Journal of the Operational Research Society (2006), 1–11 0160-5682/06 $30.00

www.palgrave-journals.com/jors

What do UK employers want from OR/MS?

C Mar Molinero* and A Xie

University of Kent, Canterbury, Kent, UK

Over 500 job advertisements in the Newsletter of the Operational Research Society are analysed to explore the skills that employers want from students and the tasks that such students are expected to perform in their jobs. There are consequences for the content of Operational Research and Management Science degrees.

Journal of the Operational Research Society advance online publication, 4 October 2006
doi:10.1057/palgrave.jors.2602286

Keywords: statistics; OR education; practice of OR

Introduction

Although there is no agreed definition of what is Operational Research/Management Science (OR/MS), few will dissent from the view that it is an applied discipline whose aim is to solve complex problems arising in the management of large organizations (Mar Molinero, 1992). If we had a clear view of what those problems are, it would be possible for universities to design course structures and syllabuses that are in touch with reality. But problems continuously change and the technical approaches needed to address them also evolve. Universities train future practitioners, but universities are not always in daily touch with practice, and it is easy for university teachers to get stuck in the teaching of nice algorithms that are intellectually satisfying and easy to examine, but whose use was long relegated to the pages of industrial history. The converse is also true, there may be algorithms or approaches that have proven valuable in industry but that are still relatively unknown in academia. All disciplines evolve, but OR/MS evolves faster than most. This is natural, given its 'research' orientation and its 'scientific approach' methodology. It is, therefore, relevant to ask from time to time what it is that practitioners do. In the past, this has led to surveys of OR groups (Mingers, 1991). Here, we depart from the survey tradition and we examine the content of advertisements in the specialized press.

In this paper, we look at 512 advertisements for jobs placed in the Newsletter of the Operational Research Society during the period December 2003–January 2005. The aim is to explore the skills that employers require from future employees in the OR/MS area, the tasks that the employees will be expected to perform, and their employment conditions.

Advertisements are short by their very nature. They have to be selective in their content and not state the obvious. Advertisements in the specialized press can be even more concise, as the reader is expected to be familiar with the concepts and share a culture. This gives a particular meaning to words. Take for example the term 'statistical modelling'. Anyone with an OR/MS background is expected to be able to conduct statistical modelling of some kind. But, by stating in an advertisement that the job involves 'statistical modelling', the prospective employer is emphasizing this aspect. It is saying that the employee will spend most of the time in statistical modelling and not in other things, such as project planning, and that the person selected to do the job will have to be specially good at statistical modelling. If it is the case that good statistical modelling skills are in short supply, this will be reflected in the salary offered. Advertisements also contain details about the work that the employee will be performing, about the place of work, and about salary and benefits, although some of this information is, at times, missing.

The information contained in the advertisements was translated into keywords and coded into an SPSS file for further analysis. Each advertisement was treated as an record and each keyword as a variable. The data are binary, indicating the presence or absence of a keyword in an advertisement. This created a file that was analysed using multivariate statistical techniques. Two multivariate statistical techniques were used: ordinal multidimensional scaling (MDS) and hierarchical cluster analysis (HCA).

The use of multidimensional scaling revealed that the job contained in an advertisement can be described in terms of four independent criteria: its orientation towards analysis inside the organization, or towards the environment in which the organization operates; the macro- (taking into account the whole picture) or micro- (concentrating on the detail) vision of the analysis; the use of technical or judgemental elements in the analysis; and the orientation towards planning.

Cluster analysis further revealed that organizations often require micro-analysis of the external environment, and that this is a consequence of the need to make sense of large data sets for marketing purposes.

**Correspondence: C Mar Molinero, Kent Business School, University of Kent, Canterbury, KentCT2 7PE, UK.*
E-mail: cm235@kent.ac.uk, c.mar-molinero@kent.ac.uk

The next section describes the data and offers some exploratory analysis. This is followed by a methodological section in which the model and the results are discussed. The conclusion summarizes the findings.

The data

The data set consists of all the job advertisements published in the Newsletter of the Operational Research Society between 17th December 2003 and 12th January 2005. No attempt was made to explore further back. If we had been going back in time, the nature of the jobs required might have changed, and the salaries might not have been comparable, since salaries were taken at face value and no inflation factor was applied. Some advertisements were placed by specialized employment agencies and some by individual firms. In total, information was collected from 540 advertisements, but some of them had appeared more than once and duplications were removed from the data set. This reduced the total number to 512.

The information contained in each advertisement was coded into three sets of variables: those related to the conditions that the person employed had to satisfy (requirements); those related to the activities to be performed in the job (activities); and other information such as salary, non-pay benefits, location, and experience required. There were instances when the same keyword appeared both as a requirement and as an activity. For example, some advertisements required experts in forecasting techniques to engage in forecasting sales. Thus, forecasting appeared both as a requirement and as an activity. In such cases, the key word was coded twice, once as an activity and again as a requirement.

The coding of salaries presented difficulties. Some salaries included unspecified 'benefits'. A variable was thus coded taking the value 1 if there were benefits and zero if there were no benefits. This was clearly unsatisfactory, as benefits can take many different forms. Sometimes, salaries were described as 'competitive' or some similar term. Rather than remove from the data such records, a 'nearest-neighbour' approach was followed to estimate a salary value for further statistical analysis. The data set was searched and jobs similar to the one with missing salary data were identified. The missing value was replaced with an estimate obtained from these nearest neighbours. Although this procedure introduces an element of error, only a small number of values were estimated in this way, and it was hoped that the statistical analysis would be robust to estimated data. Salary data were then coded into ranges (numbers in brackets indicate the number of times that this variable took the value one). The names of the variables and their definitions are as follows:

- *upto25*: up to £25 000 p.a. (134).
- *upto35*: more than £25 000 and up to £35 000 (240).
- *upto45*: more than £35 000 and up to £45 000 (101).
- *upto55*: more than £45 000 and up to £55 000 (25).
- *upto65*: more than £55 000 and up to £65 000 (10).
- *upto90*: more than £65 000 and up to £90 000 (2)
- *n-benefits*: the pay contains non-salary benefits (313).

Each one of these variables is, of course, of the zero/one type.

We can see that most advertisements are aimed at the lower end of the salary scale, the mode being between £25 000 and £35 000.

Missing value problems were also faced when looking at the experience required for the job. Again, when no information was available, a nearest-neighbour type of estimation was followed. There often was a degree of ambiguity; for example, at times 2 years experience was required, but the experience had to be in the relevant area. To capture this difference, a distinction was made between 'some experience required' and 'medium level of experience required'. Then, five binary variables were created:

- *no*: no experience required (245).
- *some*: some experience; 1 or 2 years (52).
- *moderate*: moderate experience; 2 or 3 years (104).
- *medium*: medium experience; at least 3 years (60).
- *high*: high experience; 5 years or more (51).

We see that jobs that do not specify experience are the most common. This, combined with the fact that salaries are often in the lower half of the range, indicates that most advertisements are aimed at relatively young people.

The location of the job produced eight zero/one variables:

- London (182),
- South (174),
- North (68),
- Midlands (60),
- Scotland (4),
- Wales (10),
- Ireland (1),
- UK (13).

Jobs that required the employee to travel through the country or abroad were classified as UK.

London and its surrounding area appear to be the most common location for OR/MS workers.

It is evident that two advertisements may use different words to describe the same activity or to state the same requirement. Some editing was, therefore, needed in order to translate activities and requirements into keywords. The keywords finally used were as follows:

Requirements for the job:

- *rs_STA_PACKAGES*: Skills in the use of statistical packages. This included SPSS, SAS, STATA, Eviews, R, and Ox (265).
- *rs_COMMUNICATION*: Communication skills, including client facing skills, presentation skills, documentation, languages, and liaising with others (177).

- *rs_DATABASE*: Skills in the use of database software such as SQL, Oracle, Access, Dbase, DataQuery, and OLAP (120).
- *rs_SPREADSHEET*: Skills in the use of spreadsheet software. Excel was often mentioned as a requirement (98).
- *rs_REGRESSION*: Regression and econometric modelling abilities (42).
- *rs_MARKETING*: Marketing expertise, including propensity modelling, gravity modelling, profiling, segmentation, e-commerce, and campaign management tools (36).
- *rs_COMPUTER_PROGRAMMING*: Computer programming languages. Languages specifically mentioned were VBA, MatLab, Macros, Visual Basic, C, and Java (34).
- *rs_DAT_MINING*: Skills in the use of data mining software. We classified as such CHAID, Enterprise Miner and Neural Networks tools (28).
- *rs_MS_OFFICE*: Expertise in Microsoft Office systems including Power Point, Word, and MS packages (27).
- *rs_MULT_ANALYSIS*: Expertise in multivariate statistical analysis. There was specific mention of some specialized form of multivariate analysis such as cluster analysis, and discriminant analysis (26).
- *rs_FLEXIBILITY*: Flexibility, sometimes described as mobility (16).
- *rs_SIMULATION*: Expertise in simulation modelling (14).
- *rs_BUSINESS_INTELLIGENCE*: Skills in the use of business intelligence packages including BRIO, COGNOS, VIPER, Powerplay, and Business Objects (14).
- *rs_GEO_INFO*: Geographical information systems. Sometimes the advertisement just required knowledge of geodemographical techniques, and in most cases, knowledge of a package was specified. Packages mentioned included GIS, ArcInfo, SegMentz, Insight, Mosaic, MapInfo, and GeoMedia. (12).
- *rs_FORECASTING*: Skills in building forecasting models (8).
- *rs_SIMULATION_TOOLS*: Expertise in simulation tools such as Witness, Quest, Arena, and Simul8 (5).
- *rs_STATISTICS*: Here, we classified requests for expertise in statistical modelling such as Bayesian Modelling, Inference, and Time Series (5).
- *rs_ACCOUNTING_FINANCE*: Accounting and Finance. Some advertisements mentioned net present value calculations and profitability analysis (4).
- *rs_OPTIMIZATION*: Optimization expertise. No particular techniques or approaches were mentioned (2).
- *rs_HUMAN_RESOURCE*: Human resource management expertise including reward systems (1).

The activities to be undertaken were described with much less detail. These were:

- *as_MODELLING*: Modelling (364).
- *as_STATISTICS*: Statistical analysis (187).
- *as_MARKETING*: Marketing (136).
- *as_REPORTS*: Report writing (96).
- *as_MANAGEMENT*: General management (59).
- *as_FORECASTING*: Forecasting (57).
- *as_GISDAT*: Geographical information systems data related (52).
- *as_ACCFIN*: Accounting and finance related (45).
- *as_PLANNING*: Planning (40).
- *as_INFOSY*: Information systems (34).
- *as_PRESENTATION*: Presentation (26).
- *as_CRERAT*: Credit rating (25).
- *as_RISK*: Risk analysis (19).
- *as_OPTIMIZATION*: Optimization (19).
- *as_SCHROS*: Scheduling and rostering (18).
- *as_LOGTRA*: Logistics and tracking (14).
- *as_SIMULATION*: Simulation (13).
- *as_CONSULTANCY*: Consultancy (12).
- *as_DATMIN*: Data mining (11).
- *as_ECONOMETRICS*: Econometrics (8).
- *as_BENCHMARKING*: Benchmarking (8).
- *as_SOFT*: Soft OR (7).
- *as_MANPOWER*: Manpower planning (3).
- *as_QUALITY*: Quality related (2).

As an advertisement may list several requirements or the employee is to engage in several activities, the numbers in brackets do not add up to 512. It is interesting to notice the importance given to computers, computing, and packages as a requirement for employment. Statistical software expertise, knowledge of database systems, spreadsheets, computer programming, data mining software, and Microsoft Office systems appear to be in higher demand than particular technical skills of an academic nature such as optimization. This could, of course, imply that academic skills such as optimization are presumed and do not need to be explicitly stated, but computing skills could also be presumed and they are often explicitly named as a requirement.

On the side of the activities to be performed, apart from modelling, which is clearly demanded in any OR/MS job, and often explicitly stated, what becomes clear is the dominance of statistical analysis over deterministic optimization approaches. The level of statistical expertise demanded is often very high.

Communication skills require special mention. These are often stated as requirements and as activities. It is clear, and not at all surprising, that employers do not only want people who are competent at computer tools and at statistical analysis, but they also want people who can communicate their findings. Of particular interest is the follow-up stage, which is reflected in the inclusion of 'tracking' as a keyword in the database.

This section has been purely descriptive in its treatment of the results. There are clear associations between keywords, and these need to be explored. For example, skills in database management are often required in order to target customer management. This will be explored in the next section.

Summarizing, we have produced a table of 64 zero/one variables by 512 cases. This table will be analysed with the tools of multivariate analysis.

Analysis

The data collected were kept in the form of a table of variables by cases, and it will be analysed by means of the tools of multivariate analysis, in particular MDS and HCA.

The objective of the analysis is to see the way in which location, salary, requirements, and activities are related. It is also important to explore if there are subsets of the variables that tend to appear in combination.

MDS is a proximity-based approach to data modelling (Kruskal and Wish, 1978). In our particular case, we will be modelling the proximity between variables; that is, we are interested in knowing how often two particular keywords appear together in the same advertisement.

The first step in the implementation of the algorithm is to define a measure of proximity between any two keywords. There are many measures of proximity that could be defined, but we need to take into account the fact that we are dealing with zero/one variables (Yin and Yasuda, 2005). The measure of proximity between two keywords was obtained by counting the number of times they both take the value 1 simultaneously over the 512 advertisements. This is a measure of similarity: the higher the number of times that the two variables take the value 1 in an advertisement, the more similar they are. Since there are 64 keywords, we end up with a table of 64 by 64 measures of proximity, which is used as an input to the PROXSCAL routine in SPSS, one of the two MDS options in the package SPSS.

In order to understand how MDS works, it is convenient to think in terms of a geographical map. If the map is flat, cities are points in the plane. One could add a grid to the map in such a way that we only have to refer to the coordinates in the grid of a city in order to say where it is. There are standard coordinates—longitude and latitude—that could be used, but nothing stops us from having specific coordinates for each particular map. These coordinates are the dimensions of the map. Thus, a flat map is a two-dimensional configuration. There is no need to constrain ourselves to two dimensions: we could add a third coordinate in the form of altitude over the level of the sea, and even a fourth coordinate indicating average annual rainfall. In this way, in order to give details about a particular town, we would refer to the values of four dimensions: the position on a north/south dimension, the position on an east/west dimension, the altitude over the level of the sea, and average annual rainfall. From such a map, we could derive a table of distances between any two cities. If two cities appear to be very close to each other in the map, the distance will be small, and if two cities appear to be far apart from each other in the map, the distance will be large.

The MDS algorithm works in the opposite direction. Similarities between points are taken to indicate how close any two points should be in a map. For a given number of dimensions, the points are located in the space in such a way that if the similarity is high, the points are close to each other, and if the similarity is low, the points are located far apart. There is a need to understand what the dimensions mean, and we hope that careful examination of the map will provide clues to this effect.

It is important to draw the map in the correct number of dimensions. There are various techniques that can be used to assess the appropriate dimensionality of the data. In this particular case, we have relied on the so-called ‘elbow test’. This works on the principle that the matching of similarities (the original data) with distances (as calculated from the final map) is never perfect, and a measure of goodness of fit can be calculated. This measure of goodness of fit is nothing else than a sum of squared errors, and is known as ‘stress’. As in any other algorithm, increasing the number of parameters in the model results in a decrease in stress—when the extra parameters are relevant—, or increasing the number of parameters leaves stress unchanged—when the extra parameters do no add anything to existing results. The number of parameters in the model is the number of points (n) multiplied by the number of dimensions (k), because each of the n points has to be located in the k-dimensional space. It follows that increasing the number of dimensions by one requires the inclusion of n extra parameters in the estimated model, and the value of stress should decrease until we have reached the relevant dimensionality. The plot of stress *versus* dimensionality should show a decline followed by a horizontal line. This looks like an elbow and gives its name to the test. The location of the elbow indicates the number of dimensions in which the map should be drawn.

MDS configurations were created in a six-dimensional space, the maximum value allowed by the PROXSCAL algorithm. The elbow test suggested that this is a reasonable dimensionality for the data, although experience suggests that not all the dimensions are needed in order to interpret the resulting configuration. As a map in a six-dimensional space is impossible to interpret other than mathematically, we have to work with projections into two-dimensional subspaces. The projection on to the first and the second dimensions can be seen in Figure 1. The projection on to the third and fourth dimensions can be seen in Figure 2. No attempt was made to explore higher dimensionality projections. We will comment on these figures further below.

When interpreting the maps, we must be aware of the fact that we are working with projections and it is easy to think that two points are close to each other when they are, in fact, far apart. Imagine two people, one standing at the top of a very high building and one standing at the bottom of the same building. When seen from above, these two individuals may appear to be close to each other even if they are not. For this reason, when one works with projections is its convenient to indicate how far two points are in the space. This is done by means of cluster analysis. There are various approaches to

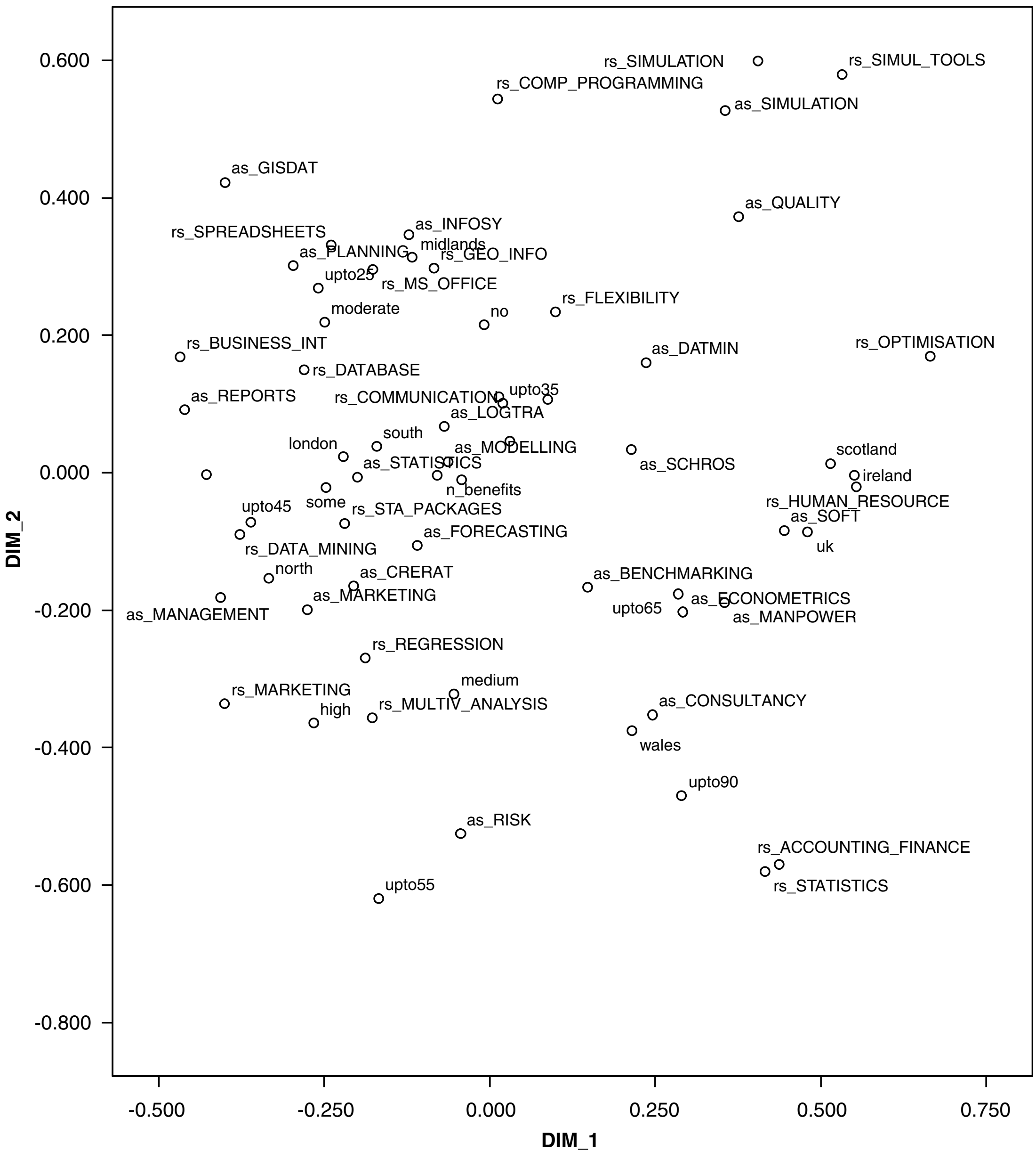

Figure 1 Multidimensional scaling configuration. Plot of Dimensions 1 and 2.

cluster analysis that could be used. Here, we have chosen to work with HCA. The method is called hierarchical because once a point joins a cluster it remains in it until the end of the algorithm.

HCA starts by exploring the distances between points until the shortest one is identified. The two points that are closest to each other are then merged into a single point forming the first cluster. The distance between this cluster and the remaining points has to be calculated. There are various options for calculating distances between the points that belong to a cluster and the points that are outside the cluster (including distances between clusters). Since the cluster is made equivalent to a new point, we are back to a matrix of distances between points, and the procedure starts again. A 'tree diagram' or dendrogram indicates the sequence in which the points are merged. Each time a merger takes place, a branch is created indicating how far apart the points merged were. When the branches are short, the points have much in common. When the branches are long, the points have little in common. Eventually, all the points merge into a single cluster, and it is reasonable to stop the procedure when the distances become large. What is meant by 'large' is a matter of judgement. As with any other statistical technique, we do not want too many clusters, in which case there would be too much

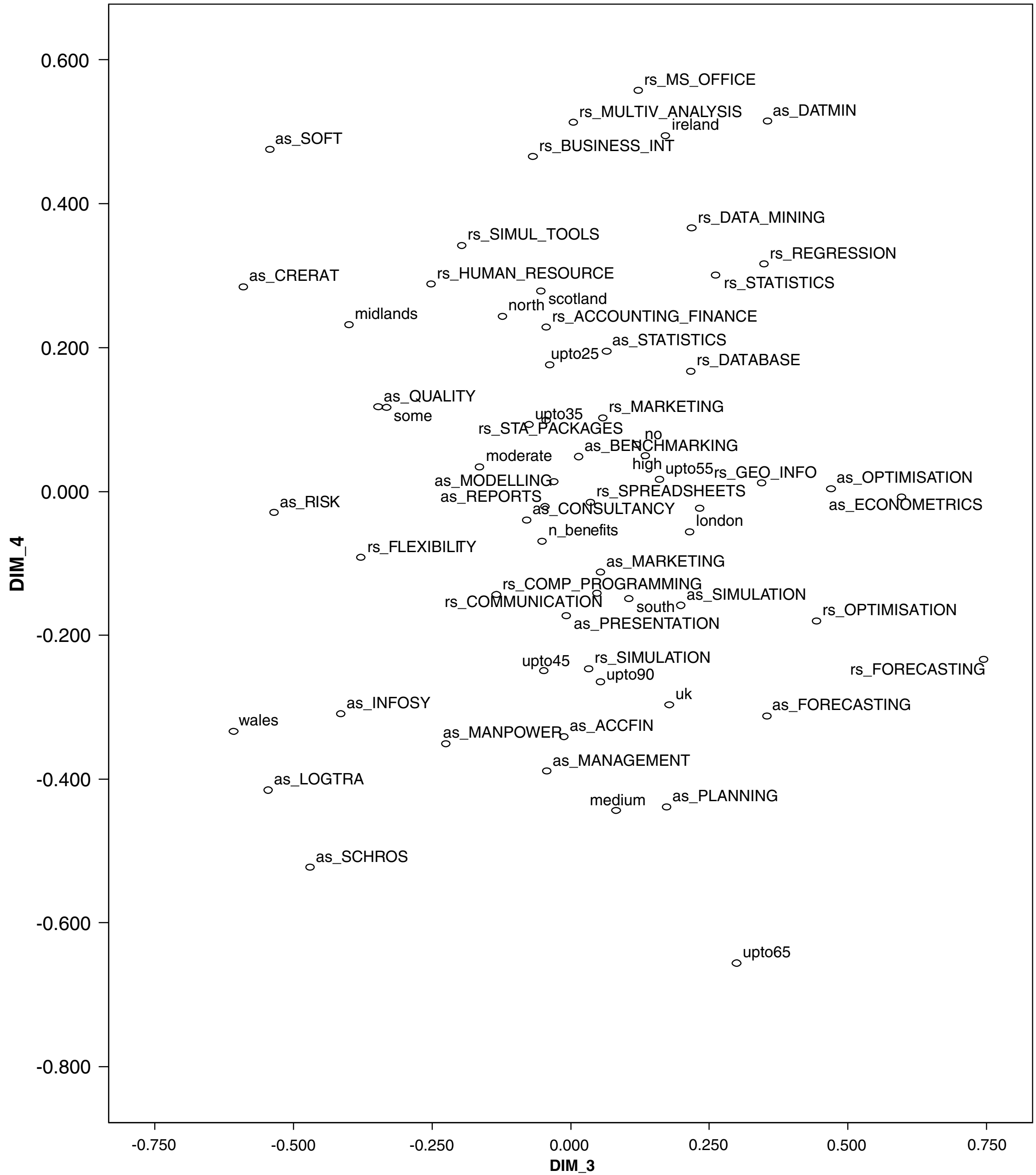

Figure 2 Multidimensional scaling configuration. Plot of Dimensions 3 *versus* 4.

detail, but we do not want clusters to contain points that are very different either.

Keywords are points in a six-dimensional configuration, but we are forced to work with projections in pairs of dimensions. In this context, HCA is of great benefit. Euclidean distances between the points in the space were calculated from their coordinates in the sixth dimensional space and used as dissimilarities in an HCA that uses Ward's agglomeration approach (Ward, 1963). HCA is known to be sensitive to the particular agglomeration schedule used, but Ward's approach attempts to maximize similarity between clusters while maximizing dissimilarity between clusters, and tends to produce reasonable results. The dendrogram can be seen in Figure 3.

Six clusters have been identified in Figure 3. Figure 4 has been obtained from Figure 1 by replacing keywords with cluster membership. The clusters are as follows:

Cluster 1 contains the following variables: South, London, Midlands, upto25, upto35, upto45, n_benefits, no, moderate, high, rs_STA_PACKAGES, rs_BUSINESS_INT, rs_DATABASE, rs_SPREADSHEETS, rs_COMP_

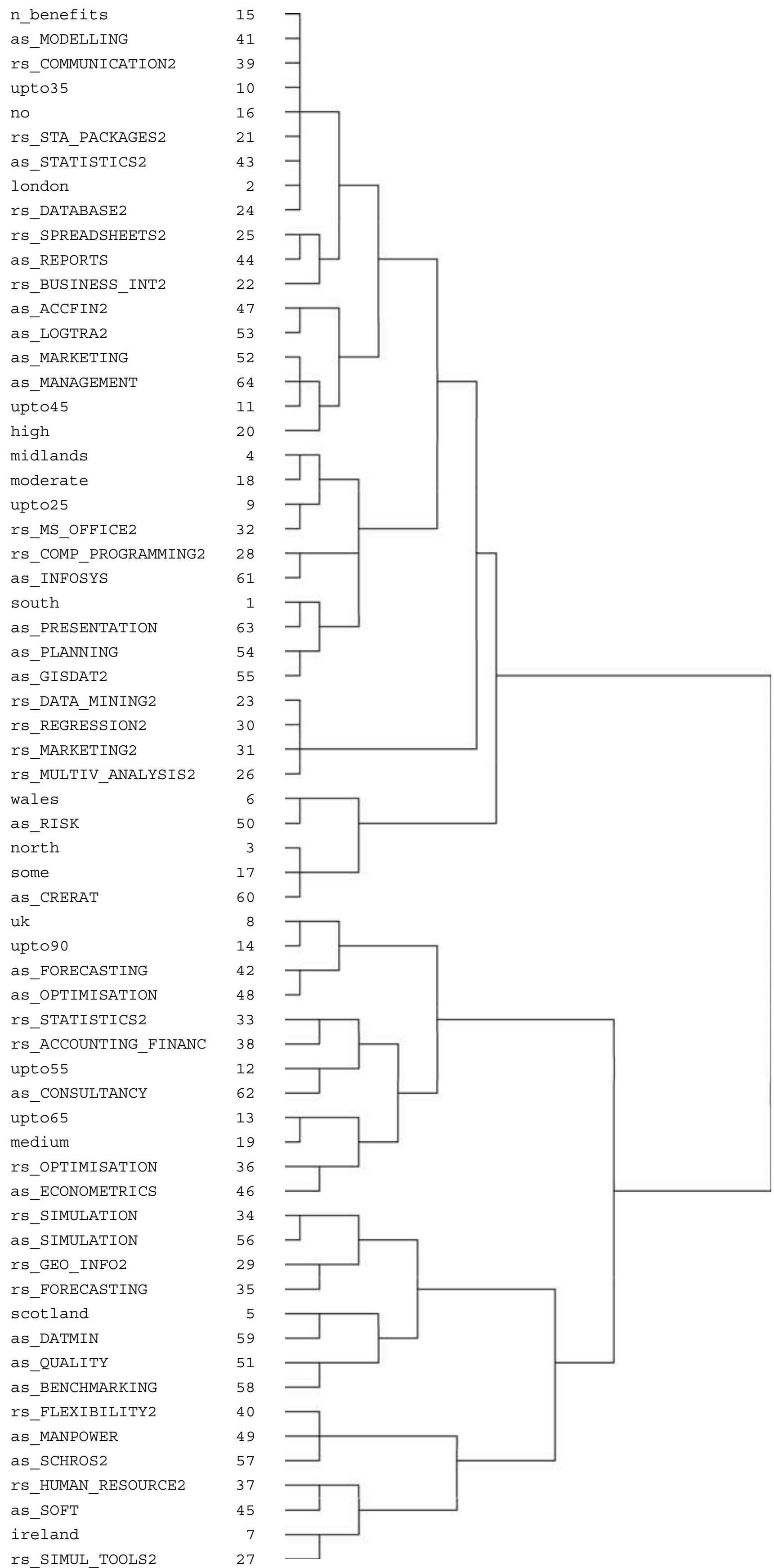

Figure 3 Dendogram using Ward's method.

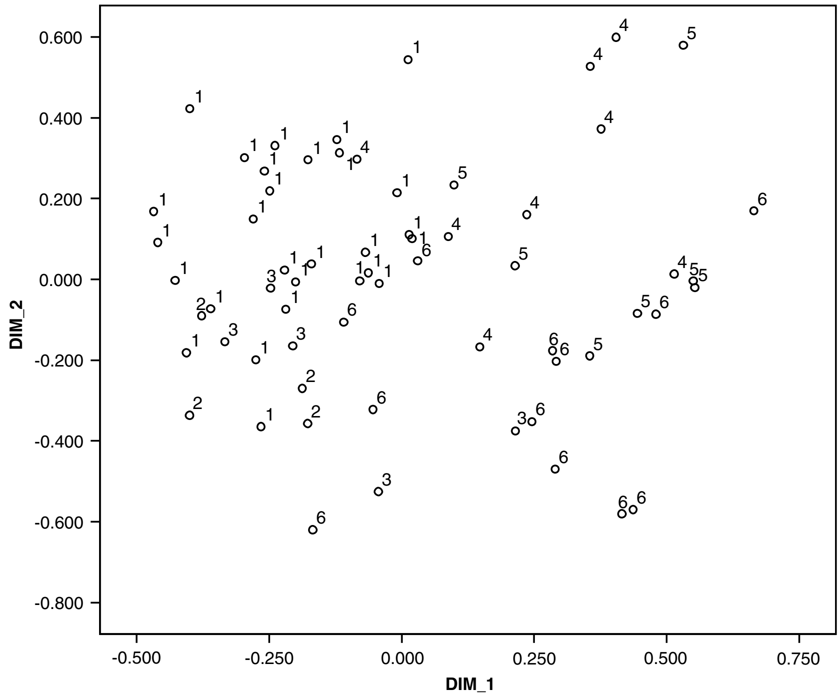

Figure 4 Projection of the MDS configuration on Dimensions 1 and 2 showing the clusters.

PROGRAMMING, rs_MS_OFFICE, rs_COMMUNICATION, as_MODELLING, as_STATISTICS, as_REPORTS, as_ACCFIN, as_MARKETING, as_LOGTRA, as_PLANNING, as_GISDAT, as_INFOSY, as_PRESENTATION, as_MANAGEMENT.

This cluster groups the demand for individuals of various levels of experience who are expected to have a fair amount of technical expertise to analyse data and to summarize its main features. It clearly reflects the need to analyse large databases for decision purposes. It is also noticeable that it groups the keywords that are most often mentioned in advertisements.

Cluster 2 contains the following variables: rs_DATA_MINING, rs_MULTIV_ANALYSIS, rs_REGRESSION, rs_MARKETING.

This cluster is not very different from the previous one, except in the sense that it requires individuals to be familiar with the tools of multivariate analysis—regression could be defined as multivariate analysis—within a marketing context.

Cluster 3 contains: North, Wales, Some, as_RISK, as_CRERAT.

This cluster shows a specialized application of OR/MS analysis: risk analysis and credit rating. This work involves the use of large data sets within a multivariate analysis context, and it is not surprising to discover that there is little difference between Cluster 2 and Cluster 3. Indeed, at a higher level of clustering they merge into a single cluster. We have preferred to keep them apart because the risk analysis activity is very different from the analysis of large databases for marketing purposes.

Cluster 4 contains: Scotland, rs_GEO_INFO, rs_SIMULATION, rs_FORECASTING, as_QUALITY, as_SIMULATION, as_BENCHMARKING, as_DATMIN.

This is a mixed bag that appears to group together the demand for statistically oriented individuals to work in statistically oriented problems. In this sense, one must remember that the now popular approach to quality control known as 'six-sigma' is only applied statistical analysis, and that benchmarking, as practised in industry, is often a statistical exercise. The knowledge of geographical information systems as a requirement for employment, and of data mining as an activity, reflects emerging trends in OR/MS.

Cluster 5 contains: Ireland, rs_SIMUL_TOOLS, rs_HUMAN_RESOURCE, rs_FLEXIBILITY, as_SOFT, as_MANPOWER, as_SCHROS.

The activities of scheduling and rostering, and manpower planning certainly go beyond mathematical and statistical analysis and require the use of soft systems methodologies. It is not surprising to find such activities associated with requirements for expertise in manpower planning, and simulation, and a flexible attitude to work.

Cluster 6 contains: UK, upto55, upto65, upto90, Medium, rs_STATISTICS, rs_OPTIMIZATION, rs_ACCOUNTING-FINANCE, as_FORECASTING, as_ECONOMETRICS, as_OPTIMIZATION, as_CONSULTANCY.

The interpretation of this cluster is probably associated with the consultancy activity. Individuals with a high level of qualifications are required in order to engage in very technical activities, being highly rewarded for doing so.

Having discussed the clusters, we are now in a position to return to the figures and attach meaning to the dimensions. This we do by looking at the points that are located at the extremes of the axes and trying to identify what it is that they have in common. For an analogy, imagine that we have a representation of cities in Great Britain, that on the left-hand side of the representation one can find Bristol, Liverpool, and Bangor, and that on the right-hand side, one can find Canterbury and Norwich. It would be sensible to conjecture that the left-hand side is associated with the geographical west and that the right-hand side is associated with the geographical east. This type of reasoning will lead us to label the various dimensions. It is, of course, possible for different people to attach different labels to the same dimensions. In what comes we will try to justify our choice of labels.

Turning now to Figure 1, we can see that most of the data-oriented activities and requirements appear on the left-hand side of Dimension 1. The presence of marketing, credit rating, business intelligence, and geographical information suggest that this side of Dimension 1 is related to data gathering in order to explore the environment in which the firm operates. The other extreme of Dimension 1 contains activities such as quality, manpower planning, and simulation; there are also requirements related to optimization, simulation tools, accounting and finance, and statistics. It is reasonable to conclude that the right-hand side of Dimension 1 is associated with studying how the firm operates internally and working out methods of optimizing operations. This is the traditional realm of operational research, and it emphasizes analysis within the firm. We suggest that Dimension 1 provides an external *versus* internal orientation of the analysis.

It is somewhat more difficult to place an interpretation on Dimension 2. At the top end, this dimension contains simulation, information systems, planning, and quality as activities. The requirements associated with this end of the dimension are computer programming, simulation tools, and geographical information systems. This suggests the idea of microanalysis; that is, the collection of detailed data in order to explore how the various parts of a model relate to each other. At the other end of this dimension is the activities of risk analysis, consultancy, econometrics, and manpower planning. The requirements associated with this end of the dimension are statistical knowledge, knowledge of accounting and finance, multivariate analysis, regression, and marketing. This suggests that the lower end of Dimension 2 focuses on analysis at the macro-level; that is, the study of aggregate data in order to explore the overall picture. Thus, an interpretation of Dimension 2 is macro- *versus* micro-analysis.

The meaning of Dimension 3 can be ascertained by looking at Figure 2. The activities associated with the left-hand side of Dimension 3 are soft analysis, credit rating, risk analysis, logistics and transportation, and scheduling and rostering. The only requirement clearly associated with this activity is flexibility. If we think in terms of the way in which credit rating, risk analysis, logistics, and scheduling take place in industry, we tend to think in terms of methods that combine formal analysis tools with the exercise of judgement. At the right-hand side of Dimension 3, we find activities such as optimization, econometrics, forecasting, and benchmarking. These activities are to a large extent guided by very technical knowledge with little room for judgement. We can, therefore, suggest that Dimension 3 captures the balance between judgement and technical skills that the job requires.

Dimension 4, on the lower end, appears to be associated with planning activities. This can be deduced from the presence of scheduling, planning, managerial activities, logistics and transportation, and manpower planning (not shown in the figure). The activities that appear at the top end of Dimension 4—such as data mining, soft, or and credit rating—do not appear to be related to a planning horizon.

We can summarize the above discussion by saying that advertisements that have appeared in the Newsletter of the Operational Research Society can be described in terms of four independent characteristics (dimensions): external *versus* internal orientation of analysis; micro- *versus* macro-orientation; technical *versus* judgement orientation; and planning oriented or not.

Within this framework, we can see that Cluster 1, situated on the North West corner of Figure 1, can be described as external orientation at the micro-level. This combination appears to have been very much demanded by employers, as the keywords relating to both activities and requirements within this cluster appeared most often in advertisements. But this cluster is associated with relatively low salaries. The experience required is also at the lower end of the range. This suggests that employers are looking for recent graduates in their first jobs, and that they expect them to be well trained in information technology in order to perform data analysis.

Cluster 2, at the South West end of Figure 1 combines externally oriented analysis with a macro-vision, a combination that describes well the marketing orientation earlier discussed.

Clusters 3 and 4 both appear in the North East corner of Figure 1, indicating that the jobs advertised have an internal orientation and require detailed data analysis. Perhaps, the difference between these two clusters can be found in the fact that Cluster 3 appears to be situated towards the right-hand side of Dimension 3, and Cluster 4 appears to be situated on the left-hand side of this dimension, indicating that Cluster 3 is oriented towards the technical while Cluster 4 is oriented towards the judgemental.

Cluster 5 appears towards the bottom end of Dimension 3 and the right-hand side of Dimension 1, indicating an orientation towards the internal and the judgemental.

Finally, Cluster 6, in the South East corner of Figure 1, indicates a combination of internal orientation and macro-analysis. These are the best paid jobs and the ones that combine the tools of hard analysis with highly technical abilities. The list of requirements and activities very much reflects the standard contents of a standard Master's courses in OR:

statistics, optimization, accounting and finance, forecasting, and problem solution (consultancy). Universities are, perhaps, being guided by the most glamorous cluster of jobs, which is in relatively low demand, when devising syllabuses and advertising courses. It could, of course, be argued that these well-paid senior jobs requiring good qualifications are not advertised, but filled internally. The question then arises on whether universities should train students for their first job, or for the job they will eventually perform if they are successful in their careers. A possible, and reasonable, scenario would be for university courses to offer the skills that industry requires from new graduates, and for the Operational Research Society, through its continuing education programme, to provide the tools that are required later on in life.

The implications for teaching of OR/MS

If university institutions are to produce the graduates that industry demands, they should very much take into account the findings of this research. A series of lessons should be taken into account in the process of educational planning.

1. Modelling is most often included as a keyword in the advertisements. In fact, more than half of them explicitly state that the person employed will be required to engage in modelling. This emphasis is surprising, given that all the advertisements are in a specialized publication aimed at professionals whose main area of expertise is precisely modelling. What are we to conclude from this? Perhaps, Industry is telling educational institutions that it does not want individuals who can reproduce the stepping stone algorithm in the transportation model, for example, but that it wants individuals who know when to apply the transportation model, who are able to modify it, and who are able to use it to inform decision making. Industry is pleading for educational institutions to train students how to think, and is not interested in elegant solutions to past problems.
2. Most advertisements require students to be familiar with computer packages, databases, and computer decision tools. This is clearly a consequence of the world in which we live. Data are easily generated and stored. There is a need for individuals to be able to process data and extract relevant information from it. Given the large amount of data, the emphasis is not in carefully developing in-house analysis methods, but in quick and, possibly, 'dirty' methods to generate intelligence. Such intelligence is often required for marketing purposes and, indeed, marketing expertise is high in the list of requirements.
3. Communication skills are crucial. As both the world of data gathering and data analysis become more and more technical, the ability to put across a complex analysis message is paramount.
4. Statistical skills are often mentioned in one form or another, both as requirements and as activities. Data mining, multivariate analysis, risk analysis, quality control, simulation, forecasting, segmentation, and geographical analysis require a clear understanding of statistical reasoning. Statistics is a hard subject to teach and to understand, but no shortcuts can be taken.
5. To train future professionals in the skills that industry demands, students should be exposed to standard data sets. This means that educational institutions should be subscribing to standard databases, such as company accounts databases, population censuses, and other specialized data sets. These should then be used within a teaching context in such a way that students learn to use them and to produce relevant analyses. This can, of course, be very expensive.
6. The traditional syllabus of an OR/MS course, including highly satisfying but seldom demanded techniques, such as Dynamic Programming and Queuing Theory, should be thoroughly revised. New subjects such as Data Mining, Heuristics, and some aspects of Multivariate Analysis should replace them. Universities should, for example, teach Supply Chain Management rather than Inventory Theory.

Conclusion and discussion

Job advertisements in the specialized press are the way in which employers try to communicate their needs to the outside world. These advertisements can be read as a cry for help to teaching institutions, as employers are telling them the skills that are needed, and the contexts in which they are to be applied. But advertisements are read by recent graduates, when there is little that can be done to match what universities produce with what employers require. Perhaps, employers and educators do not communicate in a way in which a meaningful conversation can take place. But it has been argued in this paper that the tools of multivariate statistical analysis can contribute to making communication easier by highlighting educational needs and emerging areas of research interest.

One finding of this research is that modern data storage capacity is creating new opportunities for analysis within a marketing context. Employers are telling us that there is a need to equip students with the technical ability to analyse large data sets and with the skills to communicate their findings. The emphasis appears to be shifting away from doing things optimally towards getting the best intelligence from available data, and using it in a reasonable way.

If universities decide to accept the challenges, they need to re-structure their courses away from the very detailed mathematical knowledge of how an algorithm works towards a better understanding of the opportunities that might lie hidden in a customer database. Such a database is likely to be huge and not amenable to exact analysis. Perhaps, we should discard from the classroom the stepping stone algorithm and immerse ourselves in the world of the approximate and fast heuristic.

References

Kruskal JB and Wish M (1978). *Multidimensional Scaling*. Sage University Paper series on Quantitative Applications in the Social Sciences, number 07-011. Sage Publications: Newbury Park, CA.

Mar Molinero C (1992). Operational research: From war to community. *Socio-Econom Plann Sci* **26**: 203–212.

Mingers J (1991). The contents of MSc Operational Research courses: Results of a questionnaire to OR groups. *J Opl Res Soc* **42**: 375–382.

Ward JH (1963). Hierarchical grouping to optimize an objective function. *J Amer Statist Assoc* **58**: 236–244.

Yin Y and Yasuda K (2005). Similarity coefficient methods applied to the cell formation problem: A comparative investigation. *Comput Ind Eng* **48**: 471–489.

Received June 2005;
accepted June 2006 after two revisions