Int. J. of Thermodynamics
ISSN 1301-9724

Vol. 12 (No. 1), pp. 9-16, March 2009
www.icatweb.org/journal.htm

Quantitative Causality Analysis for the Diagnosis of Energy Systems

Sergio Usón[1]*, Antonio Valero and Luis Correas

Centre of Research for Energy Resources and Consumptions (CIRCE), University of Zaragoza
María de Luna 3, 50018 Zaragoza, Spain
[1]E-mail: suson@unizar.es; Phone: +34 976 76 25 82; Fax: +34 976 73 20 78

Abstract

The aim of Thermoeconomic diagnosis is to identify malfunctioning components in energy systems and to quantify the impact on fuel consumption caused by each of them. The quantitative causality analysis is a method based on the linearization of the thermodynamic model of the system. The accuracy of the method is demonstrated by quantifying the error produced in the diagnosis of a large amount of points of a real example. Additionally, a formulation to combine this approach with thermoeconomic analysis has been developed. It allows to precisely quantify intrinsic and induced effects. Finally, the method is compared with two diagnosis approaches based on linear regression and neural networks.

Keywords: Thermoeconomic diagnosis, malfunctions, linear regression, neural networks.

1. Introduction

The term *diagnosis* refers to the activities aimed at the detection of anomalies in any system. In general, these anomalies are of interest because they can lead, if not solved, to failures. In energy systems, there are also other types of anomalies which are of interest. These situations very rarely result in a catastrophic failure but may entail huge economic expenses due to the additional fuel consumption they cause.

In this context, it is interesting not only to detect these anomalies but also to quantify the effect of each one of them in the variation of fuel consumption. This quantification is of great importance when a decision of component repairing or substitution should be taken. This task is the objective of *thermoeconomic diagnosis.*

1.1. Fuel impact formula and malfunctions

Thermoeconomics provides tools to deal with the diagnosis problem. The main tool is the fuel impact formula, which relates the variation of unit exergy consumptions of the plant components to the fuel increment they cause. It was suggested by Valero et al. (1990, 1999) and developed by Reini (1994), Lozano et al. (1994) and Torres et al. (1999).

$$\Delta F_T = \sum_{i=1}^{n}\left(\sum_{j=0}^{n} k^*_{P,j}\left(\mathbf{x}^1\right)\Delta\kappa_{ji}P_i\left(\mathbf{x}^0\right) + k^*_{P,i}\left(\mathbf{x}^1\right)\Delta\varpi_i\right) \tag{1}$$

where ΔF_T is the increment of fuel entering the plant from operation point $\mathbf{x}^1$ to reference point $\mathbf{x}^0, k^*_{P,j}$ is the unit exergy cost of the product of component $j, \Delta\kappa_{ji}$ is the variation of the unit exergy consumption of component i from component j, P_i is the product of component i and ω_i is the part of the product of the plant coming from the component i. When a unit exergy consumption κ_{ji} (the amount of product of component j entering component i divided into the product of component i) increases, the irreversibility in component i also increases in a quantity which is called *malfunction*.

$$MF_{ji} = \Delta\kappa_{ji}P_i\left(\mathbf{x}^0\right) \tag{2}$$

$$MF_i = \Delta k_i P_i\left(\mathbf{x}^0\right) = \sum_{j=0}^{n} MF_{ji} \tag{3}$$

This malfunction causes an additional amount of fuel called *malfunction cost.*

$$MF^*_{ji} = k^*_{P,j}\left(\mathbf{x}^1\right)MF_{ji} \tag{4}$$

$$MF^*_i = \sum_{j=0}^{n} MF^*_{ji} \tag{5}$$

If MF^*_0 is defined as:

$$MF^*_0 = \sum_{i=1}^{n} k^*_{P,i}\left(\mathbf{x}^1\right)\Delta\varpi_i \tag{6}$$

it can be proved that:

$$\Delta F_T = \sum_{i=0}^{n} MF^*_i \tag{7}$$

1.2. The problem of induced malfunctions

The main drawback of the application of Thermoeconomics for diagnosis is the problem of *induced malfunctions*, which appear because unit exergy costs are not always the true independent variables characterizing the components' behaviour. Due to the importance of this question, the TADEUS initiative was launched in order to serve as a common forum and test bench for researchers interested in this topic (Valero et al. 2004a,b).

The filtration of effects induced by the control system is the origin of the thermoeconomic diagnosis methodology developed by Verda (2004). Reini and Taccani (2004)

*Corresponding Author

propose a method based on the calculation of the cost of malfunction induced by the product variation.

Lazzaretto and Toffolo (2006) have made a critical review of thermoeconomic diagnosis methodologies, and point out that thermoeconomic variables are not enough to clearly distinguish intrinsic and induced effects, and they should be combined with the use of thermodynamic variables. These authors analyse the use of several indicators to identify the component where the anomalies take place (Toffolo and Lazzaretto 2004), and determine that a suitable indicator is the irreversibility of the component corrected by using the variation of local thermodynamic variables (properties of flows entering and exiting the component). The approach is applied to the TADEUS problem.

A thermodynamic model and Thermoeconomic analysis had been previously combined to determine intrinsic and induced malfunctions in a method developed by Valero et al. (1999). They propose to use a simulator of the thermal system, in order to determine the effect of the variation of an operating parameter x_r on an unit exergy consumption:

$$\Delta\kappa_{ij}^{r} = \kappa_{ij}\left(\mathbf{x}^0 + \Delta x_r\right) - \kappa_{ij}\left(\mathbf{x}^0\right) \tag{8}$$

The application of the previous equation allows one to decompose malfunctions into two terms, intrinsic and induced, depending on whether the operating parameters are associated with the studied component or not. It should be noted that these operating parameters are not only associated with the analyzed component, but also parameters characterizing the behavior of all the components of the plant. The approach has been successfully applied to a steam cycle of a power plant (Valero et al. 1999).

Other authors consider that, for an actual application in an operating power plant, these filtering techniques are not yet reliable enough, and propose to apply methods which avoid the use of Thermoeconomics. Zaleta and Muñoz (2004) propose to use directly a simulator. Correas (2004) has developed a diagnosis algorithm based on a thermodynamic description of the thermal system, which avoids the use of a fine-tuned simulator. In addition to the TADEUS problem, this methodology has been successfully applied to a combined cycle (García-Peña et al. 2001) and to a conventional coal fired power plant (Usón et al. 2006). This algorithm is the origin of the formulation proposed here. However, the quantitative causality analysis goes beyond by including, among other features, a direct connection with thermoeconomic analysis and a quantification of the effect of measurement errors (Usón et al. 2007).

1.3. Content of this paper

After presenting the formulation of the diagnosis methodology, a working example is used to demonstrate that the error produced by the method when applied to an actual power plant is very small. Afterwards, a formulation to connect this approach with the fuel impact formula is developed. This theory is a systematic approach to quantify intrinsic and induced malfunctions. Finally, two alternative methods based on the use of linear regression and neural networks are applied to the case of work in order to compare results.

To illustrate the capability of the quantitative causality analysis and to compare it with the approaches based on linear regression and neural networks, a large amount of plant data is needed. These operating points correspond to the Teruel 3x350 MW pulverised coal-fired power plant, located in Andorra, in the Teruel province (Spain) and owned by Endesa Generacion S.A. The repeated diagnosis along time (anamnesis) of this plant has been presented in a previous paper (Usón et al. 2006), where more details of the plant and the diagnosis system can be seen.

2. Formulation of quantitative causality analysis

A thermal system described by a set of n_t thermodynamic variables ($\mathbf{x}$) is considered. These variables are sufficient to characterize the matter and energy flows (pressures, temperatures, flow rates, compositions) and indicators of components behaviour, such as isentropic efficiencies for turbines or effectiveness for heat exchangers. Only n_d of these variables are actually independent (the *free diagnosis variables*): ambient conditions, fuel quality variables (composition and heating value), set points and indicators of components efficiency. In the working example, 47 free diagnosis variables have been considered. Finally, there is also a global efficiency indicator e. It should be noted that it is possible to consider more than one global efficiency indicator by repeating the procedure. For example, in the working example, three indicators are used: boiler efficiency, cycle heat rate and unit heat rate.

The goal of the diagnosis procedure is to compare two operation points (actual and reference) characterized by two variable vectors ($\mathbf{x^1}$ and $\mathbf{x^0}$) in order to divide the variation of the global efficiency indicator into a summation of terms, each one due to a free diagnosis variable:

$$\Delta e = e^1 - e^0 \cong \sum_{q=1}^{n_d} I_q \tag{9}$$

where I_q is the impact caused by the free diagnosis variable q. It should be noted that the actual and reference states of the plant are known; in other words, both vectors $\mathbf{x^1}$ and $\mathbf{x^0}$ should have been calculated prior to the diagnosis procedure by using a ‘performance test code’ (real operation) or a simulator (fictitious states of the plant).

The first step to achieve this decomposition is the introduction of the problem restrictions. There are a set of n_r restrictions which are common to the problems of diagnosis, performance test, simulation and optimization:

$$\mathbf{R}\left(\mathbf{x}\right) = \mathbf{0} \tag{10}$$

Examples of these n_r restrictions can be matter and energy balances or definitions of components’ parameters. It should be noted that values or expressions for the parameters characterizing the behaviour of the components (e.g. isentropic efficiencies of turbines or effectiveness of heat exchangers) are not needed. This is a strong difference with the development of a simulator, which reduces substantially the time needed for the development of a diagnosis system.

For small deviations, this set of restrictions can be expanded in a first order Taylor series:

$$\mathbf{R}\left(\mathbf{x}^1\right) \cong \mathbf{R}\left(\mathbf{x}^0\right) + \mathbf{JR}\left(\mathbf{x}\right)\Big|_{\mathbf{x}^0} \cdot \Delta\mathbf{x} \tag{11}$$

where $\mathbf{JR}$ is the Jacobian matrix corresponding to restrictions $\mathbf{R}$ and $\Delta\mathbf{x} = \mathbf{x}^1 - \mathbf{x}^0$. Since both $\mathbf{x}^0$ and $\mathbf{x}^1$ are solutions of the system, they accomplish the restrictions. Accordingly:

$$\begin{bmatrix} \mathbf{JR} \\ \mathbf{VD} \end{bmatrix} \cdot \Delta\mathbf{x} = \mathbf{JD} \cdot \Delta\mathbf{x} \cong \begin{bmatrix} \mathbf{0} \\ \Delta\mathbf{x}_d \end{bmatrix} \quad (12)$$

where **JD** is a n_t x n_t matrix including **JR** and **VD**. **VD** is a n_d x n_t matrix defined as:

$VD_{qm} = 1$ if the mth element of **x** is the qth element of $\mathbf{x_d}$.

$VD_{qm} = 0$ in other case.

By inverting the **JD** matrix, the variation of **x** is related to the variation of $\mathbf{x}_d$.

$$\Delta\mathbf{x} \cong \mathbf{JD}^{-1} \cdot \begin{bmatrix} \mathbf{0} \\ \Delta\mathbf{x}_d \end{bmatrix} = \mathbf{JD}^{-1} \cdot \begin{bmatrix} \mathbf{0} \\ \mathbf{U} \end{bmatrix} \cdot \Delta\mathbf{x}_d \quad (13)$$

The global efficiency indicator has a definition which relates it with some (or all) of the n_t variables. For example, the efficiency of a power plant is calculated by dividing the power produced into the product of fuel flow rate times fuel heating value. Accordingly, the increment of e can be expressed by using a Taylor series again:

$$\Delta e \cong \sum_{m=1}^{n_t} \frac{\partial e}{\partial x_m}\Big|_{\mathbf{x}^0} \cdot \Delta x_m = \mathbf{ex}^t \cdot \Delta\mathbf{x} \quad (14)$$

Finally, the variation of e can be related to the variation of the free diagnosis variables:

$$\Delta e \cong \mathbf{ex}^t \cdot \mathbf{JD}^{-1} \cdot \begin{bmatrix} \mathbf{0} \\ \mathbf{U} \end{bmatrix} \cdot \Delta\mathbf{x}_d = \mathbf{ed}^t \cdot \Delta\mathbf{x}_d \quad (15)$$

3. Analysis of non-linearities

As it can be seen in section 2, the proposed diagnosis method is based on the linearization of the restrictions of the problem and of the definition of the global efficiency indicator. In order to reduce uncertainty when non-linear behaviour appears, it is proposed to use not the derivatives in the reference state but the average value of reference and actual state: matrix **JD** and vector **ex** are evaluated in points $\mathbf{x^1}$ and $\mathbf{x^0}$ and then average values are calculated, which in turn are used in the procedure explained above to calculate **ed**. In other words, **ed** is calculated individually for each diagnosis example (variable **ed**). The strategy provides fairly good results when applied to real systems, as is demonstrated below.

The main consequence of the error caused by the linearization of the equations is a difference between the variation of the global efficiency indicator and the summation of impacts. If this fact is considered, Eqn. (9) becomes:

$$\Delta e = e^1 - e^0 = \sum_{q=1}^{n_d} I_q + \varepsilon_e \quad (16)$$

where ε_e is the error caused by linearization. If this term is small enough, the method is suitable. However, it is more interesting to use a relative value of the error, which can be obtained by dividing it into the maximum impact.

$$\varepsilon_{e,rel} = \frac{\varepsilon_e}{\max\left(I_q\right)_{q=1..n_d}} \quad (17)$$

To demonstrate that these relative errors are small enough, 4570 examples corresponding to 6 years of operation of the three units of the working example have been considered.

Figure 1 shows a histogram with the distribution of the relative residual term of cycle heat rate when two techniques are applied: i) by using quantitative causality analysis, which allows one to adapt the values of vector **ed** for each example (variable **ed**) and ii) by using an average value of **ed** vector (constant **ed**), calculated by considering the 4570 points. Figures 2 and 3 represent the same distribution for boiler efficiency and unit heat rate.

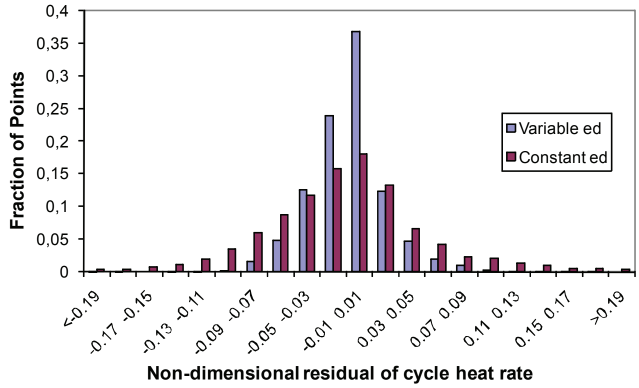

Figure 1. Distribution of non-dimensional residual of cycle heat rate.

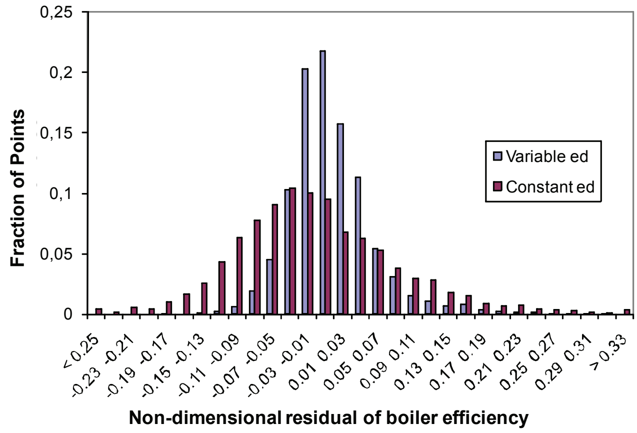

Figure 2: Distribution of non-dimensional residual of boiler efficiency.

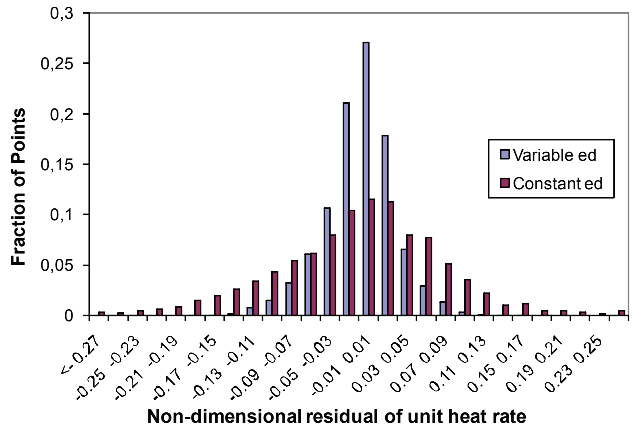

Figure 3: Distribution of non-dimensional residual of unit heat rate.

The presented bar charts show that the proposed diagnosis method provides a precision of ±3% in addressing the source of inefficiency for about 70% of the cases. This value decreases to less than 50% if constant impact factors are used. These results are very important because they prove that the method proposed presents a good compromise between accuracy and implementation cost. If a simulator were used, linearization errors would have disappeared, which would provide an unappreciable accuracy increment entailing a substantial increment in computation and implementation cost. On the other hand, accuracy decreases with the use of constant values.

4. Connection of quantitative causality analysis and the fuel impact formula

The fuel impact formula was developed to relate the variation of the exergy resources entering the plant with the variations of the independent variables of the thermoeconomic model (unit exergy consumptions and final products). Since unit exergy consumptions are not always the true variables characterizing the components' behaviour, the fuel impact formula sometimes disregards induced effects.

On the other hand, quantitative causality analysis is able to relate a set of independent variables to a user-defined efficiency indicator. However, the flexibility in being adapted to the real thermodynamic behaviour of the system makes it impossible to obtain analytical formulas, so numerical calculations are used.

The idea proposed here is to link both approaches by considering the independent variables of the thermoeconomic analysis (unit exergy consumptions and plant product) as the dependent efficiency indicators of the quantitative causality approach. This allows one to determine the value of both intrinsic and induced effects.

4.1. From free diagnosis variables to unit exergy cost and final product

In Section 2, it has been seen how, given a thermal system described by a set of thermodynamic variables $\boldsymbol{x}$, it is possible to decompose the variation of a global efficiency indicator e into a summation of terms each one corresponding to a free diagnosis variable.

The same system can be represented by a thermoeconomic model. This representation includes exergy of flows, exergy costs, unit exergy costs of flows, and unit exergy consumptions of components. However, when the fuel impact formula is applied, the only free variables are unit exergy consumptions and exergy flows leaving the component (final products). So, it is interesting to consider k_{ij} and ω_i as the indicators depending on $\mathbf{x}_d$. Each specific consumption κ_{ij} is a quotient between exergies, so that it is a function of the set of thermodynamic variables $\boldsymbol{x}$:

$$\Delta\kappa_{ij} \cong \mathbf{kx}^{ij,t} \cdot \Delta\mathbf{x} = \sum_{m=1}^{n_t} kx_{ij}^{m} \cdot \Delta x_m \tag{18}$$

where $\mathbf{kx}^{ij}$ is an n_t x 1 vector containing the partial derivatives of the unit exergy cost k_{ij} related to the set of n_t variables of the thermodynamic model of the system. It should be noted that these derivatives can be calculated numerically. By using the quantitative causality analysis (Eqn. (15)), the variation of the unit exergy consumption can be finally related to the variation of the free diagnosis variables:

$$\Delta\kappa_{ij} \cong \mathbf{kx}^{ij,t} \cdot \mathbf{JD}^{-1} \cdot \begin{bmatrix} \mathbf{0} \\ \mathbf{U} \end{bmatrix} \cdot \Delta\mathbf{x}_d = \sum_{q=1}^{n_q} kd_{ij}^{q} \cdot \Delta x_{d,q} \tag{19}$$

Each element kd_{ij}^{q} indicates the variation of the unit exergy consumption κ_{ij} when the value of the free diagnosis variable $x_{d,q}$ is increased of one unit.

A similar development can be made for each one of the products of the plant:

$$\Delta\omega_i \cong \mathbf{wx}^{i,t} \cdot \Delta\mathbf{x} \tag{20}$$

$$\Delta\omega_i \cong \mathbf{wx}^{i,t} \cdot \mathbf{JD}^{-1} \cdot \begin{bmatrix} \mathbf{0} \\ \mathbf{U} \end{bmatrix} \cdot \Delta\mathbf{x}_d = \sum_{q=1}^{n_d} wd_i^{q} \cdot \Delta x_{d,q} = \sum_{q=1}^{n_d} \Delta\omega_i^{q} \tag{21}$$

where each element wd_i^q indicates the variation of the plant product ω_i when the value of the free diagnosis variable $x_{d,q}$ is increased one unit.

The equations presented above allow one to relate the independent variables of the thermoeconomic model of a system (unit exergy consumptions of their components and final products) to the free diagnosis variables. This result is very useful for determining the effects induced by the variation of ambient conditions, fuel quality and set points and by anomalies appearing in other components.

4.2. Quantification of intrinsic and induced malfunctions

If the decomposition of $\Delta\kappa_{ij}$ and $\Delta\omega_i$ into elements corresponding to the free diagnosis variables is substituted in the fuel impact formula, it yields:

$$\Delta F_T \cong \sum_{i=1}^{n}\left[\sum_{j=0}^{n} k_{P,j}^{*}\left(\mathbf{x}^1\right)\cdot\left(\sum_{q=1}^{n_d}\Delta\kappa_{ji}^{q}\right)\cdot P_i\left(\mathbf{x}^0\right)\right] + \sum_{i=1}^{n}\left[k_{P,i}^{*}\left(\mathbf{x}^1\right)\cdot\sum_{q=1}^{n_d}\omega_i^{q}\right] \tag{22}$$

In the previous equation, the impact in fuel due to each variation of unit exergy consumption and each final product variation is decomposed into impacts due the free diagnosis variables. It can be rearranged by introducing elements kd and wd:

$$\Delta F_T \cong \sum_{q=1}^{n_d}\left[\sum_{i=1}^{n}\left(\sum_{j=0}^{n} k_{P,j}^{*}\cdot kd_{ji}^{q}\cdot P_i + k_{P,i}^{*}\cdot wd_i^{q}\right)\right]\cdot\Delta x_{d,q} \tag{23}$$

This equation shows how the fuel increment can be shared into several components, each one due to an independent free diagnosis variable.

Malfunctions can also be decomposed into a summation of several terms, each one corresponding to an impact due to a free diagnosis variable:

$$MF_{ji} = \Delta\kappa_{ji}\cdot P_i\left(\mathbf{x}^0\right) \cong \sum_{q=1}^{n_d}\Delta\kappa_{ji}^{q}\cdot P_i\left(\mathbf{x}^0\right) = \sum_{q=1}^{n_d} MF_{ji}^{q} \tag{24}$$

$$MF_i = \sum_{j=0}^{n} MF_{ji} \cong \sum_{j=0}^{n}\sum_{q=1}^{n_d}\Delta\kappa_{ji}^{q}\cdot P_i\left(\mathbf{x}^0\right) = \sum_{q=1}^{n_d} MF_i^{q} \tag{25}$$

where MF_{ji}^{q} and MF_{i}^{q} are respectively the contribution of the q[th] free diagnosis variable to MF_{ji} and MF_{i}:

$$MF_{ji}^{q} = \Delta\kappa_{ji}^{q} \cdot P_i\left(\mathbf{x}^0\right) \tag{26}$$

$$MF_{i}^{q} = \sum_{j=0}^{n} \Delta\kappa_{ji}^{q} \cdot P_i\left(\mathbf{x}^0\right) = \sum_{j=0}^{n} MF_{ji}^{q} \tag{27}$$

The malfunction cost can also be decomposed into the summation of several components, each one due to a variation of a free diagnosis variable.

$$MF_{ji}^{*} = k_{P,j}^{*}\left(\mathbf{x}^1\right) \cdot MF_{ji} \cong \sum_{q=1}^{n_d} k_{P,j}^{*}\left(\mathbf{x}^1\right) \cdot \Delta\kappa_{ji}^{q} \cdot P_i\left(\mathbf{x}^0\right) = \sum_{q=1}^{n_d} MF_{ji}^{*,q} \tag{28}$$

$$MF_{i}^{*} = \sum_{j=0}^{n} MF_{ji}^{*} \cong \sum_{j=0}^{n}\sum_{q=1}^{n_d} k_{P,j}^{*}\left(\mathbf{x}^1\right) \cdot \Delta\kappa_{ji}^{q} \cdot P_i\left(\mathbf{x}^0\right) = \sum_{q=1}^{n_d} MF_{i}^{*,q} \tag{29}$$

where $MF_{ji}^{*,q}$ and $MF_{i}^{*,q}$ are the contributions of $\Delta x_{d,q}$ to MF_{ji}^{*} and MF_{i}^{*}:

$$MF_{ji}^{*,q} = k_{P,j}^{*}\left(\mathbf{x}^1\right) \cdot \Delta\kappa_{ji}^{q} \cdot P_i\left(\mathbf{x}^0\right) \tag{30}$$

$$MF_{i}^{*,q} = \sum_{j=0}^{n} k_{P,j}^{*}\left(\mathbf{x}^1\right) \cdot \Delta\kappa_{ji}^{q} \cdot P_i\left(\mathbf{x}^0\right) = \sum_{j=0}^{n} MF_{ji}^{*,q} \tag{31}$$

Finally, the malfunction cost due to the variation of the product MF_0^{*} can be divided into impacts due to the variation of free diagnosis variables:

$$MF_{0}^{*} = \sum_{i=1}^{n} k_{P,i}^{*}\left(\mathbf{x}^1\right) \cdot \Delta\omega_i \cong \sum_{i=1}^{n}\sum_{q=1}^{n_d} k_{P,i}^{*}\left(\mathbf{x}^1\right) \cdot \Delta\omega_i^{q} = \sum_{q=1}^{n_d} MF_{0}^{*,q} \tag{32}$$

where $MF_0^{*,q}$ is the contribution of $x_{d,q}$ to MF_0^{*}.

The impact in fuel is the summation of the malfunction costs of all the components, included MF_0^{*}:

$$\Delta F_T = \sum_{i=0}^{n} MF_i^{*} \cong \sum_{i=0}^{n}\sum_{q=1}^{n_d} MF_i^{*,q} = \sum_{q=1}^{n_d} MF^{*,q} \tag{33}$$

where $MF^{*,q}$ is the fuel consumption variation due to the free diagnosis variable $x_{d,q}$:

$$MF^{*,q} = \sum_{i=0}^{n} MF_i^{*,q} \tag{34}$$

As it is shown by Eqn (33), the impact in fuel can be obtained by a double summation of the cost of malfunctions in each component due to each free diagnosis variable. Accordingly, $MF_i^{*,q}$ elements can be represented in a table, where each column corresponds to a free diagnosis variable $x_{d,q}$ and each row to a component. The summation of elements in a column is $MF^{*,q}$ and the summation of elements in a row equals approximately to MF_i^{*}. Finally, the total summation corresponds approximately to ΔF_T. It should be noted that these approximate values are not exact because of the linearization made in the quantitative causality analysis. This table can be named *table of malfunctions induced by the free diagnosis variables* (MFD, Table 1).

The previous table is very interesting because it includes the influence of each one of the n_d free diagnosis variables on each one of the n plant components and, finally, on the variation of fuel consumption. However, in a real example, the number of free diagnosis can be quite high. It may be interesting not to perform a detailed analysis including all the interdependences but to simplify the results by grouping the elements in each row in five groups:

1) Intrinsic (int): free diagnosis variables corresponding to the component with the same number as the row.

2) Induced by other components (oc): free diagnosis variables describing the behaviour of other components.

3) Induced by ambient conditions (ac).

4) Induced by fuel quality (fq).

5) Induced by Set points (sp).

Results of this grouping are summarized in a *table of intrinsic and induced malfunctions* (MFI, Table 2).

Table 1:Table of malfunctions induced by free diagnosis variables (MFD).

$MF_0^{*,1}$	$MF_0^{*,2}$	…	MF_0^{*,n_d}	MF_0^{*}
$MF_1^{*,1}$	$MF_1^{*,2}$	…	MF_1^{*,n_d}	MF_1^{*}
…	…	…	…	…
$MF_n^{*,1}$	$MF_n^{*,2}$	…	MF_n^{*,n_d}	MF_n^{*}
$MF^{*,1}$	$MF^{*,2}$	…	MF^{*,n_d}	ΔF_T

Table 2: Table of intrinsic and induced malfunctions (MFI).

$MF_0^{*,int}$	$MF_0^{*,oc}$	$MF_0^{*,ac}$	$MF_0^{*,fq}$	$MF_0^{*,sp}$	MF_0^{*}
$MF_1^{*,int}$	$MF_1^{*,oc}$	$MF_1^{*,ac}$	$MF_1^{*,fq}$	$MF_1^{*,sp}$	MF_1^{*}
…	…	…	…	…	…
$MF_n^{*,int}$	$MF_n^{*,oc}$	$MF_n^{*,ac}$	$MF_n^{*,fq}$	$MF_n^{*,sp}$	MF_n^{*}
$MF^{*,int}$	$MF^{*,oc}$	$MF^{*,ac}$	$MF^{*,fq}$	$MF^{*,sp}$	ΔF_T

5. Comparison of quantitative causality analysis with linear regression and neural networks

In some cases, when quite a big amount of operation data is available and high precision is not required, it may be interesting to try to avoid the use of quantitative causality analysis. It is possible to use linear regression in order to calculate average values of impact factors ($\mathbf{ed}_{av}$).

A difficulty may be caused by the presence of collinearity. This problem appears when the correlation factor (R_{qs}) among one or more pairs of variables has a value near to one, and is quite common when a large numbers of variables are present.

$$R_{qs} = \frac{\sum_{t=1}^{n_p}\left(x_q^t - \bar{x}_q\right)\cdot\left(x_s^t - \bar{x}_s\right)}{\left[\sum_{t=1}^{n_p}\left(x_q^t - \bar{x}_q\right)^2 \cdot \sum_{t=1}^{n_p}\left(x_s^t - \bar{x}_s\right)^2\right]^{1/2}} \tag{35}$$

To solve this problem, a variable change is performed in order to eliminate the most important interdependences between free diagnosis variables. If two variables x_q and x_s belonging to $\mathbf{x_d}$ are correlated:

$$x_q \cong a_{q0} + a_{qj} \cdot x_s \tag{36}$$

The first of them is substituted by:

$$\hat{x}_q = x_q - a_{q0} - a_{qs} \cdot x_s \tag{37}$$

Then, linear regression is applied to the new sets of variables. Finally, inverse variable change is used in order to obtain the impact factors corresponding to the initial variables, which is the objective of the procedure. This approach was applied for this test case when a correlation factor (R_{qs}) above 0.6 was found.

Another option may be to apply neural networks instead of linear regression. The idea is to use a large amount of plant data to train a neural network whose inputs are the free diagnosis variables and the output is the global efficiency indicator. Three neural networks are needed in this test case, one for each of the three global efficiency indicators.

The three networks have the same architecture. There are 47 inputs for the free diagnosis variables and only one output. The number of neurons in the hidden layer has been fixed at 3, because the form of the function required is not complex. The output function selected for these neurons is sigmoid. Outputs of these 3 neurons are the inputs of the output neuron. The output function of this output neuron is linear.

The three networks have been trained by using 70% of the 4760 operation points of the case study. Since the transformation of the free diagnosis variables has provided good results in the application of linear regression, this technique has also been applied in the neural networks.

To evaluate the suitability of linear regression and neural networks for the diagnosis problem, sensitivity of both of them when the free diagnosis variables are varied is considered and compared to results provided by the quantitative causality method. The general conclusion is the same for linear regression and for neural networks: both methodologies provide good results for the variables with high impact, and accuracy decreases for low impact variables. A difference between the two approaches is the capability of neural networks to deal with non-linear behaviour. Three examples are included to illustrate these points.

Figure 4 represents the relation between the variation of the coal high heating value and the impact that this variable originates on the boiler efficiency. Families of points represent the results provided by the quantitative causality method for the 4760 tests classified into the 3 units of the power plant, while lines are obtained by simulation using linear regression and neural networks. Coal high heating value is the variable with the highest influence on boiler efficiency, so that results provided by the three methods are practically the same.

The relation between sulphur mass fraction in coal variation and its impact on boiler efficiency is plotted in Figure 5. The impact standard deviation of this variable is around 5% of that of coal HHV, so that results provided by linear regression and neural networks are less accurate.

Figure 6 represents the influence of cooling tower effectiveness on cycle heat rate. It shows the capability of the quantitative analysis and of neural networks to deal with non linear behaviour.

6. Conclusion

Starting from a diagnosis algorithm, which had proved its suitability in real examples such as a combined cycle and a coal fired power plant, a complete diagnosis methodology has been presented.

The method is based on the thermodynamic description of the system and on the linearization of the problem's restrictions. Due to this second characteristic, an error intrinsic to the method appears. Accordingly, it is a key

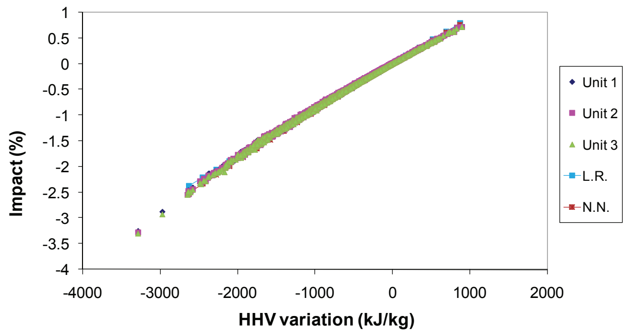

Figure 4: Coal HHV increment and its impact on boiler efficiency.

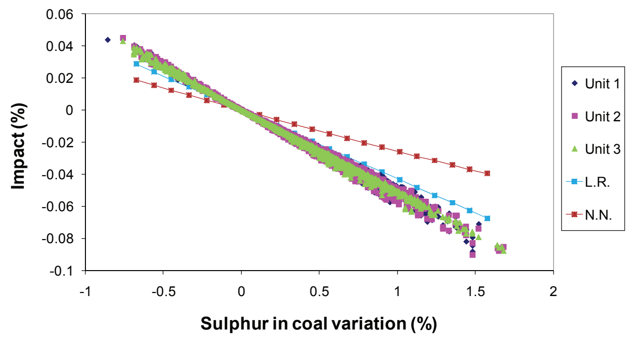

Figure 5: Sulphur in coal increment and its impact on boiler efficiency.

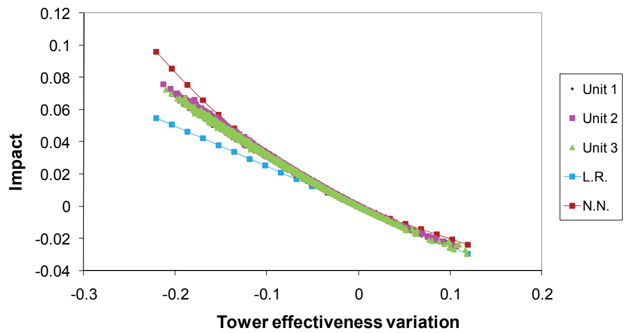

Figure 6: Cooling tower effectiveness increment and its impact on cycle heat rate.

issue to demonstrate that it is small enough. To do so, the error originating in the diagnosis of a large amount of operation periods of a working example has been analyzed. Results show that it is very small, so the method is suitable for its purpose and it is not worth to use more time consuming simulations. Additionally, an impact calculation based on the use of constant factors produces much more error. In other words, the method provides an optimum between the use of constant impact factors (low computational cost and high error) and the use of a simulator (low error but high cost of computation and implementation).

Afterwards, a connection between quantitative causality analysis and classical thermoeconomic diagnosis (fuel impact formula) has been developed. This development is able to precisely quantify intrinsic malfunctions and effects induced by other components, ambient conditions, fuel quality and set points, by using a systematic approach based on the thermodynamic restrictions of the problem. The formulation goes a step forward from the idea of using a simulator to determine the influence of variables representative of components behaviour (physical model) on the evolution of unit exergy consumption (thermoeconomic model), which was successfully applied to a steam cycle (Valero et al. 1999). Application of the method proposed here to a working example is under development.

Finally, linear regression and the use of neural networks have been presented as alternative ways to connect the variation of the free diagnosis variables with the impact they originate in the global efficiency indicator. They are not based on any description of the system, but they rely directly on the use of a large amount of operation points. Both of them provide quite good results for variables which have high impact, but error increases when impact decreases. Another important difference among the methods is that the quantitative causality analysis and neural networks are able to deal with non-linear behaviour, where linear regression does not.

Acknowledgements

The authors would like to acknowledge the personnel of Teruel Power plant (Endesa Generación), for their valuable collaboration, and Eng. Luis M. Blasco of CIRCE for developing the data treatment infrastructure. They would also like to acknowledge the Regional Government of Aragón and European Social Found for the grant for Usón (Ref. 120/2003).

Nomenclature

e	Global efficiency indicator
F_T	Fuel entering the plant [kW]
I	Impact
k^*	Unit exergy cost
MF	Malfunction [kW]
MF^*	Malfunction cost [kW]
MFD	Table of malfunctions induced by free diagnosis variables
MFI	Table of intrinsic and induced malfunctions
n	Number of components
n_d	Number of free diagnosis variables
n_r	Number of restrictions
n_t	Number of thermodynamic variables
n_z	Number of points
P	Product [kW]

Greek

Δ	Increment
ε	Error
κ	Unit exergy consumption
ω	Product of the plant

Matrices and vectors

ed	Sensitivity vector of e related to $\mathbf{x_d}$
ex	Sensitivity vector of e related to $\mathbf{x}$
JD	General restrictions matrix for the diagnosis problem
JR	Jacobian matrix of general restrictions
kd	Sensitivity vector of κ related to $\mathbf{x_d}$
kx	Sensitivity vector of κ related to $\mathbf{x}$
R	Matrix of general restrictions
U	Unit matrix
VD	Matrix of diagnosis variables
wd	Sensitivity vector of ω related to $\mathbf{x_d}$
wx	Sensitivity vector of ω related to $\mathbf{x}$
x	Vector of thermodynamic variables
$\mathbf{x_d}$	Vector of free diagnosis variables

Subscripts

av	Average
e	Error
P	Product
r	Operating parameter
rel	Relative

Superscripts

0	Reference state
1	Actual state
ac	Ambient conditions
fq	Fuel quality
int	Intrinsic
oc	Other components
r	Operating parameter
sp	Set points
t	Transpose

References

Correas, L., 2004, "On the Thermoeconomic Approach to the Diagnosis of Energy System Malfunctions. Suitability to Real-Time Monitoring", *Int. J. Thermodynamics*, vol. 7, No. 2, pp. 85-94.

García-Peña, F., Correas, L., Millán, J.L., 2001, "Improving O&M at IGCC Puertollano through Thermo-economic diagnosis", *Proceedings of PowerGen Europe 2001*, Brussels, Belgium.

Lazzaretto, A., Toffolo, A., 2006, "A Critical Review of the Thermoeconomic Diagnosis Methodologies for the Location of Causes of Malfunctions in Energy Systems", *ASME Journal of Energy Resources Technology*, Vol. 128, No. 4, (first published in 2003 as ASME IMECE2003-42688).

Lozano, M.A., Bartolome, J.L., Valero, A., Reini, M., 1994, "Thermoeconomic diagnosis of energy systems", *Proceedings of the Third Florence World Energy Research Symposium*, Florence, Italy.

Reini, M., 1994, "Analisi e Sviluppo dei Metodi Termoeconomici per lo studio degli Impianti di Conversione dell'Energia", Ph.D. dissertation, Department of Mechanical Engineering, University of Padova, Italy, [in Italian].

Reini, M., Taccani, R., 2004, "On the Thermoeconomic Approach to the Diagnosis of Energy System Malfunctions. The Role of the Fuel Impact Formula", *Int. J. Thermodynamics*, Vol. 7, No. 2, pp. 61-72.

Toffolo, A., Lazzaretto, A., 2004, "On the Thermoeconomic Approach to the Diagnosis of Energy System Malfunctions. Application of a New Indicator for the Location of Causes", *Int. J. Thermodynamics*, Vol. 7, No. 2, pp. 41-49.

Torres, C., Valero, A., Serra, L., 1999, “Structural theory and thermoeconomic diagnosis-Part I: on malfunction and dysfunction analysis”, *Proceedings of ECOS 99*, Tokyo, Japan.

Usón, S., Correas, L., Valero, A., 2006, “Anamnesis for Improving Thermoeconomic Diagnosis. The case of a 3x350 MW Coal-fired Power Plant”, *Proceedings of ESDA 2006*, Turin, Italy.

Usón, S., Valero, A., Correas, L., Verda, V. 2007, “Quantitative causality analysis for the diagnosis of energy systems”, *Proceedings of ECOS 2007*, Padova, Italy.

Valero, A., Torres, C., Lozano, M.A., 1990, “On causality in organized energy systems: Part II: symbolic exergoeconomics; Part III: Theory of perturbations”, *Proceedings of FLOWERS 90*. Florence, Italy.

Valero, A., Torres, C., Lerch, F., 1999, “Structural theory and thermoeconomic diagnosis, Part III: Intrinsic and induced malfunctions”, *Proceedings of ECOS’99*, Tokyo, Japan.

Valero, A., Correas, L., Zaleta, A., Lazzaretto, A., Verda, V., Reini, M., Rangel, V., 2004a, “On the Thermoeconomic Approach to the Diagnosis of Energy System Malfunctions. Part 1 The TADEUS Problem”, *Energy*, Vol. 29, No. 12-15, pp. 1875-1887.

Valero, A., Correas, L., Zaleta, A., Lazzaretto, A., Verda, V., Reini, M., Rangel, V., 2004b, “On the Thermoeconomic Approach to the Diagnosis of Energy System Malfunctions. Part 2 Malfunction Definitions and Assessment”, *Energy*, Vol. 29, No. 12-15, pp. 1889-1907.

Verda, V., 2004, “On the Thermoeconomic Approach to the Diagnosis of Energy System Malfunctions. From Diagnosis to Prognosis”, *Int. J. Thermodynamics*, Vol. 7, No. 2, pp. 73-83.

Zaleta, A., Muñoz, A. G., 2004, “A Reconciliation Method based on a Module Simulator. An Approach to the Diagnosis of Energy System Malfunctions”, *Int. J. Thermodynamics*,Vol. 7, No. 2, pp 51-60.