A Bayesian Network Model for Supporting School Managers Decisions in the Pandemic Era

Flaminia Musella1 · Paola Vicard2 · Maria Chiara De Angelis1

Accepted: 6 May 2022 / Published online: 2 June 2022 © The Author(s), under exclusive licence to Springer Nature B.V. 2022

Abstract

Due to the dramatic health situation caused by the COVID-19 pandemic, in Italy the emergency remote teaching lasted longer than in other countries. The mandatory teaching modalities have required digital transformation processes in a framework where digital-divide is one of the limitations to school modernization. We believe that the experience can promote a deeper formatting of organizational process. The paper shows results of a multitarget research carried out during the Italian lockdown aiming at animating the debate around school from multi-actors perspectives and at supporting policies. The paper aims at showing the potentiality of a multivariate statistical method as a tool supporting school managers in identifying those challenges they have to face to improve the setting up of internal processes. The main result is a model supporting the decision making process at orienting school managers strategies.

Keywords Bayesian networks · Digital-divide · Evidence sensitivity · Italian school · Organizational aspects · Value of information · What-if-analysis

1 Research framework and motivation

The starting raising of COVID-19 pandemic in Italy, as in most countries of the world, caused lockdown with the related introduction of the emergency remote teaching for schools and higher education. At the beginning, distance learning has been the unique solution for keeping students on learning for a while, and it has been still chosen in some time frames during the second phase of the emergency. UNESCO data about the number of weeks of school closures in EU, shows that Italy has a sad prime. Among

1 Link Campus University, Via del Casale di S. Pio V, 44, 00165 Rome, Italy

2 Università Roma Tre, Via S. D’Amico, 77, 00145 Rome, Italy
European countries, Italy has adopted one of the most lasting closing strategy between March and August 2020 (Fig. 1).

Moreover, by extending the focus until May 2021, Italy is still one of the European countries where the partial or total closure were most widely adopted in compulsory school (Fig. 2).

It’s a matter of fact that education, health and economy, are the main areas that globally have been suffering the most. Only during the first Italian lockdown - March-May 2020 - Italian students lost 32% of the attendance time for the compulsory education (computed on OECD report - https://www.invalsi.it/invalsi/ri/pisa2018/Country_Note.pdf). We believe that the effect of remote teaching has been amplified in Italy because of the well-known digital divide and the contrast between teachers’ digital skills criticism vs teachers’ promotion digital educational experimentation Capogna et al. (2020), splitting the reality in a double-speed system.

This is highlighted in Fig. 3 that shows the Italian areas where the percentage of families with internet access is above or below the national average. However, digital divide may be considered just the iceberg top of a deeper and more significant issue related to the organizational processes in the Italian school system. The phenomenon does acquire relevance in the pandemic era. The issue has been faced up by the “Digital technologies, education and society” (DiTES) centre of research, that is the principal investigator of this work.
In this paper, we discuss results of a web-survey conducted among all the actors directly and indirectly involved in the remote teaching. Due to the different stakeholders involved, the study has been conducted in a multitarget perspective and has reached 474 school managers, 3444 teachers, 787 students, 2116 parents, for a total of 6821 interviews. The questionnaires were submitted between May and July 2020. This paper focuses on the results of remote school teaching from the school managers' point of view. The survey has been carried out during the Italian lockdown with the help of different partners such as Roma Tre University, Italian digital revolution, Forum associazioni familiari and Associazione nazionale dei dirigenti pubblici e alte professionalità della scuola with the aim of informing the debate on post-emergency schools and, consequently, the education policies. Here, we mainly focus on the results of remote school teaching from the school managers' point of view. The main research question addressed is "what have been the challenges of school managers for improving organizational process in the school?". Here, due to the complexity of the phenomenon, the research question is addressed by a multivariate statistical model capable, on one hand, to catch the, possibly intricate, dependencies among variable and, on the other hand, to depict them by an easy-to-read graphical representation. The proposed statistical model, being associated with an inference engine, allows to probabilistically evaluate the effect of an informative shock on one or more variables that can be manipulated by the decision maker. Therefore various scenarios can be rigorously analysed and compared supporting school managers in the decision process.

The paper is organized as follows. Section 2 discusses the methodological approach highlighting both the literature review used to elaborate the tool of investigation and the statistical model chosen to address the research question. Section 3 presents the data and the sample. Section 4 shows the resulting model and discusses its accuracy; the use of the model is shown in Sect. 4.1. Conclusions are in Sect. 5 and discussions are addressed in Sect. 6.

2 Methodological Approach

In this Section we present: (i) the methodology adopted to handle the multitarget perspective: from the literature review to the investigating tool (Sect. 2.1); (ii) the statistical technique chosen to assess the research question (Sect. 2.2).
2.1 A Multitarget Perspective

The tackled phenomenon is a multiperspective issue involving many stakeholders with a different role. For this reason we developed a multitarget questionnaire that has been submitted to all the actors that, directly and indirectly, have had something to do with remote teaching. Areas of the questionnaire have been identified with the help of a literature review discussed in Sect. 2.1.1; the resulting questionnaire structure is argued in Sect. 2.1.2.

2.1.1 Literature Review

The research plant was built stressing the impact of distance learning upon organizational processes and management of e-leadership of the school managers (Barnard, 1968; Mintzberg, 1983). We define organizational resilience as the "capability of a system to maintain its function and structure in the face of internal and external changes and to degrade gracefully when it must" (Sutcliffe, 2007; Weick et al., 2005), moreover, we mainly consider the role of communication in safeguarding organizational resources in the face of complex and ambiguous situations (Weick et al. 2005). In particular, the issue of the school’s resilience was investigated considering the following:

(a) the ability to drive innovation in teaching-learning methodologies (Pitzalis et al., 2016), with a specific focus on the three dimensions of learning: cognitive, social and emotional (Illeris, 2003);
(b) the strategies to address the digital divide, which has increased the social gap between those who have access to the information technologies (personal computers, Internet, smartphones) and those who are partially or totally excluded (Hargittai, 2010; Jackson et al., 2008; Livingstone & Helsper, 2007) and
(c) the overall satisfaction of the various actors involved, in the perspective of promoting a culture of self-evaluation (Kyriakides et al. 2002) for a quality-oriented education system (Castoldi 2007; Allulli 2012).

2.1.2 The Resulting Questionnaire

According to the literature review discussed in the previous Subsection, the resulting questionnaire was organized in 5-cross sections such as: profiling, organizational aspects, teaching methodology, tools and access and satisfaction whose connection with the literature is synthesized in Table 1.

Each section was made by items equals and specific for every target respondents. The rationale of the choice was inherent in the need to investigte both common aspects involving all stakeholders and to observe specific matters by actor. Thus, the questionnaire has been developed with the help of different partners in order to provide an institutional proponent for each stakeholder as explained in Table 2.

Firstly, partners worked together for sharing common items worthy of investigation, then each partner highlighted those topics mostly relevant for the represented group of stakeholder.

In the present paper, we only focus on the school managers. By the questionnaire, many items have been observed for every section but a selection, eventually based on
| Section                  | Meaning                                      | References                                                   |
|--------------------------|----------------------------------------------|--------------------------------------------------------------|
| Organizational aspects   | Snapshot of organizing elements              | Barnard (1968), Mintzberg (1983), Sutcliffe (2007), and Weick et al. (2005) |
| Teaching methodology     | Investigation about teaching methodologies   | Pitzalis et al. (2016) and Illeris (2003)                    |
| Tools and access         | Account for tools and access limitations     | Livingstone and Helsper (2007), Hargittai (2010), and Jackson et al. (2008) |
| Satisfaction             | Collection of satisfaction judgement for the experience | Castoldi (2007), Alulli (2012), and Kyriakides et al. (2002) |
multivariate techniques of data reduction, brings to consider fewer variables in the analysis. Table 3 provides a synthesis of variables and corresponding modalities for questionnaire facing up scool managers.

The analysis has been carried out with the aim to explore the relations among variables affecting the challenges in organizational processes improvement for the school. To adress this complex issue, Bayesian networks, able to identify the independencies

| Table 2 | Partner involved in the research by actor in the remote teaching |
|---|---|
| Actor | Representative partner |
| Students | Roma Tre University (R3) |
| Parents | Forum Associazioni familiari (FAF) |
| School managers | Associazione Nazionale dei dirigenti pubblici e alte Professionalità della scuola (ANP) |
| Teachers | Italian Digital Revolution (AIDR) |

| Table 3 | Variables |
|---|---|
| Section | Item (Variable/Node) | Modalities |
| Profiling | Gender | M, F |
| Profiling | Age class | < 40; 40–50; 50–60; > 60 |
| Profiling | Geographical area | North, Center, Island, South |
| Profiling | School level | comprehensive school, other |
| Organizational aspects | Timeliness | less than a week, 1 week, 2 weeks, more than 2 weeks |
| Organizational aspects | All Matters | Yes, No |
| Organizational aspects | Involving teaching capability problems | none, few, quite, a lot |
| Organizational aspects | IT project | Yes, No |
| Organizational aspects | Guidelines | Yes, No |
| Teaching methodology | Interaction promotion | 5-point Likert |
| Teaching methodology | Staff-manager coordinator | 5-point Likert |
| Teaching methodology | Teachers’ appreciation technologies | 5-point Likert |
| Teaching methodology | Internal communication organization | 5-point Likert |
| Teaching methodology | Remote teaching evaluation | Yes, No |
| Teaching methodology | Teachers’ tutoring | Yes, No |
| Tools and access | Teachers’ access problems | Never, Sometimes, Always |
| Tools and access | Students’ access problem | Never, Sometimes, Always |
| Satisfaction | Priority redesign | 5-point Likert |
| Satisfaction | Decision process redesign | 5-point Likert |
| Satisfaction | Internal communication improvement | 5-point Likert |
| Satisfaction | School-family communication improvement | 5-point Likert |
| Satisfaction | Emerging teachers’ needs | 5-point Likert |
| Satisfaction | Teachers’ willing | 5-point Likert |
| Satisfaction | Families’ willing | 5-point Likert |
| Satisfaction | Teachers’ cooperation | 5-point Likert |
| Satisfaction | Organizing process improvement | 5-point Likert |
among variables, have been choosen. The method is shortly introduced in the next Subsection.

2.2 Modeling by Bayesian networks: structural learning and inference

A Bayesian network (BN) represents a multivariate probability distribution of a set of variables by means of a directed acyclic graph (DAG), i.e. a mathematical object made of a finite set of vertices (also called nodes) representing random variables, and a set of directed edges (arrows) displaying direct relevance of one variable to another. A directed graph is acyclic if it is not possible, starting from a node, to go back to it following arrows directions. A BN is thus a map embedding conditional independencies among variables such that the multivariate distribution can be read off the DAG by using the $d$-separation criterion Pearl (1986). For example, given the variables $x$, $y$ and $z$, the conditional independence of $x$ and $z$ given $y$ is graphically displayed by structures $x \rightarrow y \rightarrow z$, $x \leftarrow y \leftarrow z$ or $x \leftarrow y \rightarrow z$. Differently, the graphical structure $x \rightarrow y \leftarrow z$ denotes conditional dependence of $x$ and $z$ given $y$ and it is known as $v$−structure.

In a discrete BN, each node is associated with a probability table that is marginal in the case of founder nodes (those without incoming edges) and is conditional for other nodes.

When subject-matter knowledge is available, the BN can be built manually by experts; if the dependence structure is unknown, or partially known, the network has to be learnt directly from data by means of efficient algorithms (Buntine, 1994, 1996; Neapolitan, 2003). The structural learning phase consists in two steps: firstly the DAG is learnt, then parameters are estimated. Two main approaches are used for graph estimation: scoring and searching (Cooper & Herskovits, 1992; Heckerman, 1995; Chickering et al., 1994) spanning the space of all possible models and choosing the one maximising a given score function, and constraint-based iteratively checking (conditional) independences by performing statistical tests on the data. Two or more different DAGs may imply the same set of conditional independence relations. This happens when they have the same edges regardless of their direction, and the same $v$−structures. In such a case, the DAGs are said Markov equivalent (Verma & Pearl, 1991). Many algorithms move between independence equivalence classes and finally provide a graph representing the selected equivalence class. Such a graph, called completed partially directly acyclic graph (CPDAG), contains both direct and undirect edges. Specifically, a CPDAG has: a direct edge where all the DAGs in the equivalent class have the same directed edge; an undirected edge between two nodes, say variables $x$ and $y$, if in the equivalence class there is at least a DAG with $x \leftarrow y$ and a DAG with $x \rightarrow y$. Among constraint-based algorithms, the most popular procedure is the PC algorithm (Spirtes et al., 2000). It is a stepwise backward method that takes as input a database $D$ over a set of $K$ variables and provides, in output, a CPDAG.

In detail, the PC algorithm starts from a complete undirected graph and, given a chosen significance level, it performs a sequence of statistical tests, such as $\chi^2$ in case of categorical data, for deciding to keep or remove edges between pairs of nodes, conditioned by a subset of nodes. The output of this first step is a graph whose edges are undirected, called skeleton graph. In the second step of the algorithm, v-structures are found on the basis of the test results.

In the last few years, the PC algorithm has become a reference point and a benchmark for developing new constraint-based strategies (Steck, 2001; Kalisch et al., 2012; Marella et al., 2014; Musella, 2013). Here, we adopt Nominal-Ordinal PC algorithm, NOPC algorithm (Musella, 2015) that is suitable when the dataset comprises both nominal and ordinal
variables. The NOPC procedure consists of the following four steps: (i) set the variables type, nominal or ordinal; (ii) identify the skeleton of the graph by properly checking marginal and conditional independencies between pair of nodes, according to the variable typology as listed in Table 4; (iii) CPDAG identification and (iv) DAG extension.

Once the structure is learnt, the parameters estimation from data may be performed by the EM algorithm supporting possible missing values in the data. In case of networks learnt from data, goodness of fit can be evaluated by verifying the model accuracy on the basis of the receiver operating characteristic curve (ROC - Metz, 1978). The curve, plotting the true positive rate (sensitivity) against the true negative rate (1-specificity), describes an area (AUC) interpretable as an overall measurement of model performance and varying between 1 (a model with a perfect discriminant capacity) and 0.5 (a model with a null discriminant capacity). According to main interpretations, AUC values between 0.7 and 0.9 denote reasonable model performance, and values greater than 0.9 indicate strong predictive accuracy (Pearce & Ferrier, 2000; Swets, 1988).

A strenght of BNs is the underlying inference engine that allows users to propagate the information throughout the network in a mouse-click time and, consequently, to update the probability tables associated to the nodes. This functionality is largely appreciable among managers since it permits to perform a what-if-analysis, i.e. a process that allows the user to perform simulations by inserting the evidence (meaning a new input information) and by observing its impact on the probability table of other nodes (target variables). The tool is thus used to evaluate the probabilistic impact of an improvement action on a variable of interest. Moreover, by using the software Hugin, value of information analysis can be easily performed. This allows the user to quantify the mutal information between variables and, thus, to explore the most impacting variables for the aspect of interest. This may help in defining the improvement actions priority. Finally, the network can be used for invetigating the impact of different set of evidences on the response variable of the model (evidence sensitivity).

3 The Data

The overall web-survey consists of 6821 interviews, distributed as follows: 474 school managers, 787 students, 2116 parents, 3444 teachers. The survey has been carried out during the lockdown: a web-questionnaire has been administered between May and July 2020. Respondents are mainly women (74.58%) between 51 and 60 years of age (51.41%) and mostly working in public and comprehensive schools. These information are coherent with the population distribution of school managers. As far as geographical areas are concerned, Lazio and Lombardia are over-represented, whilst Campania, Calabria and Sicily are under-represented (see Fig. 4). The reason for this may be linked both to the greater coverage of the territory of central Italy and to the high level of

Table 4  Test automatically selected by the NOPC procedure according to the variable pair

| Class X     | Class Y     | CI test            |
|-------------|-------------|--------------------|
| Nominal     | Nominal     | CCE                |
| Nominal     | Ordinal     | Kruskal – Wallis   |
| Binary      | Ordinal     | Wilcoxon           |
| Ordinal     | Ordinal     | Jonckheere – Terpstra |
compliance on the part of Lombard school principals due to the impact of pandemic on their territory. However, we can not exclude that Regions in the South and Islands are less represented in the research (28% cumulatively) due the digital divide issue commented before (Fig. 3).

The e-research reached 4.4% of the population (considering the school principals employed in the school in the scholastic year 2019-2020 - Open Data MIUR). This can be considered a good response rate since it is significantly larger than that usually characterizing e-research, generally being around 1% (MacElroy, 2000). The predictive estimated model is shown in Sect. 4.

4 The Model

As said before, the paper deals with the school managers. Here, we focus only on fully observed units (238) that still represent 3% of the entire universe. The model has been estimated by the NOPC algorithm in $\text{R}$, and then it has been managed in $\text{Hugin}$. The learnt network is displayed in Fig. 5.

It’s worth noting that the variables directly affecting (parent nodes) organizing process improvement (response or hypothesis variable) are: (i) Internal communication improvement, (ii) Priority redesign, (iii) School-family communication improvement, (iv) Internal communication organization. Figure 6 shows the network with monitors containing the probability tables associated with the nodes (variables).

To verify the accuracy of the model, a sample of 250 observations has been generated according to the network dependencies. The simulated sample has been used to validate the model in predicting the organizational processes improvement by using the Receiver Operating Characteristic (ROC). The resulting area under the curve for the learnt model with respect to organizational processes improvement variable is 0.87 that, according to some interpretations (Swets, 1988, for instance), stands for a good accuracy of the model. As a consequence the model is considered good enough to be used for simulating scenarios.
4.1 Using the model

The goal of this research is to investigate the most impacting levels for improving the organizational processes in the school. The research question may be addressed by performing a what-if-analysis supervised by experts evaluation. We have had the opportunity to interview the vice-president of ANP to speak about the challenges she believed more significant. She declared to strategically act on the following areas: internal and external communication adequacy, the organization efficacy and the teachers’ reply. We compared the expert evaluation with the value of information analysis. The tool is available in Hugin software and it helps in measuring how much information is shared by a couple of variables. In fact, given a BN model and a response variable, this analysis allows the user to identify the most informative variables with respect to the goal.
the hypothesis variable. The informative power is measured in terms of mutual information \( i.e. \) the quantity of information shared by two variables, say \( X \) and \( Y \) denoted by \( I(X; Y) \). Results of the analysis are shown in Table 5. The first column lists all the variables the management can consider to affect the target variable, \textit{Organizing process improvement} denoted by \( H \). The second column shows the value of information quantities, denoted with \( I(T, H) \), where \( T \) denotes the variable on which the management can act (in the first column) and \( H \) the hypothesis variable (organizing process improvement). The reported results highlight that \textit{internal communication improvement}, \textit{priority redesign}, \textit{satisfaction DS-teachers communication} and \textit{school-family communication improvement} are relevant variables to be managed in a probabilistic approach, since they share the largest amount of information with the target variable. Moreover, three of them are those aspects identified by the expert as well, as shown in the last column of Table 5.

Next, we perform \textit{what-if-analysis} by instatiating (inserting evidence) those nodes that resulted to be crucial from the value of information analysis illustrated above.

**Table 5** Value of information results crossed by expert evaluation

| Variable                                    | Mutual information, \( I(T;H) \) | Key aspect by expert |
|---------------------------------------------|-----------------------------------|----------------------|
| Internal communication improvement          | 0.06                              | x                    |
| Satisfaction DS-teachers communication      | 0.04                              |                      |
| Priority strategies redesign                 | 0.04                              | x                    |
| School-family communication improvement     | 0.03                              | x                    |
| Decision process redesign                   | 0.03                              |                      |
| Emerging teachers’ needs                    | 7.73E-03                          | x                    |
| Staff-manager coordinator                    | 7.51E-03                          |                      |
| Teachers’ appreciating technologies         | 5.56E-03                          |                      |
| Interaction promotion                        | 4.47E-03                          |                      |
| Families’ willing                           | 2.07E-03                          |                      |
| Gender                                      | 1.39E-03                          |                      |
| All matters                                 | 9.40E-04                          |                      |
| Guidelines                                  | 3.59E-04                          |                      |
| Teachers’ cooperation                        | 2.83E-04                          |                      |
| Age class                                    | 9.24E-04                          |                      |
| Geographical area                           | 5.19E-04                          |                      |
| School level                                | 2.41E-06                          |                      |
| IT project                                  | 1.59E-06                          |                      |
| Involving teaching capability                | 0                                 | x                    |
| Remote teaching evaluation                   | 0                                 |                      |
| Students’ access problem                    | 0                                 |                      |
| Teachers’ access problem                    | 0                                 |                      |
| Teachers’ willing                           | 0                                 | x                    |
| Teachers’ tutoring                          | 0                                 | x                    |
| Timeliness                                  | 0                                 |                      |
Specifically, four possible scenarios are analysed and their impact on the response variable is evaluated.

4.1.1 First scenario based on the improvement of internal communication

A scenario instantiating the level of internal communication improvement to its maximum, makes the probability of the highest level of organizational processes improvement increase by 80% (see Fig. 7).

![Diagram](image-url)

**Fig. 7** Network and monitor windows: probability table of organizing process improvement expressed in percentage before (a) and after (b) a simulating action in internal communication improvement.
4.1.2 Second Scenario Based on the Improvement of Priority Design

A scenario setting the highest level of priority redesign at 100% produces an improvement of about 66% in the probability of the maximum level of organizing processes improvement (see Fig. 8).

Fig. 8 Network and monitor windows: probability table of organizing process improvement expressed in percentage before (a) and after (b) a simulating action in priority redesign
4.1.3 Third Scenario Based on the Improvement of School-Family Communication

A scenario developed with the aim to maximize the probability linked to the highest level of School-family communication allows to increase of 53% the probability of the highest level of organizing processes improvement (see Figure 9).

Fig. 9 Network and monitor windows: probability table of organizing process improvement expressed in percentage before (a) and after (b) a simulating action in school-family communication
4.1.4 Fourth Scenario Based on the Improvement of Satisfaction DS-Teachers Communication

A scenario instatiating the level of Satisfaction DS-teachers communication to its maximum, makes the probability of the highest level of the target node increase by about 36% (see Fig. 10).

4.1.5 Managing Multiple Variables Simultaneously

Moreover, scenarios may be built by manipulating more than one variable at a time in order to show the joint effect of simulated actions. Table 6 shows the probability table of organizing processes improvement, with the increasing—in percentage—of the probability associated to the highest level of the response variable, when an evidence is inserted in two or three variables simultaneously. As a matter of fact, the impact is immense when three variables are jointly managed.

Nevertheless, the best combination of improvement actions may be easily identified by performing an evidence sensitivity analysis. A BN, indeed, can be also used for identifying the most suitable scenario able to shock the maximum status of the hypothesis variable. In detail, a subset of instantiated variables is considered to be impacting if they produce a normalized likelihood (NL) greater than 1. The analysis has been carried out by instatiating all the four parents of the target node (see Table 7); it’s worth noting that the most impacting scenario doesn’t involve the entire set of variables (7th position) but only those resulting from value of information analysis and expert evaluation. Figure 11 shows that a scenario instatiating to their maximum the three variables of the NL maximizing configuration in Table 7, produces a 283% increase in the probability of the highest level of the target node.

Finally, it’s worth noting that the resulted most impacting scenario produces a NL larger when we focus on the South (NL = 3.89) rather than on the North (NL=3.83), on the Center (NL=3.84) or on the Islands (NL = 3.75). This allows to better orient the policy that, according to the results of the analysis, should rethink organizational processes by taking care of internal and external communication, by reorganizing strategies and by starting from the South of the Country.

5 Conclusions

In this paper we have shown a model for supporting school managers in their decision process. Due to the complexity of the topic, a multivariate statistical approach based on Bayesian networks has been chosen, being also proper to evaluate scenarios. A BN model has been estimated from a set of data coming from a multitarget observational study. The multivariate dependence structure learnt form the data depicts the relations among variables and highlights the relevance of one variable to another, supporting decisors in mapping the intricate influences of aspects affecting the organizational process improvement. As a consequence, the estimated model helps school managers in identifying those key aspects needing improvement actions and providing a strong impact on the target variable. Different possible decisions can be easily evaluated using the inferential engine proper of BNs. Section 4.1 has shown some simulating scenarios.
helping managers in orienting strategies for improving the organizational processes. By comparing experts’ evaluation and value of information analysis some priorities emerged and their impact, in probabilistic terms, on the variable of interest has been investigated by means of evidence propagation. In Fig. 12 a summary of the effects on

Fig. 10 Network and monitor windows: probability table of organizing process improvement expressed in percentage before (a) and after (b) a simulating action in satisfaction ds-teachers communication
the target node organizing process improvement of the simulation of the four scenarios illustrated in Sects. 4.1.1–4.1.4 is shown. It emerges that improvement in the target variable can be achieved handling different aspects. Moreover, a complex scenario involving the action on more than one variable simultaneously has been simulated in Sect. 4.1.5 showing that a joint intervention produces a dramatic impact on the target variable. Therefore, BNs help school managers both possibly confirming the strategic actions they have a priori established, and identifying additional key aspects to intervene on to achieve the planned results.

6 Discussion

Due the pandemic, Italian School managers have tackled a challenging attempt to support organisational processes in the school, which was already affected by a severe immobility, above all in the digital field. This is the reason way the school, during the pandemic, has started to think differently and to wonder what is the best strategy to accompany a drastic change postponed too long. The present work arises form this need: highlighting the main challenges faced by school managers during the lockdown. Our aim was providing school managers with a multivariate statistical tool supporting them in the probabilistic evaluation of the effect of improvement actions on the organization process. Beyond the pandemic, the illustrated approach may be extended to managers and political decisors, helping them to compare strategies that may have an impact on multiple actors. As a matter of fact, the analysis here discussed focuses on a single point of view, that of school principals. In future research it could be enriched by including the evaluation of the other actors of the educational process, such as teachers, students and families, giving rise to a complex integrated managerial tool able to represent the school system as a whole with all the stakeholders, and to manage it in a holistic way. In this context, the relationships among the different dimensions of the school system will be represented by means of a complex network structure. In such a way, probabilistic evaluation of the impact of an action of a stakeholder both on its reference dimension and on the other actors dimensions can be performed.

| Organizing Process | Baseline P (Priority strategy redesign =5) =1 | 1 | 2 | 3 | 4 | 5 |
|-------------------|---------------------------------------------|---|---|---|---|---|
| Improvement       | P (Priority strategy redesign =5) =1        | 0.1075 | 0.1083 | 0.2234 | 0.4044 | 0.1565 |
|                   | P (Satisfaction DS-teachers communication=5)=1 | 0.0617 | 0.0617 | 0.1247 | 0.4426 | 0.3092 (+97%) |
|                   | P (Internal communication improvement=5)=1 | 0.176 | 0.0176 | 0.0176 | 0.4168 | 0.5303 (+239%) |

Table 6 Probability table of the target variable before (baseline) and after simulated scenarios
### Table 7  Impact, in terms of NL, of a subset of instantiated variables (parents node of Organizing process improvement)

| Nodes                               | Instantiated |
|-------------------------------------|--------------|
| Priority strategies redesign        | x            |
| Satisfaction DS-teachers communication | x            |
| School-family communication improvement | x            |
| Internal communication improvement  | x            |
| NL                                  | 3.83 3.39 2.76 2.3 2.23 2.15 2.13 1.99 1.98 1.92 1.82 1.66 1.53 1.37 1.28 |
Fig. 11 Network and monitor windows: probability table of organizing process improvement expressed in percentage before (a) and after (b) a simulating action in the three variables elected by the sensitive analysis, i.e. Priority strategies redesign, School-family communication improvement and Internal communication improvement.
Acknowledgements

This research has been supported by the project "Didattica a distanza ai tempi del Covid-19" by DiTES research center of Link Campus University in cooperation with University of Roma Tre, ANP, Forum Associazioni Familiari and AIDR. We want to thank the anonymous referees for their useful comments and suggestions.

References

Allulli, G. (2012). Modello teorico integrato di valutazione delle strutture scolastiche e formative. FSE: ISFOL.

Barnard, C. (1968). The function of the executive. Harvard University Press.

Buntine, W. (1994). Operations for learning with graphical models. Journal of Artificial Intelligence Research, 2, 159–225.

Buntine, W. (1996). A guide to the literature on learning probabilistic networks from data. IEEE Transaction on Knowledge and Data Engineering, 8(2), 195–210.

Castoldi, M. (2007). La qualità a scuola, Roma, Ed. Carocci, Faber

Chickering, D. M., Heckerman, D., Meek, C., & Madigan, D. (1994). Learning Bayesian networks is NP-hard. Technical Report MSR-TR-94-17, Microsoft Research, Microsoft Corporation, Redmond, Washington

Cooper, G., & Herskovits, E. (1992). A Bayesian method for constructing Bayesian belief networks from databases. Machine Learning, 9(4), 309–47.

De Capogna, S., Angelis, M. C., & Musella, F. (2020). Exploring practices, problems and cultural challenges of Italian teachers in the digital era. Scuola democratica, Learning for Democracy, 2, 259–284.

Hargittai, E. (2010). Digital Na(t)ives? Variation in internet skills and uses among members of the net generation. Sociological Inquiry, 80(1), 92–113

Heckerman, D. (1995). A tutorial on Learning with Bayesian Networks. Technical Report MSR-TR-95-06, Microsoft Research

Illeris, K. (2003). Three dimensions of learning: Contemporary learning theory in the tension field between the cognitive, the emotional and the social. Paperback. Malabar.

Jackson, L. A., Zhao, Y., Kolenic III, A., Fitzgerald, H. E., Harold, R., & Von Eye, A. (2008). Race, gender, and information technology use: The new digital divide. CyberPsychology & Behavior, 11(4), 437–442.

Kalisch, M., Mächler, M., Colombo, D., Maathuis, M. H., & Bühlmann, P. (2012). Causal inference using graphical models with the R package pcalg. Journal of Statistical Software, 47(11), 1–26.

Kyrkaides, L., Campbell, R., & Christofidou, E. (2002). Generating criteria for measuring teacher effectiveness through a selfevaluation approach: A complementary way of measuring teacher effectiveness. School Effectiveness and School Improvement, 13, 291–325.

Livingstone, S., & Helsper, E. (2007). Gradations in digital inclusion: Children, young people, and the digital divide. New Media & Society, 9(4), 671–696.

MaelRoy, B.: Measuring response rates in online surveys. Quirk’s Marketing Research, 583 (2000).

Marella, D., Musella, F., & Vicard, P. (2014). Learning Bayesian networks in complex survey sampling. In Proceedings of SIS conference

Metz, C. E. (1978). Basic principles of ROC analysis. Seminars in Nuclear Medicine, 8(4), 283–298.

Mintzberg, H. (1983). La progettazione dell’organizzazione aziendale, Milano, Il Mulino

Springer
Musella, F. (2013). A PC algorithm variation for ordinal variables. *Computational Statistics, 28*(6), 2749–2759.

Musella, F. (2015). A proposal for learning Bayesian networks from categorical variables. In *SIS 2015 meeting statistics and demography: The Legacy of Corrado Gini*

Neapolitan, R. E. (2003). *Learning Bayesian networks*. Pearson Prentice Hall.

Pearce, J., & Ferrier, S. (2000). Evaluating the predictive performance of habitat models developed using logistic regression. *Ecological Modelling, 133*, 225–245.

Pearl, J. (1986). A constraint-propagation approach to probabilistic reasoning. In Kanal, L. N., & Lemmer, J. F. (Eds.), *Uncertainty in artificial intelligence*. North Holland.

Pitzalis, M., Porcu, M., De Feo, A., & Giambona, F. (2016). Innovare a scuola. Insegnanti, studenti e tecnologie digitali, Bologna, Il Mulino.

Spirtes, P., Glymour, C., & Scheines, R. (2000). *Causation, prediction, and search* (2nd ed.). MIT Press.

Steck, H. (2001). Constraint-based structural learning in Bayesian networks using finite data. PhD thesis, Institut für Informatik der Technischen Universität München.

Sutcliffe, M. K. (2007). Organizational resilience: Towards a theory and research agenda. Conference Paper, October.

Swets, K. A. (1988). Measuring the accuracy of diagnostic systems. *Science, 240*, 1285–1293.

Verma, T., & Pearl, J. (1991). Equivalence and synthesis of causal models. 255–268

Weick, K. E., & Sutcliffe K. M. & Obstfeld, D. (2005). Organizing and the process of sensemaking. *Organization Science, 16*(4), 409–421.

**Publisher’s Note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.