

ETCS Data Integration

Mark Wedley

Senior Engineer, Digital Transformation.

Introduction

We will look at Data Integration and how we use it for the European Train Control System (ETCS).

This will give you an example of how you can collect and interpret data.

We work with ETCS data to understand incidents and accidents on the railway.

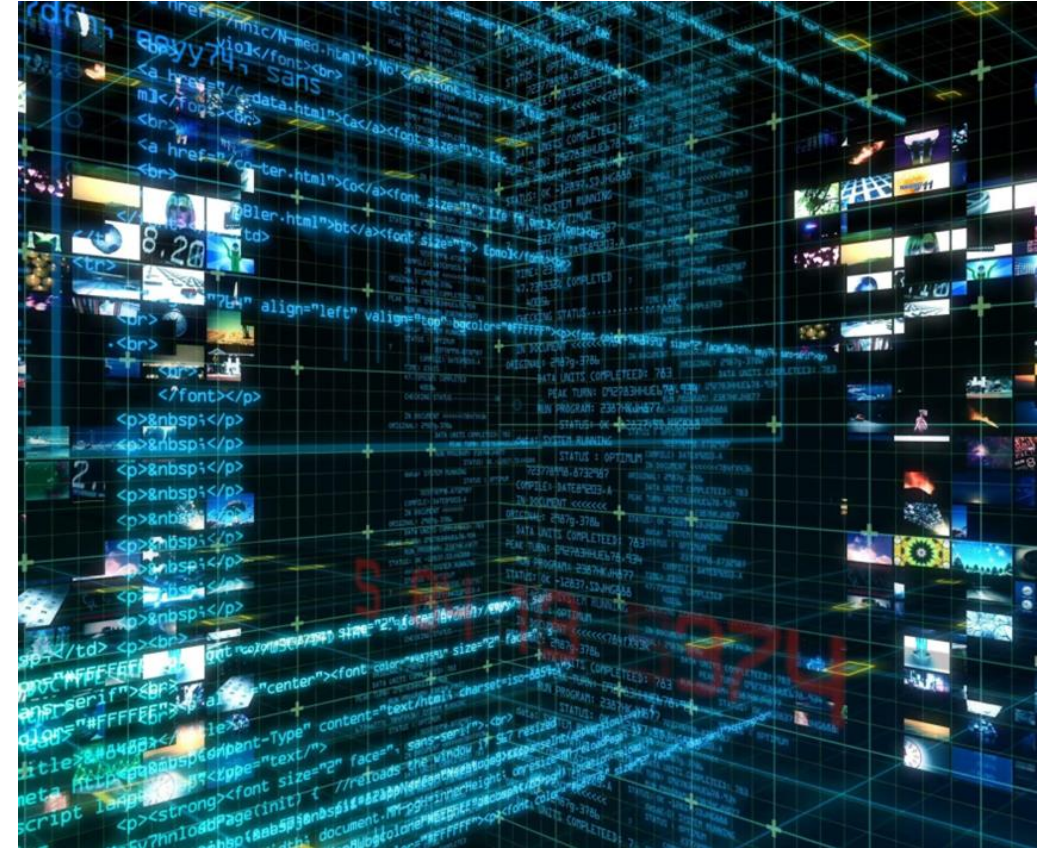


Outline

- Introduction to Data Integration
- Types of Data Integration
- Case Study on ETCS Data Interpretation
- Challenges in Data Integration
- Best practices
- Future Trends
- Conclusion

1: What is Data Integration?

- **Definition:**
- The process of combining data from multiple sources to provide a unified view.
- **Importance:**
- Ensures consistent, accurate, and accessible data for decision-making and analysis.



2. Types of data Integration

Batch

- ETCS Log files are transferred to storage Daily



Real-Time

- ETCS data is recorded in real time



Middleware

- Each ETCS system has software to carry this out



Extract, Transform, Load (ETL)

- Does ETCS do this?



3: Case Study – ETCS Data Interpretation

1	Timestamp	SequenceNo	Interface	Item	Value	LogType	Crc16
2	GeoData 1 & 0						
3	2021-06-28T03:25:07.993	015CE824	RBCEcho	RBC\TLK\Echo	Info	DESC=Train Position Report,NID_MESSAGE=136,L_MESSAGE=25,T_TRAIN=95743026,NID_ENGINE=20794,PACKETS=1,PACKET=1,DESC=Position Report,NID_PACKET=0,L_PACKET=122,Q_SCALE=1,	
4	2021-06-28T03:25:07.993	015CE825	RBCEcho	RBC\TLK\Echo	Info	unset CommonAttributeChangeEvent DB79	
5	2021-06-28T03:25:07.994	015CE826	RBCEcho	RBC\TLK\Echo	Info	DESC=Train Position Report,NID_MESSAGE=136,L_MESSAGE=25,T_TRAIN=95743026,NID_ENGINE=20794,PACKETS=1,PACKET=1,DESC=Position Report,NID_PACKET=0,L_PACKET=122,Q_SCALE=1,	
6	2021-06-28T03:25:07.994	015CE827	RBCEcho	RBC\TLK\Echo	Info	unset CommonAttributeChangeEvent 1457	
7	2021-06-28T03:25:07.999	015CE828	RBCEcho	RBC\TLK\Echo	Info	DESC=Train Position Report,NID_MESSAGE=136,L_MESSAGE=25,T_TRAIN=95743026,NID_ENGINE=20794,PACKETS=1,PACKET=1,DESC=Position Report,NID_PACKET=0,L_PACKET=122,Q_SCALE=1,	
8	2021-06-28T03:25:08.000	015CE829	RBCEcho	RBC\TLK\Echo	Info	unset CommonAttributeChangeEvent F878	
9	2021-06-28T03:25:08.000	015CE82A	RBCEcho	RBC\TLK\Echo	Info	DESC=Train Position Report,NID_MESSAGE=136,L_MESSAGE=25,T_TRAIN=95743026,NID_ENGINE=20794,PACKETS=1,PACKET=1,DESC=Position Report,NID_PACKET=0,L_PACKET=122,Q_SCALE=1,	
10	2021-06-28T03:25:08.000	015CE82B	RBCEcho	RBC\TLK\Echo	Info	unset CommonAttributeChangeEvent 1C1B	
11	2021-06-28T03:25:09.711	015CE82C	RBCServer	RBC heartbeat	OK	High MnemonicChangeEvent 6118	
12	2021-06-28T03:25:09.711	015CE82D	RBCServer	RBC\TLK Status=519280, RbcTime=528970461	CommonAttributeChangeEvent 156D		
13	2021-06-28T03:25:11.710	015CE82E	RBCEcho	RBC\TLK\Echo	Info	DESC=General Message,NID_MESSAGE=24,L_MESSAGE=10,T_TRAIN=95743306,M_ACK=0,NID_LRBG=32775,PACKETS=0 CommonAttributeChangeEvent AB72	
14	2021-06-28T03:25:11.710	015CE82F	RBCEcho	RBC\TLK\Echo	Info	unset CommonAttributeChangeEvent F058	
15	2021-06-28T03:25:11.710	015CE830	RBCEcho	RBC\TLK\Echo	Info	DESC=General Message,NID_MESSAGE=24,L_MESSAGE=10,T_TRAIN=95743306,M_ACK=0,NID_LRBG=32775,PACKETS=0 CommonAttributeChangeEvent AD1B	
16	2021-06-28T03:25:11.710	015CE831	RBCEcho	RBC\TLK\Echo	Info	unset CommonAttributeChangeEvent 8C19	
17	2021-06-28T03:25:11.911	015CE832	RBCServer	RBC heartbeat	OK	High MnemonicChangeEvent 98F9	
18	2021-06-28T03:25:11.911	015CE833	RBCServer	RBC\TLK Status=519280, RbcTime=528970701	CommonAttributeChangeEvent 9999		
19	2021-06-28T03:25:13.176	015CE834	RBCEcho	RBC\TLK\Echo	Info	DESC=Train Position Report,NID_MESSAGE=136,L_MESSAGE=25,T_TRAIN=95743545,NID_ENGINE=20794,PACKETS=1,PACKET=1,DESC=Position Report,NID_PACKET=0,L_PACKET=122,Q_SCALE=1,	
20	2021-06-28T03:25:13.176	015CE835	RBCEcho	RBC\TLK\Echo	Info	unset CommonAttributeChangeEvent 6DE6	
21	2021-06-28T03:25:13.177	015CE836	RBCEcho	RBC\TLK\Echo	Info	DESC=Train Position Report,NID_MESSAGE=136,L_MESSAGE=25,T_TRAIN=95743545,NID_ENGINE=20794,PACKETS=1,PACKET=1,DESC=Position Report,NID_PACKET=0,L_PACKET=122,Q_SCALE=1,	
22	2021-06-28T03:25:13.177	015CE837	RBCEcho	RBC\TLK\Echo	Info	unset CommonAttributeChangeEvent C2EA	
23	2021-06-28T03:25:13.182	015CE838	RBCEcho	RBC\TLK\Echo	Info	DESC=Train Position Report,NID_MESSAGE=136,L_MESSAGE=25,T_TRAIN=95743545,NID_ENGINE=20794,PACKETS=1,PACKET=1,DESC=Position Report,NID_PACKET=0,L_PACKET=122,Q_SCALE=1,	
24	2021-06-28T03:25:13.183	015CE839	RBCEcho	RBC\TLK\Echo	Info	unset CommonAttributeChangeEvent 51E5	
25	2021-06-28T03:25:13.183	015CE83A	RBCEcho	RBC\TLK\Echo	Info	DESC=Train Position Report,NID_MESSAGE=136,L_MESSAGE=25,T_TRAIN=95743545,NID_ENGINE=20794,PACKETS=1,PACKET=1,DESC=Position Report,NID_PACKET=0,L_PACKET=122,Q_SCALE=1,	
26	2021-06-28T03:25:13.183	015CE83B	RBCEcho	RBC\TLK\Echo	Info	unset CommonAttributeChangeEvent B566	
27	2021-06-28T03:25:14.110	015CE83C	RBCServer	RBC heartbeat	OK	High MnemonicChangeEvent 33B9	
28	2021-06-28T03:25:14.110	015CE83D	RBCServer	RBC\TLK Status=519280, RbcTime=528970921	CommonAttributeChangeEvent 8635		
29	2021-06-28T03:25:16.294	015CE83E	RBCServer	RBC heartbeat	OK	High MnemonicChangeEvent D106	
30	2021-06-28T03:25:16.294	015CE83F	RBCServer	RBC\TLK Status=519280, RbcTime=528971141	CommonAttributeChangeEvent 9A20		
31	2021-06-28T03:25:18.193	015CE840	RBCEcho	RBC\TLK\Echo	Info	DESC=Train Position Report,NID_MESSAGE=136,L_MESSAGE=25,T_TRAIN=95744045,NID_ENGINE=20794,PACKETS=1,PACKET=1,DESC=Position Report,NID_PACKET=0,L_PACKET=122,Q_SCALE=1,	
32	2021-06-28T03:25:18.193	015CE841	RBCEcho	RBC\TLK\Echo	Info	unset CommonAttributeChangeEvent 8E27	
33	2021-06-28T03:25:18.193	015CE842	RBCEcho	RBC\TLK\Echo	Info	DESC=Train Position Report,NID_MESSAGE=136,L_MESSAGE=25,T_TRAIN=95744045,NID_ENGINE=20794,PACKETS=1,PACKET=1,DESC=Position Report,NID_PACKET=0,L_PACKET=122,Q_SCALE=1,	
34	2021-06-28T03:25:18.193	015CE843	RBCEcho	RBC\TLK\Echo	Info	unset CommonAttributeChangeEvent F625	
35	2021-06-28T03:25:18.199	015CE844	RBCEcho	RBC\TLK\Echo	Info	DESC=Train Position Report,NID_MESSAGE=136,L_MESSAGE=25,T_TRAIN=95744045,NID_ENGINE=20794,PACKETS=1,PACKET=1,DESC=Position Report,NID_PACKET=0,L_PACKET=122,Q_SCALE=1,	
36	2021-06-28T03:25:18.199	015CE845	RBCEcho	RBC\TLK\Echo	Info	unset CommonAttributeChangeEvent D445	
37	2021-06-28T03:25:18.199	015CE846	RBCEcho	RBC\TLK\Echo	Info	DESC=Train Position Report,NID_MESSAGE=136,L_MESSAGE=25,T_TRAIN=95744045,NID_ENGINE=20794,PACKETS=1,PACKET=1,DESC=Position Report,NID_PACKET=0,L_PACKET=122,Q_SCALE=1,	
38	2021-06-28T03:25:18.199	015CE847	RBCEcho	RBC\TLK\Echo	Info	unset CommonAttributeChangeEvent AC47	
39	2021-06-28T03:25:18.494	015CE848	RBCServer	RBC heartbeat	OK	High MnemonicChangeEvent AE67	
40	2021-06-28T03:25:18.494	015CE849	RBCServer	RBC\TLK Status=519280, RbcTime=528971961	CommonAttributeChangeEvent 466C		
41	2021-06-28T03:25:20.493	015CE84A	RBCServer	RBC heartbeat	OK	High MnemonicChangeEvent 2A47	
42	2021-06-28T03:25:20.493	015CE84B	RBCServer	RBC\TLK Status=519280, RbcTime=528971562	CommonAttributeChangeEvent D95B		

GENERAL MESSAGE

```

{
  NID_MESSAGE = 1
  L_MESSAGE = 39
  DATE
  {
    YEAR = 21
    MONTH = 6
    DAY = 26
  }
  TIME
  {
    HOU = 15
    MIN = 40
    SEC = 25
    TTS = 0 ms
  }
  TRAIN_POSITION
  {
    Q_SCALE = 1 '1 m scale'
    NID_LRBG
    {
      NID_C = 2
      NID_BG = 5355 |
    }
    D_LRBG = 30061
    Q_DIRLRBG = 0 'Reverse'
    Q_DLRBG = 1 'Nominal'
    L_DOUBTOVER = 1928
    L_DOUBTUNDER = 1523
  }
  V_TRAIN = 77 km/h
  DRIVER_ID = '2r47'
  NID_ENGINE = 20910
  SYSTEM_VERSION = 33 0x21 'Reserved for future use (this is a valid value)'
  M_LEVEL = 1 'Level STM specified by NID_STM'
  M_MODE = 13 'STM National'
}

```

Background

- European Train Control System (ETCS) upgrades are happening globally.
- The data recorded in the system is vital to understanding incidents



Objective

- I chose Python as the tool to use due to a lack of experience with programming and its integrated modules.



Methodology

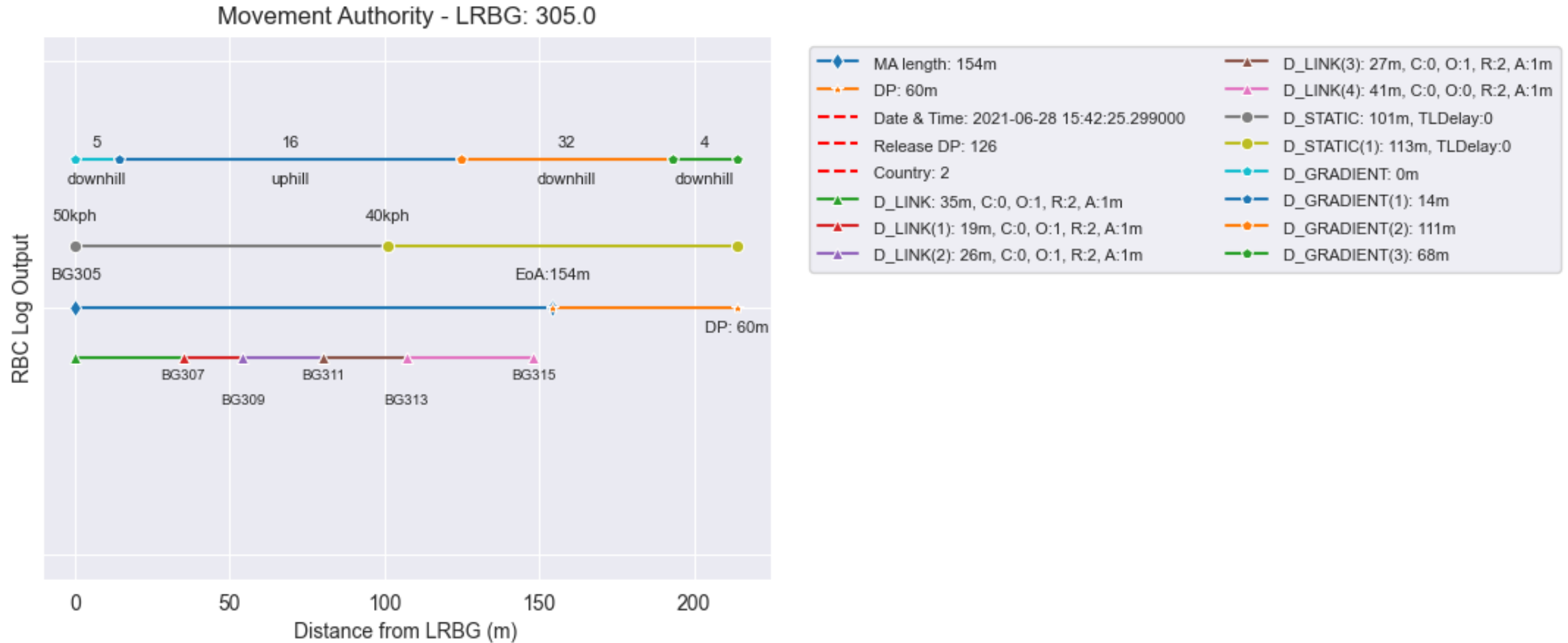
```

GENERAL MESSAGE
{
  NID_MESSAGE = 1
  L_MESSAGE = 39
  DATE
  {
    YEAR = 21
    MONTH = 6
    DAY = 26
  }
  TIME
  {
    HOU = 15
    MIN = 40
    SEC = 25
    TTS = 0 ms
  }
  TRAIN_POSITION
  {
    Q_SCALE = 1 '1 m scale'
    NID_LRBG
    {
      NID_C = 2
      NID_BG = 5355 |
    }
    D_LRBG = 30061
    Q_DIRLRBG = 0 'Reverse'
    Q_DLRBG = 1 'Nominal'
    L_DOUBTOVER = 1928
    L_DOUBTUNDER = 1523
  }
  V_TRAIN = 77 km/h
  DRIVER_ID = '2r47'
  NID_ENGINE = 20910
  SYSTEM_VERSION = 33 0x21 'Reserved for future use (this is a valid value)'
  M_LEVEL = 1 'Level STM specified by NID_STM'
  M_MODE = 13 'STM National'
}

```

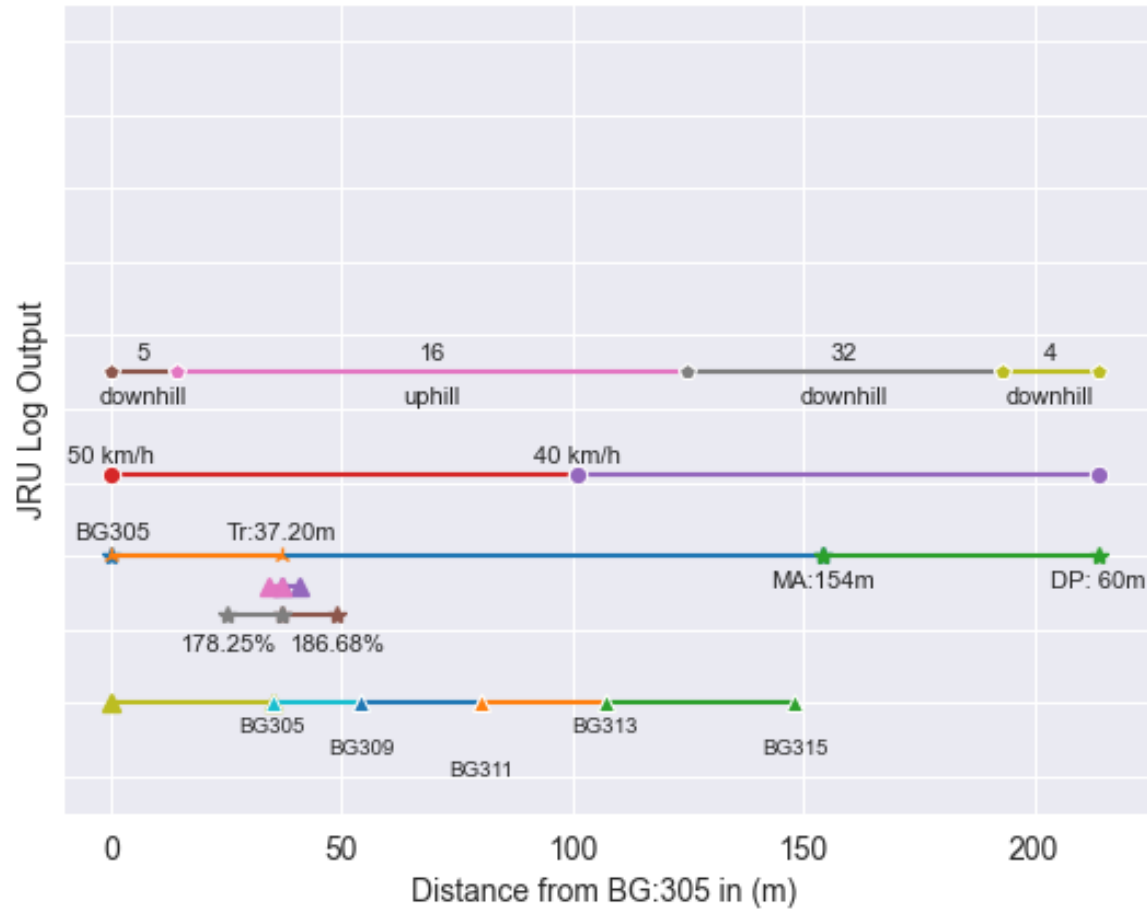
datetime	NID_MESSAGE	L_MESSAGE	Q_SCALE	NID_C	NID_BG	D_LRBG	Q_DIRLRBG	Q_DLRBG	L_DOUBTOVER	L_DOUBTUNDER	V_TRAIN	DRIVER_ID	NID_ENGINE
2021-06-28 14:42:14.250	136.0	24.0	1 '1 m scale'	2.0	301.0	7.0	1 'Nominal'	1 'Nominal'	3.0	3.0	50 km/h	'5z10'	20910.0
2021-06-28 14:42:14.200	20.0	46.0	0 '10 cm scale'	2.0	301.0	68.0	1 'Nominal'	1 'Nominal'	20.0	21.0	49 km/h	'5z10'	20910.0
2021-06-28 14:42:14.800	1.0	39.0	0 '10 cm scale'	2.0	301.0	149.0	1 'Nominal'	1 'Nominal'	24.0	25.0	49 km/h	'5z10'	20910.0
2021-06-28 14:42:15.400	20.0	46.0	0 '10 cm scale'	2.0	301.0	229.0	1 'Nominal'	1 'Nominal'	28.0	28.0	49 km/h	'5z10'	20910.0
2021-06-28 14:42:15.400	6.0	65.0	0 '10 cm scale'	2.0	303.0	227.0	1 'Nominal'	1 'Nominal'	28.0	28.0	49 km/h	'5z10'	20910.0
2021-06-28 14:42:15.650	136.0	24.0	1 '1 m scale'	2.0	303.0	8.0	1 'Nominal'	1 'Nominal'	3.0	3.0	50 km/h	'5z10'	20910.0
2021-06-28 14:42:15.800	3.0	171.0	1 '1 m scale'	2.0	303.0	103.0	1 'Nominal'	1 'Nominal'	23.0	23.0	48 km/h	'5z10'	20910.0
2021-06-28 14:42:15.850	132.0	24.0	1 '1 m scale'	2.0	303.0	11.0	1 'Nominal'	1 'Nominal'	3.0	3.0	50 km/h	'5z10'	20910.0
2021-06-28 14:42:15.850	136.0	24.0	1 '1 m scale'	2.0	303.0	11.0	1 'Nominal'	1 'Nominal'	3.0	3.0	50 km/h	'5z10'	20910.0
2021-06-28 14:42:16.200	6.0	65.0	0 '10 cm scale'	2.0	299.0	160.0	1 'Nominal'	1 'Nominal'	25.0	25.0	48 km/h	'5z10'	20910.0
2021-06-28 14:42:16.600	20.0	46.0	0 '10 cm scale'	2.0	299.0	212.0	1 'Nominal'	1 'Nominal'	27.0	27.0	48 km/h	'5z10'	20910.0
2021-06-28 14:42:16.550	6.0	65.0	0 '10 cm scale'	2.0	299.0	201.0	1 'Nominal'	1 'Nominal'	27.0	26.0	48 km/h	'5z10'	20910.0
2021-06-28 14:42:16.850	136.0	24.0	1 '1 m scale'	2.0	299.0	8.0	0 'Reverse'	0 'Reverse'	3.0	3.0	50 km/h	'5z10'	20910.0
2021-06-28 14:42:17.800	20.0	46.0	0 '10 cm scale'	2.0	299.0	208.0	0 'Reverse'	0 'Reverse'	28.0	28.0	48 km/h	'5z10'	20910.0
2021-06-28 14:42:19.000	20.0	46.0	0 '10 cm scale'	2.0	299.0	360.0	0 'Reverse'	0 'Reverse'	33.0	32.0	46 km/h	'5z10'	20910.0
2021-06-28 14:42:19.800	1.0	39.0	0 '10 cm scale'	2.0	299.0	460.0	0 'Reverse'	0 'Reverse'	36.0	35.0	44 km/h	'5z10'	20910.0
2021-06-28 14:42:20.200	20.0	46.0	0 '10 cm scale'	2.0	299.0	508.0	0 'Reverse'	0 'Reverse'	38.0	37.0	44 km/h	'5z10'	20910.0
2021-06-28 14:42:21.200	6.0	65.0	0 '10 cm scale'	2.0	305.0	625.0	0 'Reverse'	0 'Reverse'	41.0	40.0	42 km/h	'5z10'	20910.0
2021-06-28 14:42:21.450	136.0	24.0	1 '1 m scale'	2.0	305.0	8.0	1 'Nominal'	1 'Nominal'	3.0	3.0	45 km/h	'5z10'	20910.0
2021-06-28 14:42:21.400	20.0	46.0	0 '10 cm scale'	2.0	305.0	76.0	1 'Nominal'	1 'Nominal'	23.0	23.0	42 km/h	'5z10'	20910.0
2021-06-28 14:42:22.600	20.0	46.0	0 '10 cm scale'	2.0	305.0	211.0	1 'Nominal'	1 'Nominal'	28.0	28.0	40 km/h	'5z10'	20910.0
2021-06-28 14:42:23.700	9.0	21.0	1 '1 m scale'	2.0	305.0	329.0	1 'Nominal'	1 'Nominal'	32.0	32.0	38 km/h	'5z10'	20910.0
2021-06-28 14:42:23.850	138.0	28.0	1 '1 m scale'	2.0	305.0	35.0	1 'Nominal'	1 'Nominal'	4.0	4.0	40 km/h	'5z10'	20910.0
2021-06-28 14:42:23.800	20.0	46.0	0 '10 cm scale'	2.0	305.0	342.0	1 'Nominal'	1 'Nominal'	33.0	32.0	38 km/h	'5z10'	20910.0
2021-06-28 14:42:24.100	3.0	122.0	0 '10 cm scale'	2.0	305.0	372.0	1 'Nominal'	1 'Nominal'	34.0	33.0	38 km/h	'5z10'	20910.0
2021-06-28 14:42:24.250	132.0	24.0	1 '1 m scale'	2.0	305.0	39.0	1 'Nominal'	1 'Nominal'	4.0	4.0	40 km/h	'5z10'	20910.0

An output from an RBC Log



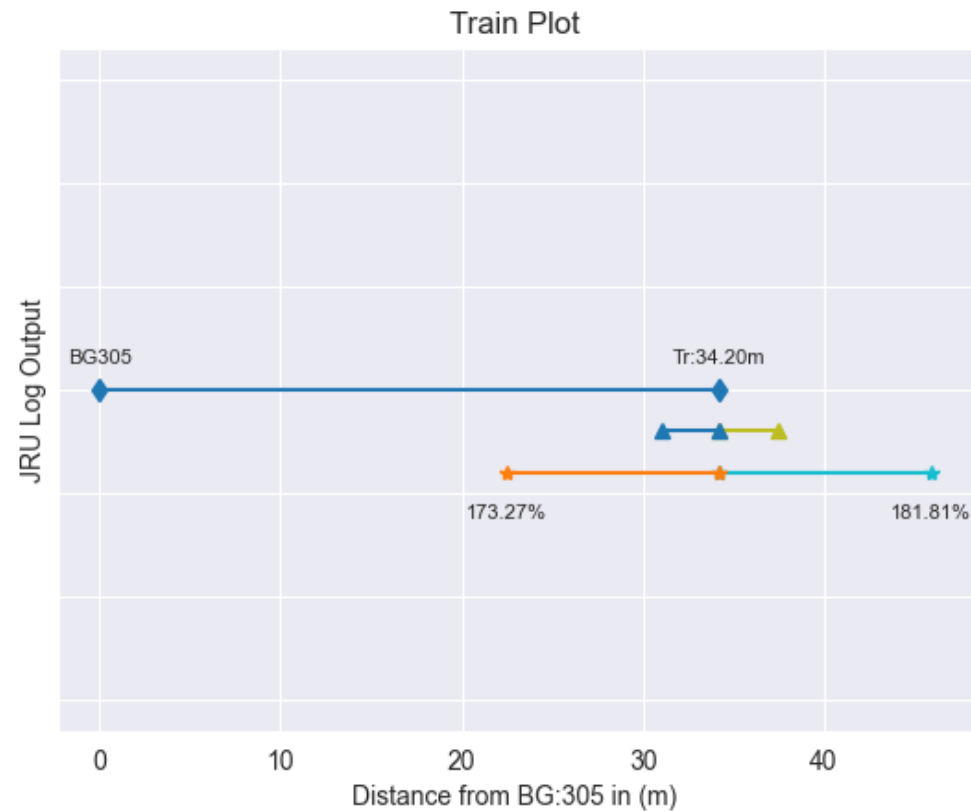
An output from a JRU Log

Train Plot



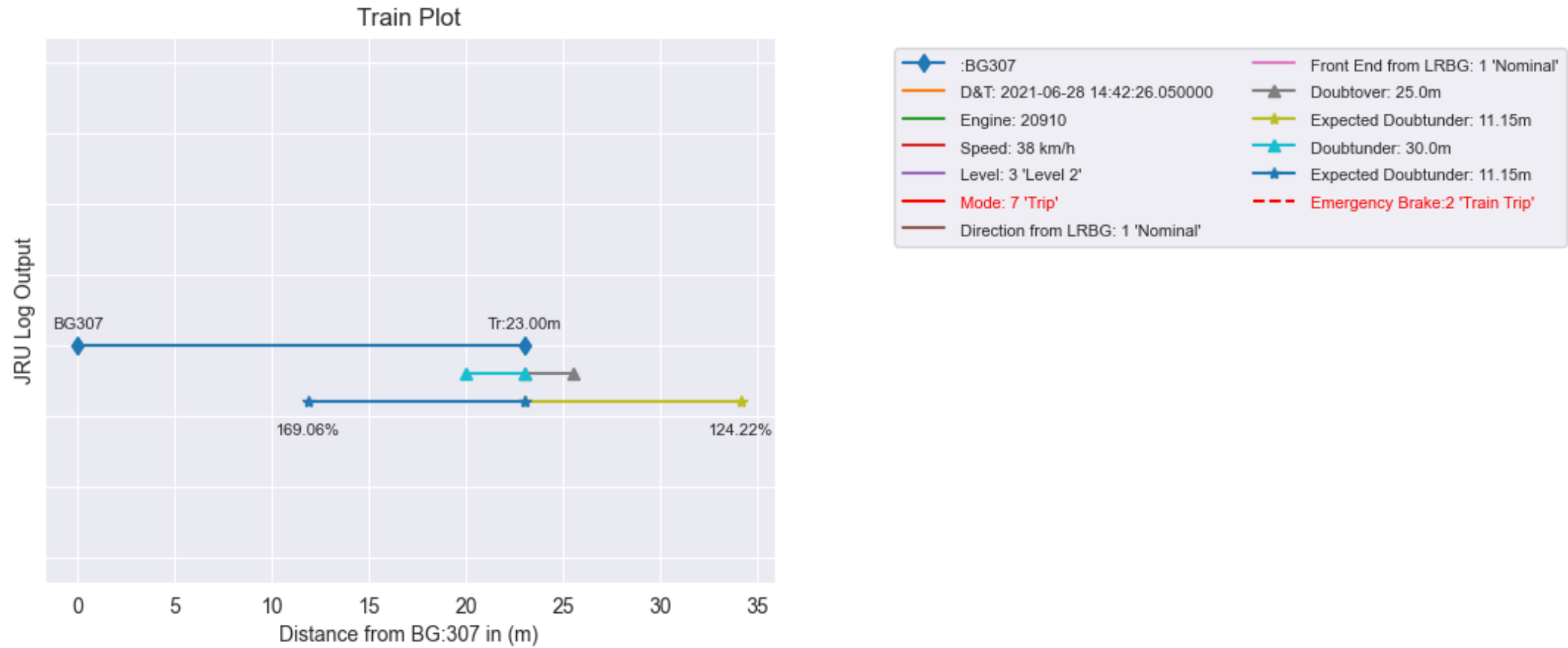
- MA:154m
- :BG305
- DP: 60m
- D&T: 2021-06-28 14:42:24.100000
- Loc ACC: 1m
- Engine: 20910
- Speed: 38 km/h
- Level: 3 'Level 2'
- Mode: 0 'Full Supervision'
- Direction from LRBG: 1 'Nominal'
- Front End from LRBG: 1 'Nominal'
- 1 'Nominal'
- RBC Identity:14937
- Ack: 0 'No acknowledgement required'
- Doubtover: 34.0m
- Expected Doubtunder: 11.86m
- Doubtunder: 33.0m
- Expected Doubtunder: 11.86m
- D_LINK:35m, C:0, O:1, R:2, A:1m
- D_LINK[0]: 19m, C:0, O:1, R:2, A:1m
- D_LINK[1]: 26m, C:0, O:1, R:2, A:1m
- D_LINK[2]: 27m, C:0, O:1, R:2, A:1m
- D_LINK[3]: 41m, C:0, O:0, R:2, A:1m
- D_STATIC: 101m, TLDelay:0
- D_STATIC:113m, TLDelay:0
- D_GRADIENT: 0m
- D_GRADIENT[0]: 14m
- D_GRADIENT[1]: 111m
- D_GRADIENT[2]: 68m

Position Report

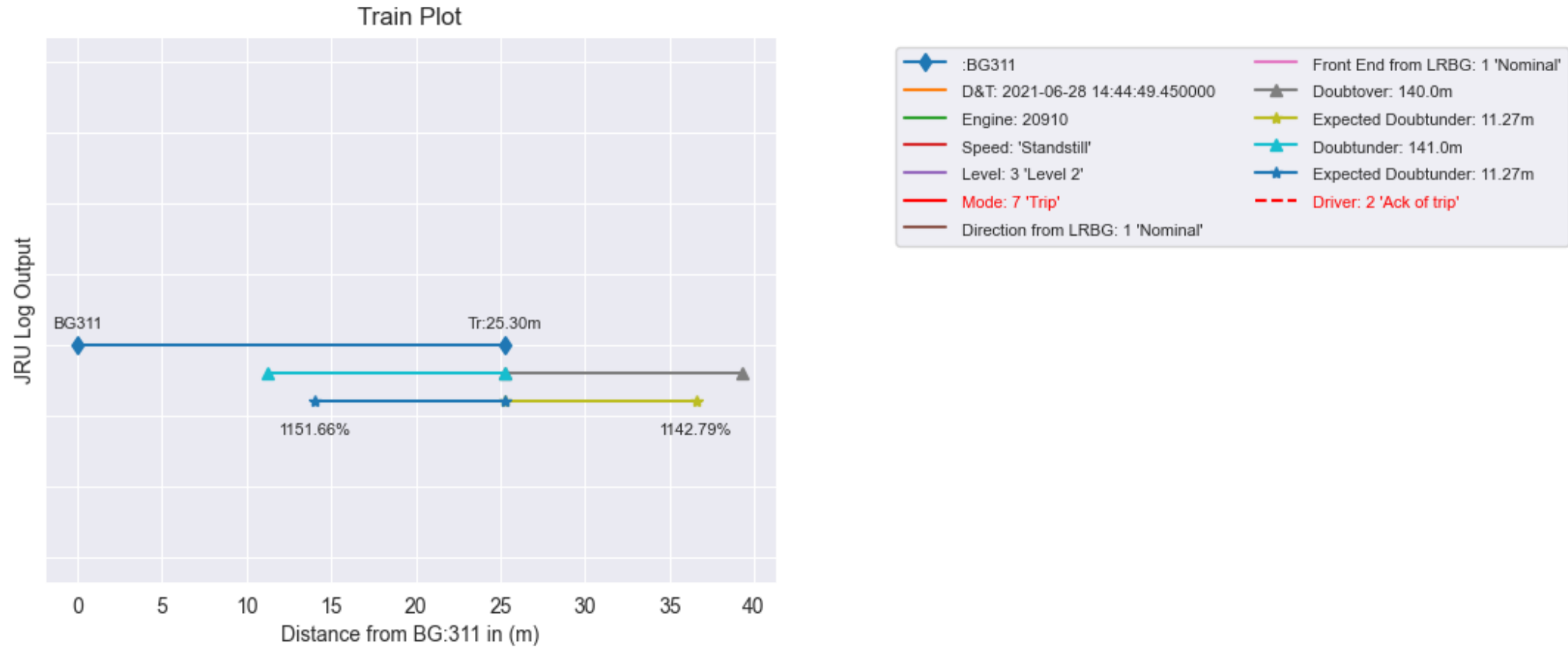


◆ :BG305	Direction from LRBG: 1 'Nominal'
— D&T: 2021-06-28 14:42:23.800000	Front End from LRBG: 1 'Nominal'
— Engine: 20910	▲ Doubtunder: 33.0m
— Speed: 38 km/h	▲ Expected Doubtunder: 11.71m
— Level: 3 'Level 2'	▲ Doubtunder: 32.0m
— Mode: 0 'Full Supervision'	★ Expected Doubtunder: 11.71m

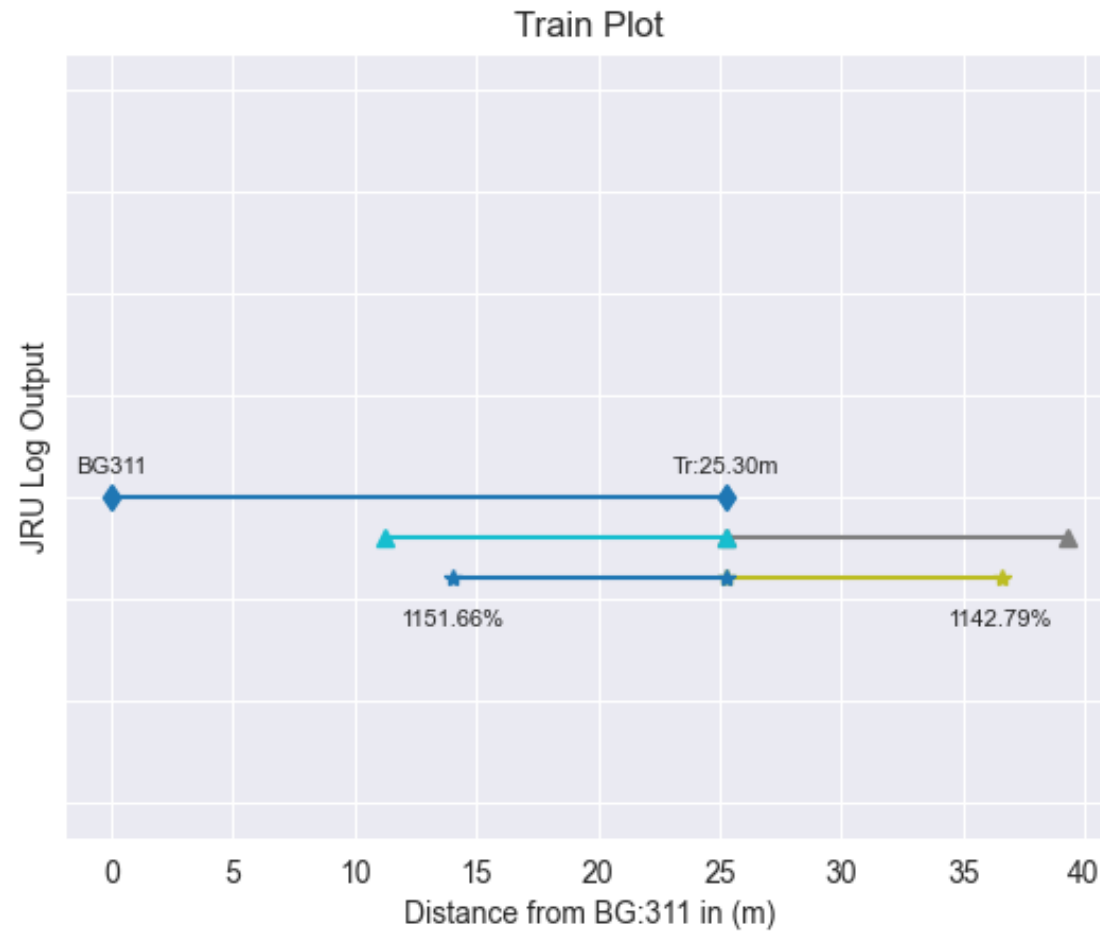
Unconditional Emergency Stop



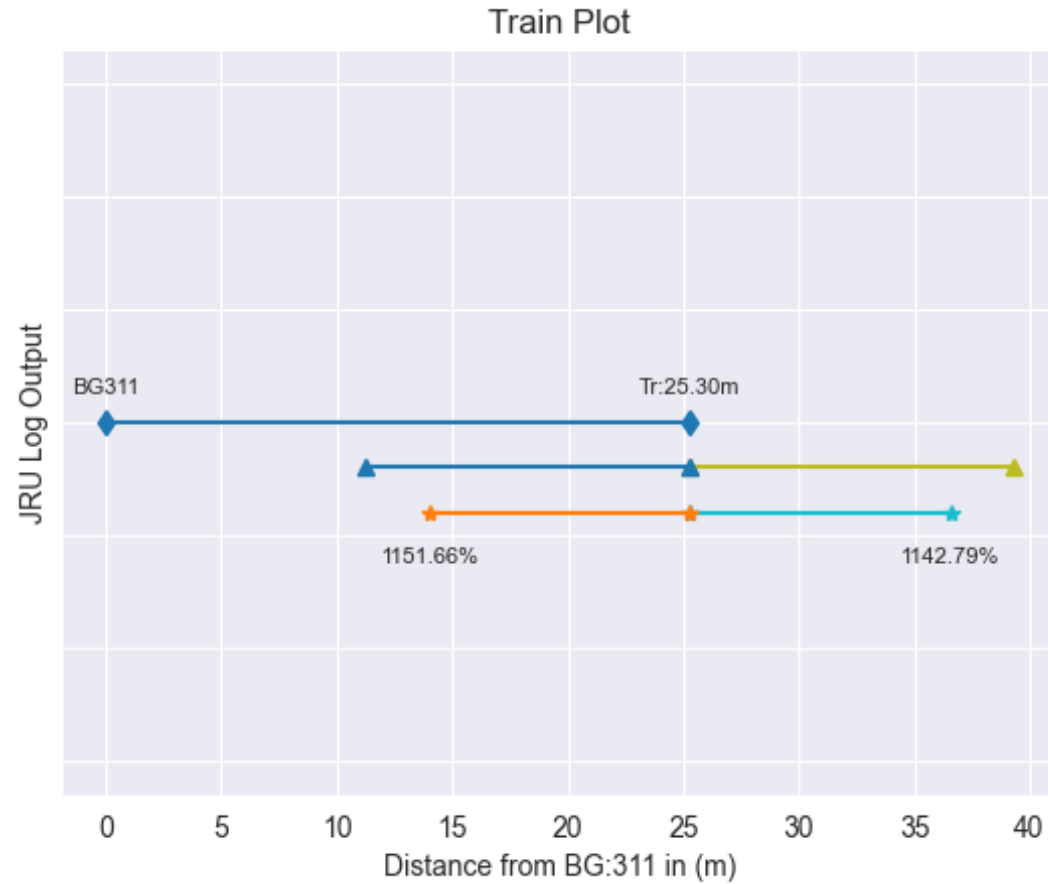
Acknowledgement of Train Trip



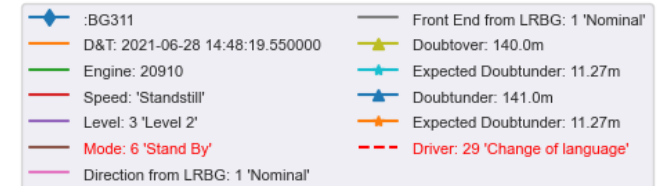
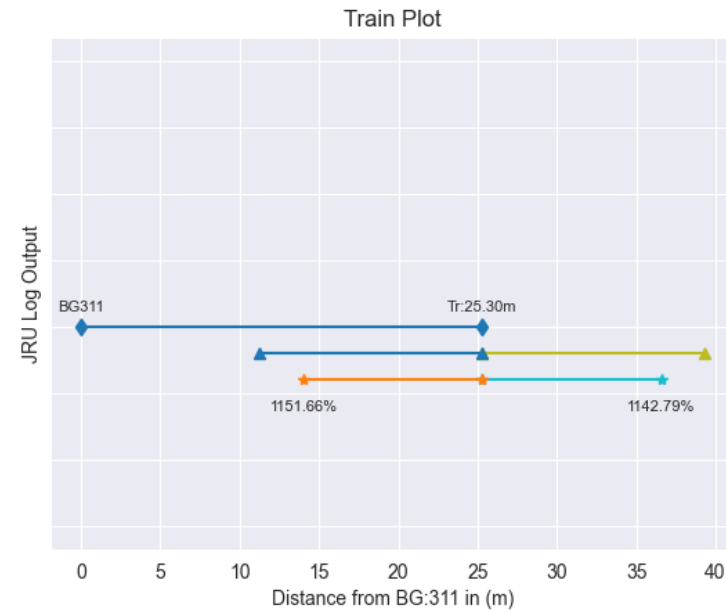
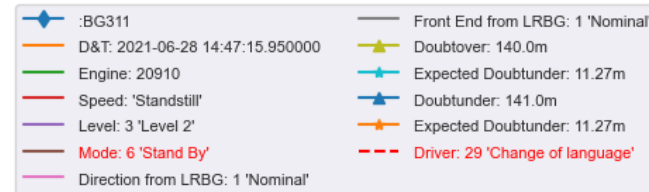
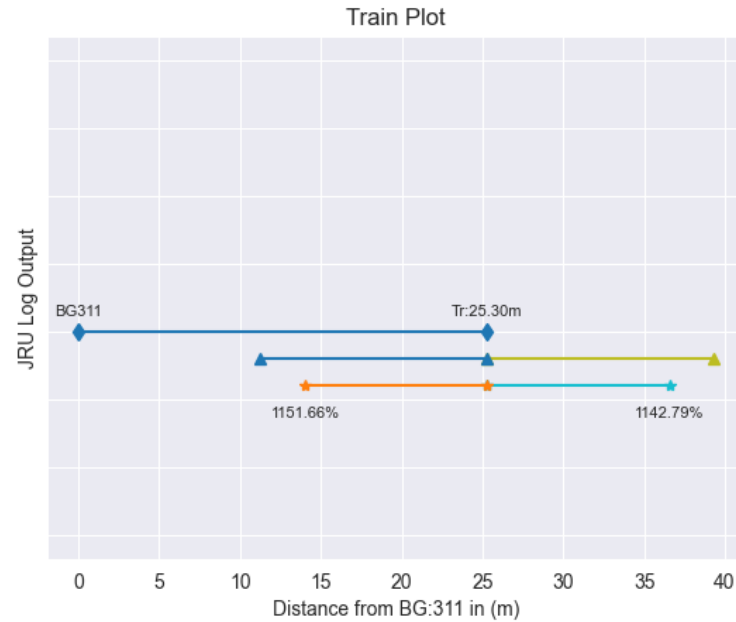
Post Trip



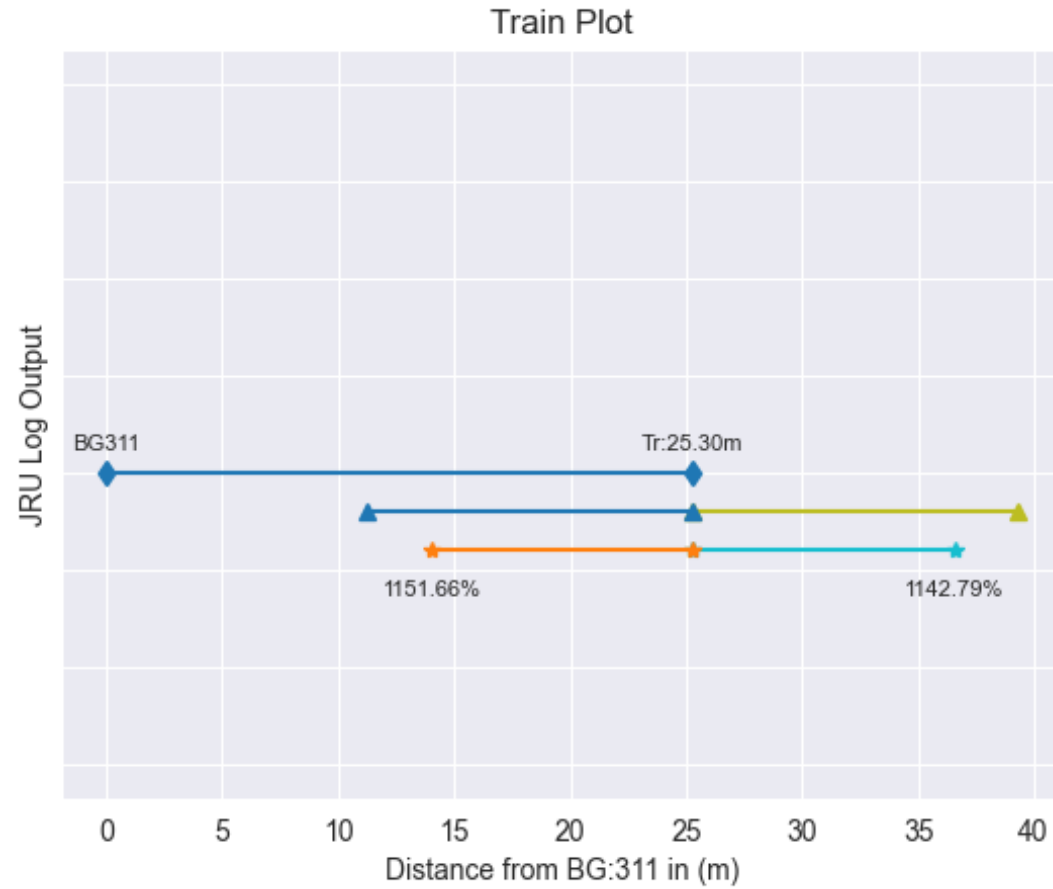
Standby Mode



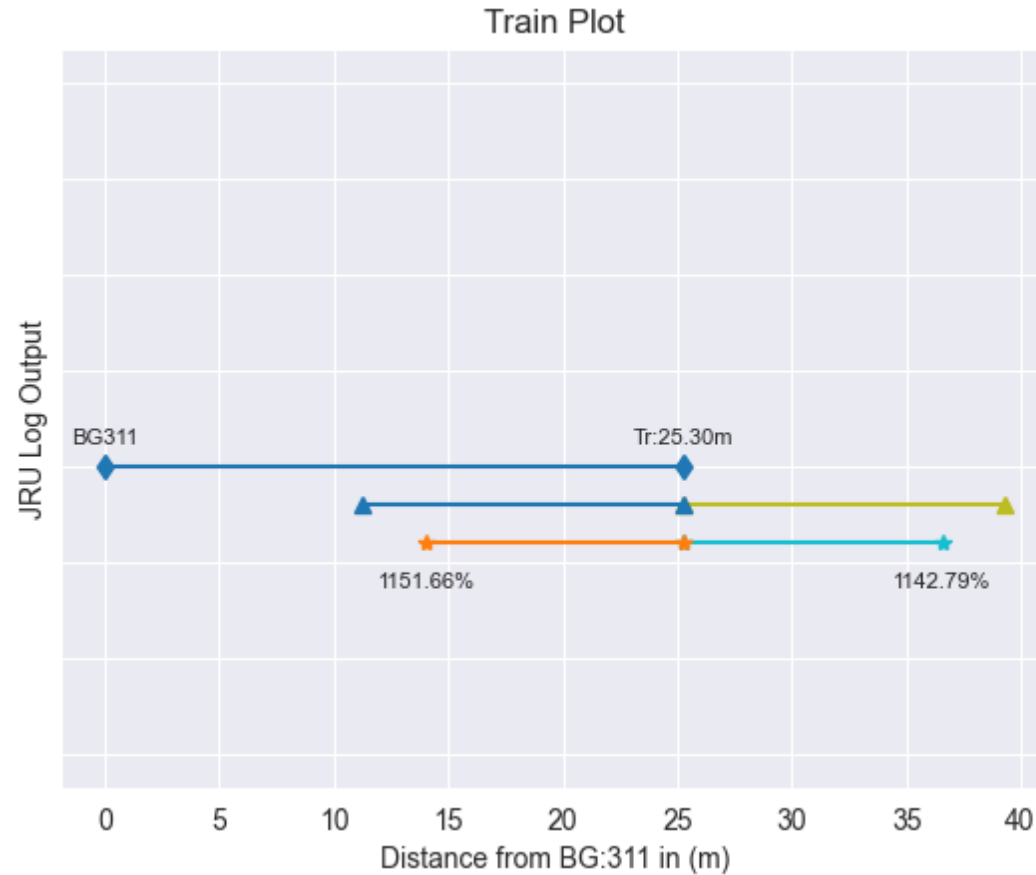
Language changes



Level Inhibition

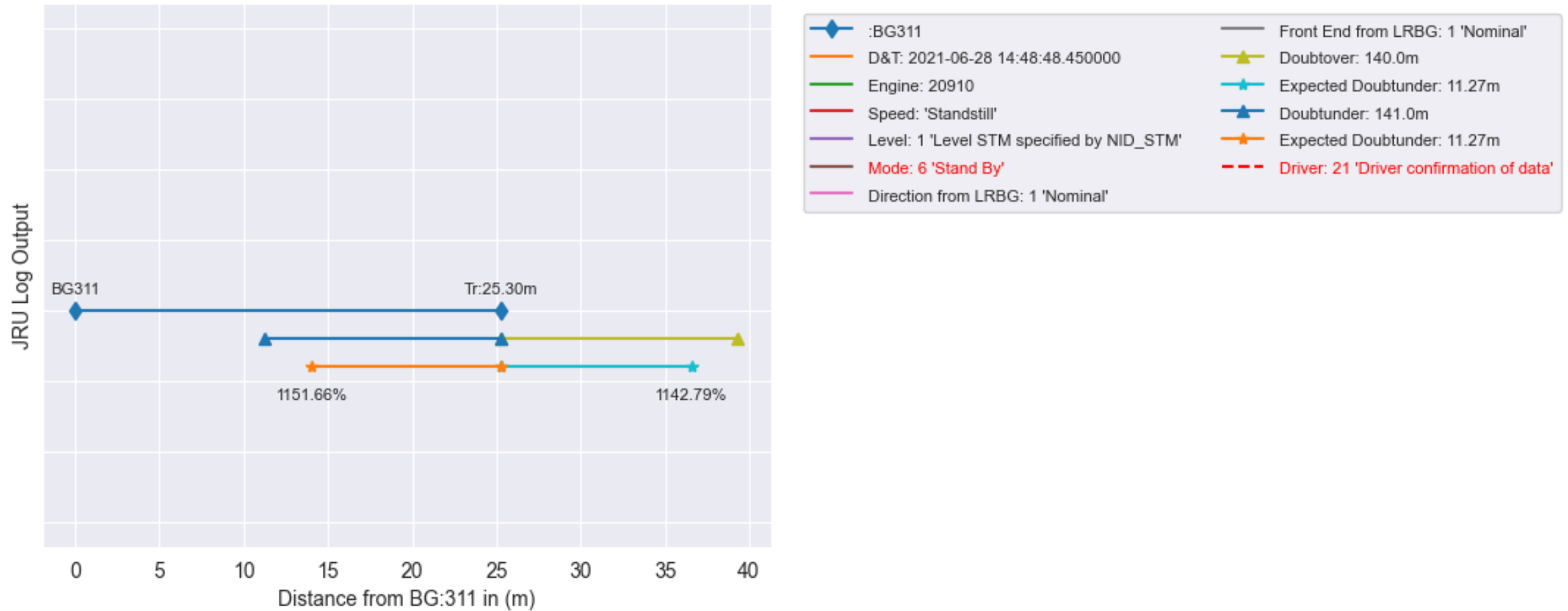


National Train Control

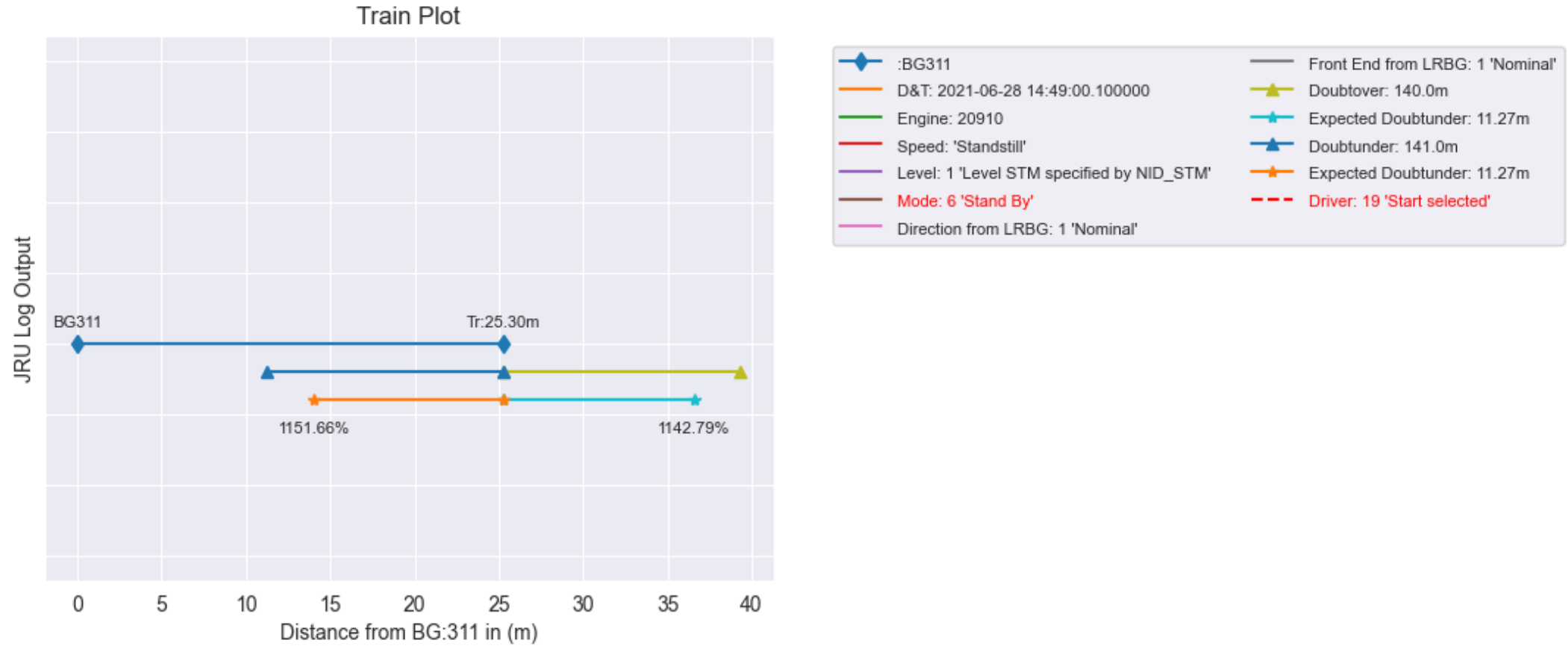


Data Entry

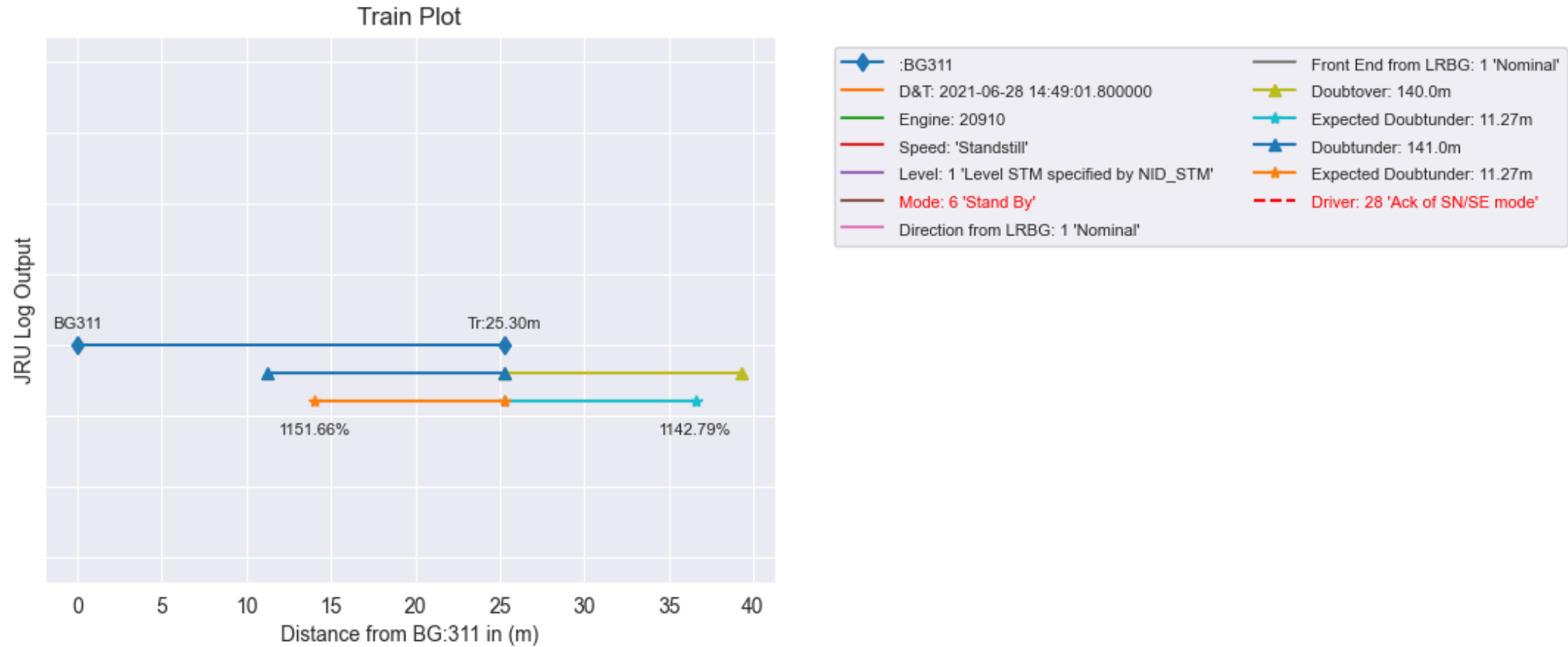
Train Plot



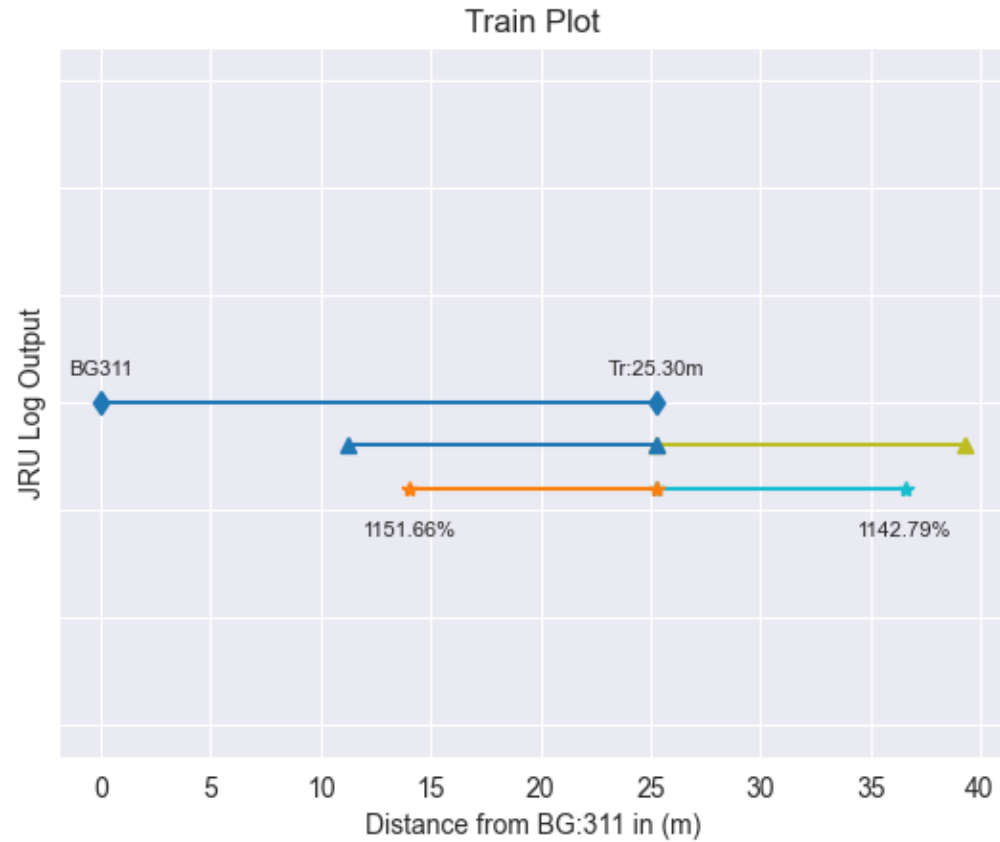
Start of Mission



Acknowledgement

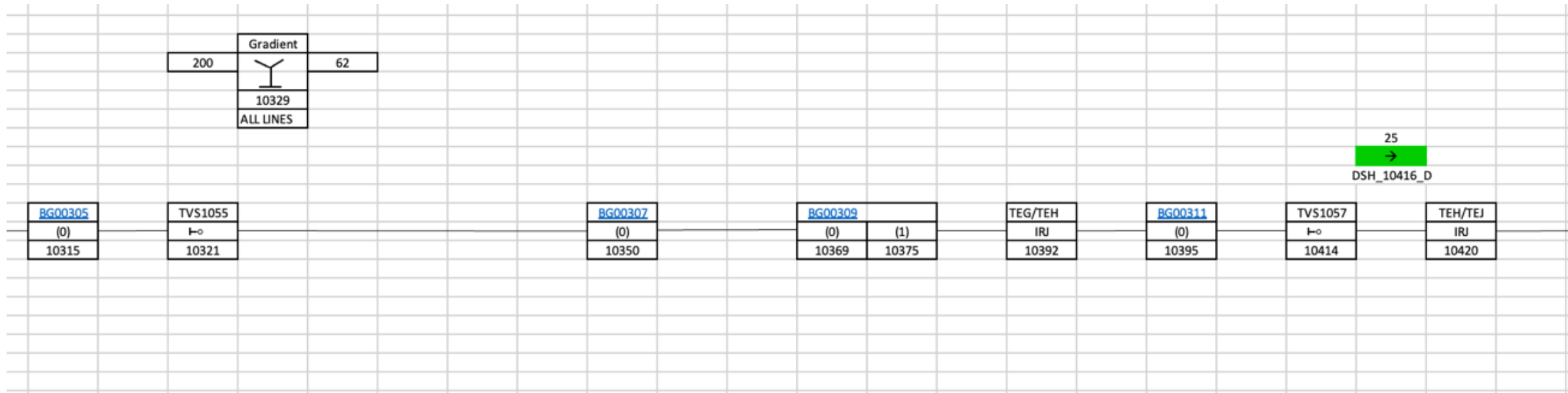


NTC Mode



- :BG311
- D&T: 2021-06-28 14:49:02
- Engine: 20910
- Speed: 'Standstill'
- Level: 1 'Level STM specified by NID_STM'
- Mode: 13 'STM National'
- Direction from LRBG: 1 'Nominal'
- Front End from LRBG: 1 'Nominal'
- Doubt over: 140.0m
- Expected Doubt under: 11.27m
- Doubt under: 141.0m
- Expected Doubt under: 11.27m

ETCS Data Generation Plan



Key Findings

Gained Efficiency in code interpretation

- From 1-3 days to minuets

Accuracy Improved

- Highlights mismatched time stamps on log files

Overlooked details

- The driver changed the language



Recommendations



Standardised File Format



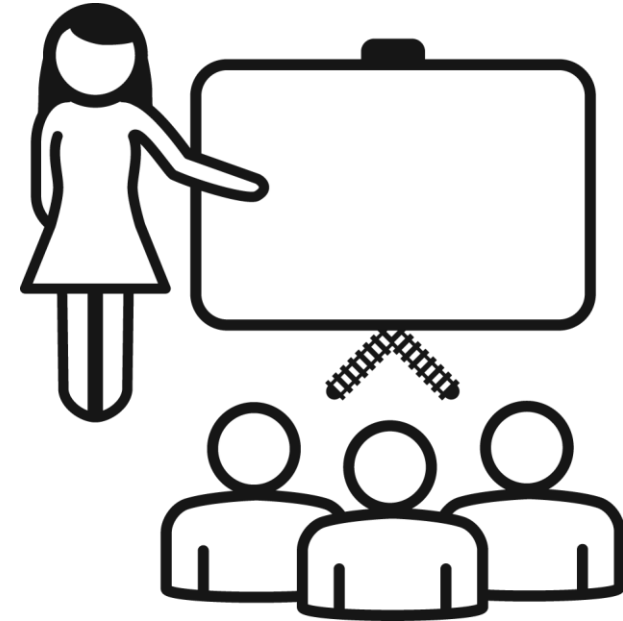
Integration with GSM-R



Enhancements



Digital Simulations



Conclusion



Feasible



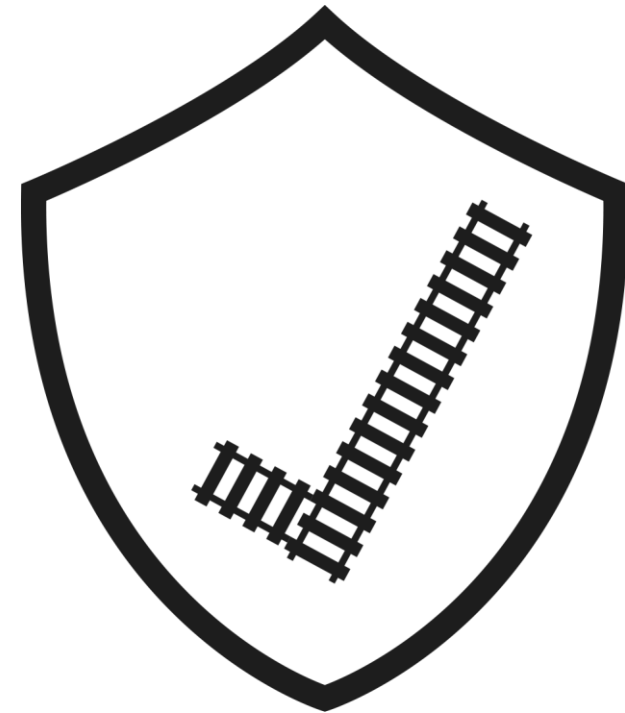
Impactful



Industry Wide



Revolutionise



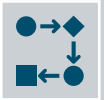
4 Challenges



Quality



Security



Compatibility



Scalability



5: Best Practice



Data Mapping

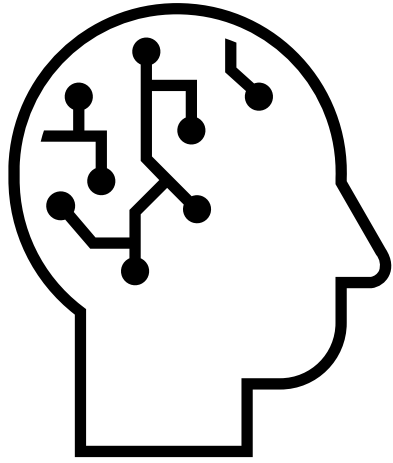


Data Governance



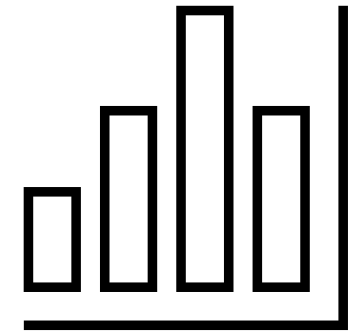
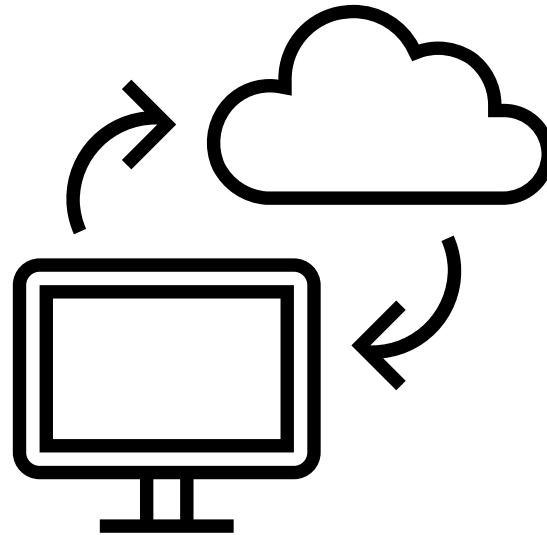
Monitoring & Maintenance

6: Future Trends



AI and Machine Learning

Cloud Integration



Self-service Data Integration

7: Conclusion

- Consistent data across all sources is vital to parsing the data.
- Using software can be more accurate the human interpretation

Do you have any questions?



Thank you