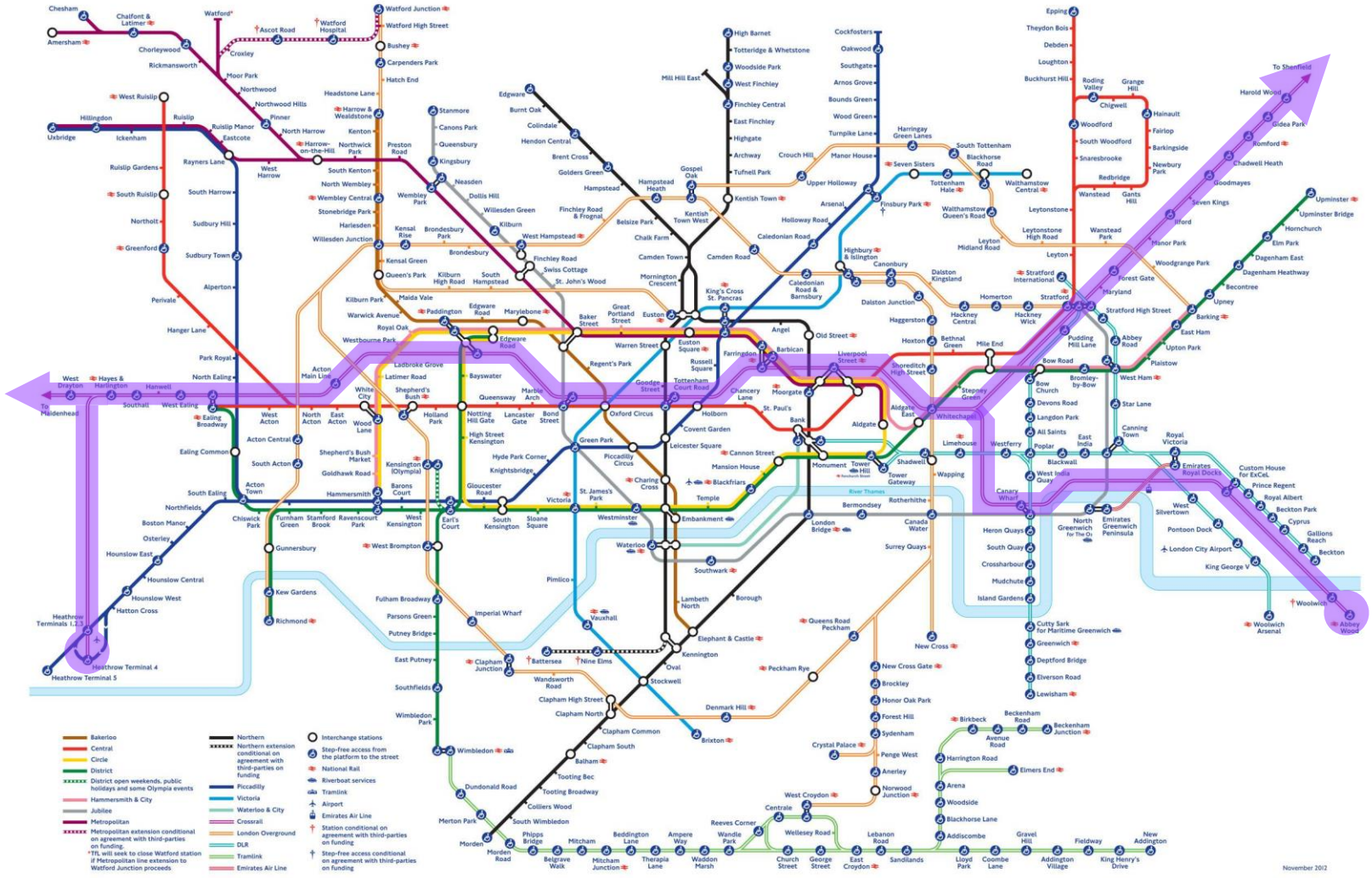
alius consulting

In collaboration Understanding Railways TRENOLab MTR

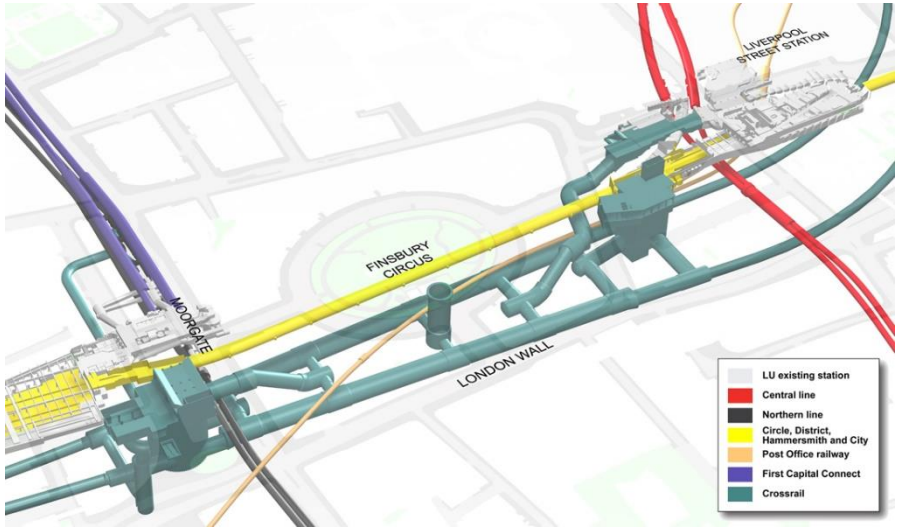
Big Data to improve the quality of operations: application to Crossrail in London

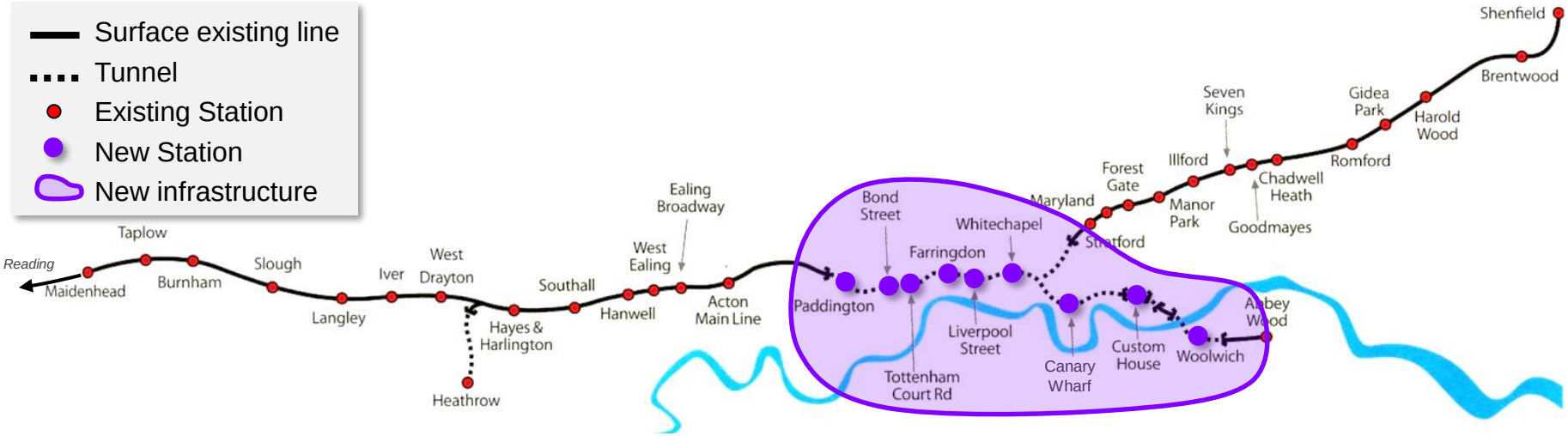
APTA 2018 RAIL CONFERENCE

A new railway service across London from Est to West, going also outside the Greater London area.

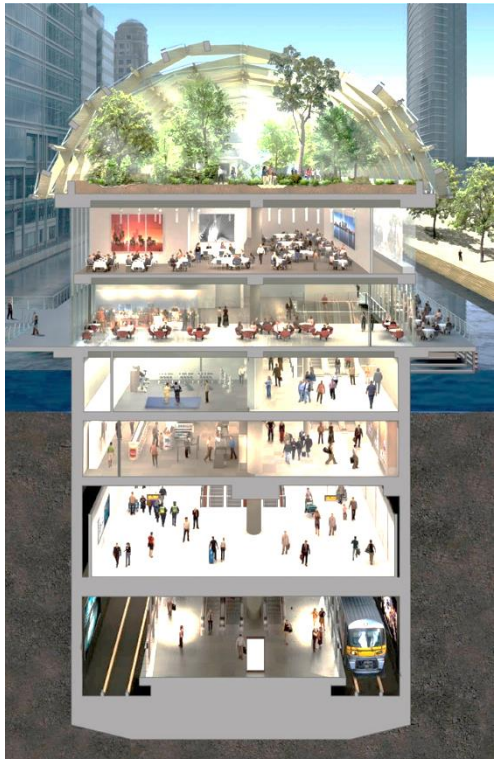


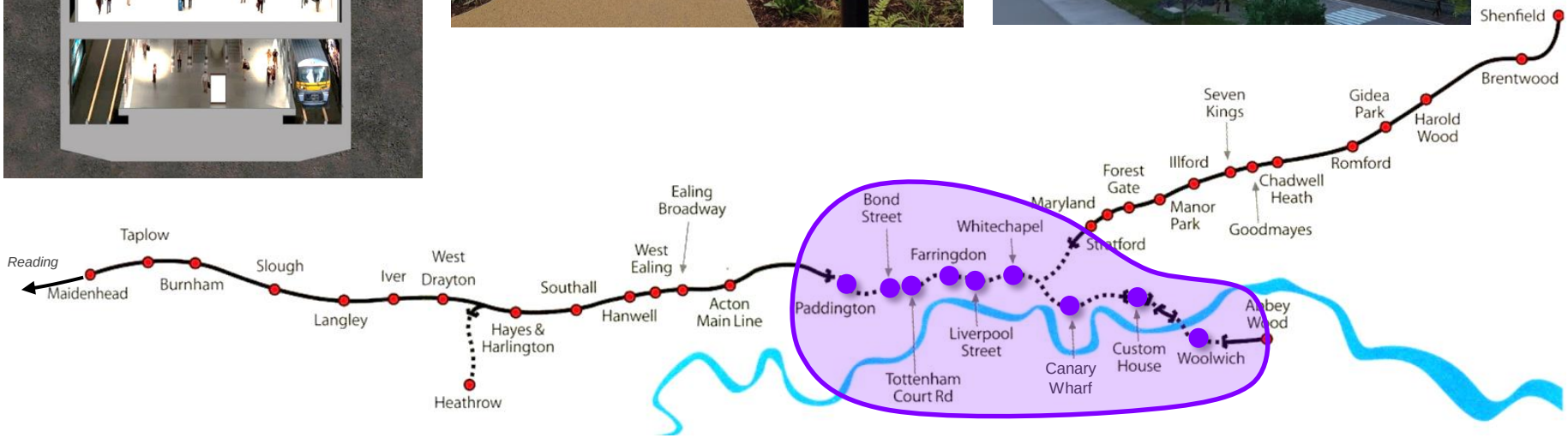
A new underground link between existing rail lines through the construction of a new tunnel in the city center.





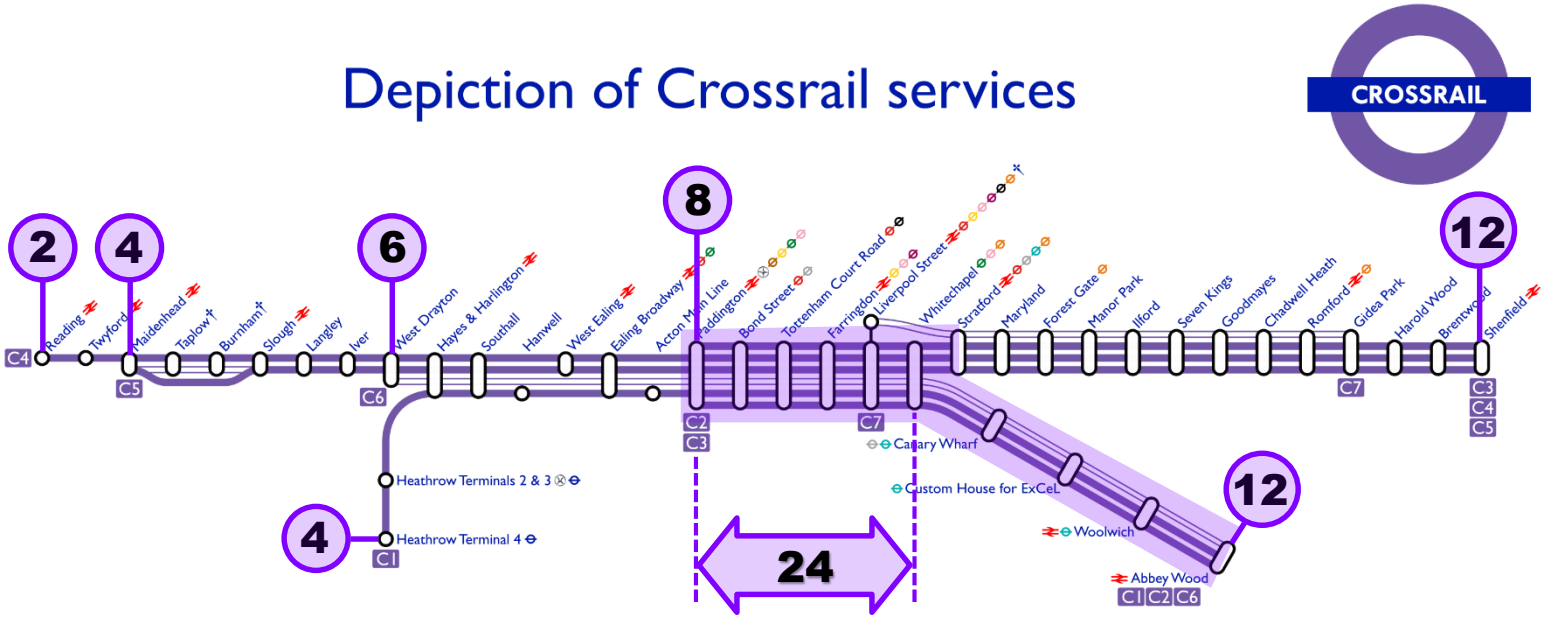
Nine new stations and new connections with the airports, Network Rail stations, the Underground and Dockland LR.





A very dense service, with 24 trains per hours and direction during the peaks through the central section.

Depiction of Crossrail services



† Notes

Taplow and Burnham stations
C5 services only stop here in peak time

Liverpool Street station
C7 services depart from the overground main line terminus

Key to symbols

- Change for Underground, Overground, DLR
- Change for National Rail
- Change for Heathrow Express
- † For this station, see notes
- Services only operate at peak time

Services

		Timetable Peak	Timetable Off-peak
C1	Abbey Wood – Heathrow	4	4
C2	Abbey Wood – Paddington	6	4
C3	Shenfield – Paddington	8	4
C4	Shenfield – Reading	2	2
C5	Shenfield – Maidenhead	2	2
C6	Abbey Wood – West Drayton	2	–
C7	Gidea Park – Liverpool St	4	–

trains per hour

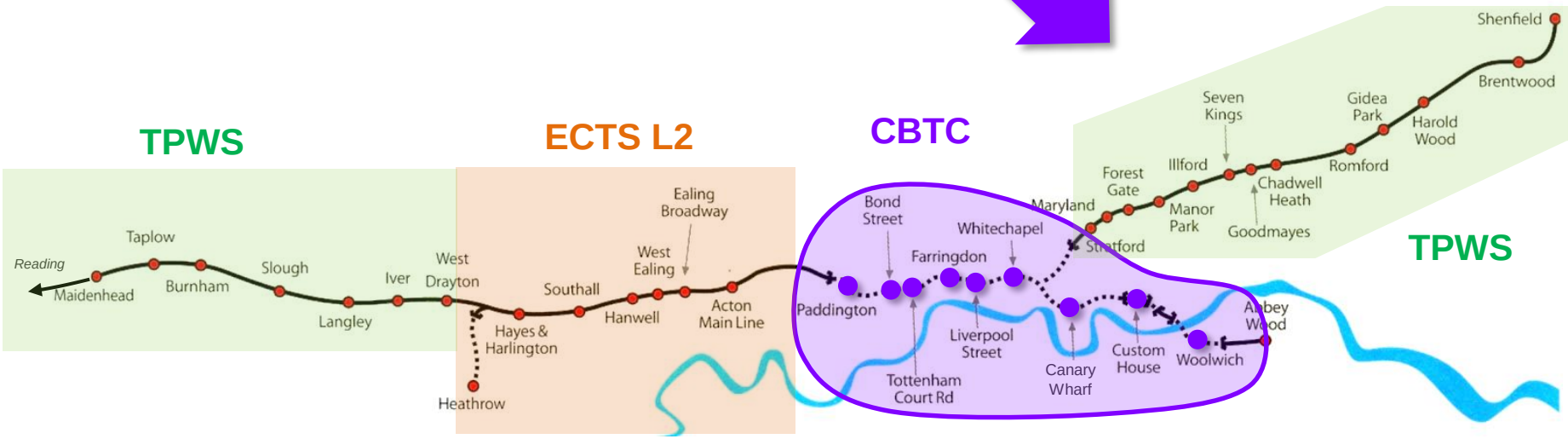
Occasional peak only services:
C13 Shenfield – Heathrow
C36 Shenfield – West Drayton
C24 Abbey Wood – Reading
C25 Abbey Wood – Maidenhead

TRANSPORT FOR LONDON
EVERY JOURNEY MATTERS

Crossrail services depicted as planed by Transport for London, line numbers are made up though.
This is not an official map. It was not authorized and is not affiliated with Transport for London, National Rail or any other transport company or organisation.

3 signalling systems to be manage, a sequential opening in 4 stages.

December 2019 | May 2018 | December 2018 | May 2017

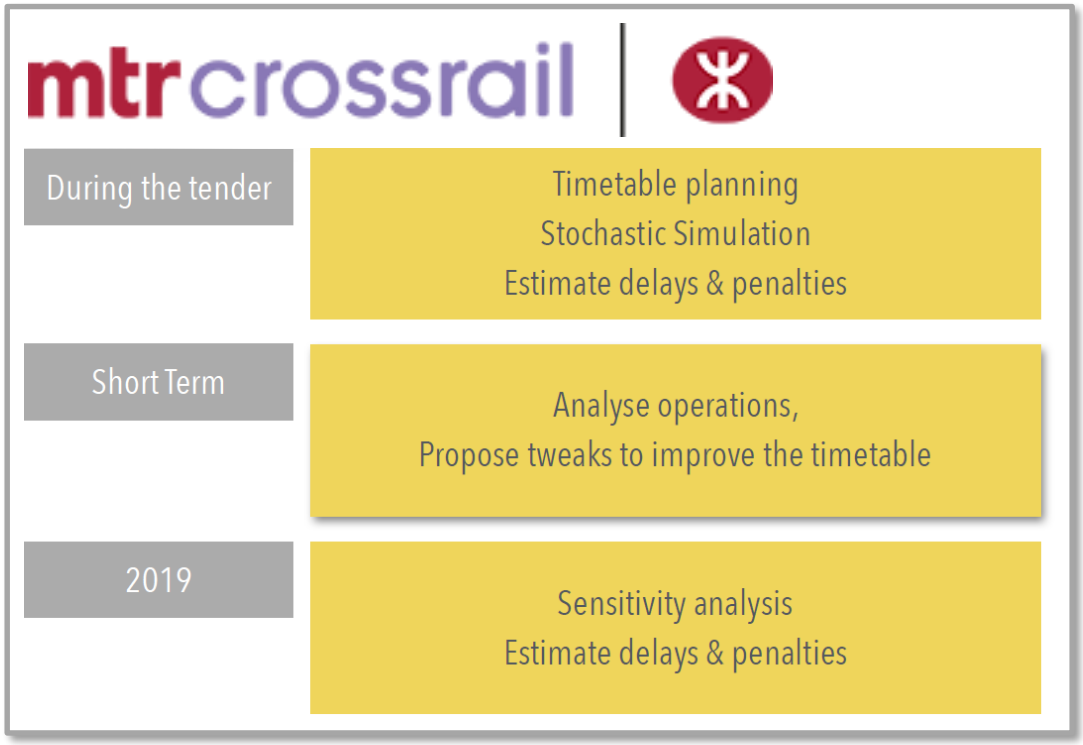


How to ensure the best results and perfect operations for such a big and complex project?

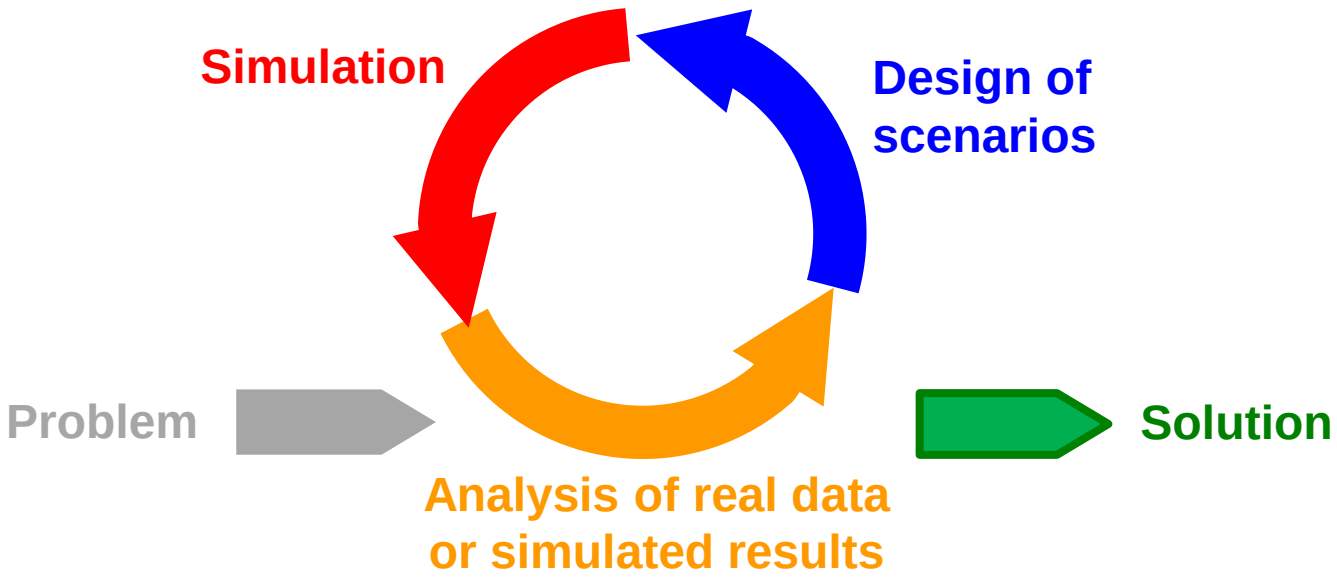
MTR – Metro Hong Kong decided to use a dynamic simulation from the early tendering stage. To goal was to reduce the risks for the company ensuring a better quality of service during the future operations. In this way it was possible to estimate (thus reducing possible malus and ensuring a better financial balance).

After winning the concession of Crossrail, MTR decided to continue to realize dynamic simulations in order to :

- Improve the quality of service on the first routes (on the existing tracks)
- better prepare the future opening of the central section.



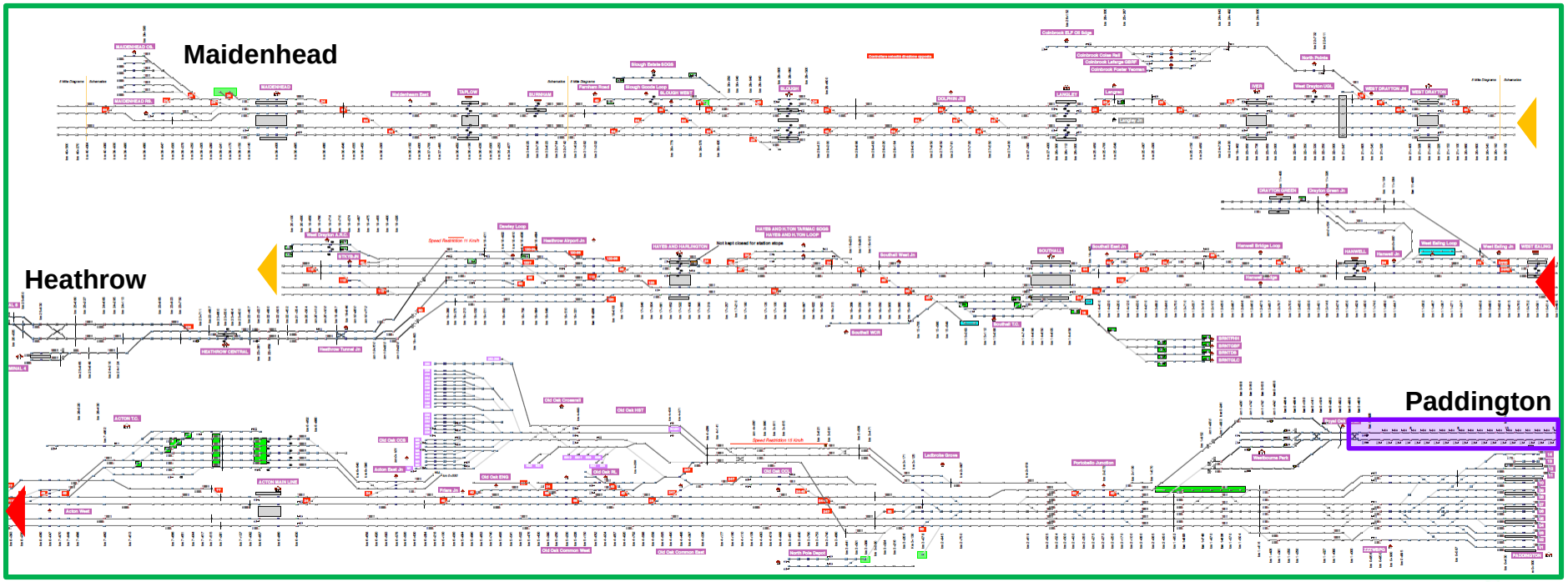
The approach for the simulation with the “planning loop” and the simulation model for Crossrail

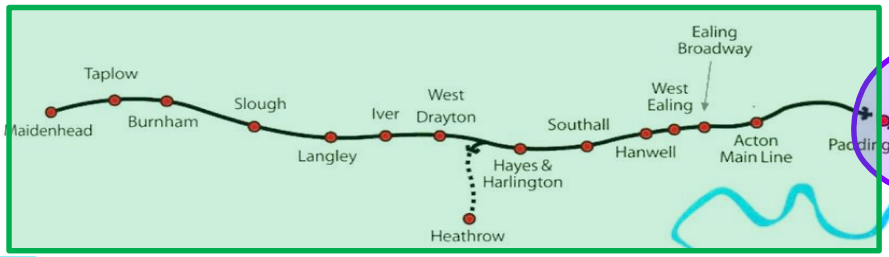


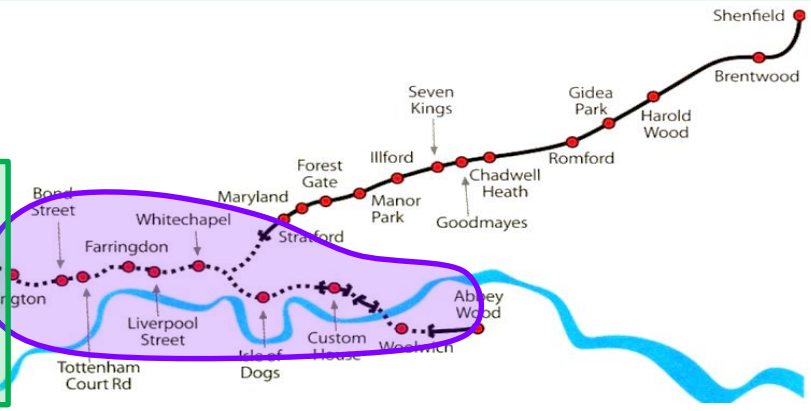
The **rail planning** can be viewed as a **loop** (*see also alius presentation for the Rail Conference 2017*), which starts from the **problem definition** and whose elements are the **analysis of real data**, the **design of scenarios**, the **simulation** and the **analysis of simulated results**.

The loop is the ideal support to timetable planning, which is becoming more demanding since punctuality has become fundamental since delays lead to higher costs and lower customer satisfaction.

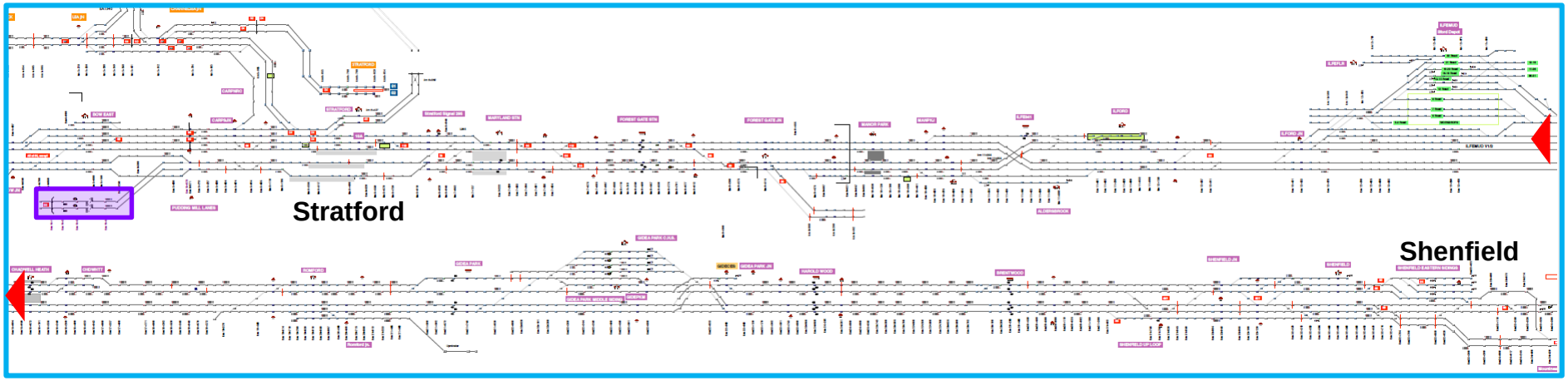
The approach for the simulation with the “planning loop” and the simulation model for Crossrail

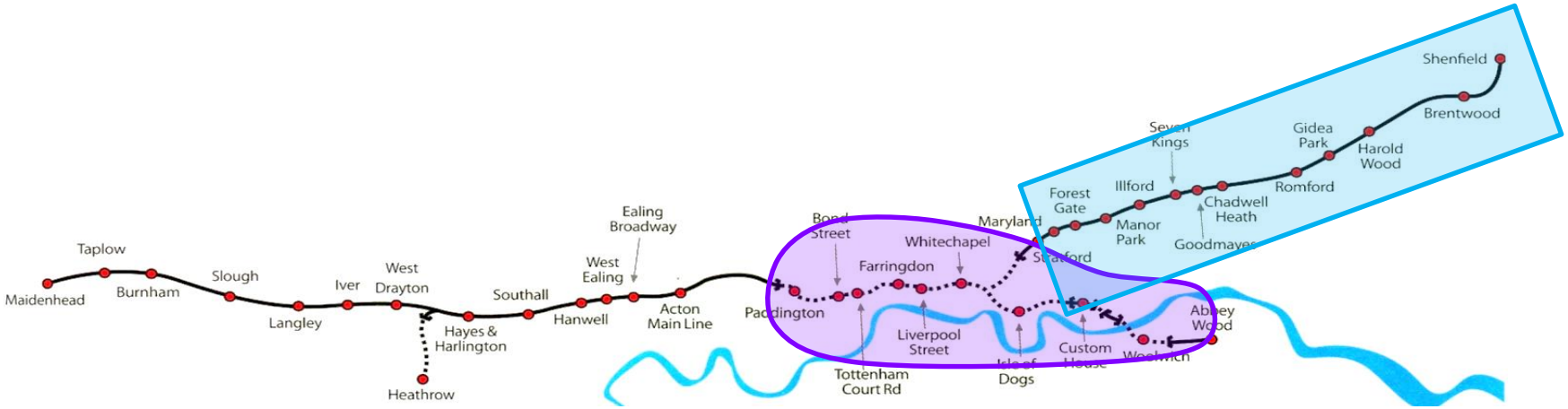






The approach for the simulation with the “planning loop” and the simulation model for Crossrail

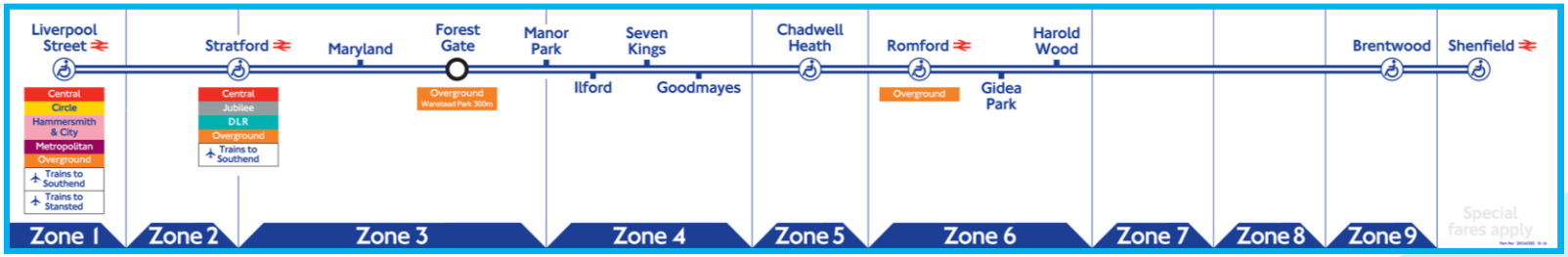




How to use Big Data to improve the operations ? The example of the Crossrail “Stage 1” (May 2017)

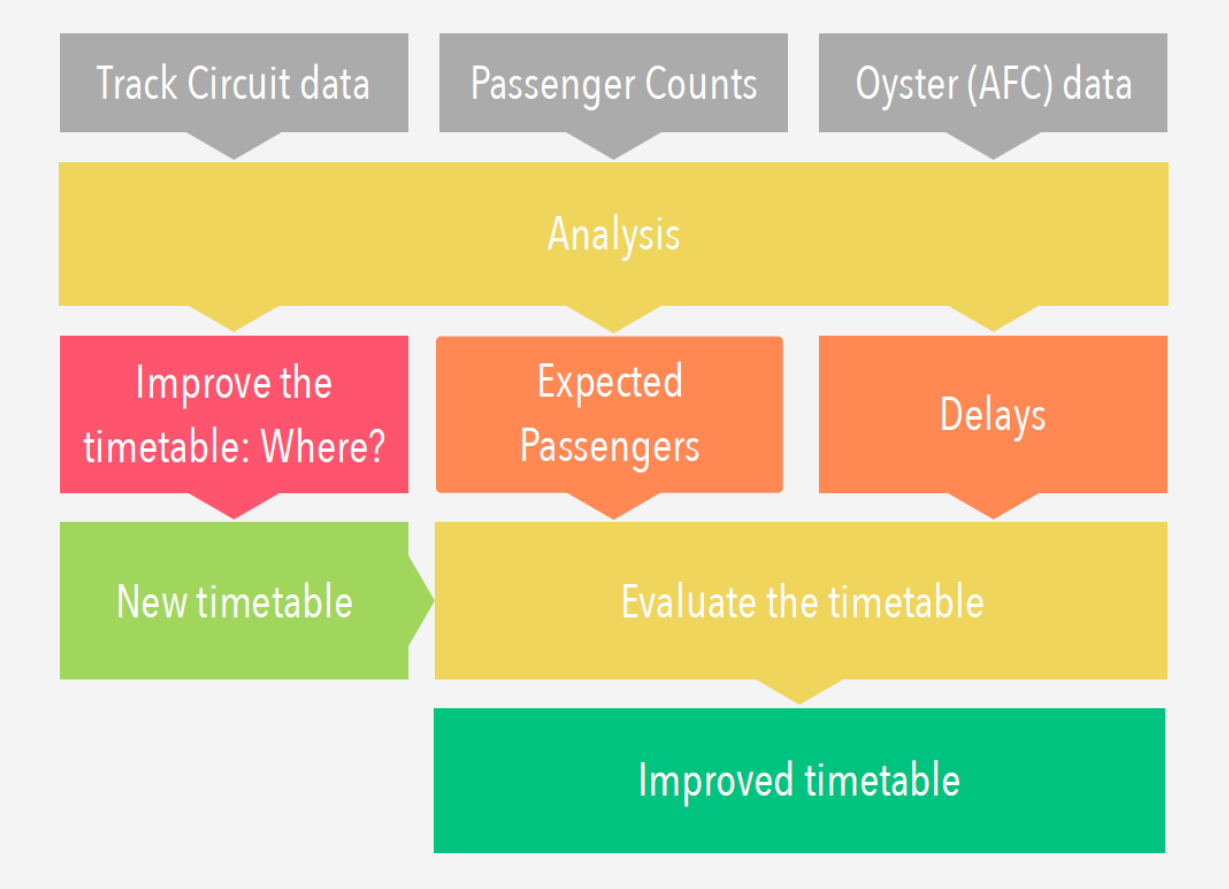
Goal for MTR in 2017 :

- ❑ **increase the quality of operations** on the East part of the network
- ❑ **reduce the risks for December 2018** (propagation of delays from the classic national network in the Crossrail tunnel)



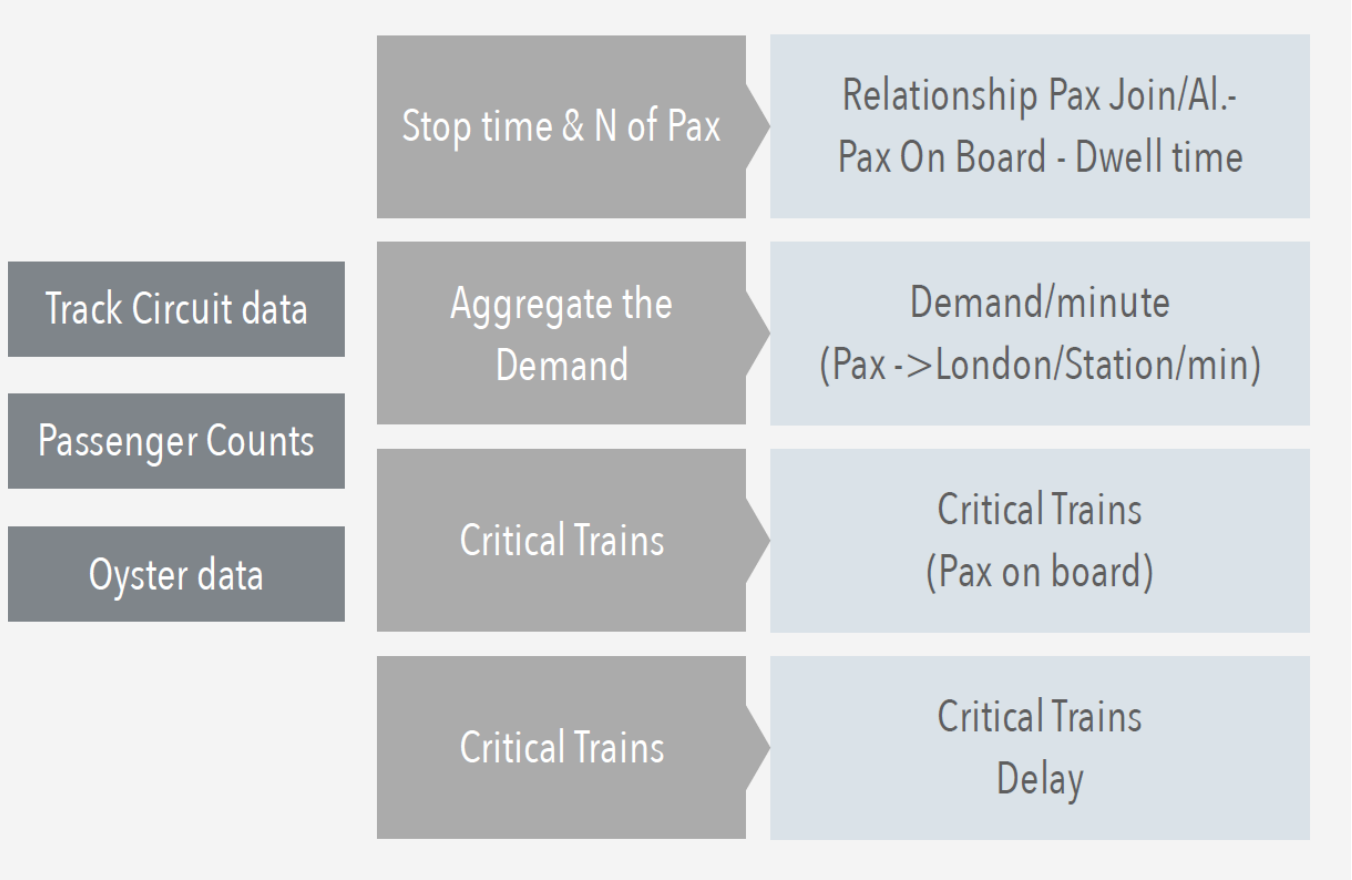
How to use Big Data to improve the operations ? Workflow: main concept

Combination of stochastic simulation, analysis of train data, analysis of pax counts and smart card data and of course … experience ;-)

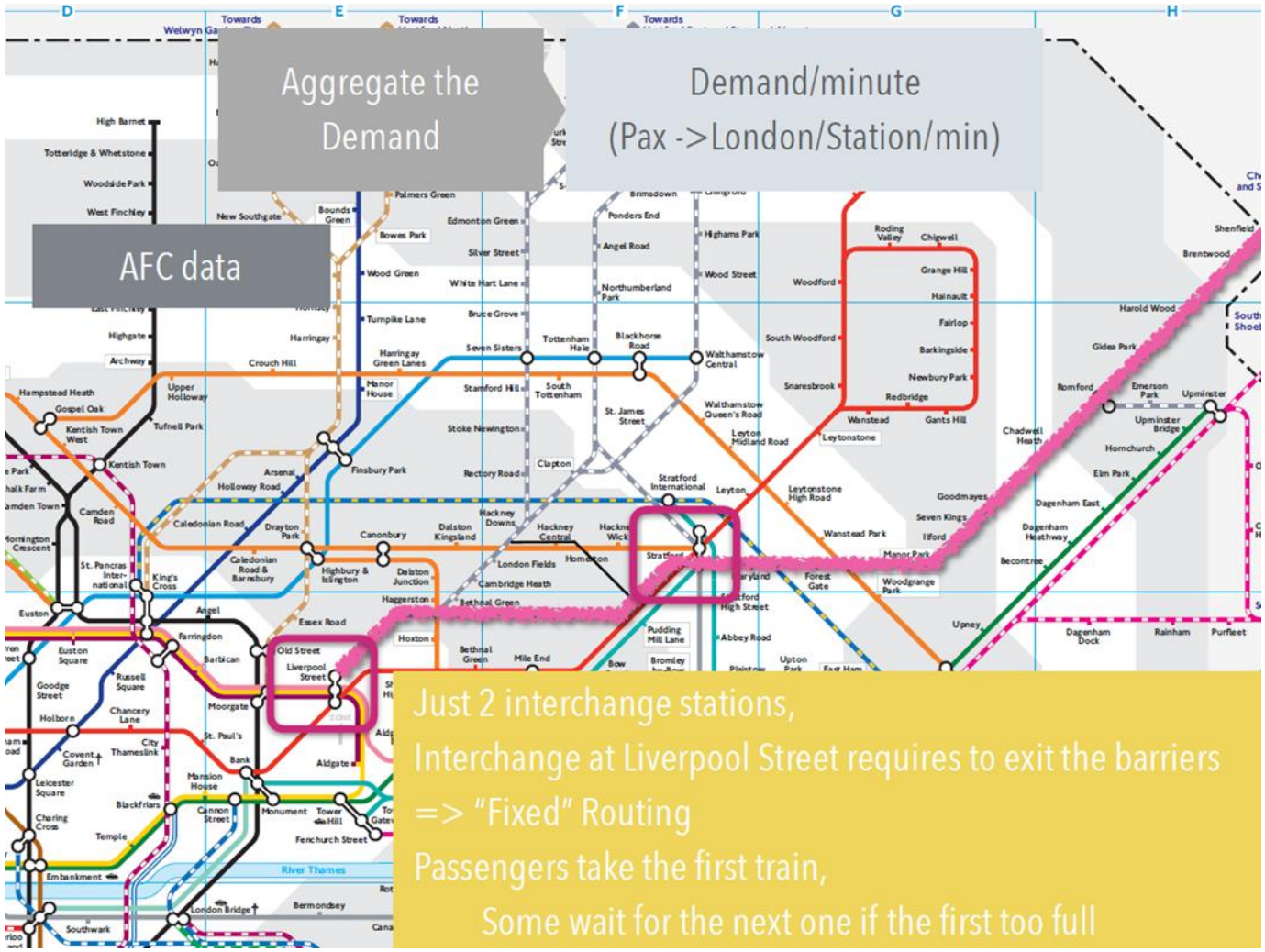


How to use Big Data to improve the operations ? Workflow: input

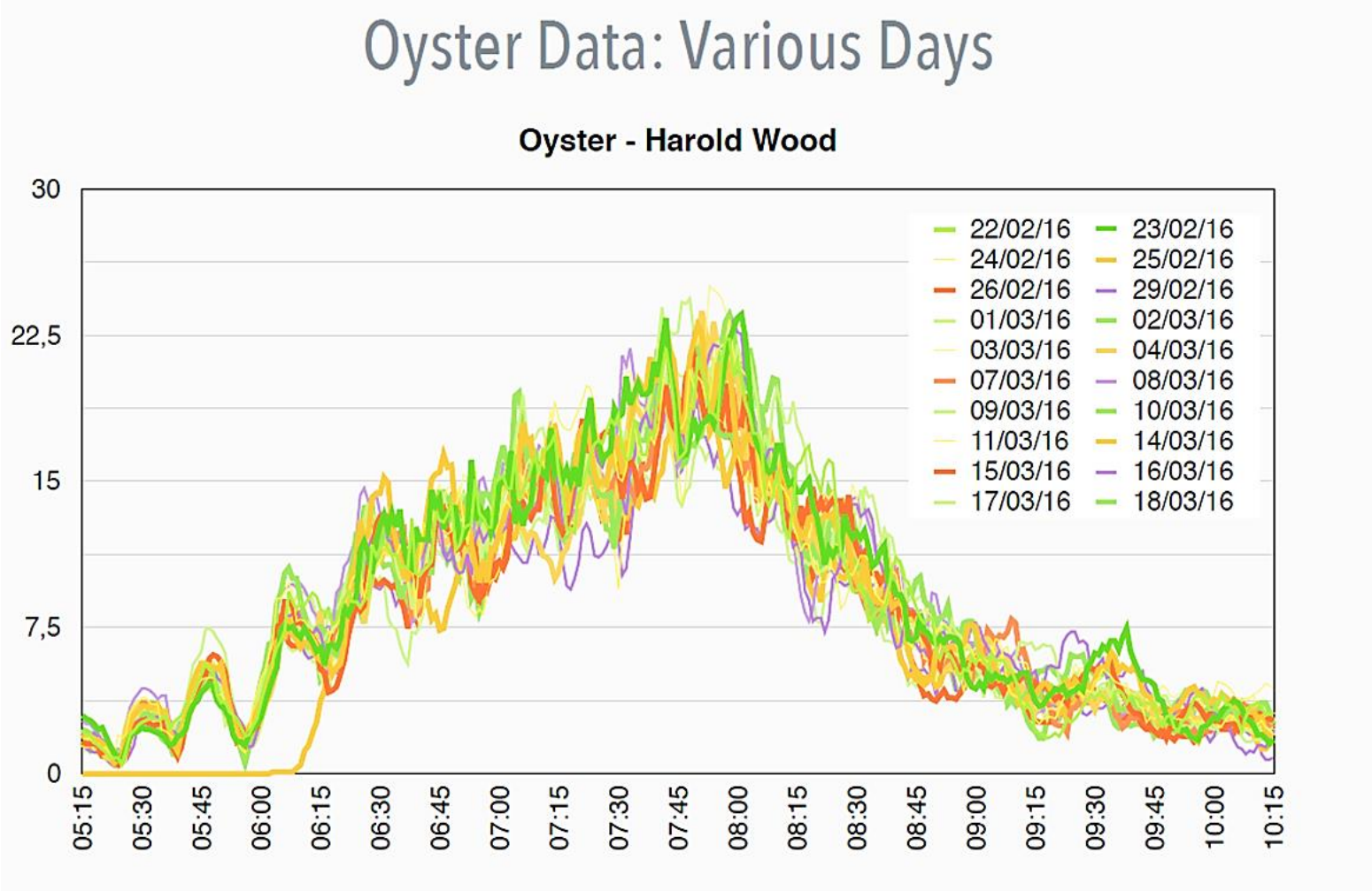
Combination of stochastic simulation, analysis of train data, analysis of pax counts and smart card data, experience



How to use Big Data to improve the operations ?
Utilization of the information about the passengers count



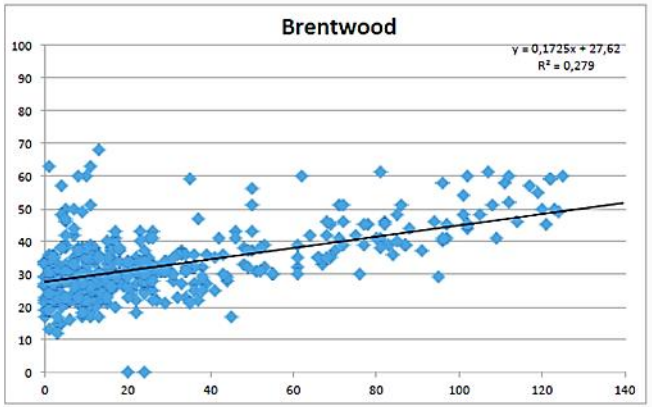
How to use Big Data to improve the operations ?
Utilization of the information about the passengers count

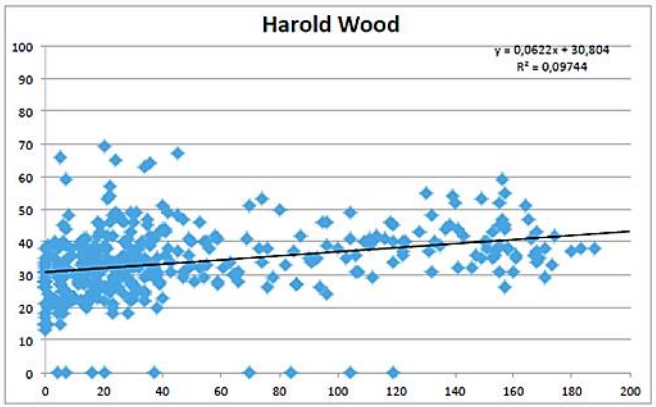


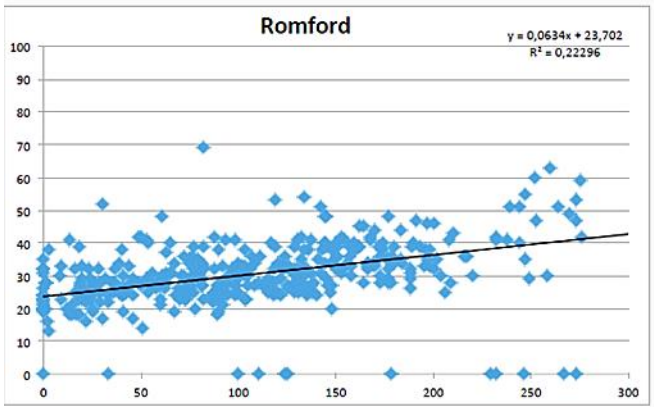
How to use Big Data to improve the operations ?
Utilization of the information about the passengers count

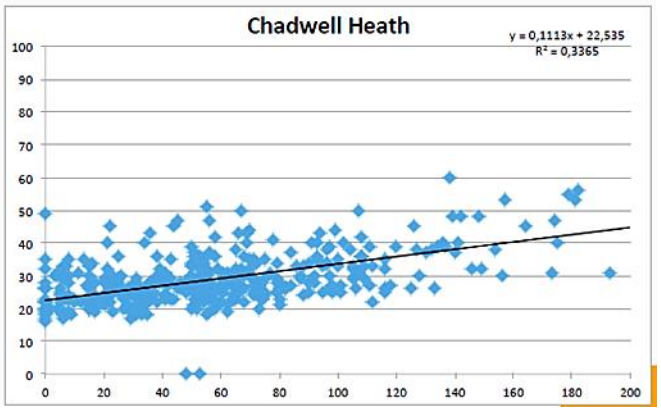
Stop time & N of Pax

Relationship Pax Join/Al.-
Pax On Board - Dwell time





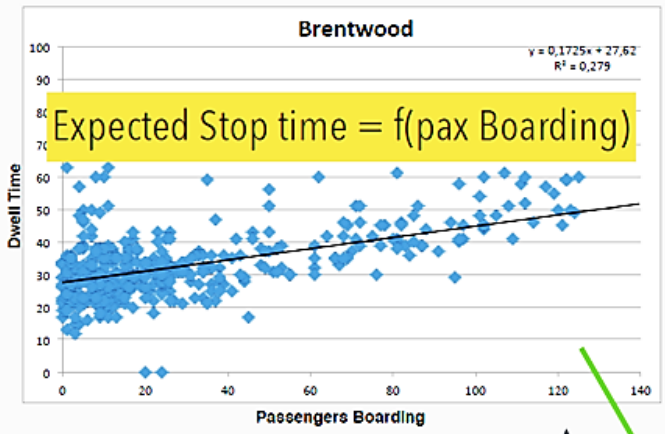




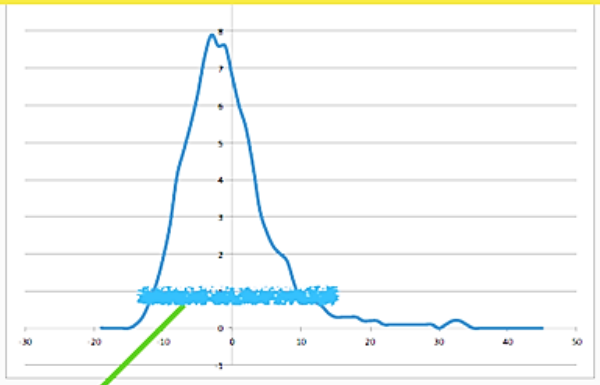
How to use Big Data to improve the operations ?

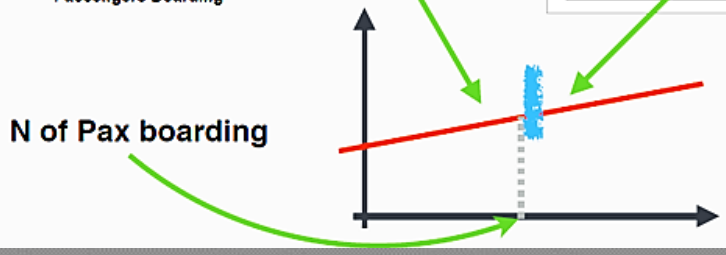
Stop time & N of Pax

Relationship Pax Join/Al.-
Pax On Board - Dwell time



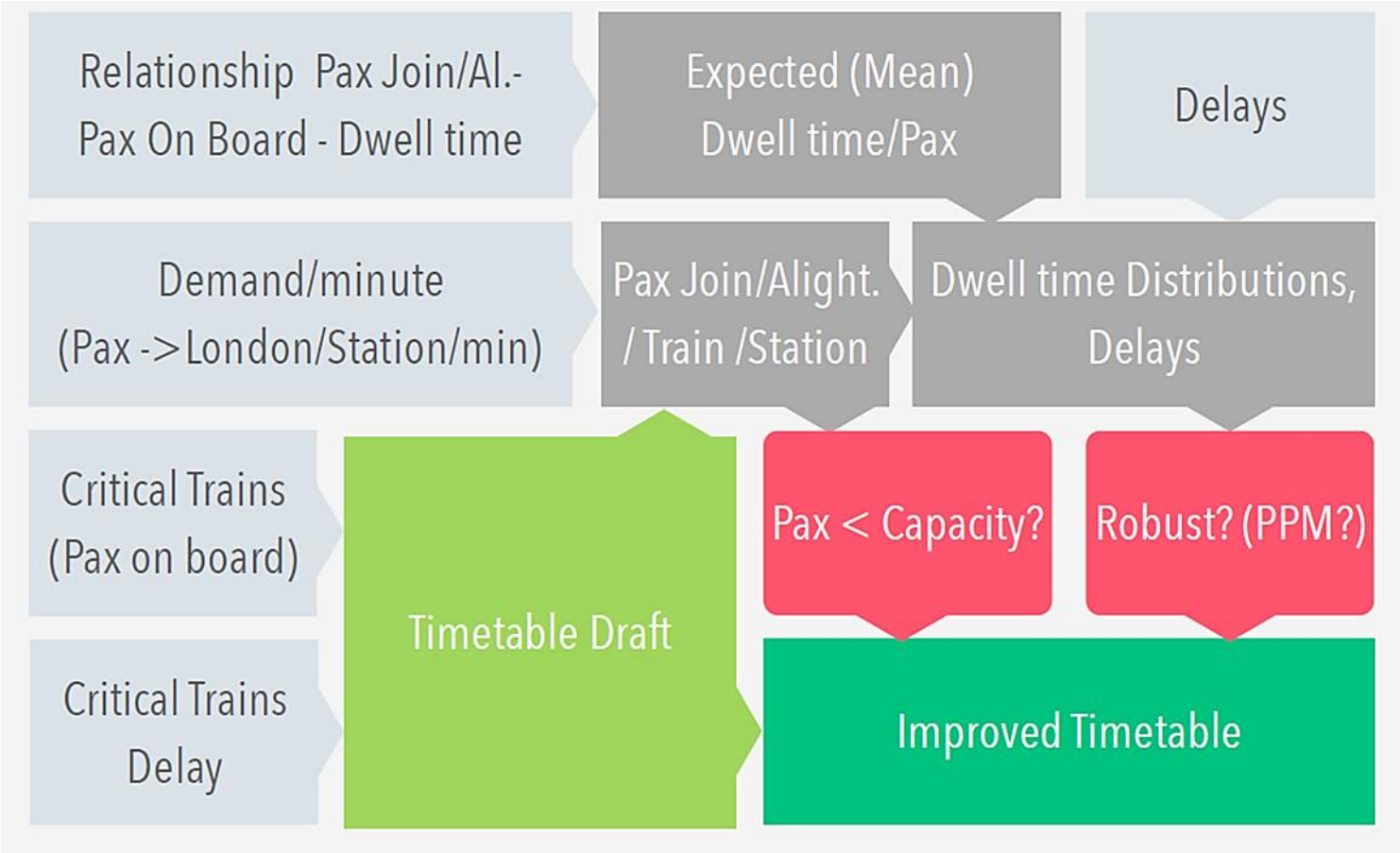
Actual stop Time = logn(Expected Stop time)





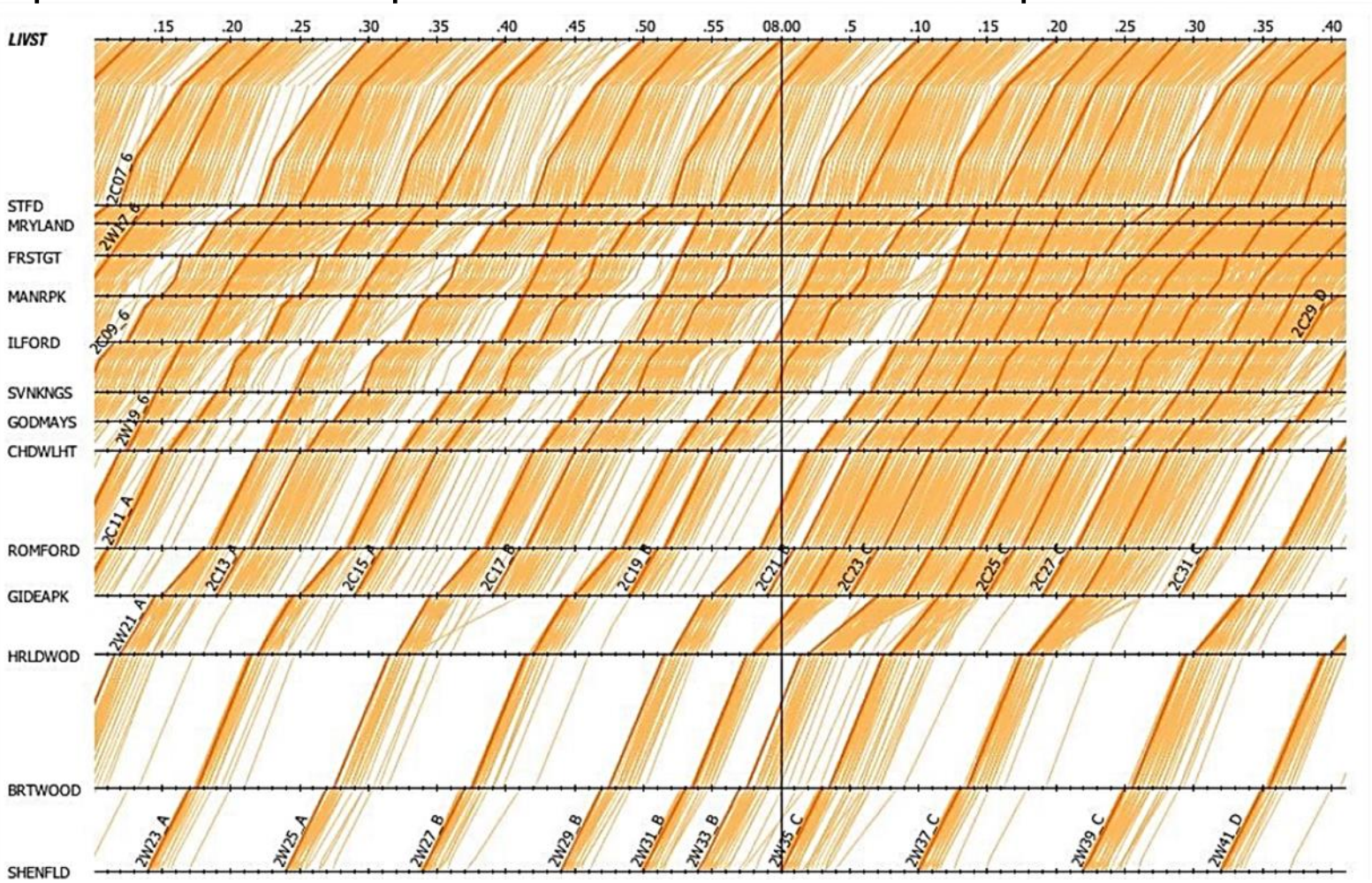
The relationships obtained for each station and displayed at the previous slide are used for the simulation.
The line obtained is combined with a log-normal distribution representing the deviation from this deterministic mean

How to use Big Data to improve the operations ? Workflow



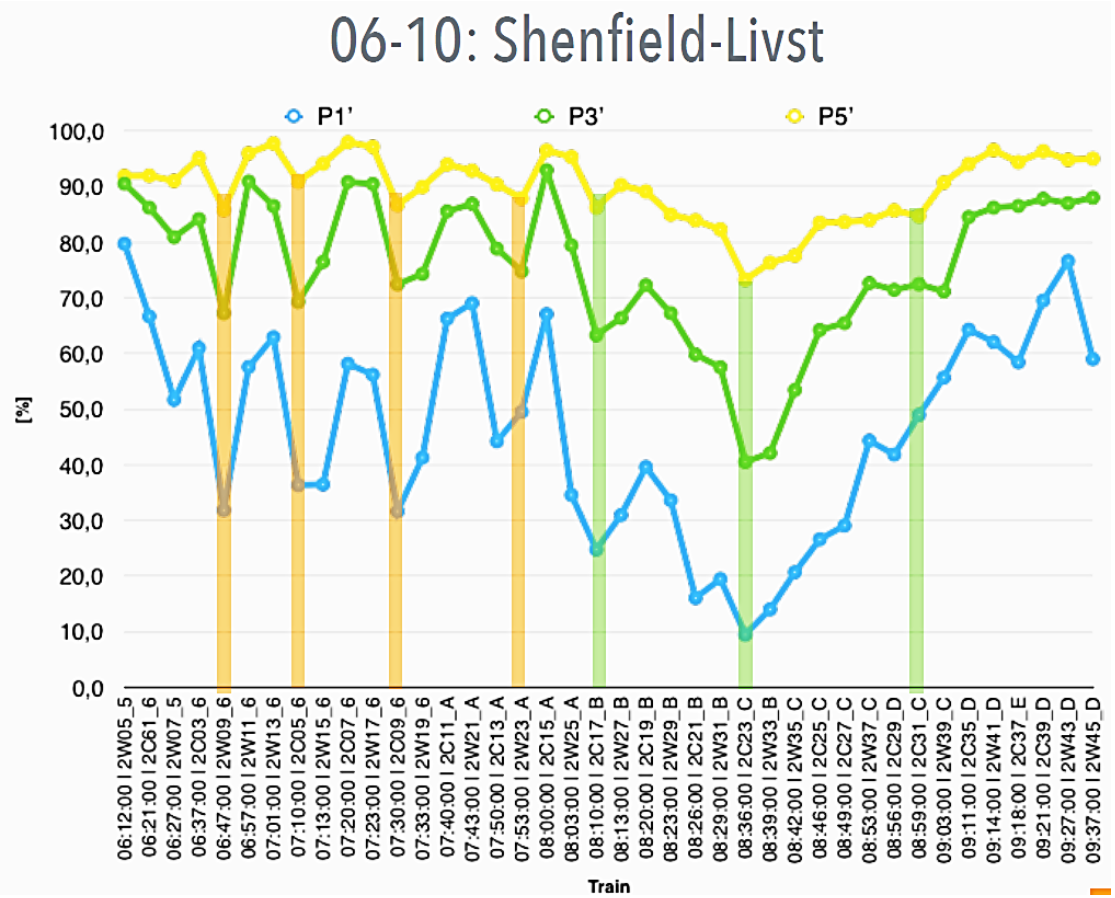
How to use Big Data to improve the operations ? Analysis of real operation data (berth level data)

Comparison between planned timetable et real occupation

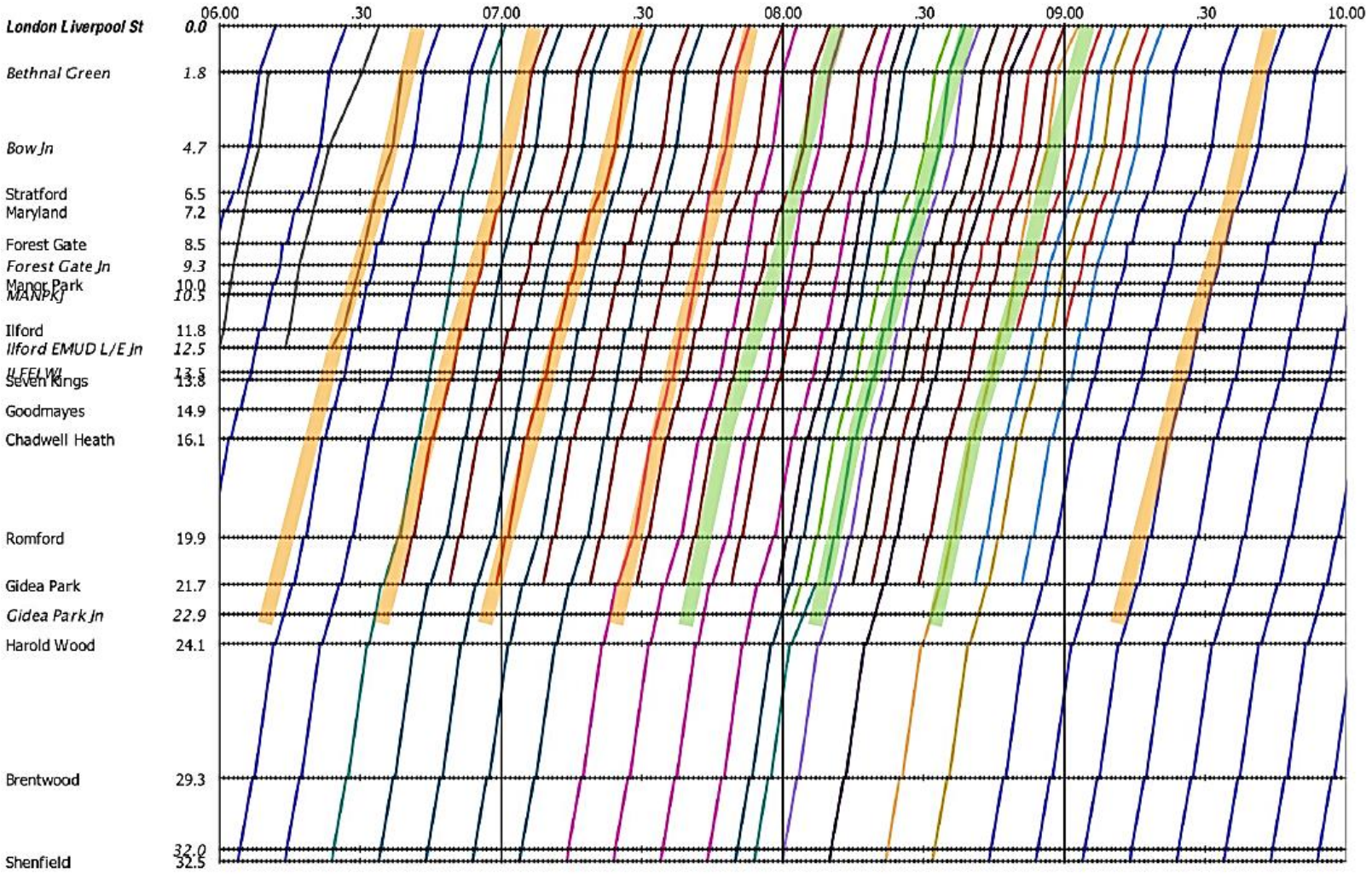


How to use Big Data to improve the operations ? Analysis of real operation data (berth level data)

Punctuality (1', 3' and 5' minutes delays)in a station during the morning peak hours

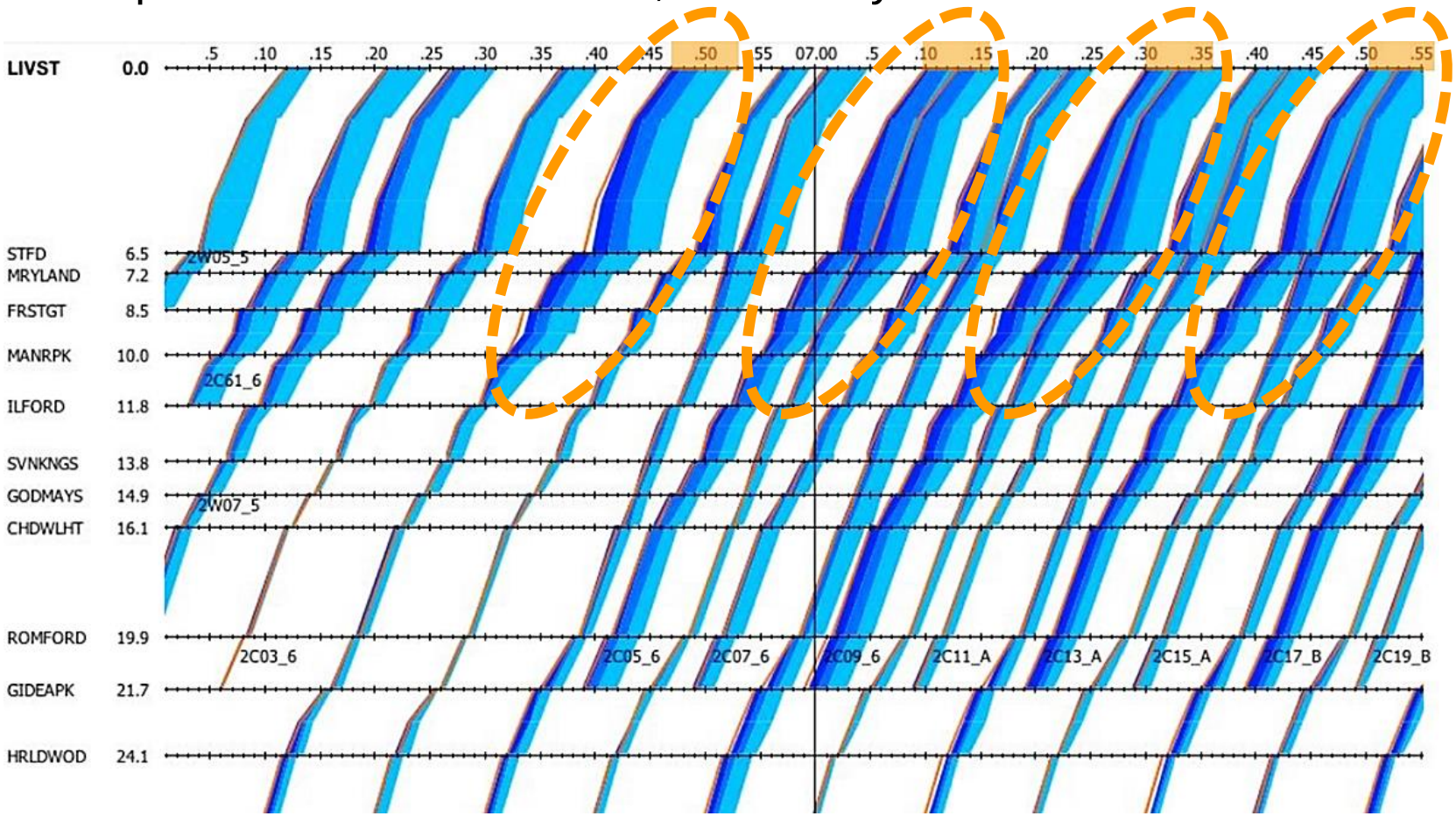


How to use Big Data to improve the operations ? Identification of the most critical trains in the timetable



How to use Big Data to improve the operations ? Identification of the most critical trains in the timetable

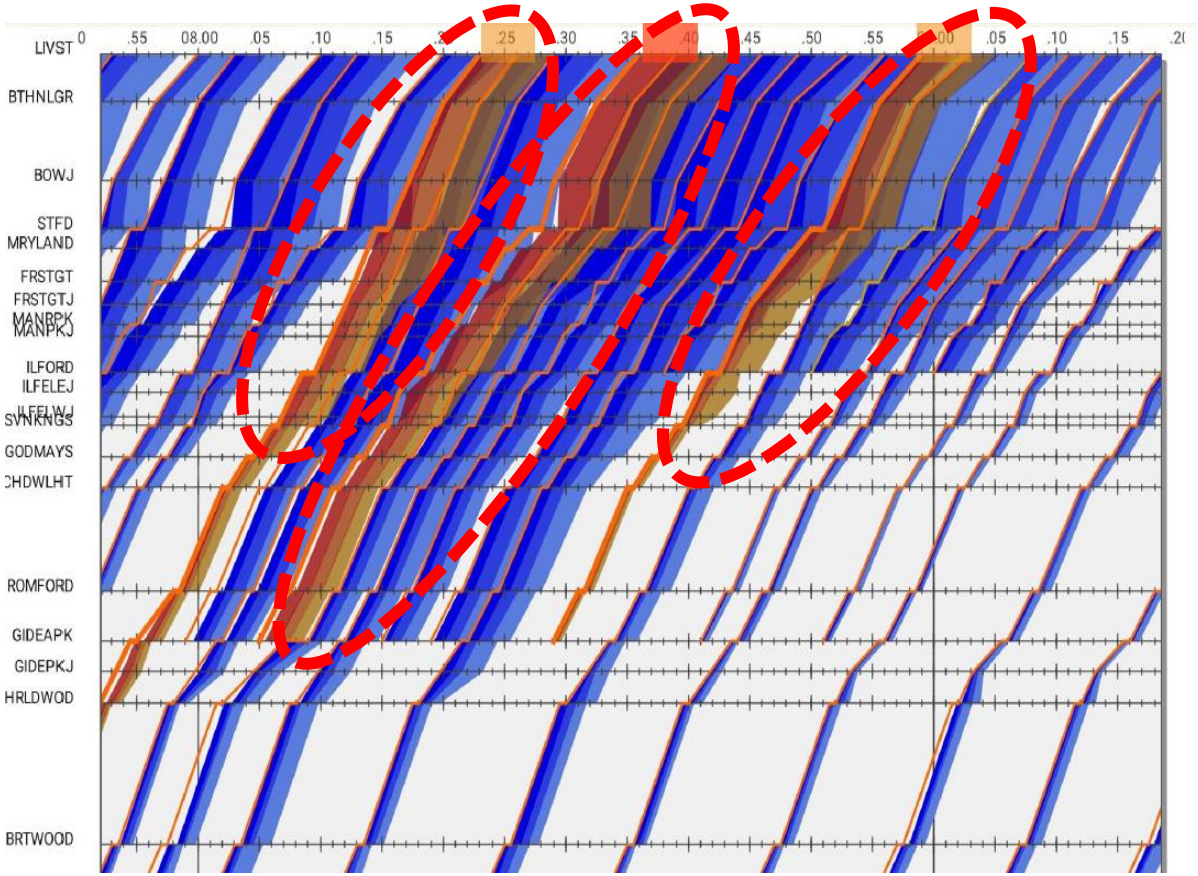
Analysis of the actual data between **06:30 and 07:30** :
late departure from GIDEA Park, then delays not recovered



How to use Big Data to improve the operations ? Identification of the most critical trains in the timetable

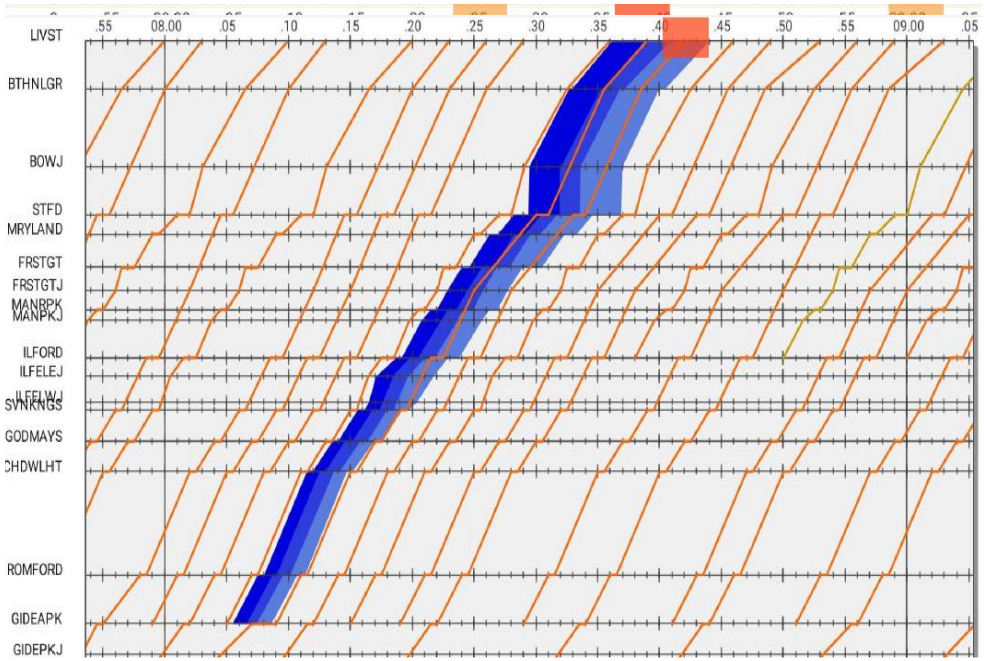
Analysis of the actual data between **08:00 and 09:00** :

Area of nearly 50% of the trains overlaps with the following one



How to use Big Data to improve the operations ? Identification of the most critical trains in the timetable

Analysis of the actual data between **08:00 and 09:00** :
Detail of a very critical train departing from GIDEA Park

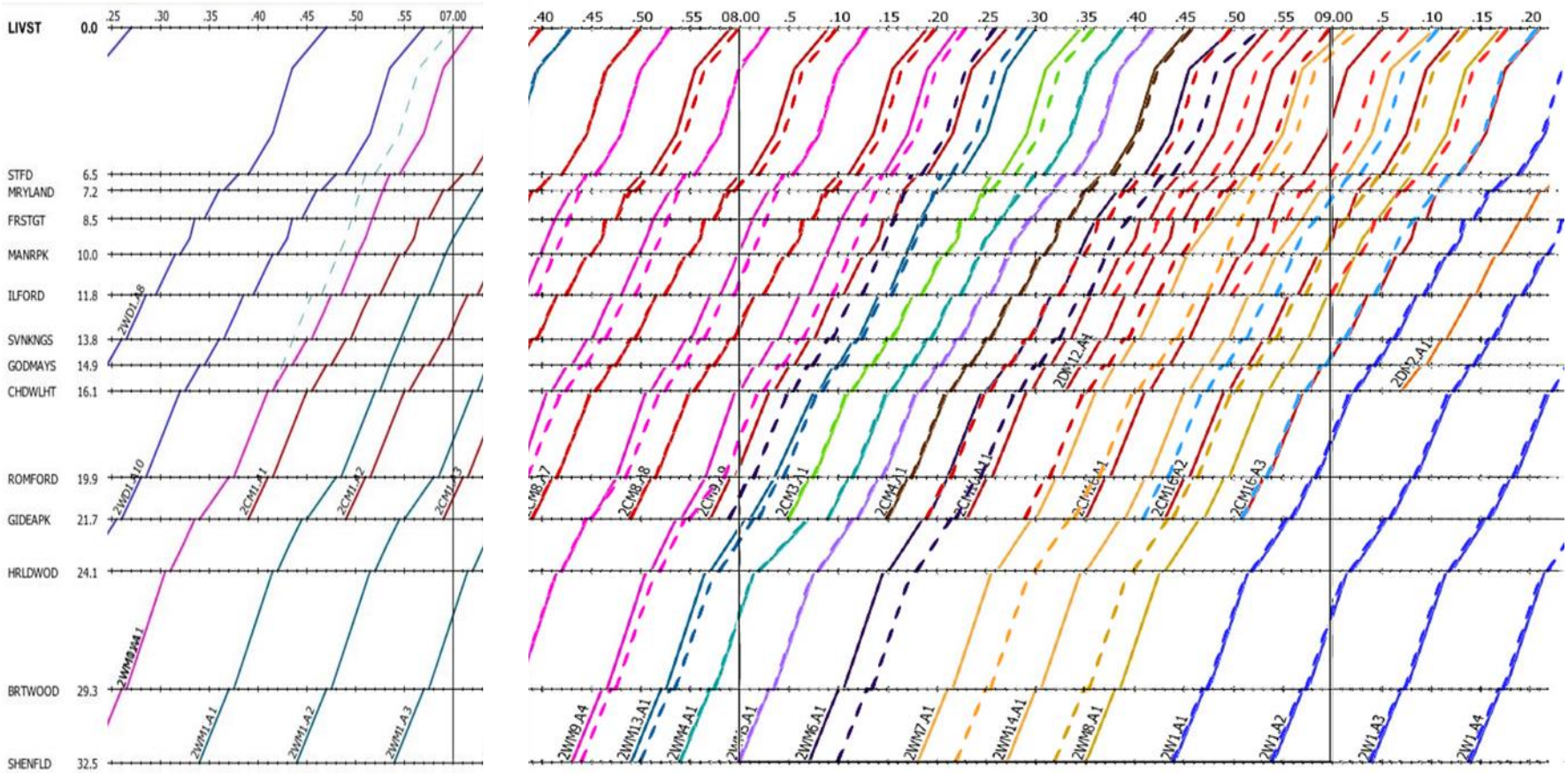


The trains **start to late** and the **running/stop time is not sufficient**.

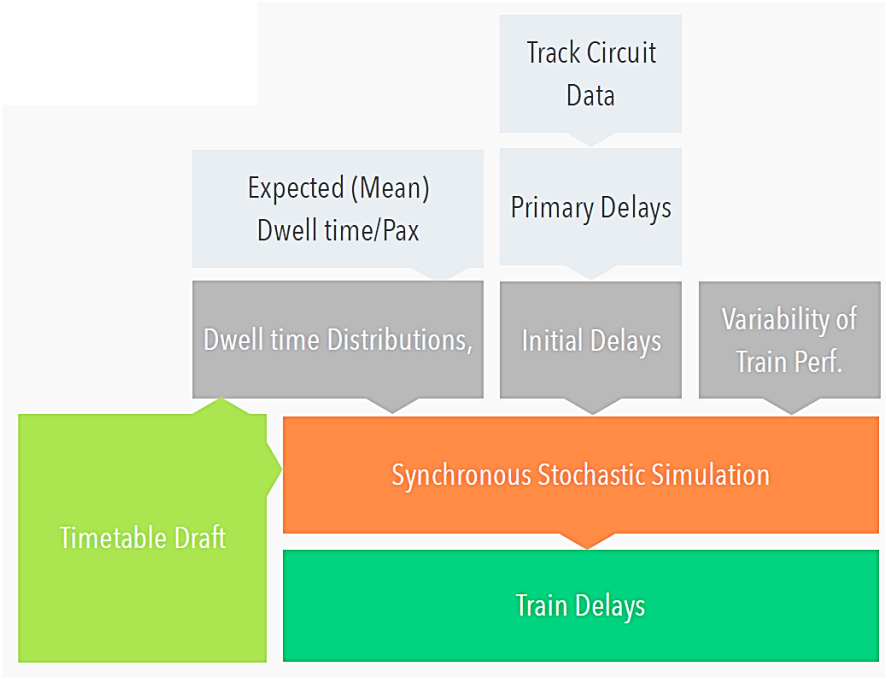
How to use Big Data to improve the operations ?
Proposed timetable improvements

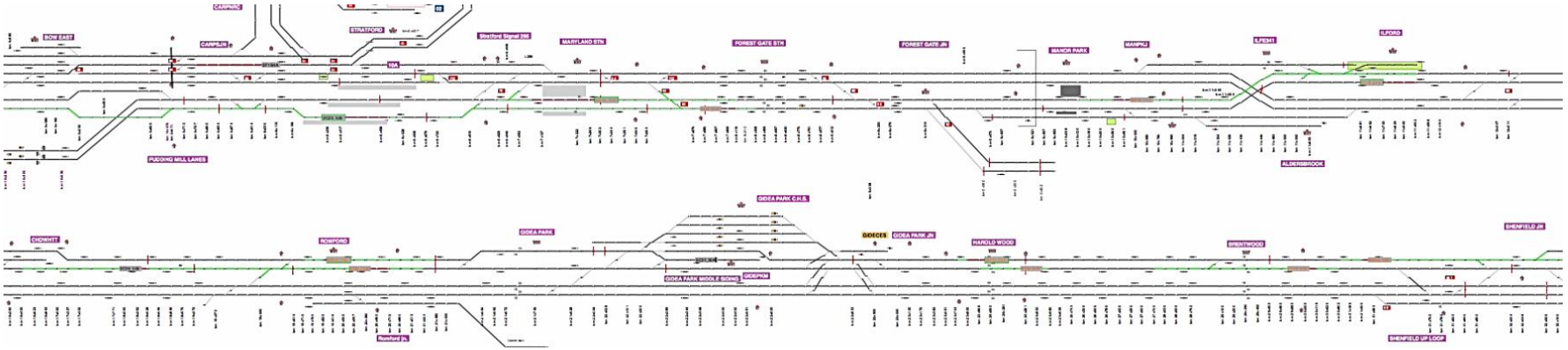
- - - - - December 2015

——— Proposal

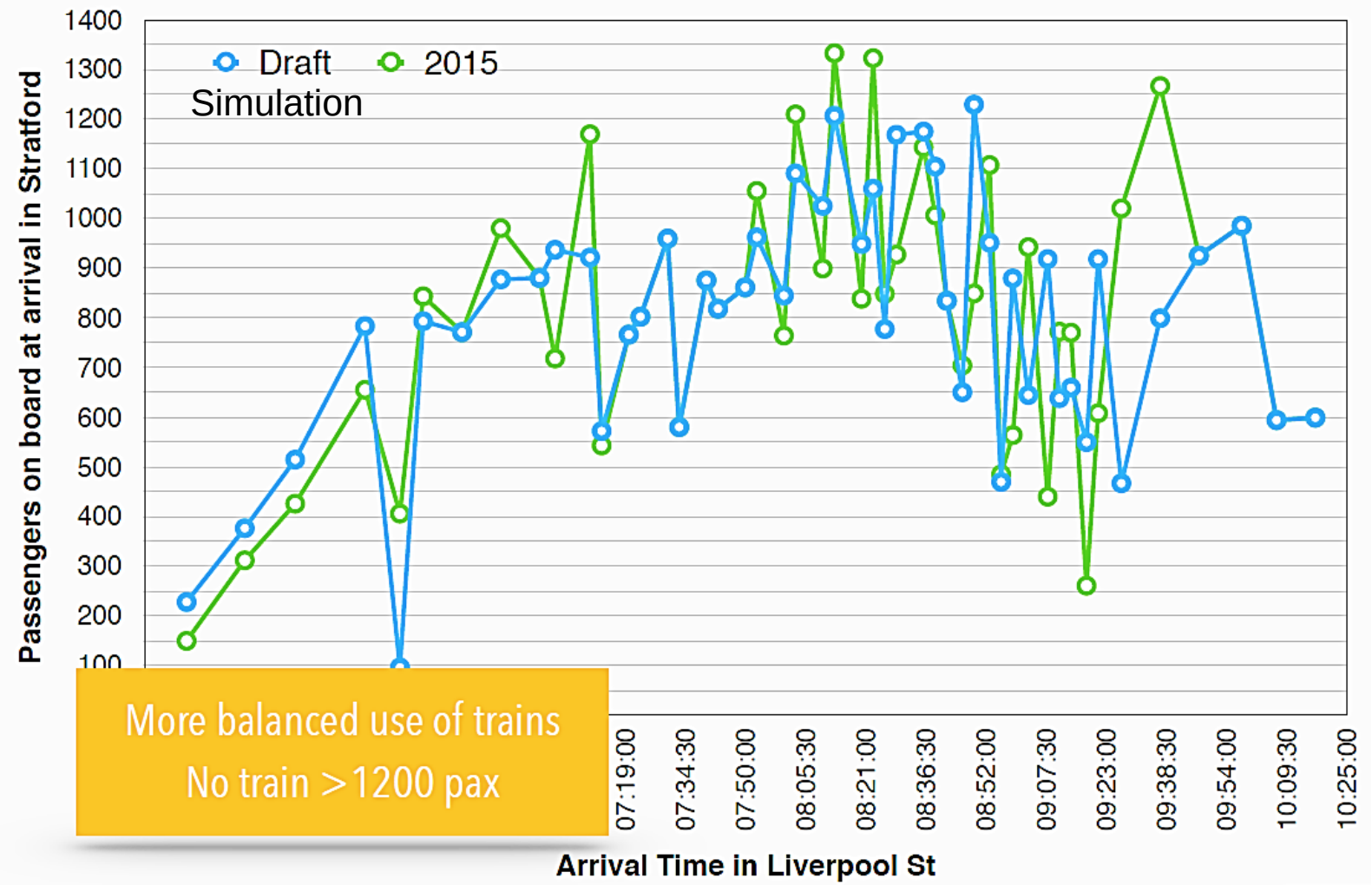


How to use Big Data to improve the operations ? Estimation of the results with the simulation





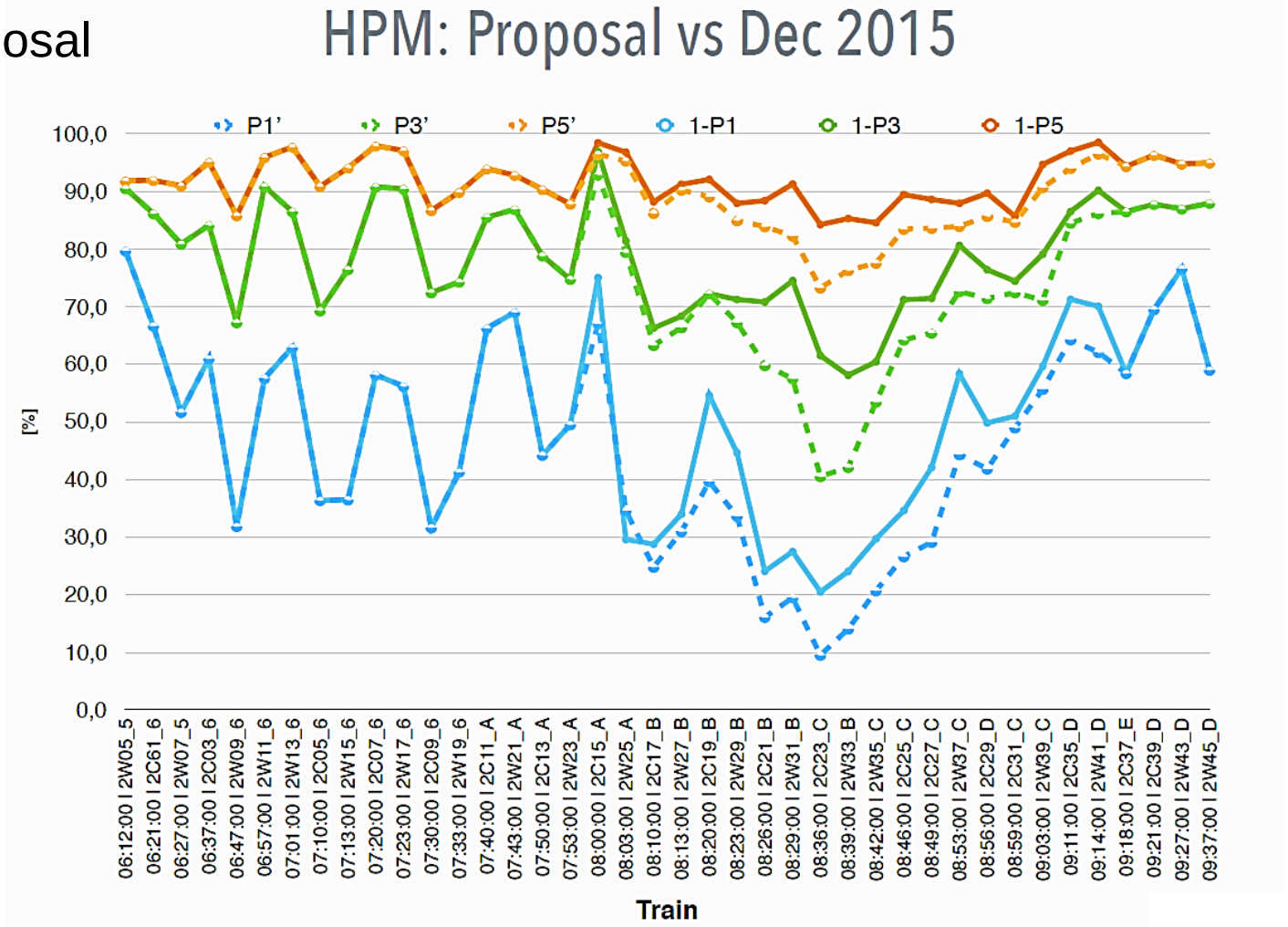
How to use Big Data to improve the operations ? Estimation of the results with the simulation



How to use Big Data to improve the operations ? Estimation of the results with the simulation

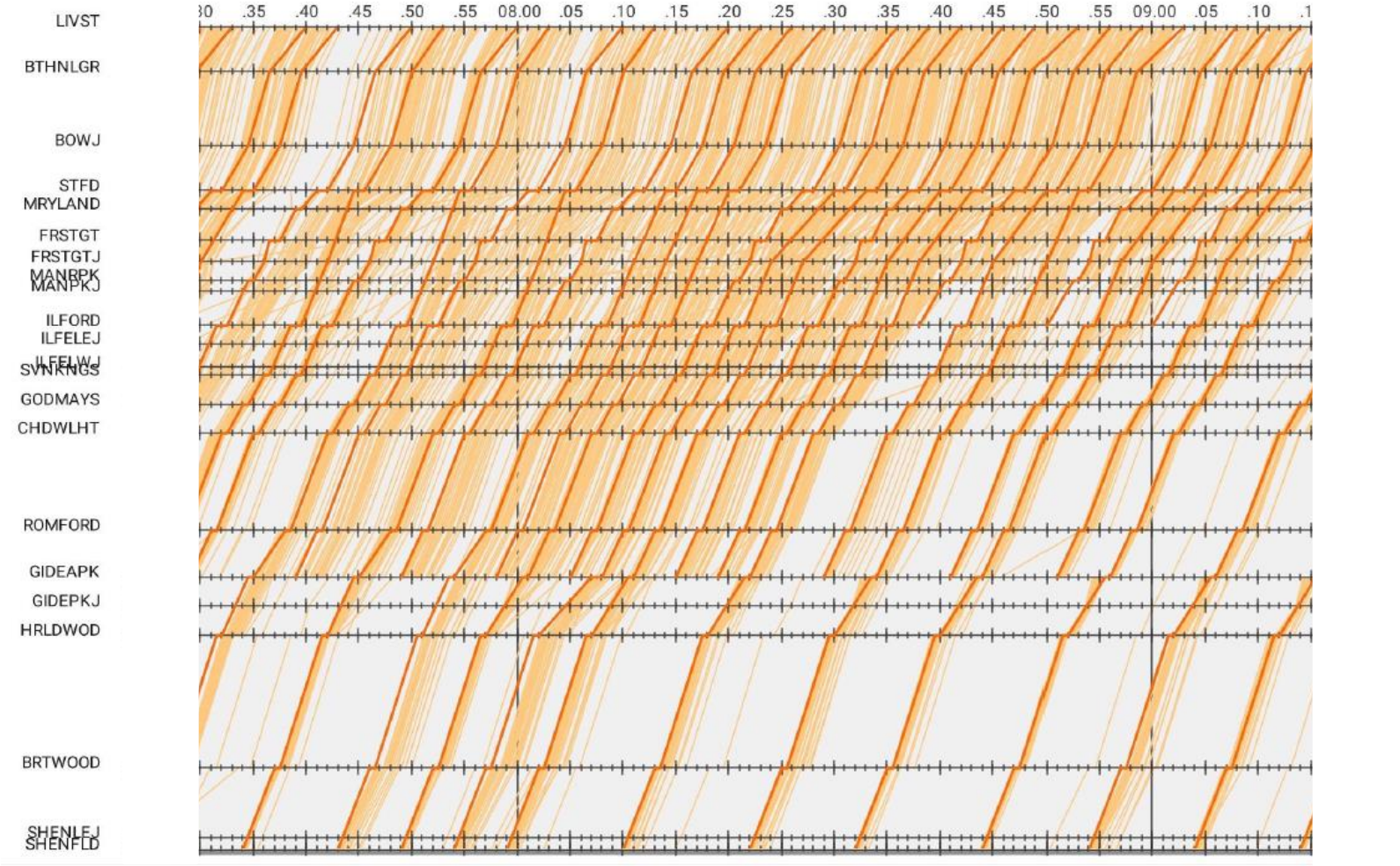
December 2015

Proposal



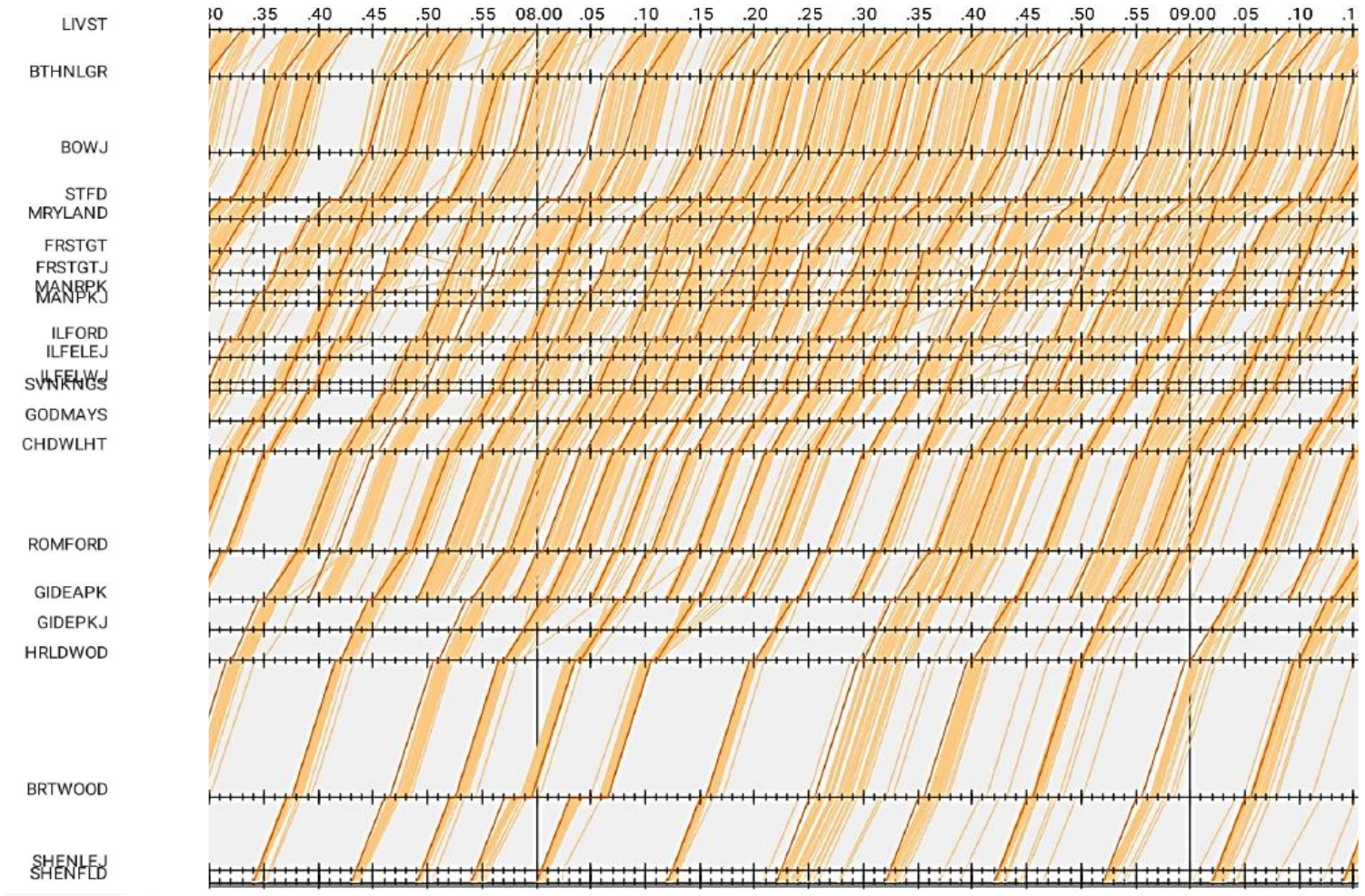
How to use Big Data to improve the operations ?
Ex-post verification : April 2017

Comparison between planned timetable and real operation



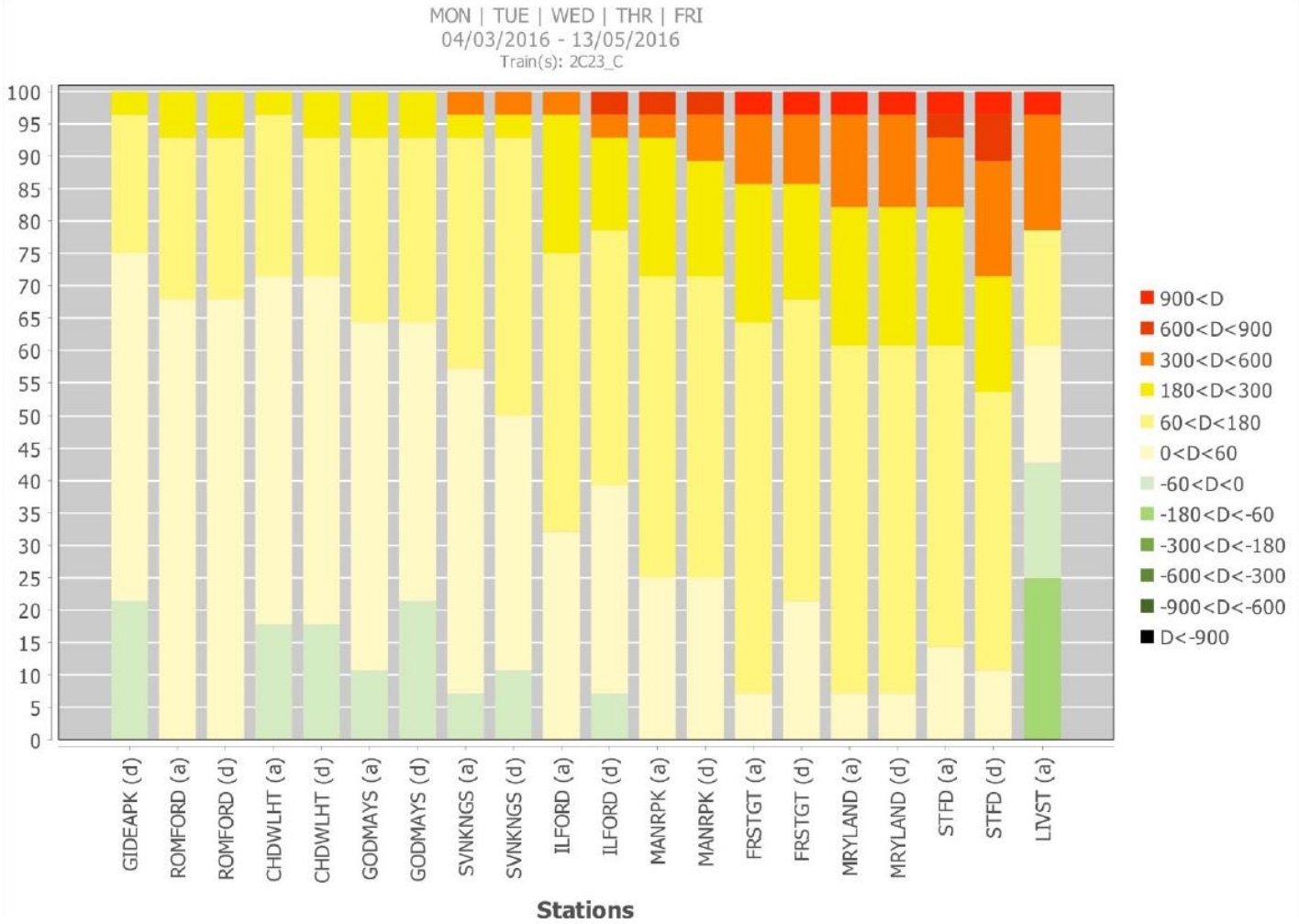
How to use Big Data to improve the operations ?
Ex-post verification : May 2017

Comparison between planned timetable and real operation



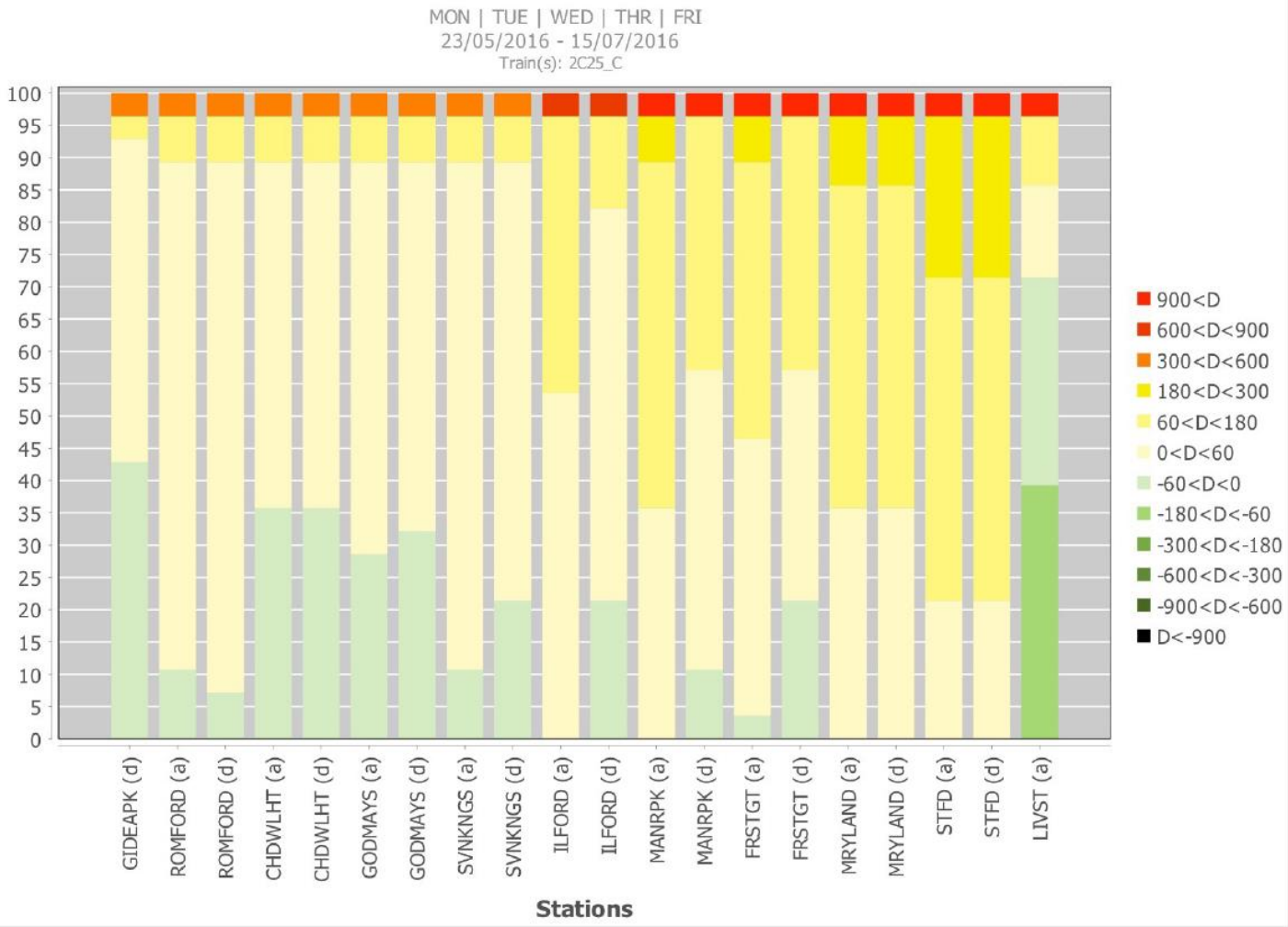
How to use Big Data to improve the operations ?
Ex-post verification : April 2017

Evolution of the performance along the line : train arriving an 08:35



How to use Big Data to improve the operations ?
Ex-post verification : May 2017

Evolution of the performance along the line : train arriving an 08:35



How to use Big Data to improve the operations ? Conclusion

- Efficient and effective utilization on a synchronic dynamic simulation
- A good calibration of the model based on real data is essential to ensure good results
- The utilization of passenger data permit to ensure a better evaluation of the new dwell time at the stations.
  This value is very important in a very dense and saturated network.
- The utilization of a powerful new “planning suite” reduce significantly the project time and increase the accuracy of the results.