



**COLLANA DEL  
DIPARTIMENTO DI ECONOMIA**

**OBJECT-ORIENTED BAYESIAN NETWORKS FOR COMPLEX  
QUALITY MANAGEMENT PROBLEMS**

Flaminia Musella - Paola Vicard

**ISSN 2279-6916** Working papers

(Dipartimento di Economia Università degli studi Roma Tre) (online)

---

Working Paper n° 174, 2013

I Working Papers del Dipartimento di Economia svolgono la funzione di divulgare tempestivamente, in forma definitiva o provvisoria, i risultati di ricerche scientifiche originali. La loro pubblicazione è soggetta all'approvazione del Comitato Scientifico.

Per ciascuna pubblicazione vengono soddisfatti gli obblighi previsti dall'art. 1 del D.L.L. 31.8.1945, n. 660 e successive modifiche.

Copie della presente pubblicazione possono essere richieste alla Redazione.

**esemplare fuori commercio  
ai sensi della legge 14 aprile 2004 n.106**

**REDAZIONE:**

Dipartimento di Economia  
Università degli Studi Roma Tre  
Via Silvio D'Amico, 77 - 00145 Roma  
Tel. 0039-06-57335655 fax 0039-06-57335771  
E-mail: [dip\\_eco@uniroma3.it](mailto:dip_eco@uniroma3.it)  
<http://dipeco.uniroma3.it>



**DIPARTIMENTO DI ECONOMIA**

**OBJECT-ORIENTED BAYESIAN NETWORKS FOR COMPLEX  
QUALITY MANAGEMENT PROBLEMS**

Flaminia Musella - Paola Vicard

*Comitato Scientifico:*

*Fabrizio De Filippis*

*Anna Giunta*

*Paolo Lazzara*

*Loretta Mastroeni*

*Silvia Terzi*

# Object-Oriented Bayesian Networks for complex quality management problems

Flaminia Musella

Paola Vicard

Sapienza Università di Roma

Università Roma Tre

## Abstract

Quality management and customer satisfaction evaluation can be difficult tasks to perform when processes involve multiple production lines or provide multichannel services. As a consequence, the top management needs to check the quality from different perspectives and to evaluate the improvement strategies at several levels. To this aim, we propose to use Object-Oriented Bayesian Networks by which different quality aspects and evaluations can be integrated in a unique framework allowing to analyse improvement strategies in real time. We show, by an application to an internal-customer satisfaction survey, how to combine several areas of satisfaction and how to evaluate the impact on the global quality of improvement actions developed in one or more areas.

**Key words:** Bayesian networks, service quality, structural learning

**JEL Classification:** C110, M300

## 1 Introduction

Nowadays goods and service production processes are increasingly complex. Complexity may be due to several different reasons depending on the features of the product/service and on the organisational structure of enter-

prises/administrations. As a consequence, monitoring and managing quality become difficult tasks.

As far as manufacturing production is concerned, the quality of products is a mix of evolving items belonging to different stages of the production process managed by different owners. Moreover, the production lines may be spread in various regions of the same country or in different countries. In order to perform quality management it is necessary to analyse both the different production lines separately and to have a clear and exhaustive picture of the overall quality of the whole production process. Only in this case the top-management can at the same time check the consistency of the final products with the declared standards and laws and propose and implement appropriate improvement strategies accounting for the expertise and the suggestions of each line specific process owner.

As far as services are concerned, in the last decades the proportion of public and private services offered via different channels (face-to-face, internet, mail) has been continuously increasing. When service quality or customer satisfaction is analysed, the different channel-specific results must be evaluated both singularly – accounting for the production process complexity – and as contributions to the overall quality/satisfaction.

When there are many actors involved in the decision making phase (because of *e.g.* supranational geographical size or multichannel delivery), top management responsibility is two-fold:

- to supervise strategies developed by every agent;
- to correctly coordinate strategies in order to define priorities.

In fact, for each channel or production line, the process owner usually has at his/her disposal a multiple set of evolving satisfaction elements contributing simultaneously and differently to determine the perceived quality. The management is interested in finding relations among these features and between them and the perceived quality. In order to fulfill the second task, it is required a multi-object tool capable to combine the various sources of information, *i.e.* the statistical models relative to each specific production area.

Recently, Bayesian networks have been proposed as a tool for managing information coming from customer satisfaction surveys (Renzi *et al.* 2009; Salini and Kenett 2009; Tarantola *et al.* 2012). Bayesian networks, thanks to their easy-to-interpret graphical representation, facilitate communication between managers and statisticians. Moreover, they are a very useful statistical tool to perform what-if analysis; a variety of scenarios can be constructed to evaluate the impact of improvement actions. In this paper we propose the use of Object-Oriented Bayesian networks (Koller and Pfeffer 1997) (from now on OOBNs). Through an application to an Italian public administration, we show that OOBNs are a suitable tool to support complex decision problems when more process owners are involved. Each decision maker can study his/her improvement strategy with respect to his/her production area, both checking the effect on the global quality and accounting for decisions in the other areas.

The paper is organized as follows. In Section 2 basics on Object-Oriented Bayesian networks together with their use in the quality management framework are introduced. Section 3 provides an application of the tool to internal-customer satisfaction real data. A variety of scenarios are shown and analysed in Section 4. Conclusions and open issues are addressed in Section 5.

## **2 Object-Oriented Bayesian network for quality management**

A Bayesian network (Cowell *et al.* 1999) is a multivariate statistical model satisfying sets of (conditional) independence statements portrayed by a directed acyclic graph (DAG). A DAG is composed of a set of nodes and a set of arrows between pairs of nodes. Each node represents a random variable and is associated with its probability distribution. The distribution is marginal for nodes having no incoming edges and is conditional for other nodes. A missing arrow between two nodes implies (conditional) independence between the corresponding variables. A BN is provided with an inference engine allowing to efficiently make inference on the model parameters.

OOBNs represent an extension of BNs by which complex and high di-

mensional problems can be modelled in terms of inter-related objects. They provide a hierarchical definition of a BN by means of building blocks (classes of BNs). An OOBN is thus a multi-instance network made of standard nodes, representing random variables, and instance nodes (objects) being BNs or OOBNs themselves. From now on nodes are indicated in **teletype** face and classes of networks in **bold** face.

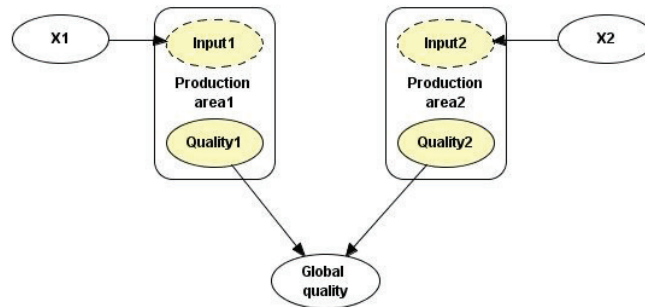


Figure 1: An OOBN example to model the global quality by considering two production areas.

The OOBN in Figure 1 models a simple process constituted by two production areas, represented by instance nodes (rounded rectangles) of class **Production area1** and **Production area2** respectively. Each instance, being a BN itself, describes the production process in a specific area and has the quality of that area as output variable. Therefore nodes **Quality1** and **Quality2** are output nodes (full line) carrying information out of their respective instances to determine (appropriately combined) the overall process quality represented by the standard node **Global quality**. The standard nodes **X1** and **X2** may represent top-management standard/constraints the local managers have to account for in organising their area production process. The arrow from node **X1** (a standard node) to the input node **Input1** is an identity link representing the fact that the top-management decision is transferred into the **Production area1**. Each area manager can decide supported by his/her area network; in fact each instance (and all its internal standard nodes) can be accessed by expanding the corresponding objects. Figure 2 shows the expansion of **Production area1** where interface nodes are characterised by an outer grey ring with a dashed line (input nodes) or

a full line (output nodes). Notice that in the example in Figure 1 the two instances are not connected; this means that the two production areas are marginally independent but become dependent when analysed conditionally on Global quality.

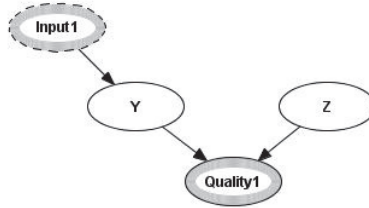


Figure 2: Expansion of the instance of class **Production area1** in Figure 1.

It is remarkable that the inference process of OOBNs is particularly efficient in spite of the complex structure since the standard nodes inside an instance are conditionally independent from the rest of the OOBN, given the interface nodes of that instance (Koller and Pfeffer 1997, for more details).

OOBNs are a suitable tool for representing large and complex domains. Thanks to modularity, increasing model refinements can be easily handled by properly adding new building blocks. Therefore, for example, production processes with more than two production areas can be modelled simply adding an instance for each area to the network in Figure 1. Notice that in this case the computational complexity grows linearly with the number of additional objects. Interdependence between production areas can be dealt with as well by appropriately linking the interface nodes of the corresponding objects.

In the literature, there are many applications of OOBNs to several fields, such as forensic statistics (Dawid *et al.* 2007) and economic decisions (Mortera *et al.* 2013). A first attempt to apply OOBNs to customer satisfaction is in Musella *et al.* (2007).

In this paper we show how OOBNs can be used to develop a flexible and easy-to-use decision support system for complex managerial problems by an application to internal customer satisfaction data from an Italian public administration.

### 3 An application to an Italian public administration

Recently an Italian public administration carried out a survey to evaluate the quality of its internal services. The staff-satisfaction survey dealt with delivered automation and information technology services relative to mail, internet and workplace care areas. These service areas belong to the information technology department but are supervised by different managers. Thus service quality must be evaluated both specifically for each area and globally for the department.

The sample is constituted by 203 units (middle managers and employees). Data have been collected by a questionnaire organised in four sections, one for each of three investigation areas and one for profiling. The variables have been measured using a five points scale where 1 denotes the lowest level and 5 the highest one, *i.e.* the worst and the best judgement respectively for satisfaction related variables. In most cases the states 1 and 5 were never or very rarely chosen by the respondents. Hence levels 1 and 2 and levels 4 and 5 have been merged so that three states are left, *e.g.* for the variables measuring satisfaction: 1 (for poorly satisfied), 2 (for moderately satisfied) and 3 (for very satisfied). This has been done to avoid table sparsity problems affecting the constraint based method used to learn the network structure.

The socio-demographic aspects surveyed, being neither directly nor indirectly significantly connected to any of the other variables of interest, are not represented in our model.

The three investigated areas are independently managed by different actors but their quality performances directly affect the overall department quality judgement; therefore the three service areas are assumed to be marginally independent but they are dependent conditionally on the overall evaluation.

The overall model, by which the top management can monitor the global quality, has been built (1) by learning a Bayesian network for each service area separately and (2) by integrating the resulting networks in a single OOBN. Our analysis has been carried out using the software HUGIN ([www.hugin.com](http://www.hugin.com)).

### 3.1 Learning and using the networks for mail, internet and workplace care areas

Each network structure has been estimated separately from data using a constraint-based algorithm, namely NPC (Necessary Path Condition) (Steck 2001). The NPC algorithm is suitable for recursively testing marginal and conditional association between categorical variables. The algorithm admits to consider logical constraints, such as for example ban/assignment of a specific direction. In addition, the NPC algorithm allows the user to choose among independence equivalent models the one most suitable for the analysed problem. This allows to integrate the automatic learning procedure and the subject-matter knowledge.

Once the BN model has been estimated, efficient information propagation techniques (Cowell *et al.* 1999) can be exploited in order to assess the effect of changes in the status of specific variables on the behaviour of the remaining variables. In this way, various improvement scenarios can be simulated and evaluated. To this aim it is sufficient to insert the evidence and propagate it throughout the network so that in a mouse-click time the effect of possible managerial decisions can be read off the network.

We next analyse the three area network construction separately.

#### 3.1.1 Mail area

The first section of the questionnaire deals with the quality of the mail service and with the employees' satisfaction about this. Several specific items are included in this section. The items resulting to be directly or indirectly connected to mail service satisfaction are listed in Table 1 (together with their corresponding node names and states).

Figure 3 shows the estimated network for the mail area. For each node, the monitor shows the states and the marginal probability distribution (expressed in percentage) of the corresponding variable.

The relation structure entails that the **Mail satisfaction** is directly influenced by the **Information quality** and the **Easy of use**, while it is indirectly affected by the other variables. The marginal probability table of **Mail satisfaction** shows that the staff is moderately or very satisfied

Variable name	Node name	States
<b>Satisfaction related variables:</b>		
Availability of the mail service	Availability	1 (poorly) 2 (moderately) 3 (very)
Easy of use of the service	Easy of use	1 (poorly) 2 (moderately) 3 (very)
Information quality about mail service	Information quality	1 (poorly) 2 (moderately) 3 (very)
Mail service overall satisfaction	Mail satisfaction	1 (poorly) 2 (moderately) 3 (very)
Usefulness of mail service	Usefulness	1 (poorly) 2 (moderately) 3 (very)
<b>Other variables:</b>		
Frequency of use of the service	Frequency of use	1 (rarely or sometimes), 2 (often), 3 (always)
Learning channel	Learning channel	1 (course), 2 (colleagues), 3 (self-educated)

Table 1: Variables of mail network and their states.

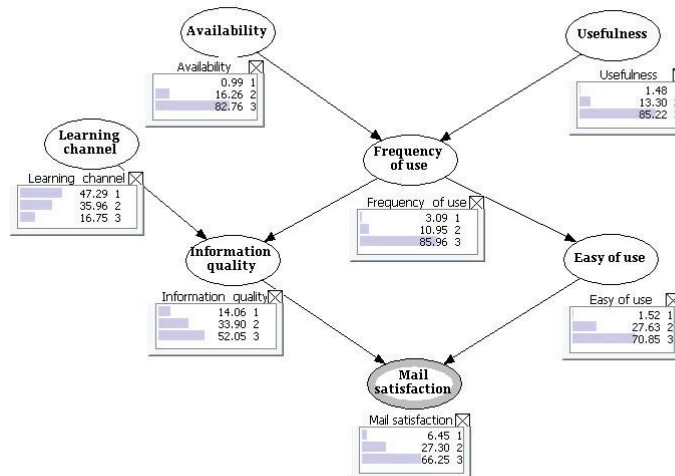


Figure 3: Bayesian network for the mail area showing the marginal probabilities, expressed in percentage, of the variables.

(levels 2 and 3) with probability 0.9355. In particular in 66.25% of the cases the employees are very satisfied. Nevertheless this result can be improved by, for example, refining the quality of information about mail service usage.

Figure 4 illustrates the scenario where information quality is never judged poorly satisfying; the probability that the mail satisfaction is evaluated with 3 rises from 0.6625 to 0.7048. This probability rises even more significantly

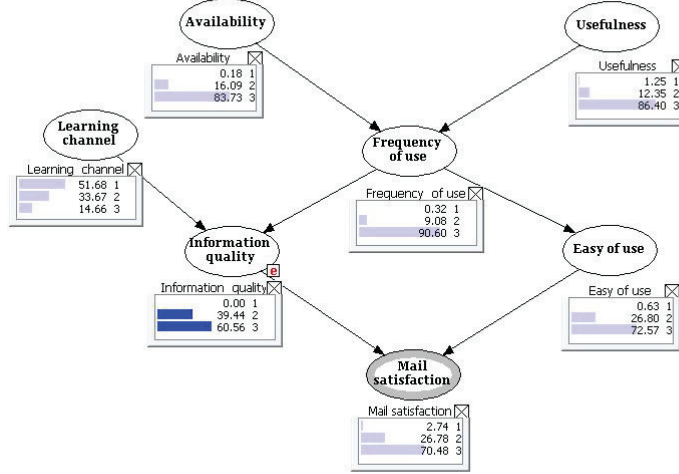


Figure 4: Marginal probabilities of mail area network when Information quality is never poorly satisfied.

(Figure 5) reaching 0.7558 when all the staff is completely satisfied with information quality (*i.e.* having inserted the evidence `Information quality=3`).

### 3.1.2 Internet area

The second section of the questionnaire is devoted to evaluate the quality of the internet service. In detail, the items and questions that, after the structural learning, are directly or indirectly connected to the internet service satisfaction are reported in Table 2.

Internet is not accessible to the whole staff; in particular, out of the 203 employees, only 163 have internet access. Therefore, the dataset comprises two groups of cases: 40 records with no answers on internet area items and 163 fully observed individuals. To simultaneously model both groups of respondents, a filter binary variable has been introduced. This variable, named *access*, takes value 1 when internet access is allowed and 0 when it is denied. As a consequence structural zeros in the corresponding contingency tables are generated, implying dependence among the filter and all the variables following it. Therefore structural learning has been performed by imposing

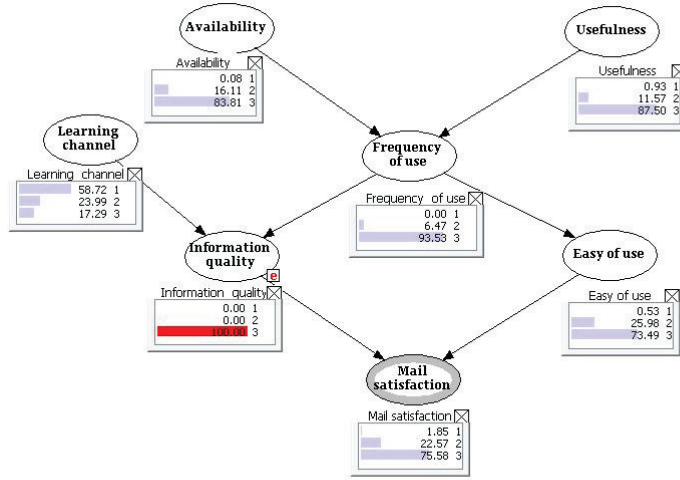


Figure 5: Marginal probabilities of mail area network when Information quality=3.

Variable name	Node name	States
<b>Satisfaction related variables:</b>		
Availability of the connection	Availability	0 (no answer) 1(poorly) 2 (moderately) 3 (very)
Easy of use of internet service	Easy of use	0 (no answer) 1(poorly) 2 (moderately) 3 (very)
Information quality about internet service	Information quality	0 (no answer) 1(poorly) 2 (moderately) 3 (very)
Internet service overall satisfaction	Internet satisfaction	0 (no answer) 1(poorly) 2 (moderately) 3 (very)
Usefulness of internet service	Usefulness	0 (no answer) 1(poorly) 2 (moderately) 3 (very)
<b>Other variables:</b>		
Access to the service	Access	0 (denied), 1 (allowed)
Frequency of use of the service	Frequency of use	0 (no answer), 1 (rarely or sometimes), 2 (often), 3 (always)
Interest in further knowledge	Further knowledge	0 (no answer), 1 (little or not), 2 (interested), 3 (very)
Learning channel	Learning channel	0 (no answer), 1 (course), 2 (colleagues), 3 (self-educated)

Table 2: Variables of internet network and their states.

the presence of a directed edge from node **Access** to all the other variables of the internet area. Figure 6 shows the estimated network together with the marginal probability tables (in percentage) in the case of allowed internet

access (*i.e.* having inserted the evidence  $\text{Access}=1$ ).

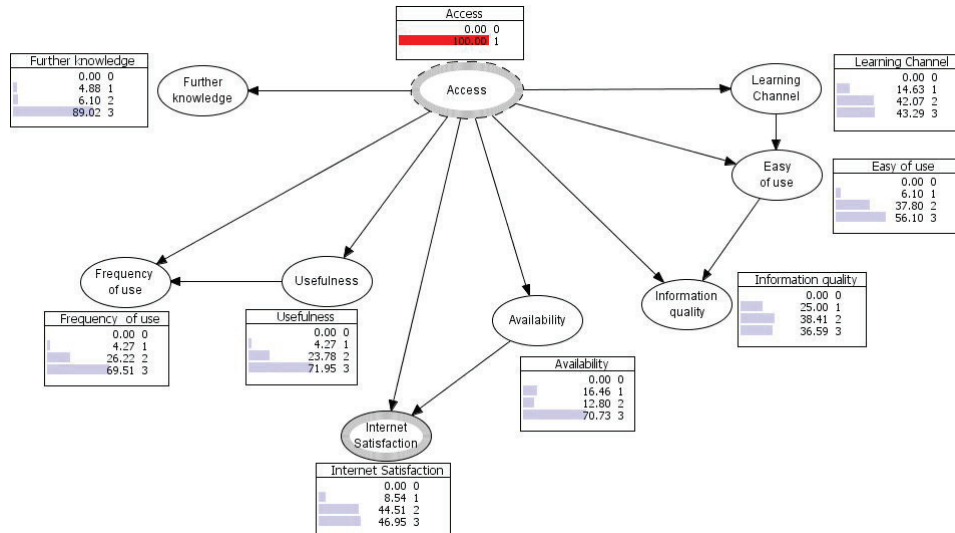


Figure 6: Bayesian network for the internet area showing the marginal probabilities, expressed in percentage, of the variables, when  $\text{Access}=1$ .

The network in Figure 6 shows that the satisfaction about internet assistance service, apart from the filter variable, is directly influenced solely by the availability of internet connection during the day (**Availability**). Moreover, the probability that an employee evaluates with 3 the assistance service is less than a half (specifically, 0.4695). If the management decides to invest in the network system, rising the connection availability to the maximum level (*i.e.*  $\text{Availability}=3$ ), then the probability of the highest internet satisfaction grade amounts to 0.5517 (see Figure 7).

### 3.1.3 Workplace care area

The third section of the questionnaire deals with the quality of the workplace care service and measures the corresponding employees satisfaction. Table 3 shows the variables present in the final network for this area.

The learnt network (Figure 8) shows that the **Workplace care satisfaction** is directly influenced by the **Courtesy** and the **Timeliness** of the workplace care operators and by the **Information quality**. The other variables affect the satisfaction indirectly.

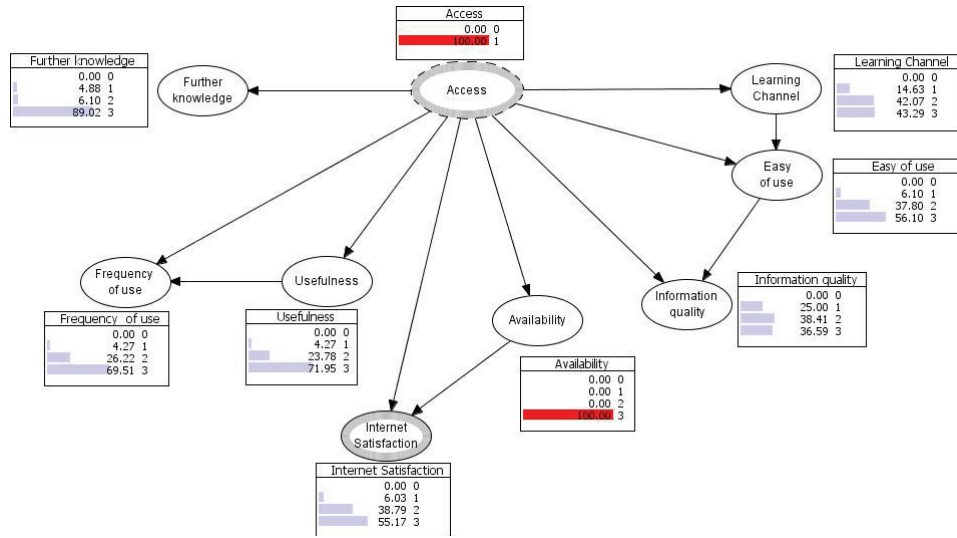


Figure 7: Marginal probabilities of internet area network when  $\text{Access}=1$  and  $\text{Availability}=3$ .

Variable name	Node name	States
<b>Satisfaction related variables:</b>		
Information quality about the service	Information quality	1(poorly) 2(moderately) 3(very)
Staff courtesy	Courtesy	1(poorly) 2(moderately) 3(very)
Staff reliability	Reliability	1(poorly) 2(moderately) 3(very)
Timeliness	Timeliness	1(poorly) 2(moderately) 3(very)
Usefulness of workplace care service	Usefulness	1(poorly) 2(moderately) 3(very)
Workplace care service overall satisfaction	Workplace care satisfaction	1(poorly) 2(moderately) 3(very)

Table 3: Variables of workplace care network and their states.

As shown in Figure 8, in most of the cases (94.43%) the employees are moderately (2) or very (3) satisfied, the probability of observing a grade 3 of satisfaction being equal to 0.5179. Among the variables directly influencing Workplace care satisfaction, Information quality shows the largest improvement margins (the probability of the highest mark is 0.4187, while the corresponding probabilities for timeliness and for staff courtesy are 0.7389 and 0.7918 respectively). We can easily simulate a scenario (see Figure 9) where the management performs an improvement action on information quality so that all the employees give mark 3 (*i.e.* set evidence  $\text{Information quality}=3$ ). In such case the probability of grade 3 for workplace care sat-

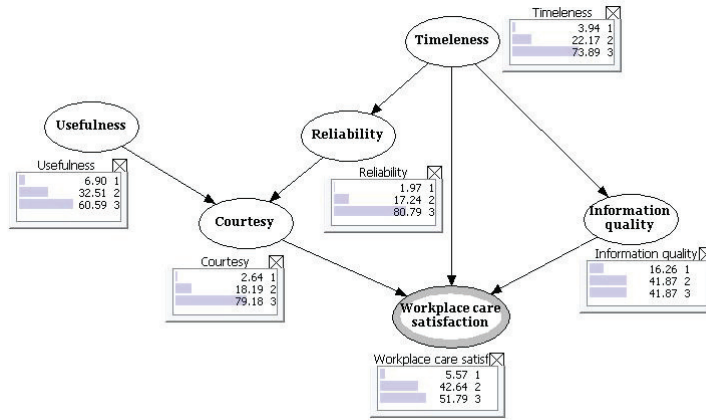


Figure 8: Bayesian network for the workplace care area showing the marginal probabilities, expressed in percent, of the variables.

isfaction rises to 0.7023.

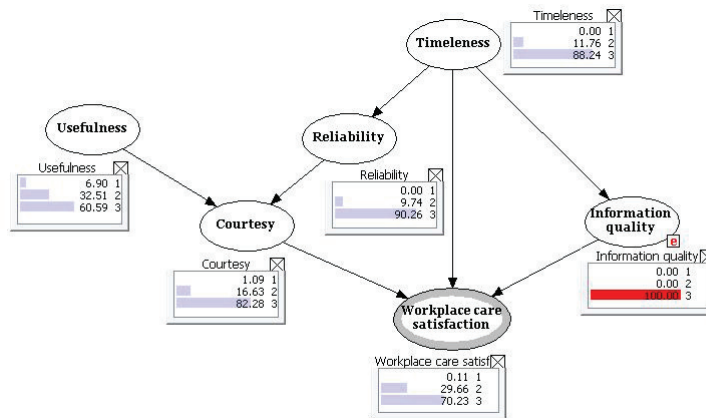


Figure 9: Marginal probabilities of workplace care area network when Information quality=3.

### 3.2 The model for global quality

Once a BN has been learnt for each service area, we need to combine their informative contributions into a single global quality indicator. The overall problem, accounting for both the area specific components and the global

quality indicator, is modelled by an OOBN where each area is represented by an instance node of its corresponding class of network, *i.e.* **Mail area**, **Internet area** and **Workplace care area**. This model can be used to strategically evaluate the effect that a variation in a single area determines in the whole system.

### 3.2.1 The global service quality

If not directly measured, an indicator of global service quality can be obtained by properly combining the satisfaction levels of each area. Although there is not a unique approach in service quality measurement, the concept of service quality indicator defined by Peña (1997) is largely shared in the literature. Below we briefly introduce this approach.

Consider a sample of  $n$  customers who express their opinion on  $p$  attributes of a service. Let  $X_{i1}, X_{i2}, \dots, X_{ip}$  be the evaluations of these attributes given by the  $i$ -th customer. The corresponding perceived service quality, denoted by  $Q_i$ ,  $i = 1, \dots, n$ , is a function of the  $p$  quality aspects. A linear quality indicator is

$$Q_i = \sum_{j=1}^p w_{ij} \cdot X_{ij}$$

where, when all variables are measured using the same scale,  $0 \leq w_{ij} \leq 1$  with  $\sum_{j=1}^p w_{ij} = 1$ . The generic  $w_{ij}$ ,  $i = 1, \dots, n$ , and  $j = 1, \dots, p$  denotes the weight of the  $j$ -th attribute for the  $i$ -th customer. Weights are assumed to be independent from the attributes. Moreover,  $w_{ij}$  can be interpreted as a measure of the relative importance, with respect to quality, of the aspect  $j$  for the  $i$ -th customer.

In the present framework, the global quality indicator represents a synthesis of  $p$  specific satisfaction measurements, denoted by  $S_j$ ,  $j = 1, \dots, p$ . Weights are assumed to be constant over respondents. In detail, denoting by  $S_{ij}$  the satisfaction measurement for unit  $i$ ,  $i = 1, \dots, n$ , about feature  $j$ ,  $j = 1, \dots, p$ , and by  $w_j$  the corresponding weight, the global quality indicator  $Q_i$  is defined as

$$Q_i = \sum_{j=1}^p w_j \cdot S_{ij} \tag{1}$$

with  $0 \leq w_j \leq 1$  and  $\sum_{j=1}^p w_j = 1$ .

In our application,  $p = 3$  and  $Q_i = \sum_{j=1}^3 w_j \cdot S_{ij}$  where:

- $S_{i1}$  is the mail overall satisfaction for the the  $i$ -th employee;
- $S_{i2}$  is the internet overall satisfaction for the the  $i$ -th employee;
- $S_{i3}$  is the workplace care overall satisfaction for the the  $i$ -th employee;
- $w_1$  is the weight of mail overall satisfaction;
- $w_2$  is the weight of internet overall satisfaction;
- $w_3$  is the weight of workplace care overall satisfaction.

Generally speaking, weights can be estimated from data or fixed by experts. In this paper, we consider both approaches also accounting for the internet access issue (allowed/denied).

### 3.2.2 The OOBN for global quality

According to previous considerations and having defined the global quality indicator formula, the overall OOBN model (Figure 10) is made of:

- an instance node for each class network **Mail area**, **Internet area** and **Workplace care area**, derived in Sections 3.1.1 3.1.2 and 3.1.3. The three instance nodes communicate with the node **Global quality indicator** by means of their respective output nodes, **Mail satisfaction**, **Internet satisfaction** and **Workplace care satisfaction**.
- the node **Access**, associated to the filter variable and splitting the sample in the units having and those not having internet access. This node directly contributes to the global quality indicator definition and is part of **Internet area** network, as shown in Figure 6. For information consistency, the arrow between the standard node **Access** and **Internet area** is an identity link.
- the **Global quality indicator** that summarises the weighted contribution of every area according to the expression (1). The global quality

index assumes three possible levels: low (coded by 1), medium (coded by 2) and high (coded by 3).

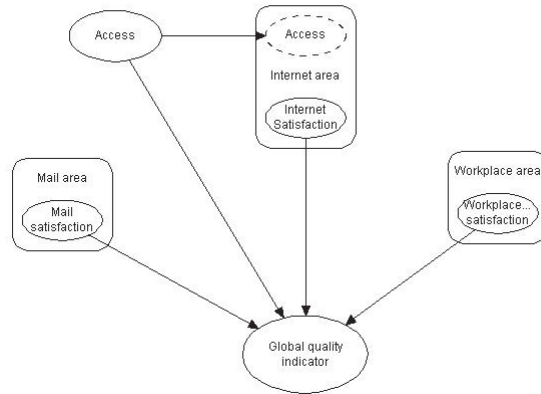


Figure 10: OOBN combining satisfaction areas.

The network in Figure 10 allows managers to easily monitor the global quality accounting for information coming from all the different service areas suitably weighted. Therefore, weights determination is a crucial phase. It can be performed either by using some appropriate statistical techniques or on the basis of expert opinions reflecting management policies or theoretical assumptions (Nardo 2005). Both methodologies are reasonable in principle and we consider both of them. Therefore we have:

- weights predetermined by management accounting for managerial strategies. They are defined by management for all three areas; hence, when internet access is denied,  $w_2 = 0$  and the other weights,  $w_1$  and  $w_3$ , are normalised so that they sum to 1.
- estimated weights. In this case, weights have been estimated separately for the employees having and for those not having the internet access. The estimation is performed by factor analysis according to the approach of Nicoletti *et al.* (2000). The method consists in (1) carrying out a factor analysis and (2) defining the weights from the matrix of factor loadings. In detail, weights are given by the proportion of the unit variance of the indicator explained by manifest variables.

Area	Preassigned weights		Estimated weights	
	Access allowed	Access denied	Access allowed	Access denied
Mail	0.40	0.44	0.37	0.50
Internet	0.10	-	0.33	-
Workplace-care	0.50	0.56	0.30	0.50

Table 4: Preassigned and estimated weights of the three service areas.

Table 4 shows the weights determined in both approaches accounting also for the case of denied internet access. In order to represent in a single network both sets of weights (preassigned and estimated), the model in Figure 10 has been modified by the addition of a new node, **Weights**, with two possible states: 1 (preassigned) and 2 (estimated). The resulting network is shown in Figure 11.

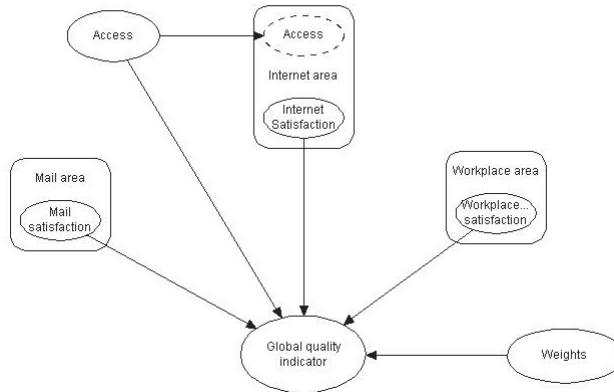


Figure 11: Refinement of the network in Figure 10 allowing for the representation of different sets of weights.

## 4 What-if analysis

The OOBN can be used to simulate or evaluate the impact on the global quality indicator of improvement actions proposed or applied in one or more information technology department areas. Moreover, having specified two

different sets of weights, we can analyse the sensitivity of the global quality indicator to them.

Consider the case where weights are estimated and the internet access is allowed ( $\text{Weights}=1$ ,  $\text{Access}=1$ ). The marginal probabilities, without any additional evidence, of the global model in Figure 11 are displayed in Figure 12(A). Suppose an improvement action is performed in the mail area clearing the probability of being poorly satisfied with this service. As a consequence (Figure 12(B)) the probability associated to the lowest grade of global quality ( $\text{Global quality}=1$ ) decreases from 0.0049 to 0.0014; the probability associated to the medium grade of global quality ( $\text{Global quality}=2$ ) decreases from 0.3901 to 0.3518; the probability associated to the highest grade of global quality ( $\text{Global quality}=3$ ) increases from 0.6050 to 0.6468.

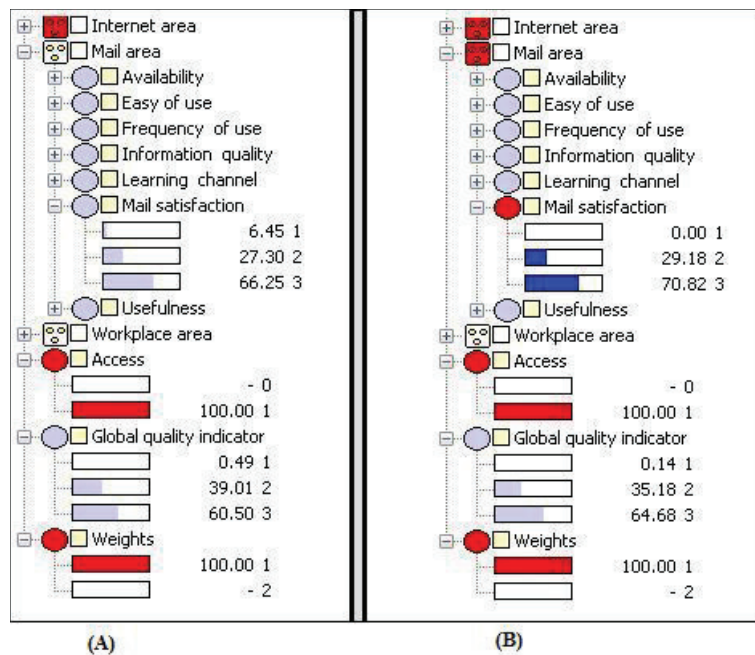


Figure 12: Marginal probabilities in the global model with preassigned weights and access to the internet before (A) and after (B) an improvement action in the mail area.

Many other different scenarios can be analysed. Here we consider the following ones.

**Scenario a.** This represents the baseline when no evidence is provided. Only

Global quality	Access=1					Access=0				
	a	b	c	d	e	a	b	c	d	e
Low (1)	0.49	0.19	0.45	0.00	0.00	1.88	0.89	-	0.04	0.00
Medium (2)	39.01	32.96	36.37	26.72	18.18	49.67	48.50	-	34.27	31.35
High (3)	60.50	66.85	63.18	73.28	81.82	48.45	50.61	-	65.69	68.65

Table 5: Global quality probability table in percentage before (**a**) and after improvement actions **b-e** when weights are preassigned.

Global quality	Access=1					Access=0				
	a	b	c	d	e	a	b	c	d	e
Low (1)	0.84	0.32	0.62	0.56	0.14	3.11	0.13	-	0.00	0.00
Medium (2)	47.88	39.70	43.76	38.66	25.32	34.33	13.41	-	14.65	6.42
High (3)	51.28	59.98	55.62	60.78	74.54	62.56	86.47	-	85.35	93.57

Table 6: Global quality probability table in percentage before (**a**) and after improvement actions **b-e** when weights are estimated.

nodes **Access** and **Weights** are instantiated according to the case of interest.

**Scenario b.** An improvement action is performed in the *Mail area*. The information quality and the easy of use of the mail service have been set as completely satisfying (**Information quality=3**; **Easy of use=3**).

**Scenario c.** The simulated action is carried out in the internet area and deals with the availability of the service that is assumed to be completely satisfying (**Availability=3**). This scenario can be analysed when **Access=1** only.

**Scenario d.** This scenario consists in a variation in the *Workplace care area*. The improvement action is simulated to rise the information quality to the maximum level of satisfaction (**Information quality=3**).

**Scenario e.** This scenario consists in actions performed in all area simultaneously. More specifically, scenarios **b**, **c** and **d** are combined, scenario **c** being added when **Access=1** only.

Tables 5 and 6 show the impact of scenarios **b-e** on the global quality when weights are assigned by management, and when weights are estimated

respectively. In both tables, scenarios are evaluated under presence/absence of internet access. Columns **a**, in italics, show the marginal probability distributions of the global quality when no evidence, apart from weights and internet access specification, is inserted.

Consider the intervention **d** on the workplace care area. From Table 5 (preassigned weights) we see that when internet access is allowed (**Access=1**) this improvement action produces a relevant change in the probability distribution of global quality: the probability of grade 3 rises from 0.6050 to 0.7328. When the weights are estimated (Table 6) the same scenario for **Access=1** does not give rise to an equally relevant increase in the probability of the maximum level of the global quality. This is due to the fact that for workplace care area the management fixed weight (0.50) is higher than the estimated one (0.30). It is also interesting to notice that in case investments on one single area are planned (scenarios **b-d**), when **Access=1** it is mostly convenient to intervene on the workplace care area (scenario **d**) having larger improvement margins both for preassigned (Table 5) and estimated (Table 6) weights. Differently, when **Access=0**, intervening on the workplace care area (scenario **d**) is the best solution only for preassigned weights (Table 5); the probability of the highest score rises from 0.4845 to 0.6569. If the estimated weights are used (Table 6) it will be slightly more convenient to invest in the mail area since scenario **b** generates an increase of the probability of the grade 3 from 0.6256 to 0.8647.

Although not surprising, it is relevant to observe that, whatever the weights determination method, the most efficient results are achieved by managing all areas simultaneously (scenario **e**).

## 4.1 Managing the quality

OoBN can help top management not only in coordinating the improvement actions in different production areas but also in calibrating them when a satisfaction level target is fixed for the global quality. In fact, when an overall objective is given, OoBNs allow to both derive the consequent achievement required for each area and have a clear picture of the best strategies to reach the target. Consider the case where weights are preassigned (**Weights=1**)

and internet access is allowed ( $\text{Access}=1$ ). The probability distribution of mail, internet and workplace satisfaction and of the global quality are shown in Figure 13.

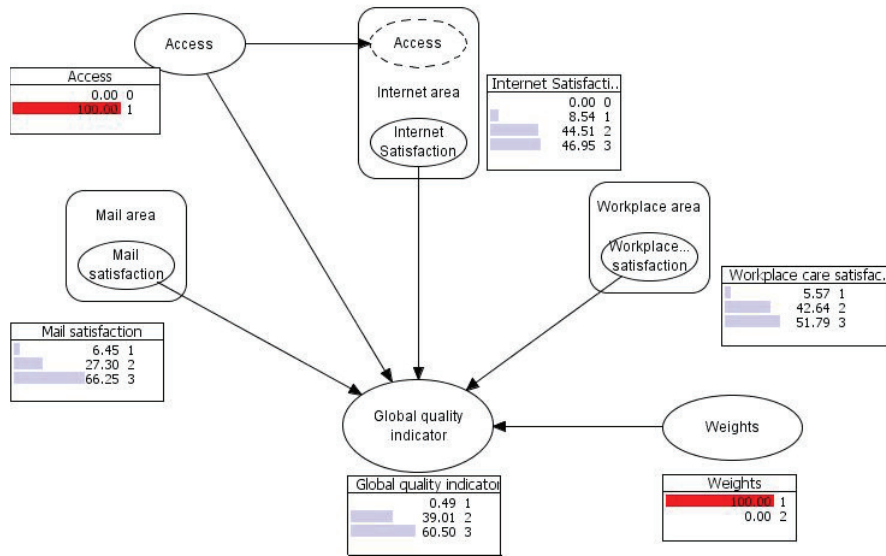


Figure 13: Network showing probability tables of overall satisfaction variables and global quality indicator when weights are preassigned ( $\text{Weights}=1$ ) and internet access is allowed ( $\text{Access}=1$ ).

Suppose that obtaining the highest level of the global quality is the overall managerial target. Inserting and propagating the evidence  $\text{Global quality}=3$  in the network in Figure 13, the probability tables of the satisfaction variables of each area are modified as shown in Figure 14.

These results support the top management in defining the specific area objectives and, in turn, each area supervisor can analyse the possible strategies for his/her area on the basis of the corresponding Bayesian network. Comparing the probability of the evaluation of level 3 in Figure 13 and 14, we observe that all areas are involved in the achievement of the overall target. More interventions are required in the workplace care area: the probability of the highest level of satisfaction has to rise from 0.5179 to 0.7808. The manager of workplace care area can obtain the required improvement by strategically intervening on (all or some) aspects affecting the workplace care overall satisfaction. Figure 15 shows the marginal probability tables of

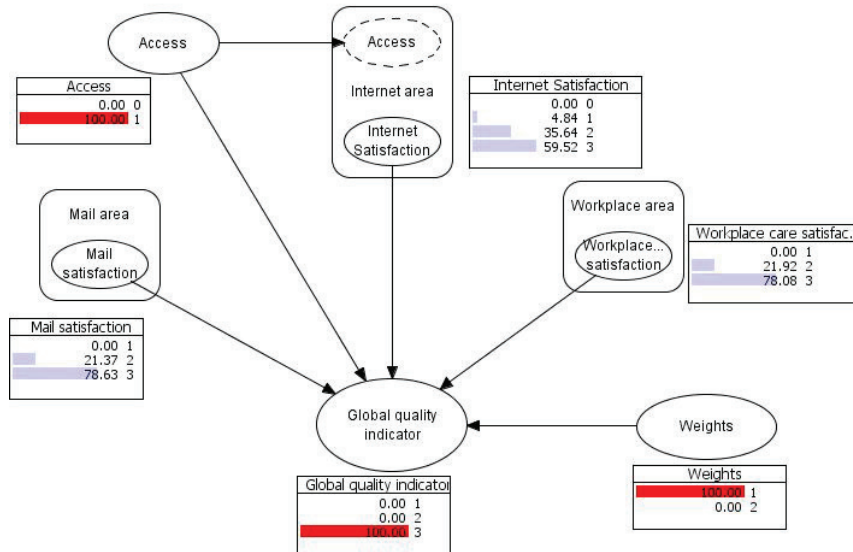


Figure 14: Network showing probability tables of overall satisfaction variables when weights are preassigned ( $Weights=1$ ), internet access is allowed ( $Access=1$ ) and an evidence on global quality indicator is provided.

every satisfaction item of workplace satisfaction area when no target is given (A) and when evidence  $Global\ quality=3$  in the top-level network is fixed (B).

By comparing tables in Figure 15 (A) and (B), it is possible to observe that a more or less important improvement should be achieved in each satisfaction aspect. The priority interventions are relative to the information quality in the first place and the timeliness since for them more significant probability tables modifications are required. In particular, the probability of the highest level of information quality has to rise from 0.4187 to 0.5072 while that of timeliness has to reach 0.8179 from 0.7389. Notice also that information quality resulted to be the variable with the largest improvement margins from the analysis in Section 3.1.3.

## 5 Conclusions and discussion

Quality management problems may be increasingly complex, the main reasons being the product features, the presence of multichannel services, the

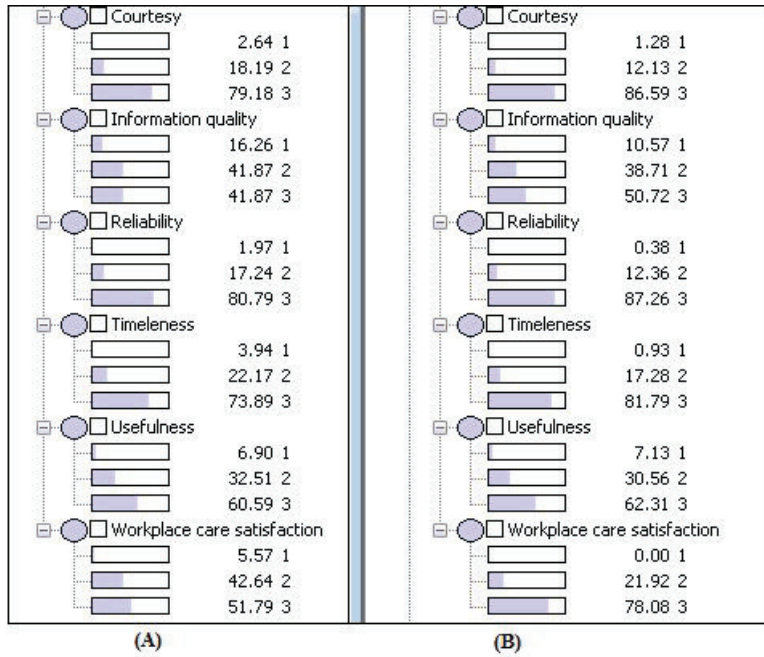


Figure 15: Marginal probability tables for variables of the workplace care satisfaction area before (A) and after (B) having fixed a global quality standard (Global quality=3).

organization of the manufacturing/delivering. In order to guarantee the declared standards, managers need to monitor the quality by analysing data coming from different sources accounting for issues relative to different goods/service production areas often managed by different owners. In this context, OOBNs represent a natural answer to managers' needs since they are able to represent large complex domain problems and allow to plan and simulate coordinated improvement actions even when there are multiple decision makers eventually at various organisational levels. By the OOBN, top management can easily both coordinate the strategies chosen at each area level and suggest the improvement actions once a global quality standard is fixed.

We have illustrated the efficacy of OOBNs in dealing with quality management problems by an application to an internal-customer satisfaction survey. Several areas of satisfaction are separately modelled and then combined in a single model for the global quality. The overall model is then used to simulate various scenarios by inserting evidence in variables belonging to one

or more service areas. In the paper, marginal independence between the different areas has been assumed. In some situations experts may suggest that the production areas are dependent; in such cases, expert knowledge can help incorporating appropriate links between them. It is also interesting to notice that additional elements (such as quality standards, external market conditions and/or regulations), contributing to determine the overall quality, can be easily included in the OOBN thanks to modularity property.

As illustrated in sections 3.2.2 and 4, the weights attributed to production/service areas play an important role for the final decision. Therefore further researches will be devoted to study and compare different statistical methodologies to determine weights.

The application of OOBNs here proposed can also be extended to all those problems where evaluations coming from different sources should be combined into a unique measure and the decision maker can be either the top management or the final user. As an example, many public and private services (such as universities, firms, banks, *etc.*) are evaluated by both national and international agencies; a consumer having to choose among them should have the possibility to base his/her decision on a single judgement obtained as a synthesis of the various indicators. In this sense OOBNs are a promising framework to model and to provide an easy-to-use support for decision.

**Acknowledgement** We thank Antonella Pennacchio, Maria Francesca Renzi and Silvia Vargiu for useful comments. We thank Caterina Conigliani for reviewing the paper.

## References

- Cowell, R. G., Dawid, A. P., Lauritzen, S. L., and Spiegelhalter, D. J. (1999). *Probabilistic Networks and Expert Systems*. Springer, New York.
- Dawid, A., Mortera, J., and Vicard, P. (2007). Object-Oriented Bayesian

- Networks for Complex Forensic DNA Profiling Problems. *Forensic Science International*, **169**, 195–205.
- Koller, D. and Pfeffer, A. (1997). Object-Oriented Bayesian Networks. In *Proc. 13th Annual Conference on Uncertainty in Artificial Intelligence, (UAI) August 1-3, 1997, Brown University, Providence, Rhode Island, USA*, pp. 302–13. Morgan Kaufmann.
- Mortera, J., Vicard, P., and Vergari, C. (2013). Object-Oriented Bayesian Networks for a Decision Support System for Antitrust Enforcement. *The Annals of Applied Statistics*. To appear.
- Musella, F., Renzi, M. F., and Vicard, P. (2007). Un Modello di Valutazione della Qualità Basato su Sistemi Esperti Probabilistici. In *Valutazione e Customer Satisfaction per la Qualità dei Servizi*, pp. 213–6. Centro stampa università, Roma – Italy.
- Nardo, M. (2005). a.o. Handbook on Constructing Composite Indicators: Methodology and User Guide. OECD Statistics Working Paper.
- Nicoletti, G., Scarpetta, S., and Boylaud, O. (2000). Summary Indicators of Product Market Regulation with an Extension to Employment Protection Legislation. OECD, Economics department working papers.
- Peña, D. (1997). Measuring Service Quality by Linear Indicators. In *Managing Service Quality*, pp. 35–51. Edited by Kunst and Lemmink, Chapman Publishing Ltd.
- Renzi, M. F., Vicard, P., Guglielmetti, R., and Musella, F. (2009). Probabilistic Expert Systems for Managing Information to Improve Services. *The TQM Journal*, **21**, 429–42.
- Salini, S. and Kenett, R. S. (2009). Bayesian Networks of Customer Satisfaction Survey Data. *J. Appl. Statist.*, **36**, 1177–89.
- Steck, H. (2001). *Constraint-Based Structural Learning in Bayesian Networks using Finite Data*. PhD thesis, Institut für Informatik der Technischen Universität München.
- Tarantola, C., Vicard, P., and Ntzoufras, I. (2012). Monitoring and Improving Greek Banking Services Using Bayesian Networks: An Analysis of Mystery Shopping Data. *Expert Systems with Applications*, **39**, 10103–11.