A measurement method of routing flexibility in manufacturing systems

F. Zammori*, M. Braglia*, M. Frosolini*

*Dipartimento di Ingegneria Meccanica, Nucleare e della Produzione Via Bonanno, 25/A 56126 Pisa, Italia

ABSTRACT

This paper focuses on routing flexibility, which is the ability to manufacture a part type via several routes and/or to perform different operations on more than one machine. Specifically, the paper presents a comprehensive method for the measurement of routing flexibility, in a generic manufacturing system. The problem is approached in a modular way, starting from a basic set of flexibility indexes. These are progressively extended to include more comprehensive and complex routing attributes, such as: the average efficiency, the range and the homogeneous distribution of the alternative routes. Two procedures are finally proposed to compare manufacturing systems in terms of routing flexibility. The first one uses a vectorial representation of the previously defined indexes and the second one is based on data envelopment analysis, a multi-criteria decision making approach. The paper concludes with a numerical example, supported by discrete event simulation, which validates the proposed approach.

1. Introduction

Nowadays the industrial field is characterized by a continuous strain to enhance manufacturing flexibility, which has become one of the main levers to succeed in an ever-changing market. Manufacturing flexibility can be generally defined as the ability of a productive system to quickly react to the changes occurring in its internal and/or external environment, and is especially important for flexible manufacturing systems (FMS) and for flexible assembly systems (FAS). Indeed, this significant attribute distinguishes advanced manufacturing systems from the traditional high-volume process-dedicated production systems, like flow shops and/or automated transfer lines (Chan, 2001; Borenstein & Rohde, 2005). In addition, FMS and FAS generally require conspicuous initial investments that are difficult to be justified, unless the tangible and intangible benefits arising from an increase of flexibility can be fully captured and quantified. As observed by Gupta and Goyal (1989), Sethi and Sethi (1990), Sarker et al. (1994) and Zhang et al. (2002), measuring flexibility is a major concern, because the use of financial evaluation methods (i.e. Net present value, internal rate of return and payback period) is seldom appropriate to take decisions on the acquisition of advanced manufacturing technologies.
Unfortunately, there is not a unified framework to quantify and measure flexibility (Slack, 1987; Beskese et al. 2004) and some authors (Gupta & Buzacott 1989) even raise doubts on the possibility of measuring it, by means of quantitative attributes only. This is due to the high number of potential environmental changes (machine breakdowns, volume and/or mix changes, introduction of new products, etc.) that make it hard or even impossible to include all the critical aspects of flexibility into a single metric (Azzone & Bertelè, 1991). Nonetheless, useful managerial insights can be obtained by dividing flexibility into independent elementary concepts and limiting the analysis to few distinctive features of a manufacturing system. As a matter of fact a great effort has been made to build quantitative analytical tools to measure different types of flexibility, such as: sequencing flexibility, machine flexibility, routing flexibility, volume flexibility, product mix flexibility, layout flexibility and labour flexibility (Brill & Mandelbaun, 1989; Hutchinson & Sinha, 1989; Taymaz, 1989; Kochikar & Narendran, 1992; Nagarur, 1992; Roll et al., 1992; Chen & Chung, 1996; Das, 1996; Chang, 2004).

The present work focuses on routing flexibility, which is the ability to process a part type via several routes. This choice is motivated by the fact that routing flexibility has been recognized as a fundamental competitive lever for advanced manufacturing systems (Caprihan & Wadhwa, 1997; Yu & Green, 2000; Chang, 2007), since it simplifies the scheduling and the balancing of the machines and facilitates the fulfillment of the customers’ requirements (Sethi & Sethi, 1990). Furthermore, in technical literature there seems to be a lack of multidimensional approaches capable of fully describe this important parameter (Chang, 2007). Therefore, the objective of this paper consists in the development of a comprehensive methodology to measure and to capture the main aspects of routing flexibility.

Nomenclature

\[ P = \text{number of part types} \]
\[ p = \text{part type subscript} \]
\[ M = \text{number of machines} \]
\[ m = \text{machine subscript} \]
\[ R_p = \text{number of routes for the } p\text{-th part type} \]
\[ \bar{R}_p = \text{number of alternative routes for the } p\text{-th part type} \]
\[ TR = \text{Total number of routes} \]
\[ r = \text{route subscript} \]
\[ M = \text{Machine-Part matrix} \]
\[ T = \text{Time matrix} \]
\[ C = \text{cost matrix} \]

Main indexes

\[ \text{ARE}_p = \text{Alternative Route Efficiency} \]
\[ \text{JRAE} = \text{Job Routing Average Efficiency} \]
\[ \text{JRR} = \text{Job Routing Range} \]
\[ \text{JRV} = \text{Job Routing Versatility} \]
\[ \text{GRE} = \text{Global Routing Efficiency} \]
\[ \text{GRR} = \text{Global Routing Range} \]
\[ \text{GRV} = \text{Global Routing Versatility} \]
\[ \text{GRF} = \text{Global Routing Flexibility} \]

2. Routing flexibility

In this paper routing flexibility follows the frequently adopted definition given by Browne et al. (1984) and it is considered as “the ability to handle breakdowns and to continue producing the given set of part types. This ability exists if either a part type can be processed via several routes or, equivalently, if each operation can be performed on more than one machine”. In other words, the
concept of routing flexibility is strictly related to the capacity to handle breakdowns by means of alternative routes, and becomes evident if production has no meaningful and dramatic downtimes.

With the terminology of Buzacott (1982), this type of flexibility is also referred as cycle time flexibility, for it protects production rate from breakdowns. Alternatively, Chen et al. (1992), Gerwin (1993) and Das (1996) named it re-routing flexibility, considering the number of viable routes as its main feature. Literature on the subject matter is rather extensive and one of the first attempts to measure routing flexibility can be found in the work by Browne et. al (1984), who proposed the use of the average variation in the production rate (when a breakdown occurs) as a first proxy of routing flexibility. Chatterjee et al. (1987) suggested a different measure based on the ratio of the number of available routes, to the total number of manufactured products. Similarly, by observing that routing flexibility (RF) is linked to the ability of the material handling system to connect two generic machines, Nagarur (1992) proposed an index based on the proportion of the available alternative routes with respect to the potential ones:

$$RF = \frac{\sum_{l} \sum_{m} X_{lm}}{M(M-1)}$$

(1)

where $X_{lm}$ equals one if the $l$-th and the $m$-th machine (with $l \neq m$) are connected and zero otherwise, and $M(M-1)$ is the total number of potential routes in a system consists of $M$ machines. An equivalent approach was proposed by Kochikar and Narendran (1992), who directly addressed the measurement of routing flexibility in FMSs with the introduction of the Producibility index $\rho_{p}(M)$ of the $p$-th part type with respect to the $M$-th machine set. This index represents the ratio of the number of available routing options at each stage, over the total number of options existing at that stage.

All the previously mentioned approaches consider a single dimension of routing flexibility; however, to get a more comprehensive measure, Das (1996) stressed the necessity to create a more complex analytical tool capable to capture the differences of alternative routes in terms of disparity and efficiency. Disparity should account for the level of machinery and equipment communality, while efficiency should address the difference in processing time between an alternative route and the shortest one. The author also observed that each flexibility dimension should be conveniently evaluated by multiple measures, rather than by a single one. Indeed a multi level approach makes it possible to separate and discriminate the different aspects of flexibility in a more precise and reliable way. A first attempt to address multiple dimensions of routing flexibility was made by Chen and Chung (1996), who proposed a set of indexes relative to a generic set of part types. In doing so the authors introduced and distinguished the concepts of potential routing flexibility, actual routing flexibility, and routing flexibility utilisation.

Recently, Chang et al. (2001) argued that a comprehensive model for the measurement of routing flexibility should consider, at least, two important attributes of a manufacturing system: the efficiency $E$ and the versatility $V$ (i.e. the numerousness) of its routes. Starting form this work, Chang (2007) proposed a three dimensional framework, which also considers routing variety $D$, an attribute that quantifies the differences of the alternative routes available for a part type $p$. Specifically, the variety $D$ is evaluated as follows,

$$D_{p} = \frac{1}{R_{p}(R_{p}-1)} \sum_{r=1}^{R_{p}} \sum_{s=1, s \neq r}^{R_{p}} d_{(rs),p}.$$  

(2)

where $R_{p}$ is the number of routes (of the $p$-th part type) and $d_{(rs),p}$ is the ratio of the number of different machines to the total number of machines of the $r$-th and $s$-th route.
To compute the efficiency $E_p$, several operating parameter are combined by means of data envelopment analysis (DEA), a multi-criteria decision making technique based on linear programming (Cooper et al., 2004). In doing so, an output oriented CCR model (Charnes et al., 1978) is used, taking manufacturing costs, set up and processing times as the input decision variables, and the output quantity and quality (of the routes) as the output decision variables. Solving a CCR model for each route, one obtains $R_p$ basic efficiencies $e_{rp}$, whose average value gives the overall efficiency $E_p$:

$$E_p = \frac{1}{R_p} \sum_{r=1}^{R_p} e_{rp}. \quad (3)$$

To evaluate the versatility $V_p$, the author proposed a method based on the entropy approach, which was firstly introduced by Shannon (1948) in information theory and then adapted to the measurement of flexibility by Yao (1985) and Kumar (1987). As demonstrated by Eq. (4), this choice is motivated by the observation that the entropy approach satisfies both the versatility and the uniformity requirement of a flexible system, as its value increases with the number of alternative routes and with the even distribution of efficient routes among different part types.

$$V_p = - \sum_{r=1}^{R_p} v_{rp} \log(v_{rp}), \quad (4)$$

where $v_{rp}$ represents the normalized efficiency of the $r$-th alternative route.

Finally, if $P$ part types are manufactured, the routing flexibility of the whole system is measured as:

$$ROFLX = \frac{1}{P} \sum_{p=1}^{P} E_p \cdot V_p \cdot D_p. \quad (5)$$

In this paper, an alternative framework for the assessment of routing flexibility is presented, with the objective to incorporate several crucial performance factors into a consistent set of routing flexibility indexes. To this aim, the problem is approached in a modular way, starting from basic flexibility indexes, which are progressively extended in order to include more detailed and complex routing aspects in a set of metrics, suitable at a different level of complexity and characterized by an easy analytical shape. The main aspects that have been considered are: (i) the efficiency (evaluated in term of cost and time), (ii) the number and (iii) the homogeneous distribution of the alternative routes. Additionally, the basic set of indexes has been expanded to capture other meaningful parameters that can affect the routing flexibility of a manufacturing system. These are: the covering degree, the production quality, the backtracking probability and the availability of the equipment installed in the plant.

To complete the work and to improve its practical utility, two approaches, that make it possible to compare manufacturing systems in terms of routing flexibility, are also presented. The first one uses a vectorial representation to visualize the global routing flexibility in a three dimensional space, and to compare it with that of an ideal manufacturing system. The second one is based on a multi-output single-input DEA model, and can be used to compare two or more manufacturing systems in relative terms. Both approaches are clearly explained and validated by means of a numerical example supported by a discrete event simulation model.

3. Operating data

In the following part of the paper we will consider a production system characterized by $P$ part types (i.e. jobs) and $M$ machines. Each product must be associated to a binary Part-Machine matrix $\bar{M} = [x_{m,r}]$ that encodes all its feasible routes. In the matrix, routes are listed in columns, machines
are listed in rows and a generic element \( x_{mr} \) is one if the \( r-th \) processing route requires the \( m-th \) machine, and zero otherwise. The convention is used to assign the first column of \( M \) to the standard route and to use the subscript \( r = 0 \) to denote the standard route.

For the sake of simplicity, all the Part-Machine matrixes can be assembled in a three dimensional structure \( M = [x_{m,r,p}] \), whose third dimension refers to the part type. In this way, a generic entry \( x_{mpr} \) is one if the \( r-th \) processing route of the \( p-th \) part type requires the \( m-th \) machine. An example (Matrix 1) of an \( M \) matrix is built considering three part types \( (p_i) \) manufactured using five different machines \( (m_i) \).

\[
\begin{array}{cccccc}
\text{p1} & | & \text{p2} & | & \text{p3} & \\
\hline
m_1 & 1 & 1 & 1 & 1 & 0 \\
m_2 & 0 & 1 & 0 & 1 & 1 \\
m_3 & 0 & 1 & 0 & 1 & 1 \\
m_4 & 1 & 0 & 0 & 1 & 1 \\
m_5 & 1 & 1 & 1 & 0 & 0 \\
\end{array}
\]

\[ M = \begin{bmatrix}
r_01 & r_{11} & r_{02} & r_{12} & r_{03}
\end{bmatrix}\]

**Matrix 1 An example of three-dimensional Part-Machine matrix**

Each column \( r_{0p} \) represents the standard route of the \( p-th \) part type, while \( r_{rp} \) denotes its \( r-th \) alternative route. For example, both \( p_1 \) and \( p_2 \) can be manufactured via an alternative route, \( r_{11} \) and \( r_{12} \), respectively.

Replacing all the values \( x_{mpr} \) of the \( M \) matrix with the corresponding processing times \( l_{mpr} \), a processing time matrix \( T = [t_{m,r,p}] \) can be obtained. In this case, the values \( l_{mpr} \) should include, at least, the cycle time and the transportation time that are needed to process and move the \( p-th \) part along its \( r-th \) route. Besides, the time values \( l_{mpr} \) could also include the set-up time, especially when the change-over tasks have a meaningful variation in relation with the adopted production sequence.

The operating data that will be used to evaluate routing flexibility include, lastly, the manufacturing cost per time unit. These costs must be computed for each machine and can be arranged within an \( M \)-dimensional array \( \tilde{C} = [\tilde{c}_1, \tilde{c}_2, \ldots, \tilde{c}_m, \ldots, \tilde{c}_M] \), whose generic element \( \tilde{c}_m \) represents the cost per unit of time of the \( m-th \) machine. Finally, if the production times are multiplied by the corresponding costs, the production cost matrix \( C = [c_{m,r,p}] = [l_{m,r,p} \cdot \tilde{c}_m] \) is obtained. From the values of \( C \), the route production cost \( R_{C rp} \) can be easily computed as follows:

\[
R_{C rp} = \sum_{m=1}^{M} c_{mpr} 
\]

\[ (6) \]

**4. Basic routing flexibility indexes**

In this section the basic metrics of alternative processing routes will be defined. Subsequently, these metrics will be aggregated to evaluate the routing flexibility of a productive system, taking into account the number, the efficiency and the homogeneous distribution of all the feasible alternative routes.
4.1. Job Routing Average Efficiency

The efficiency of the \( r \)-th alternative route of the \( p \)-th part type can be expressed in terms of the route production cost \( RC_{rp} \), taking the cost of the standard route \( RC_{0p} \) as the reference parameter. In this way, the alternative route efficiency \( ARE_{rp} \) is formally defined as:

\[
ARE_{rp} = \frac{RC_{0p}}{RC_{rp}} \quad \forall \ r \neq 0
\]  

(7)

If a part type has several alternative routes \( \tilde{R}_p \), a second index, called job routing average efficiency \( JRAE_p \), can be used to evaluate the average efficiency of its alternative routes:

\[
JRAE_p = \frac{\sum_{r=1}^{\tilde{R}_p} ARE_{rp}}{\tilde{R}_p}.
\]  

(8)

4.2. Job Routing Range

The index \( JRAE_p \) has the advantage to be very easy, but unfortunately it does not adequately consider the number of alternative routes deployed for the \( p \)-th part type and, for this reason, it can lead to misleading results.

For instance, consider the situation given in Matrix 2, which displays three part types characterized by one, two and three alternative routes, respectively. In this case, a straight comparison of the job routing average efficiencies \( JRAE_1 = 0.9, JRAE_2 = 0.8, JRAE_3 = 0.63 \) would indicate \( p_1 \) as the part type with the higher routing flexibility. This result could legitimize some doubts: as a matter of fact the routing flexibility of \( p_3 \) should be greater than that of \( p_1 \) because \( p_3 \) has three alternative routes and one of them has the same efficiency as \( r_{11} \), which is the only alternative route of \( p_1 \).

\[
C = \begin{array}{cccccc}
2 & 2 & 1 & 1 & 2 \\
1 & 2 & 1 & 1 & 4 \\
3 & 2 & 2 & 2 & 3 \\
3 & 4 & 2 & & & \\
2 & 2 & 1.25 & 2 & 4 \\
2 & 3 & 1 & 2 & 1 & \\
3 & 2 & 3 & & 3 \\
1 & 1 & 2 & 2 & 1.4 & 3 & 4 \\
\end{array}
\]

Matrix 2 An example of Cost matrix

Such problem can be overcome through the introduction of an additional index called job routing range \( JRR \), which is zero for an item without alternative routes. To compute \( JRR \) we start by observing that, the greater the number of viable routes, the greater the flexibility. However, the marginal increase of flexibility (due to the addition of an alternative route) also depends on the number of the alternative routes \( \tilde{R}_p \) available for the \( p \)-th part type. If \( \tilde{R}_p \) is low, the capacity to counteract failures is noticeably augmented as the number of alternative routes increases from \( \tilde{R}_p \) to
(R_p + 1). Vice versa if \( \hat{R}_p \) is high, the effect of an additional route has a negligible effect. To fulfill these requirements, JRR_p can be represented as an increasing function of \( \hat{R}_p \):

\[
J_{RR}(\hat{R}_p) = \begin{cases} 
1 - \left(1 - \frac{\hat{R}_p}{(TR - 1)}\right)^{\alpha_p} & \text{if } TR \geq 2 \\
0 & \text{otherwise}
\end{cases}
\]

(9)

where \( TR \) is the total number of routes of the system and \( \alpha_p \) is a shape parameter greater than one.

As clearly shown in Eq. (9), JRR_p is a nonzero positive number, provided that there is, at least, one alternative route. Specifically, JRR_p equals zero if the \( p \)-th product can be processed only via its standard route (i.e. \( \hat{R}_p = 0 \)), whereas JRR_p equals one if the part type shares all the alternative routes deployed in the system (i.e. \( \hat{R}_p = [TR-1] \)). Conversely, when there is a unique processing route (i.e. \( TR = 1 \) and \( \hat{R}_p = 0 \forall p \)) JRR_p equals zero, because the manufacturing system does not have routing flexibility. This particular condition does not have any practical interest and have been included in Eq. (9) only for the sake of clarity. Indeed, a system with a single route corresponds to a rigid production/assembly line, which cannot be adequately described with the set of indexes proposed in this paper.

To better explain the effect of the shape parameter on the job routing range, Fig. 1 shows the shape of JRR_p when \( TR \) equals ten and \( \alpha_p \) increases from one to nine.

![Fig. 1. Job routing range evaluation function](image)

To evaluate JRR_p one needs to define a sensible value for the shape parameter \( \alpha_p \). To this aim, an approach based on the complexity and on the availability of the machines of the standard route can be used. The complexity of the standard production route is an important parameter for the determination of \( \alpha_p \) because it influences the easiness to generate new alternative routes. This peculiar parameter can be captured by the standard route complexity index SRC_p, defined as the number of machines of the standard route, over the number of machines installed in the plant:

\[
SRC_p = \frac{\sum_{m=1}^{M} x_{m0p}}{M}
\]

(10)

where \( M \) is the total number of machine and \( x_{m0p} \) is a generic value of the standard route columns of the \( M \) matrix.

In the computation of \( \alpha_p \) also machine availability should be taken into account. Indeed, to avoid dramatic production downtimes, it is a good managerial practice to deploy alternative routes for jobs that are processed by machines subject to frequent breakdowns and/or to long lasting periods of
maintenance. Conversely, if the availability of the machines of the standard route is high, there is no need to introduce many alternative routes.

Under the hypothesis that the likelihood of two simultaneous failures (for the standard and for the alternative route) is negligible, the standard route availability \( S_{RA_p} \) can be expressed as:

\[
S_{RA_p} = \prod_{m=1}^{M} x_{m0p} A_m
\]  

\[
A_m = \frac{MTBF_m}{MTBF_m + MTTR_m}
\]

where \( A \) is the steady state availability, \( MTBF \) and \( MTTR \) stand for mean time between failure and mean time to repair, respectively (O'Connor, 2002).

\( SRC_p \) and \( SRA_p \) can be finally combined to obtain a meaningful value for \( \alpha_p \):

\[
\alpha_p = \left( \frac{1}{1 - SRA_p} \right)^{SRC_p}
\]

As shown by Eq. (13) \( \alpha_p \) decreases when the route availability \( SRA_p \) approaches one. This is because \( (1 - SRA_p) \) represents an estimation for the frequency of the \( p \)-th job to be processed with an alternative route, and can be seen as an indicator of the necessity to deploy additional routes. Conversely \( \alpha_p \) increases as the complexity \( SRC_p \) tends to one. Indeed, the capability to arrange alternative routes greatly diminishes with the increase in the complexity of the standard route.

To better explain the use of \( JRR_p \), we re-consider the previous example under the hypothesis that the standard routes for \( p_1, p_2 \) and \( p_3 \) are characterized by a 90% availability. In this case, remembering that \( p_1, p_2 \) and \( p_3 \) do not share any route (i.e. \( TR = 9 \)), that \( M = 8 \) and that \( \tilde{R}_1 =1, \tilde{R}_2 =2, \tilde{R}_3 =3 \), one obtains the following estimates:

\[
SRC_1 = SRC_2 = SRC_3 = \frac{\sum_{m=1}^{8} x_{m0p}}{8} = 0.5
\]

\[
\alpha_1 = \alpha_2 = \alpha_3 = \left( \frac{1}{1 - SRA_p} \right)^{SRC_p} = \left( \frac{1}{0.1} \right)^{0.5} \approx 3.16
\]

\[
JRR_1 \approx 34.4\%, \quad JRR_2 \approx 59.7\%, \quad JRR_3 \approx 77.3\%
\]

In accordance with these values, the ranking that was previously obtained by making a straight comparison of the JRAE indexes (i.e. \( p_1, p_2, p_3 \)) has been completely altered. Indeed \( p_1 \) is now ranked as the part type with the highest routing flexibility, followed by \( p_2 \) and by \( p_1 \), respectively. This seems to be correct because, as previously noted, \( p_3 \) has three alternative routes and one of them has the same efficiency as \( r_{11} \), which is the only alternative route of \( p_1 \).

### 4.3. Job Routing Versatility

In analogy with the approach proposed by Chang (2001, 2007), the information concerning the efficiency and the range of the routes can be aggregated into a single metric called job routing versatility \( JRV_p \). To this aim, for each part type we introduce two vectors, namely the efficiency vector \( E_p \) and the range vector \( R_p \), both of dimension \( (TR - 1) \). As shown in Eq. (14), the first \( \tilde{R}_p \) elements of \( E_p \) are the alternative routing efficiencies of the \( p \)-th part type, indexed in a decreasing way. All the remaining \( (TR - \tilde{R}_p - 1) \) values are zero.
\[ E_p = \left[ (\text{ARE}_{rp})_1, (\text{ARE}_{rp})_2, \ldots, (\text{ARE}_{rp})_i, \ldots, (\text{ARE}_{rp})_{R_p}, 0, \ldots, 0 \right] \]

\[ \text{with } i = 1, \ldots, (TR - 1), \quad E_p[i] \neq 0 \quad \forall i \in [1, R_p] \quad \text{and} \]

\[ (\text{ARE}_{rp})_1 \geq (\text{ARE}_{rp})_2 \geq \cdots \geq (\text{ARE}_{rp})_i \geq \cdots \geq (\text{ARE}_{rp})_{R_p} \]

Similarly, as shown in Fig. 2, \( R_p \) contains the marginal increments of flexibility \( \Delta JRR_p(i) \), that can be obtained with the introduction of an additional route for the \( p \)-th part type.

\[ R_p = [\Delta JRR_p(1), \Delta JRR_p(2), \ldots, \Delta JRR_p(i), \ldots, \Delta JRR_p(TR - 1)] \]

\[ \text{with } i = 1, \ldots, (TR - 1) \quad \text{and} \quad \Delta JRR_p(i) = \left( JRR_p(i) - JRR_p(i - 1) \right) \]

\[ \text{Fig. 2. Job Routing range marginal increments} \]

Since the flexibility rises as the number of alternative routes increases and/or if the efficient routes are evenly distributed among different part types, the versatility can be obtained as the scalar product of vector \( E_p \) and \( R_p \). In this way, the routing flexibility is evaluated as a weighted average of the efficiency of each alternative route, using the marginal increment of flexibility as weighting coefficients.

\[ JRV_p = \langle E_p, R_p \rangle . \]

Since the sum of the elements of \( R_p \) equals one (i.e. \( \sum_{i=1}^{TR-1} R_p[i] = 1 \quad \forall p \)), it is easy to see that \( JRV_p \) changes between zero and one. Specifically \( JRV_p \) can be one only if the \( p \)-th part type can be manufactured via \( (TR-1) \) alternatives routes with a 100\% efficiency. To better explain the concept of \( JRV_p \), let us consider again the numerical example introduced in section 4.2. In this case it results:

\[ JRV_1 = [0.9, 0, 0, 0, 0, 0, 0, 0] \times [0.344, 0.253, 0.176, 0.115, 0.067, 0.033, 0.011, 0.001] \approx 31\% \]

\[ JRV_2 = [0.8, 0.8, 0, 0, 0, 0, 0, 0] \times [0.344, 0.253, 0.176, 0.115, 0.067, 0.033, 0.011, 0.001] \approx 48\% \]

\[ JRV_3 = [0.9, 0.6, 0.4, 0, 0, 0, 0, 0] \times [0.344, 0.253, 0.176, 0.115, 0.067, 0.033, 0.011, 0.001] \approx 53\% \]

This result improves the ranking obtained using the JRR indexes because, as observed at the beginning of section 4.2, the routing flexibility of \( p_2 \) and \( p_3 \) must be similar and bigger than that of \( p_1 \).

5. Global routing flexibility indexes

\( JRAE_p \) and \( JRR_p \) can be used, at an aggregate level, to characterize the whole manufacturing system in terms of global routing efficiency (GRE) and global routing range (GRR). These two global
indices are defined in Eqs. (17-18) and give the average efficiency and the average range of the routes of a manufacturing system.

\[
\text{GRE} = \frac{\sum_{p=1}^{P} \text{JRAE}_p}{P},
\]

(17)

\[
\text{GRR} = \frac{\sum_{p=1}^{P} \text{JRR}_p}{P}.
\]

(18)

Unfortunately, GRE and GRR tend to neglect how evenly routes are distributed among the \( P \) part types. As a matter of fact, a system characterised by many alternative routes, but relevant to a few products only, will result less flexible than a system with an identical number of routes, but uniformly distributed among several products. Similarly, a manufacturing system where only a few products can exploit the most efficient processing routes and the majority of products rely on the least efficient ones is, by evidence, less flexible than a system which utilizes both efficient and inefficient routes with higher product uniformity.

To find a remedy to this circumstance we can use \( \text{JRV}_p \) to define an additional global index called global routing versatility (GRV), which should be zero for a system characterized by the worst possible distribution (of the alternative routes) and one in the best case. The first circumstance corresponds to the condition of \((P - 1)\) part types being processed via the standard route and one item alone being the outcome of all the alternative routes listed in the \( M \) matrix (i.e. \( \text{JRV}_1 = \text{JRV}_2 = \text{JRV}_3 = \ldots = \text{JRV}_{(P-1)} = 0; \text{JRV}_P > 0 \)). Conversely, the latter circumstance corresponds to a homogeneous distribution of routes among jobs (i.e. \( \text{JRV}_1 = \text{JRV}_2 = \ldots = \text{JRV}_P > 0 \)).

Considering these requirements, to evaluate GRV we can use the Gini concentration index (GCI), which is a common measure of statistical dispersions (Gini, 1921).

\[
\text{GCI} = \sum_{j=1}^{P} \left( \frac{j}{P} - \frac{\sum_{i=1}^{j} \text{V}[i]}{\text{JRVTOT}} \right),
\]

(19)

where \( \text{JRVTOT} = \sum_{i=1}^{P} \text{V}[i], \) and \( \text{V}[i] \) is the \( i \)-th element of the routing versatility vector \( V \), which contains the values of \( \text{JRV}_p \) for each part type, indexed in an increasing order:

\[
V = \begin{bmatrix} \text{JRV}_1, \text{JRV}_2, \ldots, \text{JRV}_i, \ldots, \text{JRV}_P \end{bmatrix}
\]

(20)

with \( i = 1, \ldots, P \) and \( (\text{JRV}_p)_1 \leq (\text{JRV}_p)_2 \leq \ldots \leq (\text{JRV}_p)_i \leq \ldots \leq (\text{JRV}_p)_P \).

When a single part type is the outcome of all the alternative routes listed in the \( M \) matrix, GCI equals \((P - 1)/2\), which is the maximum value that this index can take. Indeed, in this particular circumstance, since a single product (i.e. the one placed in the \( P \)-th position of vector \( V \)) benefits of a non null routing flexibility, we have \( \text{V}[i] = 0 \ \forall \ i \neq P \) and \( \text{JRVTOT} = \sum_{i=1}^{P} \text{V}[i] = \text{V}[P] \). Therefore Eq. (19) can be simplified as follows,

\[
\text{GCI} = \sum_{j=1}^{P} \left( \frac{j}{P} - \frac{\sum_{i=1}^{j} \text{V}[i]}{\text{JRVTOT}} \right) = \left( \frac{1}{P} - \frac{0_1}{\text{V}[P]} \right) + \left( \frac{2}{P} - \frac{0_1 + 0_2}{\text{V}[P]} \right) + \ldots + \left( \frac{P}{P} - \frac{0_1 + \ldots + 0_P + \text{V}[P]}{\text{V}[P]} \right)
\]

\[
= \left( \frac{\sum_{j=1}^{p-1} \left( \frac{j}{P} + 0 \right)}{P} - 1 \right) = \frac{1}{P} \sum_{j=1}^{p-1} j = \frac{1}{P} \frac{P(P-1)}{2} = \frac{P-1}{2}.
\]

Conversely, GCI is zero if the alternative routes are evenly distributed among part types, so that \( \text{V}[1] = \text{V}[2] = \ldots = \text{V}[i] = \ldots = \text{V}[P] \) and \( \text{JRVTOT} = P \cdot \text{V}[i] \ \forall \ i \). In this case it results:
\[ \text{GCI} = \sum_{j=1}^{P} \left( \frac{j}{P} - \frac{1}{RV_{\text{TOT}}} \sum_{i=1}^{j} V[i] \right) = \sum_{j=1}^{P} \left( \frac{j}{P} - \frac{j \cdot V[i]}{P \cdot V[i]} \right) = 0. \]

Owing to these considerations, to obtain an index that ranges in the interval between zero and one (in correspondence to the worst and to the optimum condition), we can formally define GRV in the following way,

\[ \text{GRV} = 1 - \frac{2 \text{GCI}}{(P - 1)} \] (21)

6. Routing flexibility vector

In order to come up with a single measure of routing flexibility, a three-dimensional picture may be helpful to simultaneously illustrate all the different themes previously discussed. In doing so a global routing flexibility index (GRF) can be conceived as the modulus of a three-dimensional vector \( \vec{R} \) whose coordinates are GRE, GRR and GRV,

\[ \vec{R} = \text{GRE}\hat{i} + \text{GRR}\hat{j} + \text{GRV}\hat{k}, \] (22)

\[ \text{GRF} = \sqrt{\text{GRE}^2 + \text{GRR}^2 + \text{GRV}^2}, \] (23)

with values included between 0 and \( \sqrt{3} \).

Although GRF is more synthetic, it is also a bit less selective, since the threefold information concerning number, efficiency and distribution of alternative processing routes gets lost. In other words, GRF is not much relevant by itself. Clearly, what discriminates between the flexibility of two or more alternative systems is the degree of balance among the underlying contributions given by GRE, GRR and GRV.

To maintain the selectiveness of the information, an alternative approach is to consider the cosine similarity \( \psi \) between \( \vec{R} \) and the best conceivable flexibility vector \( \vec{R}_0 = 1\hat{i} + 1\hat{j} + 1\hat{k} \). This is a frequently adopted way to compare two vectors, by measuring the cosine of the angle \( \theta \) between them (Manning et al. 2008). In other words, the cosine similarity determines whether two vectors are pointing in roughly the same direction or not. Therefore, the bigger the angle \( \theta \) between \( \vec{R} \) and \( \vec{R}_0 \) (i.e. the lower the cosine similarity \( \psi \)) the worse the situation. The cosine similarity can be easily computed as shown by Eq. (24):

\[ \psi = \cos(\theta) = \left( \frac{\langle \vec{R}, \vec{R}_0 \rangle}{|\vec{R}| \cdot |\vec{R}_0|} \right) = \left( \frac{\text{GRE} + \text{GRR} + \text{GRV}}{\sqrt{\text{GRF}^2 + \text{GRF}^2 + \text{GRF}^2}} \right). \] (24)

Finally GRF and \( \psi \) can be conveniently combined in a single metric, with values in the range between zero and one as follows,

\[ \vartheta = \frac{\text{GRR}}{\sqrt{3}} \cdot \psi \] (25)

Such a framework, based on the above-mentioned indices, is an immediate visual tool to evaluate the flexibility of a manufacturing system and to provide useful indications for corrective actions. Actually, the target should be that to get a balance among the values of GRE, GRR and GRV, which should be as close as possible to one. Suppose for example that an apparently satisfying GRF value comes from the combination of a low GRV and a valuable GRR and GRE. This means that the manufacturing system features a potential machine routing flexibility, which is poorly exploited, since several highly efficient alternative routes are available for a few products only. In this case,
priority should be given to redesigning products, rather than keeping on investing in additional flexible machines.

7. A DEA based flexibility assessment method

In this section, an alternative DEA based approach is presented to combine GRE, GRR and GRV into a global routing flexibility score $\hat{\theta}$. As well known, assessing and comparing the efficiency of $N$ decision making units (DMU) (i.e. operating systems) can be vague and subjective in that it depends on the factors and on the basis for the selected comparison. If a DMU $j$ is characterized by $I$ inputs $x_{ij}$ and $K$ outputs $y_{kj}$, one could measure its efficiency $\hat{\theta}_j$ as follows,

$$
\hat{\theta}_j = \frac{\sum_{k=1}^{K} u_k y_{kj}}{\sum_{i=1}^{I} v_i x_{ij}},
$$

(26)

where $u_1, \ldots, u_K$ and $v_1, \ldots, v_I$ are weighting factors associated with the outputs and the inputs, respectively. Unfortunately, also Eq. (26) generates some problems, since $\hat{\theta}_j$ depends strongly on the adopted set of weights. At different weights, the efficiency value may undergo relevant variations and it becomes difficult to fix a single structure of weights that might be accepted by all the DMUs.

A DEA approach can solve this problem by evaluating the efficiency of each DMU, through the weights system that is the best for the DMU itself. This implies the solution of $N$ linear programming models to find the system of weights that allows the efficiency of each DMU to be maximized. To this aim models (26-27) are frequently adopted, which are known as the standard output oriented CCR models, expressed in primal and dual form, respectively. In the CCR models the subscript $0$ represents the DMU which is being evaluated and $s^{-}_i, s^{+}_k$ are slack variables. DMUs for which the optimal value $q^* = \phi^* \neq 1$ are inefficient, whereas, provided that $q^* = \phi^* = 1$, a DMU is said technically efficient if all the slack variables equal zero and is said weakly efficient if some slack variables are greater than zero. For a further discussion concerning these and similar models, the reader is referred to the clear work by Cooper et al. (2004).

\[ \begin{align*}
\min q & = \sum_{i=1}^{I} v_i x_{i0} \\
\text{subject to} & \\
\sum_{i=1}^{I} v_i x_{ij} - \sum_{k=1}^{K} u_k y_{kj} & \geq 0 \quad \forall j \in [1, N] \\
\sum_{k=1}^{K} u_k y_{k0} & = 1 \\
u_k, v_i & \geq 0 \quad \forall k \in [1, K], \forall i \in [1, I]
\end{align*} \]

(27)

\[ \begin{align*}
\max \phi & \\
\text{subject to} & \\
\sum_{j=1}^{N} x_{ij} \lambda_j + s^{-}_i & = x_{i0} \quad \forall i \in [1, I] \\
\sum_{j=1}^{N} y_{kj} \lambda_j - s^{+}_k & = \phi y_{k0} \quad \forall k \in [1, K] \\
\lambda_j & \geq 0 \quad \forall j \in [1, N]
\end{align*} \]

(28)
Unfortunately it is sometimes difficult to recover the explicit input–output relationship among the data, as required by the standard DEA models. This typically occurs when, as in the present case, one wants to combine a set of performance indicator into a global score. In these circumstances data sets are given without inputs (i.e. performance indicators are estimates of the goodness of the outputs), or the original input–output data cannot be easily recovered. As demonstrated by (Lovell & Pastor, 1999), a standard CCR model without explicit inputs and/or outputs cannot be used, since it would rate all DMUs as inefficient. However, a possible solution can be found in the landmark work by Thompson et al. (1986), who suggest using a single constant input CCR model. In the present case, this specific model assumes the dual form given in Eq. (29).

Note that in this case, an optimal value $\phi^* = 1$ suggests maximal flexibility, conversely an optimal value $\phi^* > 1$ suggests insufficient flexibility, since it is possible to expand all performance indicators by $100 \cdot (\phi^* - 1)\%$, without exceeding the best performance observed among the $N$ DMUs. In other words, the larger the value of $\phi^*$, the weaker the routing flexibility.

\[
\begin{align*}
\max & \quad \phi \\
\text{subject to} & \quad \sum_{j=1}^{N} \text{GRE}_j \lambda_j \geq \phi \text{GRE}_0 \\
& \quad \sum_{j=1}^{N} \text{GRR}_j \lambda_j \geq \phi \text{GRR}_0 \\
& \quad \sum_{j=1}^{N} \text{GRV}_j \lambda_j \geq \phi \text{GRV}_0 \\
& \quad \sum_{j=1}^{N} \lambda_j \leq 1 \\
& \quad \lambda_j \geq 0 \quad \forall j \in [1,N]
\end{align*}
\]

(29)

To conclude this section, it is useful to note that, as shown in Liu et al. (2010), after some manipulations the dual form of (29) turns into model (30).

\[
\begin{align*}
\max & \quad \hat{\theta}_0 = u_{\text{GRE}} \text{GRE}_0 + u_{\text{GRR}} \text{GRR}_0 + u_{\text{GRV}} \text{GRV}_0 \\
\text{subject to} & \quad u_{\text{GRE}} \text{GRE}_j + u_{\text{GRR}} \text{GRR}_j + u_{\text{GRV}} \text{GRV}_j \leq 1 \quad \forall j \in [1,N] \\
& \quad u_{\text{GRE}}, u_{\text{GRR}}, u_{\text{GRV}} \geq 0
\end{align*}
\]

(30)

Since the strong duality theorem assures that the optimal value $\phi^*$ coincides with the optimal value $\hat{\theta}^*$, model (30) shows that, in the proposed approach, the global routing flexibility index $\hat{\theta}$ is obtained as a weighed sum of the index GRE, GRR and GRV; where the weightings coefficients are the optimal values obtained by solving a CCR model.

8. Efficiency analysis of alternative routes

In the following part of the paper, a more detailed evaluation of the alternative routing efficiency index $\text{ARE}_{rp}$ is presented. To this aim, some critical aspects that may reduce the actual efficiency of alternative routes are analyzed in a more sophisticated way.
The covering degree concept

ARE\textsubscript{rp} does not take into account some problems that can be better explained through the introduction of the covering degree concept. Covering degree relates to the number of machines of the standard route that can breakdown, without compromising the possibility to access an alternative one. To get a better understanding, consider the situation detailed in Matrix 3.

\begin{table}[h]
\centering
\begin{tabular}{c|cccccc}
 & \textbf{p\textsubscript{1}} & \textbf{p\textsubscript{2}} & \textbf{p\textsubscript{3}} \\
\hline
m\textsubscript{1} & 1 & 0 & 0 & 1 & 1 \\
m\textsubscript{2} & 0 & 1 & 1 & 0 & 1 \\
m\textsubscript{3} & 0 & 1 & 0 & 1 & 0 & 1 \\
m\textsubscript{4} & 0 & 1 & 0 & 1 & 0 & 0 \\
m\textsubscript{5} & 1 & 0 & 1 & 1 & 0 & 0 \\
\hline
r\textsubscript{01} & r\textsubscript{11} & r\textsubscript{02} & r\textsubscript{12} & r\textsubscript{03} & r\textsubscript{13}
\end{tabular}
\caption{An example of Part-Machine Matrix}
\end{table}

In this case three items are processed by five machines and each item has two processing routes. Still, in terms of covering degree, \( r\textsubscript{11}, r\textsubscript{12}, r\textsubscript{13} \), behave in a different way. While it is possible to activate the alternative route \( r\textsubscript{11} \) regardless of machines failure along the standard route \( r\textsubscript{01} \) (i.e. \( m\textsubscript{1} \) or \( m\textsubscript{5} \)), this is not true either for \( r\textsubscript{12} \) or for \( r\textsubscript{13} \). Actually, if \( m\textsubscript{5} \) fails, we cannot use either the standard route or the alternative one for \( p\textsubscript{2} \), but also its alternative one cannot be used. At the opposite extreme, all the viable routes of \( p\textsubscript{3} \) are conditioned to the availability of both \( m\textsubscript{1} \) and \( m\textsubscript{2} \).

It is evident that the concept of covering degree largely influences the possibility of accessing alternative routes, so it should be included in the flexibility analysis. This can be done by means of the covering efficiency index \( CE\textsubscript{rp} \):

\[
CE\textsubscript{rp} = 1 - \frac{B\textsubscript{rp}}{M\textsubscript{0p}},
\]

where \( M\textsubscript{0p} \) is the number of machines used in the standard route and \( B\textsubscript{rp} \) is the number of machines used both in the standard \( r\textsubscript{0p} \) and in the alternative route \( r\textsubscript{rp} \) of the \( p\text{-th} \) part type.

The need to consider such an index is urged by the better suitability of a relative rather than an absolute measure. In fact, covering a route that requires several operating machines is much more difficult than covering a route made by a single machine. In the extreme case, if a product requires all the machines to be operating, a real alternative route cannot be established and the problem can only be solved through the adoption of machines, redundancy.

\( CE\textsubscript{rp} \) can be incorporated in the alternative route efficiency, in the following way:

\[
ARE\textsubscript{rp}\textsuperscript{C} = CE\textsubscript{rp} \cdot ARE\textsubscript{rp} = \left( 1 - \frac{B\textsubscript{rp}}{M\textsubscript{0p}} \right) \cdot \frac{RC\textsubscript{0p}}{RC\textsubscript{rp}}
\]

Note how the optimal solution \( ARE\textsubscript{rp}\textsuperscript{C} = 1 \) is obtained when no machines pertaining to the standard route are shared with that of the \( r\text{-th} \) alternative route (i.e. maximum degree of covering) and the processing costs of the two routes are identical.

Evidently, starting from \( ARE\textsubscript{rp}\textsuperscript{C} \), all the other flexibility indexes can be adjusted to take into account the covering degree concept.

8.2. Quality of routing

Product quality is another critical factor to discriminate between alternative routes. In relative terms, the quality of an alternative route \( QR\textsubscript{rp} \) can be evaluated as:
\[ Q_R = \frac{D_{rp}}{D_{op}}, \]  

(33)

where \( D_{op} \) and \( D_{rp} \) represent the percentage of defects of the standard and of the alternative route, respectively.

Also \( Q_R \) can be easily included as an additional weight in the computation of the routing efficiency index:

\[
AR_{EC} = Q_R \cdot CE_{rp} \cdot ARE = \frac{D_{rp}}{D_{op}} \cdot \left( 1 - \frac{B_{rp}}{M_{op}} \right) \cdot \frac{RC_{rp}}{RC_{op}}
\]

(34)

8.3. The implications of layout efficiency on material handling management

Several techniques for layout optimization rely on a flow matrix \( F = |f_{ij}| \), where \( f_{ij} \) is the frequency of items travelling between the \( i \)-th and the \( j \)-th machine. Typically, the flow matrix \( F \) is based on the standard routes and so, if an alternative route is subsequently deployed, the values of \( F \) could change making the layout less efficient. To deal with this possibility, it is convenient to measure the layout efficiency in terms of the efficiency of alternative routes, with respect to the existing layout. In many automated manufacturing environment, and especially in the case of FMSs, machines are arranged along a straight track with a material handling device moving jobs from one station to another. In this condition, one can assume, without loss of generality, that materials flow along the line from left to right. If the operations sequence of a job differs from the serial sequence of the machines, the job has to travel to the left (i.e. backward) to be processed and this reverse travel is referred to as backtracking. As noted by Byrne and Chutima (1997), the performance of the system can be exploited if the workloads between machines are balanced and the distances travelled by parts are kept to a minimum. Furthermore, Hassan (1994) and Kouvelis et al. (1995) stated that backtracking has a greater impact on modern manufacturing systems and so, layout should be studied to minimize the total backtracking distance of the material-handling device. An appropriate index to express the worsening performances of the material handling device can be obtained by measuring the increment of backtracking, as compared to the standard route. In this way the layout efficiency index \( LE \) can be computed as:

\[
LE = \frac{BT_{op}}{BT_{rp}},
\]

(35)

where \( BT_{op} \) and \( BT_{rp} \) represent the backtracking entity of the standard and of the \( r \)-th alternative route, respectively.

Again, also \( LE \) can be used as an additional weight in the computation of the routing efficiency index:

\[
AR_{LQC} = LE \cdot QR \cdot CE \cdot ARE = \frac{BT_{op}}{BT_{rp}} \cdot \frac{D_{rp}}{D_{op}} \cdot \left( 1 - \frac{B_{rp}}{M_{op}} \right) \cdot \frac{RC_{op}}{RC_{rp}}
\]

(36)

9. Numerical example

This section presents a numerical application concerning two alternative productive systems. It is assumed that ten machines are used to manufacture three products via the standard routes shown in Matrix 4.
The availability of each machine has also been included in the standard route matrix $S$, as an additional column $A$. To enhance routing flexibility two alternative configurations have been deployed by the process planner, as shown in the cost matrixes 5 and 6.

9.1. Routing Flexibility: a vectorial approach

To evaluate the goodness of both solutions, starting from the alternative route efficiency, the whole set of previously introduced metrics has been evaluated, without considering (for the sake of simplicity) either backtracking or the quality of the alternative routes.
Table 1
Routing flexibility metrics

<table>
<thead>
<tr>
<th></th>
<th>Configuration #1</th>
<th>Configuration #2</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>$r_{11}$</td>
<td>$r_{12}$</td>
</tr>
<tr>
<td>$A_{RE_{rp}}$</td>
<td>0.97</td>
<td>0.95</td>
</tr>
<tr>
<td>$CE_{rp} = 1 - \frac{B_{rp}}{M_{op}}$</td>
<td>0.50</td>
<td>0.67</td>
</tr>
<tr>
<td>$A_{RE_{rp}} \cdot CE_{rp}$</td>
<td>0.48</td>
<td>0.63</td>
</tr>
<tr>
<td>$SR_{rp} = \prod_{m=1}^{M} A_{m} x_{m0p}$</td>
<td>0.90</td>
<td>0.83</td>
</tr>
<tr>
<td>$SRC_{p} = \sum_{m=1}^{M} \frac{x_{m0p}}{M}$</td>
<td>0.60</td>
<td>0.60</td>
</tr>
<tr>
<td>$\alpha_{p} = \left( \frac{1}{1 - SRA_{p}} \right)^{SRC_{p}}$</td>
<td>4.1</td>
<td>2.92</td>
</tr>
<tr>
<td>$J_{RR_{p}} = 1 - \left( 1 - \frac{R_{p}}{TR_{p} - 1} \right)^{\alpha_{p}}$</td>
<td>0.42</td>
<td>0.57</td>
</tr>
<tr>
<td>$J_{RV_{p}} = \langle E_{p}, R_{p} \rangle$</td>
<td>0.20</td>
<td>0.315</td>
</tr>
</tbody>
</table>

Table 2
Aggregate routing flexibility metrics

<table>
<thead>
<tr>
<th></th>
<th>Configuration #1</th>
<th>Configuration #2</th>
</tr>
</thead>
<tbody>
<tr>
<td>$GRE = \sum_{p=1}^{P} J_{RAE_{p}}$</td>
<td>0.498</td>
<td>0.402</td>
</tr>
<tr>
<td>$GRR = \sum_{p=1}^{P} J_{RR_{p}}$</td>
<td>0.542</td>
<td>0.563</td>
</tr>
<tr>
<td>$GRV = 1 - \frac{2GCI}{P - 1}$</td>
<td>0.864</td>
<td>0.67</td>
</tr>
<tr>
<td>$GRF = \sqrt{GRE^2 + GRR^2 + GRV^2}$</td>
<td>1.135</td>
<td>0.963</td>
</tr>
<tr>
<td>$\psi = \cos \left( \frac{\langle \bar{R}, \bar{R}_0 \rangle}{</td>
<td>\bar{R}</td>
<td>}, \psi \right)$</td>
</tr>
<tr>
<td>$\vartheta = \frac{GRF}{\sqrt{3}} \cdot \psi$</td>
<td>0.634</td>
<td>0.545</td>
</tr>
</tbody>
</table>

Although both solutions have an equal number of alternative routes, the first one is significantly better than the second one because $\psi_1 \approx 0.634$ and $\psi_2 \approx 0.545$. The superiority of the first solution is evident also from the three dimensional representation of the routing flexibility vectors $\bar{R}_0, \bar{R}_1, \bar{R}_2$ shown in Fig. 3.

Fig. 3. Visual assessment of routing flexibility
Differences are mainly due to a reduction in the global routing efficiency (GRE\(_1 = 0.498\), GRE\(_2 = 0.4\)) and to an uneven distribution of the alternative routes (GRV\(_1 = 0.86\), GRV\(_2 = 0.67\)). The reasons of these results can be better explained considering that in the second configuration:

- \(p_2\) has three alternative routes, but, due to a poor covering degree (CE\(_{32} = 0.17\)), its routing versatility is low (JRV\(_2 = 0.19\));
- there is just an alternative (inefficient) route for \(p_3\), and this has a negative impact on its routing versatility (JRV\(_3 = 0.12\)). Specifically the standard route \(r_{03}\) is characterized by the lowest availability and by the lowest complexity and so, to assure production continuity, more than a single alternative route should be deployed.

Consequently, although the alternative solutions are comparable in terms of the number of alternative routes (i.e. GRR dimension), the first one is preferable, for it gets a significant advantage considering how evenly efficient routes are distributed among the manufactured items (i.e. GRE and GRV dimensions).

This example clearly underlines the fact that the number of available routes is not sufficient to discriminate between alternative solutions and it must be integrated with additional performance factors. It also underlines the utility to keep on separated levels the threefold information concerning the number, the efficiency and the distribution of alternative routes. As a matter of fact, what discriminates between the flexibility of two or more alternative systems is the degree of balance among the underlying contributions GRE, GRR and GRV.

9.2. Routing Flexibility: a DEA based approach

A similar result can be obtained using the output oriented CCR model (29) proposed in section 7. In this case, since there are two DMUs, using data of Table 2, the single constant input CCR model (29) assumes the following primal forms, where \(P_1\) and \(P_2\) refer to configuration #1 and #2, respectively.

\[
\begin{align*}
(P_1) \max \phi_i & \quad \text{subject to} \quad 0.498\lambda_1 + 0.402\lambda_2 \geq 0.498\phi_i \\
& \quad 0.542\lambda_1 + 0.563\lambda_2 \geq 0.542\phi_i \\
& \quad 0.864\lambda_1 + 0.670\lambda_2 \geq 0.864\phi_i \\
& \quad \lambda_1 + \lambda_2 \leq 1, \quad \lambda_1, \lambda_2 \geq 0.
\end{align*}
\]

\[
\begin{align*}
(P_2) \max \phi_2 & \quad \text{subject to} \quad 0.498\lambda_1 + 0.402\lambda_2 \geq 0.402\phi_2 \\
& \quad 0.542\lambda_1 + 0.563\lambda_2 \geq 0.563\phi_2 \\
& \quad 0.864\lambda_1 + 0.670\lambda_2 \geq 0.670\phi_2 \\
& \quad \lambda_1 + \lambda_2 \leq 1, \quad \lambda_1, \lambda_2 \geq 0.
\end{align*}
\]

Solving models \(P_1\) and \(P_2\) with the dedicated linear programming solver Lindo 6.1 (2002) yields \(\phi_1 = \phi_2 = 1\). Therefore, both DMUs appear as efficient, and the DEA model is not capable to discriminate between them. As one usually does in this case, to augment the discrimination power, the cross efficiency method can be used (Cooper et al., 2004). Briefly, this analysis is based on the definition of a cross efficiency matrix \(CE_{N \times N} = [\vartheta_{jk}]\) that is a square matrix (with as many rows and columns as there are units being compared), whose generic element \(\vartheta_{jk}\) represents the efficiency of the \(j\)-th DMU evaluated through the optimal weights structure for the \(k\)-th one. If DMU\(_j\) is efficient (i.e. \(\vartheta_{jj} = 1\)), but it exhibits a behavior specialized along a given dimension with respect to the other units, \(\vartheta_{jk}\) will be less than 1 for some value of \(k\). Therefore, an interesting discriminating index \(\tilde{\vartheta}_j\) can be obtained taking the average of the values in the \(k\)-th row:

\[
\tilde{\vartheta}_j = \frac{1}{N} \sum_{k=1}^{N} \vartheta_{jk}
\]

(37)

To compute this value, it is convenient to solve the dual problem (30), because this model makes explicit use of the coefficients \(u_k\) used to weight the performance indicators of each DMU.
Solving these models yields $\theta_{12} \approx 1$ and $\theta_{21} \approx 0.78$. Therefore $\tilde{\theta}_1 \approx (1+1)/2 = 1$ and $\tilde{\theta}_2 \approx (1 + 0.78)/2 = 0.89$ and configuration #1 results as the best option. A synthesis of all the results obtained with the DEA based approach is given in Table 3.

**Table 3**

<table>
<thead>
<tr>
<th>Configuration #1</th>
<th>Configuration #2</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\phi^*$</td>
<td>1</td>
</tr>
<tr>
<td>$\lambda_1$</td>
<td>1</td>
</tr>
<tr>
<td>$\lambda_2$</td>
<td>0</td>
</tr>
<tr>
<td>$\theta^*$</td>
<td>1</td>
</tr>
<tr>
<td>$u_{GRE}$</td>
<td>0</td>
</tr>
<tr>
<td>$u_{GRR}$</td>
<td>0</td>
</tr>
<tr>
<td>$u_{GRV}$</td>
<td>1.16</td>
</tr>
<tr>
<td>$\theta_{ij}$</td>
<td>1</td>
</tr>
<tr>
<td>$\theta$</td>
<td>1</td>
</tr>
</tbody>
</table>

Note that the values obtained with the vectorial and the DEA approaches are similar. Indeed, although their solutions differ (i.e. $\theta_{1} \approx 1$ and $\theta_{2} \approx 0.64$), this difference is due to the fact that $\theta_{1}$ expresses the flexibility of the first configuration in relative terms with the second one, whereas $\theta_{2}$ expresses the flexibility of the first configuration with respect to the ideal one. As a matter of fact, computing the percentage difference between $\theta_{1}$ and $\theta_{2}$ and between $\theta_{1}$ and $\theta_{2}$ one obtains a value of 14% and of 12%, respectively, which are sufficiently aligned values.

### 10. Validation through simulation analysis

To validate the proposed approach, the operating performance of both configuration #1 and #2 has been assessed via simulation, modelling the two manufacturing systems by means of the dedicated software Simul8 Professional (2006). Making reference to the definition given by Browne et al. (1984), which correlates routing flexibility with the ability to react to manufacturing problems and to assure production continuity, the systems have been compared in terms of their average throughput (i.e. the higher the throughput the more the system is considered to be flexible).

#### 10.1. Simulation models

For the sake of simplicity we considered the same costs (per unit of time) on each machine, that is $\bar{c}_m = \bar{c} = 1$ with $m = 1,\ldots,10$. In this way the processing time matrix $T$ coincides with the cost matrix $C$ and data of Matrix 4 can be directly taken as the average processing time of each machine. Furthermore, processing times are assumed to be normally distributed, with a standard deviation equal to 0.25 of the mean. In a similar way, machines availability (see the first column of Matrix 4) is reproduced using exponential and Erlang distributions to model the time between failures and the repair time, respectively. An example of the modelled systems is given in Fig. 4, which is relative to configuration #1 (a similar model is used for Configuration #2). Note that, to improve the readability of Fig. 4, the main route $r_{01}$ of product $p_1$ has been highlighted with respect to the other routes deployed in the system.
Most of the elements of Fig. 4 are self-explaining, but some ones (i.e. the dummy machines) deserve some further explanations. In particular, to facilitate the attainment of a steady state, and to limit the variability of the throughput, a CONWIP system has been adopted (Hoop & Spearman, 2000). CONWIP stands for constant work-in-process and designates a strategy that uses cards (or other visual control methods) to limit the work in process (WIP) that can accumulate in a manufacturing system. Each job is associated with a card for the whole duration of its manufacturing cycle. As soon as the job is processed by the last machine, its cards are released and can be associated with a new job. Since no job is allowed in the system without a card, the overall amount of WIP equals the number of available cards (Braglia et al., 2010). In the simulation model, this strategy is reproduced by means of four dummy (i.e. zero processing time) machines. Specifically the End machine releases CONWIP cards as soon as a product ends its manufacturing cycle, while machines P1, P2 and P3 check the availability of CONWIP cards before pushing a new job into the system. These dummy machines perform an additional task, since they are used to specify the route (i.e. principal or alternative) assigned to each job. To make this choice, three operating rules have been conceived and implemented. Specifically, after a preliminary check, that identifies the machines (if any) stopped due to failures, jobs are assigned to the remaining available routes based on the following operating rules:

- rule #1: assign the job to the route with the minimum WIP level;
- rule #2: assign the job to the route with the minimum cumulated processing time;
- rule #3: assign the job to the route whose bottleneck is characterized by the lower saturation.

10.2. Simulation results

Three systems were compared via simulation, namely base case (i.e. the rigid system with standard routes only), configuration #1 and configuration #2. Operating performance of each one was assessed in five different settings: (i-iii) a single product is manufactured, (iv) all products are manufactured and a homogeneous product mix (i.e. 1:1:1) must be respected, (v) all products are manufactured and any mix can be used, provided that each product accounts, at least, for the 15% of the overall
production. In each instance, the WIP (i.e. the number of CONWIP cards for each product) that maximises the throughput was determined using the optimization tool (OptQuest) embedded in the simulation environment. Finally, to compare the results, the following scheme was used:

- Output parameter: Daily Throughput [jobs/day];
- Total Simulation time: 48000 [min], that is 100 working days;
- Warm Up period: 7200 [min], that is 15 working days;
- Routing Selection: use of rule #3, since a preliminary test demonstrated the superiority of this rule over the other ones;
- Execution scheme: use 30 simulation runs for each analyzed configuration;
- Result analysis: evaluation of the confidence interval for the average Daily Throughput (DT) at a 95% confidence level.

Note that, due to the use of the CONWIP strategy, a warm up period of fifteen working days was considered sufficient (to reach the steady state), as visually demonstrated by Fig. 5, which shows the evolution of the average DT during a simulation run.

The obtained results, listed in Table 4, demonstrate that, as one could have reasonably guessed, both systems are definitely more flexible that the basic (rigid) configuration. From the data of Table 4 one can also note that Configuration #1 is the best alternative with respect to products p2 and p3, but not with respect to p3. Therefore neither solution dominates the other one; this is the reason why, under the constraint of a homogeneous product mix, both configurations perform in a similar way (i.e. there is no statistical evidence that one system is better than the other one).

**Table 4**
Result of the simulation model

<table>
<thead>
<tr>
<th></th>
<th>Base Case</th>
<th>Configuration #1</th>
<th>Configuration #2</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>WIP</strong></td>
<td><strong>DT [jobs/day]</strong></td>
<td><strong>DT [jobs/day]</strong></td>
<td><strong>DT [jobs/day]</strong></td>
</tr>
<tr>
<td><strong>WIP</strong></td>
<td><strong>Low 95%</strong></td>
<td><strong>Avg. 95%</strong></td>
<td><strong>High 95%</strong></td>
</tr>
<tr>
<td>Single product p1</td>
<td>12</td>
<td>66.8</td>
<td>68.2</td>
</tr>
<tr>
<td>Single product p2</td>
<td>12</td>
<td>91.4</td>
<td>93.4</td>
</tr>
<tr>
<td>Single product p3</td>
<td>6</td>
<td>88.5</td>
<td>90.4</td>
</tr>
<tr>
<td>All products mix (1:1:1)</td>
<td>47</td>
<td>116</td>
<td>118</td>
</tr>
<tr>
<td>All products Max DT</td>
<td>80</td>
<td>138.4</td>
<td>142</td>
</tr>
</tbody>
</table>
However, following an approach similar to the DEA one, if we relax the constraints on the homogeneous product mix, and let each system adopt the product mix that maximises its global performance, we reach a clearer situation. Indeed, in this unconstrained setting, Configuration #1 and #2, obtain an average DT of $217 \pm 8$ and $200 \pm 5$ [jobs/day], respectively, with a percentage difference of 8.5%. Furthermore, since the DT confidence intervals do not intersect, the superiority of Configuration #1 over Configuration #2 is statistically significant. These results evidently agree with the conclusions obtained in the previous sections of the paper and empirically confirmation of the validity of both approaches presented in the paper.

11. Conclusions and remarks

In this paper, a theoretical framework and a set of indexes to evaluate several aspects of routing flexibility have been illustrated. The proposed indexes are intended to measure routing flexibility, by taking into account the number of viable routes and other issues that may compromise the efficiency of the system under analysis. A major peculiarity of the proposed measurement method consists in its flexibility, since a variety of aspects, which could potentially affect the overall effectiveness of alternative routes, are taken into consideration. This allows one to freely decide the extent and the level of details of the indexes included in the flexibility analysis.

Finally, to help practitioners in the planning of the right corrective actions, a graphical and a DEA based approach have also been presented and explained through a meaningful numerical example supported by a simulation analysis.

Acknowledgments

The authors are grateful to the anonymous referee(s) and to the editor of the journal for their valuable comments. Thanks to their precise and detailed suggestions, both the structure and the technical value of the paper have increased.

References


