HBM Prenscia Technology Day

Virtual Seminar Series

Data Analytics from Large Quantities of CAN Bus Data

Presented by Dr. Frédéric Kihm
Data Analytics and Vibration Product Manager, Prenscia HBK

Session #3: **Vehicle Testing & Simulation**

Data Analytics from Large Quantities of CAN Bus Data

Abstract

Engineers in the fields of test & measurement and maintenance are facing the need to process a mixture of various data types, including low-cost digital data from communication buses and connected vehicles. Bus data is an inexpensive source of readily available parameters in a complex electronic system. Accessing bus data offers huge potential benefits such as understanding customer usage from connected vehicles in order to improve the product validation process and thereby reduce unexpected failures. However, bus data also raises a number of challenges in terms of its analysis because of the quality of the data, quantity of data, inconsistency of data, lack of certain data, etc.

The solution would be a dedicated digital bus data processing tool that would perform analytics from huge quantities of unevenly sampled, heterogenous sensor data in a scalable fashion and at high speed.

The main benefit for the engineers would be for them to convert overwhelming data volumes into actionable decisions without the intervention of data scientists.

About the presenter, Dr. Frédéric Kihm

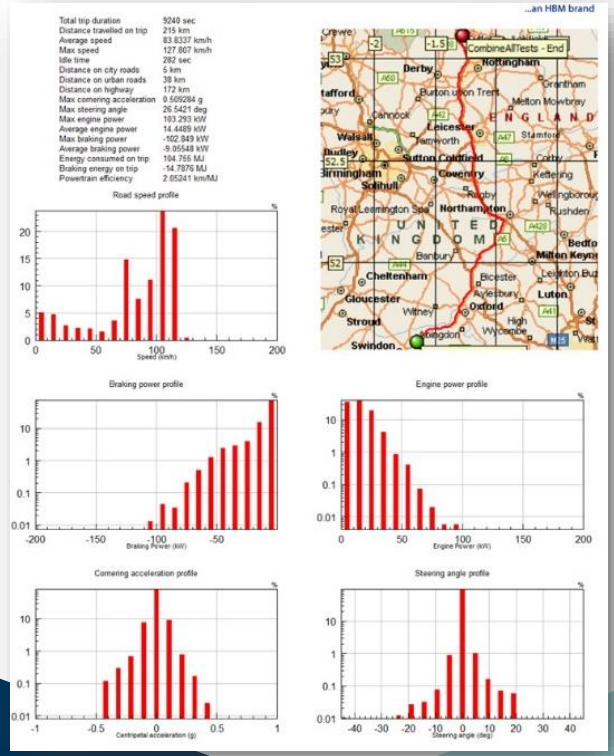
Frédéric Kihm is Product Manager at HBM Prenscia, responsible for the signal processing related software products, which includes GlyphWorks, VibeSys and nCodeDS. Frédéric previously worked as an Engineering Consultant for HBM Prenscia, involved with signal processing, fatigue/durability and vibration analyses in the automotive, aerospace, and defense industries.

Frédéric holds a MS in Mechanical Engineering from IFMA University in France and a PhD from the Institute of Sound & Vibration Research (ISVR) in Southampton, UK.

HBK
HOTTINGER BRÜEL & KJÆR

Data Analytics from Large Quantities of CAN Bus Data

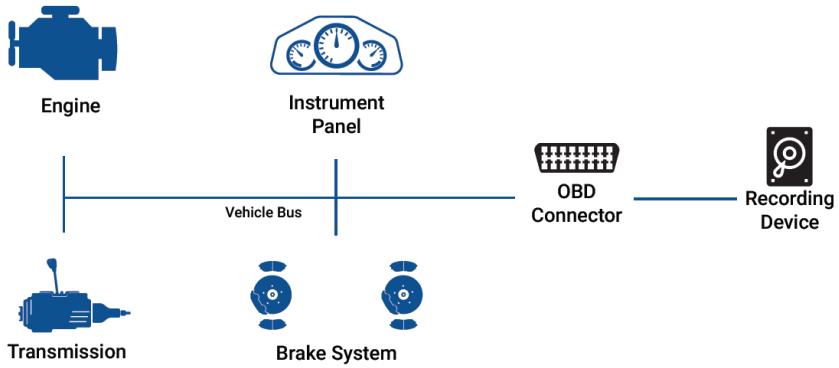
Dr. Fred Kihm
Product Manager – Analytics and Signal Processing



What is CAN data? What is it used for?

WHAT IS CAN DATA ?

- A digital bus connects several electronic systems together so they can exchange digital information.
- The CAN bus is a robust vehicle bus standard designed to allow microcontrollers and devices such as the Engine Control Unit (ECU) to communicate with other control units.
- It is a message-based protocol.



COMMON APPLICATIONS OF BUS DATA ANALYSIS

- Capture real usage to enhance design
- Validate ADAS and autonomous vehicles
- Demonstrate the reduction of energy consumption and emissions
- Enhance and validate numerical simulation

HBK
HOTTINGER BRÜEL & KJÆR

The challenges of processing CAN signals

- Non evenly spaced time-series data
  - Sparse and compressed data – avoid conversion to iso-sampled Time Series where possible!
  - Asynchronously sampled between channels
- Account for missing values, errors, NaN
  - Correct for missing data and anomalies
- Lots of Data
  - Need to be able to cope with huge quantities of data

#DSCSV				
#KEYWORDS				
Index	BreakStatus	LightStatus	TurnSignal	Engine Temp
#DATATYPES				
Double	String	String	String	Double
#DATA				
2.375271587	**ON**	NAN	**OFF**	0.144359342
2.678784655	NAN	NAN	NAN	0.149879337
5.300131698	NAN	NAN	NAN	0.728316792
6.674974961	**OFF**	NAN	NAN	0.746329859
9.047822865	**ON**	NAN	**LEFT**	2.314751245
9.894960469	NAN	NAN	NAN	3.123792842
10.67305112	NAN	NAN	NAN	4.613097105
24.3861021	NAN	NAN	NAN	5.498927251
26.69137779	NAN	NAN	NAN	5.703001804
30.74209588	NAN	NAN	**RIGHT**	5.911507592
31.24335127	NAN	**ON**	**RIGHT**	7.154817144
31.24949808	NAN	NAN	NAN	7.398343901
33.24835731	NAN	NAN	**LEFT**	7.533195658
34.52175349	NAN	**OFF**	NAN	7.718276165
37.19233678	NAN	**ON**	**OFF**	8.27933905
38.00122635	**ON**	NAN	NAN	8.683423675
40.24895699	NAN	NAN	NAN	9.051386883
40.98218412	NAN	**OFF**	**RIGHT**	10.43684393
41.27696247	**OFF**	**ON**	**RIGHT**	11.17062673

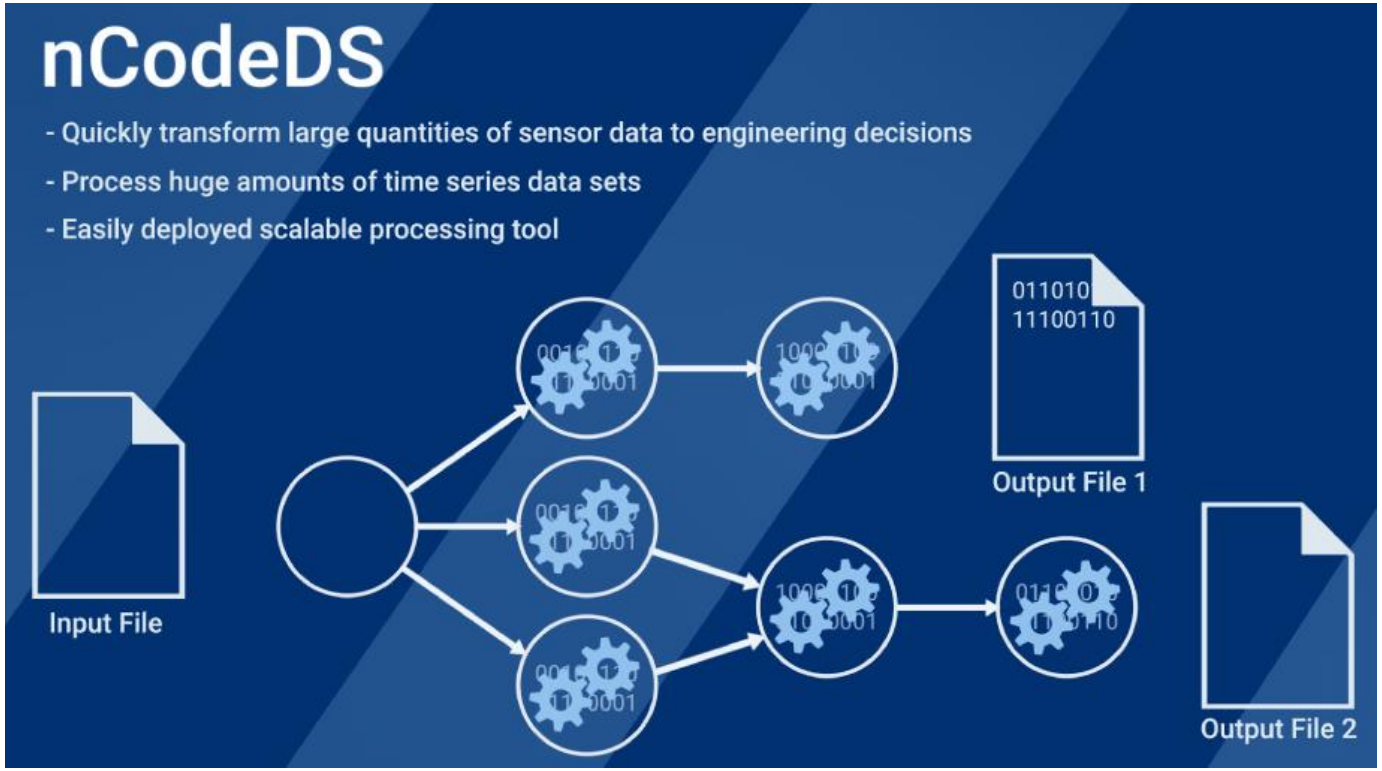
Non-numeric data

Presence of NaN (Not a Number)

Non-constant time steps

What is nCodeDS?

- nCodeDS is a server-side, scalable processing tool, which is part of Aqira.
  - Multi-threaded
  - Streaming framework
- The data can be unevenly spaced, it can contain NaNs, gaps, etc.
- nCodeDS provides a wide range of analysis capabilities.
  - Anomalies
  - Signal Analysis
  - Fatigue Analysis
- nCodeDS support a large number of file formats.
  - MDF3 / **MDF4** , CSV
  - Binary formats from traditional DAQ



HBK
HOTTINGER BRÜEL & KJÆR

How to use nCodeDS?

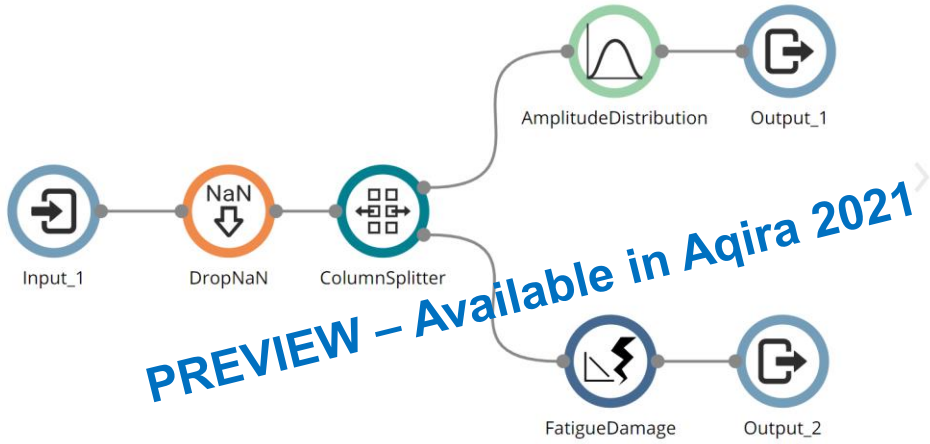
- An nCodeDS processing network can be scripted in Python
  - Widespread programming language
  - Very simple and flexible
  - Many libraries available (incl. Machine Learning)

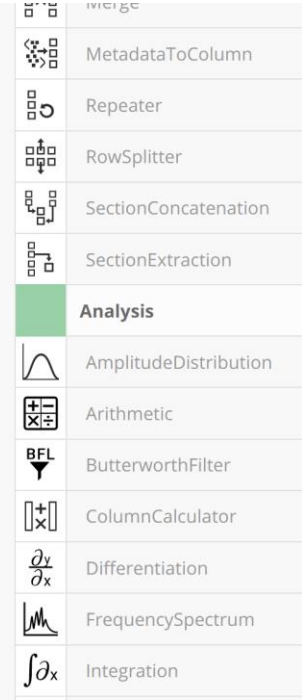
OR

- An nCodeDS processing network can be created graphically
  - Drag and drop nodes from palette
  - Fill in the property fields
  - Python node available

```
from nCodeDS.Network import *

try:
    netman = NetworkManager()
    #
    # We first declare the nodes that the network is going to use
    #
    netman["input1"]    =  Input(Connection = "dsfile://File.Filena
    netman["RStats"]    =  RunningStats(Operations = "RMS", WindowI
    netman["spikes"]    =  LimitDetection(UpperLimit = 100)
    netman["CreateMask"] = ColumnCalculator("func=MyColumnCalc", K
    netman["merge"]    =  Merge()
    netman["outputmerge"]  =  Output(Connection = "dsfile://File.F
    #
    # We then make the connections between the nodes
    #
    netman["input1"] >> netman["RStats"] >> netman["spikes"] >> ne
    netman["input1"] >> netman["merge"].inputs["Input1"]
    netman["merge"] >> netman["outputmerge"]
    #
    # Run it
    #
    netman.run()
```

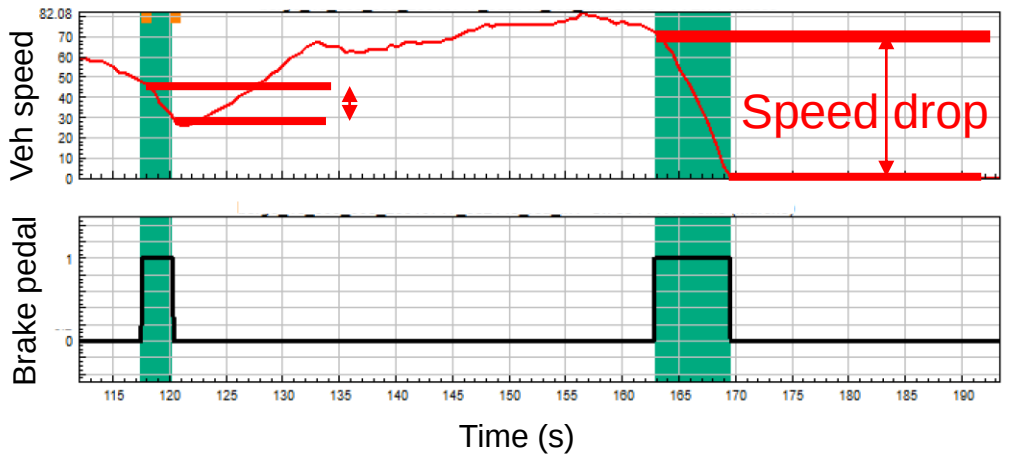


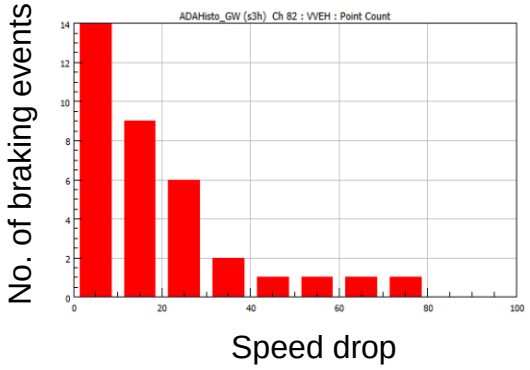


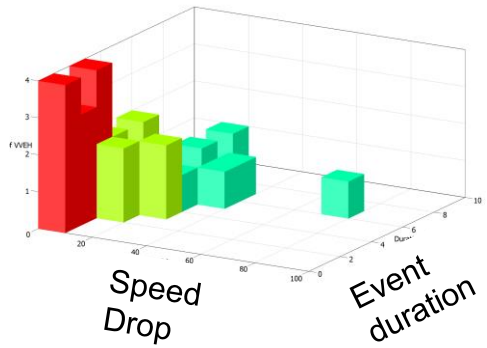
HBK HOTTINGER BRÜEL & KJÆR

Example Application : Analysis of data subsets that meet specified criteria

- Select all braking events and the corresponding vehicle speed
- Calculate vehicle speed drop everytime the brake pedal is actuated
- Build a histogram showing No. of braking events for various ranges of vehicle speed drop



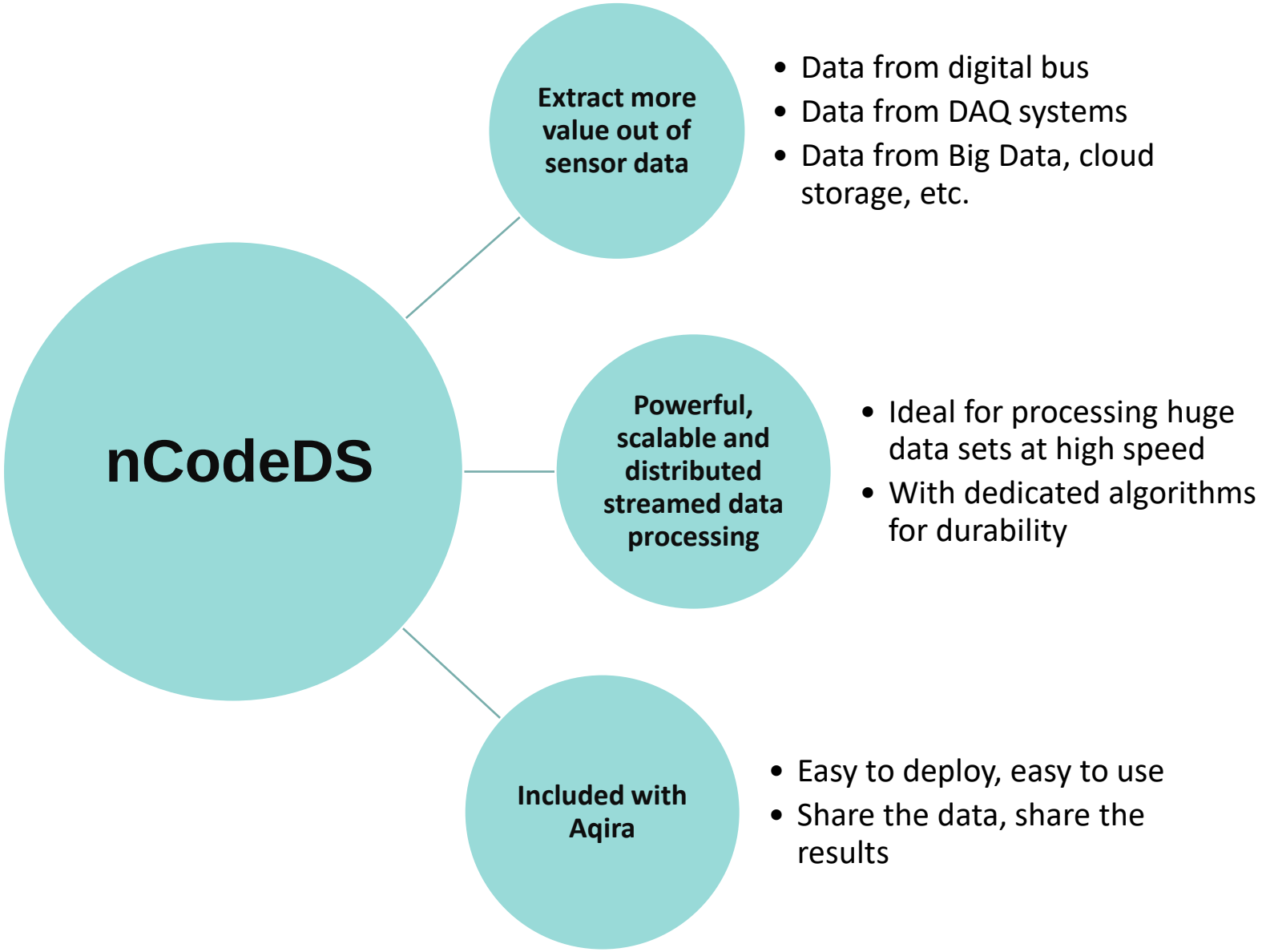




HBK HOTTINGER BRÜEL & KJÆR

Conclusion

- The availability of CAN data is beneficial to test engineers
- However, processing CAN data comes with a number of challenges
- nCodeDS can help engineers extract insights from huge quantities of CAN data



HBK
HOTTINGER BRÜEL & KJÆR

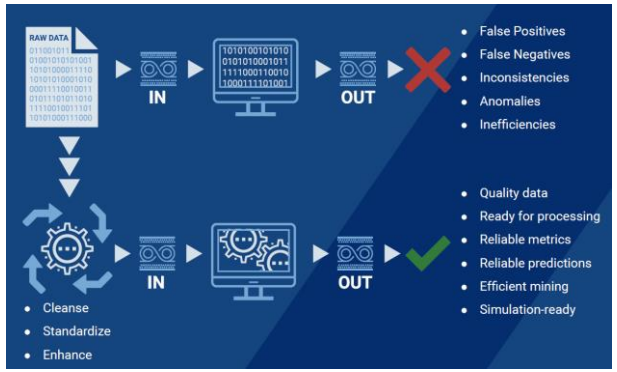
Previously Recorded Webinars – Digital Streams and Data Analytics

nCodeDS enables engineers who are overwhelmed by measured sensor data to gain **actionable insights from high performance data analytics**.

Accelerate Decision Making with Faster Analysis and Data Search

Originally presented on May 13, 2020

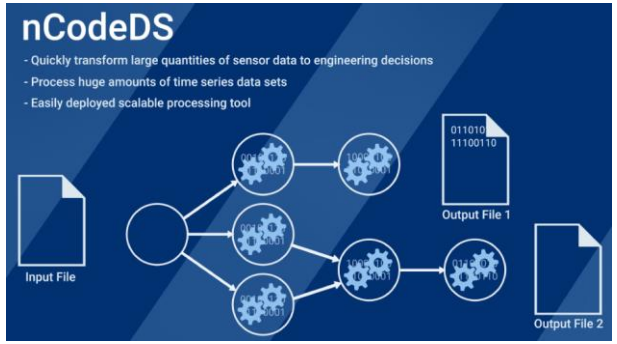
https://www.ncode.com/videos/accelerate-decision-making-with-faster-analysis-and-data-search



Big Data Analytics for Engineers

Originally presented on December 11, 2019

https://www.ncode.com/videos/big-data-analytics-for-engineers



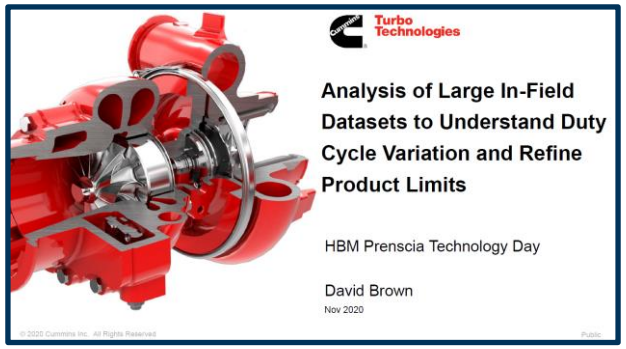
Actionable Insights from Digital Streams with nCodeDS

Originally presented on May 15, 2019

https://www.ncode.com/videos/actionable-insights-from-digital-streams

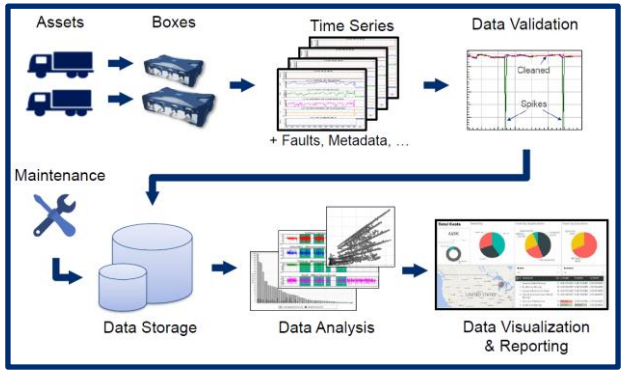
HBK HOTTINGER BRÜEL & KJÆR

Session #4: Vehicle Fleet Data Analytics



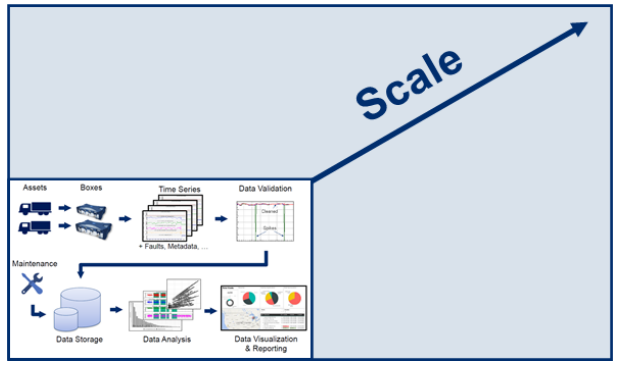
Analysis of Large In-Field Datasets to Understand Duty Cycle Variation and Refine Product Limits

Mr. David Brown,
Cummins Turbo Technologies | Group Leader – Low Cycle Fatigue



VePRO - A Program Overview – Scalable CBM

Dr. Mark A. Pompetzki,
Prenscia Engineering Solutions | Director – Engineering Solutions



Application of a Vehicle Data Simulator to Big Data CBM Systems

Mr. Iain J. Dodds,
Prenscia Engineering Solutions | Head of Integrated Solutions

Thank You

https://www.ncode.com/products/ncode-streamed-data-analytics

https://www.hbmprenscia.com/webinars

HBM

Brüel & Kjær

PUBLIC

HBK
HOTTINGER BRÜEL & KJÆR

Q&A

- Ashish Jahagirdar (Fiat Chrysler, India)
  - What do you mean by Node? -- is it using Haddop Mapreduce in the back end?
    - *In this case a "node" is a digital signal processing step within the process workflow or "network" – see the presentation slide #7 "How to use nCodeDS". This is not using "Haddop Mapreduce" in the back end. Instead, we are using our own multithreaded streaming framework which is designed to handle sensor data in a highly optimized way.*
  - Can you do time series forecasting?
    - *Yes, to the extent, that time series forecasting can use simple statistical trending models (running statistics) and use these for near future trend modelling. Python "nodes" allow the user to embed more complicated machine learning time series forecasting models as required for specific parameters. The user has to program the actual forecasting capability required for their modelling, leaving nCodeDS to handle the data import, data throughput and any pre- or post-processing. (https://en.wikipedia.org/wiki/Time_series)*
  - SPN FMI Codes is way to identify problems related with any system?
    - *Yes, there are many different fault codes for different systems. nCodeDS does not come with a library of system fault codes. A user can define the fault codes they require for their analysis in their network. Deeper analysis can then occur based on the fault code(s) of interest: the user can count them, use them as a trigger to extract specific sensor data before and after they occur to eventually understand what is causing them, etc.*

Q&A

- Eduardo Calixto (ECC, Germany)
  - Which are the usual CAN failure modes and effects ?
    - *Our customers use the information transmitted through the CAN network to do reliability on the various components in a car: they analyze the various sensor information (temperature, speed, pressure, etc.) or the error codes.*
    - *The data transmission via the CAN network can have its own reliability issues. First, it is worth noting that such issues are of paramount importance because of the role of data transmission in modern cars: it could be extremely dangerous for the driver if faulty data is transmitted. As far as I know, the CAN protocol has built-in failure detection mechanisms that were designed so that the probability of corrupted data to be undetected throughout the life of a vehicle is extremely small.*

- David Brown (Cummins, UK)
  - Very interesting presentation discussing the common challenges with CAN data - are there any studies showing the time savings of nCode DS vs nCode Glyphworks ?
    - *Yes, we have done a number of speed comparisons. The results vary depending on the characteristics of the input data (evenly or unevenly spaced data, quantity of data, no. of channels, file format), the analysis (what algorithms, how many nodes/glyphs) and the computer (number of cores, memory). Overall, nCodeDS is between 5 and 10 times faster than GlyphWorks in the majority of our test cases, increasing to 100 times faster for some specific cases, and decreasing to only 3 times faster for others.*

HBM Prenscia Technology Day

Virtual Seminar Series

Data Analytics from Large Quantities of CAN Bus Data

Presented by Dr. Frédéric Kihm
Data Analytics and Vibration Product Manager, Prenscia HBK

Session #3: **Vehicle Testing & Simulation**