Creating an Intelligent Railway System: from big data collection to information distribution
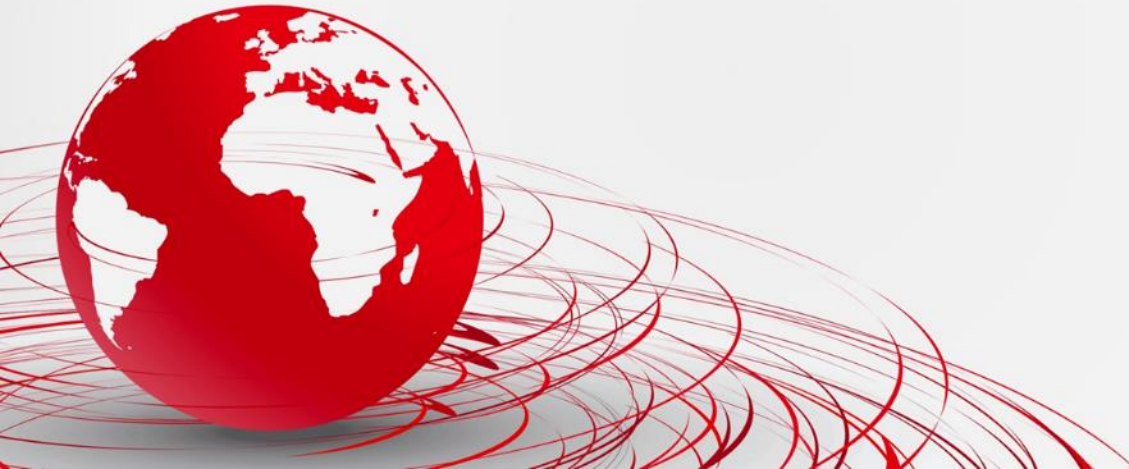
Pietro Pace
MER MEC SpA

About MERMEC

MERMEC main goal is to supply integrated solutions for railway infrastructure inspection, measuring, data analytics and decision support.
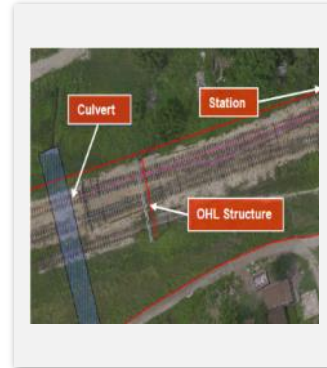

MEASURING VEHICLES AND TRAINS

MEASURING SYSTEMS

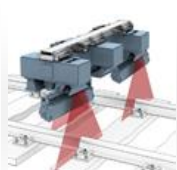**DATA ANALYTICS & DECISION SUPPORT SOFTWARE**

MEASURING SERVICE & CONSULTANCY

mermec

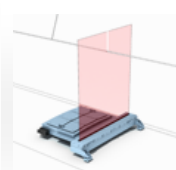# Measuring systems and vehicles product portfolio
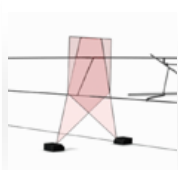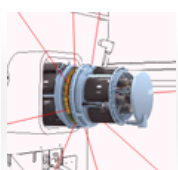
The wider and deeper portfolio on the market
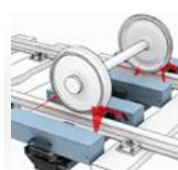
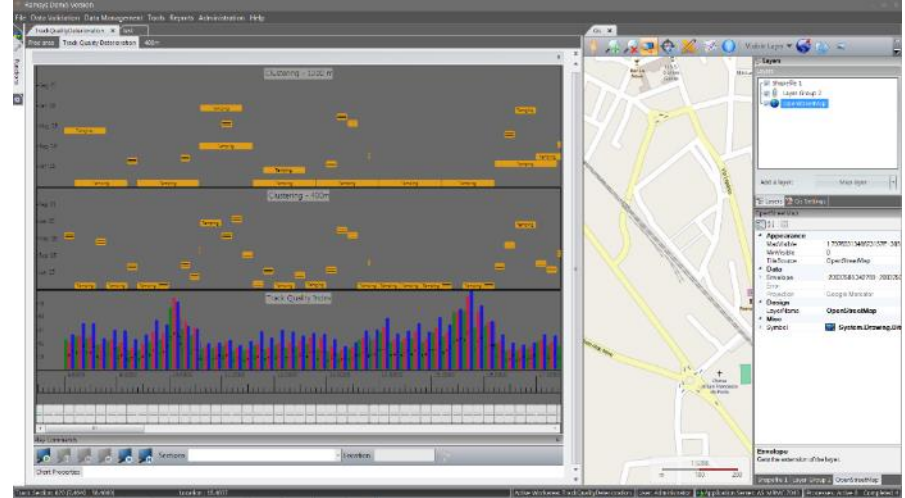| Track Measurement | Track Inspection | Catenary Measurement | Catenary Inspection | Tunnel Inspection | Train Monitoring | Measurement of Signalling & TLC |

60+ Types Measuring Systems

4 Models Recording Cars

1500+ Measuring and Inspection Systems Sold Internationally

New "big data" analytics and decision support

10+ TRACKWARE 20+ RAMSYS 10+ Countries

Agenda

1	Multiple conditioning monitoring devices
2	Big Data Characteristics (7 Vs)
3	Manage big data
4	Accurately monitor and predict asset health
5	Extraction of new information (3 use cases)
6	Conclusions

mermec

Agenda

mermec

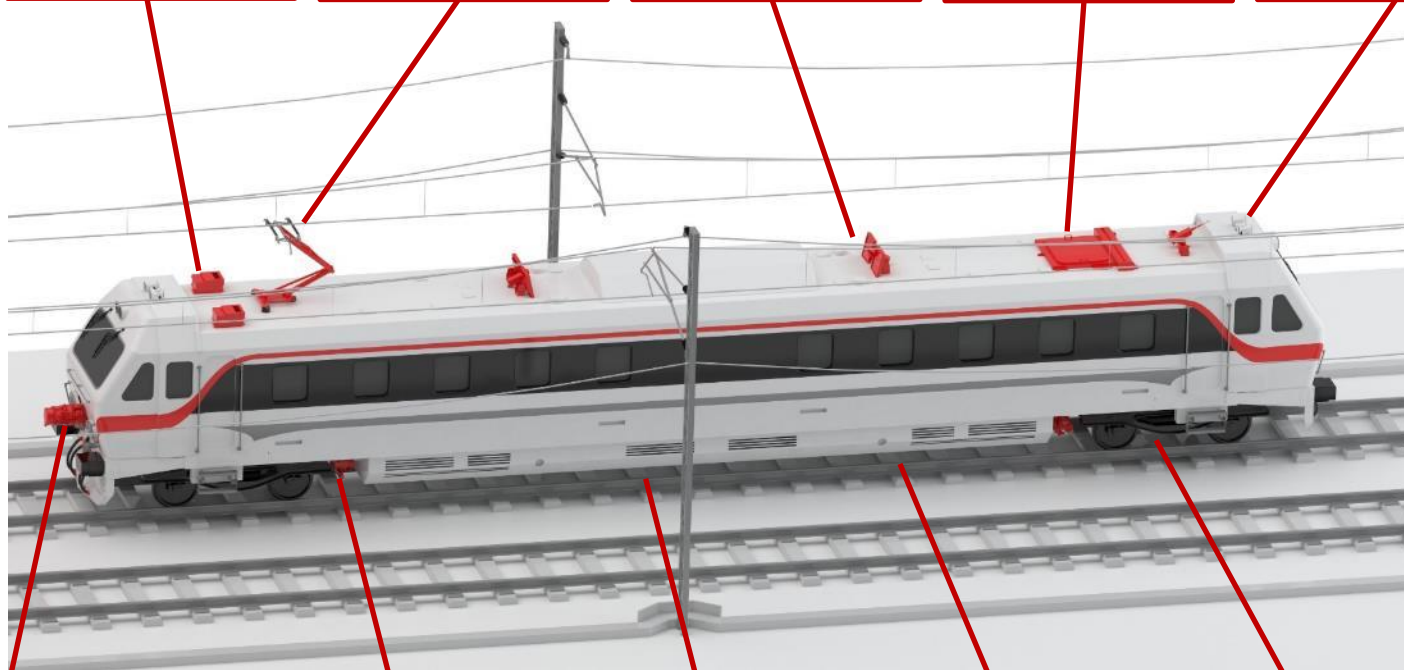# Infrastructure Condition Monitoring

POLE & TUNNELS DETECTOR

PANTOGRAPH MEASUREMENTS

OHL INSPECTION

OHL GEOMETRY & WIRE WEAR

VIDEO SURVEILLANCE

CLEARANCE GAUGE & TUNNEL INSPECTION
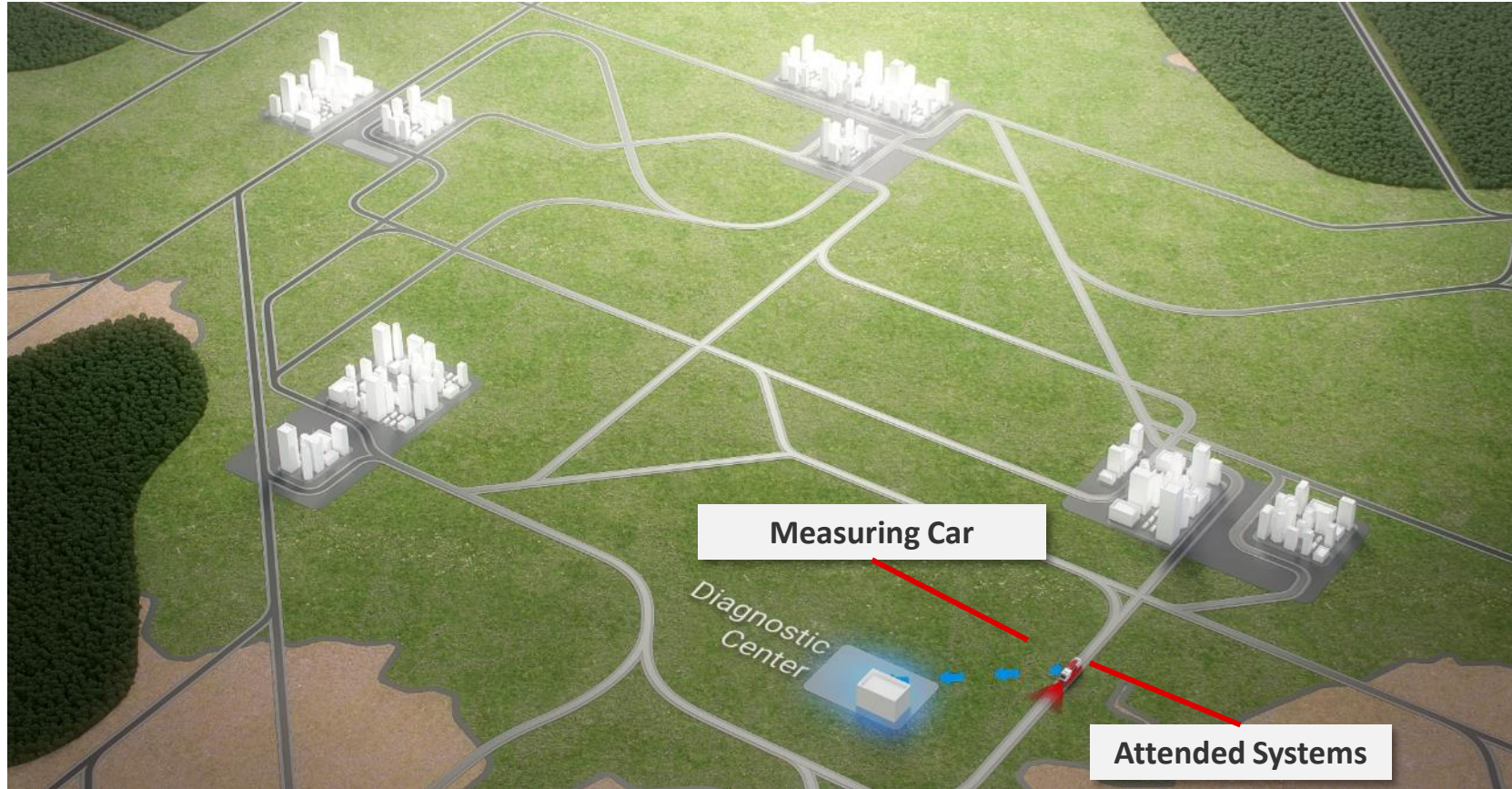
TRACK GEOMETRY & RAIL PROFILE

RAIL FLAW DETECTION
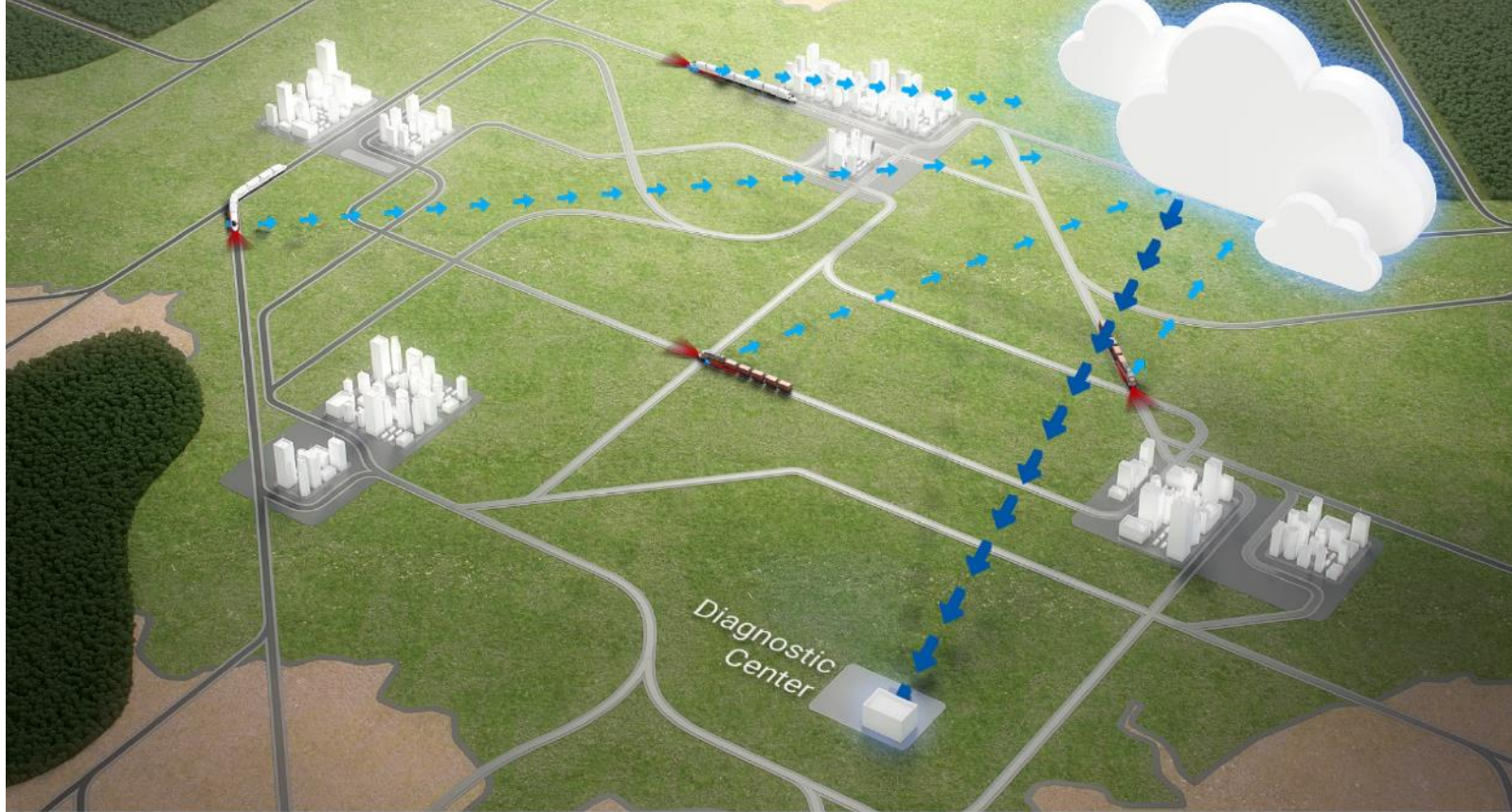
RAIL CORRUGATION

TRACK INSPECTION

mermec

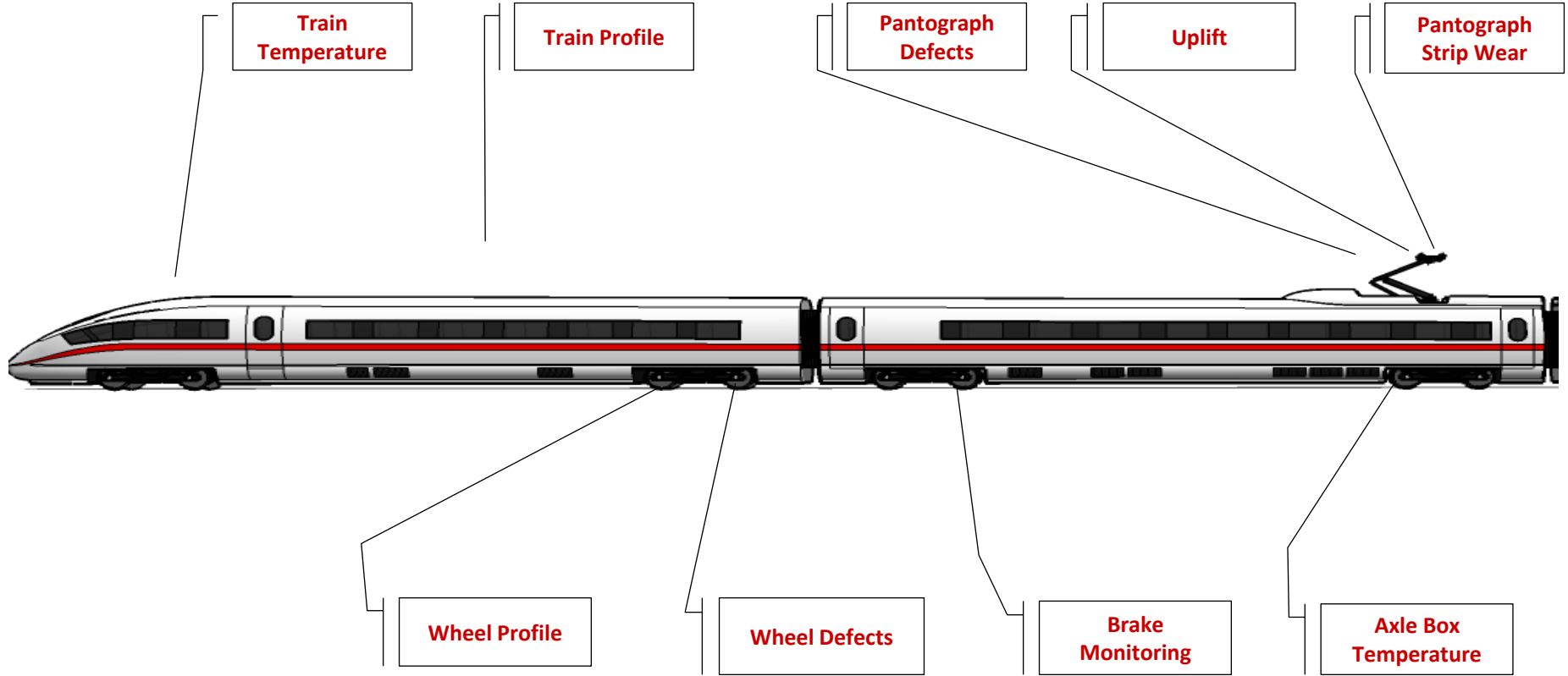# Centralised Infrastructure Condition Monitoring

Measuring Car

Diagnostic Center

Attended Systems

mermec

Distributed Infrastructure Condition Monitoring

Rolling Stock Condition Monitoring (Way Side)

Train Temperature

Train Profile

Pantograph Defects

Uplift

Pantograph Strip Wear

Wheel Profile

Wheel Defects

Brake Monitoring

Axle Box Temperature

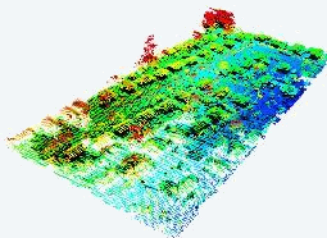mermec

Lidar-Scanning Airplanes Surveys

1.

LIDAR SURVEY

- LIDAR points
- Georeferenced Photos and Videos (aerial view)
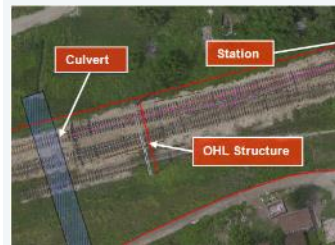

2.

DIAGNOSTIC VEHICLE SURVEY

- Georeferenced Localisation
- Condition Data
- Photos and Videos (infrastructure view)

3.

GEOGRAPHICAL DATA PROCESSING

- Corrected Geographical Data
- Digital Terrain Model
- Orthophotos

4.

OBJECTS EXTRACTION

- GPS and Track Referenced Objects including Track centerline

5.

ASSET CHARACTERISTICS POPULATION AND DATA TESTING

- Tested and validated Asset Data

ᗠ mermec

Agenda

mermec

Volume: How big are these data?

1 Bit = Binary Digit
8 Bits = 1 Byte
1024 Bytes = 1 Kilobyte
1024 Kilobytes = 1 Megabyte
1024 Megabytes = 1 Gigabyte
1024 Gigabytes = 1 Terabyte
1024 Terabytes = 1 Petabyte
1024 Petabytes = 1 Exabyte
1024 Exabytes = 1 Zettabyte
1024 Zettabytes = 1 Yottabyte
1024 Yottabytes = 1 Brontobyte
1024 Brontobytes = 1 Geopbyte
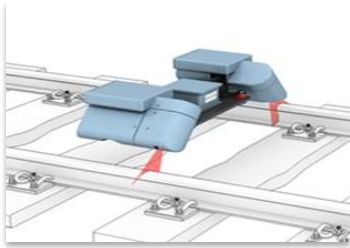1024 GeopBytes = 1 Saganbyte
1024 Saganbytes = 1 Jotabyte

Year	Vehicle/Train Systems Configuration	GB/Mile	* Required Storage (TB)
1994	Track Geometry and Rail Profile	0.003	~0.5
2010	Several systems (no automatic inspection)	0.079	~14
2014	Several systems and automatic inspection	15.716	~2900

* 5 years of data for 15000 km/~9320 miles (inspection frequency 4 times per year)

≋ mermec

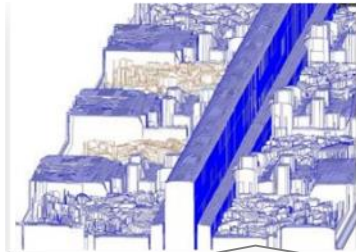# Velocity: Inspecting and Measuring ...

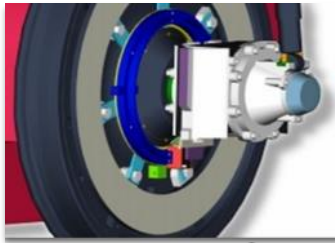**Track Geometry**
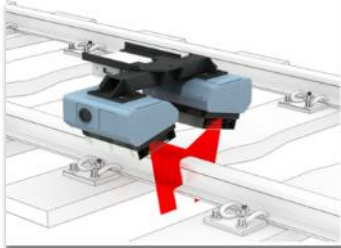
Rail Profile
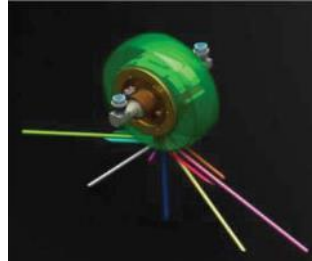
Ultrasonic Inspection
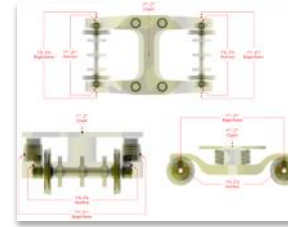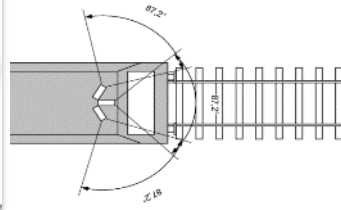

Running Dynamics
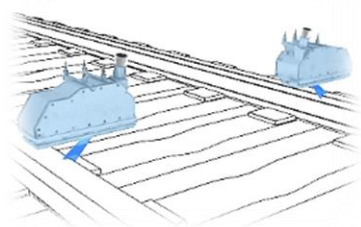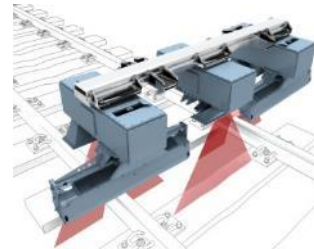

RCF Detection

Ballast Measurement

Wheel/Rail Interaction
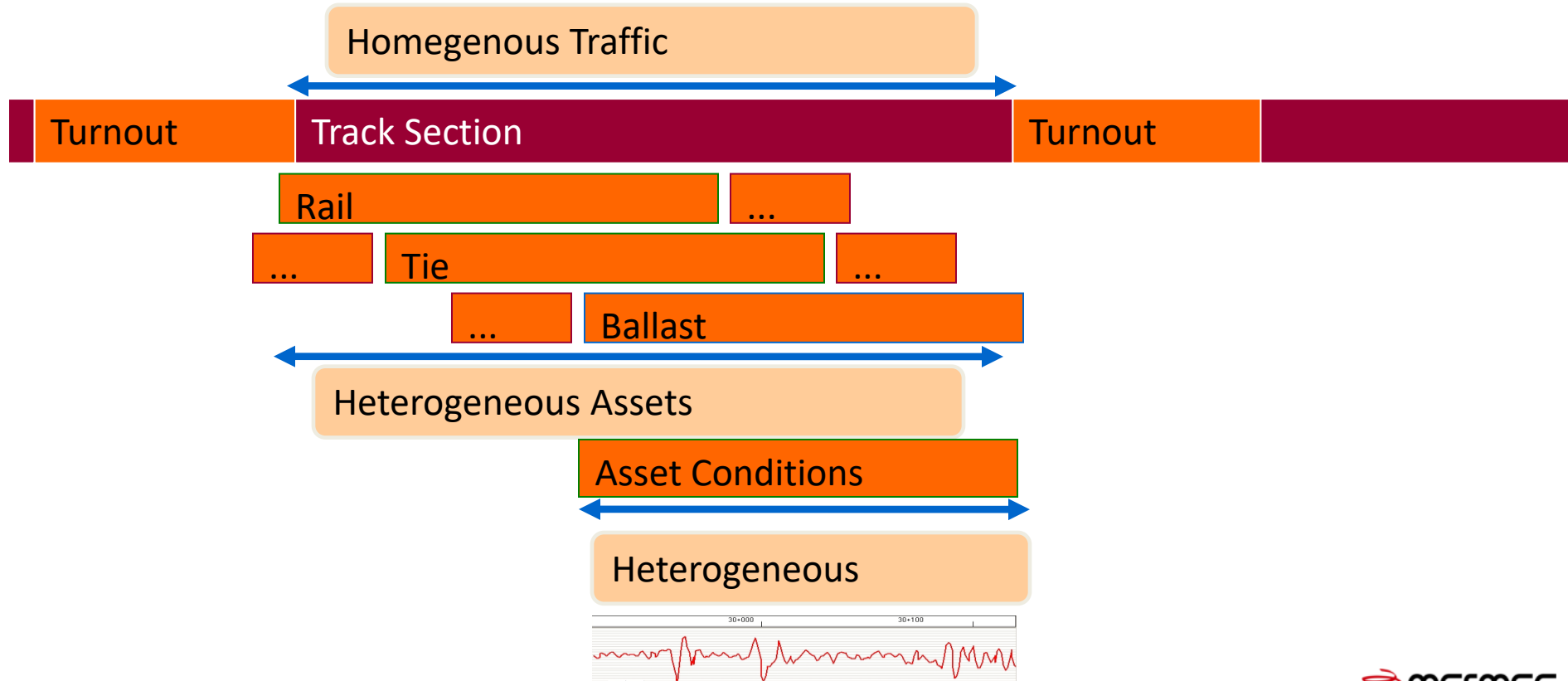

Track Video Surveillance
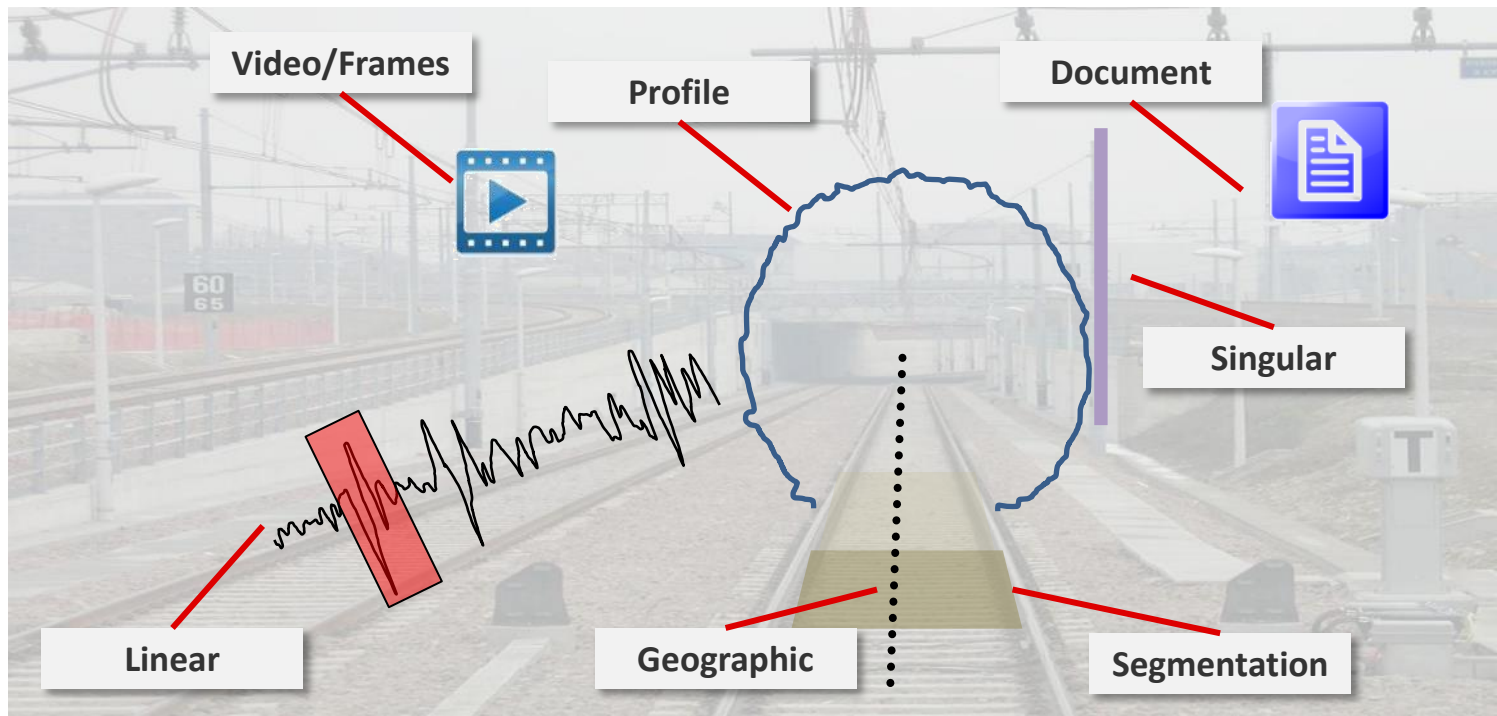

Vision Inspection

Tunnel

- **Measuring Speed**: Recording vehicle when hauled can reach 200 km/h
- **High Speed:** High Speed Recording Trains up to 400 km/h
- **High Frequency**: Autonomous/Unattended Systems (at each train run)

mermec

Variability: Data at Different Levels

Homegenous Traffic

| Turnout | Track Section | Turnout | |

Rail | ...

... | Tie | ...

... | Ballast

Heterogeneous Assets

Asset Conditions

Heterogeneous

MERMEC

Variety: Different Data Entity Types

Veracity: Data Quality

- Diagnostic data are affected by typical uncertainty and in some cases by errors (e.g. Peaks, Noise, Miss-Calibrations, Wrong Master Data, etc.)
- Invalid data leads to invalid information (GIGO effect), in particular invalid defects/alerts so called "**False Positives**"
- **False Negatives/Gaps** can be generated by missed inspections (e.g. unavailable track, RCF, system failure, etc.)
- **"Unknown Positive or Negative"** for data that have invalid localization
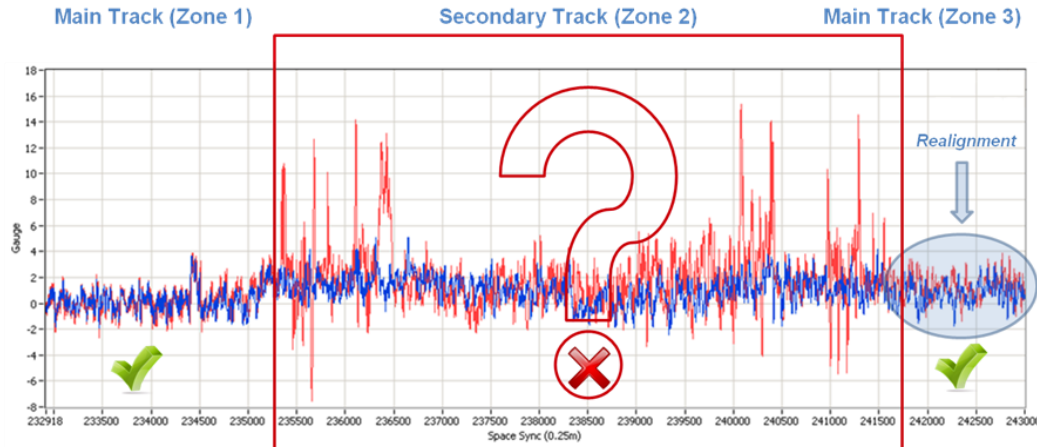
Visualisation: Common User Interface

False Positive

Valid Defect

Ballast

Track Geometry

MERMEC

Value: Why using big data?
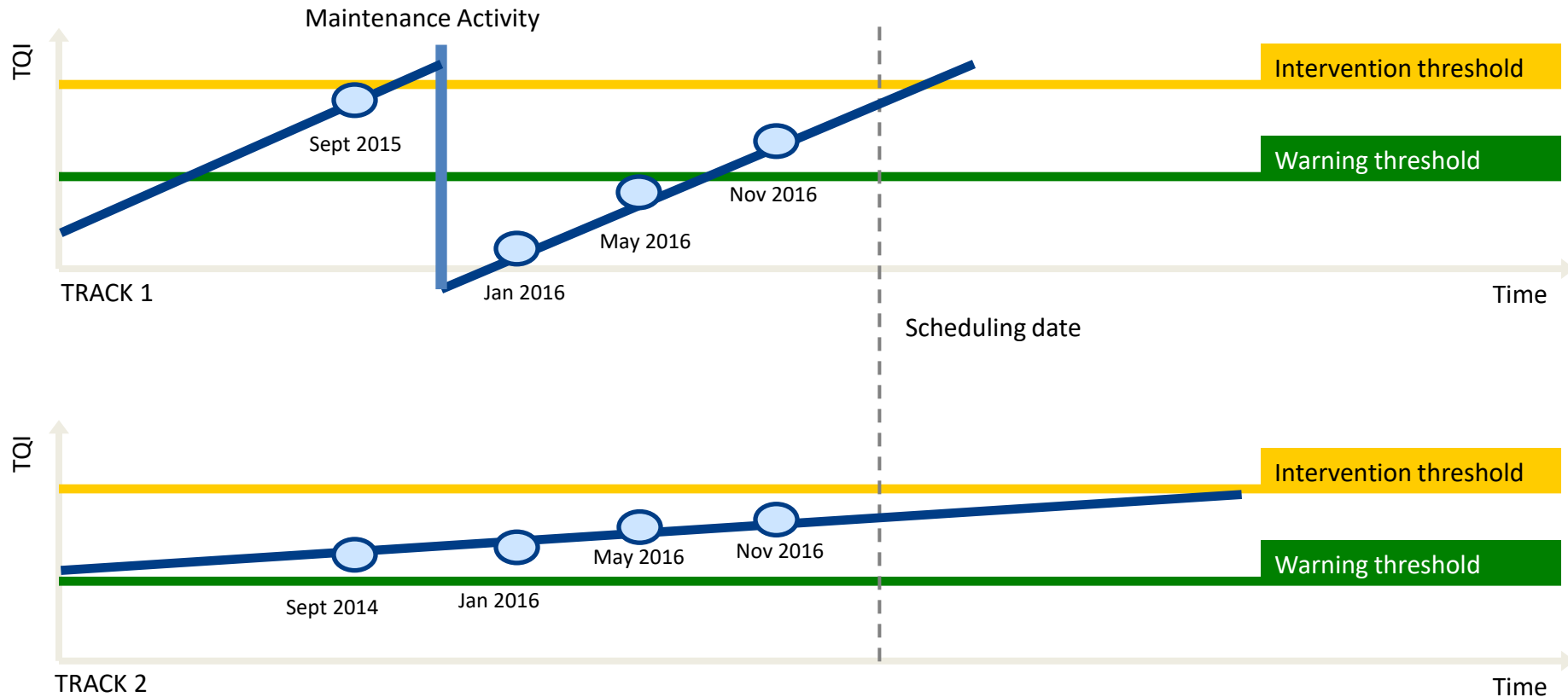
Agenda

mermec

Managing Big Data and Data Integration
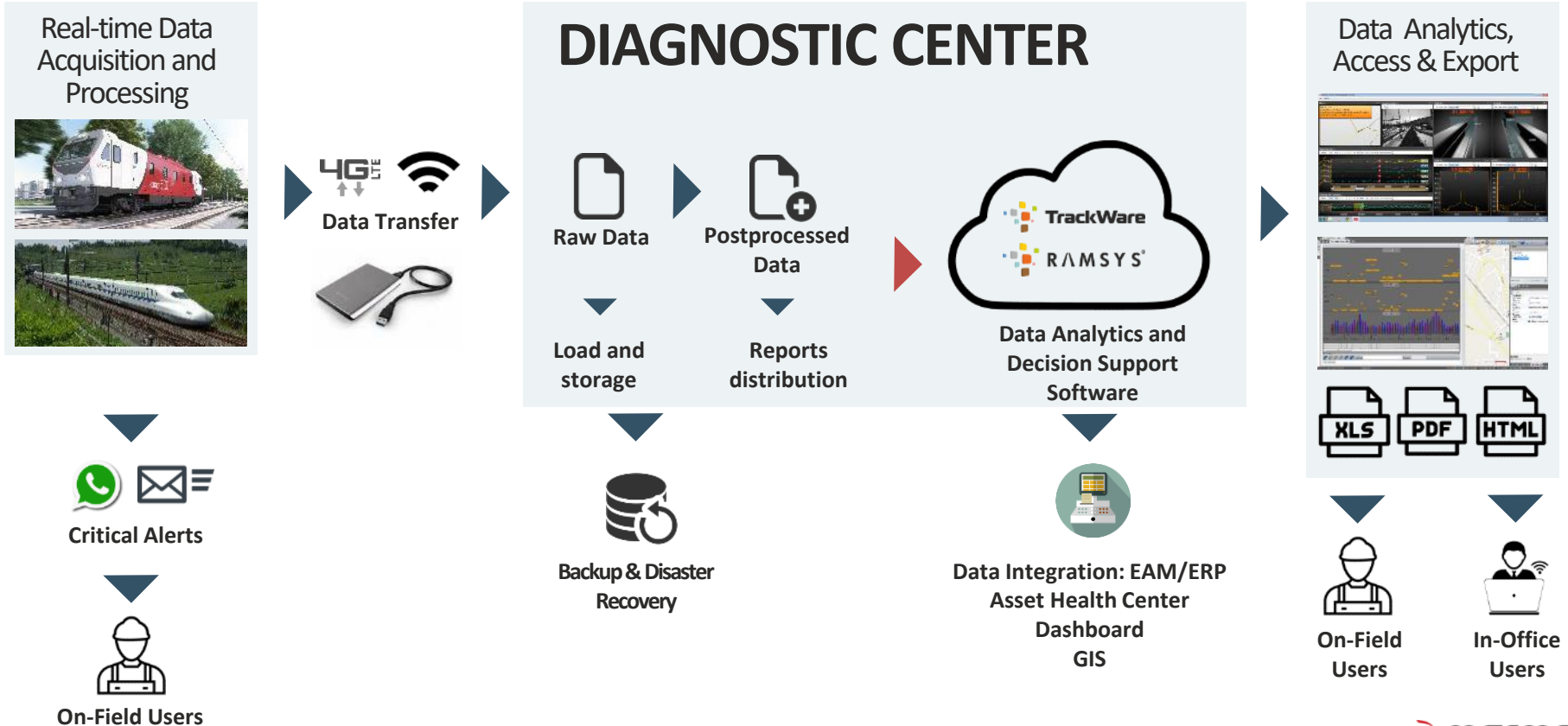
Real-time Data Acquisition and Processing

Data Transfer

DIAGNOSTIC CENTER

Raw Data

Postprocessed Data

Load and storage

Reports distribution

TrackWare

RAMSYS

Data Analytics and Decision Support Software

Data Analytics, Access & Export

XLS   PDF   HTML

Critical Alerts

On-Field Users

Backup & Disaster Recovery

Data Integration: EAM/ERP Asset Health Center Dashboard GIS

On-Field Users

In-Office Users

MERMEC

Agenda

mermec

From data collection to data distribution

Data Import → Data Alignment → Measurement Validation → Data Processing → Defects Review → Data Distribution

Data Alignment and Validation

Position Shift
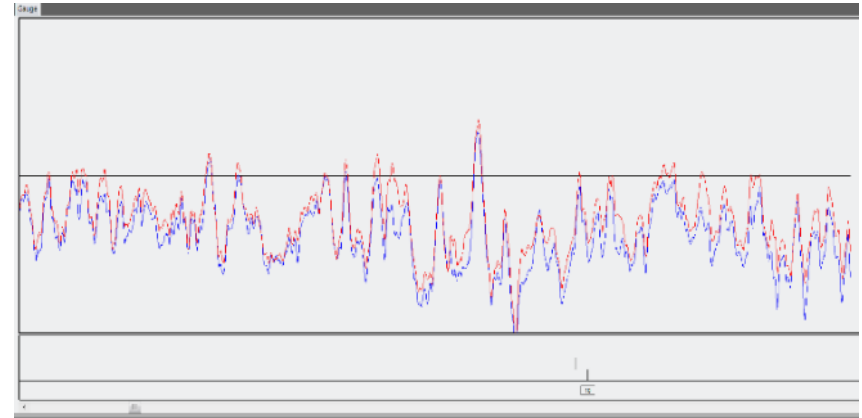
Noise
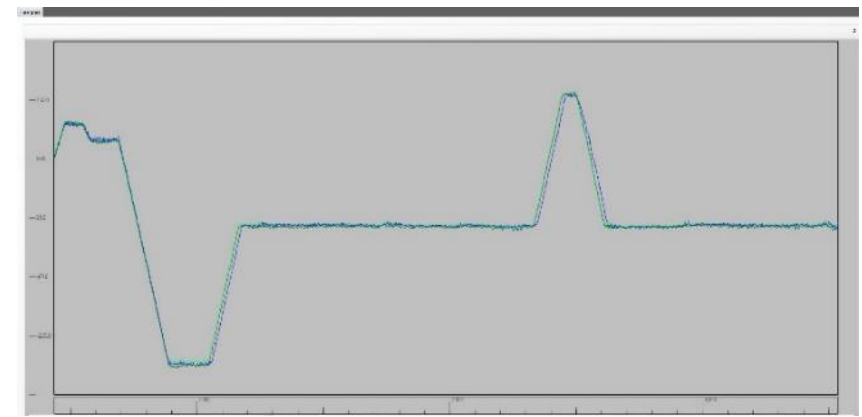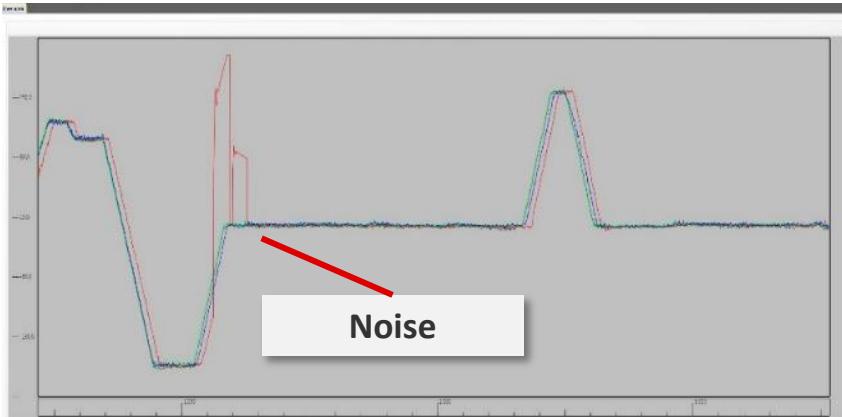
Syncronise different data in space

Planned Works

Key Performance Index/Score

Work History

Track Quality Index

Defect History

Infrastructure Features
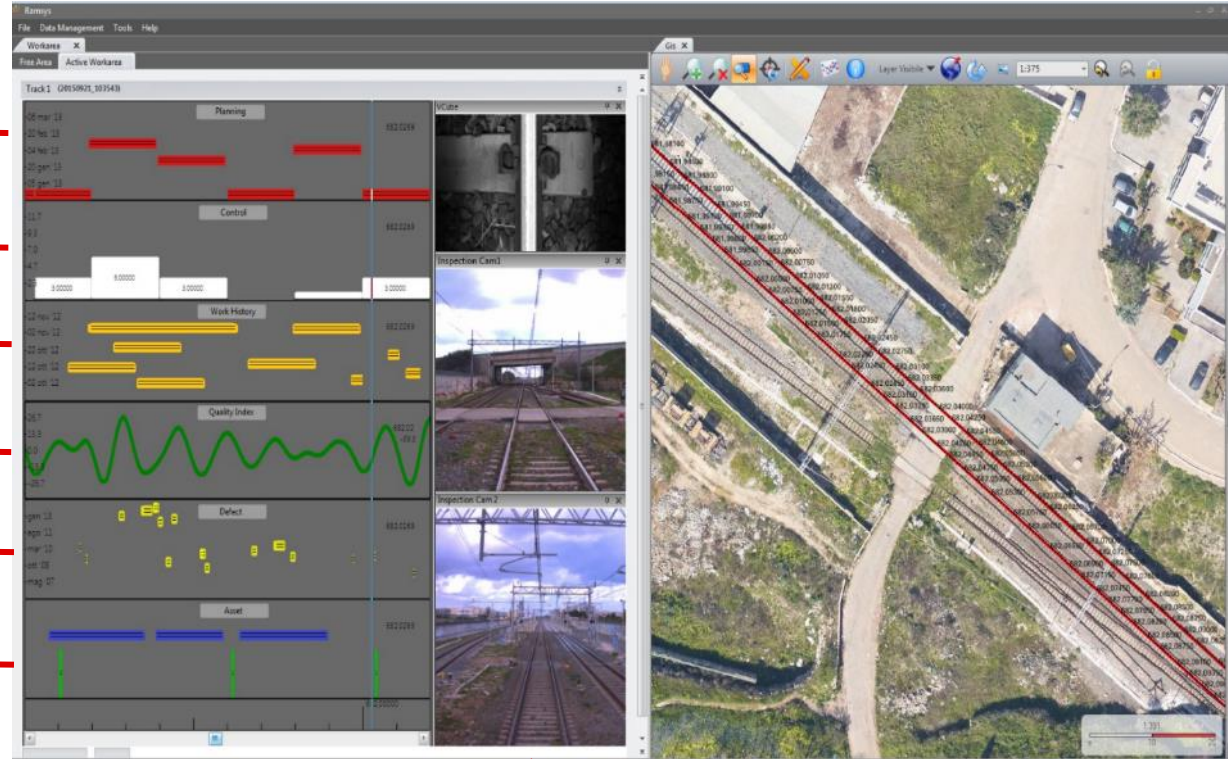
Video Frames

Geographical Data

MERMEC

Process data in time

TQI Deterioration

TQI Improvement
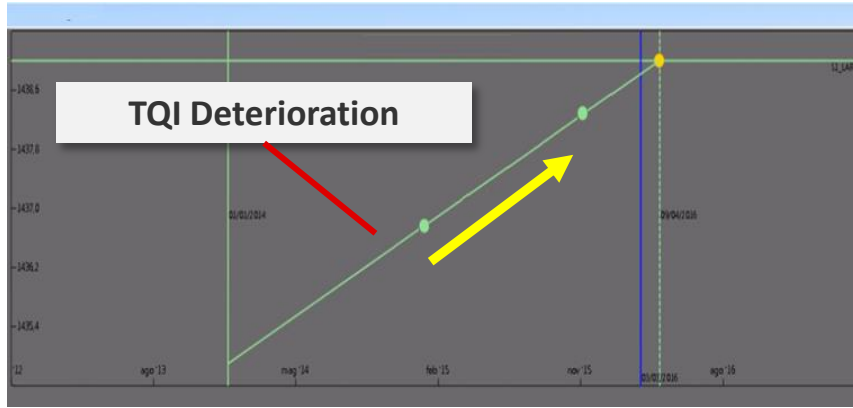
The rate of change can be calculated with only two Track Quality Indexes (TQIs).

Accuracy of the rate of change calculated with only two TQIs?
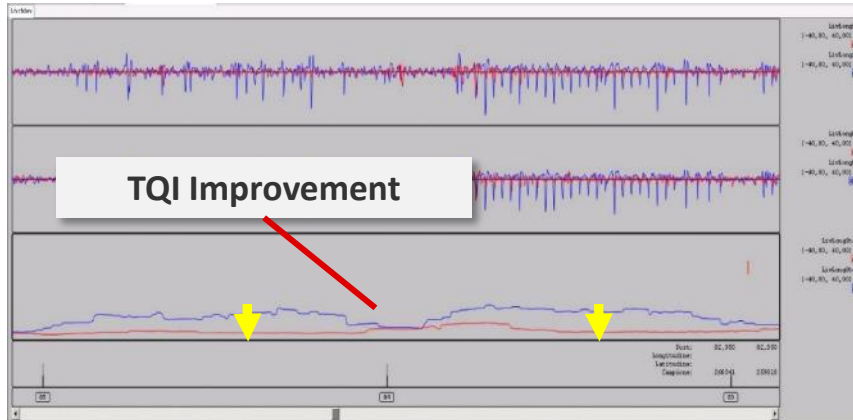
Prediction time interval
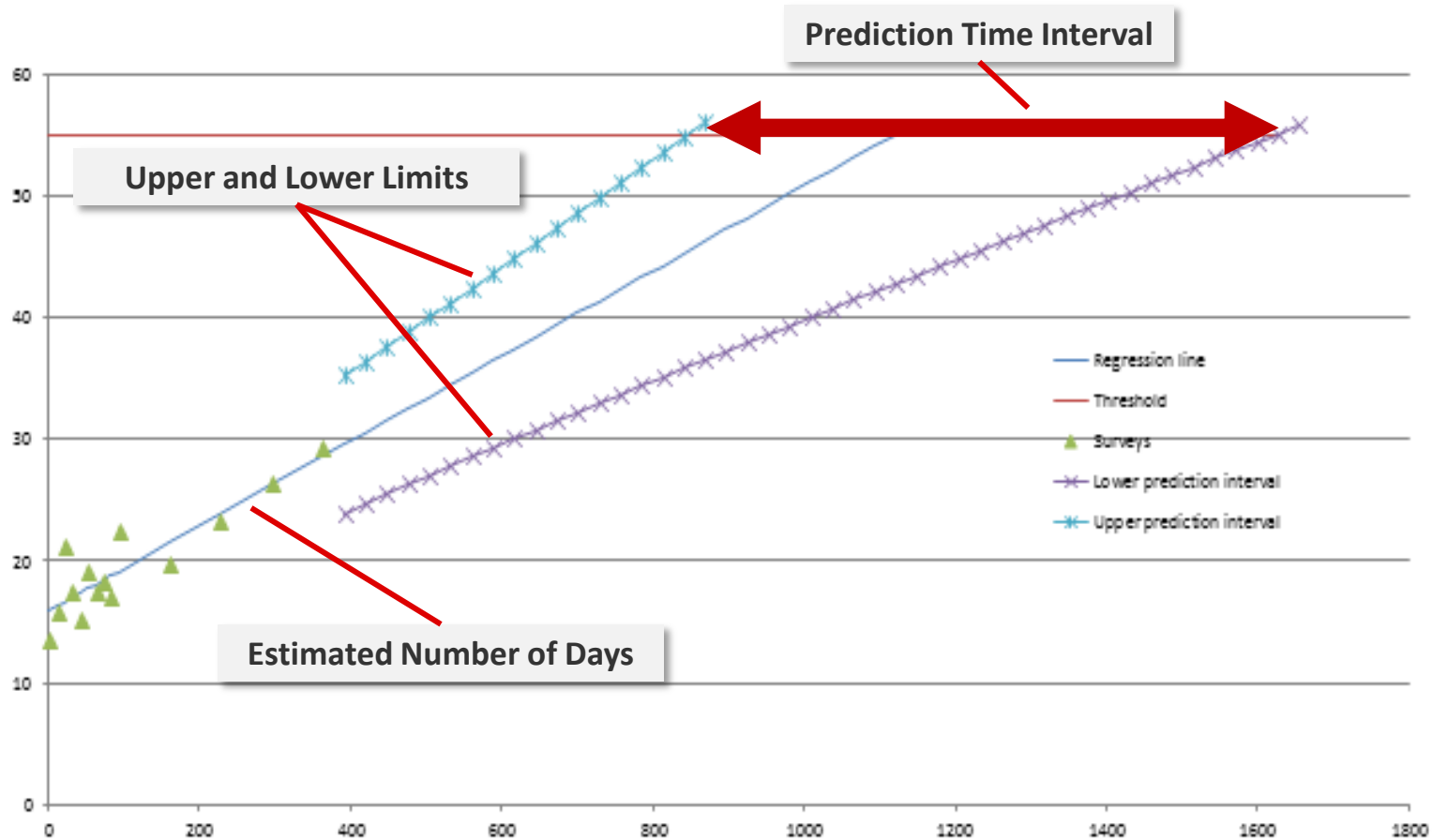
Improved accuracy in using bigger data
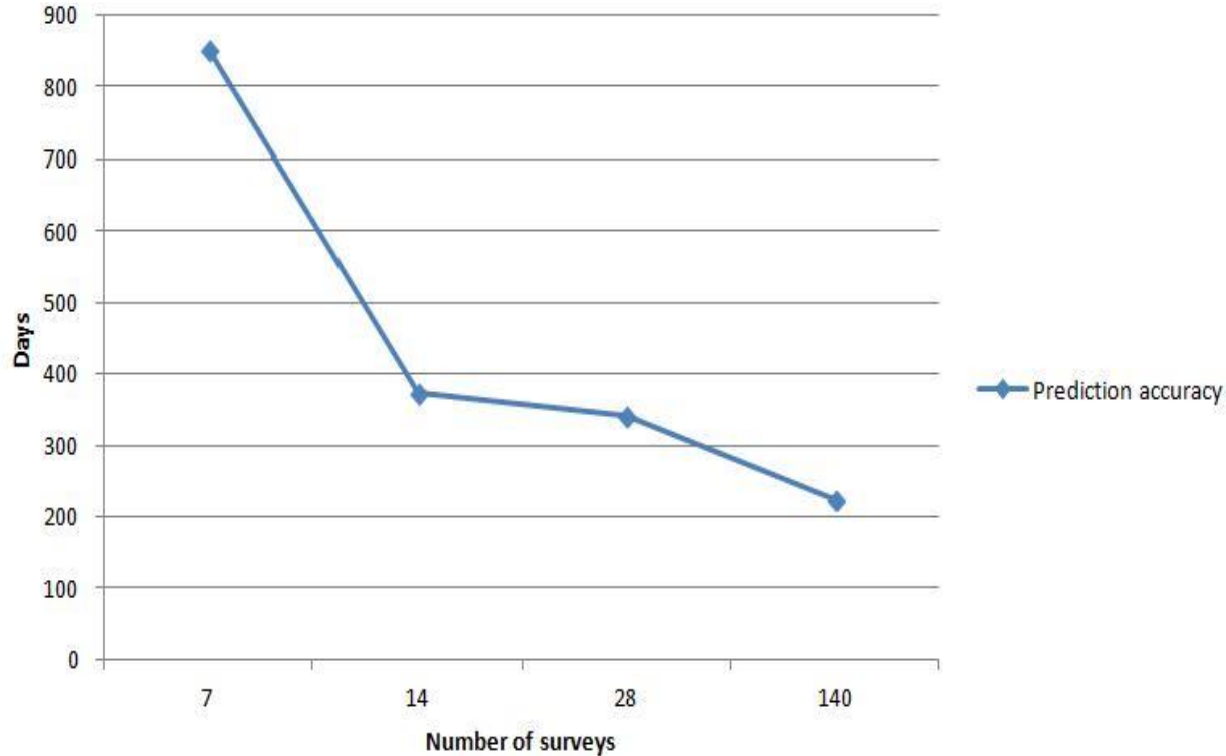
Agenda

mermec

Correlate ride quality with track data

Roughness of Twist

Repetition of Ride Dynamics Defects

Asset Data (Curve)
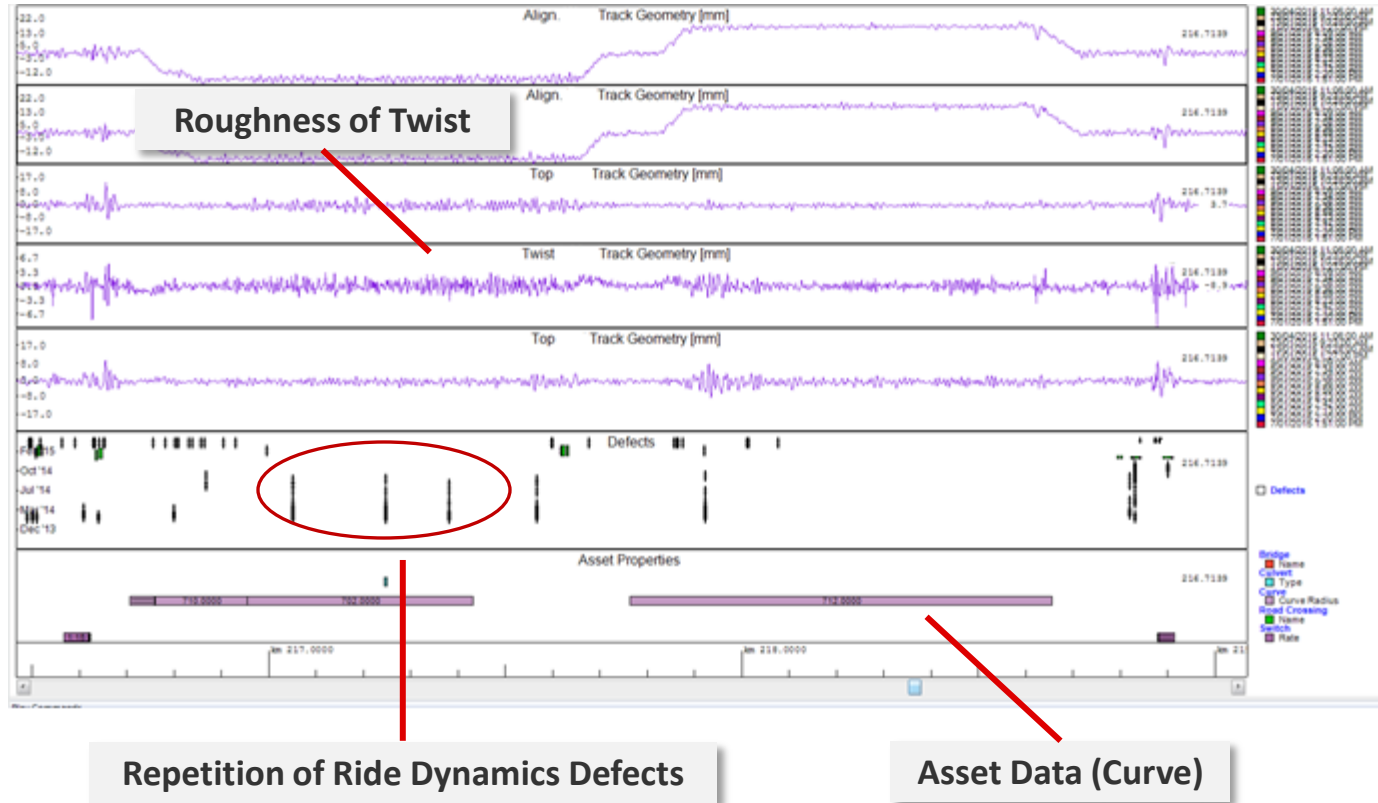
Create scenario of maintenance & renewal works

Future Works Scenario

Past Works

Efficiency & Efficacy of works
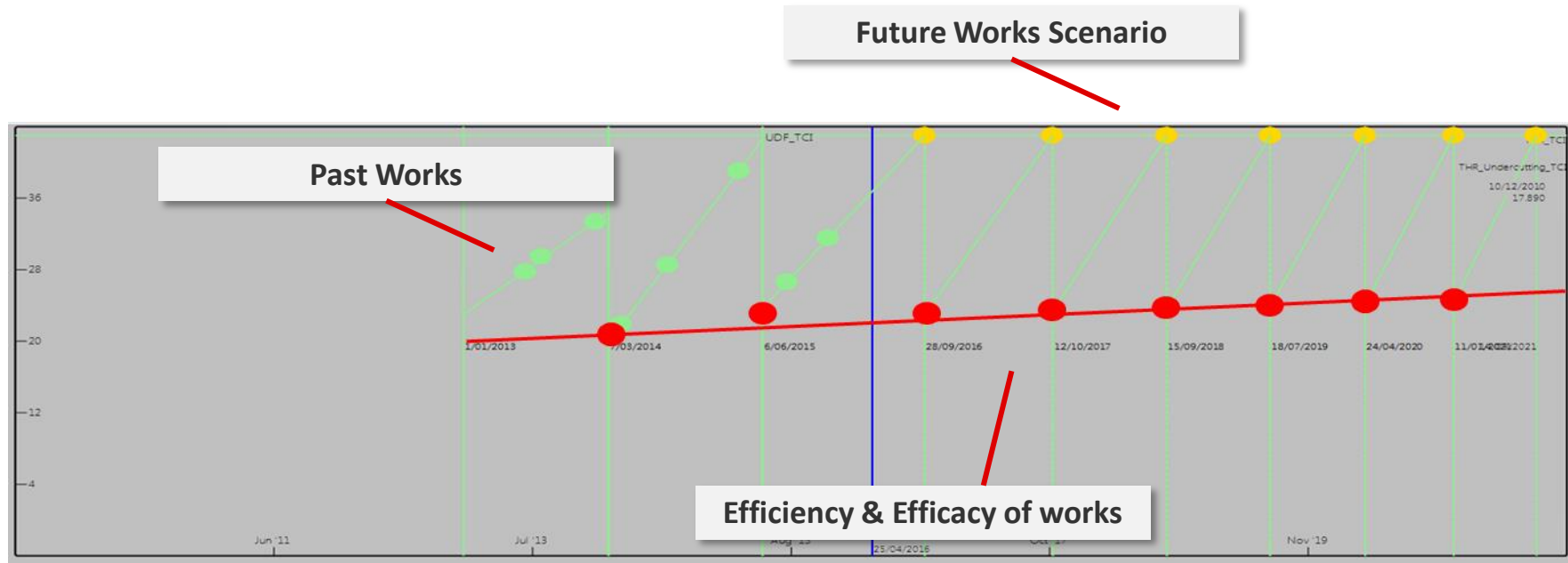
From prediction to prescription of a work plan

About 30% saving in life cycle costs

Total Actual Cost (NPV) = 590,000

Total Actual Cost (NPV) = 840,000

Agenda

1	Multiple conditioning monitoring devices
2	Big Data Characteristics (7 Vs)
3	Manage big data
4	Accurately monitor and predict asset health
5	Extraction of new information (3 use cases)
6	Conclusions

mermec

Conclusions

❑ Key characteristics of BIG DATA

❑ Key user requirements from current practices

❑ Big data collection (hardware) is not sufficient ....

❑ Two key questions for any railway:

(1) "How BIG is the decision support you are getting from current

SMALL and BIG DATA?

(2) Are you sure you cannot make it even BIGGER?"

Thanks for your kind attention!

"People who are really serious about
software
should make their own
hardware"
Alan Kay

mermec