Preprints of the 21st IFAC World Congress (Virtual)
Berlin, Germany, July 12-17, 2020

Design of a neural virtual sensor for the air and charging system in a Diesel engine

Vincenzo Alfieri *, **Carmen Pedicini** *, and **Corrado Possieri** **

* *PUNCH Torino Spa (former GM Global Propulsion Systems - Torino S.r.l.), 10129 Turin, Italy (e-mails: [vincenzo.alfieri, carmen.pedicini]@punchtorino.com).*
** *Istituto di Analisi dei Sistemi ed Informatica "A. Ruberti", Consiglio Nazionale delle Ricerche (IASI-CNR), 00185 Roma, Italy (e-mail: corrado.possieri@iasi.cnr.it).*

Abstract: The main objective of this work is to design a virtual sensor capable of estimating variables that are unmeasurable on-line in the air and charging subsystem of a Diesel engine. In order to achieve this objective, a data-driven approach is pursued. In particular, we show that combining high-gain observers and feed-forward neural networks, it is possible to design an observer for the air and charging system of a Diesel engine on the basis of data acquired via a test bench. The performance of this observer is evaluated in a real experimental setting.

Keywords: Observers, Diesel engines, data-drive approaches.

1. INTRODUCTION

Diesel engines are facing an increased hardware complexity with the aims of reducing their emissions and of improving fuel economy (Verschaeren and Verhelst, 2018). Since this complexity alone is not capable of improving efficiency enough, this led to the necessity of developing sophisticated control systems capable of dealing with such plants in an optimized way. Among the various subsystems that compose a Diesel engine, the air and charging system is among the ones that received more interest in the literature due to the fact that classical control structures, such as map-based PID, do not properly address the control problem (Wahlstrom et al., 2010). Several control architectures have been proposed for this subsystem, such as predictive control (Ortner and Del Re, 2007), control Lyapunov functions (Jankovic and Kolmanovsky, 2000), sensor-based approaches (Amstutz and Del Re, 1995), and feedback theory (Park et al., 2014; Alfieri et al., 2018). A common aspect to some of this control structure is that they require the current value of the state of the plant in order to determine the current control action. However, during normal functioning, Diesel engines are usually not fully sensorized, *i.e.*, some states are actually not directly measurable. This leads to the necessity of developing virtual sensors capable of determining unmeasurable state variables from available ones. An effort in this direction has been made in Harder et al. (2018) by using a square-root unscented Kalman filter. The main objective of this work is to develop a virtual sensor to achieve this goal without relying in a model for the Diesel engine. In order to pursue this objective, a data-driven observer has been designed by relying on data acquired via an engine test bench and coupling a high-gain "practical" observer with an artificial neural network (see Section 6 for further details). In order to evaluate the performance of this observer, it has been tested on real experimental data (which have not been used to train the network) acquired via a test bench.

2. PLANT DESCRIPTION

The system considered in this work is a four cylinders Diesel engine with single stage variable geometry turbocharger (VGT), equipped with a high-pressure exhausted gas rejection (EGR) loop as well as a low-pressure EGR loop. A representation of the system is in Fig. 1.

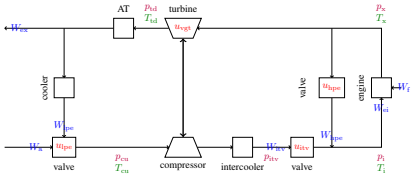

Fig. 1. Graphical representation of the Diesel engine.

The fresh air is mixed through a three-way valve with exhaust gases and is pressurized by the compressor, which is powered by the turbine through a common shaft. Afterwards, the mixture of fresh air and exhaust gases is cooled by the intercooler. In order to control the intake mass flow, which is a mixture of the flow coming from the compressor and the flow coming from a second high-pressure EGR (exhaust gas recirculation) valve, a controlled throttle valve is present. From the intake manifold the mixture of gases enters into the engine together with the fuel and takes part in the combustion process. The product of the combustion is then expelled into the exhaust manifold, where it splits in two parts: one that is recirculated through the high-pressure EGR valve and one that goes through the turbine. The latter flow provides the power to drive the compressor passing through the VGT. Finally, the gas mixture passes through the after-treatment and splits again in two: one part goes out of the engines and the other one is firstly

cooled and then goes through the three-way valve to be mixed with fresh air.

3. NOMENCLATURE

Let $\mathbb{R}$ and $\mathbb{N}$ denote the sets of real an natural numbers, respectively. The symbol $\|\cdot\|_p$ denotes the p-norm of the vector at argument. The symbol $\oplus$ denotes the direct sum, *i.e.*, given $A \in \mathbb{R}^{n_1 \times n_2}$, $B \in \mathbb{R}^{n_3 \times n_4}$,

$$A \oplus B = \begin{bmatrix} A & 0_{n_1,n_4} \\ 0_{n_3,n_2} & B \end{bmatrix},$$

where $0_{a,b}$ denotes the $a \times b$-dimensional matrix whose entries are all 0. The symbol I denotes the identity matrix.

In order to improve readability of this work, we first introduce a legend of the symbols used through the discussion:

$\beta_{\rm c}$: compressor pressure ratio (−),
$c_{\rm p}$: specific heat capacity at constant pressure (J/K)
η_ν: engine volumetric efficiency multiplier (−),
γ: ratio of specific heats (1.3946),
$N_{\rm e}$: engine speed (rpm)
$p_{\rm cu}$: upstream compressor pressure (Pa),
$p_{\rm ex}$: downstream after-treatment pressure (Pa),
$p_{\rm i}$: intake manifold pressure (Pa),
$p_{\rm itv}$: upstream ITV pressure (Pa),
$p_{\rm td}$: downstream turbine pressure (Pa),
$p_{\rm x}$: exhaust manifold (upstream turbine) pressure (Pa),
R_a: specific gas constant ($287.058\,\mathrm{J\,Kg\,K^{-1}}$),
$T_{\rm cu}$: upstream compressor temperature (K),
$T_{\rm hpe}$: high-pressure exhausted temperature (K),
$T_{\rm i}$: intake manifold temperature (K),
$T_{\rm lpe}$: low-pressure exhausted temperature (K),
$T_{\rm td}$: downstream turbine temperature (K),
$T_{\rm x}$: exhaust manifold temperature (K),
$u_{\rm itv}$: ITV (intake throttle valve) position (%),
$u_{\rm lpe}$: LPE (low-pressure exhausted) valve position (%),
$u_{\rm hpe}$: HPE (high-pressure exhausted) valve position (%),
$u_{\rm vgt}$: position (%) of the moveable vanes in the VGT (variable geometry turbocharger),
$W_{\rm a}$: clean air mass flow rate (Kg/s),
$W_{\rm ei}$: engine-in mass flow rate (Kg/s),
$W_{\rm ex}$: mass flow rate that goes out the engine (Kg/s).
$W_{\rm f}$: fuel mass flow rate (Kg/s),
$W_{\rm hpe}$: HPE valve mass flow rate (Kg/s),
$W_{\rm itv}$: ITV mass flow rate (Kg/s),
$W_{\rm lpe}$: LPE valve mass flow rate (Kg/s),
$V_{\rm cu}$: upstream compressor volume (m^3),
$V_{\rm d}$: engine displacement volume (m^3),
$V_{\rm i}$: intake manifold volume (m^3),
$V_{\rm td}$: downstream turbine volume (m^3),
$V_{\rm x}$: exhaust manifold volume (m^3).

4. PHYSICAL MODEL OF THE AIR & CHARGING SYSTEM

In this section, by using the same modeling strategy employed in Alfieri et al. (2015, 2018), we derive a simplified model to describe the dynamical behavior of the air and charging system in the Diesel engine described in Section 2. This model is used in the subsequent Section 6 to identify the input, the state, and the output of the plant.

By using the standard model (Guzzella and Onder, 2009), which is based on the ideal gas law, the principle of mass conservation, and neglecting the contribution given by temperature variations, the dynamics of $p_{\rm cu}$, $p_{\rm i}$, $p_{\rm x}$, and $p_{\rm td}$ are given by

$$\dot{p}_{\rm cu} = \frac{R_a\, T_{\rm cu}}{V_{\rm cu}} \left(W_{\rm a} + W_{\rm lpe} - W_{\rm itv}\right), \tag{1a}$$

$$\dot{p}_{\rm i} = \frac{R_a\, T_{\rm i}}{V_{\rm i}} \left(W_{\rm itv} + W_{\rm hpe} - W_{\rm ei}\right), \tag{1b}$$

$$\dot{p}_{\rm x} = \frac{R_a\, T_{\rm x}}{V_{\rm x}} \left(W_{\rm ei} - W_{\rm hpe} - W_{\rm itv}\right), \tag{1c}$$

$$\dot{p}_{\rm td} = \frac{R_a\, T_{\rm td}}{V_{\rm td}} \left(W_{\rm itv} - W_{\rm lpe} - W_{\rm a}\right). \tag{1d}$$

The flows through the EGR valves, the throttle valve, and the turbine can be modeled by using the orifice equation for compressible gases (Guzzella and Onder, 2009),

$$W_{\rm itv} = \mathrm{CdA}_{\rm itv}\left(u_{\rm itv}\right) \Xi\left(\frac{p_{\rm i}}{\beta_{\rm c}\, p_{\rm cu}}\right) \beta_{\rm c}\, p_{\rm cu} \sqrt{\frac{1}{R_a\, T_{\rm itv}}},$$

$$W_{\rm hpe} = \mathrm{CdA}_{\rm hpe}\left(u_{\rm hpe}, \frac{p_{\rm i}}{p_{\rm x}}\right) \Xi\left(\frac{p_{\rm i}}{p_{\rm x}}\right) p_{\rm x} \sqrt{\frac{1}{R_a\, T_{\rm hpe}}},$$

$$W_{\rm lpe} = \mathrm{CdA}_{\rm lpe}\left(u_{\rm lpe}, p_{\rm cu} - p_{\rm ex}\right) \Xi\left(\frac{p_{\rm cu}}{p_{\rm ex}}\right) p_{\rm ex} \sqrt{\frac{1}{R_a\, T_{\rm lpe}}},$$

where the compressor pressure ratio is defined as $\beta_{\rm c} = \frac{p_{\rm itv}}{p_{\rm cu}}$, the flow effective areas $\mathrm{CdA}_{\rm itv}(\cdot)$, $\mathrm{CdA}_{\rm lpe}(\cdot,\cdot)$, and $\mathrm{CdA}_{\rm hpe}(\cdot,\cdot)$ can be identified from experimental data as functions of the corresponding arguments, and $\Xi(\cdot)$ denotes the pressure correction factor, that is given by

$$\Xi(x) = \sqrt{\frac{2\gamma}{\gamma - 1}\left(\Pi^{\frac{2}{\gamma}}(x) - \Pi^{\frac{\gamma+1}{\gamma}}(x)\right)},$$

$$\Pi(x) = \begin{cases} \left(\frac{2}{\gamma+1}\right)^{\frac{\gamma}{\gamma-1}}, \text{ if } x < \left(\frac{2}{\gamma+1}\right)^{\frac{\gamma}{\gamma-1}} & \text{(chocked)}, \\ x, \quad \text{if } \left(\frac{2}{\gamma+1}\right)^{\frac{\gamma}{\gamma-1}} \leqslant x \leqslant 1 & \text{(subsonic)}, \\ 1 \quad \text{if } x > 1 & \text{(no backflow)}. \end{cases}$$

Since the function $x \mapsto \Xi(x)$ is continuous but not differentiable, in the remainder of this work such a function is substituted by its kth moving average, that is

$$\tilde{\Xi}(x) = \int_{x-s}^{x+s} \int_{\xi_{k-1}-s}^{\xi_{k-1}+s} \cdots \int_{\xi_2-s}^{\xi_2+s} \Pi(\xi_1)\, \mathrm{d}\xi_1 \cdots \mathrm{d}\xi_k,$$

where s is a small positive real number. Note that, since $x \mapsto \Xi(x)$ is continuous, the function $x \mapsto \tilde{\Xi}(x)$ is C^k.

On the other hand, the engine-in mass flow rate is modeled by using the speed-density equation $W_{\rm ei} = \frac{\eta_\nu\, V_{\rm d}\, N_{\rm e}}{120\, R_a\, T_{\rm i}}\, p_{\rm i}$.

Finally, we need to account for the dynamics of the compressor pressure ratio, which can be modeled as (Watson and Janota, 1982)

$$\dot{\beta}_{\rm c} = c\,(P_{\rm t} - P_{\rm c}), \tag{2}$$

where c is a constant coefficient, $P_{\rm t}$ and $P_{\rm c}$ are the turbine and compressor powers, respectively, which are given by

$$P_t = c_{\rm p}\left(W_{\rm itv} + W_{\rm f}\right) T_{\rm td}\, f_{\rm t}\left(\frac{(W_{\rm itv} + W_{\rm f})\sqrt{T_{\rm x}}}{p_{\rm td}}, u_{\rm vgt}\right),$$

$$P_c = c_{\rm p}\, W_{\rm itv}\, T_{\rm cu}\, f_{\rm c}\left(\frac{W_{\rm itv}\sqrt{\frac{T_{\rm cu}}{287}}}{p_{\rm cu}}, \beta_{\rm c}\right),$$

where $f_{\rm t}(\cdot,\cdot)$ and $f_{\rm c}(\cdot,\cdot)$ denote the turbine and compressor efficiencies, which can be determined using supplier maps.

Hence, letting $x = [\, p_{\rm cu}\ p_{\rm i}\ p_{\rm x}\ p_{\rm td}\ \beta_{\rm c}\,]^\top$ denote the state vector, letting $u = [\, u_{\rm itv}\ u_{\rm lpe}\ u_{\rm hpe}\ u_{\rm vgt}\,]^\top$ denote the vector of control inputs, letting

$$w = [\, T_{\rm i}\ T_{\rm cu}\ T_{\rm x}\ T_{\rm td}\ T_{\rm itv}\ T_{\rm lpe}\ W_{\rm a}\ W_{\rm f}\ N_{\rm e}\,]^\top$$

denote the vector of exogenous inputs, letting $v = [\, u^\top\ w^\top\,]^\top$ denote the vector of the overall inputs of the plant (assumed to be measurable or modeled), and letting $y = p_{\rm i}$ denote the measurable output, combining (1) and (2), the dynamics of the plant depicted in Fig. 1 can be modeled as

$$\dot{x} = F(x, v), \qquad y = H(x). \tag{3}$$

The main objective of this work is to design a virtual sensor for the plant, *i.e.*, a tool that, given measurements of the input and output vectors is able to reconstruct the current state of the system. This objective is pursued by designing a data-driven observer for system (3).

5. REVIEW OF NONLINEAR OBSERVABILITY CONCEPTS

Consider a system in the form

$$\dot{x} = f(x, v), \qquad y = h(x, v), \tag{4}$$

where $x \in \mathbb{R}^n$ is the state vector, $u \in \mathbb{R}^\ell$ is the input vector, $y \in \mathbb{R}^m$ is the output vector, $f(\cdot)$ and $h(\cdot)$ are C^k functions for some sufficiently large $k \in \mathbb{N}$. Assume that for each initial condition $x_0 \in \mathbb{R}^n$ and any measurable, bounded input vector $v(\cdot)$, the solution $x(t) = \phi(t, x_0, v(\cdot))$ of system (4) is well-defined for all times $t \geqslant 0$. System (4) is *observable* if, for any input $v(\cdot)$, there do not exist two different initial conditions $x_{0,1}$ and $x_{0,2}$ such that $h(\phi(t, x_{0,1}, v(\cdot))) = h(\phi(t, x_{0,2}, v(\cdot))$ for all times $t \geqslant 0$ (Inouye, 1977). This condition entails with the indistinguishability of two different states. Namely, another interpretation of the above definition is that, given its dynamics, system (4) is observable if and only if its initial condition can be reconstructed by using only measurements of the input $v(\cdot)$ and of the output $y(\cdot)$.

Let $v^{(i)}(t) = \frac{\mathrm{d}^i}{\mathrm{d}t^i} v(t)$ denote the ith time derivative of the input function and let $v_{e,k}(t) = [\, (v^{(0)}(t))^\top\ \cdots\ (v^{(k)}(t))^\top\,]^\top$ denote the vector of the time derivatives of the input. Hence, define the *observability map of order* k of (4),

$$O_k(x, v_{e,k}) = \begin{bmatrix} L_f^0 h(x, v_{e,k}) \\ \vdots \\ L_f^k h(x, v_{e,k}) \end{bmatrix},$$

where the functions $L_f^j h(x, v_{e,k})$, $j = 0, \ldots, k$, are given by $L_f^0 h(x, v_{e,k}) = h(x)$, $L_f^{j+1}(x, v_{e,k}) = \frac{\partial L_f^j(x, v_{e,k})}{\partial x} f(x, v^{(0)}) + \sum_{i=0}^{k-1} \frac{\partial L_f^j(x, v_{e,k})}{\partial v^{(i)}} v^{(i+1)}$. The key role of the observability map of order k is to link the current state $x(t)$ of system (4) and the current time derivatives of the input $v_{e,k}(t)$ up to order k with the vector of the time derivatives of the output up to order k,

$$y_{e,k}(t) = [\, (y^{(0)}(t))^\top\ (y^{(1)}(t))^\top\ \cdots\ (y^{(k)}(t))^\top\,]^\top,$$

where $y^{(i)}(t) = \frac{\mathrm{d}^i}{\mathrm{d}t^i} y(t)$ denote the ith time derivative of y. In fact, the following relation holds for all $t \geqslant 0$,

$$y_{e,k}(t) = O_k(x(t), v_{e,k}(t)).$$

System (4) is *strongly k-differentially observable* if the suspension $(x, v_{e,k}) \mapsto O_k(x, v_{e,k}) \times v_{e,k}$ is an injective immersion, *i.e.*, it is a function that preserves distinctness (it never maps distinct elements of its domain to the same element of its codomain), it is differentiable, and its derivative is everywhere injective (Bishop and Crittenden, 2011). By Aeyels (1981); Gauthier and Kupka (2001), strong k-differential observability is a *generic property*, *i.e.*, almost all systems that can be written in the form (4) are strongly k-differentially observable for some sufficiently large $k \in \mathbb{N}$ (generically, for $k = 2\,n + 1$ Sontag, 2002). Furthermore, the motivation to consider differential observability rather than plain observability is that it can be readily used to design observers for system (4). In fact, if system (4) is k-differentially observable, then there is $Q_k(\cdot, \cdot)$ such that

$$x = Q_k(O_k(x, v_{e,k}), v_{e,k}), \tag{5}$$

for all x and $v_{e,k}$. Therefore, in such a case, an observer for system (4) can be designed by coupling the function $Q_k(\cdot, \cdot)$ with any tool that is able to estimate the time derivatives of the output, such as high-gain observers (Tornambe, 1992), sliding mode differentiators (Shtessel et al., 2014), and super-twisting algorithms (Moreno and Osorio, 2012). This is exactly the rationale used in Section 6 to design a data-driven observer for the diesel engine.

6. DATA-DRIVEN OBSERVER DESIGN

In this section we design a data-driven observer for the plant depicted in Fig. 1 without requiring the knowledge of the functions $\mathrm{CdA}_{\rm itv}(\cdot)$, $\mathrm{CdA}_{\rm lpe}(\cdot, \cdot)$, $\mathrm{CdA}_{\rm hpe}(\cdot, \cdot)$, $f_T(\cdot, \cdot)$, and $f_c(\cdot, \cdot)$ appearing in (3). These dynamics are used to identify the inputs, the state, and the outputs of the plant.

The rationale behind the proposed observer is the following: estimate the vector of the time derivatives of the input $v_{e,k}$ and of the output $y_{e,k}$ (that is the set of variables that can be measured on-line during normal functioning) and use a function $Q_k(\cdot, \cdot)$ that satisfies (5) to obtain an estimate of the state x_k of the plant (that cannot be measured during normal functioning).

Two problems have to be faced in order to design the observer described above:

- (P1) a system that computes exactly the time derivatives of a signal is not causal (Kailath, 1980) and hence there is the need of a tool capable of estimating them;
- (P2) the function $Q_k(\cdot, \cdot)$ is difficult to compute even in the case that a model of the plant is available (Gauthier and Kupka, 2001).

However, some tools in the literature can be used to address the problems reported above. Indeed, Problem (P1) can be faced using some tools that are able to "practically" estimate (*i.e.*, with arbitrarily small estimation error) the time derivative of a signal, such as high-gain observers (Carnevale et al., 2018; Astolfi et al., 2018), sliding mode (Shtessel et al., 2014), or super-twisting algorithms (Moreno and Osorio, 2012).

On the other hand, Problem (P2) can be addressed by using algebraic geometry tools (Menini et al., 2016), sliding mode approaches (Menini et al., 2018), or machine learning methods (Friedman et al., 2001).

Since the virtual-sensor has to be integrated in an engine control module (ECM), we designed the observer by using the least (on-line) computationally demanding tools among the ones that are recalled above. Namely, the time derivatives of the vector v and of the signal y have been estimated by using a discretization of the high-gain observer given in Tornambe (1992), while the function $Q_k(\cdot,\cdot)$ has been estimated from real data using feed-forward neural networks (FFNN) (Friedman et al., 2001). Fig. 2 depicts a representation of the resulting data-driven observer.

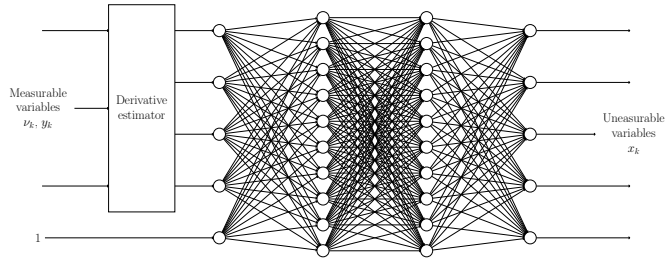

Fig. 2. Graphical representation of the data-driven observer.

Note that, if the derivative estimator estimates the time derivatives of the measurable variables up to order $k = 2\,n+1$ and the FFNN is suitably trained to approximate the function $Q_{2\,n+1}(\cdot,\cdot)$ that satisfies (5), then the observer depicted in Fig. 2 is a generic observer for plants that can be written in the form (4) (see Aeyels, 1981; Sontag, 2002). Therefore, it appears particularly suitable to achieve the goal of this work.

In order to detail how the observer depicted in Fig. 2 has been designed, in Section 6.1, we review the high-gain practical observer given in Tornambe (1992). The problem of determining a discrete-time system that approximate such an observer is dealt with in Section 6.2. On the other hand, in Section 6.3, we show how to design the FFNN that approximates the function $Q_k(\cdot,\cdot)$. These tools are coupled in Section 6.4 to design a data-drive observer for the plant depicted in Fig. 1. Finally, Section 6.5 reports the results of some experiments in which such an observer has been used to estimate the state of the plant.

6.1 "Practical" high-gain observer

In this section, we review the high-gain observer proposed in Tornambe (1992), that, under mild assumptions, is able to estimate the time derivatives of a signal with arbitrarily small estimation error, with arbitrarily fast transient rate, and without any knowledge of the signal to be estimated.

Let $k \in \mathbb{N}$ be given and consider the following system

$$\dot{\eta} = A_k\,\eta + B_k\,\nu, \tag{6a}$$

where $\eta \in \mathbb{R}^{k\,z}$ is the state, $\nu = \left[\,\nu_1\ \cdots\ \nu_z\,\right]^\top \in \mathbb{R}^z$ is the signal whose time derivatives have to be estimated,

$$A_k = \bigoplus_{i=1}^{z} \begin{bmatrix} -\frac{\mu_{i,1}}{\varepsilon} & 1 & 0 & \cdots & 0 \\ -\frac{\mu_{i,2}}{\varepsilon^2} & 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ -\frac{\mu_{i,k}}{\varepsilon^k} & 0 & 0 & \cdots & 1 \\ -\frac{\mu_{i,k+1}}{\varepsilon^{k+1}} & 0 & 0 & \cdots & 0 \end{bmatrix}, \quad B_k = \bigoplus_{i=1}^{z} \begin{bmatrix} \frac{\mu_{i,1}}{\varepsilon} \\ \frac{\mu_{i,2}}{\varepsilon^2} \\ \vdots \\ \frac{\mu_{i,k}}{\varepsilon^k} \\ \frac{\mu_{i,k+1}}{\varepsilon^{k+1}} \end{bmatrix}, \tag{6b}$$

the coefficients $\mu_{i,k}$ are such that the polynomial $\varrho^{k+1} + \mu_{i,1}\,\varrho^k + \cdots + \mu_{i,k}\,\varrho + \mu_{i,k+1}$ is Hurwitz, $i = 1,\ldots,z$, and ε is a sufficiently small, positive real parameters. By Tornambe (1992); Carnevale et al. (2018), system (6) is a "practical" high-gain observer for the time derivatives of the input $\nu \in \mathbb{R}^z$, *i.e.*, if there exists $M > 0$ such that $\|\frac{\mathrm{d}^{k+1}}{\mathrm{d}t^{k+1}}\nu(t)\|_\infty \leqslant M$ for all $t \geqslant 0$, then the norm of the estimation error $\eta(t) - \left[\,\nu_1(t)\ \cdots\ \frac{\mathrm{d}^k}{\mathrm{d}t^k}\nu_1(t)\ \cdots\ \frac{\mathrm{d}^k}{\mathrm{d}t^k}\nu_z(t)\,\right]^\top$ can be made arbitrarily small in an arbitrarily small amount of time by letting the design parameter ε be sufficiently small. Thus, system (6) can be used in the observation scheme depicted in Fig. 2 as derivative estimator. In order to mitigate the effects of noise on the estimates of the time derivatives of ν and y, such signals have been filtered with a low-pass filter before being fed into system (6).

6.2 Discretization of the high-gain observer

The high-gain observer (6) is designed assuming that the input ν (that is the signal whose time derivatives up to order k have to be estimated) is available at continuous-time. However, in the considered application, just discrete-time samples of such a signal are available. Therefore, we discretized such a system by using the blilinear method (Franklin et al., 1998).

Let $k \in \mathbb{N}$ be given and consider the matrices A_k and B_k defined in (6b). Hence, define

$$E_k = \left(I + \frac{\tau}{2}\,A_k\right)\left(I - \frac{\tau}{2}\,A_k\right)^{-1}, \tag{7a}$$

$$G_k = \sqrt{\tau}\,\left(I - \frac{\tau}{2}\,A_k\right)^{-1} B_k, \tag{7b}$$

$$V_k = \sqrt{\tau}\left(I - \frac{\tau}{2}\,A_k\right)^{-1}, \tag{7c}$$

$$D_k = \frac{\tau}{2}\left(I - \frac{\tau}{2}\,A_k\right)^{-1} B_k. \tag{7d}$$

where τ is the sampling time. Thus, letting $\nu_D(\kappa) = v(\kappa\,\tau)$, $\kappa \in \mathbb{N}$, the discrete-time system

$$\chi(\kappa+1) = E_k\,\chi(\kappa) + G_k\,\nu_D(\kappa), \tag{8a}$$

$$\eta_D(\kappa) = V_k\,\chi(\kappa) + D_k\,\nu_D(\kappa), \tag{8b}$$

constitutes a discretization of the high-gain observer (6).

6.3 Approximation of the function $Q_k(\cdot,\cdot)$

In Sections 6.1 and 6.2, we proposed a tool to estimate the time derivative of unknown signals. In this section, such a tool is used to find an approximation of the inverse of the suspension of the observability map by relying on data. We performed several experiments on the engine in different working conditions (including transient behaviors – load steps at different constant engine speed, tip-in and tip-out in partial and full load – so to properly analyze the dynamics of the plant) and, in each experiment, we collected the dataset

$${}^{\jmath}\mathcal{H} = \left\{{}^{\jmath}v(\kappa), {}^{\jmath}y(\kappa), {}^{\jmath}x(\kappa)\right\}_{\kappa=0}^{{}^{\jmath}K},$$

where ${}^{\jmath}v(\kappa)$, ${}^{\jmath}y$, and ${}^{\jmath}x$ are samples of the input, output, and state vectors, respectively, ${}^{\jmath}K\,\tau$ is the length of the experiment (in seconds), τ is the sampling time (the same in all the experiments), and $\jmath$ is the experiment number.

In order to use the reasoning outlined at the beginning of this section, in each experiment $\jmath$, we computed the vectors ${}^{\jmath}v_{e,11}(\kappa)$ and ${}^{\jmath}v_{e,11}(\kappa)$ of the time derivatives up to order 11 of the inputs v and of the output y of the plant

via the discretized "practical" high-gain observer given in Section 6.2. Hence, the extended datasets

$$ {}^{J}_{e}\mathcal{H} = \{({}^{J}v_{e,11}(\kappa), {}^{J}y_{e,11}(\kappa), {}^{J}x(\kappa)\}_{\kappa=0}^{{}^{J}K}, $$

have been used to train a FFNN. Indeed, in view of the discussion reported in Section 5, if $n = 5$, then there generically exists a function $W_{11}(\cdot,\cdot)$ such that

$$ {}^{J}x(\kappa) = W_{11}({}^{J}v_{e,11}(\kappa), {}^{J}y_{e,11}(\kappa)). $$

Moreover, since $W_{11}(\cdot,\cdot)$ is differentiable, by the universal approximation theorem (Cybenko, 1989), there exists a FFNN that approximates it with arbitrarily small approximation error over compact sets. Hence, we can attempt at determining an approximation of such a function by training a FFNN using the extended datasets ${}^{J}_{e}\mathcal{H}$.

In order to determine the simplest structure of the FFNN that provides the smallest estimation error, we trained several neural networks using different structures and evaluating their MSE performance via 10-fold cross validation repeated 100 times. Fig. 3 depicts the results of such tests.

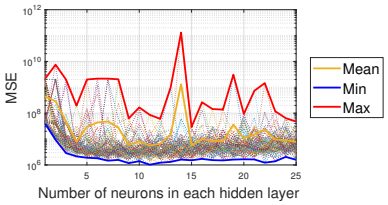

Fig. 3. Results of the cross validation.

By considering the results depicted in such a figure, a FFNN with *two hidden layers* and *eleven neurons in each layer* is capable of approximating with sufficiently small approximation error the function $W_{11}(\cdot,\cdot)$. Thus, such an approximation can be readily used to design an observer for the plant, as detailed in the following section.

6.4 Data-driven observer

In this section, the discretized high-gain observer given in Section 6.2 and the FFNN trained by using the method described in Section 6.3 are coupled to design an observer, thus obtaining the observer depicted in Fig. 2.

Let $k = 11$, let E_{11}, G_{11}, V_{11}, and D_{11} be defined as in (7), and let $\Theta(v_{e,11}, y_{e,11})$ be the input-output function synthesized by the FFNN trained by using the method given in Section 6.3. Thus, a data-driven observer is

$$ \chi(\kappa+1) = E_{11}\,\chi(\kappa) + G_{11}\,\nu_D(\kappa), \tag{9a} $$
$$ \eta_D(\kappa) = V_{11}\,\chi(\kappa) + D_{11}\,\nu_D(\kappa), \tag{9b} $$
$$ \hat{x}(\kappa) = \Theta(\Pi_1\,\eta_D(\kappa), \Pi_2\,\eta_D(\kappa)), \tag{9c} $$

where $\hat{x}(\kappa)$ is an estimate of $x(\kappa)$, Π_1 and Π_2 are projection matrices such that

$$ v_{e,11} = \Pi_1\,\chi, \qquad y_{e,11} = \Pi_2\,\chi. $$

Note that the observer (9) can be easily implemented in an ECM since it is a discrete-time system with linear, time-invariant dynamics and nonlinear output function and hence it does not require much computational power. In fact, in order to execute one step of (9), one has to perform $57.6 \cdot 10^3$ elementary operations (*i.e.*, sums and products) to compute $\chi(\kappa)$ from $\chi(\kappa-1)$ and $57.6 \cdot 10^3$ elementary operations to compute the vectors $\Pi_1\,\eta_D(\kappa)$ and $\Pi_2\,\eta_D(\kappa)$ that are the inputs of the function $\Theta(\cdot,\cdot)$. On the other hand, the complexity of executing (9c) depends on the structure of the FFNN (activation function, number of hidden layers, *etc.*) that is employed and can be computed by means of the tools given in Orponen (1994). For instance, by using $x \mapsto \tanh(x)$ as activation function, two hidden layers and eleven neurons in each hidden layer (that is the structure of the FFNN used in the following Section 6.5), requires $2.992 \cdot 10^3$ plus $O(M(\aleph)\,\aleph^{\frac{1}{2}})$ elementary operations, where $\aleph$ denotes the number of digits of precision at which the function is to be evaluated and $M(\aleph)$ stands for the complexity of the chosen multiplication algorithm.

6.5 Experimental results

The data-driven observer (9) designed as detailed in Sections 6.2 and 6.3 has been used on some experimental data that have not been used to train the FFNN, but that correspond to the same working conditions of the training dataset. Since. in these experiments, measurements of the estimated variables were available, we compared them with the estimates obtained by using the data-driven observer (9) in order to verify its effectiveness.

Fig. 4 depicts the results of one of such experiments, showing the estimation error $\hat{x}(\kappa) - x(\kappa)$.

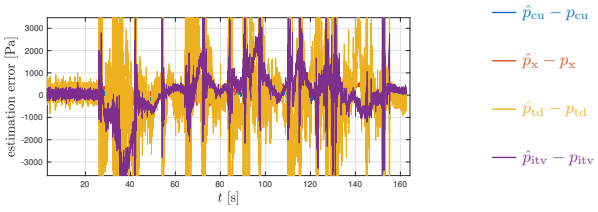

Fig. 4. Results of the experiment with the observer (9).

As shown by such a figure, the data-driven observer (9) is able to "practically" reconstruct the state of the plant, with an average error of 2.6299 KPa. This is essentially due to two main facts: on one hand the plant depicted in Fig. 1 is strongly 11-differentially observable (recall that, by Aeyels (1981); Sontag (2002), this property is generic for plants whose state dimension is $n = 5$), on the other hand, the inverse $Q_{11}(\cdot,\cdot)$ of the suspension of observability map $O_{11}(\cdot,\cdot)$ is a differentiable function of its entries. This implies that $Q_{11}(\cdot,\cdot)$ is also an absolutely continuous function of its entries and hence if the time derivatives of the output y and of the input v are estimated with a sufficiently small estimation error (by decreasing the parameter ε), then the estimate $\hat{x}$ obtained by using (9c) is close to the actual value of x.

To further corroborate the effectiveness of the proposed observation structure, the following Table 1 reports the estimation errors obtained by using the proposed observation scheme in 12 different experimental settings that have not been used for training.

7. CONCLUSIONS

The tests reported in this work showed that the data-driven observer presents good estimation capabilities, being able to estimate the state of the plant both during normal functioning and at rest conditions. The drawbacks

Table 1. Error on total validation data.

	Max Abs.	Max Rel.	Mean Rel.	Mean Sq.	r^2
p_{cu}	2.91	0.0302	0.00106	0.0208	0.997
p_{x}	68.7	0.379	0.0165	19.5	0.995
p_{td}	9.08	0.0851	0.00406	0.497	0.992
p_{itv}	19.2	1.14	0.00444	3.03	0.999

of this approach, however, are that the physical meaning of the involved variables is somehow lost and that its effectiveness is strongly dependent on the training set that is used to train the FFNN. However, if the training set is sufficiently large to take into account all the functioning scenarios, then the considered approach seems to be a good tool to design virtual sensors for Diesel engines.

The proposed data-driven observer presents several advantages with respect to other state-of-the-art tools. Indeed, the high-gain observer (6) requires minimal tuning (essentially just the selection of the parameter ε), whereas the FFNN can be trained by using any algorithm capable of training static neural networks. Therefore, the training phase is much simpler than the one that has to be carried out to train a convolutional neural network (Schmidhuber, 2015), or, more generally, a structure accounting for the previous input/output values, since just a static function has to be approximated. Finally, it is worth stressing that, in the proposed data-driven approach, the FFNN can be substituted by any tool capable of approximating nonlinear functions, such as the set membership tool given in Milanese and Novara (2004). The choice of using a FFNN is due to the availability of a wide number of training algorithms specifically developed for such functions.

Future work will deal with the comparison of the proposed observation scheme with others available in the literature, such as unscented Kalman filter obtained using FFNNs to approximate the unknown functions $\mathrm{CdA}_{\mathrm{itv}}(\cdot)$, $\mathrm{CdA}_{\mathrm{lpe}}(\cdot,\cdot)$, $\mathrm{CdA}_{\mathrm{hpe}}(\cdot,\cdot)$, $f_T(\cdot,\cdot)$, and $f_c(\cdot,\cdot)$.

REFERENCES

Aeyels, D. (1981). Generic observability of differentiable systems. *SIAM J. Control Optim.*, 19(5), 595–603.

Alfieri, V., Conte, G., and Pedicini, C. (2015). Feedback linearization control for the air & charging system in a Diesel engine. In *Eur. Control Conf.*, 2565–2570. IEEE.

Alfieri, V., Conte, G., and Pedicini, C. (2018). Nonlinear model-based multivariable control for air & charging system of Diesel engine with short and long route EGR valves. *Int. J. Autom. Tech.*, 19(3), 405–412.

Amstutz, A. and Del Re, L.R. (1995). EGO sensor based robust output control of EGR in diesel engines. *IEEE Trans. Control Syst. Tech.*, 3(1), 39–48.

Astolfi, D., Marconi, L., Praly, L., and Teel, A.R. (2018). Low-power peaking-free high-gain observers. *Automatica*, 98, 169 – 179.

Bishop, R.L. and Crittenden, R.J. (2011). *Geometry of manifolds*. Academic Press.

Carnevale, D., Possieri, C., and Tornambe, A. (2018). On "'over-sized" high-gain practical observers for nonlinear systems. *IEEE/CAA J. Automatica Sinica*, 5(3), 691–698.

Cybenko, G. (1989). Approximations by superpositions of a sigmoidal function. *Math. Control Sign. Syst.*, 2, 183–192.

Franklin, G.F., Powell, J.D., and Workman, M.L. (1998). *Digital control of dynamic systems*. Addison-wesley.

Friedman, J., Hastie, T., and Tibshirani, R. (2001). *The elements of statistical learning*. Springer.

Gauthier, J.P. and Kupka, I. (2001). *Deterministic observation theory and applications*. Cambridge University Press.

Guzzella, L. and Onder, C. (2009). *Introduction to modeling and control of internal combustion engine systems*. Springer.

Harder, K., Buchholz, M., Späder, T., and Graichen, K. (2018). A real-time nonlinear air path observer for off-highway diesel engines. In *Eur. Control Conf.*, 237–242.

Inouye, Y. (1977). On the observability of autonomous nonlinear systems. *J. Math. Anal. Appl.*, 60(1), 236–247.

Jankovic, M. and Kolmanovsky, I. (2000). Constructive Lyapunov control design for turbocharged diesel engines. *IEEE Trans. Control Syst. Tech.*, 8(2), 288–299.

Kailath, T. (1980). *Linear systems*. Prentice-Hall.

Menini, L., Possieri, C., and Tornambe, A. (2016). Switching signal estimator design for a class of elementary systems. *IEEE Trans. Autom. Control*, 61(5), 1362–1367.

Menini, L., Possieri, C., and Tornambe, A. (2018). A Newton–like algorithm to compute the inverse of a nonlinear map that converges in finite time. *Automatica*, 89, 411–414.

Milanese, M. and Novara, C. (2004). Set membership identification of nonlinear systems. *Automatica*, 40(6), 957 – 975.

Moreno, J.A. and Osorio, M. (2012). Strict Lyapunov functions for the super-twisting algorithm. *IEEE Trans. Autom. Control*, 57(4), 1035–1040.

Orponen, P. (1994). Computational complexity of neural networks: a survey. *Nordic J. Comput.*, 1(1), 94–110.

Ortner, P. and Del Re, L. (2007). Predictive control of a diesel engine air path. *IEEE Trans. Control Syst. Tech.*, 15(3), 449–456.

Park, I., Hong, S., and Sunwoo, M. (2014). Robust air-to-fuel ratio and boost pressure controller design for the EGR and VGT systems using quantitative feedback theory. *IEEE Trans. Control Syst. Tech.*, 22(6), 2218–2231.

Schmidhuber, J. (2015). Deep learning in neural networks: An overview. *Neural networks*, 61, 85–117.

Shtessel, Y., Edwards, C., Fridman, L., and Levant, A. (2014). *Sliding mode control and observation*. Springer.

Sontag, E.D. (2002). For differential equations with r parameters, $2r+1$ experiments are enough for identification. *J. Nonlinear Sci.*, 12(6), 553–583.

Tornambe, A. (1992). High-gain observers for non-linear systems. *Int. J. Syst. Sci.*, 23(9), 1475–1489.

Verschaeren, R. and Verhelst, S. (2018). Increasing exhaust temperature to enable after-treatment operation on a two-stage turbo-charged medium speed marine Diesel engine. *Energy*, 147, 681–687.

Wahlstrom, J., Eriksson, L., and Nielsen, L. (2010). EGR-VGT control and tuning for pumping work minimization and emission control. *IEEE Trans. Control Syst. Tech.*, 18(4), 993–1003.

Watson, N. and Janota, M. (1982). *Turbocharging the internal combustion engine*. Macmillan Int.