Data description

For the simulation baseline of NYC, there are five input files in total: config file, network file, schedule file, population file and vehicle file.

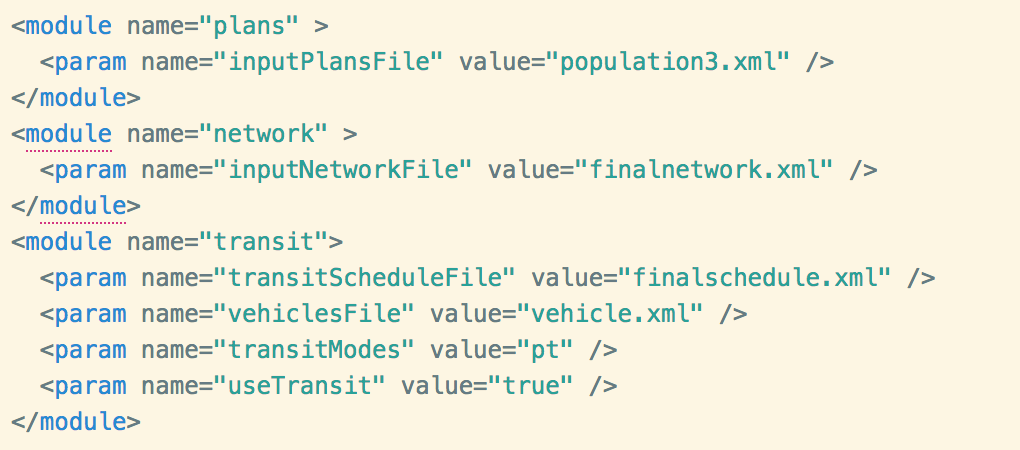
1. Config file

The config file includes the basic information to run simulation. You can set the parameters by yourself to simulate a specific scenario.



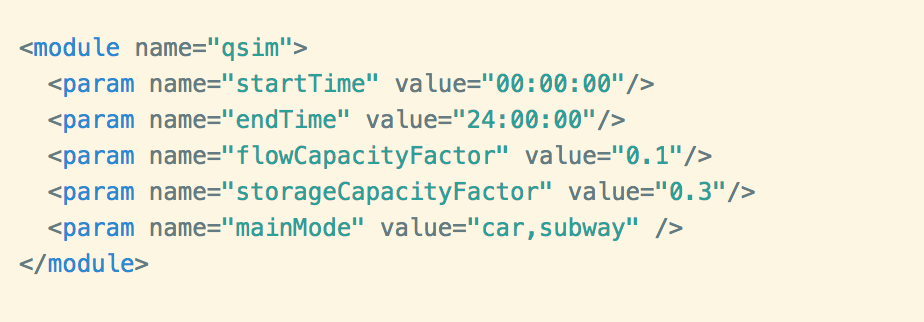
**Fig.1 Parameters of controller**

Fig. 1 shows the parameters of controller. The “outputDirectory” is the location where you want to output. Just keep it as Fig. 1 and it will generate an output folder in the same folder of your config file. The “firstIteration” and “lastIteration” will decide how many iterations of simulation you want. The setting of Fig. 1 will generate 10 iterations.



**Fig. 2 Parameters of input files**

Fig. 2 shows the locations of other input files. Just keep them as Fig. 2 and put all these files in the same folder of your config file.

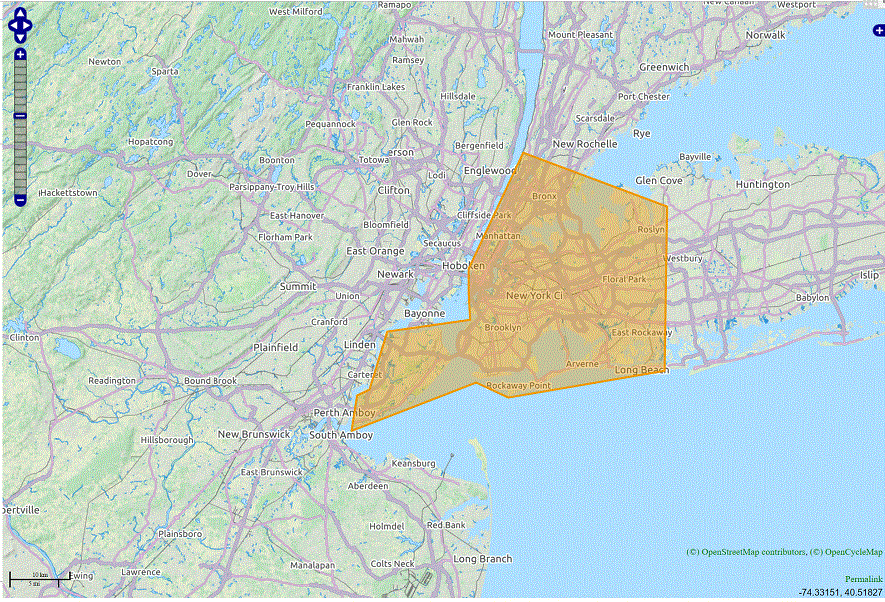


**Fig. 3 Parameters of simulation**

Fig. 3 shows the parameters of simulation. By setting “flowCapacityFactor”, MATSim will scale down the flow capacity to the same factor. Here we use 0.1, which means the scenario we running is 10% of the real world. (It scales down the model to fit the population rather than duplicate the population to fit the model.)

1. Network

The network file was generated from the data extracted from OpenStreetMap. It includes roads information of five boroughs of NYC.



**Fig. 4 Area of data extracted from OpenStreetMap**

1. Schedule

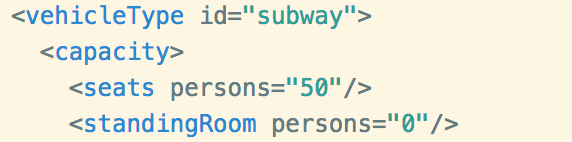
Schedule files area generated from historical GTFS data. It includes transit stops, transit routes and departure times of each subway line in both weekdays and weekends. There are two schedule files in our experiment. “finalschedule.xml” is the original transit schedule file, while the “finalschedule-half.xml” is the transit schedule that reduced the headway of A-line to half.

1. Population

The population file consists of the travel plan for each agent. It contains the activity types, locations, and end times. We generated this population file from NYMTC 2010/2011 household travel survey, by filtering the trips both started and ended in five boroughs of NYC and obtaining 8487 agents in total. Actually, they are only representatives of whole population. In the future, we need to duplicate it into a full population file by multiplying the zone-to-zone factor to each agent.

1. Vehicle

The vehicle file is also generated from the historical GTFS data. It is corresponding to the schedule file. You could change the capacity of trains in this file.



**Fig. 5 Capacity parameters**

1. Running instruction

To run MATSim with these input files, modelers don’t need to change a lot of parameters. First, they need to set the value of “transitScheduleFile” in Fig. 2 to “finalschedule.xml” and run the simulation to obtain the results with the regular schedule. Then, change the same value to “finalschedule-half.xml” and also the output file name in Fig. 1 to obtain the results with the schedule of reduced-headway. Those two results are what we shown in the book.