Architecture for Scalable, Self-*, human-centric, Intelligent, Secure, and Tactile next generation IoT

ASSIST-IoT Technical Report #6

From OBD to Connected Diagnostics: A Game Changer at Fleet, Vehicle and Component Level

Carlos Guardiola, Christian Vigild, Frederik De Smet, Klaus Schusteritz

Submitted to: 6th IFAC Conference on Engine and Powertrain Control, Simulation and Modeling (E-COSM 2021)
From OBD to connected diagnostics: a game changer at fleet, vehicle and component level

Carlos Guardiola ∗ Christian Vigild ∗∗ Frederik De Smet ∗∗ Klaus Schusteritz ∗∗

Abstract: Early on-board diagnostics (OBD) standards were enforced in 1988 and, by the beginning of the XXI century, all major automotive markets require some sort of OBD. Over the years, the diagnostics software layer has grown in complexity, yet robust fault detection remains a challenging task: insufficient memory and computation power, suboptimal calibration, and the lack of sufficient real-life operation data for model development are some of the limiting factors. The connected vehicle paradigm allows a complete reshaping of the vehicle diagnostics: real-life data feeds provide operation data of the vehicle fleet; artificial intelligence assists data clustering and model development; and over-the-air update tools allow the deployment of new software components and optimized calibration. A smart combination of embedded and cloud components seems to be a major step forward for the next generation of vehicles, allowing the determination of the in-service emissions at vehicle and fleet level. While many challenges are still to be solved, connectivity offers a giant leap in the area of vehicular diagnostics.

Keywords: internal combustion engine, engine control, diagnostics, connectivity, IoT, OBD

1. INTRODUCTION

Information and communications technologies (ICT) are playing a major role in the current technological scenario: last few years have seen a plethora of new services being developed, and some classical, mature markets are being significantly transformed by the ICT penetration. This is not the case for the propulsion system control and diagnostics, since there has been a limited impact of ICT to date, and control and diagnostics software is basically designed as it was done 10 years ago. It is easy to predict that this will change in near future, but as the so-called Amara’s Law (Amara, 1984) states that we tend to overestimate the effect of a technology in the short run and underestimate the effect in the long run. In this paper we focus on some foreseen implementations that could effectively be adopted within the next few years. Specifically, we will focus on propulsion system diagnostics because, to the contrary of low-level control algorithms, it does not directly compromise drivability, thus allowing a higher flexibility in over-the-air actuations. In this introduction, some of the drivers of the proposed transformation are first reviewed. Arguably, emission legislation has been one of the major drivers for the technological evolution of the automotive propulsion systems. While regional differences exist, three major action priorities may be identified from a general perspective:

(1) The search of low emission technologies. For that, worldwide legislations have applied a progressive and persistent decrease of the limit of the regulated emissions, combined with the inclusion of additional species along the years (e.g. particulate number for diesel engines in Euro 5b and for gasoline engines in Euro 6). As shown in left hand plot in Figure 1, EU regulated limits for diesel engines have been cut more than 90% for NOx and 97% for particulate matter from Euro 2 to Euro 6d. Additionally, CO2 emissions are since a few years being regulated at a fleet level.

(2) Ensuring that the emissions in real life operation are according to the legislated values. Initial efforts were based on the increase of the representativeness of the certification cycle for real-life operation (Mock et al., 2012). As a consequence, old cycles as the NEDC (top-central plot in Figure 1), originally designed for comparing emissions and consumption between different models, have been replaced by more representative cycles as the WLTC (bottom-central plot). This trend has resulted in the later introduction of real driving emission (RDE) verification using portable emissions measurement systems (PEMS), where the vehicle is tested in real life conditions (right plot in Figure 1). Here, the driving itself is unknown, yet the order and the street characteristics are known (urban, rural district and highway sections are included).

(3) Ensuring that the emission level does not vary along service life. Here again a couple of mechanism have been enacted. On one hand, onboard diagnostics targets the identification of faults impacting vehicle
While ICT themselves cannot modify the bare limit of a given technology for satisfying a given emission target, they can effectively change the way the system is monitored, diagnosed and controlled. As it will be discussed in this paper, it seems now possible to have an onboard system able to provide an estimation of the in-service emissions at an individual level, and propagate it to a cloud service for getting metrics at fleet level. This is a game changer in the sense that a real-time feedback will exist and will not be restricted to discontinuous testing of fleet samples (as done in ISC verification). Such connected vehicle paradigm allows reshaping the vehicle diagnostics: real-life data feeds provide operation data of the vehicle fleet; artificial intelligence assists data clustering and model development; and over-the-air update tools allow the deployment of new software components and optimized calibration. A smart combination of embedded and cloud components seems to be a major step forward for the next generation of vehicles. While many challenges are still to be solved, connectivity offers a giant leap for vehicular diagnostics.

2. ENABLERS.

A series of technologies will be necessary for the effective implementation of this new approach to the diagnostic of the powertrain. Note that a given technology can be considered an enabler or an inhibitor depending of its state of maturity. Next, some of the main contributing technologies (and their current limitations) are highlighted. The implementation of the framework proposed in this paper will need:

1. Communication infrastructure with sufficient bandwidth for an effective communication in the different dimensions: Vehicle to vehicle (V2V), Infrastructure to vehicle (I2V), and Vehicle to infrastructure (V2I). Advantages are gained even from lower data intensity implementations (2G—3G), and may be maximized for data rich scenarios (4G—5G). From the propulsion system perspective, connectivity provides the possibility of propagating emission and sensor data from the vehicle to the OEM, and of deploying new software components on demand. Note that connectivity may be intermittent and, in the limit, some vehicles can stay off connection for most of their service life. This compromises cloud-based implementation of critical tasks with real-time requirements. On the contrary, over-the-air update of control and diagnostics firmware could be periodically done without necessarily affecting drivability.

2. On-board computational power and data acquisition capabilities. Modern ECUs have significantly overcome some of the past limitations concerning data acquisition of analog channels. As an example, knock sensor is usually connected to high speed software configurable analog acquisition channel, able to provide a proper angle windowing and filtering of the signal; on the other hand, last generation ECUs have implemented with success in-cylinder pressure measurement, where crank-synchronous acquisition must be satisfied. In both cases the analysis of the signal is cycle-to-cycle and cylinder-to-cylinder, which demonstrate the capability of the production ECUs to handle demanding acquisition tasks. Concerning computational power, modern ECUs have moved to multi core architecture, and are now able to offer a computational power sufficient for implementing complicated signal processing and model-based optimization. Most of the control and OBD tasks are usually scheduled in a real time approach, needed in many cases for ensuring the safe operation of the system.

3. Cloud computing power and ability to deal with big, growing quantity of data. Artificial Intelligence (AI) techniques applied to the collected data allow the (semi-) automated development of diagnostic models. Specific methods are needed for dealing with unstructured data: on one hand, the fleet may be comprised of vehicles with varying sensor sets and different control firmware versions; on the other hand, because of communication bandwidth constraints and the possibility of intermittent connectivity, the system must be able to integrate incomplete data, and data streams are expected to vary in the number of channels and frequency.

4. Reliable and cost-effective sensors for the more important emissions and operation variables of the internal combustion engine. Last years have seen a significant advancement in automotive sensor development: in addition to classical lambda sensors, air mass flow meters and various temperature and pressure sensors, some of the currently available sensors cover more relevant emissions as NOx and particulate matter (Payri et al., 2014). However, the accuracy and stability of the different sensors should in this context always be kept in mind. Without a periodic recalibration procedure, sensors may be subject to significant drift adding to the sensor dispersion (Mora et al., 2018), while some of the sensor exhibit significant cross sensitivity to other species as it is the case for NOx sensor which is affected by NH3 (Hommen et al., 2017).

5. System models. In addition to sensors, system modeling for the air-path is state-of-the-art, and model-based control has been implemented to a high extent.
Although gas pressure, air mass flow and concentration sensors are widespread, observers and system models are widely used, and sometimes used as feedback quantities to improve the overall system accuracy (Guardiola et al., 2019c,a) or for implementing model based failure detection (Angelov et al., 2006). Many of these models run in modern ECUs, and the results of the models are stored as internal ECU variables.

Because of the latter two points, it is expected to have a rich data scenario, where a combination of sensed and calculated values will be available. However, because of cost constraints and existence of a large amount of engine variations, not all engines may feature the same set of sensors. Thus, engines with different quantity and quality of sensors may be present for the same vehicle variant.

### 3. CONNECTED DIAGNOSTICS

**Classical development process.** The classical development process for the ICE control and diagnostics system is sketched in Figure 2. The whole design approach is essentially heuristic, and based on a pre-defined static knowledge of the system.

The design approach is affected by the following issues:

- **System performance.** The classical control system is hardly ever optimal as it has to cover all thinkable operating conditions with a fixed calibration set. Thus, the same control system should perform the same in Finland, or for that matter Spain which could differ significantly in its requirement to optimality. Thus, the classical method may not adapt to varying use cases and constraints when new information is present.

- **Robustness and adaptability.** Adaptability would be a key factor to improve engine performance; indeed, the same engine is often used in different variants to fulfil specific market and use requirements. Engines are subject to manufacturing discrepancies and ageing, but also exposed to different driving and environmental conditions. Fuels are also different and multi-fuel engines are expected to play an important role in the future. For the case of the OBD algorithms, a large number of variants may be required to cover the different use cases.

- **Time to market.** The SW structure is often incrementally updated by adding independent control loops for each subsystem and their individual new requirements. This way the control structure becomes larger and more complicated with each new powertrain generation. Thus, state-of-the-art engine control systems contain several thousands of tunable parameters in their control structures, contributing to the overall increase of the system complexity and intractability. Additionally, system calibration must be finished before the engine gets to the market, thus precluding the evolution of the control and OBD system along the vehicle life. The latter requirement is dictated by the current homologation procedure.

**Connected diagnostics.** Harnessing ICT, embedded system can be complemented with a remote system, as sketched in Figure 3. There the vehicle is integrated in a decentralized system allowing a secured communication with the system and the different actors. The vehicle acts as an intelligent far-edge node, with computing, data processing capabilities, and in charge of the low-level control of the system. ECU software can be updated on demand, and new software functions deployed, catering for a dynamic modification of the code in the embedded system throughout the lifetime of the vehicle. Furthermore, the system provides a framework for systematic data collection. Machine learning techniques applied to such a large data collection will provide growing system knowledge. Areas such as performance and operational characteristic should be mentioned in this context too. From there on it is possible to derive conclusions on the specific vehicle under consideration, but also on the vehicle series, the driver or for that matter the traffic or use conditions.

The combination of a high-fidelity data scenario, sufficient computational power and data transfer bandwidth, along with the ability to deal with a huge amount of data, caters for a completely recast the current onboard diagnostics (OBD) system. In the proposed approach, whose main elements are shown in Figure 4, a part of the diagnostics software is run on the embedded controller. The rest of the system, however, runs on the distributed nodes located in the edge system infrastructure. The system connects via 12V and V2I to the individual powertrain controllers, dynamically. Since most of the diagnostics in the OBD layer are not compromising drivability nor system safety,
requirements to real-time performance may be relaxed. This allows the decision of the part of the diagnosis methods to be taken in distributed nodes in the infrastructure. The selection of the parts of code to be run on-board and on the edge node services must be done according to the trade-offs between available computational power, bandwidth of the data transmission, and the foreseen benefits of using machine learning techniques. In this sense, transferring the diagnostics to a distributed edge service allows a drastically different approach: while ECU software usually works like a microcontroller (memory is scarce, and deterministic time step is used), the service will allow more processing power and capabilities thanks to its decentralized nature and the ability to storing and exchange secure data for every vehicle and from the beginning of its life. On the other hand, streaming a huge quantity of data may be challenging. In the case of high-frequency signal sampling, some preprocessing and parameterization of the signal in the embedded controller may be required before broadcasting to the infrastructure. An example: the embedded controller may run cycle-to-cycle and cylinder-to-cylinder diagnostics, but the data size must be reduced before it can be streamed effectively.

On the decentralized system, machine learning tools will be used for deriving new diagnostics methods, which can be run on historic data on the database or may request data streaming from the different units in use. Parts of code may be downloaded to the different embedded controllers for fast signal processing. The system may also integrate information coming from technical inspections facilities (as emission measurement tests or mechanic verifications), from engine workshops (as information on performed maintenance operations), and from visual inspection facilities (which detect the existence of damages on the vehicle surfaces).

In the next sections a few applications of such new framework are highlighted, and the advantages over the current static onboard approach for emission monitoring and system diagnostics identified.

4. MONITORING FLEET IN-SERVICE EMISSION

As discussed in the introductory section, emission level should be kept within the legal limits along the vehicle life. While ISC mechanism offers a verification of this by periodic testing of a limited sample of the fleet, the proposed connected diagnostics system offers a more universal approach – taking a larger portion of vehicles into account and considering real-life use of the fleet. How a cost effective, efficient, yet sufficiently precise method of fleet data acquisition could look like, which at the same time is less dependent on ISC test procedures, remains open. A consequent solution to address these demands is to:

1. Focus on fleet level emissions rather than single vehicle emissions, since emission level will vary depending on the individual circumstances of a given vehicle.
2. Increase fleet wide emission sensing capabilities, using additional sensors when considered for a part of the fleet.
3. Use statistical methods to prove the fleet in total is meeting ISC requirements within a desired confidence interval.

Current propulsion systems are already equipped with several emission related sensors, designed to provide lifelong data to operate the vehicle within its legal limits. While legal vehicle operation is guaranteed, the accuracy over lifetime might decrease, due to known aging effects like sensor drift and of course the accuracy level is not comparable to research-grade test equipment like Portable Emissions Measurement Systems (PEMS) in general. One obvious solution would be to design a new generation of emission sensors, able to overcome aging effects within the harsh propulsion system environment, however this might not be the preferred solution in a cost sensitive environment like the automotive industry. A potential solution is the usage of selected standard production sensors (for example standard NOx sensors), which are on the upper end of the allowed production accuracy spread and thus increasing emission sensing capabilities, compared to their counterparts at the lower end of the production spread.

High-fidelity (HiFi) sensors can easily be identified with the end-of-line data already being provided by the sensor supplier, thus allowing a separation of standard and highly accurate sensors from the same batch of sensors in the vehicle manufacturing plant. In order to provide an accurate yet cost efficient solution all vehicles are equipped with a standard sensor, which is sufficient for vehicle operations during the whole vehicle life, even considering the previously mentioned aging effects. Only a small yet statistically sufficient amount of vehicles will be equipped with a HiFi sensor in addition (of course the electronic system design for the vehicle needs to allow installation of optional sensors, and physical space constraints must also be considered). In order to ensure unbiased HiFi sensor operation, these sensors are subject to exchange on a regular basis, e.g. during standard vehicle maintenance appointments without bothering the owner. This drift compensated operation of HiFi sensors ensures a sufficient number of vehicles run in parallel to their standard counterparts. On the basis of the vehicles with both the standard and the HiFi sensor, statistical information on...
standard sensor drift and spread can be computed. Such information can be used in the observer design for the emission of the vehicles using standard sensors, which in turn will allow to get fleet wide emission distribution based on real world driving.

Obvious benefits are:

- All vehicles are monitored permanently, thus the amount of dedicated, costly and time consuming ISC tests can be reduced significantly, if not abandoned completely.
- Harsh driving situations can be put into a statistical context. Both legislator and OEM can focus on statistically relevant driving situations to further improve fleet level emissions reduction efficiency, instead of spending resources and time potentially on very harsh yet statistically irrelevant driving situations.
- Data which is transparent and trustworthy can be shared between the OEM, the legislator and any third party, e.g. environmental NGOs to further increase trust and allow discussions on a par. In this sense, data integrity technologies may be used to secure the data transfer from the initial sensor measurement and the final receiver (e.g. an external party).
- Fault detection monitors of emission devices have varying capabilities. Whereas electrical system diagnostics can often be very precise (±10% degradation can be robustly detected), for devices like catalysts and EGR valves, the degradation levels that can be detected robustly are much higher and vary strongly from device to device. For some devices only a total failure can be detected. A fleet monitoring approach guarantees meeting an average emission level, which allows to push individual monitors to their limits, while avoiding requesting a repair at a fixed vehicle emission level without any fault detection.

With the gained understanding of the fleet level emission distribution and expected over-the-air (OTA) capabilities also in the propulsion system perimeter, even corrective actions are within reach, if needed. Up till now only service appointments allow an update of the PCM calibration, either done after a dedicated recall or – most of the times – in the background during a standard service appointment. However, both options are not ideal from the time and cost point of view. OTA allows a convenient approach for control software and calibration update, and it may avoid recall actions which negatively impact customer satisfaction. The above mentioned data gathering techniques give the OEM the data and the tools to provide and distribute new propulsion system calibrations, on-demand and nearly in real-time, depending only on connectivity and computing power, for example to compensate vehicle aging effects, tightened regulations, or unforeseen driving situations which might have negative impact on fleet emissions. In case a given emission species is beyond its expected value, different mechanism may be implemented in order to modify the emission level. Note this may not always be possible: in case a given subsystem performance has been significantly deteriorated (e.g. catalyst ageing affecting one or several of their properties, as in Guardiola et al. (2019b)) the only possibility may be substitution of the part to regain the overall required system performance back. In other cases, it is possible to restore the system performance by correcting the expected sensor or actuator drift; in this sense the existence of the high-fidelity sensors is key, since they allow to obtain statistical models for the expected variation of the series sensor along the vehicle life.

An important feature of the proposed framework is the validation of the new calibrations over a subset of the fleet. Such subset may be used for verifying that the effect of the proposed calibration operates as expected, before all vehicles are updated. As it was presented in Guardiola et al. (2016), it is possible to balance between different emissions by deploying a number of individual NOx-optimal and CO2-optimal calibrations and balancing between these calibrations. Figure 5 shows an example of the effect of closed-loop control of the emission level for different feed gas NOx objectives, along with the associated fuel (i.e. CO2) benefit or penalty.

While in the case of the Figure a closed loop was implemented, it is also possible to deploy the calibration in a subset of fleet to monitor the mid term effects through the connected monitoring service. The system is then able to assess the effect of the individual calibrations from a statistical point of view and the best calibration may be selected for a fleet wide deployment. Additionally, it is possible to implement AI assisted optimization methods where the system is able to incrementally vary the calibration; since the number of individual vehicles in the fleet is high, federated learning allows accelerating the optimization phase when compared with traditional methods (i.e. testing the new calibrations on 1000 vehicles against using a few testing vehicles). It must be noticed, however, that a certain level of supervision is needed, since drastic modifications of the calibration could compromise the operation of the vehicle.

5. DIAGNOSTICS ON DEMAND

With the proposed framework, it is possible to deploy diagnostic methods on demand which may be updated over the air. While it can be done to the complete fleet, the deployment may be restricted to vehicles fulfilling a given condition, as per example those which behave as outliers of the in-service emission distribution, or those which incorporate a given part batch whose performance could be compromised by ageing. The latter may serve for
verifying part performance and save recalls of the complete series.

For example, in Guardiola et al. (2020) a method for detecting the ageing of diesel oxidation catalyst was presented. The method relies on using post-injection pulses and to evaluate the value of the lambda sensor downstream of the catalyst package (or the NOx/lambda integrated sensor). In Figure 6 the rational of the method is shown: post-injection fuel pulses need a higher temperature in the catalyst for being oxidized when the catalyst is aged. While scheduling such method for new parts unnecessarily complicates the diagnostic software, the method is able to provide a good estimate of the catalyst performance, and can serve as a good tool for verifying the catalyst performance on vehicles with a certain mileage.

Another important aspect is the possibility of developing new diagnostic algorithms while the vehicle is in its service life. In this sense, several mechanisms are possible. For example, the OEM control software engineer can introduce small size faults in the calibration (e.g. slightly modifying a sensor calibration) in order to create models of the effect of such fault over the engine operation. Such models are later used for developing fault specific methods which may be used for diagnosing other vehicles. In this sense, vehicles fitted with an extended sensor set may serve for gaining insight into the fault effect. Another possibility is the use of past records of a vehicle affected by a fault which has been later identified in the garage. This is a feature difficult to implement in classical OBD systems because of memory scarcity, and many times complicates replicating problems which appear during normal driving. For the case of connected systems, the information from vehicles affected by faults of unknown origin may be registered for later analysis and exploitation. Compared with classical OBD approach, connected diagnostics allows continuously evolving the vehicle software during the service life, and provides a framework for the development of new methods. When AI is added, the cumbersome task of generating and calibrating new models may be simplified and partly automated, while systematic real-life data records of the fleet allows the determination of statistical thresholds for failure detection.

6. CONCLUSION

The combination of connectivity, accurate onboard sensors, and a smart balance of embedded and cloud components will allow a significant evolution of the diagnostic software of the propulsive system. On one hand, this will add the possibility of systematic data collection during vehicle life, thus allowing a continuous update of the diagnostic software. New diagnostic models and methods will be developed as real-life data is available, and it will be possible to deploy methods on demand for checking specific vehicles gathering certain conditions, and to install corrective actions involving software updates. On the other hand, it will be possible to monitor the in-service emission distribution for the fleet. Beyond the detection of outliers, emission data feedback will allow implementing strategies for the calibration optimization in real-life operation.

REFERENCES

Amara, R. (1984). New directions for futures research—setting the stage. Futures, 16(4), 401–404.

Angelov, P., Giglio, V., Guardiola, C., Lughofer, E., and Luján, J.M. (2006). An approach to model-based fault detection in industrial measurement systems with application to engine test benches. Measurement Science and Technology, 17(7), 1809.

Guardiola, C., Pla, B., Bares, P., and Mora, J. (2020). An on-board method to estimate the light-off temperature of diesel oxidation catalysts. International Journal of Engine Research, 21(8), 1480–1492.

Guardiola, C., Pla, B., Bares, P., and Peyton Jones, J. (2019a). Integration of intermittent measurement from in-cylinder pressure resonance in a multi-sensor mass flow estimator. Mechanical Systems and Signal Processing, 131, 152–165.

Guardiola, C., Pla, B., Bares, P., and Waschl, H. (2016). Adaptive calibration for reduced fuel consumption and emissions. Proc. of the Inst. of Mech. Eng. Part D, Journal of Automobile Engineering, 230(14), 2002–2014.

Guardiola, C., Pla, B., Mora, J., and Lefebvre., D. (2019b). Experimental determination and modelling of the diesel oxidation catalysts ageing effects. Proc. of the Inst. of Mech. Eng. Part D, Journal of Automobile Engineering, 233, 3016–30294.

Guardiola, C., Pla, B., Bares, P., and Stefanopoulou, A. (2019c). Cylinder charge composition observation based on in-cylinder pressure measurement. Measurement, 131, 559–568.

Hommen, G., Kupper, F., and Seykens, X. (2017). Robust, model-based urea dosing control for SCR aftertreatment systems using a cross-sensitive tailpipe nox sensor. SAE Paper 2017-01-0938.

Mock, P., German, J., Bandivadekar, A., and Riemersma, I. (2012). Discrepancies between type-approval and real-world fuel consumption and CO₂ values in 2001-2011 European passenger cars. Technical report, The International Council on Clean Transportation, London.

Mora, J., Willems, F., Seykens, X., and Guardiola, C. (2018). An OBD strategy to estimate SCR ageing and detect urea injection faults. IFAC-PapersOnLine, 51(31), 369–376. 5th IFAC Conference on Engine and Powertrain Control, Simulation and Modeling E-COSM 2018.

Payri, F., Luján, J., Guardiola, C., and Pla, B. (2014). A challenging future for the ic engine: New technologies and the control role. Oil & Gas Science and Technology, 70, 15–30.