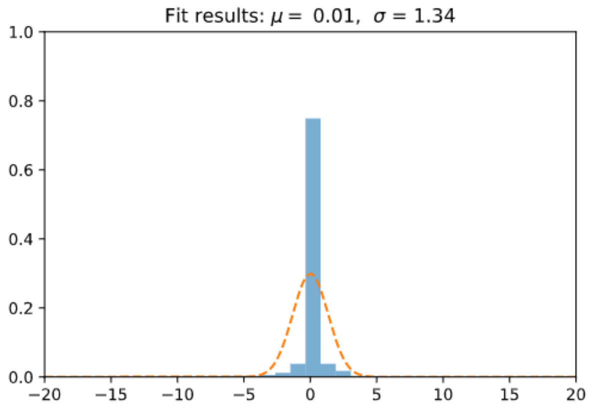

Figure 1. Process noise ω_m distribution on SF

The correlogram of ω_m is shown in the following figure 2, where all spikes are within the 95% confidence interval. The correlogram does not show any obvious autocorrelation pattern of ω_m. Therefore, ω_m is assumed to be the white noise on San Francisco dataset.

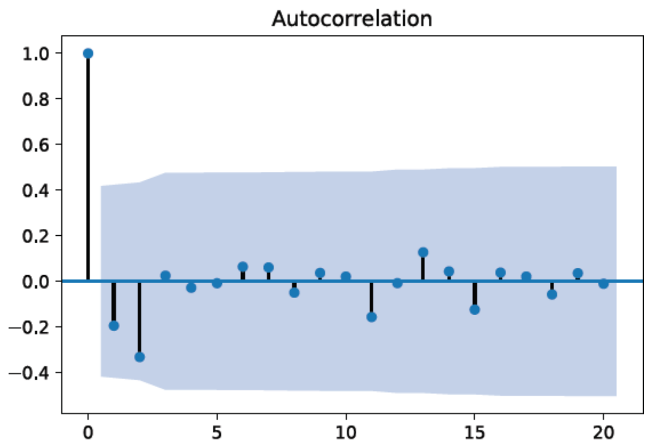

Figure 2. Process noise ω_m autocorrelation on SF

Similarly, we checked v_m on the San Francisco dataset. The distribution and correlogram are shown in Figure 3.

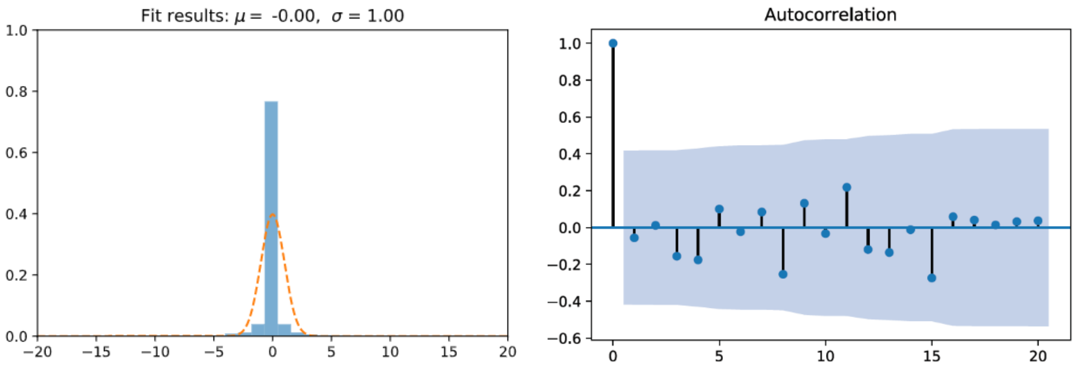

Figure 3. measurement noise v_m distribution and autocorrelation on SF

Since v_m is also identically distributed with a mean of zero and is not autocorrelated, it is assumed to be the white noise on San Francisco dataset.

The distribution and autocorrelation check of ω_m and v_m on Wuhan dataset is shown in Figure 4 and Figure 5, respectively.

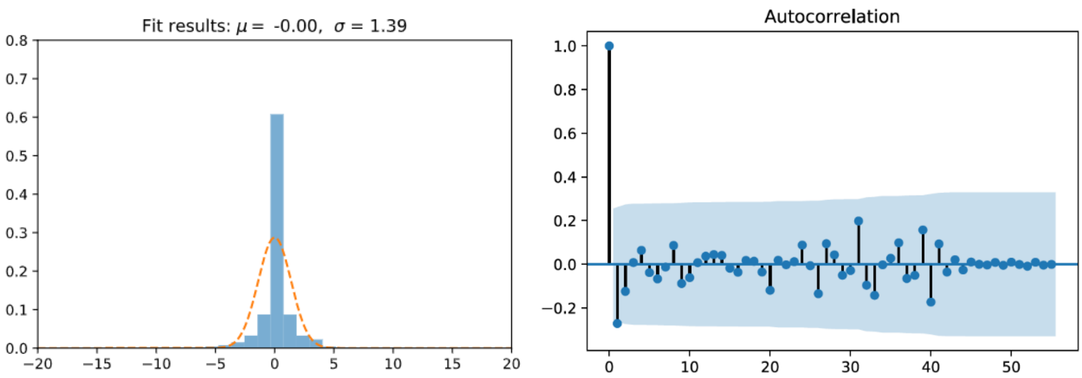

Figure 4. Process noise ω_m distribution and autocorrelation on WH

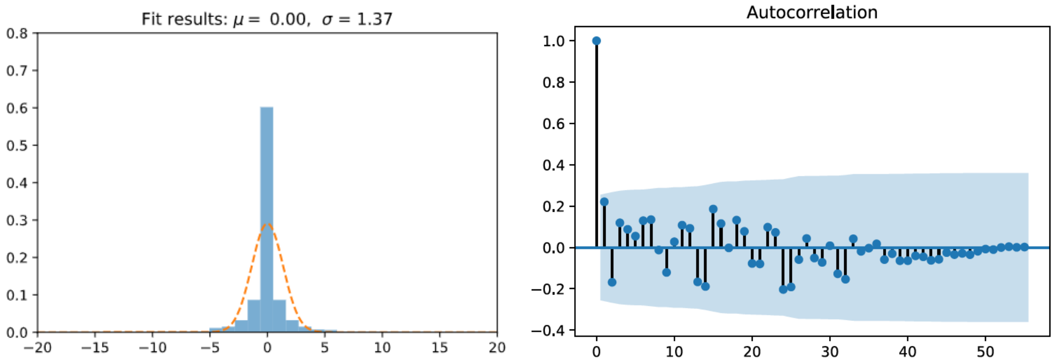

Figure 5. measurement noise v_m distribution and autocorrelation on WH

Based on the distribution and autocorrelation, it is assumed ω_m and v_m are white noise on the Wuhan dataset.

Therefore, we assume ω_m and v_m are white noise.