Chapter

Ecological Predictive Maintenance of Diesel Engines

António Simões, José Torres Farinha and Inácio Fonseca

Abstract

The ecological predictive maintenance (EPM) of diesel engines is a great contribution to improve the environment and to stimulate good practices with good impact in the human health. The ecology is a rapidly developing scientific discipline with great relevance to a sustainable world, whose development is not complete as a mature theory. There are, however, general principles emerging that may facilitate the development of such theory. In the meantime, these principles can serve as useful guides for EPM. According to the state of the art, it can be stated that through prediction algorithms, the equipment's performance can be improved. To support this approach, it is necessary to implement a good condition monitoring maintenance. The result permits to maximise the time spacing between interventions and to increase the reliability levels. The condition variables of each equipment can be monitored according to their specificity, such as temperature, humidity, pollutant emissions (NO_x, CO_2, HC and PM), emitted noise, etc. The environment where the equipment is inserted also must be considered. The assessment of the equipment's condition can be done by Hidden Markov Models (HMM), namely diesel engines. This chapter presents two algorithms—*Viterbi* and *Baum-Welch* algorithms—that, through the prediction of the equipment's condition, help to increase the efficiency of the maintenance planning.

Keywords: ecological maintenance, predictive maintenance, ecological predictive maintenance, diesel engines, Hidden Markov Model

1. Introduction

The e-maintenance is expanding in many industrial domains. Diesel engines segment is being revealed as one of the most important in its application. e-Maintenance is synonymous of effectiveness and efficiency, both by operators and the society in general. It permits to make anticipate diagnosis and prognosis conducting to a better organisation and people performance. This is the way to increase data precision and the decision confidence level.

To summarise, it can be said that e-maintenance allows to reduce human intervention, because it is a powerful mechanism, used in automatic mode to determine on-the-fly changes in the engine condition. This approach uses information and communication technologies (ICT) to provide, automatically, logistical support for the technicians.

The development presented in this chapter is supported on a distributed acquisition system with intelligent processing. This process helps to reduce engine emissions and, simultaneously, decreases costs for people, families, companies and organisations.

The solution can be extended to alert in general for the need of inspection of the vehicles that circulate in a certain street arch. This is one way to contribute to reduce the environmental impact and to improve the life quality in the urban centres.

An application was developed using MatLab, in which the algorithms based on HMM were coded. Other specific developments have been added by other languages such as PHP, SQL, etc. The sensing equipment includes vibration monitoring, emission gas analysis, sonometer, opacimeter and high-resolution digital camera. Some of these devices already integrate a remote transmission system.

Detection and decision of the action to be done and the alarm triggering are made in worthy time, allowing to schedule interventions before failures and high damages.

The equipment that emits pollutants can be monitored through the measurement of HC, CO_2, NO_x, PM, vibration and noise, allowing to determine the state of the equipment [1].

The focus is to implement a model that aggregates the contributions of the information obtained from the monitored variables and their behaviour over time, in order to determine the current state and the next one. Usually, the following situations may happen: the sequence of states is linked; the current state can be predicted by the previous ones. The precedent cases are suitable to be managed by HMM that can handle many variables.

The positioning among well-defined objectives, like maintenance, costs and operational processes, is fundamental.

In the case of diesel engines, the usual is to use schedule maintenance based on distance travelled or time elapsed. Another less frequent, but available, hypothesis is to check the quality of the oil and replace it if necessary. Recent models of some brands already have an oil sensor that gives this information to the vehicle's computer and the driver's panel.

e-Maintenance is an e-management supporting tool based on several views and perspectives. It can be defined as:

- Intelligent system with resources for information gathering, processing and decision-making. It includes transmission technology, sensor technology, maintenance activities—logistics, maintenance plans, etc., as well as diagnostic and prognostic capabilities—sensors, computer power, digital information and smart algorithms.

- e-Maintenance gives access to maintenance aspects like the following:

  i. Remote maintenance

  ii. Predictive maintenance

  iii. Real-time maintenance

  iv. Collaborative maintenance

The increasingly technological development allows the construction of intelligent equipment with the capacity for sensing, processing and transmission, allowing alarms and real-time events depending on the diesel engine's state.

This digital information allows to use condition-based maintenance (CBM) as part of the e-maintenance solution.

Maintenance information enables not only to increase the effectiveness and efficiency of the diesel engines but also to provide the persons with informative internet services. For example, the information about unexpected maintenance activities within a diesel engine and maintenance process enables opportunistic maintenance in order to reduce the negative impact [2].

Simultaneously, the same information can be correlated to spatial data in order to provide better decision support for a route planning aimed for the consumer—offering good information is essential about the services that support the performance contributing to greater consumer satisfaction of services. Managers and technicians recognise the relevance of the oil analysis applied to fleet predictive maintenance based on condition monitoring [3]. This considerably decreases the reaction time to solve critical problems of diesel engines and optimises overall equipment and vehicle trustworthiness [4].

The projected model (**Figure 1**) describes an integrated platform that is called the diesel engine e-maintenance (DEEM) that includes the items of the subsystem visible in **Figure 1**.

This model contains a sub-item called ecological predictive maintenance (EPM) based on environmental indicators [5].

In the EPM, the emission spectrum and the HMM are the innovative matter presented in this chapter. The motivation is based on the usefulness of the emission spectrum and its coherence, which can be used in a viable way by specialists who perform condition monitoring through wireless technology.

At present time, the new hardware and software solutions require a more complex integration and communication among the several pieces of this complex "puzzle".

One thing that concerns since the first time is the simplicity, friendliness and low cost of all system. The central system is based on a Linux server running Apache web server and PostgreSQL database. All system is available through IPv4 connectivity from the acquisition system level to the Linux server. Data acquisition can be done using special low-cost hardware, as also by high-performance acquisition systems, like National acquisition hardware using LabView, obtainable by IPv4 connectivity and Ethernet PLC's. It is also available by Transmission Control Protocol/Internet Protocol (TCP/IP) server for reception of data acquired from different

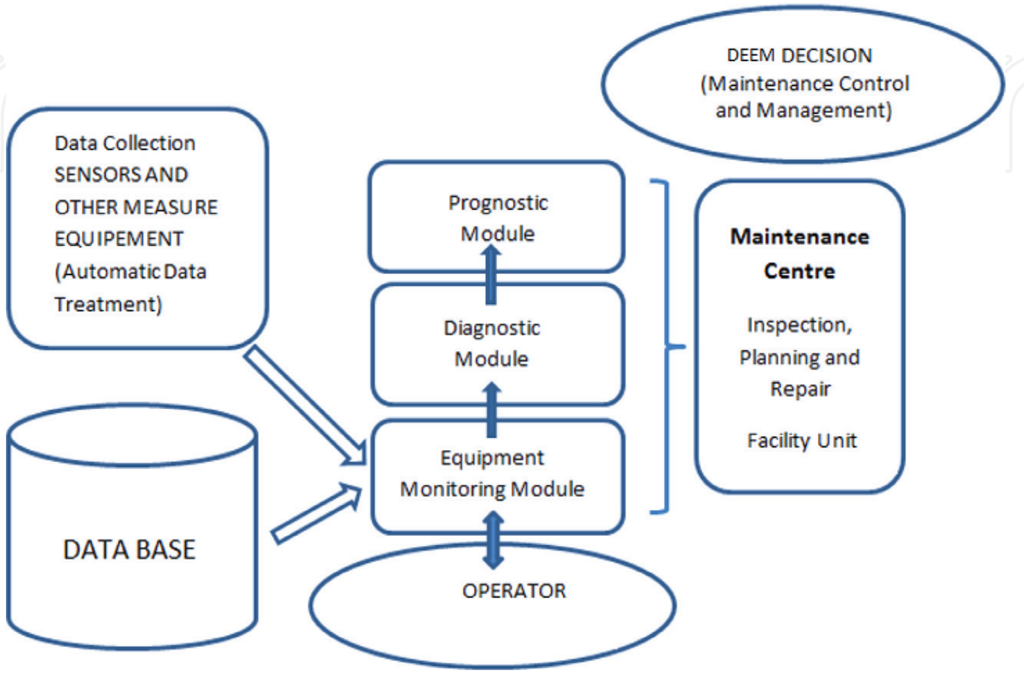

Figure 1.
Diesel engines e-maintenance.

acquisition hardware, using Unreliable Datagram Protocol (UDP) packets with acknowledgement.

Nowadays, e-maintenance systems have new adding, like the following:

- Wireless communication to IP devices to receive measurements from diesel engines or any others
- Condition monitoring modules to predict interventions based on variables that are regularly measured by remote way, by physical connecting or by human reading

2. Ecological diesel engines e-maintenance

New trends in this new paradigm are to diminish faults and respect the environment.

The system has included a prediction algorithm for condition monitoring maintenance that uses a new forecast paradigm based on HMM [6, 7].

The use of artificial intelligence, as neural networks with the objective to maximise chances of success, is a great challenge. The presented model begins with the measure of condition variables as source data that will permit to forecast the condition monitoring indicators, now through HMM models, and after, predicts the new state.

To be able to measure the condition monitoring signals in an optimal approach and to support the diagnostic and prognostic phases of e-maintenance, the analysis of refined signals must be used.

The Ecological Maintenance Performance Indicators (EMPI) are based on safety, reduction of downtime, health, pollution mitigation, costs and waste and on improving productivity, capacity utilisation and quality [8].

Thus, using a specific group of statistics and reference conditions (requirements/targets), the EMPI evaluates the actual conditions.

Figure 2 shows an e-maintenance framework proposed for maintenance management. The local platform involves the condition monitoring system and the monitored systems and vehicles. Performance and health data for vehicles and systems are logged by the condition monitoring system [3].

The CPU processing power is very important, but some sensors already integrate processing capabilities, what provides information already processed for the analysis.

Based on reviews, it appears that, despite much research ongoing, several articles on various signal analysis techniques have been published, on essential methods, and have achieved viable success.

In order to implement the acquisition of monitored signals, an acquisition system is proposed in **Figure 3**.

The back-office server system runs Linux and MySQL to store the values of the physical variables in their digital format. The acquisition system installed on the vehicles has four possible options:

1. Beckhoff PLC with Ethernet and acquisition cards
2. ARM microcontroller with Ethernet and ADCs
3. National CompactRio with LabView for prototyping research
4. Compact PC for local installation

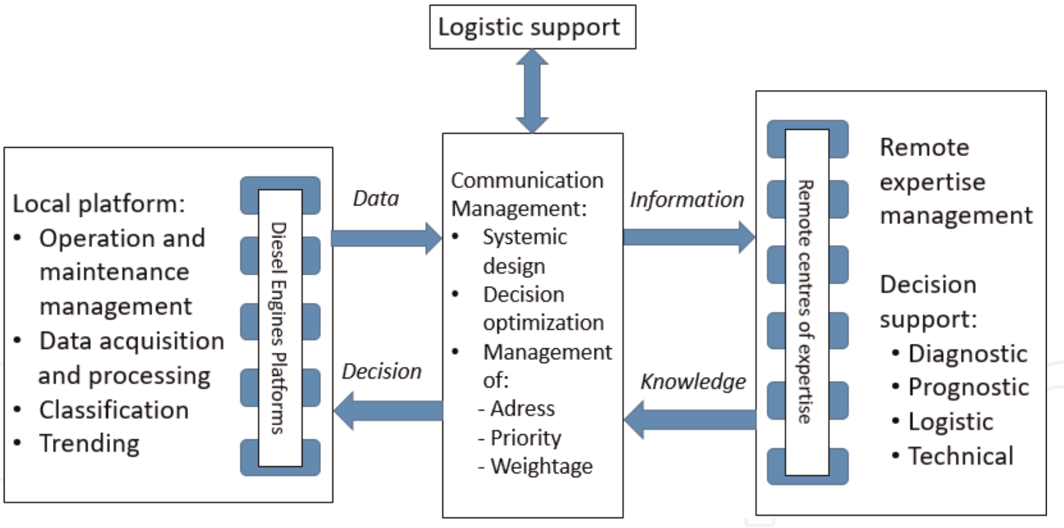

Figure 2.
Framework for e-maintenance [9].

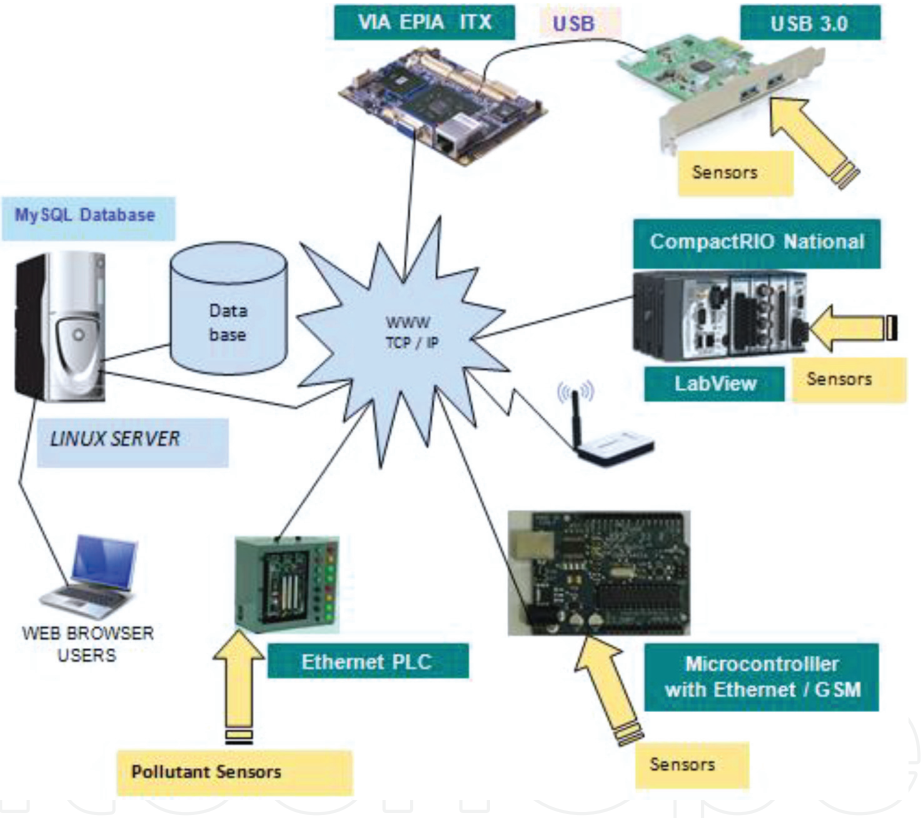

Figure 3.
Data acquisition system.

The physical variables can be acquired by each solution and transmitted via Ethernet to the back-office Linux server.

Typical key performance indicators used in fleet management include operating cost, asset availability and lost-time injuries.

In our analyses, we have defined an emission spectrum to characterise the pollution impact and to infer the engine vehicle class state (**Figure 4**). This matrix includes effluents and noise. At the same time, the Vehicle Specific Power (VSP) is used.

To reduce downtime, improve the environment, reduce waste and costs and increase process capability, emission spectra (ES) and overall equipment efficiency (OEE) evaluate the performance, as the main key performance indicators that the fleet maintenance needs for continuous improvement.

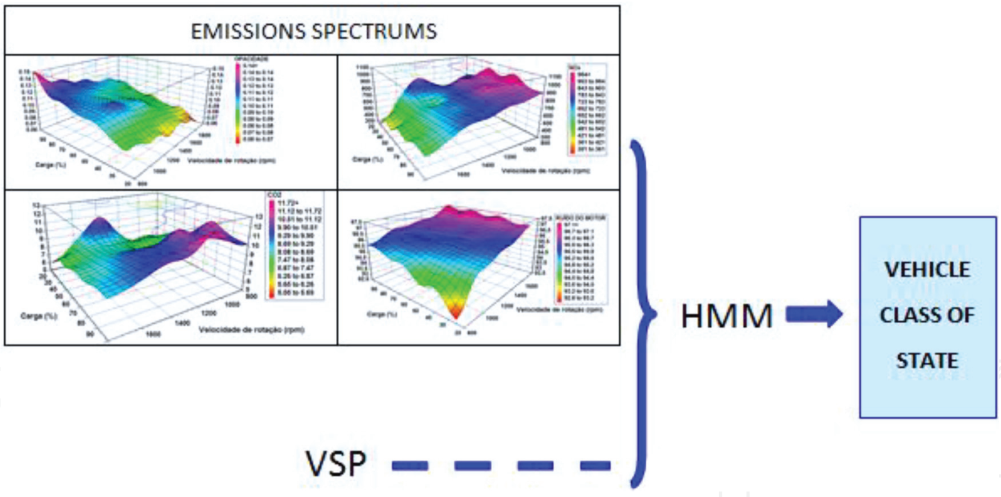

Figure 4.
Vehicle class of state evaluation.

To determine the impact on the environment and performance for each individual vehicle, maintenance, OEE, the four ES elements (Noise, NO_x, Opacity (or PM) and CO_2), quality of service and availability are very important.

The important variable CO was not used because this pollutant is not significant in diesel engines. However, it is very relevant in spark ignition vehicles.

Preventive maintenance services and activities are driven by road side sensors or vehicle metre readings. To help and simplify the fleet maintenance management process, all e-maintenance features are combined.

The platform provides the fleet managers with the necessary tools to supervise the operation, such as report generation and others, anytime, anywhere. All those involved in the fleet management increase their degree of satisfaction.

In the development of diesel engines e-maintenance (DEEM), the platform uses the following modules/resources:

- Technical management: methods of condition monitoring and tools, prognosis and state diagnosis.

- Data acquisition and reporting: use of wireless internet when possible, the concepts of e-maintenance management, the newest information technologies, mobile communications for data collection and reporting of malfunctions.

- Analysis and modelling of technical data: tools with advanced techniques that allow to perform environmental impact analysis, HMM analysis, maintenance engineering analysis, reliability block diagram, cause and consequence analysis, life cycle cost, FTA, FMECA, ETA, part counting analysis and tolerance analysis. MatLab was used to construct the model matrix and the response matrix with the predicted values.

- Logistics and systems design: tools that streamline data and information between institutions and departments, aiming to assist management and business decisions. They also assure the project selection of variables and maintenance levels and evaluate diversity and redundancy, levels and factors, maintainability versus reliability, diagnosis and modularity, trade-off studies, ecological predictive maintenance *versus* preventive maintenance programme, replacement and part control programme.

The relationship between dysfunctions and symptoms plays a relevant role in the understanding (i.e. diagnosis) of the health condition. A lot of work has been implemented, in order to anticipate the demand for spare parts in the future, almost always automatically.

The environmental specifications and regulations applicable to each case, as well as international standards, help to define the classes and thus adopting a set of limits.

Figure 5 shows DEEM's architecture.

There is a high correlation between the diesel engine states, the diagnosis obtained and the emissions spectra data. However, automatic diagnosis is still not the rule, despite countless successful research applications.

In the future, with data available from numerous sensors, their analysis will be done by automatic programmes to support diagnosis and prognosis. Determining the physical and ecological state of the fleets based on the mixing of emission signals and the dynamic signals of vehicles combined with the OEE is a major challenge. With the automatic acquisition of emissions spectra and traffic signals and advanced processing, even because the volume of data and its relationship cannot be perceived by humans in time, there is a new step in maintenance. So, for humans, this intelligent processing is like a black box: the inputs and the result are known—inside there are algorithms and technologies not noticeable. The future

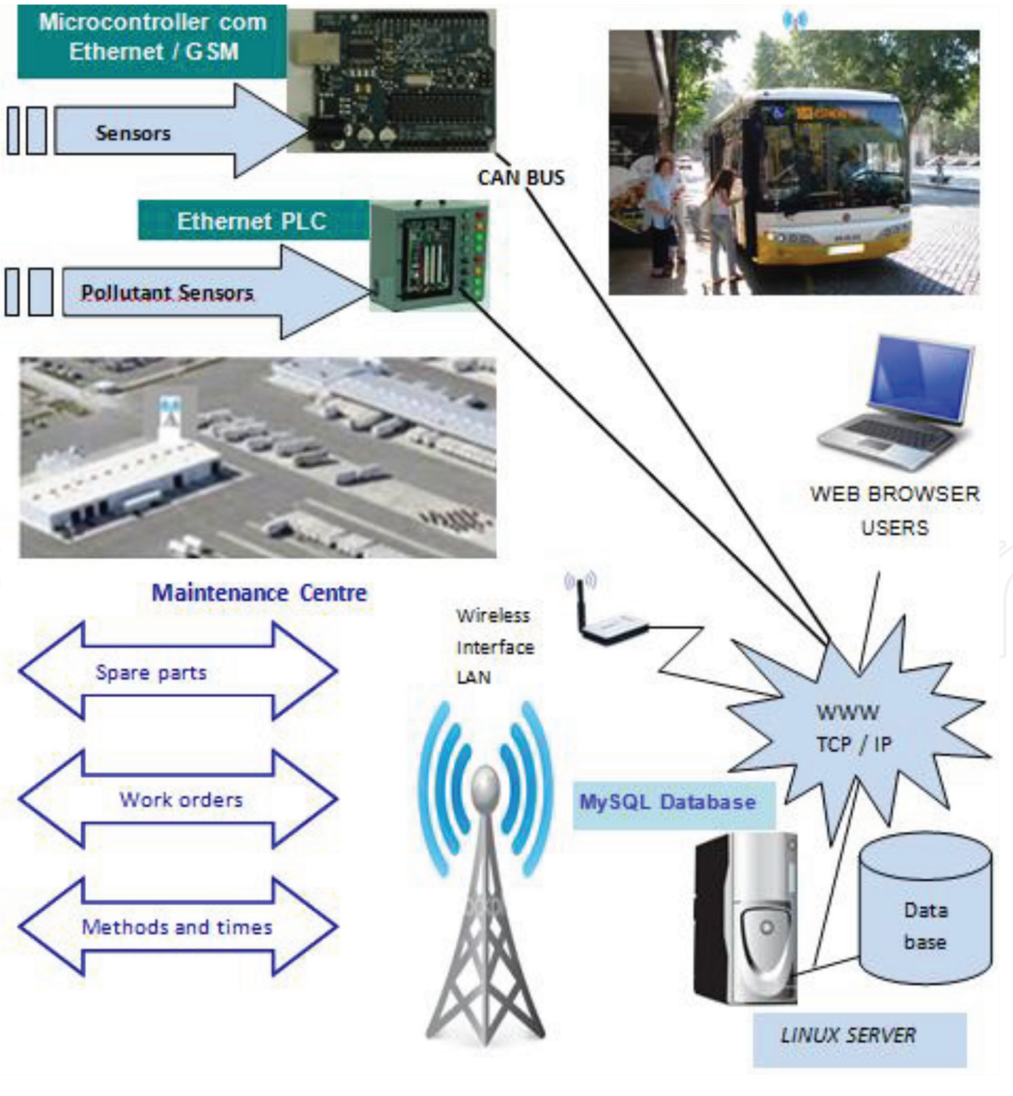

Figure 5.
Diesel vehicle e-maintenance layout.

goes in this direction of all digitization and entry of artificial intelligence in the study of the resolution of these problems.

Surprisingly, automatic diagnostics are little used in transport, although much research is done to solve the problem.

The central and municipal administrations have still adhered little, although there are strong indications of new ways. The big challenge is to tailor the entire infrastructure to install sensors and tools, communications and signal analysis to produce useful results.

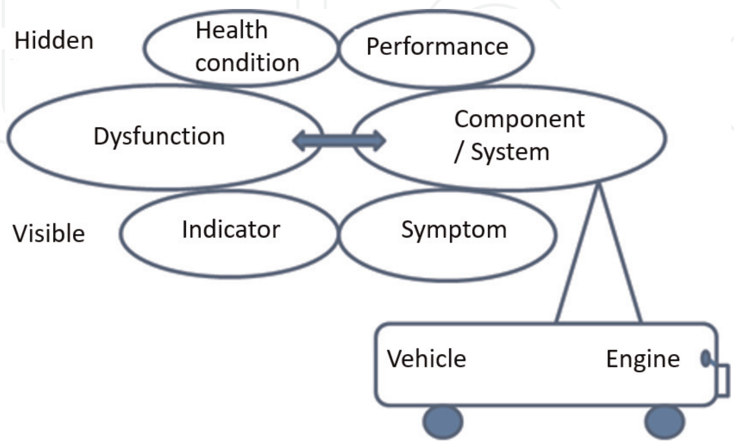

Figure 6.
Fleet model conceptualization.

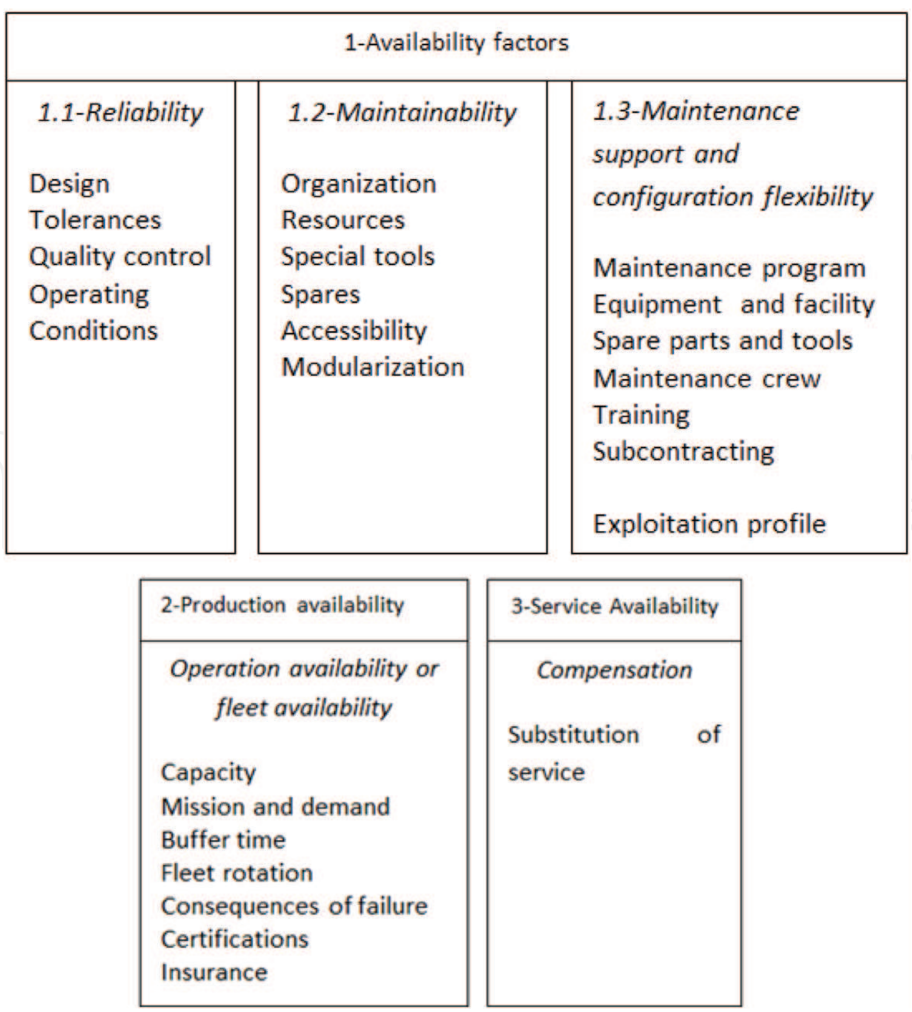

Figure 7.
Availability factors groups.

Obviously, the correct sensors must be used with appropriate techniques, although the operators have not yet adhered to and facilitated this path.

As a result, continuous research and the obtained results point to obtaining usable knowledge in the scope of the component wear and the correct detection of the prognosis, with the possibility of automating the process. In the future, both the actual data and the simulations for the situations experienced will be available to support the prognosis, diagnosis and monitoring of the condition state.

Maintenance beyond tomorrow solutions can carry out the requests, the acts of management and allocation of the human resources, automatically. If there are errors, neural networks or HMM are important because they evaluate and adapt.

Each fleet has its own context, operational service and performance [10].

The corresponding component and context can define the meaningful health indicator. For example, the system environment may cause abnormal behaviour. In this case, contextual information does not only use technical or service level issues but also environmental issues.

Different criteria or contextualization can help assess the health status of a component. Abnormal behaviours can be defined by symptom indicators, representing the health condition.

The performance context is one of the main objectives in a fleet and is associated to the optimisation capacity (costs and equipment availability).

The relation between dysfunctions and symptoms helps to understand the health condition of the equipment designated by diagnosis (**Figure 6**).

Achieving the objectives of a company goes through the fleet to meet performance and availability goals, both now and in the future, and for this e-platform, there is a programme for systematic monitoring and evaluation of various aspects of technical activities and services for diesel engine management (**Figure 7**).

In the operation phase, observations of the fleet availability performance should be used to evaluate the need for corrective actions, that is, improvements and modifications. Feedback should identify [11]:

1. The performance of the fleet according to the established goals and requirements

2. In the case of goals and requirements not reached, where and why the deviations occur

3. The improvements implemented through reliable and low-cost solutions

The main conclusions reported in this article are the following: Why use HMM? How and what model to build? How the values observed can be read? How the model is trained? Why the model is appropriate and how to assess the perplexity index [12].

3. Hidden markov model

Markov processes are a class of probabilistic models used to study the evolution of systems along time. Transition probabilities help to identify the evolution of the system between periods. The Markov chain, within its logic, characterises the temporal behaviour of the system, as described in the probability matrix in the first state.

In order to apply an HMM to the process, it must fall in the next requirements:

- The process must be stochastic.
- The probability to move from one to another state does not depend on the transitions of earlier times. Thus, in calculations only the previous state is necessary.

A Hidden Markov Model inherits from the Markov Model (first order) [13]:

- The states of the model are only observable indirectly and must be inferred (hidden states).
- The measurable variables (emissions, in the present case) depend on the probability of the hidden states.
- Hidden states inherited the behaviour of states from a first-order Markov process.

The next set of probabilities and states/emissions (in this approach) is a set of correlated states of the Hidden Markov Model, that is:

- The hidden states: represent the true state, referring to the Markov process.
- The observable variables (emissions): the visible face of the process, which allow to infer the hidden states.
- The probabilities of the states hidden in the initial step.
- The matrix of transition probabilities: related to hidden states.
- The probability matrix of the output symbol of the model: each element exposes the probability of generating an emission symbol, considering the state in which the model is the probability matrix of the observations.

The present chapter, in the next sections, both in the simplicity and performance of the model, shows that the Hidden Markov Model (HMM) can be used to predict the equipment condition.

There is an initial state probability vector, which represents the probability of the system starting in a given class that links to a state.

In practice, the emissions of a vehicle depend on many factors (variables), such as temperature and engine condition, but also driving, operating conditions and maintenance.

System planning performs cycles between interventions, starting at zero (new vehicle or repaired as new), evolving over time with maintenance and exploration leading to evolution in the states.

In some cases, it is interesting to apply the HMM to continuous densities of observations. For this to be possible, some constraints have to be applied to the probability density function of the model, thus ensuring a consistent re-estimation of the parameters of this function.

A discrete or continuous HMM have the following properties:

- The time of permanence in a state is Markovian (process without memory).
- The next state depends only on the instant of transition and the current state: Markovian property.

The process can only be discrete (HMM) due to the following two reasons:

1. The dwell time in a state does not need to follow an exponential distribution.

2. In the present approach, this time depends on the climate, the acceleration and other changeable variables. In addition, it is not possible to continuously measure emissions.

The classification of the data introduces difficulties. Researchers using complex classifiers admit that they can solve almost the entire problem. It is a signal that a good classifier gives better performance and avoids system malfunction—which is not always the case.

The present work uses statistics, but the use of neural networks may be a good option.

3.1 Emissions and states

Emissions from a diesel engine is a specific situation assumed that there is an initial state (when the engine is new or rebuilt like new), referenced as the state at instant zero. Next states evolve until the limits imposed by international standards, environmental rules and some requisites of each real situation are reached. **Figure 8** shows a generic evolution of states and observations in diesel engines context [14].

- Q: $\{q_1 \ldots q_N\}$—Set of possible values for the hidden states, designated as "states vector" or "states library"

- V: $\{V_1 \ldots V_M\}$—Set of possible values for the observations, designated as "observations vector" or "emissions collection"

The time evolution of vectors associated with the emission matrix and transition matrix is presented in **Figure 9**, that is:

- $A = \{a_{ij}\}$—Probability matrix (state transition)

- $B = \{b_{ik}\}$—Probability matrix (state observation)

The vector for the initial instant must be known as the start model:

- $\prod = \{\prod_i\}$—Probability matrix (initial state)

The HMM Model comes as:

$$\left\{Q, V, \prod, A, B\right\} \tag{1}$$

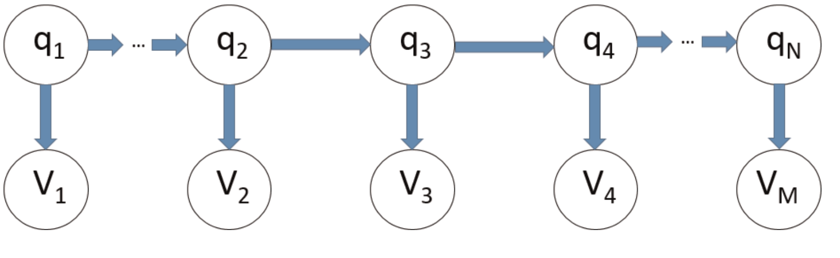

Figure 8.
Hidden states and observed symbols in a diesel engine.

The implementation of HMM follows the following points:

1. To determine the probability of the observation sequence

2. To know the sequence of observations and to determine the most appropriate sequence of hidden states

3. To know the set of possible models and the observation sequence and to determine which model best fits the data

For a model and an observation sequence, the corresponding sequence probability is as follows:

$$O = (o_1 \ldots o_T), \lambda = (A, B, \Pi) \tag{2}$$

One should calculate $P(O|\lambda)$. A hypothesis for this calculation, the sum of the probabilities of the possible state sequences (which lead to this state), is equal to the probability of the observation sequence. However:

- The calculation to the gross force takes time, because in the observations T and N, there is N^T sequence of possible states.

- If the HMM is small, e.g. $T = 9$ and $N = 9$, there are approximately 387.5 million hypotheses.

Dynamic programming is one way of solving the problem. The steps to achieve this solution are the following:

$$P(O|S, \lambda) = b_{s_1 o_1} b_{s_2 o_2} \ldots b_{s_T o_T} \tag{3}$$

$$P(S|\lambda) = \pi_{s_1} a_{s_1 s_2} a_{s_2 s_3} \ldots a_{s_{T-1} s_T} \tag{4}$$

$$P(O|\lambda) = \sum_S P(O|S, \lambda) P(S|\lambda) \tag{5}$$

$$P(O, S|\lambda) = P(O|S, \lambda) P(S|\lambda) \tag{6}$$

that can be resumed by:

$$P(O|\lambda) = \sum_{\{S_1 \ldots S_T\}} \pi_{s_1} b_{s_1 o_1} \prod_{t=1}^{T-1} a_{s_t s_{t+1}} b_{s_{t+1} o_{t+1}} \tag{7}$$

which is the basis of the *forward-backward* algorithm described in the next section.

The HMM algorithm has quite a few unsolved problems that lead to many answers, giving to the succeeding actions:

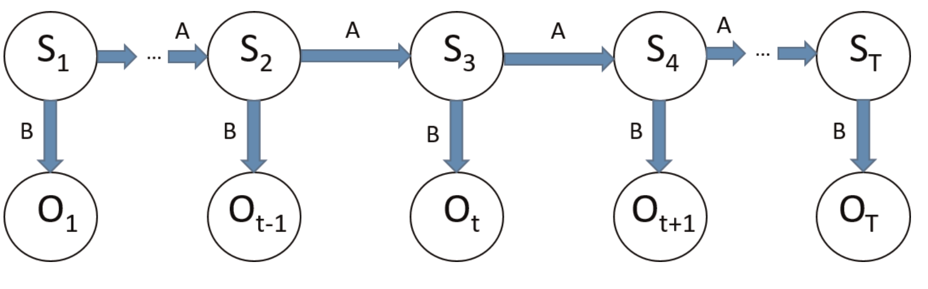

Figure 9.
HMM with transition probability matrixes.

I. Assessment:

1. Problem: Are there several possible HMMs and a number of observations where HMM probably generate the given sequence?

2. Solution

- For each HMM calculates the probability of the observed sequences.
- The most likely sequence must be chosen.
- Reduce complexity through the *Forward* algorithm.

II. Interpreting:

1. Problem: For an observation sequence and a given HMM, determine the most likely sequence (underlying hidden sequence) that gave rise to the observations sequence.

2. Solution:

- In a possible sequence of underlying hidden states, determine the probability of the observed sequences.
- The most likely sequence must be chosen.
- Reduce complexity with the *Viterbi* algorithm.

III. Knowledge:

1. Problem: It does estimate the probabilities of the HMM from the training data.

2. Solution:

a. Train with labelled data [15]:

- To evaluate the transition probability
- $P(q_i,q_j)$ = (number of transitions from q_i to q_j)/(total number of transitions of q_i)
- The observation probability matrix $P(q_i, V)$ = (number of symbol V occurrences in state q_i)/(number of all symbol occurrences in state q_i)

b. Train with the unlabelled data:

- *Baum-Welch* algorithm basic idea:
    i. To maximise the probability of the observation sequence, given the model
    ii. To estimate new probability from the previous HMM until P (current HMM) – P (previous HMM) $< e$ (a small number)

3. Problem: Use training data to estimate HMM probabilities.

4. Solution:

a. Train using marked data [15]:

- Transition probability
- P(q_i,q_j) = (transitions number from q_i to q_j)/(entire number of transitions of q_i)
- Observation likelihood matrix P(q_i, V) = (number of symbol V incidences in state q_i)/(number of all symbol incidences in state q_i)

b. Train with unmarked data:

- Basic idea—*Baum-Welch* algorithm:

i. For the model, maximise the probability of the observation sequence.

ii. Evaluate new probability from the preceding HMM until P (current HMM) – P (previous HMM) < e (a lesser number).

For unmarked and marked data [16], the following sections provide a solution to solve this problem.

3.2 *Forward-backward* algorithm

Predicting hidden states is complex and requires efficient algorithms to solve the problem, which can be solved by applying the principles of *dynamic programming* through the *forward-backward algorithm*.

It uses the auxiliary variable forward α_t (i) that is the probability of observing the partial sequence “O1, O2, ..., Ot” and that at time t we have state Si. This can be defined by Eq. (5):

$$\alpha_t(i) = P(O_1, O_2, O_3, \ldots\ldots, O_t, S_t = q_i|\lambda) \tag{8}$$

that is the basis of the *forward* procedure.

The calculation of $\alpha_t(i)$is achieved by the following steps:

$$\begin{aligned}
&\alpha_1(i) = \pi_i b_i(O_1), \quad 1 \le i \le N \\
&\alpha_{t+1}(j) = P\left(O_1, \ldots\ldots, O_t, O_{t+1}, S_{t+1} = q_j \middle| \lambda\right) = \\
&\sum_{i=1}^{N} P\left[O_1, \ldots., O_t, O_{t+1}, S_{t+1} = q_j \middle| S_t = q_i, \lambda\right].P[S_t = q_i|\lambda] = \\
&\sum_{i=1}^{N} P[O_1, O_2 \ldots., O_t|S_t = q_i, \lambda].P[S_t = q_i|\lambda].P\left[O_{t+1}, S_{t+1} = q_j \middle| S_t = q_i, \lambda\right]. = \\
&\sum_{i=1}^{N} P[O_1, \ldots., O_t, S_t = q_i|\lambda].P\left[O_{t+1}|S_{t+1} = q_j, S_t = q_i, \lambda\right].P\left[S_{t+1} = q_j \middle| S_t = q_i, \lambda\right]
\end{aligned} \tag{9}$$

$$\alpha_{t+1}(j) = b_j(O_{t+1})\left[\sum_{i=1}^{N}\alpha_t(i)a_{ij}\right] \tag{10}$$

$$\alpha_{t+1}(j) = \left[\sum_{i=1}^{N}\alpha_t(i)a_{ij}\right]b_j(O_{t+1}), \quad 1 \leq t \leq T, \quad 1 \leq j \leq N \tag{11}$$

Given a known state at time t, the next step is the *backward* algorithm, which calculates the sequence of probabilities in the observations. The following equations formalise the process:

$$\beta_t(i) = P\left(O_{t+1}, O_{t+2}, O_{t+3}, \ldots\ldots, O_T | S_t = q_i, \lambda\right) \tag{12}$$

- Initialization:

$$\beta_T(i) = 1, \quad 1 \leq i \leq N \tag{13}$$

- Induction:

$$\beta_t(i) = \sum_{j=1}^{N} a_{ij} b_j(O_{t+1})\beta_{t+1}(j), \quad t = T-1, T-2, \ldots 1 \text{ and } 1 \leq i \leq N \tag{14}$$

The final steps for decoding the solution are

$$P(O|\lambda) = \sum_{i=1}^{N}\alpha_T(i) \tag{15}$$

That is, the forward procedure, and

$$P(O|\lambda) = \sum_{i=1}^{N}\pi_i b_i(O_1)\beta_1(i) \tag{16}$$

that is, the backward procedure.

3.3 *Viterbi* algorithm

Now we will try to find the states sequence that best explains the observations. This can be done by the *Viterbi* algorithm, synthesised by Eq. (12):

$$S^* = \arg \underset{S}{\acute{m}ax} \quad P(S, O|\lambda) \tag{17}$$

To evaluate this equation, an auxiliary variable $\delta_t(i)$is defined. This variable corresponds to the maximum result (or higher probability) of the recorded sequence of observations, assuming that the end state is stage q_i:

$$\delta_t(i) = \underset{S_1, S_2, S_3, \ldots S_{t-1}}{\acute{m}ax} \quad P\left[S_1 S_2 S_3 \ldots\ldots S_{t-1}, \ S_t = q_i, \ \mathrm{O}_1 O_2 O_3 \ldots\ldots O_t | \lambda\right) \tag{18}$$

The algorithm may be compactly stated by:
- Initialization:

$$\begin{aligned}&\delta_1(i) = \pi_i b_i(O_1), \quad 1 \le i \le N\\&\psi_1(i) = 0\end{aligned} \tag{19}$$

• Recursive computation:

$$\delta_t(j) = \underset{1 \le i \le N}{\acute{m}ax} \left[\delta_{t-1}(i)a_{ij}\right]b_j(O_t), \quad 2 \le t \le T \quad \text{and} \quad 1 \le j \le N \tag{20}$$

$$\psi_t(j) = \arg \underset{1 \le i \le N}{\acute{m}ax} \left[\delta_{t-1}(i)a_{ij}\right], \quad 2 \le t \le T \quad \text{and} \quad 1 \le j \le N \tag{21}$$

• Termination:

$$P^* = \underset{1 \le i \le N}{\acute{m}ax} \left[\delta_T(i)\right] \tag{22}$$

$$S_T{}^* = \arg \underset{1 \le i \le N}{\acute{m}ax} \left[\delta_T(i)\right] \tag{23}$$

Next step consists of computing the most likely state sequence by backtracking:

$$S_t{}^* = \arg \underset{1 \le i \le N}{\acute{m}ax} \left[\delta_t(i)a_{iq^*_{t+1}}\right] \tag{24}$$

$$S_t{}^* = \psi_{t+1}\left(S^*_{t+1}\right), \quad t = T-1, T-2, \ldots.., 1 \tag{25}$$

The result is the generation of the sequence:

$$S_1, S_2, S_3 \ldots\ldots S_{t-1}, \quad S_t \tag{26}$$

3.4 *Baum-Welch* algorithm

The final step is the calibration of the parameters to fully define the HMM; for a sequence of observations, the question is: what is the corresponding model with the same behaviour? For a model and a sequence of observations, adjust the parameters of this model in order to more accurately approximate the observations.

$\xi_t(i,j)$ defines the probability of the system being in the state q_i at time t and in state q_j at time $t + 1$:

$$\xi_t(i,j) = \frac{\alpha_t(i)a_{ij}b_j(O_{t+1})\beta_{t+1}(j)}{\sum_{i=1}^{N}\sum_{j=1}^{N}\alpha_t(i)a_{ij}b_j(O_{t+1})\beta_{t+1}(j)} \tag{27}$$

Eq. (19) gives the probability of to occur an *i-j* transition.

$$\gamma_t(i) = \sum_{j=1}^{N} \xi_t(i,j) \quad \mathrm{t = 1....T} \tag{28}$$

Eq. (20) gives the probability of being in state i at time t.
Now it is possible to compute new estimates for the model parameters:

$$\hat{\pi}_i = \gamma_1(i) \tag{29}$$

$$\hat{a}_{ij} = \frac{\sum_{t=1}^{T-1}\xi_t(i,j)}{\sum_{t=1}^{T-1}\gamma_t(i)} \tag{30}$$

$$\hat{b}_j(k) = \frac{\sum_{t=1}^{T}\gamma_t(j)_{\mathrm{For}\ \ O_t=V_k}}{\sum_{t=1}^{T}\gamma_t(j)} \tag{31}$$

3.5 Perplexity measurement

The performance of the HMM model can be measured in two ways:

1. By classification accuracy

2. By perplexity

The ratio of the number of correct predictions of the state vector hidden by the total number of hidden states (the analysed candidates) defines the accuracy of the classification. For a small sample, the classification accuracy is unreliable and is contaminated with noise, since a sample can be assigned to a single class. For this reason, the perplexity of the data set is a more appropriate alternative. This metric measures the confidence of the predictors of the classifier through the function of the average log-likelihoods L (of N data sequence), described by Eq. (24):

$$L_i = \log p\left(q_i | o^i_{1\dots T_i}, \lambda\right) \; perp = e^{-\frac{1}{N_s}\sum_{i=1}^{N_s} L_i} \tag{32}$$

where $o^i_{1\dots T_s}$ represents for the *ith* sequences of observations of length T_s and q_i is the type of hidden states i; N_s is the number of sequences and λ the model parameters. The value one is the best quantity for the perplexity, giving probability 1 for the correct task type. A perplexity of 3 means a random decision, with a probability of one-third for each hidden state [17].

4. Implementation

Maintenance performance indicators (MPI) are a metric to gauge how the system behaves, that is, measure availability, costs and wastes, process capacity, productivity, quality, health, safety and environmental impact. The objective allows a macro idea of the state of operation of the fleet [18].

The targets for each of the MPIs used can be established as a requirement to compare each measure during the operation period in the future.

The metric associated with OEE, consisting of three elements, Quality, Performance and Availability, allows to calculate the influence of the performance of an equipment's part. The concept of e-maintenance was used in cars and buses to automatically manage the fleet. Their particularities are the following:

a. Notifications of anomalies by *email*.

b. Detect problems in advance before they occur and undetectable for drivers.

c. Automated system, either at the level of requirements or at the level of consumables, indicates a loss of performance, parts to be replaced and so on.

It is possible to make a remote access to the computers of the buses—the technical assistance can unlock some anomalous situations.

In the case of diesel engines, as mentioned above, they can be represented by a matrix of states taking into account the following variables: {CO_2, NOx, PM, NOISE, HC}; the combinations/intercepts designate the observable emission variables {V_1, V_2, V_3, V_4 .... V_M}.

Thus, including the initial, there are five possible states, and the boundaries to segment the states are:

1. Bad
2. Dysfunctional
3. Tolerable
4. Good
5. Excellent

As the transition matrix is 5 × 5, q1, q2, q3, q4 and q5, t is necessary to identify the four state classes.

Of course, the decision thresholds are four for each variable:

1. Poor or bad
2. Acceptable
3. Good
4. Excellent

At the moment, the model incorporates the variables {PM, NO_x and $NOISE$}.

If the equipment is new or after a maintenance intervention has gone into effect, the hidden and actual initial states may overlap. Generally:

$$\pi_i = P\left[S_1 = q_i \middle| \right], \quad 1 \leq i \leq N \tag{33}$$

where $\pi_i = Number\ of\ times\ in\ state\ q_i\ at\ time\ t = 1$.

The elements of the matrix q must be calculated using

$$a_{ij} = P\left[S_{t+1} = \mathrm{q}_j \middle| S_t = \mathrm{q}_i\right] = \frac{number\ of\ transitions\ from\ \mathrm{q}_i\ to\ \mathrm{q}_j}{number\ of\ times\ in\ state\ \mathrm{q}_i} \tag{34}$$

$$a_{ij} = P\left[S_{t+1} = \mathrm{q}_j \middle| S_t = \mathrm{q}_i\right] = \frac{number\ of\ transitions\ from\ \mathrm{q}_i\ to\ \mathrm{q}_j}{number\ of\ times\ in\ state\ \mathrm{q}_i} \tag{35}$$

There are 64 combinations for the 4 levels associated with each. These 64 combinations are grouped into 11 emission classes, namely: V_1, V_2, V_3, V_4, ……V_{11}.

$$A = \begin{bmatrix} a_{11} & a_{12} & a_{13} & a_{14} & a_{15} \\ a_{21} & a_{22} & a_{23} & a_{24} & a_{25} \\ a_{31} & a_{32} & a_{33} & a_{34} & a_{35} \\ a_{41} & a_{42} & a_{43} & a_{44} & a_{45} \\ a_{51} & a_{52} & a_{53} & a_{54} & a_{55} \end{bmatrix}$$

The library of emissions is much more complex, since it integrates the environmental impact of different variables:

$$B = \begin{bmatrix} b_1(1) & b_1(2) & b_1(3) & b_1(4) & b_1(5) & b_1(6) & b_1(7) & b_1(8) & b_1(9) & b_1(10) & b_1(11) \\ b_2(1) & b_2(2) & b_2(3) & b_2(4) & b_2(5) & b_2(6) & b_2(7) & b_2(8) & b_2(9) & b_2(10) & b_2(11) \\ b_3(1) & b_3(2) & b_3(3) & b_3(4) & b_3(5) & b_3(6) & b_3(7) & b_3(8) & b_3(9) & b_3(10) & b_3(11) \\ b_4(1) & b_4(2) & b_4(3) & b_4(4) & b_4(5) & b_4(6) & b_4(7) & b_4(8) & b_4(9) & b_4(10) & b_4(11) \\ b_5(1) & b_5(2) & b_5(3) & b_5(4) & b_5(5) & b_5(6) & b_5(7) & b_5(8) & b_5(9) & b_5(10) & b_5(11) \end{bmatrix}$$

Next step is to test the model as exemplified in (**Table 3**).

The challenge is to diagnose not just the state associated to a combination of emissions, but also to the prognoses of the next emission level and correspondent states of health.

Using the operating conditions, one can obtain the degradation models and thus update the initial state probability matrix [8].

The current monitoring systems are enough from the point of view of the predictive diagnosis because, depending on the case, they can anticipate the failure.

As soon as the monitoring system gives the alarm, according to what you read from the sensors, depending on the case, the time to correct the problem is short, otherwise the fault will happen—it requires a detailed analysis of the alarms before or after.

The present model can determine this monitoring data with the corresponding situation at different levels.

About diagnostic results for some of the buses of a fleet are proposed in summary form. A specialist, when using a system like this, in a specific situation, should first focus on the modes of degradation already observed that are re-incidents and/or associate already solved cases that are similar to the case at hand. The system helps to perceive the behaviour without hiding its complexity, whenever different contextual information is available.

For each model a transition state matrix is created. At the same time, another matrix is created, which associates the probabilities of different classes of the emission indicators to each state.

As an example, if we consider four states and six different emission scenarios, the instructions in MatLab and outputs will be two matrixes “4×4” and “4×6”:

- seq1 = xlsread(‘inputsallbuses’,'ENGINE’,'b72:ao72’)
- states1 = xlsread (‘inputsallbusesenginesandparts’,'b18:ao18’)
- [TRANS_EST, EMIS_EST] = hmmestimate (seq1, states1)

In the present case study, the transition matrix is shown in **Table 1**.

And the emission matrix in **Table 2**.

0,350	0,450	0,100	0,100
0,125	0,000	0,875	0,000
0,286	0,000	0,429	0,286
0,571	0,000	0,000	0,429

Table 1.
Transition matrix.

0,900	0,000	0,000	0,000	0,100	0,000
0,300	0,200	0,000	0,250	0,000	0,250
0,429	0,143	0,286	0,143	0,000	0,000
0,125	0,000	0,000	0,000	0,375	0,500

Table 2.
Emission matrix.

Emission time	t = 1	t = 2	t = 3	t = 4	t = 5	t = 6	t = 7
Emissions classification	V_4	V_5	V_7	V_7	V_4	V_9	V_{10}
Real hidden state	$S_1 = q_1$	$S_2 = q_4$	$S_3 = q_4$	$S_4 = q_5$	$S_5 = q_2$	$S_6 = q_5$	$S_7 = q_5$
Hidden state classification	q_1	q_4	q_5	q_5	q_2	q_5	q_5

Table 3.
Main outputs of ecological HMM.

Actual State Probabilities, defined by nine HMM models									
	HMM1	HMM2	HMM3	HMM4	HMM5	HMM6	HMM7	HMM8	HMM9
More probable state	2	2	1	3	3	3	3	4	3
State 3 probability	$P_{3,1}$	$P_{3,2}$	$P_{3,3}$	$P_{3,4}$	$P_{3,5}$	$P_{3,6}$	$P_{3,7}$	$P_{3,8}$	$P_{3,9}$
State 4 probability	$P_{4,1}$	$P_{4,2}$	$P_{4,3}$	$P_{4,4}$	$P_{4,5}$	$P_{4,6}$	$P_{4,7}$	$P_{4,8}$	$P_{4,9}$

Table 4.
Outputs of actual states.

In the present case, it is intended to determine the next health state and the respective level of emissions and also to diagnose the state associated to the combination of emissions as shown in **Table 3**.

We can to test several models to choose the better one. **Table 4** exemplifies how different models generate different outputs of actual states. Each tested model generates the probability of the system to be in one of the possible four states. At the same time is finding the occurrence probability of unfavourable different states (Ps, m – s-state; m-model).

Therefore, the calibration of the model assumes as an important step in the construction of an asset management solution.

The model ends with a part of the forecast, in which the most probable sequence of future emissions and the corresponding states are generated. It calculates until two periods after the current remote sensing reading.

5. Conclusions

The chapter gives an overview of an open system approach to e-maintenance with an innovative proposal of how a conventional enterprise can be transformed to a fully automated e-maintenance solution based on an ecologic condition monitoring model.

The chapter corresponds to a basis for a more in-depth study based on extensive literature review and a case study from which demonstrates how e-maintenance can influence the business model of an organisation and the improvement of the environment in the cities.

The application context of the model and case study presented allow to consider the fleet components' similarities and heterogeneities. Data of the monitored diesel vehicles are considered within their context and enhance the identification of the corresponding health condition. The case study points new ways for the future.

All the developments presented are supported on an information system. The new hardware and software solutions require complex integration and communication among the several pieces of these complex technological devices but are presented to the final user as a friendly solution.

The chapter also mentions the capabilities of the proposed architecture and addresses certain challenges faced in order to enable an open framework.

The findings suggest that the different traditional practices used in preventive and condition monitoring maintenance strategies require the building of customised solutions, according to the specificity of each organisation. An open system solution can tackle the associated problem in a fairly cost-effective way.

A new development, ubiquitous and intelligent system of e-maintenance using standards "on demand", shall facilitate interoperability with existing legacy systems.

The interval between interventions in diesel engines was improved through condition monitoring maintenance planning with the input of condition variables: CO_2, NO_x, HC, NOISE and PM. The HMM model has been shown to be adequate for maintenance planning based on these variables despite the complexity of it. The *Viterbi* and *Baum-Welch* algorithms are used in the present model.

For automatic detection of environmental impacts, the efficiency of the prediction model depends on the characteristics of the system and the sampling frequency of the measured physical variables.

Urban areas can improve environmental quality, if we reduce emissions—this is a new paradigm for building a better world.

The model has a potential versatility to be applied in various fields to evaluate the health status of the equipment.

The upcoming work will involve research towards a decision tool to assess the need for vehicle maintenance, as well as actions to be taken—an integrated application for urban transport. Remote sensor devices will be used by the system to measure emissions. The chapter also mentions the capabilities of the proposed architecture and addresses some challenges in order to enable an open framework.

Acronyms

CBM	condition-based maintenance.
CO_2	carbon dioxide
DEEM	diesel engines e-maintenance
EMPI	ecological maintenance performance indicators
EPM	ecological predictive maintenance
ES	emissions spectrums
ETA	estimated time for accomplishment
FMECA	failure mode, effects and criticality analysis
FTA	fault tree analysis
HC	hydrocarbons
HMM	Hidden Markov Model

ICT	information and communication technology
NO_x	nitrogen oxides
OEE	overall equipment effectiveness
PM	particulate matter
TCP/IP	transmission control protocol/internet protocol
UDP	unreliable datagram protocol
VSP	vehicle specific power

Author details

António Simões[1,3]*, José Torres Farinha[1,3] and Inácio Fonseca[2,3]

1 Department of Mechanical Engineering, Coimbra Polytechnic—ISEC, Coimbra, Portugal

2 Department of Electrical Engineering, Coimbra Polytechnic—ISEC, Coimbra, Portugal

3 CEMMPRE (Centre for Mechanical Engineering, Materials and Processes), Coimbra, Portugal

*Address all correspondence to: assimoes@isec.pt

References

[1] Clark N, Gajendran P. Predictive Tool for Emissions from Heavy-Duty Diesel Vehicles. Department of Mechanical and Aerospace Engineering, Argonne National Laboratory; 2003

[2] Biswal M, Parida A. An integrated approach for open e-Maintenance: Opportunities and challenges. In: 1st international workshop and congress on eMaintenance; Luleå, Sweden. 2010

[3] Monnin M et al. Fleet-wide health management architecture. In: Annual Conference of the Prognostics and Health Management Society; Montreal, Canada. 2011

[4] Frey C et al. Comparing Real-World Fuel Consumption for Diesel and Hydrogen Fuelled Transit Buses and Implication for Emissions. Transportation Research – Part D. ScienceDirecct/Elsevier; 2007. DOI: 10.1016/j.trd.2007.03.003. www.elsevier.com/locate/trd

[5] Simões A et al. Ecological predictive maintenance in hospital equipment. In: Proceedings of 20th Congress of the International Federation of Hospital Engineering (IFHE), Barcelona; 19–22 October 2008

[6] Rabiner L et al. On the application of vector quantization and hidden markov models to speaker-independent, isolated word recognition. The Bell System Technical Journal. 1983;**62**(4):1075-1105

[7] Rabiner L. A tutorial on Hidden Markov Models and selected applications in speech recognition. Proceedings of the IEEE. 1989;**77**(2):257-286

[8] Simões A et al. Buses degradation based on exploration conditions. In: Proceedings of the 3rd WSEAS International Conference on Energy Planning, Energy Saving, Environmental Education; 2009. p. 141-149. ISSN: 1790-5095, ISBN: 978-960-474-093-2

[9] Verma AK et al. A systemic approach to integrated E-maintenance of large engineering plants. International Journal of Automation and Computing. 2010;**7** (2):173-179

[10] Monnin M et al. Fleet historical cases for predictive maintenance. In: PREDICT 19. Nancy Cedex, France: Centre de Recherche en Automatique de Nancy (CRAN), Nancy Université, UMR 7039 CNRS-UHP-INPL, Faculté des Sciences-1er Cycle - BP239; France: Vandoeuvre-Les-Nancy Cedex; 2012

[11] Ahmadi A et al. Prerequisites for a Business-oriented Fleet Availability Assurance Program in Aviation. In: 1st International Workshop and Congress on e-Maintenance; Luleå, Sweden. 2010

[12] Bilmes J. What HMMs Can Do. UWEE Technical Report Number WEETR-2002-0003. Seattle, Washington: Department of Electrical Engineering, University of Washington; 2009

[13] Cunha A. Métodos Probabilisticos para Reconhecimento de Voz. Technical Report TR-03-04. Brasil: Instituto de Matemática Pura e Aplicada, Laboratório VISGRAF; 2003

[14] Simões A. Manutenção Condicionada às Emissões Poluentes em Autocarros Urbanos—Diagnóstico por Cadeias Escondidas de Markov [Doctor Doctoral Thesis]. Lisbon's High Institute of Technology (IST); 2011

[15] Zhou D et al. Learning from Labeled and Unlabeled Data Using Random Walks. Tuebingen, Germany: Max Planck Institute for Biological Cybernetics; 2005

[16] Seeger M. Learning with Labeled and Unlabeled Data. Edinburgh: Institute for Adaptive and Neural Computation. University of Edinburgh; 2002

[17] Simola J et al. Using hidden Markov Model to uncover processing states from eye movements in information search tasks. Cognitive Systems Research. 2008;**9**(4):237-251. Elsevier Science

[18] Simões A et al. Ecological predictive maintenance in urban fleets. In: Proceedings of the 5th IASME/WSEAS International Conference on Energy & Environment (EE'10); 2010. pp. 46-53. ISSN: 1790-5095 384, ISBN: 978-960-474-159-5