Journal of Applied Statistics
Publication details, including instructions for authors and subscription information:
http://www.tandfonline.com/loi/cjas20

Exploring the characteristics of rotating electric machines with factor analysis

Carlos Maté [a] & Rafael Calderón [a]
[a] Universidad Pontificia Comillas , Madrid, Spain
Published online: 02 Aug 2010.

To cite this article: Carlos Maté & Rafael Calderón (2000) Exploring the characteristics of rotating electric machines with factor analysis, Journal of Applied Statistics, 27:8, 991-1006, DOI: 10.1080/02664760050173319

To link to this article: http://dx.doi.org/10.1080/02664760050173319

Journal of Applied Statistics, Vol. 27, No. 8, 2000, 991–1006

Exploring the characteristics of rotating electric machines with factor analysis

CARLOS MATÉ & RAFAEL CALDERÓN, *Universidad Pontificia Comillas, Madrid, Spain*

ABSTRACT *Applications of multivariate statistics in engineering are hard to find, apart from those in quality control. However, we think that further insight into some technological cases may be gained by using adequate multivariate analysis tools. In this paper, we propose a review of the key parameters of rotating electric machines with factor analysis. This statistical technique allows not only the reduction of the dimension of the case we are analysing, but also reveals subtle relationships between the variables under study. We show an application of this methodology by studying the interrelations between the key variables in an electric machine, in this case the squirrel-cage induction motor. Through a step-by-step presentation of the case study, we deal with some of the topics an applied researcher may face, such as the rotation of the original factors, the extraction of higher-order factors and the development of the exploratory model. As a result, we present a worthwhile framework to both confirm our previous knowledge and capture unexplored facts. Moreover, it may provide a new approach to describing and understanding the design, performance and operating characteristics of these machines.*

1 Introduction

Applications of multivariate statistical methods are widespread in some areas, such as psychology, economics, business, marketing, chemistry, biology and social sciences. The multivariate technique of factor analysis is particularly suitable for analysing the patterns of complex and multidimensional relationships in a large set of variables and for deciding the possibility of reducing that information to a smaller number of factors or latent variables. This is why it has found increasing use during the last decade in these areas. Moreover, applications in new fields,

Correspondence: Carlos Maté, Departamento de Organización Industrial, Escuela Técnica Superior de Ingeniería—ICAI, Universidad Pontificia Comillas, 28015, Madrid, Spain.

ISSN 0266-4763 print; 1360-0532 online/00/080991-16

such as technological ones where this technique was not usually applied, have emerged.

Goronzy (1969) provides a classical and interesting example of the factor analysis methodology in management, an area close to some issues arising in the engineering sciences. Eom & Farris (1996), Johannsen (1996) and Shin & Kim (1997) develop new ways of using this method in the same area, but in more modern fields, such as information science and quality control. Khots *et al.* (1994) and Vinzanekar *et al.* (1996) are references showing the usefulness and benefits of factor analysis as a valuable tool in engineering developments. In this paper we present another application of this multivariate method in a technological area—electric machines.

When an engineer requests information about any electric machine, e.g. an induction motor from a manufacturer, she or he receives a catalogue or list with the main characteristics of the device. The key parameters that most manufacturers employ to describe their machines are: rated power, speed, efficiency, power factor, rated current, locked-rotor current, torque, locked-rotor torque, inertia and weight. Thus, a scenario with several interrelated variables arises making the application of factor analysis feasible.

The physical modelling of these machines has been fully accomplished and simple models, e.g. equivalent circuits, provide an easy way to get fairly accurate results. However, one could wonder whether an alternative analysis, which is focused on the statistical information contained in the variables describing the machine operation and its characteristics, is adequate.

We propose factor analysis as a statistical approach to gain further insight into the design and operation of these devices and to analyse what manufacturers offer to the engineer. In this paper, we have employed this methodology using the key variables and parameters of several groups of induction motors from different manufacturers. Section 2 deals with a brief description of the induction machine. Section 3 is devoted to a concise review of factor analysis, pointing out the main issues that arise when this methodology is considered. In Section 4 we will outline the data sampling process. We present the way to apply this method to rotating electric machines in Section 5, where special emphasis will be given to the adequacy of the sample to factor analysis, the number of factors or components to be retained, the rotational methods to be applied and the extraction of higher-order factors. Section 6 concerns the interpretation of the factors. We will also propose a framework for the exploratory model and its robustness and stability will be commented on.

In the final section, the conclusions of this research will be stated and new proposals on how to advance in the statistical knowledge of electric machines will be put forward.

2 The induction machine. Key variables

These machines are suitable for driving industrial processes, moving a lift or a cargo-ship, or pumping water, so manufacturers offer a wide range in power and speed. We can find two well-differentiated parts in the machine, separated by the air gap. The rotor, which is the moving part, and a firm mechanical structure called stator, which is normally the outer frame of the machine. Usually, these are classified according to the constructive form of the rotor, the squirrel-cage type or the wound-rotor type.

In the motor mode, electric energy must be supplied by a three-phase AC source.

The current is carried by the stator conductors, inducing voltages in the rotor windings. These voltages produce currents that interact with the rotating magnetic field, generating the electromechanical torque that causes the rotating motion. It can be shown that the induced magnetic field is rotating with a speed identical to the speed of the supply signal and that its magnitude is constant. However, the rotor does not revolve at the network fixed speed (*Ws*) because of the bounds set by the physical laws that rule the operation. Thus, this machine is called either an 'asynchronous' or an induction machine, because it turns at a slightly lower speed.

To know more about operation modes, physical principles and the design of any of these machines see, for example, Nasar (1987), Sen (1989) or Hamdi (1994).

Some of the key parameters manufacturers use to describe induction motors are:

- power (*P*): the mechanical output power that the machine delivers to the axis when it operates in motor mode at steady-state conditions.
- speed (*W*): the speed of the rotor measured in rpm (revolutions per minute), under normal operation mode.
- efficiency (*E*): the ratio, in percentage, of the input power that is delivered in the output.
- power factor (*PF*): the quotient between the active power and the total power.
- current (*I*): the rated current absorbed from the electric source by the machine if it operates in motor mode, or vice versa in the generator mode.
- locked-rotor current (*ILK*): the intensity of the current when the machine is starting. This intensity is greater than the one in permanent operation mode, so it is usually expressed as the quotient between it and the steady-state current.
- torque (*M*): a measure of the torque the machine delivers under normal operation conditions.
- locked-rotor torque (*MLK*): a measure of the torque the machine delivers at starting. It is usually expressed as a per unit value of the rated torque.
- breakdown torque (*MBD*): the maximum torque available. It is also expressed as a per unit value of the rated torque.
- inertia (J): the inertia of the rotating member. It indicates the reaction velocity of the rotor speed when a change in operation conditions occurs.
- weight (*WG*): the weight of the machine.

Moreover, other parameters may be calculated from those mentioned above to acquire a more complete description of the machine. For example:

- slip (*S*): the quotient between the synchronous speed minus the rotor speed and the synchronous speed. It describes how far behind the rotor speed is from the speed of the rotating synchronous field, which is fixed by the network.
- slope of the *M*-*s* curve (*M_S*): it is estimated from the torque (*M*) and the slip (*S*) values. Figure 1 shows that the *M_S* curve is almost linear when approaching zero. As the slip is usually close to zero, the slope of the curve in this region may be approximated by using the quotient between the torque and the slip.

A concise but accurate description of the machine operation should be completed with two widely used performance characteristics: the speed-torque and the speed-current curves shown in Fig. 1.

In the motoring region, the sign of the torque is positive because the machine develops it. Once the machine is started, the locked-rotor current flows through

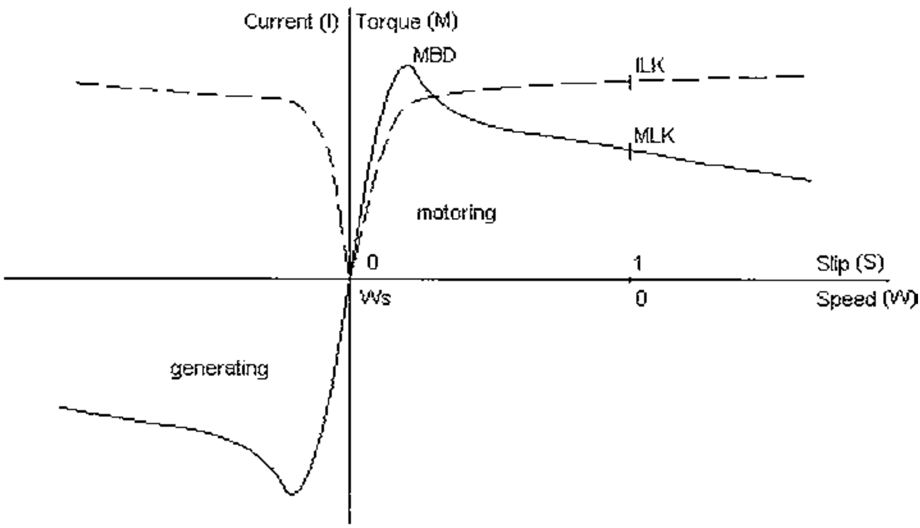

FIG. 1. Torque-speed and current-speed (dashed line) curves.

the stator and generates the locked-rotor torque. When the speed increases, the current decreases and the torque increases until it reaches a peak, which is called the breakdown torque. After that, the torque decreases and the relationship between torque and speed is almost linear in this zone. This is normally the steady-state operation zone, so the slope of the curve in this zone is a parameter to keep in mind in our study.

In the generating region, the sign of the torque is negative because we have to apply a mechanical torque to the machine in order to generate electric energy. The behaviour of the variables is similar to that observed in the motoring region.

3 An overview of factor analysis

Factor analysis is concerned with the identification of an underlying structure within a set of observed variables. This method was designed to simplify the complicated interrelations between a large number of observed or manifest variables, searching for certain unobservable common factors unrelated to each other. Thus, the aim of this technique is to find the relevant characteristics of the original set of variables through the construction of a smaller number of new common factors or latent variables with a minimum loss of information. Generally, this implies a reduction in the data dimensionality, which may be needed for further multivariate analysis such as regression or classification.

3.1 The exploratory factor analysis model

According to Johnson & Wichern (1998), Jobson (1992) and Basilevsky (1994), the general model of factor analysis has three sets of variables: a set of p observed variables $X_1, \dots X_p$ with mean vector μ and covariance matrix Σ; a set of m unobserved variables called common factors $F_1, \dots, F_m$ where $m \leqslant p$, and a set of p unique but unobserved factors $e_1, \dots, e_p$. The model is given by the p equations

$$X_1 - \mu_1 = l_{11}F_1 + l_{12}F_2 + \ldots + l_{1m}F_m + e_1$$

$$X_2 - \mu_2 = l_{21}F_1 + l_{22}F_2 + \ldots + l_{2m}F_m + e_2$$

$$\begin{matrix} \circ \\ \circ \\ \circ \end{matrix}$$

$$X_p - \mu_p = l_{p1}F_1 + l_{p2}F_2 + \ldots + l_{pm}F_m + e_p$$

or in matrix notation,

$$\underset{(p\times 1)}{\mathbf{X} - \boldsymbol{\mu}} = \underset{(p\times m)(m\times 1)}{\mathbf{LF}} + \underset{(p\times 1)}{\mathbf{e}}$$

The coefficient l_{ij} accompanying to F_j in the linear combination that describes $X_i - \mu_i$ is called the loading of the *i*th variable on the *j*th common factor, so that the matrix L is the matrix of factor loadings or the factor pattern matrix.

In this model, it is assumed that

$$E(\mathbf{F}) = \mathbf{0},\ Cov(\mathbf{F}) = \mathbf{I}$$

$$E(\mathbf{e}) = \mathbf{0},\ Cov(\mathbf{e}) = \boldsymbol{\Psi}$$

$$E(\mathbf{e}\cdot\mathbf{F}) = \mathbf{0}$$

Then, it is shown that

$$\Sigma = \mathbf{LL}' + \boldsymbol{\Psi}$$

Hence,

$$Var(X_i) = \sum_{j=1}^{m} l_{ij}^2 + \Psi_i;\ \mathrm{i} = 1, \ldots, p$$

$$Cov(X_i, X_k) = \sum_{j=1}^{m} l_{ij}l_{kj}$$

This provides a decomposition of the variance of the *i*th variable into two parts, the first part $\sum_{j=1}^{m} l_{ij}^2$ denotes the portion of $Var(X_i)$ related to the common factors and it is called the *i*th communality. The second term Ψ_i is the unique variance, which is unrelated to the common factors.

3.2 The factor analysis methodology with data

From the n observations on the above p variables, noted by $\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_n$, the factor analysis method tries to find a small number of factors that adequately represent the data according to the linear model discussed previously. Noting the sample covariance matrix by **S**, we have an estimation of the unknown population covariance matrix Σ. If the off-diagonal elements of **S** are small or those of the sample correlation matrix **R** essentially zero, the variables are not related and factor analysis would not be adequate. Otherwise, a factor model can be developed and the first problem is to estimate the matrices **L** and $\boldsymbol{\Psi}$.

Methods giving solutions to this problem are principal components analysis (PCA), principal factors (PF), maximum likelihood factoring, alpha factoring, unweighted least square factoring and image factoring. The most commonly used are PCA and PF. As stated by Tabachnick & Fidell (1996), it is quite common to employ PCA as a preliminary extraction technique, followed by one or more of the other procedures, perhaps varying the number of factors, the communality estimates, and the rotational methods in each run. The basic difference between PCA and PF is that the former focuses on variance and the latter on covariance (communality). The goal of PCA is to extract maximum variance from a set of data with a few orthogonal components, while the aim of PF is to reproduce the correlation matrix with a few orthogonal factors. Velicer & Jackson (1990a, b) compare the two methods. Zwick & Velicer (1986) analyse different rules to choose the number of components that are to be retained.

Once the factors have been extracted they are usually transformed by a rotation of the factor matrix to help the researcher in interpreting them. In orthogonal rotation, the factors remain uncorrelated, becoming easier to interpret, while in oblique rotation the factors may be correlated, taking into account possible relationships between them. Hair *et al.* (1998) state that when summarizing a large number of variables into a smaller number of uncorrelated variables for subsequent use, such as in prediction techniques, orthogonal rotation is preferred, whereas oblique rotation is more suitable when the objective is to extract some meaningful grouping while permitting the factors to be correlated. There are several techniques to perform orthogonal rotation, such as varimax, quartimax or equimax; and for oblique rotation, direct oblimin and quartimin, among others. Harman (1976) and Gorsuch (1983) describe the numerous procedures for factor extraction and rotation.

At this stage, we could think that since the induction motor is an 'electromechanical' device, we would extract an electric factor and a mechanical factor. We could also think, at first, that there is a certain degree of dependence between the two factors, so the factors would not be orthogonal. Note this hypothesis is more flexible than supposing the factors to be orthogonal.

Throughout the next sections, the results will be computed using the SPSS software. There are, however, a variety of software packages and we do not wish to give the impression that we endorse one over any other.

4 Data

When selecting the sample, one must choose its appropriate size carefully. Hair *et al.* (1998) indicate that more than 100 cases, as a rule of thumb, are needed to extract significant conclusions from the sample. It is also convenient to have a high number of cases per variable, at least 10. The samples we have worked with in our study were extracted from different manufacturer catalogues. In this paper, we present the results we have found by analysing a group of squirrel-cage motors from a catalogue published by Siemens (1988).

The sample consists of 134 cases with no missing values and 13 variables; thus, it is acceptable according to the rules found in the previous paragraph. Nevertheless, we performed some exploratory data analysis before applying factor analysis. Outliers were detected through univariate and multivariate methods, but we have retained them because they represent a segment of the population providing generalizability in the final model. The price paid for this is the possible lack of invariance and replicability of the results obtained with other samples.

In the original data, some variables were expressed in ratios instead of in their nominal values. Although we could have estimated the nominal values, we decided that the ratios were more appropriate for measuring the variables. Our concern is to capture the differences and variability between the cases, and the nominal values

TABLE 1. Correlation matrix of the variables

	E	I	ILK	J	M	M_S	MBD
E	1.00000						
I	0.53223	1.00000					
ILK	0.90605	0.47011	1.00000				
J	0.41524	0.89030	0.35289	1.00000			
M	0.49502	0.88562	0.40796	0.97623	1.00000		
M_S	0.44387	0.96174	0.38971	0.96992	0.94589	1.00000	
MBD	0.63461	0.17121	0.82751	0.02174	0.04575	0.08528	1.00000
MLK	0.34258	−0.19824	0.45100	−0.12334	−0.06942	−0.16979	0.47782
P	0.52360	0.99894	0.46545	0.87287	0.86428	0.95288	0.17676
PF	0.85051	0.50287	0.87653	0.31063	0.35824	0.38103	0.66222
S	−0.97238	−0.51094	−0.91874	−0.39313	−0.46612	−0.42412	−0.66237
W	0.20673	0.11117	0.43356	−0.09547	−0.14243	−0.00154	0.52030
WG	0.57056	0.97967	0.48118	0.92951	0.94303	0.96153	0.13726
	MLK	P	PF	S	W	WG	
MLK	1.00000						
P	−0.21130	1.00000					
PF	0.30939	0.50675	1.00000				
S	−0.35185	−0.50347	−0.85668	1.00000			
W	0.04517	0.13319	0.58786	−0.29322	1.00000		
WG	−0.15786	0.97093	0.49444	−0.54390	0.02595	1.00000	

are affected by the base magnitude, which could distort the results. Furthermore, different analyses using only the nominal values were performed with poor results, validating our decision.

5 Extraction of Factors

5.1 Preliminary tests

Before carrying out factor analysis, one must be sure about the adequacy of the data. Remember that it is a technique reflecting interdependence, so the higher the interrelation between the variables, the more accurate the conclusions obtained. Table 1 shows the correlation matrix and, as a preliminary analysis, the correlation between the variables is satisfactory, following the standard statistical practices.

On the other hand, both the Bartlett test and the KMO test confirm that the data correlation structure is suitable for factor analysis (Table 2). Moreover, an inspection of the measures of sampling adequacy (MSA) reports the same conclusion for each variable. According to Hair *et al.* (1998), MSA must be higher than 0.5 and there is only one variable, the speed, below 0.5 but close to it. In spite of this value, we decided to keep the variable in the analysis, although we must consider this decision when interpreting the results.

5.2 Number of factors

A set of criteria can be used as guidelines to determine the proper number of factors that should be extracted, see Zwick & Velicer (1986). The most widely employed criteria are the scree plot and the latent root criterion. In the case at hand, these criteria indicate it is feasible to reduce the variables to three factors, see Fig. 2 and Table 3.

TABLE 2. Measures of the adequacy of factor analysis to the sample

Kaiser-Meyer-Olkin Measure of Sampling Adequacy = 0.77340
Bartlett Test of Sphericity = 4471.2937, Significance = 0.00000
Anti-image Correlation Matrix:

	E	I	ILK	J	M	M_S	MBD	MLK	P	PF	S	W
E	0.75041											
I	−0.22079	0.73337										
ILK	−0.23042	−0.01576	0.86207									
J	−0.12747	0.35261	−0.10903	0.77557								
M	0.23032	−0.75781	−0.12176	−0.39635	0.75817							
M_S	0.11194	−0.22785	0.13231	−0.80873	−0.01339	0.81675						
MBD	−0.04509	0.00215	−0.69276	0.10577	0.08106	−0.12835	0.79245					
MLK	0.18863	−0.08651	−0.20351	−0.02039	−0.01043	−0.01720	−0.12827	0.73577				
P	0.21979	−0.99171	0.00032	−0.27784	0.79463	0.11966	−0.00006	0.09236	0.73894			
PF	−0.51859	0.27760	−0.16742	0.23580	−0.22819	−0.11643	0.29532	−0.30435	−0.27440	0.75830		
S	0.74127	−0.06302	0.19919	−0.12787	0.04560	0.12336	−0.13998	0.04924	0.04706	−0.20230	0.84936	
W	0.65209	−0.21356	−0.17749	−0.28680	0.30318	0.15680	−0.17093	0.33060	0.20029	−0.73012	0.31793	0.39485
WG	−0.01302	−0.32228	0.14604	−0.45500	−0.16738	0.60762	−0.00345	0.13692	0.21270	−0.15401	0.10544	0.10650
	WG											
WG	0.87202											

Measures of Sampling Adequacy (MSA) are printed on the diagonal.

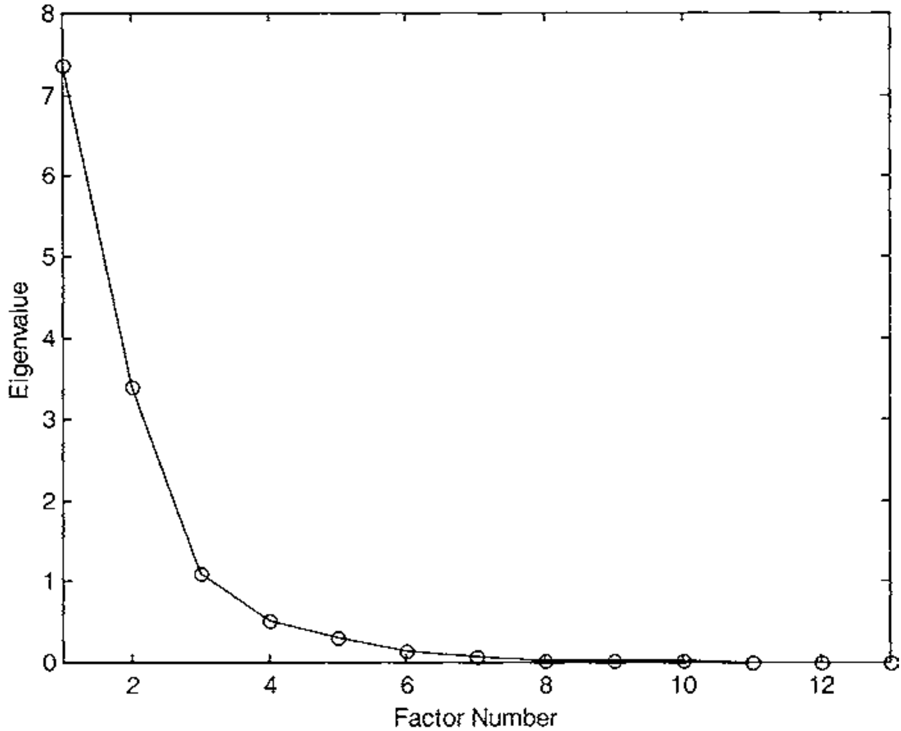

FIG. 2. Factor scree plot.

TABLE 3. Eigenvalues

Variable	Communality	*	Factor	Eigenvalue	% of Var	Cum %
E	1.00000	*	1	7.35962	56.6	56.6
I	1.00000	*	2	3.40760	26.2	82.8
ILK	1.00000	*	3	1.10164	8.5	91.3
J	1.00000	*	4	0.52127	4.0	95.3
M	1.00000	*	5	0.31098	2.4	97.7
M_S	1.00000	*	6	0.14584	1.1	98.8
MBD	1.00000	*	7	0.06430	0.5	99.3
MLK	1.00000	*	8	0.03274	0.3	99.6
P	1.00000	*	9	0.02799	0.2	99.8
PF	1.00000	*	10	0.01743	0.1	99.9
S	1.00000	*	11	0.00801	0.1	100.0
W	1.00000	*	12	0.00252	0	100.0
WG	1.00000	*	13	0.00005	0	100.0

These methods have been severely criticized, and other methods have been proposed, such as Velicer's or Horn's parallel analysis. A parallel analysis of the sample, following Glorfeld (1995), also concluded with three as the number of factors to be extracted.

However, this solution should be checked by extracting two or three factors more before assessing the final solution. This procedure suggests a heuristic approach in determining the number of factors that should be retained, generally leading to a more valuable and better-defined structure of the solution.

5.3 Extraction

Once the communalities were computed by a principal components extraction, we decided to continue with this method because all communalities were close to 1.0 and the differences with other methods would not be significant. This decision is based on the comments of Gorsuch (1983), who suggests that all the extraction procedures lead to similar results as the communalities approach 1.0. PCA offers other advantages including an exact determination of the factor scores and easier, safer computing than other methods in which the communalities are estimated and there may be computing problems, e.g. some communalities greater than 1.0.

After trying different numbers of factors, five were extracted and the factor matrix was rotated for a better understanding of the meaning of the factors.

5.4 Rotation

An orthogonal rotation was carried out in accordance with the varimax criterion. Bearing in mind the sample size and following Hair *et al.* (1998), only loadings higher than 0.5 were acceptable. From Table 4, the variables inertia (J), current (I), slope of the M-s curve (M_S), torque (M), weight (WG) and power (P) load on the first factor; while power factor (PF), locked-rotor current (ILK), efficiency (E) and slip (S) load on the second one. Each of the other three factors each load on a single variable. The breakdown torque (MBD) presents similar loadings in the second and fifth factors, but we retained it in the fifth factor due to the results obtained in other samples.

An oblique rotation was tried to confirm the distribution of the variables in the factors and to compute the correlation between them. A non-orthogonal rotation is more flexible than an orthogonal one because it does not constrain the factors to be uncorrelated. It may also help in determining the appropriateness of the orthogonal rotation: if the coefficients of the factor correlation matrix approach 0, the factors have a low correlation and the orthogonal rotation is adequate. Table 5 gives the new matrix with an oblimin rotation, showing nearly the same structure as that corresponding to the above-mentioned method. The main difference is that the breakdown torque (MBD) clearly loads on the second factor. There is another

TABLE 4. Varimax rotated factor matrix. (Salient loadings are printed in bold type and those lower than 0.4 have not been printed)

	Factor 1 (Fact1_1)	Factor 2 (Fact2_1)	Factor 3 (Fact3_1)	Factor 4 (Fact4_1)	Factor 5 (Fact5_1)
E		**0.93540**			
I	**0.93314**				
ILK		**0.84356**			
J	**0.97035**				
M	**0.94582**				
M_S	**0.98255**				
MBD		0.62605			0.67072
MLK				**0.91674**	
P	**0.92265**				
PF		**0.81847**	0.45516		
S		**−0.92848**			
W			**0.95222**		
WG	**0.94153**				

TABLE 5. Oblimin rotation

Structure Matrix:

	Factor 1 (Fact1_1)	Factor 2 (Fact2_1)	Factor 3 (Fact3_1)	Factor 4 (Fact4_1)	Factor 5 (Fact5_1)
E	0.49442	0.61244			**−0.98634**
I	**0.97218**				−0.52828
ILK	0.42641	0.81559		0.42124	**−0.92092**
J	**0.97002**				
M	**0.96440**				−0.46487
M.S	**0.99204**				−0.42491
MBD		**0.99367**	0.46331	0.43126	−0.63740
MLK		0.43787		**0.99359**	
P	**0.96184**				−0.52247
PF	0.41968	0.60934	0.60906		**−0.90958**
S	−0.47023	−0.64156			**0.98194**
W		0.48911	**0.98846**		
WG	**0.98555**				−0.55900

Factor Correlation Matrix:

	Factor 1	Factor 2	Factor 3	Factor 4	Factor 5
Factor 1	1.00000				
Factor 2	0.10220	1.00000			
Factor 3	0.01656	0.42474	1.00000		
Factor 4	−0.14926	0.39070	−0.02124	1.00000	
Factor 5	−0.47670	−0.60852	−0.29735	−0.29298	1.00000

difference concerning the output of the program: a swap of the label ‘Factor 2’ for the label ‘Factor 5’, but we have preferred to keep the output as is. A change of sign can also be noted in the coefficients of the last column.

In the new factor correlation matrix shown in Table 5, the highest coefficient corresponds to the value of the correlation between the second and the fifth factor. The closeness of the values of the loadings associated with the breakdown torque (*MBD*) on the second and fifth factors of the orthogonal rotation in Table 4 is therefore confirmed.

It is worth noting that the loadings in the oblique rotation are higher than in the orthogonal one, leading to a better distinction of the factors.

5.5 Higher-order factors

If the starting point for the extraction of factors is the correlation matrix of the variables and we have estimated the factor correlation matrix, a question may emerge about applying factor analysis to the extracted factors. This methodology is recommended by Gorsuch (1983), who states that ‘the higher-order factors should be extracted and examined so that the investigator may gain the fullest possible understanding of the data’.

In the following sections, for the sake of simplicity, Fact$i$_j refers to the factor i belonging to the set of jth-order factors, with $i, j \in \mathbb{N}$.

The extraction of second-order factors was accomplished following the same steps we have described previously for the extraction of the first-order factors. From Table 6, both the KMO test and the Bartlett test indicate the factor analysis technique is adequate. However, it should be noted that the KMO test has decreased, but it is still good enough to carry out the analysis. Accordingly, three factors were retained using the number of eigenvalues greater than 1.0 and the cumulative percentage of variance as guidelines. Nevertheless, other solutions with a different number of factors were tried before reaching the final solution. Once the factor matrix has been rotated following the varimax criterion, it can be clearly determined that the first-order factors load on the second-order ones. However, the first-order factor Fact2_1 loads in a very similar way on Fact1_2 and Fact3_2. On the other hand, the second-order factor Fact2_2 is clearly loaded on Fact1_1 and Fact5_1, while the other two second-order factors are not well defined due to the loadings of Fact2_1.

An oblique rotation with the oblimin criterion was performed, trying to solve the indeterminacy that factor Fact2_1 leads to in the orthogonal rotation, but there were no significant differences between the two rotations. This situation will be discussed in the next section with regard to the meaning of the variables.

We might continue extracting higher-order factors, but this may be more appropriate with a larger set of factors. In this case, there are only three second-order factors and only one or two third-order factors may be expected, which would not contribute to a better understanding of the underlying structure of the original variables.

6 Interpretation of the factors

6.1 Suggested structure

First, the meaning and labelling of the first-order factors will be considered. It is important to remember that, in section 5.4, five factors were identified from the original set of variables. Therefore, the analysis will be carried out following the oblique rotation in Table 5.

Fact1_1: this factor is loaded on the variables current (I), inertia (J), torque (M), slope of the M-s curve (M_S), power (P), and weight (WG). This factor is termed as 'size', because it is an index of the electric, mechanical and physical size of the machine. The higher the value of the inertia, the weight or the output power, the larger is the 'size' of the machine.

Fact2_1: this factor is only loaded on the variable breakdown torque (MBD). This result must be analysed with care. A possible interpretation of this single loading is that the breakdown torque is not highly related to the other factors and it could represent a degree of freedom in the design stage. However, Fact2_1 is correlated with Fact5_1, Fact3_1 and Fact4_1 (Table 5).

Fact3_1: this factor is only loaded on the variable speed (W). Although the same considerations as mentioned above may hold, it should be noted that the MSA value of the speed was the lowest, but it was decided to continue the analyses keeping the speed in the original set. A low value in the MSA may indicate a poor relation of the variable with the others. Hence, the speed could be considered as a nearly independent variable. This means that, for a given speed, the manufacturer could vary other characteristics of the machine within a broad range, and vice versa.

Fact4_1: this factor is only loaded on the variable locked-rotor torque (MLK). This is the same as in the two previous factors and analogous conclusions may be reached.

Fact5_1: this factor is loaded on the efficiency (E), locked-rotor current (ILK), power factor (*PF*) and slip (*S*) variable. The positive loading in slip and the negative loading in efficiency seem to underline that the higher efficiency of the machine is associated with lower slip. This may be attributed to the fact that high-power motors are usually designed to work as near as possible to the network frequency (low slip) and these motors are often more efficient than lower or medium power ones. The positive relationship between the efficiency and the power factor could mean that higher efficiency leads to higher power factor, because the higher the power factor, the more active power will be delivered, given a certain output. Thus, we suggest naming the factor '– global efficiency' because it is a measure of the efficiency achieved by the machine. The negative sign in the label indicates that global efficiency grows when the efficiency, the locked-rotor current and the power factor increase and the slip decreases.

Once we have discussed the interpretation of the first-order factors, we can move on to the second-order factors and try to give them meaning. From Table 6, Fact2_2 is clearly defined by the first-order factors Fact1_1 ('size') and Fact2_1 ('global efficiency'). Fact4_1 and Fact3_1 load on Fact1_2. The loadings of Fact2_1 on Fact1_2 and Fact3_2 are quite similar. From the previous discussion about speed, we decided to assign Fact2_1 to Fact1_2 because there is stronger

TABLE 6. Higher-order factors extraction

Kaiser-Meyer-Olkin Measure of Sampling Adequacy = 0.52546
Bartlett Test of Sphericity = 173.87346, Significance = 0.00000

Initial Statistics:

Variable	Communality	*	Factor	Eigenvalue	% of Var	Cum %
FACT1_1	1.00000	*	1	2.14467	42.9	42.9
FACT2_1	1.00000	*	2	1.22085	24.4	67.3
FACT3_1	1.00000	*	3	0.98393	19.7	87.0
FACT4_1	1.00000	*	4	0.38348	7.7	94.7
FACT5_1	1.00000	*	5	0.26708	5.3	100.0

VARIMAX Rotated Factor Matrix:

	Factor 1 (Fact1_2)	Factor 2 (Fact2_2)	Factor 3 (Fact3_2)
FACT1_1		**0.94064**	
FACT2_1	0.60017		0.60352
FACT3_1			**0.95146**
FACT4_1	**0.93586**		
FACT5_1	−0.45486	**−0.71221**	

Structure Matrix:

	Factor 1	Factor 2	Factor 3
FACT1_1		**−0.91811**	
FACT2_1	0.67259		0.68063
FACT3_1			**0.93709**
FACT4_1	**0.91861**		
FACT5_1	−0.52235	**0.77292**	−0.47046

Factor Correlation Matrix:

	Factor 1	Factor 2	Factor 3
Factor 1	1.00000		
Factor 2	−0.11212	1.00000	
Factor 3	0.20836	−0.20461	1.00000

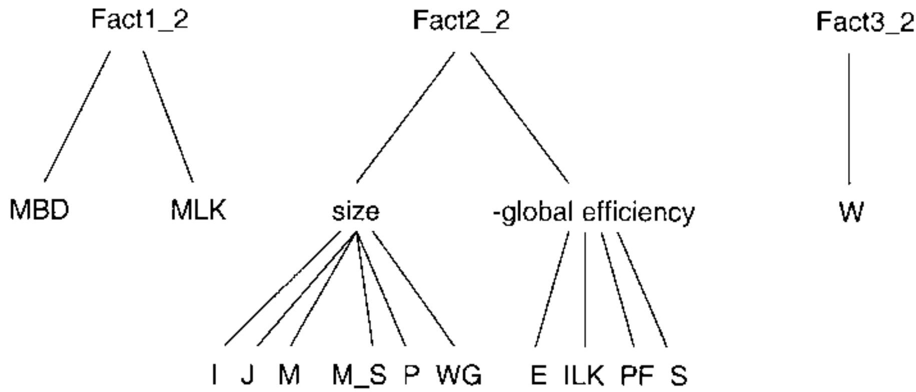

FIG. 3. Diagram of the final solution.

evidence for speed being a nearly independent variable than for the speed and the breakdown torque being related.

The final structure of the solution is represented in Fig. 3. Note that because of the different signs, there is an inverse relationship between the factor 'size' and the factor '– global efficiency', so we can say that the factor 'size' and the factor 'global efficiency' are directly related. Thus, the higher the 'size', the higher the 'global efficiency' of the motor. This statement seems to be logical because, in machines with outputs of hundreds of kilowatts, you cannot afford to work inefficiently due to heat losses. In these machines the 'global efficiency' is a critical design parameter because small differences in it may lead to increasing losses, making the design of the cooling system more difficult.

6.2 Robustness of the solution

In Section 5, different rotational methods have led to very similar results. We have also applied different extraction methods successfully, so the final solution presented is quite stable. But, can we generalize these results to other samples?

To confirm this possibility, other samples have been analysed and it was found that the basic structure replicates across them. For example, in one of the samples, factor analysis led to a similar factor structure but with a difference in the efficiency variable that loaded on the factor 'size' instead of loading on the factor 'global efficiency'. That was because the studied motors had very high output power (some of them, thousands of kilowatts), so the efficiency was a critical variable with a strong dependence on the 'size'. This may mean that, depending on the manufacturer, the kind of motors or their application, there might be differences in the solution obtained.

7 Conclusions

After examining the study presented above, we can conclude that it is valuable to apply multivariate methods to electric machines in order to confirm our previous knowledge and to discover subtle aspects that other methods may not capture. Specifically, it has been shown reasonable to summarize the key parameters of the induction motors in a few meaningful factors, using the factor analysis technique.

This has required constructing a model with linear relationships between the manifest and latent variables.

Over the different samples analysed from several manufacturers, the following may be stated:

(1) A minimum of three factors is required to summarize the main variables in a set of induction motors.
(2) At most, up to five factors might be extracted.
(3) The most stable solution, in the sense that there are no strong changes in the factor loadings across different methods of factor extraction and several rotation procedures, is obtained with three factors. However, further insight may be gained extracting more than three factors.
(4) The suggested structure for the interrelations of the main characteristics of the induction motors has shown that the final solution is more complex than a two-factor solution.
(5) A methodology quite different and independent from the traditional analysis, which is based on the physical laws that rule the induction machine, has confirmed some of the known facts about its operation. Moreover, relations that simpler models, such as equivalent circuits, may not take into account have been observed. It should be noticed that in this paper the extraction of higher-order factors has allowed a more complete understanding of the latent factor structure.

More research is required in order to find the coefficients that might define the relationship between factors and key variables for every kind of electric machine. Another interesting study might be the statistical analysis of internal variables that are not published in the catalogues, for instance, those concerning the magnetic core. On the other hand, the adequacy of more complex models with non-linear terms remains an open issue.

Acknowledgements

Our thanks are due to Dr Christine Heller and Mariano Ventosa who have kindly revised our work and have especially contributed to its enhancement in those sections concerning electric machines. We also appreciate the referee's comments, which have enriched the final version of this paper.

References

Basilevsky, A. (1994) *Statistical Factor Analysis and Related Methods. Theory and Applications* (New York, Wiley).

Eom, S. B. & Farris, R. S. (1996) The contributions of organizational science to the development of decision support systems research subspecialties, *Journal of the American Society for Information Science*, 47, pp. 941–952.

Glorfeld, L. W. (1995) An improvement on Horn's parallel analysis methodology for selecting the correct number of factors to retain, *Educational and Psychological Measurement*, 55, pp. 377–393.

Goronzy, F. (1969) A factor analysis of selected variables of manufacturing business enterprises, *Management International Review*, 6, pp. 71–96.

Gorsuch, R. L. (1983) *Factor Analysis*, 2nd edn (Hillsdale, New Jersey, Lawrence Erlbaum Associates).

Hair, J. F. Jr., Anderson, R. E., Tatham, R. L. & Black, W. C. (1998) *Multivariate Data Analysis with Readings*, 5th edn (Englewood Cliffs, New Jersey, Prentice Hall).

Hamdi, E. S. (1994) *Design of Small Electrical Machines* (Chichester, Wiley).

Harman, H. H. (1976) *Modern Factor Analysis*, 3rd edn (Chicago, University of Chicago Press).

Jobson, J. D. (1992) *Applied Multivariate Data Analysis. Vol. II: Categorical and Multivariate Methods* (New York, Springer-Verlag).

Johannsen, C. G. (1996) Strategic issues in quality management: 2. Survey analysis, *Journal of Information Science*, 22 (4), pp. 231–245.

Johnson, R. A. & Wichern, D. W. (1998) *Applied Multivariate Statistical Analysis*, 4th edn (New Jersey, Prentice Hall).

Khots, M. S., Nazarov, V. I. & Rudyak, K. B. (1994) Calculation of diesel fuel motor characteristics, *Chemometrics and Intelligent Laboratory Systems*, 22, pp. 265–271.

Nasar, S. A. (1987) *Handbook of Electric Machines* (New York, McGraw-Hill).

Sen, P. C. (1989) *Principles of Electric Machines and Power Electronics* (Singapore, Wiley).

Shin, J. K. & Kim, K. J. (1997) Restructuring a house of quality using factor analysis, *Quality Engineering*, 9(4), pp. 739–746.

Siemens (1988) *Catalog M11. Low-Voltage Three-Phase Motors* (Erlangen, Siemens).

Tabachnick, B. G. & Fidell, L. S. (1996) *Using Multivariate Statistics*, 3rd edn (New York, Harper Collins College Publishers).

Velicer, W. F., & Jackson, D. N. (1990a) Component analysis versus common factor analysis: some issues in selecting an appropriate procedure, *Multivariate Behavioral Research,* 25, pp. 1–28.

Velicer, W. F. & Jackson, D. N. (1990b) Component analysis versus common factor analysis: some further observations, *Multivariate Behavioral Research*, 25, pp. 97–144.

Vinzanekar, S. G., Jogdeo, A. G., Joshi, P. R., Sundaram, V., Parthasarathy, M. S. & Srinatham, B. (1996) A study of the influence of fiber properties on the characteristics of rotor-spun yarns by factor analysis, *Journal of Textile Institute*, 87 Part 1, No. 1, pp. 68–77.

Zwick, W. R. & Velicer, W. F. (1986) Comparison of five rules for determining the number of components to retain, *Psychological Bulletin*, 99, pp. 432–442.