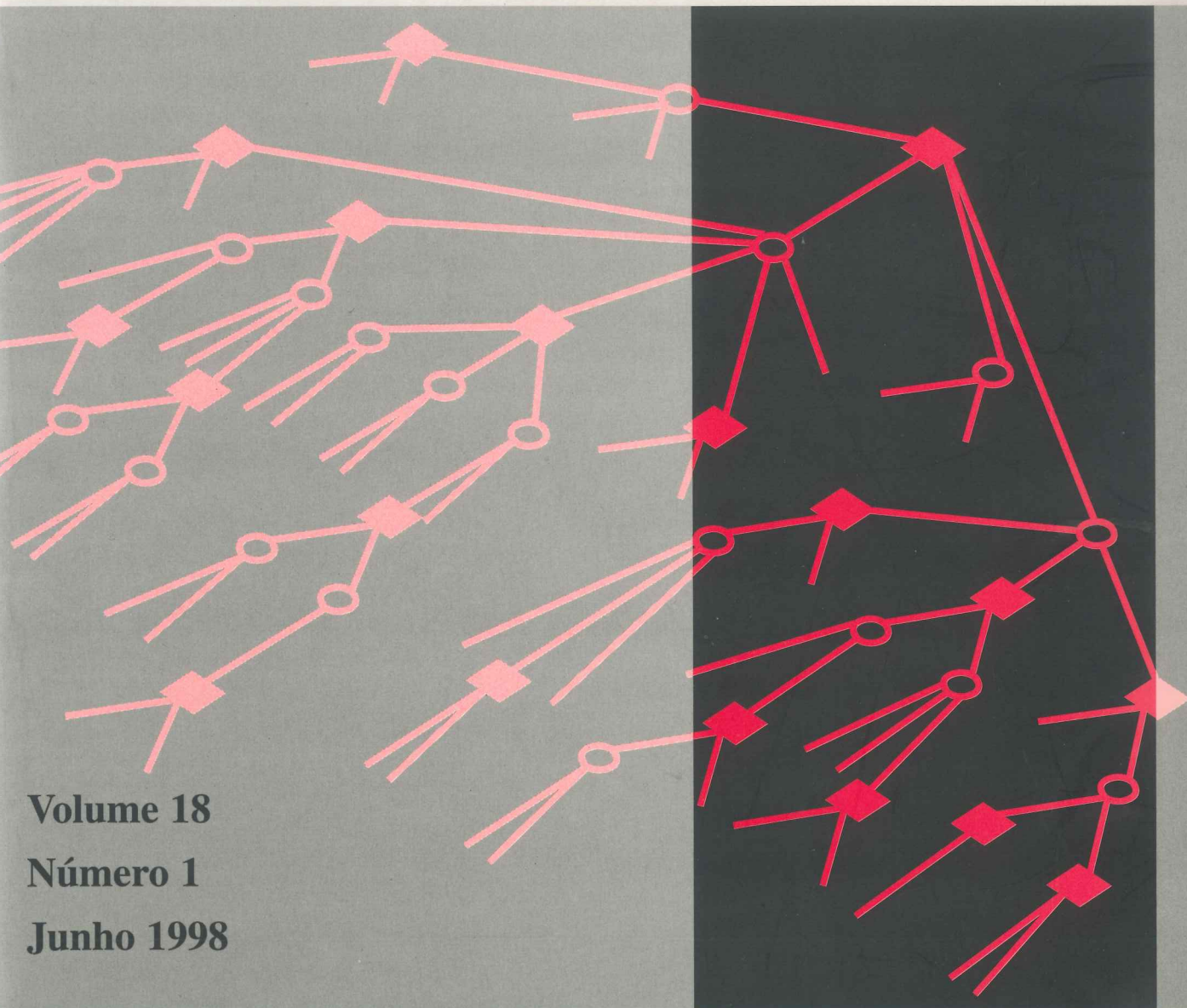


Investigação Operacional



Volume 18
Número 1
Junho 1998

INVESTIGAÇÃO OPERACIONAL

Propriedade:

APDIO — Associação Portuguesa de Investigação Operacional

ESTATUTO EDITORIAL

<<Investigação Operacional>>, órgão oficial da APDIO cobre uma larga gama de assuntos reflectindo assim a grande diversidade de profissões e interesses dos sócios da Associação, bem como as muitas áreas de aplicação da I. O. O seu objectivo primordial é promover a aplicação do método e técnicas da I.O. aos problemas da Sociedade Portuguesa.

A publicação acolhe contribuições nos campos da metodologia, técnicas, e áreas de aplicação e software de I. O. sendo no entanto dada prioridade a bons casos de estudo de carácter eminentemente prático.

Distribuição gratuita aos sócios da APDIO

INVESTIGAÇÃO OPERACIONAL

Volume 18 - nº 1 - Junho 1998

Publicação Semestral

Editor Principal: Joaquim J. Júdice
Universidade de Coimbra

Comissão Editorial

M. Teresa Almeida Inst. Sup. Economia e Gestão	Laureano Escudero IBM, Espanha	J. Pinto Paixão Univ. de Lisboa
Jaime Barceló Univ. de Barcelona	J. Soeiro Ferreira Univ. do Porto	M. Vaz Pato Inst. Sup. Economia e Gestão
Paulo Barcia Univ. Nova de Lisboa	J. Fernando Gonçalves Univ. do Porto	Celso Ribeiro Univ. Católica, Rio Janeiro
Isabel Branco Univ. de Lisboa	Clóvis Gonzaga Univ. Fed., Rio Janeiro	A. Guimarães Rodrigues Univ. do Minho
António Câmara Univ. Nova de Lisboa	Luís Gouveia Univ. de Lisboa	Mário S. Rosa Univ de Coimbra
C. Bana e Costa Inst. Superior Técnico	Rui C. Guimarães Univ. do Porto	J. Pinho de Sousa Univ do Porto
M. Eugénia Captivo Univ. de Lisboa	J. Assis Lopes Inst. Superior Técnico	Reinaldo Sousa Univ. Católica, Rio Janeiro
Jorge O. Cerdeira Inst. Sup. de Agronomia	N. Maculan Univ. Fed., Rio Janeiro	L. Valadares Tavares Inst. Superior Técnico
João Clímaco Univ. de Coimbra	Ernesto Q. Martins Univ. de Coimbra	Isabel H. Themido Inst. Superior Técnico
J. Dias Coelho Univ. Nova de Lisboa	Vladimiro Miranda Univ. do Porto	B. Calafate Vasconcelos Univ. do Porto
J. Rodrigues Dias Univ. de Évora		José M. Viegas Inst. Superior Técnico

A Revista "INVESTIGAÇÃO OPERACIONAL" está registada na Secretaria de Estado da Comunicação Social sob o nº 108335.

Esta Revista é distribuída gratuitamente aos sócios da APDIO. As informações sobre inscrições na Associação, assim como a correspondência para a Revista devem ser enviadas para a sede da APDIO - Associação Portuguesa de Investigação Operacional - CESUR, Instituto Superior Técnico, Av. Rovisco Pais, 1000 Lisboa.

Este Volume foi subsidiado por:

Fundação para a Ciência e Tecnologia

Fundação Calouste Gulbenkian

Para efeitos de dactilografia e composição, foram utilizados equipamentos gentilmente postos à disposição pelo Centro de Investigação Operacional (DEIO - FCUL).

Assinatura: 5.000\$00

BRANCH-AND-BOUND PARA O PROBLEMA DE JOB-SHOP: NOVAS REGRAS DE RAMIFICAÇÃO

Paula Marta Bruno

Faculdade de Motricidade Humana
Universidade Técnica de Lisboa
Lisboa - Portugal

Helena Ramalhinho Lourenço

Centro de Investigação Operacional, DEIO
Faculdade de Ciências da Universidade de Lisboa
1700 Lisboa - Portugal

Abstract

The job-shop scheduling problem consists in scheduling a set of jobs on a set of machines, in order to minimize the makespan.

In this paper, we develop and implement a variation of a branch-and-bound method for solving the job-shop scheduling problem, presented by Brucker, Jurisch and Sievers [1994]. The modifications refer to the lower bound method and to the branching scheme. The lower bound is obtained by solving an one-machine scheduling problem with lags. A relaxation of this problem can be solved in polynomial time by the *early-late* algorithm, developed by Lourenço [1993]. We studied some branching rules that can be adapted to this new lower bound method, in away that the successive partition of feasible solutions set gives good lags.

We present some computational results in order to compare the different versions of the branch-and-bound method.

Resumo

O problema de *job-shop* consiste num conjunto de tarefas a ser processadas por um conjunto de máquinas, com objectivo de encontrar uma sequência de tarefas em cada máquina que minimize a data global de conclusão.

Neste trabalho, desenvolve-se e implementa-se uma variante do método de *branch-and-bound* para o problema de *job-shop* proposto por Brucker, Jurisch e Sievers [1994]. As alterações aqui apresentadas dizem respeito ao método de limite inferior e à regra de ramificação. O limite inferior é obtido com base na resolução do problema de sequenciamento numa única máquina com intervalos de tempo. Uma relaxação deste problema pode ser resolvida em tempo polinomial utilizando o algoritmo *early-late*, desenvolvido por Lourenço [1993]. Construímos, então, regras de ramificação que se adaptam a este novo método de limite inferior, tendo em conta que a partição sucessiva do conjunto de soluções admissíveis deve ser feita de modo a criar bons intervalos.

Apresentamos diversos resultados computacionais, com o objectivo de comparar as diferentes versões do método de *branch-and-bound*.

Keywords

Job-shop, branch-and-bound, branching scheme.

1. Introdução

O problema de *job-shop* consiste num conjunto de tarefas a ser processadas por um conjunto de máquinas com o objectivo de encontrar uma sequência de tarefas em cada máquina que minimize a data global de conclusão "*makespan*". Apresentamos, em seguida, uma breve descrição do problema, bem como algumas definições e conceitos a ele associados.

Considere-se, um conjunto de n tarefas $\{J_1, J_2, \dots, J_n\}$ a ser processadas por um conjunto de m máquinas $\{M_1, M_2, \dots, M_m\}$. O processamento de uma tarefa numa máquina é designado por operação. A operação O_{ij} designa o processamento da tarefa J_j ($j = 1, \dots, n$) na máquina M_i ($i = 1, \dots, m$).

Para cada tarefa, é conhecida a ordem por que deve passar pelas máquinas e o correspondente tempo de processamento dado por p_{ij} ($i = 1, \dots, m$; $j = 1, \dots, n$). O conjunto de máquinas que processa cada tarefa e a respectiva ordem de processamento são fixos e conhecidos à partida, mas podem variar de tarefa para tarefa.

Uma máquina não pode realizar mais do que uma operação, simultaneamente. De uma mesma tarefa não podem ser processadas mais do que uma operação, em cada instante, e nenhuma operação pode ser interrompida.

O nosso objectivo é encontrar uma sequência de operações em cada máquina, de modo a minimizar a data global de conclusão. Sejam C_{ij} o instante de conclusão do processamento da operação O_{ij} e $C_{\max} = \max_{i=1, \dots, m, j=1, \dots, n} C_{ij}$. Então, minimizar a data global de conclusão corresponde a minimizar C_{\max} .

Uma representação comum do problema é através de um grafo disjuntivo, Roy e Sussmann [15]. O grafo disjuntivo designa-se por $G = (O, A, E)$, onde: O representa o conjunto de vértices, A o conjunto de arcos e E o conjunto de arestas; sendo a representação do problema feita do seguinte modo:

- para cada operação O_{ij} , existe um vértice com peso p_{ij} ;
- para cada par de operações consecutivas da mesma tarefa O_{ij} e O_{lj} , existe um arco (O_{ij}, O_{lj}) ;
- para cada par de operações O_{ij} e O_{lk} , a ser processado na mesma máquina, existe uma aresta $[O_{ij}, O_{lk}]$ (sem orientação associada);
- existem ainda dois nodos suplementares: θ e $*$, com peso zero, e aos quais se associam os seguintes arcos: do nodo θ para os nodos correspondentes à primeira operação de cada tarefa e, dos nodos correspondentes à última operação de cada tarefa para $*$.

Deste modo, os arcos representam as restrições de precedência de cada tarefa e as arestas representam as restrições de capacidade das máquinas.

No grafo disjuntivo, encontrar a ordem pela qual as operações devem ser processadas em cada máquina, corresponde a orientar cada aresta num dos dois sentidos possíveis. Designa-se por selecção e representa-se por E' , um conjunto de arestas para as quais se definiu uma

orientação. A uma selecção está associada uma solução parcial, pois somente algumas arestas estão orientadas. Uma selecção define uma sequência admissível quando a todas as arestas foi atribuído um sentido e o grafo não contém ciclos. Neste caso, a selecção diz-se completa.

A distância do caminho mais longo entre 0 e $*$ neste grafo, também designado por caminho crítico, corresponde à data global de conclusão desta sequência. Assim, o problema de *job-shop* pode ser formulado como o problema de encontrar uma selecção completa que minimize a distância do caminho mais longo entre 0 e $*$.

O problema de *job-shop* é um problema de difícil resolução, tendo sido provado que pertence à classe *NP-hard* [8]. Neste contexto, os algoritmos mais usuais para obter soluções são: as heurísticas (algoritmos aproximativos) e os métodos de *branch-and-bound* (algoritmos exactos baseados em técnicas de enumeração).

Diversos métodos de *branch-and-bound* têm sido desenvolvidos para obter o valor óptimo do problema. Os mais recentes foram propostos por Carlier e Pinson [6], [7], Applegate e Cook [1], Brucker, Jurisch e Sievers [5] e Boyd [4]. Estes métodos exigem elevado tempo computacional e não podem ser aplicados a instâncias do problema com elevadas dimensões. Carlier e Pinson [6] encontraram a solução para a instância com 10 tarefas e 10 máquinas proposta por Muth e Thompson [13], que permaneceu por resolver durante um longo período de tempo e que motivou o desenvolvimento de algoritmos para este problema.

Os métodos de *branch-and-bound* têm como suporte uma árvore de pesquisa. Inicialmente, a árvore tem um único nodo, a raiz, que representa todas as soluções admissíveis para o problema. Os métodos que vamos referir associam, a cada nodo da árvore, um grafo $G' = (O, AUE')$. Assim, os sucessores de cada nodo da árvore de pesquisa, são obtidos definindo a orientação de alguma ou algumas arestas no grafo disjuntivo associado.

Carlier e Pinson [6] introduzem a seguinte regra de ramificação: em cada nodo da árvore de pesquisa, escolher uma aresta ainda não orientada no grafo disjuntivo e acrescentar dois nodos à árvore, associando a cada um deles um dos possíveis sentidos para essa aresta. Utilizam a regra *EDD preemptive* [9] para obter o limite inferior.

Mais tarde, apresentam outro trabalho [7], onde introduzem um algoritmo polinomial que otimiza o cálculo de datas de disponibilidade e tempos de entrega. Este algoritmo permitiu desenvolver um novo esquema de ramificação.

Applegate e Cook [1] propuseram também um método de *branch-and-bound* baseado no primeiro trabalho de Carlier e Pinson. A principal diferença entre este algoritmo e o apresentado por Carlier e Pinson é ao nível das regras de eliminação, nomeadamente, a introdução de algumas regras mais fortes. O método de obtenção do limite inferior baseia-se também na utilização da regra *EDD preemptive*.

Posteriormente, Lourenço [11] trabalha sobre este método de *branch-and-bound*, propondo uma alteração relativa à obtenção do limite inferior. Neste caso, utiliza o algoritmo *early-late*, em vez da regra *EDD preemptive*.

Brucker, Jurisch e Sievers [5] propuseram um método baseado no segundo trabalho de Carlier e Pinson [7], introduzindo alterações ao nível da regra de ramificação. Este método apresenta um conjunto de regras de eliminação mais forte.

Dada a estreita ligação do problema de sequenciamento numa única máquina com o *job-shop* vamos apresentar uma breve descrição desse problema. No capítulo seguinte, descrevemos as regras de ramificação que propomos para o método de *branch-and-bound*. No capítulo 4, referimos os testes computacionais realizados, apresentamos os resultados obtidos para algumas das instâncias mais conhecidas do problema e elaboramos algumas conclusões.

2. Sequenciamento numa única máquina

O problema de sequenciamento numa única máquina surge, frequentemente, associado ao problema de *job-shop*. Um dos interesses deste problema reside no facto de muitos dos limites inferiores para o problema de *job-shop* serem obtidos através da sua resolução, quando se considera a relaxação das restrições de capacidade em todas as máquinas excepto numa.

O problema de sequenciamento numa única máquina pode ser descrito do seguinte modo: suponha-se que é dado um conjunto de tarefas $\{J_1, \dots, J_n\}$ para ser escalonado numa única máquina. Cada tarefa J_j ($j = 1, \dots, n$) pode começar o seu processamento a partir do instante \bar{r}_j , designado por data de disponibilidade “*release date*”, tem um tempo de processamento “*processing time*” \bar{p}_j e deverá ser entregue no instante $\bar{L}_j = \bar{C}_j + \bar{q}_j$ onde \bar{q}_j é o tempo de entrega “*delivery time*” e $\bar{O}(C, \bar{\cdot})_j = \bar{O}(S, \bar{\cdot})_j + \bar{O}(p, \bar{\cdot})_j$ a data de conclusão do processamento “*completion time*”, sendo \bar{S}_j o instante de início do processamento “*starting time*”.

O objectivo é sequenciar todas as tarefas $\{J_1, \dots, J_n\}$ na máquina, respeitando as datas de disponibilidade \bar{r}_j , e tal que o máximo atraso L_{\max} seja mínimo, onde $L_{\max} = \max_{j=1, \dots, n} L_j$. Este problema é designado por $1|\bar{r}_j|L_{\max}$, segundo a notação de Rinnooy Kan [14].

Podemos ainda considerar algumas generalizações do problema, tais como: a existência de relações de precedência entre as tarefas, isto é, restrições do tipo $J_k \rightarrow J_j$, significando que J_j não pode começar antes de J_k ter terminado. Este problema é designado por $1|prec, \bar{r}_j|L_{\max}$.

Este problema é *NP-hard*, contudo Jackson [9] verificou que relaxando a restrição que não admite interrupção de tarefas, o problema resultante $1|pmtn, prec, \bar{r}_j|L_{\max}$ pode ser resolvido em tempo polinomial obtendo-se, assim, um limite inferior para o problema de sequenciamento numa única máquina com restrições de precedência. Esse algoritmo designa-se por regra *EDD preemptive*.

Outra generalização para o problema consiste em considerar intervalos de tempo $T_{kj} \geq 0$, ou seja é imposto que entre o instante de conclusão do processamento, \bar{C}_k , da tarefa J_k e o instante de início do processamento de J_j , \bar{S}_j , tem de existir um intervalo de tempo de, pelo menos, T_{kj} . Este problema designa-se por $1|prec, \bar{r}_j, lags|L_{\max}$.

Lourenço [11] provou que o problema de sequenciamento numa única máquina com intervalos de tempo é *NP-hard*, mesmo quando se admite interrupção do processamento das

tarefas. Contudo, apresentou um caso especial do problema, em que todas as relações de precedência consistem em uma cadeia e que pode ser resolvido em tempo polinomial usando o algoritmo *early-late*. Este algoritmo é uma extensão do algoritmo de Jackson em que sucessivamente se vai identificando uma tarefa crítica e modificando a sua data de disponibilidade ou tempo de entrega dependendo se a tarefa está escalonada cedo ou tarde. A identificação da tarefa crítica, da sua tardeza ou prontidão é efectuada através de várias propriedades que estão descritas em Lourenço [12], assim como a descrição completa do algoritmo.

No contexto da relaxação do problema de *job-shop*, e tendo em conta a representação do problema por grafo disjuntivo $G = (O, A, E)$ bem como a selecção E' , a tarefa J_j corresponde à operação O_{ij} , a ser processada na máquina M_i (máquina cuja restrição de capacidade não é relaxada), $\bar{p}_j = p_{ij}$ e o conjunto de arcos AUE' definem as datas de disponibilidade $\bar{r}_j = r_{ij}$ e os tempos de entrega $\bar{q}_j = q_{ij}$ associados àquelas operações. Para cada operação O_{ij} , a sua data de disponibilidade é dada pelo comprimento do caminho mais longo entre o nodo 0 e o correspondente nodo no grafo $G' = (O, AUE')$, sendo o tempo de entrega dado pelo comprimento do caminho mais longo entre o nodo correspondente e o nodo $*$ do mesmo grafo, menos p_{ij} . Para obter estes valores é necessário resolver problemas de caminho mais longo. O método utilizado foi o algoritmo de Bellman [16].

Lageweg, Lenstra e Rinnooy Kan [10] observaram que AUE' pode definir restrições de precedência entre as operações a ser processadas na mesma máquina M_i : se as operações O_{ij} e O_{ik} são processadas na mesma máquina, então temos uma restrição de precedência $O_{ij} \rightarrow O_{ik}$ desde que exista em $G' = (O, AUE')$ um caminho entre os nodos O_{ij} e O_{ik} .

Posteriormente, Balas, Lenstra e Vazacopoulos [2] verificaram que, caso $O_{ij} \rightarrow O_{ik}$, então pode definir-se um intervalo de tempo $l_{ij,ik}$, entre essas operações, como sendo o comprimento do caminho mais longo entre os nodos O_{ij} e O_{ik} em $G' = (O, AUE')$ menos p_{ij} .

3. Regras de ramificação

As regras de ramificação aqui apresentadas tentam combinar o esquema de ramificação proposto por Brucker, Jurisch e Sievers [5] com o método de obtenção de limite inferior que se baseia na resolução de problemas de sequenciamento numa única máquina com intervalos de tempo, através do algoritmo *early-late* [11].

O problema de sequenciamento numa única máquina com intervalos de tempo considera que, entre o instante de conclusão do processamento de uma operação O_{ij} e o instante de início do processamento de outra operação O_{ik} , existe um intervalo de $l_{ij,ik}$ unidades de tempo. Portanto, o nosso principal objectivo vai ser construir um método de ramificação que privilegie operações que forneçam estes intervalos.

Para introduzir este método, comecemos por apresentar o conceito de bloco. Seja E' uma selecção completa, seja P um caminho crítico de uma solução admissível em $G' = (O, AUE')$ e seja $L(E')$ o comprimento de P . Uma sequência de nodos sucessivos em P é designada por

bloco em $G' = (O, AUE')$ se a sequência contém, pelo menos, dois nodos e as respectivas operações são processadas na mesma máquina. Um caminho crítico em $G' = (O, AUE')$ define os blocos B_1, \dots, B_k . Na Figura 1 apresenta-se um caminho crítico constituído por operações a processar em quatro máquinas distintas; neste caminho crítico definem-se três blocos.

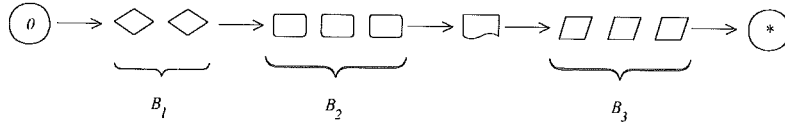


Figura 1 - Caminho crítico com três blocos

Um bloco B_p é definido pela sequência $O_{i_p p_1} \rightarrow O_{i_p p_2} \rightarrow \dots \rightarrow O_{i_p p_u}$, de u operações a processar na máquina i_p , em que $O_{i_p p_1}$ é a primeira operação a ser processada e $O_{i_p p_u}$ a última. Os conjuntos da forma $E_p^B = E_p \setminus \{O_{i_p p_1}\}$ ($p = 1, \dots, k$) designam-se por candidatos-antecedentes “Before-candidates” e os da forma $E_p^A = B_p \setminus \{O_{i_p p_u}\}$ ($p = 1, \dots, k$) por candidatos-sucessores “After-candidates”, segundo a notação de Brucker, Jurisch e Sievers [5].

A regra de ramificação consiste no seguinte: em cada nodo r , da árvore de pesquisa, obter uma solução admissível utilizando um processo heurístico e seja E_r' a selecção completa associada a essa solução; depois determinar, em $G' = (O, AUE_r')$, um caminho crítico P , os respectivos blocos B_1, \dots, B_k e, os conjuntos $E_1^B, \dots, E_k^B, E_1^A, \dots, E_k^A$. Para cada operação $O_{i_p q} \in E_p^V$ $p = 1, \dots, k; v = B, A$ obter um novo nodo, s , na árvore de pesquisa, descendente de r , a que associamos o grafo $G' = (O, AUE_s')$. Assim, em $G' = (O, AUE_s')$, $O_{i_p q}$ é colocada antes de todas as operações em B_p se $O_{i_p q} \in E_p^B$ ou, depois de todas as operações em B_p se $O_{i_p q} \in E_k^A$.

3.1 Regra “comprimento da maior cadeia” (CMC)

A título de exemplo, considere-se a instância do problema de *job-shop* proposta por Muth e Thompson, conhecida por MT10. Para essa instância é obtido o caminho crítico inicial que dá origem a três blocos: B_1, B_2 e B_3 , sendo o último destes blocos constituído por dez operações a ser processadas numa mesma máquina, a máquina 8. Na Figura 2 apresenta-se um grafo parcial contendo todas as operações deste bloco e às quais associamos o seu tempo de processamento. Além disso, indicamos as operações anterior e seguinte de cada uma delas.

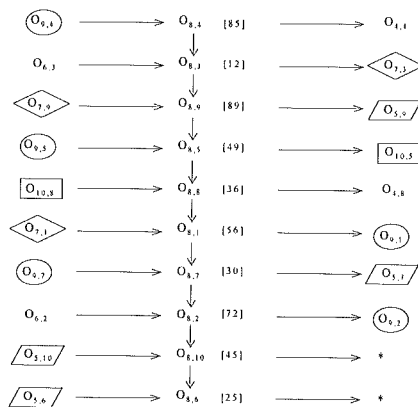


Figura 2 - Grafo parcial relativo às operações de B_3

O conjunto E_3^B é o seguinte: $E_3^B = \{O_{8,3}, O_{8,9}, O_{8,5}, O_{8,8}, O_{8,1}, O_{8,7}, O_{8,2}, O_{8,10}, O_{8,6}\}$. Consideremos a operação $O_{8,9}$, se passar a ser a primeira a processar neste bloco, serão definidas as orientações $O_{8,9} \rightarrow O_{8,j}$ ($j = 4,3,5,8,1,7,2,10,6$), em particular, é definida a orientação $O_{8,9} \rightarrow O_{8,3}$. Como se pode verificar na Figura 2, na tarefa, a operação anterior a $O_{8,9}$ e a seguinte a $O_{8,3}$ são ambas processadas na máquina 7. Pelo que, neste caso, é criada uma cadeia para essa máquina: $O_{7,9} \rightarrow O_{8,9} \rightarrow O_{8,3} \rightarrow O_{7,3}$, então, entre o instante de conclusão da operação $O_{7,9}$ e o de início do processamento de $O_{7,3}$ existe um intervalo de tempo que é no mínimo igual à soma dos tempos de processamento das operações $O_{8,9}$ e $O_{8,3}$, isto é, um intervalo de tempo de 101 unidades.

Considerando, agora, a operação $O_{8,5}$, verificamos que, caso esta operação venha a ser a primeira a processar neste bloco, ela dá origem a duas cadeias:

$$\begin{cases} O_{9,5} \rightarrow O_{8,5} \rightarrow O_{8,1} \rightarrow O_{9,1} \\ O_{9,5} \rightarrow O_{8,5} \rightarrow O_{8,2} \rightarrow O_{9,2} \end{cases}$$

De modo análogo, podemos atribuir à cadeia $O_{9,5} \rightarrow O_{8,5} \rightarrow O_{8,1} \rightarrow O_{9,1}$ um intervalo de 105 unidades de tempo e, à cadeia $O_{9,5} \rightarrow O_{8,5} \rightarrow O_{8,2} \rightarrow O_{9,2}$ um intervalo de 121 unidades de tempo. Como esta segunda cadeia fornece um maior intervalo, podemos associar à operação $O_{8,5}$ este intervalo.

Deste modo, associamos a cada operação um intervalo de tempo correspondente à cadeia com maior comprimento, assim:

- a uma operação O_{ij} à qual esteja associada apenas uma cadeia $O_{aj} \rightarrow O_{ij} \rightarrow O_{ib} \rightarrow O_{ab}$ atribuímos um intervalo de tempo correspondente a $p_{ij} + p_{ib}$;
- a uma operação O_{ij} à qual estejam associadas duas ou mais cadeias: $O_{a_1j} \rightarrow O_{ij} \rightarrow O_{ib_1} \rightarrow O_{a_1b_1}, \dots, O_{a_cj} \rightarrow O_{ij} \rightarrow O_{ib_c} \rightarrow O_{a_cb_c}$, em que c é o número de cadeias associado a O_{ij} , atribuímos um intervalo de tempo correspondente a $p_{ij} + \max_{p=1, \dots, c} p_{ib_p}$;
- a uma operação O_{ij} à qual não esteja associada qualquer cadeia atribuímos um intervalo de tempo de zero unidades.

Para as operações candidatas a última do respectivo bloco procedemos, de forma análoga, para obter o correspondente intervalo de tempo.

Considere-se a seguinte notação:

(nº de blocos) - número de blocos obtido no caminho crítico

$$B_k = \{O_{ij_1}, \dots, O_{ij_k}\} \quad (k = 1, \dots, (\text{nº de blocos}))$$

(nº de cadeias) $_{ij}^v$ - número de cadeias gerado pelo conjunto E_k^v

$$(k = 1, \dots, (\text{nº de blocos}); v = A, B)$$

(comp.cadeia) $_{ij}^v$ - comprimento da cadeia associado à operação $O_{ij} \in E_k^v$

$$(k = 1, \dots, (\text{nº de blocos}); v = A, B)$$

maq[pred(O_{ij})] - máquina em que é processada a operação que precede O_{ij}

maq[succ(O_{ij})] - máquina em que é processada a operação sucessora de O_{ij}

Para determinar o comprimento de cadeia a que cada operação dá origem caso passe a ser a primeira ou a última do bloco, pode então utilizar-se o seguinte algoritmo:

Regra - "comprimento da maior cadeia" (CMC)

Para cada bloco B_k ($k = 1, \dots, (\text{n}^\circ \text{ de blocos})$)
 $(\text{n}^\circ \text{ de cadeias})_k^v = 0$ ($v = A, B$)
 Para cada operação $O_{ij} \in E_k^B$ ($j = j_2, \dots, j_k$)
 $(\text{comp.cadeia})_{ij}^B = 0$
 Para cada operação $O_{il} \in B_k \setminus \{O_{ij}\}$
 Se $(\text{maq}[\text{pred}(O_{ij})] = \text{maq}[\text{succ}(O_{il})])$
 e $((\text{comp.cadeia})_{ij}^B < p_{il})$
 então $(\text{comp.cadeia})_{ij}^B = p_{il}$
 $(\text{comp.cadeia})_{ij}^B = (\text{comp.cadeia})_{ij}^B + p_{ij}$
 $(\text{n}^\circ \text{ de cadeias})_k^B = (\text{n}^\circ \text{ de cadeias})_k^B + 1$
 Para cada operação $O_{ij} \in E_k^A$ ($j = j_1, \dots, j_{k-1}$)
 $(\text{comp.cadeia})_{ij}^A = 0$
 Para cada operação $O_{il} \in B_k \setminus \{O_{ij}\}$
 Se $(\text{maq}[\text{succ}(O_{il})] = \text{maq}[\text{pred}(O_{ik})])$
 e $((\text{comp.cadeia})_{ij}^A < p_{il})$
 então $(\text{comp.cadeia})_{ij}^A = p_{il}$
 $(\text{comp.cadeia})_{ij}^A = (\text{comp.cadeia})_{ij}^A + p_{ij}$
 $(\text{n}^\circ \text{ de cadeias})_k^A = (\text{n}^\circ \text{ de cadeias})_k^A + 1$

Como, para utilizar o algoritmo *early-late*, é conveniente ter muitas cadeias, então uma boa opção será ramificar primeiro, a partir dos blocos que dão origem a maior número de cadeias. Deste modo, vamos tomar $\{(\text{n}^\circ \text{ de cadeias})_k^B, (\text{n}^\circ \text{ de cadeias})_k^A\}$ como o número de cadeias associado ao bloco k e designar esse valor por $(\text{n}^\circ \text{ de cadeias})_k$. Sendo a ordenação dos blocos efectuada, por ordem não crescente de $(\text{n}^\circ \text{ de cadeias})_k$ ($k = 1, \dots, (\text{n}^\circ \text{ de blocos})$).

Quanto à ordenação das operações, podemos também ordená-las, por ordem não crescente de $(\text{comp.cadeia})_{ij}^v$ ($j = j_1, \dots, j_k; v = A, B$).

Exemplo

Ilustremos o método, com o exemplo já utilizado:

$$B_1: O_{3,8} \rightarrow O_{3,6} \rightarrow O_{3,4} \rightarrow O_{3,5} \rightarrow O_{3,2} \rightarrow O_{3,10} \rightarrow O_{3,3}$$

$O_{3,1}$	$O_{3,8}$	$O_{3,6}$	$O_{3,4}$	$O_{3,5}$	$O_{3,2}$	$O_{3,10}$	$O_{3,3}$	$(\text{n}^\circ \text{ de cadeias})_1^n$
$(\text{n}^\circ \text{ de cadeias})_{3,1}^B$	-	0	179	0	185	156	0	3
$(\text{n}^\circ \text{ de cadeias})_{3,1}^A$	121	179	185	104	0	0	-	4
$(\text{n}^\circ \text{ de cadeias})_1$	4							

B₂: O_{9,3} → O_{9,4}

O _{9,1}	O _{9,3}	O _{9,4}	(nº de cadeias) ₂ ^v
(comp.cadeia) _{9,1} ^B	-	0	0
(comp.cadeia) _{9,1} ^A	0	-	0
(nº de cadeias) ₂		0	

B₃: O_{8,4} → O_{8,3} → O_{8,9} → O_{8,5} → O_{8,8} → O_{8,1} → O_{8,7} → O_{8,2} → O_{8,10} → O_{8,6}

O _{8,1}	O _{8,4}	O _{8,3}	O _{8,9}	O _{8,5}	O _{8,8}	O _{8,1}	O _{8,7}	O _{8,2}	O _{8,10}	O _{8,6}	(nº de cadeias) ₃ ^v
(comp.cadeia) _{8,1} ^B	-	0	101	121	85	68	102	0	134	114	7
(comp.cadeia) _{8,1} ^A	0	101	134	85	0	141	75	157	0	-	6
(nº de cadeias) ₂		7									

De acordo com o critério definido a ordem dos blocos é: B₃, B₁, B₂ e a das operações: O_{3,4}^A, O_{3,2}^B, O_{3,6}^A, O_{3,4}^B, O_{8,2}^A, O_{3,10}^B, O_{8,1}^A, O_{8,9}^A, O_{8,10}^B, O_{8,5}^B, O_{3,8}^A, O_{8,6}^B, O_{3,5}^A, O_{8,7}^B, O_{8,3}^A, O_{8,9}^B, O_{8,5}^A, O_{8,8}^B, O_{8,1}^A, O_{8,4}^B, O_{8,3}^A, O_{8,8}^B, O_{8,2}^A, O_{8,10}^B, O_{3,6}^B, O_{3,5}^A, O_{3,2}^A, O_{3,10}^B, O_{3,3}^A, O_{9,3}^A, O_{9,4}^B em que O_{ij}^v indica que O_{ij} ∈ E_k^v (v = B, A).

Neste método, a uma mesma operação pode estar associado um número reduzido de cadeias e um comprimento elevado; por exemplo, à operação O_{3,4}^A está associada uma única cadeia com comprimento 185. Esta operação é preferida neste método, em vez de O_{8,1}^A à qual estão associadas três cadeias, mas em que a de maior comprimento mede 141 unidades. No entanto, as três cadeias tomadas em conjunto, provavelmente, geram um intervalo com comprimento superior a 185 unidades. Por esta razão, construímos outra regra, com base nesta, mas alterando o comprimento do intervalo de tempo a associar a cada operação.

3.2 Regra “comprimento total de cadeia” (CTC)

Dado que a cada operação pode estar associada mais que uma cadeia com diferente comprimento. Ao escalonar uma operação como a primeira ou a última de um bloco, o comprimento total das cadeias obtidas pode ter maior importância e influência nos nodos a analisar posteriormente do que considerando apenas a maior cadeia. Por exemplo, consideremos a operação O_{8,5}. Já sabemos que, caso esta operação passe a ser a primeira do bloco, lhe serão associadas as cadeias:

$$\begin{cases} O_{9,5} \rightarrow O_{8,5} \rightarrow O_{8,1} \rightarrow O_{9,1} \\ O_{9,5} \rightarrow O_{8,5} \rightarrow O_{8,2} \rightarrow O_{9,2} \end{cases}$$

Vimos também que a cadeia O_{9,5} → O_{8,5} → O_{8,1} → O_{9,1} gera um intervalo de 105 unidades de tempo e, que O_{9,5} → O_{8,5} → O_{8,2} → O_{9,2} gera um intervalo de 121 unidades de tempo. Na regra CMC escolhemos, para associar a esta operação, a cadeia com maior comprimento. Vamos agora associar, a esta operação, um intervalo de 105 + 121 = 256 unidades de tempo.

Este método diverge do anterior apenas, no caso em que, a uma operação O_{ij} estão associadas duas ou mais cadeias: $O_{a_1j} \rightarrow O_{ij} \rightarrow O_{ib_1} \rightarrow O_{a_1b_1}, \dots, O_{a_cj} \rightarrow O_{ij} \rightarrow O_{ib_c} \rightarrow O_{a_cb_c}$, sendo c o número de cadeias associado a O_{ij} ; assim, em vez de um intervalo de tempo correspondente a $p_{ij} + \max_{p=1, \dots, c} p_{ib_p}$, associamos a O_{ij} um intervalo de $\sum_{p=1}^c (p_{ij} + p_{ib_p})$ unidades.

Pelo que, o novo método, pode ser descrito como se segue:

Regra - "comprimento total de cadeia" (CTC)

Para cada bloco B_k ($k = 1, \dots, (\text{n}^\circ \text{ de blocos})$)

$(\text{n}^\circ \text{ de cadeias})_k^v = 0$ ($v = A, B$)

Para cada operação $O_{ij} \in E_k^B$ ($j = j_2, \dots, j_k$)

$(\text{comp.cadeia})_{ij}^B = 0$

Para cada operação $O_{il} \in B_k \setminus \{O_{ij}\}$

Se $(\text{maq}[\text{pred}(O_{ij})] = \text{maq}[\text{succ}(O_{il})])$

então $(\text{comp.cadeia})_{ij}^B = (\text{comp.cadeia})_{il}^B + (p_{ij} + p_{il})$

$(\text{n}^\circ \text{ de cadeias})_k^B = (\text{n}^\circ \text{ de cadeias})_k^B + 1$

Para cada operação $O_{ij} \in E_k^A$ ($j = j_1, \dots, j_{k-1}$)

$(\text{comp.cadeia})_{ij}^A = 0$

Para cada operação $O_{il} \in B_k \setminus \{O_{ij}\}$

Se $(\text{maq}[\text{succ}(O_{il})] = \text{maq}[\text{pred}(O_{ik})])$

então $(\text{comp.cadeia})_{ij}^A = (\text{comp.cadeia})_{il}^A + (p_{ij} + p_{il})$

$(\text{n}^\circ \text{ de cadeias})_k^A = (\text{n}^\circ \text{ de cadeias})_k^A + 1$

Mais uma vez, a opção tomada corresponde a ramificar primeiro, a partir dos blocos que dão origem a maior número de cadeias, portanto, por ordem não crescente de $(\text{n}^\circ \text{ de cadeias})_k = \max \{ (\text{n}^\circ \text{ de cadeias})_k^B, (\text{n}^\circ \text{ de cadeias})_k^A \}$ ($k = 1, \dots, (\text{n}^\circ \text{ de blocos})$). E a ordenação das operações em cada bloco, por ordem não crescente de $(\text{comp.cadeia})_{ij}^v$ ($j = j_1, \dots, j_k; v = A, B$).

4. Resultados computacionais

Vamos estudar o comportamento computacional dos esquemas de ramificação acabados de descrever. Os programas foram implementados em C, sobre o sistema operativo LINUX (versão 1.2.3) e executados num PC 486 DX2/66Mhz.

Estes programas foram elaborados a partir dos programas escritos por Brucker, Jurisch e Sievers [5], uma vez que, gentilmente, nos cederam o seu código. Deste modo, apenas foram alteradas as unidades relativas ao esquema de ramificação e ao método de limite inferior.

As regras CMC e CTC foram construídas de forma a utilizar o algoritmo *early-late*. Pelo que, nos testes computacionais combinámos cada uma dessas regras de ramificação com o método de obtenção do limite inferior através do algoritmo *early-late* e também com o método

de obtenção do limite inferior através da regra *EDD preemptive* (método utilizado por Brucker, Jurisch e Sievers).

Com o objectivo de comparar as nossas regras de ramificação com a desenvolvida por Brucker, Jurisch e Sievers, estudámos ainda o comportamento da regra BJS (vamos designar por BJS a regra de ramificação desenvolvida por Brucker, Jurisch e Sievers) com ambos os métodos de obtenção do limite inferior.

Assim, o estudo computacional foi feito para as seguintes seis situações:

BJS		CMC		CTC	
EDD	early-late	EDD	early-late	EDD	early-late

Para cada uma destas combinações avaliamos os parâmetros:

- número de nodos pesquisados,
- tempo de execução,

até à obtenção da solução óptima.

O estudo foi feito para instâncias conhecidas da literatura, de acordo com os seguintes grupos de instâncias:

- de dimensão 10×10,
- de dimensão elevada (15×10, 20×10, 30×10).

Os dados destas instâncias podem ser obtidos na base de dados Beasley [3].

	Solução óptima	BJS		CMC		CTC	
		nº de nodos	tempo (seg.)	nº de nodos	tempo (seg.)	nº de nodos	tempo (seg.)
MT10	930	4249	513	11724	1493	<u>4959</u>	545
ABZ5	1234	2147	244	<u>972</u>	120	<u>972</u>	103
ABZ6	943	135	15	168	18	<u>153</u>	14
LA16	945	252	26	353	38	<u>351</u>	33
LA17	784	64	7	128	13	<u>122</u>	11
LA18	848	271	30	<u>821</u>	100	<u>821</u>	85
LA19	842	1463	165	<u>1937</u>	236	1986	207
LA20	902	1381	165	<u>1415</u>	175	<u>1415</u>	150
ORB1	1059	25301	3124	<u>37643</u>	4983	45159	5083
ORB2	888	2317	261	<u>1284</u>	156	<u>1284</u>	134
ORB3	1005	59309	7424	<u>112023</u>	15170	137422	15909
ORB4	1005	7848	941	9963	1362	<u>8875</u>	1030
ORB5	887	3135	342	4086	506	<u>3861</u>	411

Tabela 1 - Instâncias de dimensão 10×10

Da análise dos resultados obtidos verificámos, com alguma surpresa, que o número de nodos necessários à obtenção da solução óptima utilizando o algoritmo *early-late* ou utilizando a

regra *EDD preemptive* foi sempre o mesmo. No entanto, em termos de tempo o algoritmo *early-late* é mais moroso. Por esse facto, apenas vamos apresentar os resultados relativos à regra *EDD preemptive*.

Para instâncias de dimensão 10×10 apresentam-se, na Tabela 1, os valores dos parâmetros em estudo. Os valores sublinhados indicam o método que encontrou a solução óptima pesquisando o menor número de nodos. Os valores a itálico indicam as instâncias para as quais o método BJS pesquisa menos nodos que qualquer um dos apresentados, na obtenção da solução óptima.

Comparando os métodos CMC e CTC não podemos afirmar que um deles seja melhor que o outro, embora o método CTC pareça favorável. Na comparação entre estes e o método BJS, este parece superiorizar-se.

Quanto ao tempo de execução, podemos verificar que varia no mesmo sentido do número de nodos, e de uma forma aproximadamente proporcional.

Na Tabela 2 apresentamos os valores dos parâmetros em estudo, para instâncias de elevada dimensão. Para algumas destas instâncias foi necessário limitar o tempo de execução, fixado num máximo quatro dias (345600 segundos). Vamos identificar com * soluções correspondentes a pesquisas não concluídas; geralmente, estas soluções não são as óptimas. E os valores apresentados para os parâmetros em estudo são os necessários até à obtenção dessa solução. Por exemplo, consideremos a instância LA26, onde podemos verificar que o método que utiliza a regra CMC não concluiu a pesquisa, tendo obtido como melhor solução 1266. No período de quatro dias, foram pesquisados 531613 nodos; no entanto, o valor apresentado na tabela (484788), diz respeito ao número de nodos necessários até se obter a solução 1266 e o mesmo relativamente ao tempo de execução. Nesta Tabela os valores sublinhados indicam a melhor solução encontrada.

	Dimensão	BJS				CMC			CTC		
		Solução óptima	Solução obtida	nº de nodos	tempo (seg.)	Solução obtida	nº de nodos	tempo (seg.)	Solução obtida	nº de nodos	tempo (seg.)
LA21	1510	1046	*1057	765206	225890	* <u>1046</u>	208073	249365	*1046	242062	84934
LA22		927	927	10529	3488	<u>927</u>	6365	2041	927	7993	2169
LA23		1032	1032	6619	1829	<u>1032</u>	519	145	1032	3157	863
LA24		935	935	136630	45524	935	92736	33479	<u>935</u>	76831	24517
LA25		977	977	429524	135658	<u>977</u>	406705	136075	977	408647	118233
LA26	2010	1218	<u>1218</u>	56639	24076	*1266	484788	313729	*1292	23394	11305
LA27		1235-1256	* <u>1270</u>	185116	116043	*1293	314154	201967	*1391	543788	332863
LA28		1216	* <u>1273</u>	20042	7889	*1309	27141	16803	*1278	7821	3050
LA29		1120-1164	*1202	326106	182291	* <u>1195</u>	453193	291783	*1225	628420	331613
LA30		1355	<u>1355</u>	368	126	*1379	437	187	*1399	16296	6635
LA31	3010	1784	<u>1784</u>	8	3	*1797	11232	19339	*1788	20	9
LA32		1850	<u>1850</u>	1	0	<u>1850</u>	1	0	<u>1850</u>	1	0
LA33		1719	<u>1719</u>	77	41	*1753	2	1	*1729	1	0
LA34		1721	<u>1721</u>	15	6	*1798	3	1	*1823	10567	16321
LA35		1888	1888	24	11	1888	8	4	<u>1888</u>	7	3

Tabela 2 - Instâncias de dimensão elevada

Para as instâncias estudadas de dimensão 15×10, as novas regras de ramificação obtêm resultados significativamente melhores que a regra BJS, especialmente a regra CMC. Já para as instâncias de dimensão 20×10 e 30×10 não podemos dizer o mesmo, verificando-se neste caso, que a regra BJS é mais forte.

Em termos gerais e comparando as três regras, não podemos afirmar que alguma delas seja melhor que as restantes. Comparando apenas as duas regras desenvolvidas neste trabalho, parece-nos que a regra CTC será mais adequada para instâncias de média dimensão, enquanto que a CMC para instâncias de elevada dimensão.

As novas regras de ramificação obtêm melhores resultados para algumas instâncias mas não são sempre superiores às utilizadas por Brucker, Jurisch e Sievers [5]. A principal razão deve-se a que sendo a dimensão das instâncias pequena, o aparecimento das cadeias tem uma influência menor na prática que a esperada teoricamente.

Os bons resultados obtidos para as instâncias de dimensão elevada são indicativo que um método de *branch-and-bound* com estas regras de ramificação pode conduzir a melhores resultados em instâncias reais e de elevada dimensão. Este será um dos trabalhos a prosseguir no futuro em conjunto com a obtenção de melhores limites inferiores baseados na resolução de

problemas de uma única máquina com intervalos de tempo e sua combinação com as regras de ramificação.

5. Referências

- [1] D. Applegate, W. Cook, A computacional study of the job-shop scheduling problem, *ORSA - Journal on Computing* 3 (1991) 149-156.
- [2] E. Balas, J. K. Lenstra, A. Vazacopoulos, One machine scheduling with delayed with precedence constraints, Technical report, GSIA, Carnegie Mellow University, Pittsburgh, PA, USA, 1992.
- [3] J. E. Beasley, OR-Library: Distributing Test Problems by Electronic Mail, *Journal of Operational Research Society*, vol 41, nº11 (1990) 1069-1072.
- [4] E. A. Boyd, Integer Programming for job-shop scheduling and a related problem, preprint, 1995.
- [5] P. Brucker, B. Jurish, B. Sievers, A branch and bound algorithm for the job-shop scheduling problem, *Discrete Applied Mathematics* 49 (1994) 107-127.
- [6] J. Carlier, E. Pinson, An algorithm for solving the job-shop problem, *Management Science* 35 (1989) 164-176.
- [7] J. Carlier, E. Pinson, A practical use of Jackson's preemptive schedule for solving the job-shop problem, *Annals of Operations Research* 26 (1990) 269-287.
- [8] M.R. Garey, D.S. Johnson, R. Sethi, The complexity of flowshop and jobshop scheduling, *Mathematics of Operations Research* 1 (1976) 117-129.
- [9] J. R. Jackson, Scheduling a production line to minimize maximum tardiness, Research Report 43, Management Science Research Project, University of California, Los Angeles, 1955.
- [10] B.J. Lageweg, J.K. Lenstra, A.H.G. Rinnooy Kan, Minimizing maximum lateness on one machine: computational experience and some applications, *Statistica Neerlandica* 30 (1996) 25-41.
- [11] H. Lourenço, A computacional study of job-shop and the flow-shop scheduling problems, Ph.D. thesis, School of OR&IE, Cornell University, Ithaca, NY, USA, 1993.
- [12] H. Lourenço, The one-machine scheduling problem with lags, working paper nº 4/94 do Centro de Investigação Operacional da Universidade de Lisboa, 1994.
- [13] J.F. Muth, G.L. Thompson, *Industrial Scheduling*, Engewood Cliffs, N.J., Prentice Hall, 1963.
- [14] A.H.G. Rinnooy Kan, *Machine scheduling problems: classification, complexity and computations*, Martinus Nijhoff, The Hague, Holland, 1976.
- [15] B. Roy, B. Sussmann, Les problèmes d'ordonnement avec contraintes disjonctives, Note DS no. 9 bis, SEMA, Paris, 1964.
- [16] M.M. Syslo, N. Deo, J.S. Kowalik, *Discrete Optimization Algorithms with Pascal Programs*, Prentice-Hall, 1983.

SOLUTION OF SOME LINEAR COMPLEMENTARITY PROBLEMS ARISING IN VARIATIONAL MODELS OF MECHANICS

Joaquim J. Júdice*

Departamento de Matemática
FCT - Universidade de Coimbra
3000 Coimbra - Portugal

M.C.Soares

Departamento de Matemática
FCT - Universidade Nova de Lisboa
2825 Monte Caparica - Portugal

Abstract

In this paper we investigate the use of a block principal pivoting and an interior-point algorithms for the solution of some linear complementarity problems arising in some models of mechanics. We describe the extensions of the algorithms that deal with these special linear complementarity problems. Computational experience with MATLAB implementations of these algorithms is included and shows the appropriateness of these techniques to deal with these models.

Resumo

Neste artigo é investigada a utilidade dos algoritmos pivotais principais por blocos e de pontos interiores para a resolução de problemas lineares complementares que surgem em alguns modelos de mecânica. Extensões desses processos que tiram partido das estruturas dos problemas lineares complementares são descritos. Alguma experiência computacional com implementações MATLAB desses algoritmos é apresentada e indica a eficiência destas técnicas na resolução desses problemas.

Keywords

Variational inequalities, linear complementarity problems, partial differential equations, sparse matrices.

1. Introduction

The Linear Complementarity Problem (LCP) consists of finding vectors $z \in \mathbb{R}^n$ and $w \in \mathbb{R}^n$ such that

$$w = q + Mz, z \geq 0, w \geq 0, z^T w = 0 \quad (1)$$

for a given vector $q \in \mathbb{R}^n$ and a square matrix M of order n . This problem has been studied intensively during the past several years. The existence and uniqueness of a solution to the LCP is usually associated with classes of matrices [5,14]. One of the most important results of this type states that the LCP has a unique solution for each vector $q \in \mathbb{R}^n$ if and only if M is a P matrix [5,14]. We recall that M is a P matrix if and only if its principal minors are positive [5,14]. The class of P matrices contains the so-called Positive Definite (PD) and Nonsingular M (NSM) matrices, which are defined by the following equivalences

$$M \in \text{PD} \Leftrightarrow x^T M x > 0 \text{ for all } x \neq 0$$

$$M \in \text{NSM} \Leftrightarrow M \in \text{P} \text{ and } m_{ij} \leq 0 \text{ for all } i \neq j$$

* Supported by project PRAXIS XXI 2/2.1/MAT/346/94

It is also important to add that these two classes have a nonempty intersection and the LCP can be solved in polynomial-time when M belongs to each one of them [12,14].

Several algorithms have been designed for the LCP when M is a P matrix [5,14]. Interior-Point methods have shown to be quite efficient for the solution of LCPs with P matrices [6,10,12,19,22]. Special forms of these procedures are even polynomial when M is a PD matrix [12,22]. On the other hand, the LCP can be processed by a polynomial block principal pivoting algorithm when M is a NSM matrix [14]. This last procedure is also quite efficient to deal with large LCPs with sparse NSM matrices [6].

A large number of applications of the LCP in science, economics and engineering has been mentioned in the literature [5,14]. In many situations the complementarity problem takes the following more general form

$$\left. \begin{aligned} w &= q + Mz \\ a_i &\leq z_i \leq b_i \\ z_i = a_i < b_i &\Rightarrow w_i \geq 0 \\ z_i = b_i > a_i &\Rightarrow w_i \leq 0 \\ a_i < z_i < b_i &\Rightarrow w_i = 0 \\ z_i = a_i = b_i &\Rightarrow -\infty < w_i < +\infty \end{aligned} \right\} i = 1, 2, \dots, n \quad (2)$$

where a_i and b_i are real numbers or $\pm \infty$ satisfying $a_i \leq b_i$ for all $i = 1, \dots, n$. This problem is usually called the Mixed Linear Complementarity Problem (MLCP) and reduces to the LCP when $a_i = 0$ and $b_i = +\infty$ for all $i = 1, 2, \dots, n$.

Let \bar{M} be the principal submatrix of M associated with the indices i such that $a_i < b_i$. Then it has been shown that the MLCP has a unique solution for each vector $q \in \mathbb{R}^n$ if \bar{M} is a P matrix [5]. Furthermore the MLCP can be processed by special interior point [19,22] and block principal pivoting [11,17] algorithms that are polynomial when \bar{M} is a PD or a NSM matrix, respectively.

In this paper we investigate the solution of MLCPs arising in variational models of mechanics. These models are defined by a variational inequality in a special Hilbert space [9]. Its discretization by appropriate finite-elements leads into an Affine Variational Inequality Problem, that is in turn equivalent to the MLCP (2). Furthermore the principal submatrix \bar{M} of M mentioned above belongs to both the classes PD and NSM and this recommends the use of interior-point and block principal pivoting algorithms for the solution of the MLCP. Some simple modifications of these methods are required to deal with the so-called boundary nodes. The algorithms and these simple extensions are fully described in this paper, together with some hints about their implementations in a finite-element code.

The algorithms have been tested on the solution of three well-known models of mechanics, namely an obstacle problem [8,18], a problem of flow through a porous dam [15] and an elastic plastic torsion problem [3,9]. A MATLAB implementation including the finite-element discretization and the complementarity algorithms has been designed to process these models.

Computational experience with some instances of these models is included and shows that the complementarity algorithms are quite efficient to deal with these problems. In fact, they usually require a small number of iterations to get a solution that is numerically as good as the one that would be given by the MATLAB implementation for a similar model without constraints in the nonboundary nodes.

The structure of the paper is as follows. In section 2 the variational model is introduced together with the three applications mentioned before. The discretization of the model that leads to the MLCP is discussed in section 3. The interior-point and the block principal pivoting algorithms are described in section 4. Computational experience with these procedures on the three representative models is presented in section 5. Finally some conclusions are drawn in the last section of the paper.

2. The Variational Model

Let Ω be a bounded domain of \mathbb{R}^2 with boundary $\partial\Omega$ and let $H^1(\Omega)$ be the well-known Sobolev space [4] of functions $u : \mathbb{R}^2 \rightarrow \mathbb{R}^1$. Then for all $u, v \in H^1(\Omega)$ we can consider the bilinear form

$$\begin{aligned} c(u,v) &= \int_{\Omega} \left\{ \frac{\partial u}{\partial x_1} \cdot \frac{\partial v}{\partial x_1} + \frac{\partial u}{\partial x_2} \cdot \frac{\partial v}{\partial x_2} \right\} dx_1 dx_2 \\ &= \int_{\Omega} \nabla u \nabla v dx \end{aligned} \tag{3}$$

where ∇z represents the gradient of z . Let a and b be two given functions from \mathbb{R}^2 to \mathbb{R}^1 belonging to $H^1(\Omega)$ or $\pm\infty$ and let

$$K = \left\{ v \in H^1(\Omega) : a \leq v \leq b \text{ in } \Omega, v = g \text{ in } \partial\Omega \right\} \tag{4}$$

where $g \in H^1(\partial\Omega)$ is also a given function. For a function $f \in H^1(\Omega)$, consider the following optimization problem

$$\text{Min}_{v \in K} J(v) = \frac{1}{2} c(v,v) - L(v) \tag{5}$$

where $c(v,v)$ is defined by (3) and

$$L(v) = \int_{\Omega} f v dx = \int_{\Omega} f(x_1, x_2) v(x_1, x_2) dx_1 dx_2 \tag{6}$$

As is discussed in [8], this optimization problem has a unique solution $\bar{u} \in K$ satisfying the variational inequality:

Find $u \in K$ such that

$$c(u, v-u) \geq L(v-u) \text{ for all } v \in K \tag{7}$$

The purpose of this paper is to numerically solve this variational inequality (7) for c , L and K given by (3), (6) and (4) respectively. This is the subject of the next sections of the paper. Next, we introduce three problems that have been mentioned in the literature and can be seen as special cases of the model stated before for some appropriate choices of the function f and set K .

(i) Obstacle Problem

Consider an elastic membrane Ω fixed on its boundary $\partial\Omega$ ($v = 0$ on $\partial\Omega$) and an obstacle given by a function $\Psi : \mathbb{R}^2 \rightarrow \mathbb{R}^1$ of $H^1(\Omega)$. The problem consists of finding the vertical displacement u of this membrane under the effects of a vertical force and forced to lie over the obstacle. As is discussed in [8,18] this model is defined by (7), where

$$K = \{v \in H^1(\Omega) : v \geq \Psi \text{ in } \Omega, v = 0 \text{ in } \partial\Omega\} \tag{8}$$

(ii) Problem of Flow through a Porous Dam

Consider a reservoir full of water with height h_1 and a porous rectangular dam Ω of width w which separates the reservoir from another one of height $h_2 < h_1$ (see figure 1). The problem [1,2,13,15] consists of finding the pressure $p : \mathbb{R}^2 \rightarrow \mathbb{R}^1$ of the water that allows its movement from one reservoir to the other through the dam. By introducing the function

$$u(x_1, x_2) = \int_{x_2}^{h_1} p(x_1, x_2) dx_2$$

it is possible to show [2,21] that this problem is a special case of the model (7) with $f(x_1, x_2) = -1$ and

$$K = \{v \in H^1(\Omega) : v \geq 0 \text{ in } \Omega, v = g \text{ in } \partial\Omega\} \tag{9}$$

where $g : \mathbb{R}^2 \rightarrow \mathbb{R}^1$ is given by

$$g(x_1, x_2) = \begin{cases} \frac{(h_1 - x_2)^2}{2} & \text{in AB} \\ \frac{(h_2 - x_2)^2}{2} & \text{in CD} \\ \frac{h_1^2(w - x_1) + h_2^2 x_1}{2w} & \text{in BC} \\ 0 & \text{in DFA} \end{cases} \tag{10}$$

and the points A,B,C,D,E,F are displayed in figure 1.

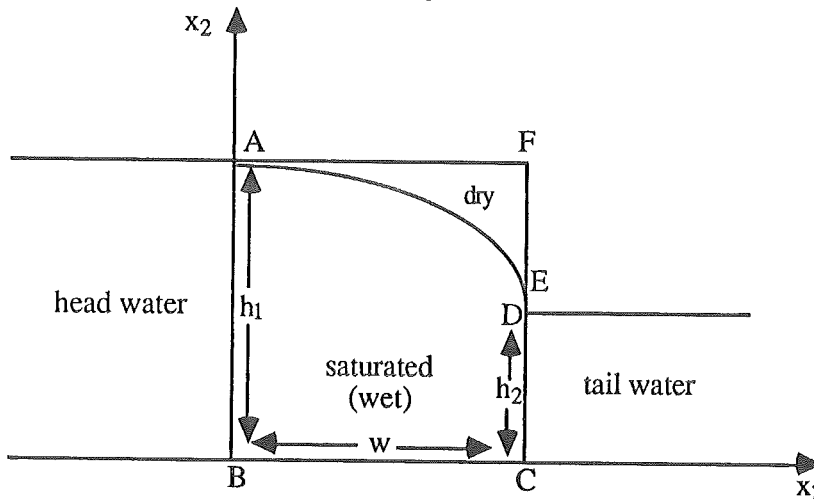


Figure 1 - Flow through a porous rectangular dam

(iii) Elastic-Plastic Torsion Problem

Consider a long cylindrical bar of cross section Ω made of an isotropic, elastic and perfectly plastic material. This bar is subject to a stress given by a vector δ with a constant C torsion angle. The problem consists of finding the stress vector δ that maintains the bar in equilibrium [3,9]. This vector δ is given by the gradient ∇u of the solution u of the variational inequality (7), where $f(x_1, x_2) = C$,

$$K = \left\{ v \in H^1(\Omega) : |v(x_1, x_2)| \leq d((x_1, x_2), \partial\Omega) \text{ in } \Omega, v = 0 \text{ in } \partial\Omega \right\} \quad (11)$$

and $d((x_1, x_2), \partial\Omega)$ represents the distance from the point (x_1, x_2) to the boundary $\partial\Omega$ of Ω .

3. Discretization of the Model and the MLCP

Consider the general model described in the previous section:

Find $z \in K$ such that

$$\int_{\Omega} \nabla z \nabla (v-z) \, dx_1 \, dx_2 \geq \int_{\Omega} f(v-z) \, dx_1 \, dx_2 \text{ for all } v \in K \quad (12)$$

As is discussed in [4,16], the finite-element method is used to find a discretization of this model. A standard triangulation T_h over $\bar{\Omega} = \Omega \cup \partial\Omega$ is considered and $H^1(\Omega)$ is approximated by the space V_h of the piecewise polynomials v_h of degree less than or equal to one and K is replaced by K_h of the elements $v_h \in V_h$ satisfying

$$\begin{aligned} a(x_1, x_2) \leq v_h(x_1, x_2) \leq b(x_1, x_2) \text{ for all } (x_1, x_2) \in T_h \cap \Omega \\ v_h(x_1, x_2) = g(x_1, x_2) \text{ for all } (x_1, x_2) \in T_h \cap \partial\Omega \end{aligned}$$

Instead of solving the model (12), the following variational inequality problem in the space V_h is considered:

Find $z_h \in K_h$ such that

$$\int_{\Omega} \nabla z_h \nabla (v_h - z_h) \, dx_1 \, dx_2 \geq \int_{\Omega} f(v_h - z_h) \, dx_1 \, dx_2 \text{ for all } v_h \in K_h \quad (13)$$

Then V_h is a finite dimensional space of dimension n equal to the number of nodes of the triangulation T_h . If $\{\Phi_1, \dots, \Phi_n\}$ is the basis for this space discussed in [4,16], we can write

$$v_h(x_1, x_2) = \sum_{j=1}^n v_j \Phi_j(x_1, x_2), \quad z_h(x_1, x_2) = \sum_{j=1}^n z_j \Phi_j(x_1, x_2)$$

for each $(x_1, x_2) \in \Omega \cup \partial\Omega$. Substituting these functions v_h and z_h in (13), we obtain

$$\begin{aligned} \int_{\Omega} \left(\sum_{j=1}^n z_j \frac{\partial \Phi_j}{\partial x_1} \sum_{i=1}^n (v_i - z_i) \frac{\partial \Phi_i}{\partial x_1} + \sum_{j=1}^n z_j \frac{\partial \Phi_j}{\partial x_2} \sum_{i=1}^n (v_i - z_i) \frac{\partial \Phi_i}{\partial x_2} \right) dx_1 \, dx_2 \geq \\ \int_{\Omega} f \sum_{i=1}^n (v_i - z_i) \Phi_i \, dx_1 \, dx_2 \end{aligned}$$

If we write

$$m_{ij} = \int_{\Omega} \left(\frac{\partial \Phi_i}{\partial x_1} \frac{\partial \Phi_j}{\partial x_1} + \frac{\partial \Phi_i}{\partial x_2} \frac{\partial \Phi_j}{\partial x_2} \right) dx_1 \, dx_2 \quad (14)$$

$$q_i = - \int_{\Omega} f \Phi_i(x_1, x_2) dx_1 \, dx_2$$

for $i, j = 1, \dots, n$, $M = [m_{ij}] \in \mathbb{R}^{n \times n}$ and $q = [q_i] \in \mathbb{R}^n$, we get to the following Affine Variational Inequality Problem in \mathbb{R}^n :

Find $z \in K$ such that

$$(v-z)^T (q+Mz) \geq 0 \text{ for all } v \in K \quad (15)$$

where

$$K = \left\{ v \in \mathbb{R}^n : a_i \leq v_i \leq b_i, i = 1, \dots, n, v_i = g_i, \text{ in } \partial\Omega \cap T_h \right\} \quad (16)$$

Note that a_i, b_i and g_i are the values of the functions a, b and g at the nodes of the triangulation.

This variational inequality problem is in turn equivalent to the following MLCP [5,14]

$$\left. \begin{array}{l} w = q + Mz \\ a_i \leq z_i \leq b_i \\ z_i = a_i \Rightarrow w_i \geq 0 \\ z_i = b_i \Rightarrow w_i \leq 0 \\ a_i < z_i < b_i \Rightarrow w_i = 0 \\ z_i = g_i, -\infty < w_i < +\infty, i \in L \end{array} \right\} i \in J \quad (17)$$

where J and L are the sets of nonboundary and boundary nodes, respectively.

As is discussed in [4,8,9] the elements of the matrix M satisfy

$$m_{ij} = m_{ji} \leq 0 \text{ for all } i \neq j$$

Furthermore the principal submatrix M_{JJ} of M is PD. So M_{JJ} is a PD and a NSM matrix [5] and this property suggests the use of polynomial block principal pivoting and interior-point algorithms for the solution of the MLCP (17). This is the topic of the next section.

4. Algorithms for the MLCP

4.1 Interior-Point Algorithm

Consider the MLCP (17) and suppose initially that all the lower and upper bounds a_i and b_i are finite. Due to the equivalence between a MLCP and an augmented LCP discussed in [5], we can write the MLCP (17) in the following equivalent form

$$\begin{bmatrix} w \\ v \end{bmatrix} = \begin{bmatrix} q \\ b \end{bmatrix} + \begin{bmatrix} M & I \\ -I & 0 \end{bmatrix} \begin{bmatrix} z \\ \lambda \end{bmatrix} \quad (18)$$

$$z_i = g_i, \lambda_i = 0, i \in L \quad (19)$$

$$(z_j - a_j)w_j = 0, j \in J \quad (20)$$

$$v_j \lambda_j = 0, j \in J \quad (21)$$

$$z_j \geq a_j, w_j \geq 0, v_j \geq 0, \lambda_j \geq 0, j \in J \quad (22)$$

where I is the identity matrix of order n . We note that the values of the variables $\lambda_i, i \in L$, have no effect on the solution of the MLCP. Hence we have set them equal to zero for simplicity.

The interior-point algorithm is an iterative method that attains to find a solution for the system of nonlinear equations (18)-(21) by maintaining the inequalities (22) strict (hence the name interior-point) and forcing the products (20) and (21) to go to zero at the same rate. To fulfil this last goal, a central parameter μ_k is introduced in each iteration k to replace the equations (20) and (21) by

$$(z_j - a_j) w_j = \mu_k, \quad j \in J \tag{23}$$

$$\lambda_j v_j = \mu_k, \quad j \in J \tag{24}$$

Next, we describe an iteration k of the algorithm. Suppose that a point $(z^k, w^k, \lambda^k, v^k)$ satisfying (19) and (22) strictly is at hand. Then the central parameter μ_k is computed by

$$\mu_k = \delta \frac{\sum_{j \in J} [(z_j^k - a_j) w_j^k + \lambda_j^k v_j^k]}{\bar{n}} \tag{25}$$

where \bar{n} is the number of elements of the set J and $0 < \delta < 1$. The Newton's direction $(\alpha^k, \beta^k, \gamma^k, \theta^k)$ for the system of equations (18), (19), (23) and (24) is next found by

$$\left| \begin{array}{cccc|cccc} I & & & 0 & & -M & & -I \\ & & & & & & & \\ & & & & & & & \\ & 0 & & I & & I & & 0 \\ \hline Z_J^k - A_J & 0 & 0 & 0 & W_J^k & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & I_L & 0 & 0 \\ 0 & 0 & \Lambda_J^k & 0 & 0 & 0 & V_J^k & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & I_L \end{array} \right| = \left| \begin{array}{c} \alpha^k \\ \beta^k \\ \gamma_J^k \\ \gamma_L^k \\ \theta_J^k \\ \theta_L^k \end{array} \right| = \left| \begin{array}{c} -(w^k - q - Mz^k - \lambda^k) \\ -(v^k - b + z^k) \\ \mu_k e_J - (Z_J^k - A_J) W_J^k e_J \\ 0 \\ \mu_k e_J - V_J^k \Lambda_J^k e_J \\ 0 \end{array} \right| \tag{26}$$

where $Z_J^k, A_J, \Lambda_J^k, V_J^k$ and W_J^k are diagonal matrices with diagonal elements equal to the components of the vectors z^k, a, λ^k, v^k and w^k respectively, I_L is the identity matrix of order equal to the number of elements $|L|$ of the set L and e_j is the vector of dimension $|J|$ with components equal to one.

After computing the direction $(\alpha^k, \beta^k, \theta^k, \gamma^k)$, it is necessary to calculate the stepsize $t_k > 0$ such that the next iterate also satisfies the conditions (19) and (22) strictly. It is easy to see that t_k should take the form

$$t_k = \eta_k \min \left\{ \min \left\{ \frac{z_j^k - a_j}{-\gamma_j^k}, \gamma_j^k < 0, j \in J \right\}, \min \left\{ \frac{w_j^k}{-\alpha_j^k}, \alpha_j^k < 0, j \in J \right\}, \right. \\ \left. \min \left\{ \frac{v_j^k}{-\beta_j^k}, \beta_j^k < 0, j \in J \right\}, \min \left\{ \frac{\lambda_j^k}{-\theta_j^k}, \theta_j^k < 0, j \in J \right\} \right\} \tag{27}$$

for some $0 < \eta_k < 1$. After t_k is found, a new point $(z^{k+1}, w^{k+1}, \lambda^{k+1}, v^{k+1})$ is obtained from the search direction and the stepsize t_k . This completes the description of the iteration k . The process is repeated until an approximate solution of the nonlinear system (18)-(21) is at hand.

It follows from the description of the interior-point algorithm that the main work in each iteration relies on the solution of the linear system (26) for the computation of the Newton's direction. On the other hand, in the solution of variational models such as (7) by the finite-element method, the matrix M is usually computed automatically. Furthermore finite-element

codes incorporate linear equations solvers with this matrix. So in the solution of the model (7) it is recommendable the Newton's direction to be computed by using this linear equations solver with this matrix M. Next, we describe a procedure that achieves this goal.

Consider again the system (26). It immediately follows that

$$\left. \begin{aligned} \gamma_1^k = 0, \quad \gamma_j^k &= \frac{-(z_j^k - a_j)(w_j^k + \alpha_j^k) + \mu_k}{w_j^k}, \\ \theta_1^k = 0 \text{ and } \theta_j^k &= \frac{-\lambda_j^k(v_j^k + \beta_j^k) + \mu_k}{v_j^k}, \quad l \in L \text{ and } j \in J. \end{aligned} \right\} \quad (28)$$

Substituting these values in the first set of equalities, we get

$$\left(\begin{array}{c} \left[\begin{array}{cc} I & 0 \\ 0 & I \end{array} \right] - \left[\begin{array}{cc} -M & -I \\ I & 0 \end{array} \right] \left[\begin{array}{ccc|ccc} (W_J^k)^{-1}(Z_J^k - A_J) & 0 & 0 & 0 & 0 & 0 \\ \hline 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & (V_J^k)^{-1}\Lambda_J^k & 0 & 0 \end{array} \right] \end{array} \right) \begin{bmatrix} \alpha \\ \beta \end{bmatrix} =$$

$$\begin{bmatrix} -w^k + q + Mz^k + \lambda^k \\ -v^k + b - z^k \end{bmatrix} + \left[\begin{array}{cc} M & I \\ -I & 0 \end{array} \right] \begin{bmatrix} \mu_k(W_J^k)^{-1}e_J - (Z_J^k - A_J)e_J \\ 0 \\ \mu_k(V_J^k)^{-1}e_J - \Lambda_J^k e_J \\ 0 \end{bmatrix}.$$

This system can be rewritten as

$$\begin{bmatrix} M_{JJ}(W_J^k)^{-1}(Z_J^k - A_J) + I_J & (V_J^k)^{-1}\Lambda_J^k & 0 & 0 \\ -(W_J^k)^{-1}(Z_J^k - A_J) & I_J & 0 & 0 \\ M_{LJ}(W_J^k)^{-1}(Z_J^k - A_J) & 0 & I_L & 0 \\ 0 & 0 & 0 & I_L \end{bmatrix} \begin{bmatrix} \alpha_J^k \\ \beta_J^k \\ \alpha_L^k \\ \beta_L^k \end{bmatrix} = \begin{bmatrix} f_J \\ d_J \\ f_L \\ d_L \end{bmatrix}, \quad (29)$$

where $f = \begin{bmatrix} f_J \\ f_L \end{bmatrix}$ and $d = \begin{bmatrix} d_J \\ d_L \end{bmatrix}$ are defined as follows

$$f_j = -w_j^k + q_j + \tilde{h}_j - \frac{\mu_k}{v_j^k}, \quad j \in J; \quad f_l = -w_l^k + q_l, \quad l \in L \quad (30)$$

$$d_j = -v_j^k + b_j - a_j + \frac{\mu_k}{w_j^k}, \quad j \in J; \quad d_l = -v_l^k + b_l, \quad l \in L$$

with $\tilde{h} = Mh$ and h the vector given by

$$h_j = \frac{\mu_k}{w_j^k} + a_j, \quad j \in J, \quad h_l = 0, \quad l \in L \quad (31)$$

The second and fourth sets of equations in (29) and the definition of the vector d imply that β^k satisfies

$$\beta_j^k = -v_j^k + b_j - a_j + \frac{(z_j^k - a_j)\alpha_j^k - \mu_k}{w_j^k}, \quad j \in J, \quad \beta_l^k = -v_l^k + b_l, \quad l \in L \quad (32)$$

In order to complete the computation of the Newton's direction, we have to finally calculate the vector α^k . To do this, let

$$\tilde{M} = \begin{bmatrix} M_{JJ} & 0 \\ M_{LJ} & 0 \end{bmatrix} + \begin{bmatrix} (V_J^k)^{-1} \Lambda_J^k + (Z_J^k - A_J)^{-1} W_J^k & 0 \\ 0 & I_L \end{bmatrix} \tag{33}$$

$$\tilde{\alpha} = \begin{bmatrix} (W_J^k)^{-1} (Z_J^k - A_J) & 0 \\ 0 & I_L \end{bmatrix} \begin{bmatrix} \alpha_J^k \\ \alpha_L^k \end{bmatrix}$$

and \tilde{f} be the vector defined by

$$\tilde{f}_j = f_j - \frac{\lambda_j^k d_j}{v_j^k}, j \in J, \quad \tilde{f}_l = f_l, l \in L \tag{34}$$

Hence, it is easy to see that α^k can be computed by first solving the system

$$\tilde{M} \tilde{\alpha} = \tilde{f} \tag{35}$$

and then set

$$\alpha_j^k = \frac{\tilde{\alpha}_j w_j^k}{z_j^k - a_j}, j \in J, \quad \alpha_l^k = \tilde{\alpha}_l, l \in L \tag{36}$$

It is also important to add that the matrix \tilde{M} of the linear system (35) can be easily computed in each iteration k of the algorithm. The following procedure performs this task.

PROCEDURE COMPMAT

Let k be the iteration number.
 If $k = 0$, set for each $l \in L$
 $m_{jl} = 0, m_{il} = 0, j \in J, i \in L$
 If $k \geq 1$, do

$$\tilde{m}_{jj} = m_{jj} + \frac{\lambda_j^k}{v_j^k} + \frac{w_j^k}{z_j^k - a_j}, j \in J$$

$$\tilde{m}_{ll} = 1, l \in L$$

$$\tilde{m}_{ij} = m_{ij}, i \neq j$$

The previous discussion leads to the conclusion that the Newton's direction $(\alpha^k, \beta^k, \theta^k, \gamma^k)$ can be computed by the following procedure:

PROCEDURE FINDDIRECTION

- 1 - Update \tilde{M} by procedure COMPMAT.
- 2 - Compute d, f, h and \tilde{f} by (30),(31) and (34) with $\tilde{h} = \tilde{M}h$
- 3 - Solve $\tilde{M}\tilde{\alpha} = \tilde{f}$
- 4 - Compute:
 - α^k by (36)
 - β^k by (32)
 - θ^k, γ^k by (28)

We are now in position to present the steps of the interior-point algorithm for the solution of the MLCP (17).

INTERIOR-POINT ALGORITHM

Let $(z^0, w^0, v^0, \lambda^0)$ be an initial point such that $z_1^0 = g_1$, $\lambda_1^0 = 0$ for all $l \in L$ and $z_j^0 > a_j$, $w_j^0 > 0$, $v_j^0 > 0$ and $\lambda_j^0 > 0$ for all $j \in J$. Let $k = 0$ and ϵ_i , $i = 1, 2$ be two tolerances.

For $k = 0, 1, \dots$

- Compute μ_k by (25).
- Find $(\alpha^k, \beta^k, \theta^k, \gamma^k)$ by procedure FINDDIRECTION.
- Compute stepsize t_k by (27) with $0 < \eta_k < 1$
- Set $z^{k+1} = z^k + t_k \gamma^k$, $w^{k+1} = w^k + t_k \alpha^k$, $\lambda^{k+1} = \lambda^k + t_k \theta^k$ and $v^{k+1} = v^k + t_k \beta^k$
- Terminate if

$$\begin{aligned} & \| (w^{k+1} - q - Mz^{k+1} - \lambda^{k+1})_J \| < \epsilon_1, \| (v^{k+1} - b + z^{k+1})_J \| < \epsilon_1 \\ & \sum_{j \in J} [(z_j^{k+1} - a_j) w_j^{k+1} + \lambda_j^{k+1} v_j^{k+1}] < \epsilon_2 \end{aligned}$$

It is possible to show that this algorithm possesses global convergence, provided η_k is chosen by a special line-search procedure in each iteration k [19,22]. However, extensive computational experience with linear complementarity problems has shown that the algorithm usually works quite well in practice when η_k has a fixed value η near 1 (usually $\eta = 0.99995$) in each iteration k [6,10].

Some of the models discussed in the previous section lead into a MLCP with $b_i = +\infty$ for all i . The algorithm simplifies in this case, as the variables λ and v do not exist. The simplified algorithm follows easily from the method described before and is not discussed in this paper. We suggest [20] for a detailed explanation of this procedure. This report also contains the description of a MATLAB implementation that was linked to the Partial Differential Equations MATLAB toolbox in order to process the variational models discussed in section 2.

4.2 Block Principal Pivoting Algorithm

This method has been proposed in [11] for the solution of a MLCP with a NSM matrix. We start its description by recalling the concept of basic solution to the MLCP. Consider a subset F of $\{1, \dots, n\}$ such that M_{FF} is nonsingular. If \emptyset represents the empty set,

$$T = \{1, \dots, n\} - F, \quad T = T_1 \cup T_2 \quad \text{and} \quad T_1 \cap T_2 = \emptyset,$$

then $w = q + Mz$ is equivalent to

$$\begin{bmatrix} z_F \\ w_{T_1} \\ w_{T_2} \end{bmatrix} = \begin{bmatrix} \bar{q}_F \\ \bar{q}_{T_1} \\ \bar{q}_{T_2} \end{bmatrix} + \begin{bmatrix} \bar{M}_{FF} & \bar{M}_{FT_1} & \bar{M}_{FT_2} \\ \bar{M}_{T_1F} & \bar{M}_{T_1T_1} & \bar{M}_{T_1T_2} \\ \bar{M}_{T_2F} & \bar{M}_{T_2T_1} & \bar{M}_{T_2T_2} \end{bmatrix} \begin{bmatrix} w_F \\ z_{T_1} \\ z_{T_2} \end{bmatrix},$$

where

$$\bar{M}_{FF} = M_{FF}^{-1}, \quad \bar{M}_{TT} = M_{TT} - M_{TF} M_{FF}^{-1} M_{FT},$$

$$\bar{M}_{TF} = M_{TF} M_{FF}^{-1}, \quad \bar{M}_{FT} = -M_{FF}^{-1} M_{FT},$$

$$\bar{q}_F = -M_{FF}^{-1} q_F, \quad \bar{q}_T = q_T - M_{TF} M_{FF}^{-1} q_F.$$

A basic solution to the MLCP (17) is a solution (\bar{z}, \bar{w}) of the previous system such that

$$F = \{i \in J : \bar{w}_i = 0\}, T_1 = \{i \in J : \bar{z}_i = b_i < +\infty\}, T_2 = LU\{i \in J : \bar{z}_i = a_i > -\infty\}$$

If $T = T_1 \cup T_2$, then the variables $w_i, i \in F$ and $z_j, j \in T$, are called nonbasic, while the remaining variables are said to be basic. If we set $a_l = g_l, l \in L$, the nonbasic variables satisfy

$$\bar{z}_{T_1} = b_{T_1}, \bar{z}_{T_2} = a_{T_2}, \bar{w}_F = 0 \tag{37}$$

These values and the equalities stated before imply that the values of the basic variables can be found by

$$M_{FF}\bar{z}_F = -(q_F + M_{FT_1}\bar{z}_{T_1} + M_{FT_2}\bar{z}_{T_2}) \tag{38}$$

$$\bar{w}_T = q_T + M_{TT_1}\bar{z}_{T_1} + M_{TT_2}\bar{z}_{T_2} + M_{TF}\bar{z}_F$$

The basic solution defined by (37) and (38) is a solution of the MLCP (17) if and only if

$$a_F \leq \bar{z}_F \leq b_F, \bar{w}_{T_1} \leq 0, \bar{w}_{T_2-L} \geq 0 \tag{39}$$

The Block Principal Pivoting algorithm is a procedure that uses basic solutions of the MLCP in each iteration and terminates when such a solution satisfies the condition (39). As is discussed in [11,17], the algorithm possesses finite convergence and is strongly polynomial. Its steps are presented below.

BLOCK PRINCIPAL PIVOTING ALGORITHM

Step 0 - Set $F = T_1 = \emptyset, T_2 = \{1, \dots, n\} = JUL$

Step 1 - Compute $\bar{z} = (\bar{z}_F, \bar{z}_{T_1}, \bar{z}_{T_2})$ and \bar{w}_{T_2} by (37) and (38). If $\bar{w}_{T_2-L} \geq 0$, stop: \bar{z} is a solution of the MLCP.

Step 2 - (i) Set

$$T_2 = T_2 - I, T_1 = T_1 \cup I$$

where $I = \{i \in T_2-L : \bar{w}_i < 0\}$

(ii) Compute \bar{z}_F by (38) and set

$$T_1 = T_1 \cup H, F = F - H$$

where $H = \{i \in F : \bar{z}_i > b_i\}$

Step 3 - Compute \bar{w}_{T_1} by (38) and let $G = \{i \in T_1 : \bar{w}_i > 0\}$. If $G = \emptyset$ go to Step 1.

Otherwise set $T_1 = T_1 - G$ and $F = F \cup G$. Repeat this step.

It follows from the description of its steps that the major computational effort of this algorithm relies on the solution of linear systems with the matrix M_{FF} . As before, this system can be processed by the linear equations solver of the finite element code. Furthermore the representation of the sets F, T_1 and T_2 can be easily done by using a linked list. A MATLAB implementation of this algorithm has been developed and linked to the Partial Differential Equations toolbox. The whole procedure is fully described in [20].

5. Computational Experience

In this section we report an investigation of the efficiency of the interior-point and block principal pivoting algorithms for solving MLCPs associated with free-boundary models discussed in section 2. All the experiences have been performed in a SPARC Station 10-52, with 64 Mb RAM and 50Mhz. The results of these experiences are displayed in three tables, where the following notations have been used:

n - dimension of the MLCP

it - number of iterations of the algorithm

T - CPU time in seconds

In the first experience, we have solved two obstacle problems, where f , Ψ and Ω are given below.

$$\text{obs1} - f = 1, \Psi = \frac{1}{2} - (x_1^2 + x_2^2), \Omega = \{(x_1, x_2) : x_1^2 + x_2^2 \leq 1\}$$

$$\text{obs2} - f = 2, \Psi = 2 - (x_1^2 + x_2^2), \Omega = \{(x_1, x_2) : x_1^2 + x_2^2 \leq 4\}$$

In order to process these two models we have used four different triangulations of the domain differing on the number of elements. The results of the performances of the algorithms for four different number n of nodes are displayed in the table 1.

		Interior-Point		Block Pivoting	
Model	n	it	T	it	T
obs1	143	21	5.49	5	0.34
	537	21	26.61	10	2.62
	2081	24	198.5	18	23.77
	8193	23	1578.9	36	325.0
obs2	143	23	5.43	7	0.61
	537	20	24.09	13	4.58
	2081	20	170.2	24	44.39
	8193	21	1402.2	51	596.3

Table 1 - Obstacle Problems

It is important to note that the number of iterations of the interior-point algorithm corresponds to the number of systems of linear equations that are required for finding the Newton's directions used by the procedure. On the other hand, the number of iterations for the block principal pivoting method represents the number of modifications in the sets F , T_1 and T_2 and is usually smaller than the number of linear systems required to be solved by the algorithm. Furthermore the dimensions of the linear systems for the block principal pivoting algorithm are usually much smaller than those of the interior-point method. This explains why a superior iteration count for the former method leads to a much smaller computational time. So the block principal pivoting method is usually more efficient in terms of computer time. However, the

number of iterations for the interior-point algorithm seems to be more stable as the dimension increases. Hence this last method may be more competitive with the block algorithm for finer triangulations that lead to huge MLCPs.

In the second experience, we have investigated the accuracy of the solutions found by the two complementarity algorithms. To do this, we have generated the second obstacle problem *obs2* in such a way that the unique solution of the discretized problem (13) coincides with the unique solution of the same variational inequality on the set

$$K_h = \{v_h \in V_h : v_h(x_1, x_2) = g(x_1, x_2), (x_1, x_2) \in \partial\Omega\}$$

Since the algorithms for computing an approximate solution by assuming or omitting the obstacle are different, then the procedures should compute different solutions. In table 2, we compare the accuracy of these solutions by displaying the values of the norm $\|\bar{x} - \hat{x}\|$, where \bar{x} is the solution computed by the complementarity algorithm and \hat{x} is the solution found by using the MATLAB toolbox for solving the model without assuming the obstacle.

	Interior-Point	Block Pivoting
n	$\ \bar{x} - \hat{x}\ $	$\ \bar{x} - \hat{x}\ $
143	1.26E - 14	9.07E - 15
537	7.64E - 14	5.58E - 14
2081	2.58E - 13	2.20E - 13
8193	1.03E - 12	1.57E - 12

Table 2 - Accuracy of solutions

The results show that there is no significant difference between the solutions \bar{x} and \hat{x} stated before. Hence, it seems that both the complementarity algorithms are able to compute solutions as accurate as the ones found by the finite-element code for free-boundary problems without constraints.

The two experiences described before have indicated that the complementarity algorithms are quite useful tools, as they are able to find sufficient accurate solutions in a relatively reduced amount of iterations and time. As is discussed in [20], experiences with the two remaining models described in section 2 have led to the same conclusions. In table 3 we illustrate the performances of the complementarity algorithms on the solution of MLCPs associated with a model of flow through a porous dam and an elastic-plastic torsion model. As stated in section 2, we simply need to indicate the definition of the function g for a full description of the first model. This function is given by

$$g(x_1, x_2) = \begin{cases} \frac{(2-x_2)^2}{2} & \text{in AB} \\ \frac{(0.5-x_2)^2}{2} & \text{in CD} \\ \frac{4(2-x_1)+0.25x_1}{4} & \text{in BC} \\ 0 & \text{in DFA} \end{cases}$$

where the points A,B,C,D and F are represented in Figure 1 of section 2. On the other hand, the elastic-plastic torsion problem is characterized by the domain

$$\Omega = \{(x_1, x_2) : x_1^2 + x_2^2 \leq 1\}$$

and by the functions $f(x_1, x_2) = 2$ and $g(x_1, x_2) = 0$. The performances of the complementarity algorithms on these two models are displayed in table 3. The results confirm the appropriateness of these algorithms to deal with these variational problems.

		Interior-Point		Block Pivoting	
Model	n	it	T	it	T
Flow through dam	198	25	8.45	11	0.91
	747	27	53.03	24	8.61
	2901	28	337.4	45	95.61
	11433	32	3203.3	87	1290.2
Elastic Plastic Torsion	143	27	10.19	5	0.46
	537	23	42.80	9	3.17
	2081	20	217.7	16	28.59
	8193	19	1587.8	31	376.5

Table 3 - Flow through a porous dam and elastic plastic torsion models

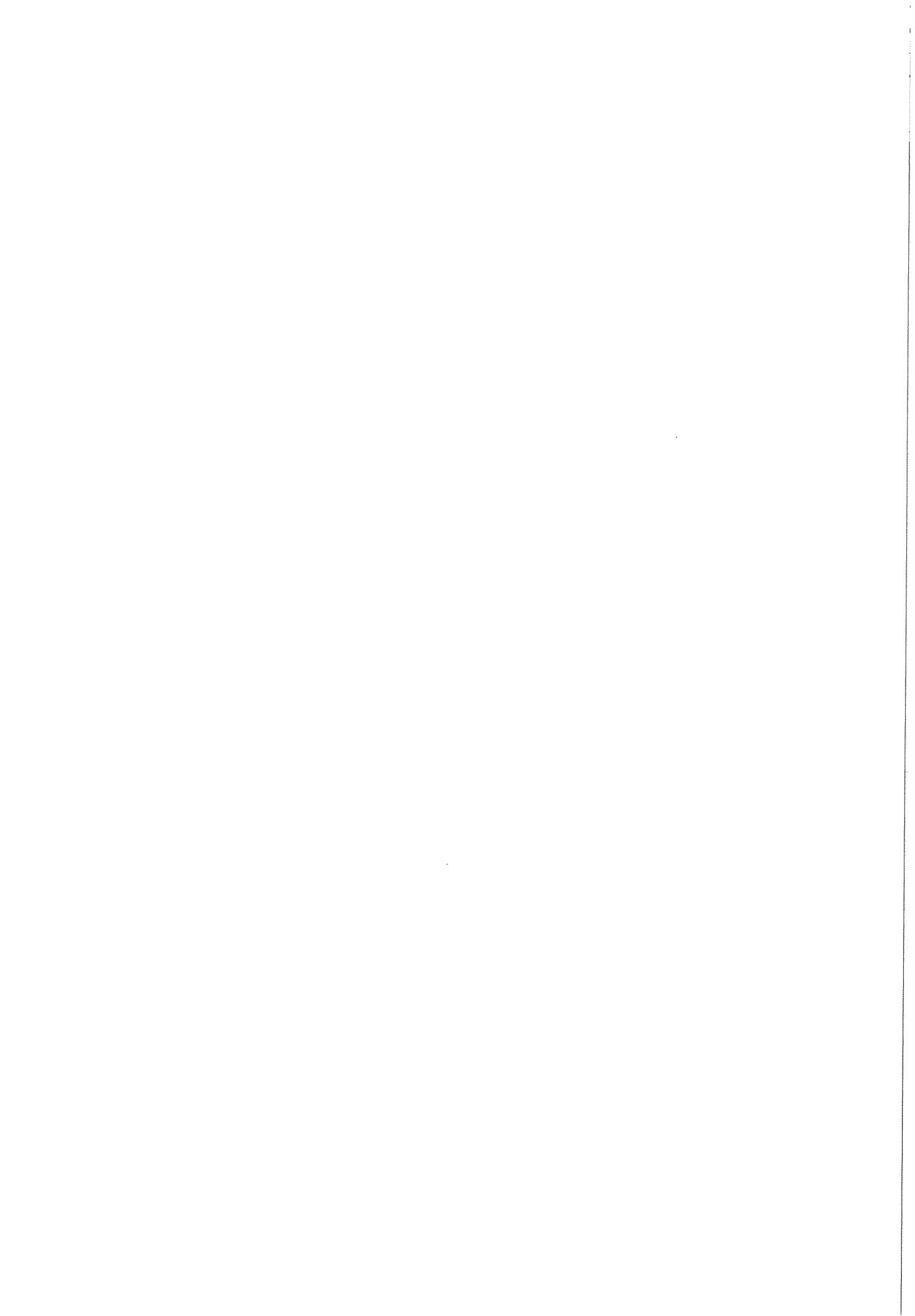
6. Conclusions

In this paper we have investigated the solution of linear complementarity problems associated with free-boundary models of mechanics. An interior-point and a block principal pivoting algorithms have been adapted to deal with these complementarity problems and have been implemented within a MATLAB finite element code. The resulting code has been used for solving special instances of obstacle, flow through a porous dam and elastic plastic torsion models. The numerical results have shown the appropriateness of these complementarity algorithms to deal with this type of models.

As is discussed in [7], many linear and nonlinear complementarity problems have been associated with variational models of mechanics. Some of these models are still to be efficiently processed. We believe that the use of techniques similar to those discussed in this paper, either alone or incorporated in sequential algorithms can help in this extent. This is certainly one of our major research topics.

References

- [1] Baiocchi, C., *Sur un problème à frontière libre traduisant le filtrage de liquides à travers des milieux poreux*, C. R. Acad. Sc. Paris 273 (1971) 1215-1217.
- [2] Chipot, M., *Variational Inequalities and Flow in Porous Media*, Springer-Verlag, New York (1984).
- [3] Chipot, M., *Some results about an elastic-plastic torsion problem*, Nonlinear Anal. Theory Methods Appl. 3 (1979) 261-269.
- [4] Ciarlet, P., *The Finite Element Method for Elliptic Problems*, North-Holland Publishing Company, New York (1978).
- [5] Cottle, R., Pang, J. and Stone, R., *The Linear Complementarity Problem*, Academic Press, New York (1992).
- [6] Fernandes, L., Júdice, J. and Patrício, J., *An investigation of interior point and block pivoting algorithms for large scale symmetric monotone linear complementarity problems*, Comp. Optim. and Applications 5 (1996) 49-77.
- [7] Ferris, M. and Pang, J.S., *Engineering and economic applications of complementarity problems*, SIAM Review 39 (1997) 669-713.
- [8] Glowinski, R., *Numerical Methods for Nonlinear Variational Problems*, Springer-Verlag, New York (1984).
- [9] Glowinski, R., Lions, J. and Trémolières, R., *Analyse Numérique des Inéquations Variationnelles*, tome 1 and 2, Bórdas, Paris (1976).
- [10] Júdice, J., Fernandes, L. and Lima, A., *On the solution of the monotone and nonmonotone linear complementarity problem*, in Nonlinear Optimization and Applications, edited by G. Di Pillo, F. Giannessi, Plenum Press, New York (1996) 213-227.
- [11] Júdice, J. and Pires, F., *A polynomial-type method for a generalized linear complementarity problem with a nonsingular M-matrix*, IMA Journal of Mathematics Applied in Business and Industry 4 (1992) 211-224.
- [12] Kojima, M., Megiddo, N., Noma, T. and Mizuno, S., *A Unified Approach to Interior-Point Algorithms for Linear Complementarity Problems*, Lecture Notes in Computer Science 538 (1991), Springer-Verlag, Berlin.
- [13] Lin, Y. and Cryer, C., *An alternating direction implicit algorithms for the solution of linear complementarity problems arising from free boundary problems*, Appl. Math. Optim. 13 (1985) 1-17.
- [14] Murty, K., *Linear Complementarity, Linear and Nonlinear Programming*, Heldermann Verlag, Berlin (1988).
- [15] Oden, J. and Kikuchi, N., *Theory of variational inequalities with applications to problems of flow through porous media*, Int. J. Eng. Sci. 18 (1980) 1173-1284.
- [16] Oden, J. and Reddy, J., *An Introduction to the Mathematical Theory of Finite Elements*, John Wiley and Sons, Inc., New York (1976).
- [17] Pang, J.S., *On a class of least-element complementarity problems*, Mathematical Programming 16 (1979) 111-126.
- [18] Rodrigues, J.F., *The Obstacle Problem in Mathematical Physics*, North-Holland Publishing Company, New York (1987).
- [19] Simantiraki, E. and Shanno, D., *An infeasible-interior-point algorithm for solving mixed complementarity problems*, in Complementarity and Variational Problems, edited by M. Ferris and J.S. Pang, SIAM (1997) 386-404, Philadelphia.
- [20] Soares, M.C., *Linear Complementarity in Variational Models of Mechanics*, Masters Thesis, Department of Mathematics, University of Coimbra (1996) (in portuguese).
- [21] Westbrook, D., *A mixed variational inequality boundary iteration method for some free boundary problems*, SIAM J. Sci. Stat. Comput. 5 (1984) 192-202.
- [22] Wright, S., *Primal-Dual Interior-Point Methods*, SIAM (1997) Philadelphia.



UM MODELO CAUSAL DE PREVISÃO DE VENDAS PARA UM PRODUTO DE GRANDE CONSUMO

Júlio Filipe

Johnson & Johnson, Lda.
Est. Consiglieri Pedroso, 69A
Queluz de Baixo
2745 Barcarena - Portugal

Isabel Hall Themido

IST/CESUR
Universidade Técnica de Lisboa
Av. Rovisco Pais
1000 Lisboa - Portugal

Abstract

This paper presents a forecasting model for bimester trade sales of Johnson & Johnson, Lda. (J&J) product in Portugal. The trade sales of this market leader show great variability, as a result of seasonality, promotional activity and retailer's stock policies. It was necessary to understand the logistic chain, from when the product leaves J&J until it reaches final consumer, and to model final consumer sales before trade sales could be modelled. A causal model was adopted, including three groups of explanatory variables - marketing, seasonality and trend, and a variable that represents the product's trade stock level. This variable has a different effect in the two models; if trade stock is high, sales of the final consumer increase reflecting retail promotional activities, but trade sales decrease as retailers have full store rooms.

Models to forecast sales by size, as a function of total product sales, are also presented. A discussion of the effect of promotional sales on the total product sales is also included.

Resumo

Neste trabalho, apresenta-se um modelo de previsão das vendas bimestrais ao retalho de um produto da Johnson & Johnson, Lda. (J&J). As vendas desse produto, *leader* no mercado dos produtos de higiene pessoal, apresentam grande variabilidade, resultante de sazonalidade, de acções promocionais e, principalmente, dos níveis de stock do produto existente no mercado. A construção do modelo passa pela compreensão de todo o processo de vendas, desde que o produto sai da J&J até que chega ao consumidor. Optou-se pela utilização de modelos causais, devido à existência de factores controláveis que influenciam o comportamento das vendas e à facilidade de os estender a outros produtos. Foram considerados três tipos de variáveis explicativas - variáveis de marketing, variáveis representativas da sazonalidade e tendência, e uma variável que representa o stock de produto no mercado e que traduz algumas acções promocionais dos retalhistas. Esta modelação, para além de permitir estimar as vendas totais do produto aos retalhistas e ao consumidor final, também fornece a previsão das vendas ao retalho por tamanho. Como o modelo é do tipo causal é ainda possível discutir o efeito das vendas promocionais nas vendas totais do produto.

Keywords

Forecasting, sales modelling, regression, marketing.

1. Introdução

A J&J (Consumo) é uma multinacional, que se dedica ao marketing e venda de produtos de grande consumo na área dos cuidados do bebé e de adulto e higiene feminina. Estas três áreas de negócio agrupam produtos comercializados em diversos tamanhos, ultrapassando 200 referências. Em Portugal o número de clientes directos aproxima-se de 800.

A temática dos modelos de previsão reveste-se de grande importância prática, uma vez que as empresas necessitam de planear a sua produção, ou caso não produzam, como acontece com a J&J Portugal, necessitam de calendarizar as suas compras. Nesta última situação, é ainda mais importante antever as vendas pois os tempos de entrega são bastante elevados, cerca de três meses para a maioria dos produtos.

A modelação de vendas traz assim vantagens evidentes como referem Lilien et al, 1992, nomeadamente:

- a redução dos custos de inventário;
- a redução das rupturas de stock;
- o conhecimento do impacto das variáveis de marketing sobre as vendas.

Este trabalho teve como objectivo desenvolver um modelo, para prever, no curto e médio prazo, as vendas globais ao nível do produto, bem como as vendas por tamanho comercializado. Para além disso, estabeleceram-se relações de causalidade entre diversas variáveis de marketing e as vendas.

O estudo incidiu sobre um produto *leader* no mercado, de ora em diante designado por *PRODUTO X*, estabelecido no mercado, desde há vários anos. Este produto tem 4 apresentações diferentes, designadas por *tamanho1*, ..., *tamanho4*. O *PRODUTO X* refere Assim, o conjunto destas quatro referências.

A análise que vamos efectuar será bimestral, na medida em que não dispomos de informação mais detalhada sobre algumas variáveis que se revelaram determinantes.

Antes de iniciar a modelação, é necessário localizar o problema no âmbito da cadeia logística, e compreender o processo de vendas num contexto integrado. Olhemos então para o esquema simplificado apresentado na Figura 1.

O fluxo do produto pode ser decomposto em duas fases: as vendas da J&J aos seus clientes directos (grossistas e retalhistas) e as vendas destes ao consumidor final (que são as vendas dos retalhistas). Embora o fluxo do produto seja o mesmo em ambas as fases, existe um desfasamento e um comportamento diferenciado das duas séries de vendas. Como podemos verificar na Figura 2, a série correspondente às compras do consumidor (vendas Nielsen) apresenta sazonalidade marcada e pequena variabilidade, ao contrário das vendas da J&J onde a sazonalidade é menos marcada, mas a variabilidade é muito maior.

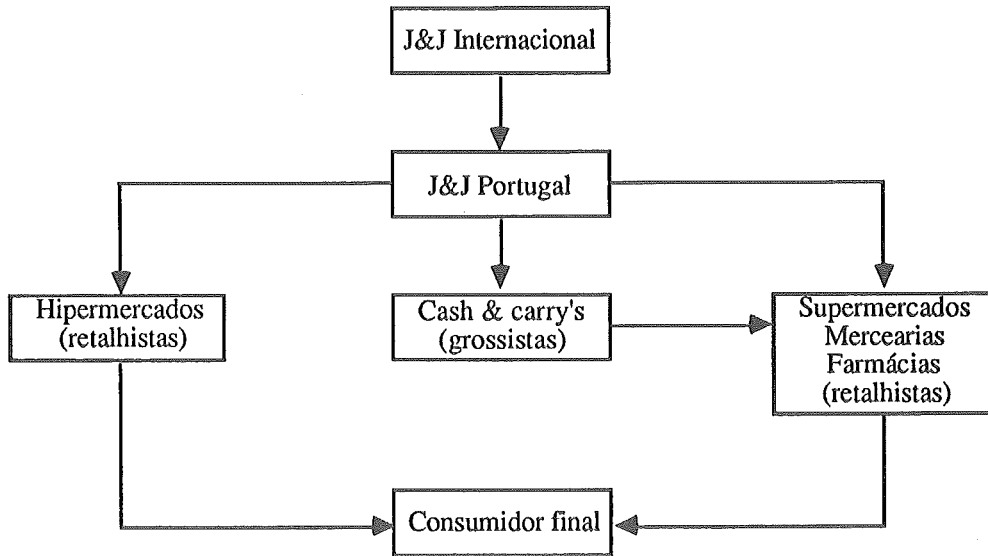


Figura 1 - Fluxos principais dos produtos comercializados

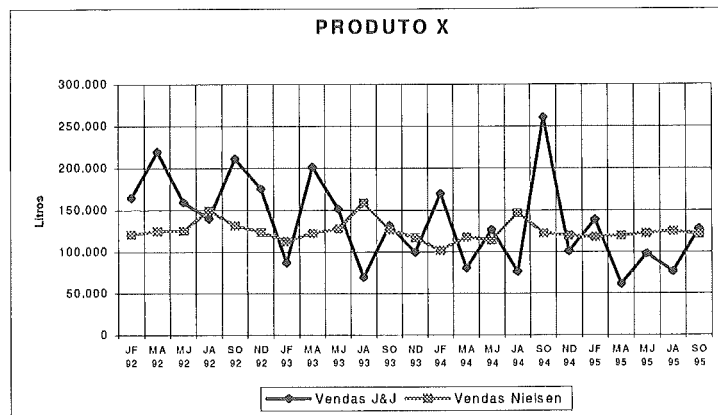


Figura 2 - Comparação entre vendas ao retalho e vendas ao consumidor

A diferença entre as duas séries resulta da sua localização na cadeia logística. O consumidor tem um comportamento regular comprando quando necessita, pelo que as suas compras apresentam menor variabilidade. A sazonalidade patente resulta de hábitos de consumo, com o Verão a suscitar maior procura. Já os retalhistas não compram regularmente, criando stocks e reagindo mais agressivamente a alterações nas variáveis de marketing. Quando estas lhes são favoráveis, configurando uma boa oportunidade de comprar a melhores preços, aumentam os seus stocks, por vezes o suficiente para só voltar a comprar em alturas favoráveis. Estamos assim perante um factor altamente perturbador das vendas da J&J (Blattberg e Levin, 1987).

As vendas da J&J são conhecidas rigorosamente, ao passo que as compras do consumidor final só podem ser estimadas através de dados recolhidos pela Nielsen, uma empresa multinacional que presta serviços na área do consumo, fornecendo informação sobre o comportamento dos produtos e marcas no ponto de venda.

Infelizmente a J&J só tem ao seu dispor dados bimestrais das vendas ao consumidor, representando estes dados, em média, cerca de 85% do mercado Nacional. Uma vez que a Nielsen não cobre todos os pontos de vendas do país, teremos que proceder à correcção desses

dados. Para isso, é necessário comparar as compras da Nielsen (que traduzem os resultados das compras dos retalhistas à J&J, recolhidas pela Nielsen nos pontos de venda) com as vendas da J&J aos retalhistas. Como as compras da Nielsen são inferiores às vendas da J&J, calcula-se a taxa de cobertura da Nielsen, fazendo o quociente entre os valores anuais daquelas duas variáveis, que representa a percentagem do produto vendido pela J&J que é detectado pela Nielsen. Para obter as vendas totais ao consumidor, basta então dividir as vendas da Nielsen por aquele factor. Como a Nielsen não recolhe informação junto dos grossistas, as vendas da J&J aquele sector só aparecem como compras Nielsen na leitura dos pequenos retalhistas, que se abastecem nos grossistas. Assim, parte da informação da Nielsen está desfasada mais do que um período. Este tipo de situações leva a que, por vezes, seja necessário o conselho da própria Nielsen sobre os valores das taxas de cobertura a adoptar.

2. Construção do modelo

Como já vimos, as vendas do *PRODUTO X* apresentam uma grande variabilidade em consequência da influência das variáveis de marketing e da localização da J&J na cadeia logística.

Os modelos causais representam nestas condições uma abordagem conveniente, por se adaptarem bem à variabilidade da série e à influência das variáveis controláveis. Adicionalmente, constituem uma abordagem simples o que é importante face à quantidade de produtos existentes.

No decorrer deste trabalho utilizaremos variáveis que representam as vendas da J&J aos retalhistas e grossistas, bem como as vendas de produtos da J&J ao consumidor final. Essas variáveis são:

- **VEND_SBP** que representa as vendas da J&J, não incluindo nem o produto oferecido em bónus de quantidade (ver variáveis de marketing apresentadas seguidamente) nem o produto promocionado.
- **VEND_NCC** que representa as compras do consumidor final, obtidas corrigindo os valores recolhidos pela Nielsen com a taxa de cobertura.

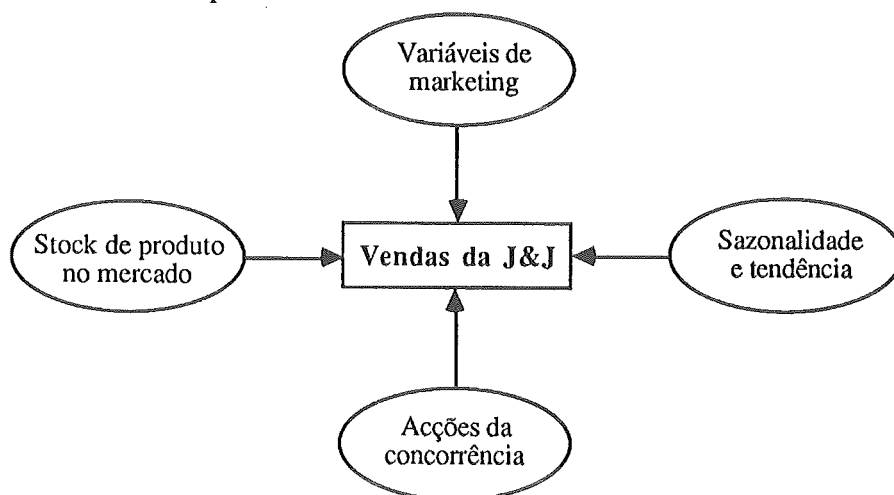


Figura 3 - Variáveis que influenciam as vendas

As variáveis potencialmente explicativas das vendas podem ser organizadas em quatro grupos da forma apresentada na Figura 3. Vejamos agora, dentro de cada um dos quatro grupos referidos, quais as variáveis que foram usadas neste trabalho.

a) **Variáveis de Marketing** - O conjunto das variáveis de marketing é numeroso, justificando que se faça uma escolha prévia, identificando as mais importantes e considerando apenas aquelas que será possível quantificar quer no presente quer no futuro.

- **BONUS(t)** representa a percentagem de bónus, em produto, dada aos retalhistas e grossistas. Por exemplo, se a percentagem de bónus for de 50%, então diz-se que a J&J dá 3=2, ou seja, na compra de 3 embalagens o retalhista paga apenas 2. Deste modo, espera-se que esta variável influencie as vendas directas da J&J, assim como as vendas ao consumidor, uma vez que, se os retalhistas compram em melhores condições de preço, poderão fazer repercutir esse desconto nos seus preços de venda.
- **BONUS(t-1)** representa a percentagem de bónus, em produto, desfasada de um bimestre.
- **BONUS(t-2)** representa a percentagem de bónus, em produto, desfasada de dois bimestres.
- **PROM(t)** representa a quantidade total de produto vendido em promoção, no bimestre t, aos retalhistas e grossistas. Embora esta promoção seja dirigida ao consumidor final, é de esperar que possa afectar os clientes directos da J&J. Estas promoções são vulgarmente designadas por "Pai e filho", porque em conjunto com o produto normal, oferece-se um produto mais pequeno ou um produto de outro tipo. Incluímos ainda nesta variável os produtos que trazem um brinde. São promoções do tipo:

ProdutoX + produtoX(pequeno), ProdutoX + ProdutoY ou ProdutoX + brinde.

Como o *PRODUTO X* é comercializado em quatro tamanhos, e somente os dois tamanhos maiores são promocionados como "Pai e filho", consideram-se as seguintes variáveis:

- **PROM3(t)** que representa a quantidade do *PRODUTO X* e tamanho 3, vendida com promoção no bimestre t.
- **PROM4(t)** que representa a quantidade do *PRODUTO X* e tamanho 4 vendida com promoção no bimestre t.
- **PREÇO(t)** que representa o preço unitário médio deflacionado de venda aos retalhistas e grossistas. O processo de deflação dos preços baseou-se nos índices de preços anuais para a classe de artigos "cuidados pessoais não duráveis", fornecidos pelo Instituto Nacional de Estatística, determinando-se o valor da inflação em cada ano e calculando seguidamente os preços por unidade, a preços de 1995.
- **TD(t)** que representa a percentagem média de desconto no preço de venda aos retalhistas e grossistas (Trade Deal) no bimestre t.
- **PUB(t), PUB(t-1), PUB(t-2)** que representam a percentagem de despesa da J&J em publicidade, em relação à despesa publicitária total do mercado em que o *PRODUTO*

X está inserido. $PUB(t)$ representa a publicidade no próprio bimestre, $PUB(t-1)$, a publicidade desfasada de um bimestre e $PUB(t-2)$, a publicidade desfasada de dois bimestres.

- $TEL(t)$, $TEL(t-1)$ e $TEL(t-2)$ que representam a despesa, em valor absoluto, da J&J em publicidade nos bimestres t , $t-1$ e $t-2$ respectivamente.

b) Variável que representa o stock de produto no mercado

Revelou-se de extrema importância quantificar o stock de *PRODUTO X* existente no mercado:

- $STOCK(t-1)$, representa o stock no mercado, ou seja, a diferença entre a quantidade total de *PRODUTO X* que sai da J&J e a quantidade de *PRODUTO X* que é comprada pelo consumidor final, acrescida da quantidade de stock já existente anteriormente:

$$STOCK(t-1) = (VEND_SBP(t) + BÓNUS(t) + PROM(t)) - VEND_NCC(t) + STOCK(t-2) \quad (1)$$

O stock de *PRODUTO X* existente no mercado vai ter impactos distintos na cadeia logística a montante e a jusante. Quanto à sua influência nas vendas directas da J&J, o nível de stock influenciará negativamente as vendas, pois quanto maior for o stock nos retalhistas e grossistas tanto menor será a propensão destes para comprar produto no bimestre seguinte, uma vez que tal representaria um acréscimo no custo de stock. É fácil de confirmar na Figura 4 a existência de uma correlação directa negativa entre estas duas variáveis.

Passa-se o contrário no que diz respeito à influência desta variável nas vendas ao consumidor final. Quando existe um nível elevado de stocks no mercado, os Retalhistas e Grossistas implementam acções de promoção de vendas, com o objectivo de reduzirem os seus stocks. Trata-se de acções dirigidas ao consumidor final, no caso dos retalhistas e aos retalhistas, no caso dos grossistas. Deste modo, a influência do nível de stock nos retalhistas e grossistas nas vendas ao consumidor final será positiva. No entanto, esta relação só se manifesta significativa quando se considera o conjunto de influências, não havendo portanto correlação directa.

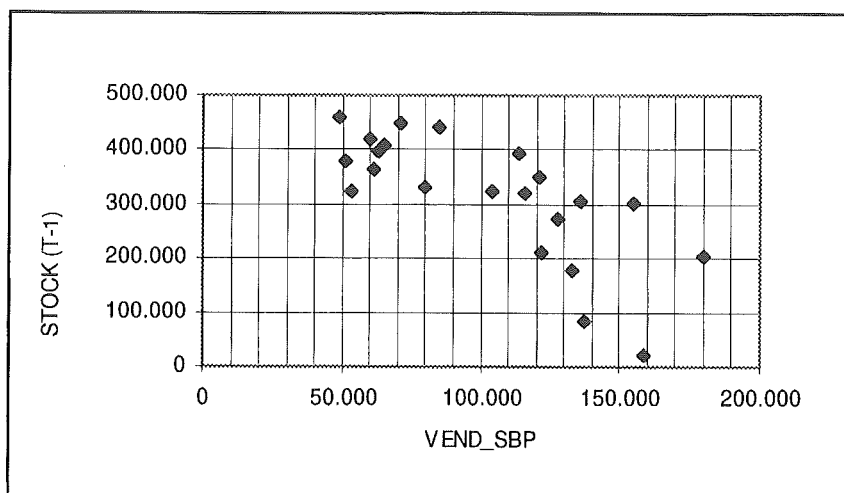


Figura 4 - Relação existente entre Stock(t-1) e vendas da J&J

c) Variáveis que reflectem a tendência e a sazonalidade

- **PERIODO(t)**, variável que permite o ajuste de uma tendência, caso se verifique a sua existência. Foi atribuído o valor 1 ao primeiro período de vendas disponível, o bimestre JF/92, o valor 2 ao 2º período MA/92, o valor 3 ao período seguinte e assim sucessivamente.
- **JF, MA, MJ, JA, SO, ND**, variáveis binárias que representam os bimestres Janeiro/Fevereiro, Março/Abril, Maio/Junho, Julho/Agosto, Setembro/Outubro e Novembro/Dezembro, assumindo o valor 1 quando estamos no bimestre correspondente e 0 no caso contrário. Teremos que excluir um dos bimestres para ultrapassar o problema da multicolinearidade (Murteira et al, 1993). Assim, o número de variáveis a utilizar foi $(n-1)$, sendo n o número de bimestres do ano, uma vez que a sazonalidade é anual ($n=6$). Os coeficientes destas variáveis representarão a sazonalidade do produto.

d) Acções da concorrência

Este é um conjunto de variáveis que influenciam potencialmente as vendas da J&J, mas que não foi possível quantificar directamente. Contudo, ao utilizar as variáveis $PUB(t)$, e a variável $STOCK(t-1)$, estamos a considerar indirectamente algumas acções importantes da concorrência e dos retalhistas.

2.1 Modelação

2.1.1 Modelos explicativos

Considerámos dois tipos de modelos explicativos, o aditivo e o multiplicativo.

MODELO ADITIVO:

$$V(t) = C + \sum_i \alpha_i X_i + \sum_j \beta_j D_j + \varepsilon \quad (2)$$

onde: $V(t)$ = volume de vendas no bimestre t

C = constante

X_i = variável explicativa i

D_j = variável muda que poderá representar a sazonalidade

ε = erro (média = 0)

MODELO MULTIPLICATIVO:

$$V(t) = C \prod_i X_i^{a_i} \varepsilon \quad (3)$$

onde: $V(t)$ = volume de vendas no bimestre t

C = constante

X_i = variável explicativa i

ε = erro associado a cada variável (média = 1)

Neste caso é necessário retirar previamente a sazonalidade da série, recorrendo ao método da decomposição clássica (Makridakis et al, 1983).

Logaritimizando $V(t)$, obtemos (4), que será tratado como um modelo aditivo:

$$\text{Ln } V(t) = \text{Ln } C + \sum_i a_i \text{Ln } X_i + \text{Ln } (\epsilon) \quad (4)$$

Outra forma de considerar a sazonalidade é proceder de forma semelhante ao modelo aditivo, utilizando variáveis mudas, ficando o modelo multiplicativo com a forma (5):

$$V(t) = C e^{\sum_j \alpha_j D_j} \prod_i X_i^{\beta_i} \epsilon \quad (5)$$

onde D_j são as variáveis mudas já descritas anteriormente.

Tal como fizemos em (4), podemos transformar (5) num modelo aditivo (6), o que facilita a estimação dos parâmetros:

$$\text{Ln } V(t) = \text{Ln } C + \sum_i \alpha_i D_i + \sum_j \beta_j \text{Ln}(X_j) + \text{Ln } (\epsilon) \quad (6)$$

O nosso objectivo último é modelar as vendas directas da J&J. Contudo, uma vez que a variável $\text{STOCK}(t-1)$ se revelou fundamental para o sucesso da modelação e que de futuro a teremos que estimar também para efectuar previsões, desenvolvemos paralelamente um modelo de previsão das vendas ao consumidor final. A variável $\text{STOCK}(t-1)$ será quantificada depois de estimadas as vendas de toda a cadeia logística, a jusante da J&J Portugal.

O modelo que melhor se adaptou às duas séries de vendas foi do tipo aditivo, pelo que somente o resultante deste é apresentado.

2.1.2 Selecção das variáveis em cada modelo

Após um estudo prévio das variáveis que poderão influenciar as vendas, já anteriormente referidas, houve que seleccionar, de entre elas, um conjunto mais pequeno (*Makridakis et al*, 1983). Neste processo, tivemos em atenção o nível de significância da estimativa de cada coeficiente, considerando apenas variáveis para as quais este valor é inferior a 5%, embora, excepcionalmente, se tenham considerado estimativas com níveis de significância que atingem 15% desde que a sua introdução não causasse instabilidade no modelo. Esta prática é frequente na área de Marketing para dar visibilidade a variáveis que são conceptualmente importantes, embora estatisticamente pouco relevantes face aos limitados dados disponíveis. Outro aspecto cuidadosamente tomado em consideração foi a validação do sinal de cada coeficiente, através do conhecimento adquirido no estudo das variáveis candidatas ao modelo e da ajuda dos gestores da empresa.

2.1.3 Modelo de vendas ao consumidor final

A obtenção deste modelo revelou-se bastante problemática face à dicotomia existente entre a concepção teórica do modelo e a sua resolução matemática, que rejeitava variáveis que nos pareciam fazer todo o sentido. De qualquer forma foi possível incluir duas variáveis de marketing, $\text{PROM}(t)$ e $\text{BONUS}(t-2)$, e concluir que o efeito das promoções ao consumidor, nas vendas ao consumidor final é mais rápido do que o efeito dos bónus. As restantes variáveis de marketing não se mostraram significativas. Podemos observar no Quadro 1 os coeficientes do modelo.

R² Ajustado	90,2%
Erro absoluto médio	2,5%

variável	coeficiente	SE(coeficiente)	t	Sig t
Constante	154929,99	6102,72	25,39	0,00%
³ √(t)	-35337,73	5360,49	-6,59	0,00%
JF	-6437,26	3000,09	-2,15	5,31%
JA	26150,28	2702,08	9,68	0,00%
SO	15310,08	3334,42	4,59	0,06%
STOCK(t-1)	0,10	0,03	3,69	0,31%
PROM(t)	0,18	0,08	2,19	4,90%
BONUS(t-2)	238,26	81,49	2,92	1,27%

Quadro 1 - Valores estimados para o modelo de vendas ao consumidor final

Apesar do coeficiente da variável JF ter um nível de significância acima de 5%, a sua inclusão acrescentou 3% ao R² ajustado o que justifica a sua inclusão. Os coeficientes das variáveis mudas que representam os bimestres Março/Abril e Junho/Julho não são apresentados por não serem significativamente diferentes de zero. O valor médio das vendas nesses bimestres é idêntico ao de Novembro/Dezembro. As vendas são abaixo da média em Janeiro/Fevereiro e acima desse valor em Julho/Agosto e Setembro/Outubro, períodos em que o consumo deste produto de higiene pessoal sobe.

Na Figura 5 pode verificar-se um bom ajuste do modelo às vendas ao consumidor final. A exiguidade dos dados disponíveis não permite testar o modelo com valores de vendas não utilizados para o ajuste, como seria desejável.

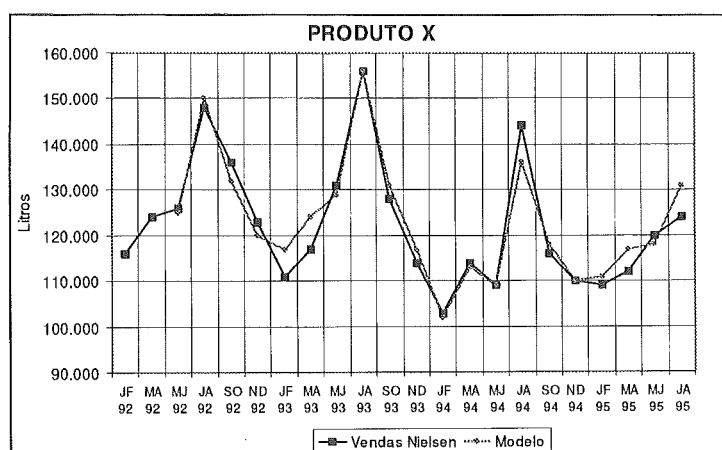


Figura 5 - Qualidade do ajuste do modelo de vendas ao consumidor final

2.1.4 Modelo de vendas da J&J ao retalho

A série de vendas da J&J ao retalho apresenta um pico no bimestre Setembro/Outubro que resulta, provavelmente, da forma como os retalhistas reagem à procura. Depois do pico de

vendas no bimestre Julho/Agosto, repõem os seus stocks no bimestre seguinte. Também os Cash & Carry, que estão, na sua maioria, encerrados em Agosto, repõem os seus stocks em Setembro. Por último, o mês de Setembro é, tradicionalmente, um mês com forte actividade promocional. É de realçar que, apesar do bimestre Setembro/Outubro ser tradicionalmente o bimestre mais alto de vendas para a J&J, no futuro este efeito pode deixar de ser tão evidente, bastando para tal que os retalhsitas antecipem as compras ao pico do consumo, em vez de responderem com atraso.

O modelo desenvolvido tem um R^2 ajustado de 86,8%. As variáveis utilizadas, bem como os seus coeficientes e níveis de significância, são apresentados no Quadro 2.

A variável STOCK(t-1) tem coeficiente negativo porque, como já foi dito, quanto maior é o stock no mercado menores serão as necessidades dos retalhistas comprarem no período seguinte.

O coeficiente da variável PROM(t) também é negativo pois quando fazemos uma promoção, as vendas do produto sem promoção (variável independente) sofrem uma redução.

A variável BONUS(t) tem, como seria de esperar, coeficiente positivo porquanto representa um incentivo para os retalhistas comprarem o PRODUTO X. No entanto verificamos que o efeito dos bónus nas vendas ao retalho é muito superior ao seu efeito nas vendas ao consumidor (tendo como coeficiente 983.93 e 238.26 respectivamente), contribuindo para o aumento do stock de produto do mercado.

Atendendo aos coeficientes da variável STOCK(t-1), no modelo das vendas ao retalho (-0.32) e no modelo das vendas ao consumidor (0.10), uma análise do efeito dos bónus num período de tempo mais alargado, mostra que afinal o seu efeito sobre as vendas não é tão positivo como uma análise apenas circunscrita ao bimestre em que o bónus foi concedido sugere.

A tendência não se mostrou significativa.

R^2 Ajustado	86,8%
Erro absoluto médio	10,0%

variável	coeficiente	SE(coeficiente)	t	Sig t
Constante	185211,96	11097,12	16,69	0,000
MJ	37013,24	8926,16	4,15	0,001
SO	45677,89	8336,22	5,48	0,000
STOCK(t-1)	-0,32	0,03	-10,56	0,000
PROM(t)	-0,54	0,20	-2,76	0,014
BONUS(t)	983,93	216,86	4,54	0,000

Quadro 2 - Valores estimados para o modelo de vendas da J&J ao retalho

Conseguiu-se melhorar este modelo decompondo a variável PROM(t) em duas variáveis: uma relativa à promoção do *tamanho3* e outra à promoção do *tamanho4*, respectivamente

PROM3(t) e PROM4(t). Podemos ver os resultados no Quadro 3 e a qualidade do ajuste na Figura 6.

R ² Ajustado	86,9%
Erro absoluto médio	9,1%

variável	coeficiente	SE(coeficiente)	t	Sig t
Constante	190497,51	10608,31	17,96	0,000
MJ	33033,22	8133,94	4,06	0,001
SO	39631,53	7765,64	5,10	0,000
STOCK(t-1)	-0,33	0,03	-11,64	0,000
PROM3(t)	-0,85	0,33	-2,60	0,019
PROM4(t)	-0,52	0,19	-2,74	0,015
BONUS(t-2)	998,83	212,75	4,70	0,000

Quadro 3 - Valores estimados para o modelo 2 de vendas da J&J ao retalho

Pensamos que esta última abordagem, para além de conduzir a um melhor ajuste, permite uma comparação interessante entre o impacto das duas promoções. O modelo sugere que 52% da quantidade vendida em promoção do *tamanho4* resulta de transferência de vendas do produto normal para o produto promocionado (ou seja, se vendermos 100 unidades do *tamanho4* em promoção, deixamos de vender 52 unidades do produto normal, havendo portanto um ganho imediato de 48 unidades). A promoção do *tamanho3* é menos positiva que a promoção do *tamanho4*, pois uma maior percentagem das vendas da promoção (85%) corresponde a transferência directa de vendas do produto normal para a promoção.

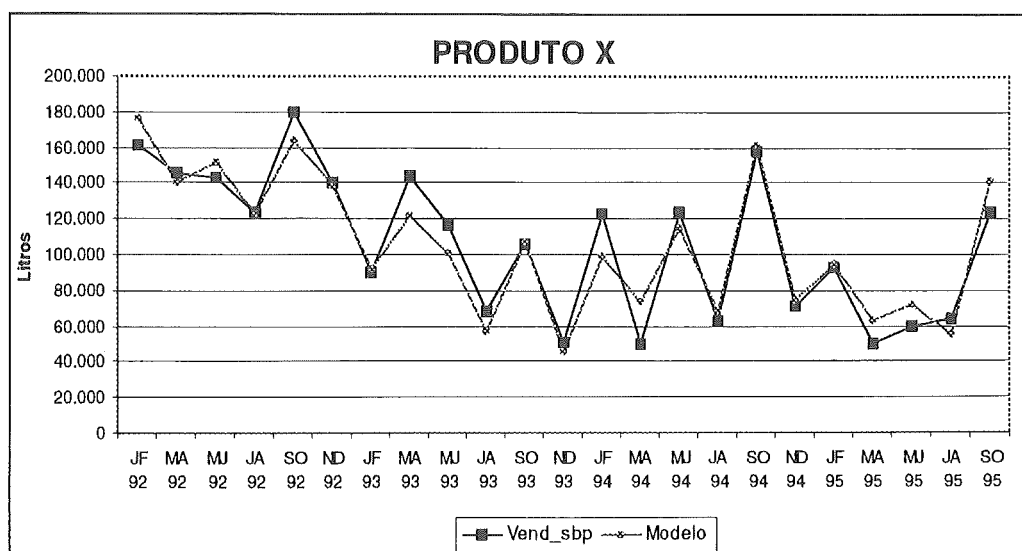


Figura 6 - Qualidade do ajuste do modelo 2 das vendas da J&J ao retalho

Na figura 7, podemos comparar as vendas, em volume, sem bônus e sem produto promocionado, com o erro cometido com este modelo. Podemos concluir que o gráfico correspondente ao erro é estacionário na média (que é zero) e estacionário na variância, ao contrário da curva correspondente às vendas da J&J.

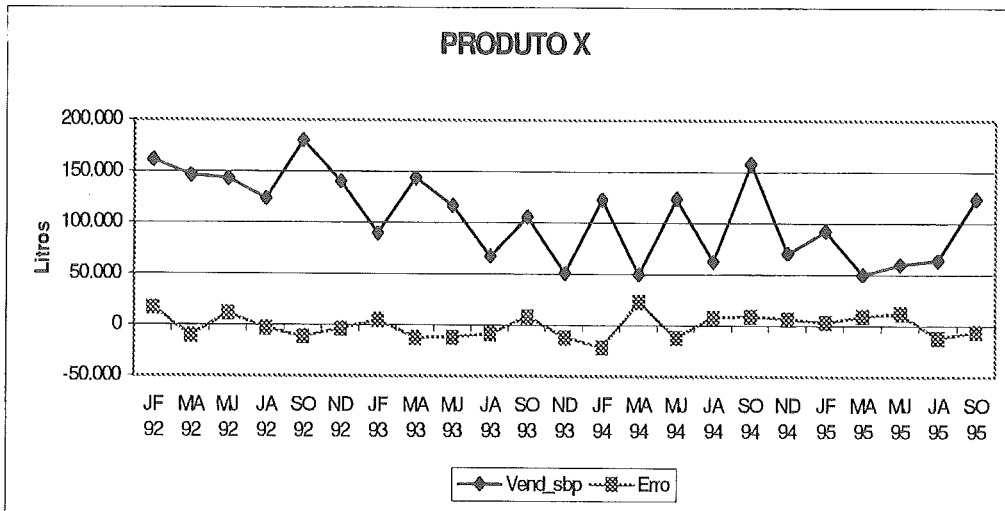


Figura 7 - Comparação entre as vendas da J&J ao retalho e o erro cometido com o modelo 2

3. Modelação das vendas do *PRODUTO X* por tamanho

O *PRODUTO X* apresenta-se, como referido no início deste trabalho, em quatro tamanhos (*tamanho1*, ..., *tamanho4*), repartindo-se as vendas totais do *PRODUTO X*, em litros, de forma equilibrada pelos vários tamanhos.

Como o *PRODUTO X* é importado já embalado, para que a modelação de vendas tenha efeitos práticos, em termos de diminuição de stocks e melhoria do nível de serviço, é necessário prever as vendas de cada um dos tamanhos. Ora as tentativas de modelar directamente as vendas ao retalho de cada um dos tamanhos mostraram-se totalmente infrutíferas, o que não é de estranhar, já que o mesmo se verificou inicialmente com as vendas totais até se considerar o stock no retalho como variável explicativa. Não sendo possível implementar para cada um dos tamanhos um modelo semelhante ao desenvolvido para as vendas totais do *PRODUTO X* - por não existirem vendas ao consumidor final estimadas pela Nielsen, desagregadas ao nível do tamanho - houve que tirar partido do modelo das vendas totais para estimar as vendas desagregadas por tamanho.

Comparando o comportamento das vendas de cada embalagem com o total (sem considerar os bônus e as promoções), pôde verificar-se que o *tamanho1* e *otamanho2*, mais pequenos e vendidos especialmente no retalho tradicional, têm um comportamento muito parecido, bastante

correlacionado com o total, enquanto que o *tamanho3* apresenta uma correlação um pouco mais fraca. Finalmente, o *tamanho4* é o que apresenta um comportamento mais díspar, como podemos verificar no Quadro 4 que fornece as correlações entre vendas totais e vendas por tamanho. Isto resulta provavelmente da maior actividade promocional que caracteriza estes dois últimos tamanhos, vendidos sobretudo através do grande retalho.

tamanho 1	95,2%
tamanho 2	95,3%
tamanho 3	78,3%
tamanho 4	62,4%

Quadro 4 - coeficientes de correlação das vendas de cada tamanho com as vendas totais do produto

O modelo mais simplista de vendas de cada tamanho resultaria de multiplicar as vendas totais estimadas para o *PRODUTO X* pela percentagem média do referido tamanho no total. No entanto, essa percentagem apresenta variações significativas, pelo que se optou por um modelo causal onde as vendas totais, *TOTAL_SBP*, figurassem como variável independente na modelação das vendas dos tamanho individuais, a par de algumas variáveis já utilizadas na modelação das vendas totais do produto e do *BONUS* de produto de cada tamanho, única variável disponível com esse nível de desagregação.

Apresentam-se, no Quadro 5, os modelos para cada tamanho de *PRODUTO X*. Só no *tamanho4* a variável *TOTAL_SBP* não se mostrou significativa, o que não surpreende, visto ser menor a sua correlação com esta variável (ver Quadro 4). O *tamanho2* encontra-se no extremo oposto, já que apenas a variável *TOTAL_SBP* se mostrou significativa, resumindo-se o modelo à versão mais simples em que as vendas do tamanho são obtidas multiplicando as vendas totais pelo peso médio do tamanho.

As variáveis *BONUS1(t)* e *BONUS4(t)* também são significativas para os modelos dos respectivos tamanhos. Os coeficientes são, como seria de esperar, positivos porque os bonus representam um incentivo à compra pelo retalho.

A variável *PROM3(t)*, que já figurava no modelo 2 para as vendas totais ao retalho, também é significativa para os modelos do *tamanho3* o que se fica a dever ao facto da promoção do *tamanho3* ter um impacto diferente quando estamos a considerar como variável dependente o total do produto ou quando só estamos a considerar o próprio tamanho que estamos a promover.

A variável *PROM3(t)* é ainda significativa no modelo do *tamanho4*. Isto resulta certamente da transferência de vendas do tamanho maior para o imediatamente inferior quando este é vendido com promoção. O mesmo não se verificará em sentido inverso já que *PROM4(t)* não contribui significativamente para explicar as vendas do *tamanho3*.

MODELO TAMANHO 1	MODELO TAMANHO 2	MODELO TAMANHO 3	MODELO TAMANHO 4
------------------	------------------	------------------	------------------

R ² ajustado	92,6%	90,3%	80,8%	72,9%
Erroabolusto médio	18,2%	18,9%	15,9%	19,8%

VARIÁVEIS	Coef	Signif	Coef	Signif	Coef	Signif	Coef	Signif
CONSTANTE	---	---	---	---	20362,44	0,006	49556,68	0,001
PERIODO(t)	---	---	---	---	779,38	0,000	---	---
³ √PERIODO(t)	---	---	---	---	---	---	-8636,50	0,055
MJ	---	---	---	---	---	---	19022,89	0,000
SO	---	---	---	---	---	---	18773,44	0,000
STOCK(t-1)	-0,03	0,000	---	---	-0,06	0,000	---	---
VEND_SBP(t)	0,29	0,000	0,32	0,000	0,13	0,001	---	---
VEND_TAM4(t-2)	---	---	---	---	---	---	-0,34	0,040
PROM3(t)	---	---	---	---	-0,18	0,053	-0,34	0,033
PROM4(t)	---	---	---	---	---	---	-0,50	0,003
BONUS1(t)	123,89	0,008	---	---	---	---	---	---
BONUS4(t)	---	---	---	---	---	---	192,79	0,092

Quadro 5 - Valores estimados para os modelos de vendas da J&J ao retalho por tamanho

As vendas dos *tamanhos* 3 e 4 apresentam tendências opostas reflectindo eventualmente uma transferência de vendas de um tamanho para o outro.

Apresentamos na Figura 8 os gráficos que mostram a qualidade do ajuste de cada modelo.

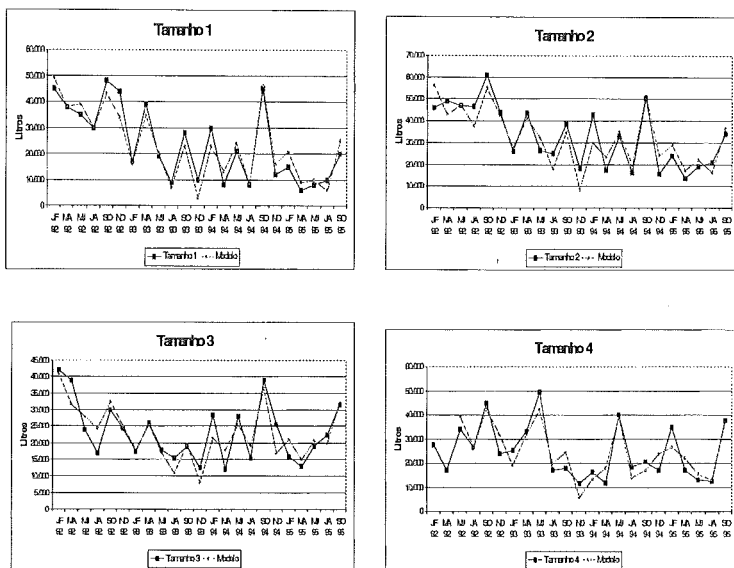


Figura 8 - Qualidade do ajuste dos modelos de previsão de vendas ao retalho por tamanho

A abordagem da modelação adoptada justifica-se do ponto de vista prático, fornece resultados bastante ajustados como se verifica na Figura 8, mas não dá qualquer garantia de consistência das previsões por tamanho com as previsões totais do *PRODUTO X*.

Com o objectivo de verificar se o somatório das previsões por tamanho está ou não desfasado da previsão do total de *PRODUTO X*, somamos as previsões por tamanho e observamos que o resultado está muito próximo da previsão total, como se pode observar na Figura 9, o que nos dá alguma garantia da robustez dos modelos desenvolvidos.

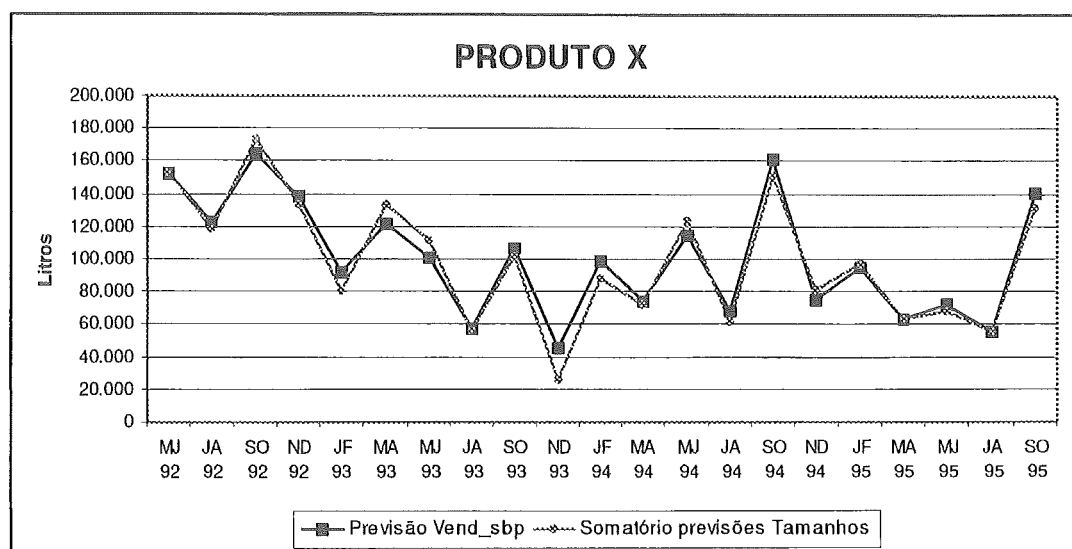


Figura 9 - Previsão do total de *PRODUTO X* vendido ao retalho vs somatório das previsões por tamanho

Por não dispormos das vendas ao consumidor por tamanho, a modelação das vendas da J&J por tamanho não pôde seguir metodologia idêntica à adoptada para estabelecer o modelo para o total de *PRODUTO X*. O facto da soma das previsões por tamanho ser muito próximo da previsão total é sinal de que os modelos referentes a cada tamanho estão a produzir bons resultados, apesar das suas limitações.

4. Conclusões

A temática dos modelos de previsão reveste-se de grande importância prática. A redução dos stocks evitando rupturas, a necessidade de ter produto suficiente para efectuar promoções e a avaliação do impacto da actividade promocional nas vendas são factores que motivam a investigação nesta área.

Neste estudo piloto foi possível desenvolver modelos causais simples para as vendas, ao retalho e ao consumidor final, de um produto de higiene pessoal, com uma posição importante no seu mercado. Isto só foi possível depois de compreender o papel do canal logístico e de estimar uma variável que funciona como barómetro desse canal, o stock no mercado.

Há no entanto factores que dificultam muito este tipo de abordagens tais como o desconhecimento das vendas por produto e por tamanho para as várias tipologias de retalhista. A abertura das fronteiras comunitárias, onde empresas multinacionais comercializam produtos

com embalagens exactamente iguais, em países vizinhos, com preços de venda ao retalho por vezes bastante diferentes, permitiu que o canal grossista de um país venda produtos ao país vizinho e veio dificultar ainda mais o conhecimento dos valores das vendas. O exemplo do que acontece entre Portugal e Espanha é, a este respeito, elucidativo. Outro factor altamente perturbador é a grande variabilidade das séries de vendas, causada pela fortíssima actividade promocional, muito superior à verificada em mercados do norte da Europa.

Por último, o facto de só existirem dados bimestrais relativamente às vendas ao consumidor torna o número de observações pequeno, obrigando a que se considere como hipótese de trabalho que o efeito das variáveis de marketing não depende do tempo, o que não é completamente realista.

5. Bibliografia

- [1] Blattberg, R.C. and Levin, A., Modelling the effectiveness and profitability of trade promotions, *Marketing Science Journal* 6 (1987) 124.
- [2] Lilien, G., Kotler, P. and Moorthy, K., *Marketing Models*, Prentice-Hall (1992).
- [3] Madridakis, S. and Wheelwright, S., *Forecasting, Methods for Management*, 5th edition, John Wiley & Sons (1989).
- [4] Murteira, B., Müller, D. and Turkman, K., *Análise de Sucessões Cronológicas*, McGraw-Hill (1993).

MODELAÇÃO DE UM SISTEMA DE COMERCIALIZAÇÃO DE ESPECIALIDADES FARMACÊUTICAS NO DISTRITO DE CASTELO BRANCO

Rute Maria Gomes Abreu

Instituto Politécnico da Guarda
Escola Superior de Tecnologia e Gestão
Dept. de Contabilidade e Auditoria, Gab. 50
6300 Guarda - Portugal

Ruy Araújo Costa

Universidade Nova de Lisboa
Faculdade de Ciências e Tecnologia
Dept. de Matemática
2825 Monte de Caparica - Portugal

Abstract

In this paper we present a modelling methodology, applied to a pharmaceutical products wholesale system, that defines inventory control rules for those products.

Statistical analysis and ABC analysis were carried out. Afterwards, a global classification was defined, assigning one of three classes to each product.

Inventory control rules were defined to each class and some "historical data simulations" were carried out to assess the adequacy of those rules.

Limitations, important characteristics and developments of this methodology are also referred.

Resumo

Apresenta-se a metodologia de modelação de um sistema de comercialização de especialidades farmacêuticas num grossista do distrito de Castelo Branco, com vista à definição de regras de gestão de *stocks* dessas especialidades.

Após o carregamento informático da informação disponível, foram desenvolvidas análises estatísticas e análises ABC que permitiram hierarquizar as apresentações em três classes.

Foram propostas regras de gestão de *stocks* das especialidades pelas três classes e levadas a cabo "simulações históricas" para aferir da adequabilidade dessas regras.

Refere-se ainda limitações, potencialidades e possíveis desenvolvimentos do modelo apresentado.

Keywords

Modelling, Inventory Control, Pharmaceutical Products.

1. Introdução

Apresentar-se-á uma metodologia de modelação do sistema de comercialização de especialidades farmacêuticas, com vista à definição de regras de gestão dos correspondentes *stocks*.

Para o desenvolvimento deste trabalho foram utilizados os dados cedidos por uma empresa, sediada no distrito de Castelo Branco, cuja actividade económica é o comércio por grosso de especialidades farmacêuticas e acessórios de farmácia.

Na figura 1 seguinte apresenta-se um esquema da abordagem efectuada:

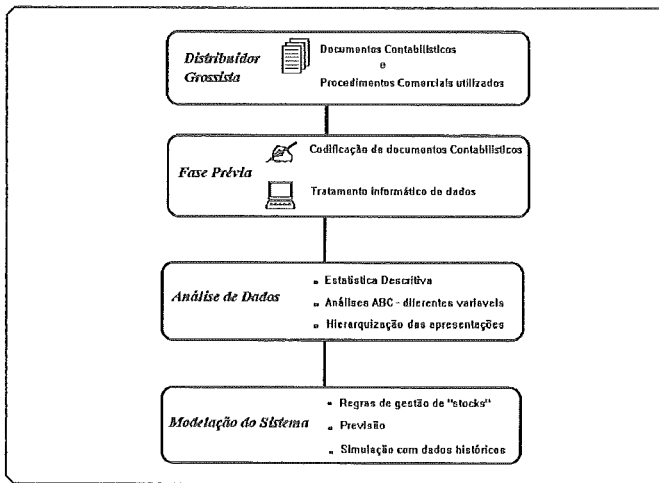


Figura 1 - Metodologia de modelação do Sistema

Ainda que a comercialização de especialidades farmacêuticas seja claramente dependente das particularidades de cada distrito (nomeadamente, em relação às especificações da respectiva carta sanitária e, obviamente, à influência das políticas de *marketing* farmacêutico utilizadas), considera-se que a metodologia empregue é genérica, podendo ser adaptada para qualquer outra região do país.

O processo de distribuição das especialidades farmacêuticas centra-se no Estado como principal regulador e controlador do mercado (apresentações, regime de preços, margens de comercialização, comparticipações, relações comerciais dos intervenientes e documentação oficial obrigatória).

A figura 2 seguinte apresenta um esquema do processo de distribuição das especialidades farmacêuticas.

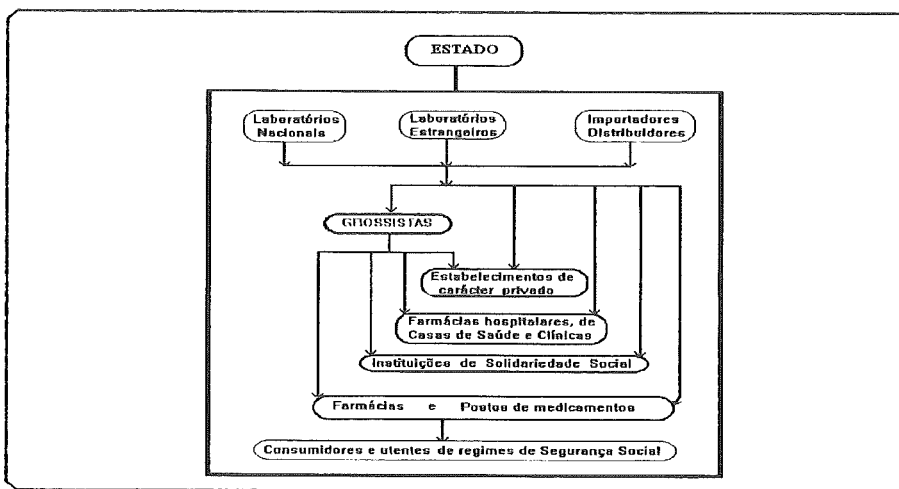


Figura 2 - Processo de distribuição de especialidades farmacêuticas

Os laboratórios nacionais e estrangeiros são os produtores. Estes e os importadores/distribuidores abastecem directamente as farmácias de oficina, hospitais, casas de saúde e outros estabelecimentos de carácter privado (lares de terceira idade, infantários,

colégios e escolas). As condições de compra são, normalmente, vantajosas ao nível de descontos comerciais, mas implicam quantidades elevadas.

Os armazenistas adquirem as apresentações aos importadores/distribuidores, laboratórios nacionais e estrangeiros, que fornecem, dentro de regiões mais ou menos restritas, às farmácias de oficina e postos de medicamentos.

Ao público em geral e, em particular, aos utentes dos diferentes regimes de Segurança Social são as farmácias de oficina e os seus postos de medicamentos que fornecem os medicamentos prescritos e sem prescrição médica obrigatória (ou venda livre), especialidades de veterinária e acessórios de farmácia.

Nas especialidades farmacêuticas não actuam as regras normais de mercado, ou seja, a procura não é influenciada pelo preço, mas sim pelo médico que prescreve ao doente (consumidor final) que, geralmente, aceita a decisão médica.

A figura 3 seguinte apresenta os principais factores que influenciam a procura:

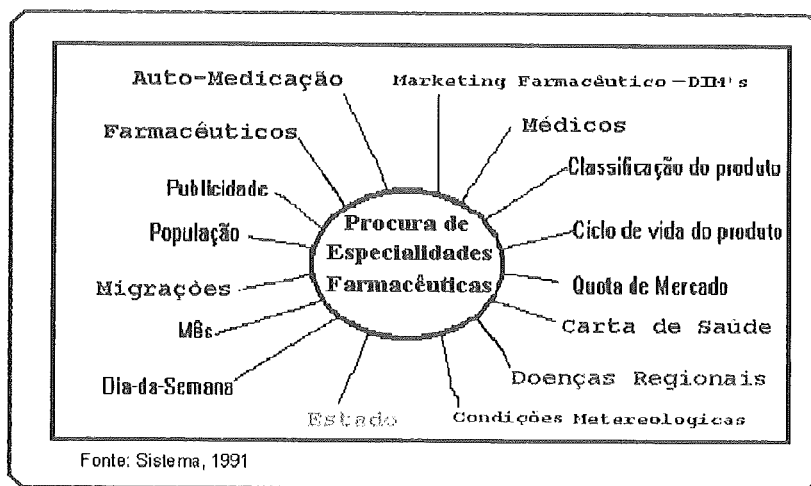


Figura 3 - Principais factores de influência na Procura das especialidades farmacêuticas

Factor essencial no ciclo de vida dos medicamentos é o *marketing* farmacêutico. Este baseia-se numa estratégia de diálogo dos delegados de informação médica (DIM) ou representantes dos laboratórios com o médico, no sentido de este prescrever os medicamentos que promovem e, de forma indirecta, na relação médico-doente pois o primeiro actua como decisor da compra.

2. Fase prévia

A empresa grossista disponibilizou os seus documentos contabilísticos (facturas de compras, vendas, notas de débito e de crédito e folhas de inventário) relativos aos anos de 1991 e 1992.

Uma primeira análise destes documentos permitiu constatar que a codificação dos dados seria extremamente lenta. Com efeito, todos os documentos disponibilizados eram manuscritos, podendo, desde logo, detectar-se algumas incorrecções. Refira-se ainda que o elevado número de apresentações de medicamentos, especialidades de veterinária e acessórios de farmácia

constantes no ficheiro informático (32129) ainda tornou mais moroso o processo de codificação.

Iniciou-se o processo de codificação a partir do inventário de 1990 e documentos de 1991, tendo-se processado informaticamente os dados relativos a 1991. De referir que este processo de codificação e tratamento informático decorreu durante catorze meses.

Assim, a maior parte das análises referidas neste trabalho têm como base os dados de 1991. Para algumas especialidades farmacêuticas foi, ainda, possível disponibilizar-se a informação relativa a 1992.

Actualmente continua a ser levado a cabo o processo de codificação e tratamento informático de documentos relativos a 1992. Em relação aos dados posteriores a 1992 pode referir-se que já se encontram disponíveis em suporte informático, embora ainda sem qualquer tratamento analítico. Espera-se, assim, que num futuro próximo se possa dispor de um maior historial de dados tratados.

3. Análise de dados

3.1 Caracterização base do mercado regional e do Sistema

No Sistema estudado, em relação a 1991, observamos que existem 3461 apresentações de medicamentos, 182 de especialidades de veterinária e 97 de acessórios de farmácia, que perfaz um total de 3740 apresentações. Refere-se que no mercado não há controlo do número de apresentações de uso veterinário, nem de acessórios de farmácia.

	Medicamentos						Veterinária	Acessórios de Farmácia	TOTAL		
	Comparticipáveis			Não Comparticipáveis							
	Nacionais			Importações			Éticos	Venda Livre			
	50%	80%	100%	50%	80%	100%					
Apresentação	1398	652	99	238	175	53	465	381	182	97	3740
	2149			466							

Quadro 1 - Número de apresentações do Sistema em 1991

No mercado relativamente às especialidades farmacêuticas e segundo informações do Infarmed, em 1991, encontramos 4198 apresentações com marca registada distinta. Estas desagregam-se em 117 formas farmacêuticas diferenciadas que implicam 6026 apresentações. Considerando que existem diferentes dosagens e tamanhos podemos então obter um total de 10838 apresentações. Representando estas, apenas medicamentos compartilhados e não compartilhados existentes no mercado.

	Medicamentos							Veterinária	Acessórios de Farmácia	TOTAL	
	Comparticipáveis						Não Comparticipáveis				
	Nacionais			Importações			Éticos				Venda Livre
	50%	80%	100%	50%	80%	100%					
Marcas	1058	469	54	339	157	40	1541	540	n.d.	n.d.	4198
	1581			536							
Formas Farmacêuticas	1562	772	70	438	218	50	2235	681	n.d.	n.d.	6026
	2404			706							
Apresentações	2541	1467	144	671	406	95	4602	912	n.d.	n.d.	10838
	4152			1172							

Fonte: Infarmed, 1991

Legenda: n.d. = não disponível

Quadro 2 - Número de apresentações do Mercado em 1991

Com vista à determinação da influência da empresa no mercado regional, elaboraram-se os Quadros 3 e 4 seguintes, relativos ao número de embalagens e ao volume de vendas totais de especialidades farmacêuticas prescritas, respectivamente.

Escala: Número de embalagens e %					
Mês (1)	ARS (2)	CLIENTES		CLIENTES	
		ARS (3)	Sistema (4)	%ARS (5)=(3)/(2)	% Sistema (6)=(4)/(3)
Janeiro	174716	31577	9453	18,1	29,9
Fevereiro	154693	29013	8423	18,8	29,0
Março	163668	32129	10627	19,6	33,1
Abril	160440	30956	8263	19,3	26,7
Maio	163652	31424	9089	19,2	28,9
Junho	154499	29366	7845	19,0	26,7
Julho	155516	27850	9283	17,9	33,3
Agosto	132463	21233	4141	16,0	19,5
Setembro	163152	33017	8595	20,2	26,0
Outubro	186781	35207	10366	18,9	29,4
Novembro	164957	31757	9159	19,3	28,8
Dezembro	148047	28234	7451	19,1	26,4
TOTAL	1922584	361763	102695	18,8	28,4

Quadro 3 - Número total de embalagens de especialidades farmacêuticas prescritas em 1991

Mês (1)	ARS (2)	CLIENTES		CLIENTES	
		ARS (3)	Sistema (4)	%ARS (5)=(3)/(2)	% Sistema (6)=(4)/(3)
Janeiro	239.909.263\$	42.485.079\$	11.237.066\$	17,7	26,6
Fevereiro	215.006.083\$	39.247.621\$	10.406.568\$	18,3	26,5
Março	226.027.462\$	43.425.642\$	11.471.977\$	19,2	26,4
Abril	231.280.830\$	43.340.739\$	10.970.184\$	18,7	25,3
Mai	240.845.078\$	45.314.322\$	11.336.371\$	18,8	25,2
Junho	230.320.544\$	42.798.219\$	10.818.573\$	18,6	25,3
Julho	235.373.211\$	40.250.771\$	10.989.533\$	17,1	27,3
Agosto	202.761.792\$	31.676.787\$	5.300.029\$	15,6	16,7
Setembro	251.511.877\$	49.691.466\$	11.737.255\$	19,8	23,6
Outubro	293.430.940\$	53.409.087\$	13.332.021\$	18,2	25,0
Novembro	257.561.908\$	48.063.437\$	11.670.817\$	18,7	24,3
Dezembro	230.592.896\$	42.217.473\$	9.469.985\$	18,3	22,4
TOTAL	2.854.621.981\$	521.920.641\$	128.830.380\$	18,3	24,7

Fonte: ARS-Estatísticas, 1991 e Sistema, 1991

Quadro 4 - Volume de vendas totais de especialidades farmacêuticas prescritas em 1991

A análise dos quadros anteriores permite determinar a quota de mercado do Sistema para os seus clientes que é cerca de 28% em número de embalagens prescritas e 25% em volume de vendas. Estes valores não incluem nem especialidades de veterinária nem acessórios de farmácia. Por outro lado, pode ainda constatar-se que a quota de mercado dos clientes do Sistema face ao mercado total representa em termos globais 18%.

Com vista à hierarquização dos fornecedores do Sistema foram desenvolvidas análises ABC por "volume de vendas" e "unidades vendidas". De entre o total de fornecedores (211), destacam-se 22 que pertencem à classe A, simultaneamente, nas duas análises, correspondendo assim aos "fornecedores mais importantes". De notar que só se pode efectuar tal análise porque apenas existe um fornecedor para cada apresentação.

A figura 4 seguinte representa o "volume de vendas" e as "unidades vendidas" para os 22 fornecedores classificados como "mais importantes".

Como se pode observar, o maior "volume de vendas" não está associado ao maior número de "unidades vendidas".

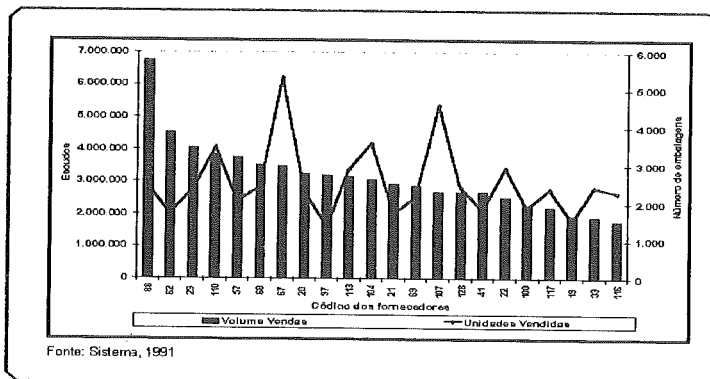
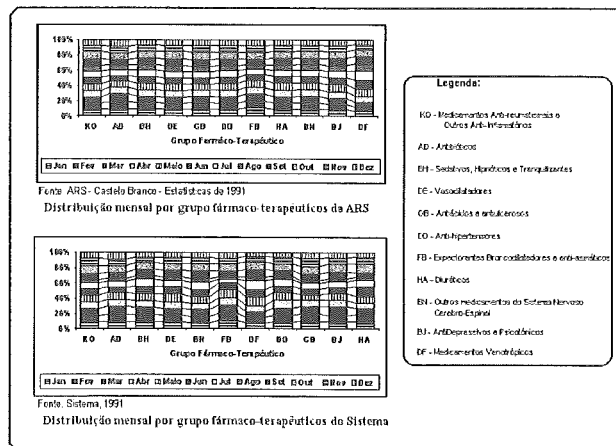


Figura 4 - Distribuição do Volume e Unidades Vendidas por 22 Fornecedores com ponderação "mais importante"

Sintetizou-se no quadro 5 seguinte a análise da distribuição da procura das especialidades farmacêuticas por grupo fármaco-terapêutico:



Quadro 5 - Análise de Distribuição Mensal da Procura por Grupo Fármaco-Terapêutico de Especialidades Farmacêuticas

É interessante mostrar que a distribuição da procura relativa aos quatro principais grupos fármaco-terapêuticos quer na ARS de Castelo Branco, quer no Sistema, é idêntica, o que se explica pelas doenças mais frequentes no distrito. Contudo, a partir do quinto grupo pode observar-se uma divergência entre os valores da ARS e o Sistema, o que é devido à Carta Sanitária dos Concelhos da Covilhã e Fundão (i.e., às doenças específicas nestes concelhos).

A agregação por grupos fármaco-terapêuticos permite apenas algumas análises mais abrangentes, dado que num mesmo grupo se poderão encontrar especialidades cujos comportamentos, em termos de procura, sejam completamente distintos. Assim, tornou-se necessário, classificar cada especialidade farmacêutica individualmente, com vista ao estabelecimento de regras de gestão dos respectivos *stocks*.

3.2 Classificação hierárquica das apresentações

Com vista à classificação hierárquica das apresentações no Sistema, poderíamos considerar diferentes variáveis de classificação: unidades vendidas, volume de vendas, unidades compradas, volume de compras, unidades de *stock* médio e valor de *stock* médio.

Classificar uma dada especialidade farmacêutica com base nestas variáveis não é uma tarefa "linear". Com efeito, não só as escalas envolvidas são distintas, como também a importância relativa dos diferentes variáveis não é idêntica. Assim, inicialmente, foram desenvolvidas seis análises ABC relativamente as variáveis anteriormente referidas.

Em cada análise ABC, a classe A engloba as apresentações que correspondem aproximadamente a 60% do valor acumulado relativo da variável em estudo. Normalmente, o número de apresentações pertencentes a esta classe, ronda os 10% do total. A classe B é definida pelas "primeiras" apresentações que não pertencendo a classe A, correspondem aproximadamente a 30% do valor acumulado relativo da variável em estudo. O número de apresentações pertencentes a esta classe ronda 30% do total. A classe C inclui todas as

apresentações não incluídas nas classes A e B (correspondendo o número de apresentações nesta classe aproximadamente 60% do total).

Nas figuras 5, 6 e 7 seguintes são representadas as correspondentes curvas de análise ABC.

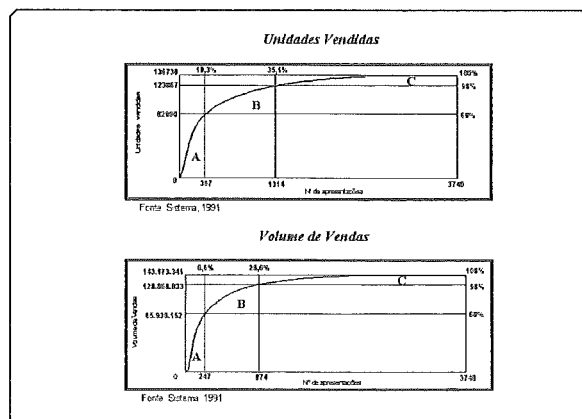


Figura 5 - Curvas de análise ABC por número de "Unidades Vendidas" e "Volume de Vendas"

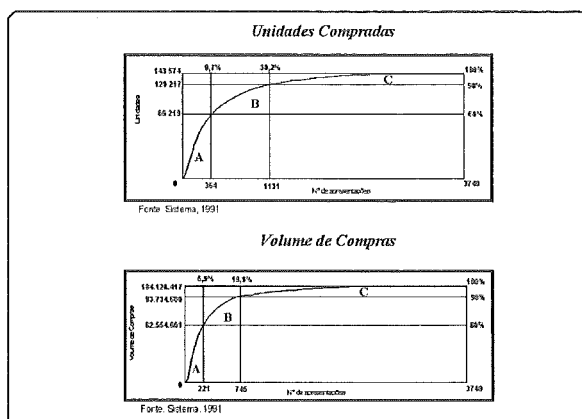


Figura 6 - Curvas de análise ABC por número de "Unidades Compradas" e "Volume de Compras"

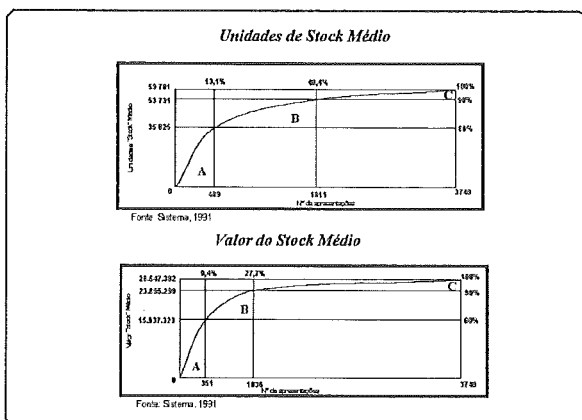


Figura 7 - Curvas de análise ABC por número de "Unidades de Stock médio" e "Valor do Stock médio"

Para hierarquizar as apresentações foram adoptadas três classes de importância: I (a mais importante), II (a intermédia) e III (a menos importante). A "importância" de uma especialidade farmacêutica está associada as seis variáveis anteriormente referidas.

Inicialmente, converteu-se, para cada apresentação, os resultados das análises ABC para uma "escala numérica", atribuindo o valor 1, 2 e 3, respectivamente, a um resultado A, B e C.

Para determinar a hierarquização das classes foram definidas ponderações que reflectem o grau de importância relativa das análises ABC desenvolvidas. Tendo em conta as particularidades deste sistema, é relativamente fácil, constatar que o número de "unidades vendidas" é uma variável muito importante face as demais e que o "volume de vendas" se deve destacar claramente de entre as cinco variáveis restantes.

Tendo sido desenvolvidas várias experiências com objectivo de determinar os factores de ponderação referidos, constatou-se que era preciso algum cuidado para garantir uma distribuição relativamente *fina* da classificação média ponderada dos seis resultados das análises ABC.

Se é óbvio que uma especialidade pertencente à Classe A nas seis análises é classificada com I e que uma especialidade pertencente à Classe C nas seis análises é classificada com III, não é imediata a classificação de uma especialidade com "resultados intermédios" nas diferentes análises ABC.

As ponderações indicadas no Quadro 6 seguinte, foram obtidos após o processo de calibração referido e permitem responder adequadamente às preocupações referidas:

<i>Variável</i>	<i>Ponderação</i>
1. Unidades Vendidas	0,5000
2. Volume de Vendas	0,3500
3. Unidades Compradas	0,0375
4. Volume de Compras	0,0375
5. Unidades de <i>stock</i> médio	0,0375
6. Valor do <i>stock</i> médio	0,0375

Quadro 6 - Distribuição das Análises ABC por variável e respectiva ponderação

Para cada especialidade farmacêutica, e com base nos seis resultados das análises ABC (convertidos para a "escala numérica"), torna-se agora possível determinar a **classificação global da apresentação**, calculando a média ponderada dos resultados referidos com a ponderação. A classificação global é um valor real do intervalo [1,0; 3,0].

Com base na classificação global de cada especialidade farmacêutica, teremos agora de atribuir a classificação I, II e III. Quanto mais baixo for o valor da classificação global, maior é a importância da apresentação correspondente. Assim, é imediata a atribuição da classificação I a especialidades com classificação global igual a 1,0, bem como a atribuição da classificação III a especialidades com classificação global 3,0.

A hierarquização das apresentações nos três níveis visa responder a várias preocupações:

- Pretende-se identificar especialmente as apresentações "mais importantes", atribuindo-lhes a classe I.

O processo de classificação a adoptar deverá ser confrontado com a realidade, garantindo-se que na classe I estão integradas a maioria das apresentações que a prática já havia levado a concluir tratar-se de "especialidades mais importantes".

Tendo em conta que um dos principais objectivos da identificação das apresentações da classe I é a sua futura monitorização mais cuidada, a nível de gestão de *stocks*, parece aceitável limitar a um máximo de 10% do número total, o número de apresentações da classe I.

- Relativamente às especialidades classificadas na classe II, recomendar-se-á a adopção de uma política de gestão de *stocks* mais flexível do que a que virá a ser proposta para a classe I, mas, que ainda assim, envolve algum cuidado.

Assim, considera-se desejável que o processo de classificação a desenvolver não origine um número de apresentações na classe II superior a 30% do seu número total.

- Espera-se que o processo de classificação inclua na classe III as apresentações que a prática demonstrou serem apresentações "menos importantes".

Uma segunda fase do processo de calibração permitiu satisfazer as condições referidas e propor três níveis de valor de classificação global (para atribuição das classificações I, II e III) apresentadas no quadro 7 seguinte.

Classe	Pontuação Ponderada
I	[1,0, 1,5[
II	[1,5, 2,5[
III	[2,5, 3,0]

Quadro 7 - Níveis de classificação global por Classes I, II e III

A figura 8 seguinte apresenta a curva de distribuição da classificação hierarquizada por Classes das 3740 apresentações.

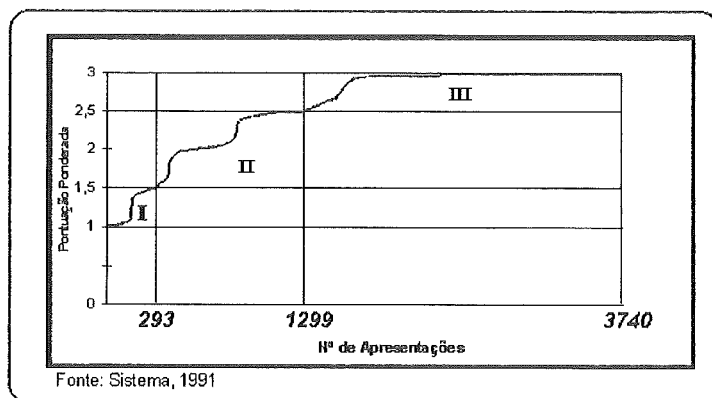


Figura 8 - Curva de distribuição da Classificação Hierárquica por Classes das 3740 apresentações do Sistema em 1991

O quadro seguinte permite relacionar a atribuição das classificações I, II e III com o número de apresentações e com as seis variáveis de classificação referidas.

Classe	Número apresentações		Vendas		Compras		Stock"médio	
			Unidades	Volume	Unidades	Volume	Unidades	Valor
I	293	8%	50%	53%	51%	54%	29%	35%
II	1006	27%	33%	35%	32%	37%	40%	47%
III	2441	65%	17%	12%	17%	9%	31%	18%

Quadro 8 - Distribuição das classes hierarquizadas por variável

A análise do quadro anterior permite constatar que apenas 8% das apresentações mereceram a classificação I ("mais importante"), correspondendo a aproximadamente 50% das unidades ou volume de vendas e compras e a aproximadamente 30% do nível médio do *stock* anual (quer em termos de unidades, quer de valor).

A classe II corresponde a 27% das apresentações, a aproximadamente 30% das vendas e compras, e a aproximadamente 45% do *stock* médio.

A classe III corresponde a 65% das apresentações, tendo uma importância claramente residual relativamente as seis variáveis referidas.

Uma questão importante pode ser levantada: "Qual o grau de (in)dependência entre as seis variáveis envolvidas no processo de classificação?"

Para reforçar a importância desta questão, recorde-se que se atribuiu o peso de 0,5 à variável "Unidades Vendidas" e 0,35 à variável "Volume de Vendas". Se essas variáveis apresentassem uma óbvia relação entre elas, de algum modo se poderia considerar que se tinha atribuído um peso total de 0,85 a um "único factor".

Antes de se proceder a uma análise quantitativa relativa a esse par de variáveis, refira-se que, a priori, *pareceria* possível conceber uma relação entre elas. Com efeito, o "Volume de Vendas" associado a uma dada apresentação, mais não é do que o correspondente número de "Unidades Vendidas" multiplicado pelo preço de venda unitário. No entanto, a posteriori, pode-se constatar que algumas das apresentações que correspondem a um maior número de "Unidades Vendidas" estão associadas a um baixo preço de venda unitário, pelo que não correspondem a elevados "Volumes de Vendas" (ocorrendo também "situações inversas").

A análise estatística levada a cabo para averiguar do grau de (in)dependência entre esse par de variáveis conduziu a duas conclusões:

- Em primeiro lugar, existe evidência que permite rejeitar a hipótese de independência (ao nível de significância de 5%) e,
- Em segundo lugar, não existe nenhuma relação de dependência forte ou óbvia entre essas variáveis, observando-se antes uma dependência "difusa".

Idênticos resultados podem ser referidos para outros pares de variáveis.

Crê-se que, ainda assim, o processo de classificação proposto se revelou muito "ajustado à realidade". Com efeito, o processo de classificação levado a cabo (com a variação dos pesos atribuídos às variáveis), que terminou com a propositura dos pesos referidos no Quadro 6, permitiu cumprir as principais expectativas no tocante a algumas "apresentações características". Assim, certas apresentações que a experiência indicou serem "pouco relevantes" apareceram incluídas na classe III, enquanto que outras apresentações, claramente importantes no dia-a-dia, surgiram na classe I.

4. Regras de Gestão de *Stocks*

A gestão de *stocks* de especialidades farmacêuticas não pode pautar-se apenas (nem mesmo primordialmente) por factores de ordem económica. Com efeito, as razões de saúde pública condicionam fortemente esta gestão. No entanto, nem todas as especialidades têm igual importância para o sistema, motivo porque se desenvolveu a classificação hierarquizada apresentada.

Se para todas as especialidades é necessário contemplar um *stock* mínimo que vise minimizar a possibilidade de ruptura, esta questão torna-se especialmente relevante para as especialidades classificadas como "mais importantes". Assim, considera-se que, para um dado instante t , se define para cada apresentação:

- S_t - nível de existências, que corresponde às venda médias dos últimos doze meses;
- S_k - nível de existências de um determinado mês k ;
- $S_t [(1/12) \times \sum_{K=t-12}^{t-1} S_k]$, com $t > 12$
- SM_t - nível de existência máximo recomendado;
- Sm_t - nível de existências mínimo permitido;
- V_t - número de unidades vendidas nos últimos doze meses.

Com base nos valores de classificação hierárquica de S_t e V_t pode propor-se a determinação do valor de SM_t e de Sm_t de acordo com as regras apresentadas no Quadro 9 seguinte:

Classe	I	II		III	
		$S_t \leq 8$	$S_t > 8$	$1 \leq V_t \leq 12$	$V_t > 12$
Stock Máximo Recomendado	S_t^*	8	S_t^*	7	$0,5V_t^*$
Stock Mínimo Permitido	$0,45S_t^*$	3	$0,3S_t^*$	2	1
Quantidade a Encomendar	$0,55S_t^*$	5	$0,7S_t^*$	5	$\text{Max}(0,5V_t^* - 1, 1)$

Legenda: * = arredondamento simétrico

Quadro 9 - Resumo de Regras de Gestão de *Stocks*

Nas regras de gestão de *stocks* apresentadas procurou-se adequar a experiência empírica do sector, com as regras impostas pelo mercado, utilizando relações matemáticas simples.

Assim, relativamente às apresentações "menos importantes", isto é, classificadas com III, propõe-se um nível mínimo do *stock* muito baixo (1 unidade se o número de unidades vendidas nos últimos doze meses não exceder 12 unidades, ou 2 unidades, caso contrário). Se o número de unidades vendidas nos últimos doze meses não exceder as 12 unidades, propõe-se um nível máximo de *stock* igual a 50% desse valor com eventual arredondamento simétrico. Tal traduz-se num nível máximo que varia entre 1 a 6 unidades, pelo que a encomenda a efectuar poderá ser entre 1 a 5 unidades (diferencial entre o nível mínimo e máximo do *stock*, com o mínimo de 1 unidade). Quando as vendas do último ano ultrapassam as 12 unidades fixa-se o nível máximo do *stock* em 7 unidades, encomendando-se, então, 5 unidades.

Refira-se que a prática demonstra que só a título excepcional, se contempla a aquisição de menos de 5 unidades de uma dada apresentação.

Na classe I propõe-se um nível máximo do *stock* igual ao nível médio das vendas dos últimos doze meses, o que reflecte a importância das apresentações desta classe.

O *stock* mínimo igual a 45% do valor médio das vendas dos últimos doze meses assegura uma margem de manobra de armazenagem de 55% do nível médio das vendas desse período. Assim, normalmente, serão feitas duas encomendas (uma "normal" e outra "extraordinária").

A política proposta para a classe II reflecte o carácter de "transição" dessa classe relativamente às classes I e III. Para apresentações com o nível médio de *stock* dos últimos doze meses não superior a 8 unidades, recomenda-se uma política "idêntica" à proposta para as "melhores" apresentações da classe III: um nível mínimo de 3 unidades e um nível máximo de 8 unidades, a que correspondem encomendas de 5 unidades. De notar que estes valores são superiores em 1 unidade aos propostos para as "melhores" apresentações da classe III, reflectindo a maior importância das apresentações da classe II e o maior cuidado para evitar situações de ruptura.

Para as apresentações "mais importantes" da classe II foi sugerida uma política idêntica à proposta para a classe I: o nível máximo é igual ao *stock* médio dos últimos doze meses (à semelhança do que se propôs para a classe I) e o nível mínimo é igual a 30% desse nível. De notar que o nível mínimo proposto para a classe I é superior a este valor, dada a maior importância desta classe.

As regras de gestão de *stocks* apresentadas no quadro 9 subentendem a **monitorização** permanente dos *stocks*, sendo colocadas encomendas das apresentações da classe I, sempre que se atinja o nível mínimo estabelecido. Adicionalmente, e tendo em conta as especificidades deste mercado, devem ser desenvolvidas **inspecções periódicas** dos *stocks* que reflectam o grau de preocupação com a sua eventual ruptura dependendo da classificação das apresentações.

No quadro 10 seguinte apresenta-se a calendarização proposta para as revisões dos *stocks* das apresentações das três classes:

<i>Mês</i>	<i>J</i>	<i>F</i>	<i>M</i>	<i>A</i>	<i>M</i>	<i>J</i>	<i>J</i>	<i>A</i>	<i>S</i>	<i>O</i>	<i>N</i>	<i>D</i>
<i>Classe</i>												
<i>I</i>	√	√	√	√	√	√	√	√	√	√	√	√
<i>II</i>		√		√			√			√		
<i>III</i>					√							

Nota: assinalou-se com √ a realização de uma inspecção aos *stocks* no início do mês correspondente

Quadro 10 - Calendarização das Inspeções por classe

Com base na inventariação directa das existências que ocorre no mês de Dezembro do ano anterior, no início de cada ano desenvolve-se um novo processo de classificação hierarquizada, que permite, eventualmente, alterar as anteriores classificações obtidas, por uma da apresentação.

As classificações estabelecidas no início de cada ano implica a definição dos correspondentes "calendários de inspecção".

Como é evidente, não seria possível inspecionar as 2441 apresentações da classe III e as 1006 da classe II, com a periodicidade mensal proposta para as apresentações da classe I, sem que tal se traduzisse numa pesada (e, provavelmente, menos inteligente) utilização dos recursos humanos.

Os *stocks* correspondentes às apresentações da classe I deverão ser inspecionados todos os inícios dos meses, sendo então determinados os valores de SM_t e Sm_t . Nessa altura são feitas as respectivas encomendas "normais". Como se referiu, o *stock* destas apresentações é vigiado em permanência, pelo que, se se atingir o nível mínimo permitido (Sm_t) é, de imediato, feita uma encomenda "extraordinária" correspondente ao diferencial entre SM_t e Sm_t .

Relativamente às apresentações da classe II, os *stocks* deverão ser inspecionados no início dos meses de Fevereiro, Abril, Julho e Outubro, porque os consumos destas apresentações são, normalmente, alteráveis com as estações do ano, tal como as campanhas efectuadas pelos fornecedores, sendo então estabelecidos os valores de SM_t e Sm_t . Um controlo informático destes *stocks* permitirá detectar situações em que se atinja o nível mínimo permitido (Sm_t), sendo então o responsável pela gestão de *stocks* alertado para essa situação. Nessas circunstâncias, caberá a este responsável decidir qual a quantidade a encomendar, tendo em conta a data da próxima inspecção e o facto de se pretender limitar o crescimento do *stock* deste tipo de apresentação.

Para garantir uma inspecção anual ao *stock* das apresentações da classe III, dever-se-ia optar por um mês em que não decorressem inspeções de apresentações da classe II. Será desejável evitar, ainda, os períodos de férias (Junho a Agosto), do inventário global (Dezembro) e da classificação hierarquizada (Janeiro). De entre os meses possíveis constatou-se

que o final da Primavera se reflectia nos padrões de consumo de muitas apresentações, pelo que se optou por sugerir a inspecção dos *stocks* das apresentações da classe III para o mês de Maio, sendo então determinados os valores de SM_i e Sm_i . Relativamente a estas apresentações não há qualquer outro controlo adicional, dado tratarem-se de especialidades com muita baixa rotação, sendo elevado o número de apresentações que praticamente não têm movimento. Refira-se que, no entanto, dado o elevado número de apresentações nesta classe é significativo o investimento de capital correspondente.

A verificação da adequabilidade das regras enunciadas poderia ser feita, por exemplo, com o recurso à simulação. No entanto, dado o baixo volume de dados tratados disponíveis (apenas correspondentes ao ano de 1991 para todas as apresentações), não foi possível caracterizar estatisticamente a procura para posteriormente proceder à simulação).

Com o aumento do volume de dados tratados disponíveis tornar-se-ia possível simular efectivamente a aplicação das Regras de Gestão de *Stocks* propostas e avaliar correctamente da sua adequabilidade, nomeadamente, da existência de situações de rotura e dos níveis médio, máximo e mínimo do *stock*. Poder-se-ia ainda efectuar análises de sensibilidade à variação de parâmetros das regras definidas, com vista à escolha dos valores mais adequados.

5. Conclusão

Apresentou-se uma metodologia de modelação do sistema de comercialização de especialidades farmacêuticas que se traduziu na classificação das diferentes especialidades em três classes hierarquizadas e na adopção de regras de gestão de *stocks* por cada classe.

O trabalho levado a cabo foi limitado por um processo de codificação muito moroso, pela falta de colaboração da indústria farmacêutica e pela utilização de um sistema informático que se veio a revelar menos versátil e possante do que seria desejável.

O relativamente baixo volume de dados disponível em suporte informático (após o processo de codificação) não permitiu a caracterização estatística da procura e eventual recurso à simulação para testar efectivamente as regras de gestão de *stocks* propostas.

Dado que os resultados obtidos têm como base uma amostra relativa ao ano de 1991, recomenda-se particular cuidado na sua generalização. Com efeito, esses resultados reflectem a situação relativa ao ano de 1991 no distrito de Castelo Branco. No entanto, e tal como se referiu anteriormente, crê-se que a metodologia utilizada pode ser generalizada a outros horizontes temporais e/ou espaciais.

É importante ter-se presente que as peculiaridades da comercialização de especialidades farmacêuticas (nomeadamente, explicadas pela contínua investigação científica nos laboratórios e a decorrente alteração das especialidades disponíveis), aconselha alguma prudência na consideração das amostras de grande dimensão. Ainda que tal pudesse ser possível, não se considera aconselhável proceder à modelação de um sistema de comercialização de especialidades farmacêuticas com base numa amostra de, por exemplo, cinco anos.

O simples processo de hierarquização das apresentações em três classes desenvolvido, permitiu identificar as especialidades que justificariam um maior cuidado na gestão dos *stocks* correspondentes.

Embora se encontre ainda em fase de avaliação, pode, desde já, referir-se que a implementação no sistema real, das regras de gestão de *stocks*, permitiu observar uma clara melhoria na gestão global do Sistema, face ao passado.

A continuação do processo de codificação e correspondente tratamento da informação, a caracterização estatística da procura, a simulação da gestão de *stocks* com as regras propostas de modo a avaliar a eficácia, a avaliação das carteiras de encomendas, o estudo do sistema nos períodos de férias dos laboratórios e a consideração de restrições financeiras e de espaço de armazenagem são algumas vias a seguir com vista ao desenvolvimento do trabalho realizado.

Referências

- [1] Abreu, R., *Introdução à Modelação de um sistema de comercialização de especialidades farmacêuticas*, Dissertação de Mestrado em Engenharia Industrial, FCT-UNL (1995) Lisboa.
- [2] Braga, M., *Gestão do Aprovisionamento*, Editorial Presença (1991).
- [3] *Cedime Informação*, Associação Nacional de Farmácias, nº 1 e 2 do ano de 1990 e nº 3, 4, 5, 6, 7 e 8 do Ano de 1991.
- [4] *Estatísticas da facturação das Farmácias - SIMS*, Administração Regional de Saúde de Castelo Branco (1991).
- [5] Hax, A. e Candea, D., *Production and Inventory Management*, Prentice-Hall, Inc. (1984).
- [6] Hillier, F. e Lieberman, G., *Introduction to Operations Research*, McGraw-Hill Internacional Series (1990).
- [7] *Índice Nacional Terapêutico*, Tupam Editores, Janeiro, Abril, Julho e Outubro (1990,1991).
- [8] Lewis, C., *Scientific Inventory Control*, Butterworth & Co Publishers (1981).
- [9] Makridakis, S. e Wheelwright, S., *Forecasting Methods for Management*, John Wiley & Sons, Inc. (1989).
- [10] *Plano Integrado de Saúde do distrito de Castelo Branco - Diagnóstico da Situação*, Administração Regional de Saúde de Castelo Branco (1987).
- [11] Reis, L. e Paulino, A., *Gestão de Stocks e Compras*, Editora Internacional (1994).
- [12] *Risco de Morrer em Portugal em 1991*, Direcção de Serviços de Informação de Saúde, Direcção Geral dos Cuidados de Saúde Primários (1992).
- [13] Silver, E. e Peterson, R., *Decision Systems for Inventory Management and Production Planning*, John Wiley & Sons, Inc. (1985).
- [14] *Síposium Terapêutico - Enciclopédia de Especialidades Farmacêuticas*, Edições Simposium, Lda. (1991).
- [15] *Síntese de Informação Estatística - 1991*, Direcção Geral de Assuntos Farmacêuticos, Direcção de Serviços de Planeamento e Estatística.
- [16] Tadeu, L. e Assis Lopes, J.A., *1º Ciclo de cursos de formação sobre "A Investigação Operacional na Empresa" - III - Previsão e Controle para a Gestão de Empresas*, Centro de Sistemas Urbanos e Regionais da Universidade Técnica de Lisboa.
- [17] Zermati, P., *A Gestão de Stocks*, Editorial Presença (1990).

PROCESSOS ESTOCÁSTICOS DE APOIO À MANUTENÇÃO PREDITIVA PARA AMBIENTES INDUSTRIAIS

R. Jardim-Gonçalves

UNINOVA - Inst.Desenv.Novas Tec. - Centro de Robótica Inteligente (CRI)
Quinta da Torre - Campus Fac.Ciências e Tec. - Univ.Nova de Lisboa
2825 Monte Caparica - Portugal

M. Martins-Barata

UNINOVA - Inst.Desenv.Novas Tec. - Centro de Robótica Inteligente (CRI)
Quinta da Torre - Campus Fac.Ciências e Tec. - Univ.Nova de Lisboa
2825 Monte Caparica - Portugal

A. Steiger-Garção

UNINOVA - Inst.Desenv.Novas Tec. - Centro de Robótica Inteligente (CRI)
Quinta da Torre - Campus Fac.Ciências e Tec. - Univ.Nova de Lisboa
2825 Monte Caparica - Portugal

J. Assis-Lopes

Instituto Superior Técnico - Secção Autónoma de Economia e Gestão
Av. Rovisco Pais
1096 Lisboa Codex - Portugal

Abstract

This paper presents and discusses the usage of stochastic processes to support predictive maintenance for industrial environments. The aim of the work executed is to give a contribution for the development of supervision and monitoring systems, attempting to improve and automate the conformity quality control during production, mainly in companies where this task is computer assisted.

The automated forecast of production fails and defects is a strategy that helps to achieve real time quality control and optimises the machine maintenance intervals, minimising unscheduled stops. The developed work covers sensorial information acquisition, forecast modeling and monitoring system parameters' identification, estimation and validation.

This paper begins presenting how important is the usage of sensorial systems for numeric control machine monitoring and supervision. Then, a detailed description of the study and experimental analyses done is also presented. This text finishes showing the achieved results, conclusions and perspectives for future work.

Resumo

A previsão automatizada de falhas e de defeitos de fabrico, é uma estratégia que ajuda a obter controlo de qualidade em tempo real e otimizar os períodos de manutenção das máquinas envolvidas no processo de fabrico, minimizando o número de paragens não previstas.

O objectivo do trabalho apresentado nesta comunicação visa contribuir para o desenvolvimento de sistemas de monitoração e supervisão de máquinas ferramenta de controlo numérico (CN), de maneira a melhorar e automatizar o controlo de qualidade de conformidade durante a produção, especialmente em fábricas onde esta é assistida por computador.

Desenvolvido no âmbito do trabalho de dissertação de Mestrado em Investigação Operacional e Engenharia de Sistemas (Instituto Superior Técnico) [7], este trabalho compreende: (1) recolha da informação sensorial; (2)

identificação, validação e estimativa dos coeficientes dos modelos de previsão; (3) estimativa dos parâmetros do sistema para monitoração e (4) identificação do estado de funcionamento da máquina.

Esta comunicação começa por apresentar a relevância dos sistemas sensoriais e perceptuais na monitoração e supervisão de máquinas ferramenta CN, descrevendo a sua relação com a qualidade de fabrico quando aplicados aos sistemas flexíveis de produção. De seguida, descreve a forma como o estudo, a modelação e a análise experimental foram conduzidos, explicando como os sistemas para monitoração e identificação foram desenvolvidos. Por fim apresenta os resultados obtidos, conclusões e perspectivas de trabalho futuro.

Keywords

Predictive maintenance, Sensorial systems, Monitoring.

1. Introdução e apresentação do problema

As necessidades vitais das tecnologias e ciências modernas, aplicadas aos sistemas de produção, têm considerado prioritária a tarefa do desenvolvimento de uma aproximação sistemática para o estudo dos fenómenos e processos. Tal estudo tem considerado crítico o problema da fiabilidade dos sistemas de produção, uma vez que é ele o responsável pela qualidade do produto final [17].

É de extrema importância para o gestor de um sistema de produção ter conhecimento dos aspectos relacionados com o nível de qualidade da produção que está a ser realizada, bem como do estado de funcionamento das máquinas, para que possa efectuar o controlo de qualidade adequado. No contexto dos Sistemas Flexíveis de Produção (SFP), estes dois aspectos assumem uma importância ainda maior porque, enquanto que a qualidade da produção está relacionada com o processo e meios utilizados, o estado de funcionamento das máquinas está relacionado principalmente com a quantidade de trabalho e os esforços a que elas são sujeitas.

Um rigoroso processo de controlo dos componentes fabricados, por forma a que estes sejam produzidos dentro dos parâmetros de fabrico exigidos (e.g., que as tolerâncias especificadas pelo desenho sejam cumpridas), é necessário para se encontrar qualidade de conformidade [12]. Este controlo pode ser feito "a posteriori", em que se analisa o componente após o fabrico, ou "a priori", através de um processo preditivo dos agentes participantes no processo de fabrico (por exemplo, prever que uma ferramenta de uma máquina está a ficar gasta e que dentro de um período de tempo não estará em condições, permitirá que se evite o fabrico de componentes defeituosos).

Nos SFP, as máquinas ferramenta de Controlo Numérico (CN) desempenham um papel muito importante, uma vez que são elas que produzem os componentes que irão fazer parte dos produtos a comercializar. A produção de componentes sem defeitos traz, sem dúvida, um grande contributo para a melhoria da produtividade e qualidade do produto final.

É também um facto que, à medida que uma máquina se vai degradando, a qualidade do seu trabalho também diminui. Assim, pode-se dizer que a qualidade do produto e o estado de funcionamento das máquinas intervenientes na sua manufactura estão directamente relacionados.

Numa máquina CN as falhas podem ocorrer de uma forma aleatória (i.e., ocasionais), ou então como consequência da degradação sucessiva da sua estrutura (i.e., não ocasionais). No primeiro caso a sua previsão apresenta-se difícil, porém o segundo caso presta-se à aplicação das metodologias ligadas ao campo da previsão, permitindo, por exemplo, a utilização de uma estratégia de manutenção programada.

Sempre que seja possível, a previsão de avarias é uma estratégia que permite uma intervenção antecipada, manual ou automática, que pode evitar, por exemplo, que uma máquina pare em plena laboração, poupando os custos associados à sua paragem. Outro aspecto, não menos importante, diz respeito à produção de produtos com defeitos, aspecto directamente relacionado com o funcionamento das máquinas que produzem os componentes das peças. Os efeitos resultantes do seu mau funcionamento podem ser bastante minimizados se for possível proceder-se a uma previsão antecipada de falhas.

A ideia de monitoração é aplicável a um diverso conjunto de situações. O seu papel é o de ajudar a identificar a ocorrência de variações não aleatórias de padrões, ou relações, existentes numa sucessão de valores representativa de uma determinada situação, podendo ser aplicada a todos os casos em que se deseje saber quando variações não aleatórias ocorrem na sucessão. O objectivo da utilização das técnicas de monitoração prende-se com o facto de se tentar descobrir o mais cedo possível quando um processo se está a desviar de uma forma não aleatória do seu comportamento habitual.

Dependendo dos objectivos a atingir, um sistema para monitoração pode englobar, ou não, o refinamento automático dos mecanismos de decisão. Esta possibilidade realça o facto de que a validação e acumulação estatística são funcionalidades que devem ser incorporadas nos sistemas com carácter evolutivo que se venham a desenvolver. Outro factor também relacionado com este aspecto diz respeito à incorporação da aprendizagem artificial, baseando-se esta na síntese de informações já acumuladas [20].

A monitoração permite então determinar se o estado actual da máquina está ou não dentro dos parâmetros aceitáveis. Com base no estado actual e no estudo do passado de uma máquina, pode-se fazer prognóstico sobre o seu estado provável para um determinado intervalo de tempo futuro. Dois resultados imediatos são esperados: (1) obter controlo de qualidade em tempo real e (2) otimizar os períodos de manutenção da máquina, minimizando paragens não programadas.

Assim, uma arquitectura integrada que represente um sistema de monitoração pode ser vista como sendo constituída pela cooperação de vários subsistemas especializados comandados por um supervisor, ou então, como um sistema monolítico baseado num modelo único [15]. Estes dois conceitos extremos representam linhas de referência importantes que não comprometem outras soluções intermédias.

Duas razões podem ser apontadas para se querer modelar sucessões cronológicas como suporte à tarefa de monitoração. A primeira é disponibilizar uma descrição da sucessão em

termos dos seus componentes de interesse (i.e., comportamento da sucessão). A outra relaciona-se com a utilização do modelo da sucessão cronológica, utilizado para descrever o comportamento da sucessão, como suporte para a previsão de observações futuras e monitoração das actuais [8].

Grande parte da construção de modelos de sucessões cronológicas, utilizando métodos estatísticos, está baseada na teoria dos processos estocásticos estacionários. A maioria destes processos pode ser modelada a partir de um elemento da classe dos modelos denominados "AutoRegressivos Média Móvel" (AutoRegressive Moving Average - ARMA) [8]. Uma estratégia para selecção do modelo mais adequado para uma sucessão que pertença a uma destas duas classes foi desenvolvido por Box e Jenkins [4].

O trabalho apresentado nesta comunicação, debruçou-se fundamentalmente sobre a automatização de alguns dos aspectos relacionados com a qualidade de conformidade, particularmente sobre aqueles que se aplicam aos SFP. Para a sua implementação, experimentação e demonstração, escolheram-se as máquinas ferramenta CN como as entidades produtoras envolvidas no processo de produção.

O objectivo principal deste trabalho foi avaliar a utilização de processos estocásticos no apoio à manutenção preditiva para ambientes industriais. A previsão de avarias, através de monitoração e identificação dos estados de funcionamento de máquinas CN, será o suporte para a obtenção dos objectivos. Para a realização deste trabalho utilizou-se um dos laboratórios do Centro de Robótica Inteligente (CRI) do UNINOVA - Instituto de Desenvolvimento de Novas Tecnologias, compreendendo um torno e uma fresa CN.

2. A informação sensorial na supervisão

Uma das tarefas dos operadores de máquinas é observarem as tendências de desempenho do sistema e fazerem as modificações que considerem necessárias com o objectivo de evitarem a ocorrência de falhas. Com o objectivo de os ajudar, sensores podem ser instalados sobre as máquinas ferramentas para detectarem a quebra de uma ferramenta, embora sejam poucos aqueles que são suficientemente sensíveis para detectar a gradual redução do desempenho da máquina (e.g., gasto continuado de uma ferramenta). Isto significa que, na maior parte das situações, tem-se informação "post mortem" que não evita a paragem da máquina, o dano no componente parcialmente maquinado, e o tempo perdido na reinicialização da mesma, com todos os custos associados.

A abordagem que tem sido seguida para efeitos de monitoração tem por base a adopção de um conjunto de estados conhecidos em que uma máquina pode funcionar. A informação sensorial permite, desta forma, caracterizar individualmente cada um desses estados por meio de um conjunto de valores que os identificam de uma forma unívoca, i.e., por meio da sua assinatura sensorial. O reconhecimento de uma assinatura sensorial, acompanhada da informação sobre o estado em que a máquina deve estar a funcionar, permite a monitoração da máquina. Uma vez identificados os sensores a usar para a caracterização do estado de

funcionamento da máquina, tem que se reconhecer quais são os seus padrões típicos para posterior referência.

A informação sensorial torna-se assim no ponto de partida da supervisão. Como tal, a aprendizagem baseada em informação sensorial, assim como a caracterização dos vários processos participantes, serão o suporte dos mecanismos de decisão utilizados.

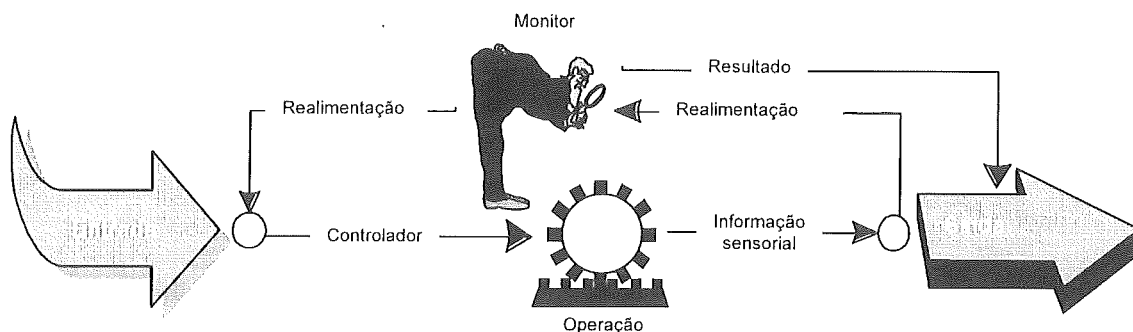


Figura 1 - Relação entre supervisor, operação, entrada e saída de um sistema de produção

A figura 1 apresenta a relação existente entre supervisor, operação, entrada e saída de um sistema de produção. Esta figura mostra que uma operação é conduzida pelo seu controlador, disponibilizando informação sobre o seu estado através de sensores acoplados à máquina que realiza a operação. O supervisor, através da análise da informação sensorial, comunica o resultado da sua supervisão ao exterior e, se conveniente, actua sobre o controlador, corrigindo-o.

2.1 Supervisão e monitoração

A aplicação do termo monitor sub-entende uma certa preocupação sobre um conjunto de objectivos a atingir. Em situações mais complexas, surge a necessidade de se considerar o monitor como sendo composto por um conjunto de Monitores Específicos (ME), em que os seus resultados são sintetizados pela entidade usualmente designada por supervisor [1] [2].

2.1.1 Arquitectura de um monitor específico

A figura 2 apresenta uma arquitectura de um ME, compreendendo sensores, o processo de aquisição de dados sensoriais, a base de conhecimento e o programa que permite efectuar a monitoração. A estrutura do ME é concebida, fundamentalmente, com base numa estratégia, sendo o seu comportamento programado de acordo com a tarefa, i.e., conjunto de acções em execução. A possibilidade de programação do ME destina-se a permitir a aplicação do conhecimento adquirido no controlo do seu comportamento.

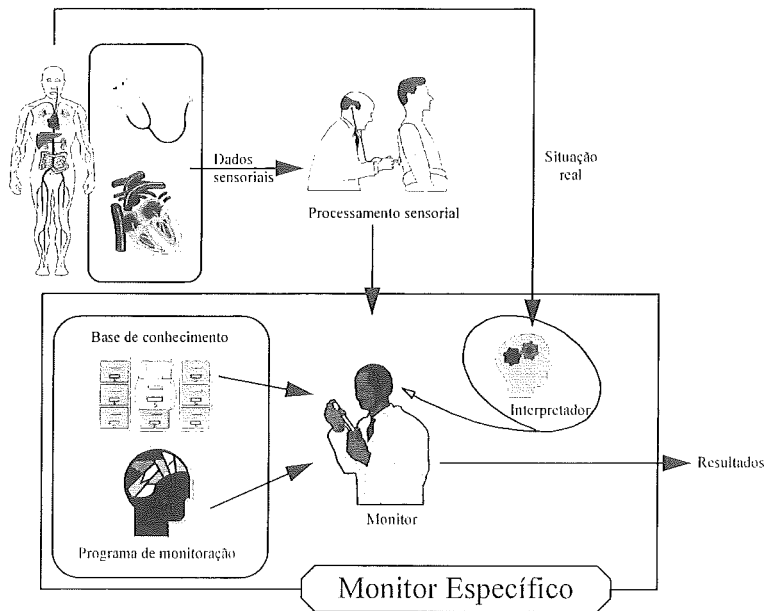


Figura 2 - Arquitectura de um Monitor Específico (ME)

Em relação às ocorrências que desconhece, o ME pode tomar uma de duas atitudes: (1) não fazer nada ou (2) proceder à aquisição dos dados sensoriais para efectuar um estudo posterior e adquirir conhecimento sobre a nova situação. Basicamente, um ME comporta funções de recolha de dados sensoriais e de identificação e verificação de situações, fornecendo resultados ao exterior.

O modelo da figura 2 considera somente a função executora dos ME, não incluindo a função de aprendizagem ou treino do mesmo. Considera-se neste modelo que, por razões de eficiência, a base de conhecimento é construída previamente e disponibilizada conjuntamente com o programa de monitoração [14] [22].

Os ME devem saber identificar novas ocorrências afim de obterem a informação necessária à sua aprendizagem. A implementação do subsistema do monitor relacionado com a identificação pode ser baseado, por exemplo, em redes neuronais, árvores de decisão, classificadores de Bayes ou análise de séries cronológicas [5]. Não sabendo à partida qual das técnicas é a mais adequada na monitoração da situação, uma solução é experimentar vários métodos, seleccionando-se o que apresenta melhor desempenho.

A implementação de um ME segue as seguintes fases:

1. Escolher o conjunto de situações que se pretende monitorar.
2. Seleccionar, consoante a especificidade do objectivo a ser monitorado, os métodos de processamento e os sensores para aquisição de dados.
3. Submeter, após a construção de um monitor que unicamente sabe recolher dados sensoriais, os dados recolhidos aos diferentes métodos de identificação e elege o melhor. Desta forma fica identificada a estrutura do ME.
4. Sempre que sejam detectadas novas situações, saltar para a fase 3 considerando a nova ocorrência.

2.1.2 Arquitectura de um sistema para monitoração, apoiado num supervisor

A figura 3 apresenta a arquitectura completa de um sistema para monitoração, apoiado num supervisor. Esta arquitectura compreende um conjunto de MEs, o agente executor de acções e o módulo supervisor. Os MEs têm a possibilidade de enviar directamente comandos de emergência ao agente executor, para que, ao detectarem situações perigosas (e.g. eminência de ruptura), possam avisar de uma forma mais rápida o agente executor. Garantir sincronização entre os vários MEs e a acção em execução, e permitir ao supervisor tomar o controlo do agente executor para efectuar testes ou correcções de recuperação sobre ele, é fundamental para o bom desempenho de um sistema para monitoração apoiado num supervisor.

Para efeitos de implementação, o supervisor é um Sistema Pericial (SP), dependente dos resultados dos vários MEs, que deve ser desenvolvido utilizando uma linguagem que permita, de uma forma facilitada, tratar o conhecimento de um especialista na área (e.g., operador de máquina). Exemplos destas linguagens são o Prolog e as aplicações vocacionadas para o desenvolvimento de sistemas baseados em conhecimento, vulgarmente conhecidas como "Expert System Shell".

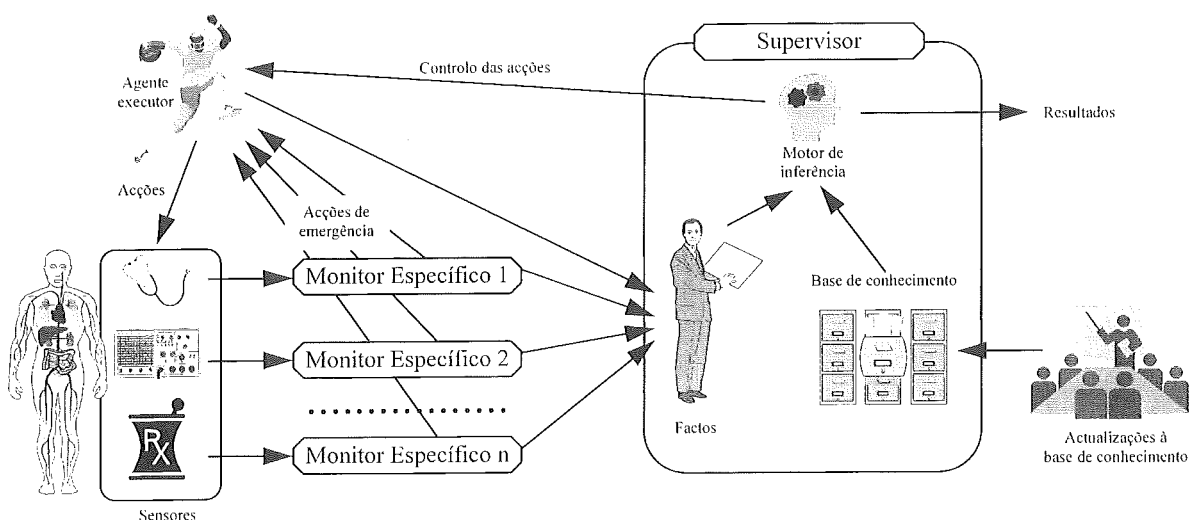


Figura 3 - Arquitectura completa de um sistema para monitoração, apoiado num supervisor

A fusão da informação sensorial é feita de acordo com as regras existentes na base de conhecimento, organizadas, por exemplo, em enquadramentos (o termo original em inglês é "frame", surgindo na bibliografia portuguesa como enquadramento) [19]. A sua estrutura permite armazenar objectos e factos sobre uma dada situação, assim como as acções que descrevem o procedimento perante a mesma, i.e., permite o conhecimento baseado em factos, regras e procedimentos [2]. A herança dinâmica, ou em tempo de execução, é outra característica importante dos enquadramentos porque, quando na base de conhecimento existem objectos derivados de outros também aí existentes, evita redundância, facilita a integridade e realça a sua estrutura.

Porque a aprendizagem artificial pode ser usada como meio para aperfeiçoamento e refinamento contínuo de um sistema pericial, o supervisor, para além de suportar as tarefas programadas de monitoração, diagnóstico e prognóstico, também recolhe informação sobre as situações ocorridas. Isto permite-lhe efectuar uma análise posterior utilizando ferramentas próprias e obter, assim, um maior conhecimento sobre o processo. "Data Base Learning", "Explanation-Based Learning", "Inductive Machine Learning" e "Integrated Machine Learning" são paradigmas de aprendizagem aplicáveis nestes casos [3] [18] [22].

3. Análise experimental

3.1. Descrição dos componentes do método utilizado

Pode-se entender o método utilizado no desenvolvimento deste trabalho como constituído pelos componentes (processos) identificados na figura 4.

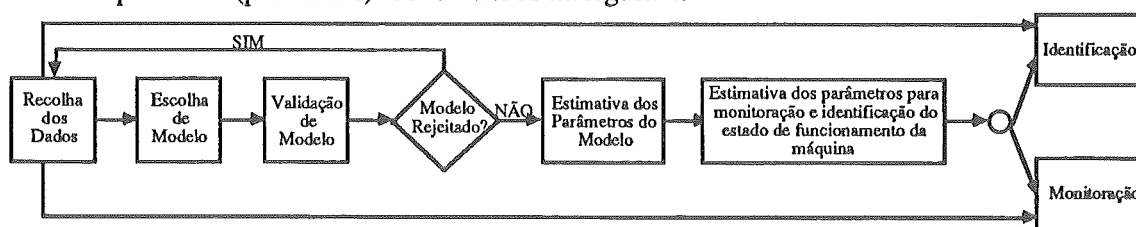


Figura 4 - Componentes do método utilizado

Cada um dos processos tem um conjunto de acções bem definido, tendo sido acoplados de uma forma sequencial, possibilitando que os resultados de uns possam ser a informação de entrada dos seguintes. A descrição que a seguir se apresenta permite ter uma visão sintética, i.e., de alto nível, do método utilizado no desenvolvimento do trabalho, assim como a forma como os seus componentes se inter-relacionam.

3.1 Recolha de dados

A primeira fase do método consiste na recolha dos dados através dos sensores instalados na máquina ferramenta CN. Esta recolha sensorial permite disponibilizar as sucessões cronológicas numa forma que servirá de base ao estudo para determinar os modelos. A recolha é efectuada pondo a máquina a funcionar em cada um dos estados que se pretende estudar, recolhendo-se para cada um deles, através dos sensores instalados, sucessões de dimensão 512.

Os estados de funcionamento da máquina que foram considerados para estudo foram: máquina a rodar em vazio a 100, 300, 500, 700 e 900 Rotações Por Minuto (RPM) e a cortar uma peça de referência (para o efeito utilizou-se um cilindro de latão).

3.1.2 Escolha do modelo

Utilizando a informação sensorial armazenada na forma de sucessão cronológica, faz-se uma análise no sentido de se determinar o modelo mais adequado para cada uma das situações de funcionamento da máquina em estudo. A metodologia seguida foi a sugerida por Box-Jenkins [4] [10] [13], em que, como forma de tentar automatizar o mais possível o processo de

identificação, se considerou um conjunto pre-estabelecido de modelos ARIMA admissíveis, escolhendo-se o mais adequado segundo esta metodologia.

A escolha do modelo mais adequado para representar um estado pré-definido de funcionamento da máquina, está baseada num critério pre-estabelecido que considera alguns dos parâmetros estimados durante a análise de cada um dos possíveis modelos. Os parâmetros considerados para o critério de escolha do modelo foram: (1) o valor da estatística Q; (2) a soma dos quadrados dos desvios; (3) o desvio padrão do erro; (4) o resultado da função da verosimilhança da estimativa dos coeficientes do modelo; (5) AIC - Akaike's Information Criterion; (6) SBC - Schwartz's Bayesian Criterion, (7) significância dos coeficientes estimados; e (8) a correlação existente entre os coeficientes do modelo.

O processo de selecção está dividido em três fases sequenciais:

Fase 1. "Rejeição" - Nesta fase rejeitam-se todos os modelos que não são considerados elegíveis, i.e., em que a sucessão dos desvios não é ruído branco, ou em que os coeficientes estimados são correlacionados ou não significativos.

Fase 2. "Seleção de modelos não dominados" - Na fase 2 a rejeição de modelos é feita através de uma tomada de decisão multi-critério, baseada nas estatísticas dos parâmetros dos modelos que passaram a fase 1. Para o efeito, definiu-se uma função que, recebendo como argumento dois modelos, indica se algum dos dois é não dominado, i.e. todos os seus parâmetros são simultaneamente melhores que os do outro. Todos os modelos dominados são rejeitados.

Fase 3. "Escolha do melhor modelo" - A última fase consiste em escolher o melhor modelo, dentro dos não dominados. O critério utilizado para desempate é o da parcimónia (i.e.; selecciona-se o modelo que tiver menor d , e menor $(p+q)$).

De maneira a tornar mais eficiente a execução do procedimento "Escolha de modelo", evitando assim o estudo de todos os modelos do conjunto preestabelecido, pode-se efectuar uma pré-selecção elegendo apenas aqueles que têm maior probabilidade de serem aceites, utilizando para o efeito a análise de Tsay-Tian. Esta análise permitirá concluir quanto à expectativa de encontrar um modelo aceitável dentro do conjunto considerado.

3.1. Validação do modelo

Após seleccionado o modelo mais adequado, passa-se à etapa de validação do mesmo. Esta fase confronta os modelos escolhidos com outras sucessões recolhidas da mesma forma, mas noutra janela temporal. Para o efeito, utiliza-se nova informação recolhida pelos sensores em condições idênticas às iniciais, para cada um dos estados de funcionamento da máquina considerados.

Esta validação é feita através de uma análise estatística, o que permite concluir quanto à adequabilidade dos modelos identificados aplicados a outras sucessões, que representam exactamente o mesmo estado da máquina noutra janela temporal, utilizando para isso um nível de significância de 10%. No caso de se concluir que o modelo escolhido não modela de forma

consistente um determinado estado de funcionamento da máquina (modelo rejeitado), considera-se que a sucessão inicial recolhida não é representativa desse estado de funcionamento, voltando-se ao ponto inicial tantas vezes quantas as necessárias, até que se encontre um modelo que o seja.

3.1.4 Estimativa dos coeficientes do modelo

Após seleccionado o modelo mais adequado para um determinado estado de funcionamento da máquina, estimam-se os seus coeficientes. Esta estimativa é efectuada utilizando o conjunto de amostras representativas do estado de funcionamento da máquina, recolhidas nas mesmas condições das da fase de validação do modelo.

Para cada uma das amostras consideradas, calculam-se os coeficientes dos modelos, realizando-se de seguida uma análise estatística sobre cada conjunto de coeficientes estimados, tomando-se um valor que se considera conveniente. Para cada coeficiente, foram experimentados a média, a mediana e a moda da distribuição empírica dos dados agrupados, tendo sido a moda a que apresentou melhores resultados experimentais.

A figura 5 apresenta uma exemplo do estudo da distribuição empírica dos dados agrupados relativa ao coeficiente Φ_1 de um modelo AR(1).

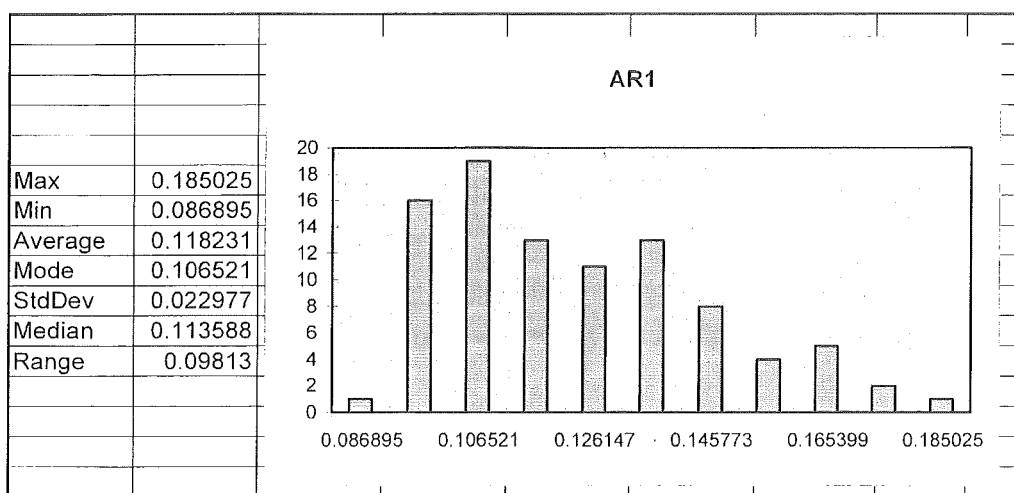


Figura 5 - Exemplo do estudo da distribuição dos coeficientes estimado de um modelo AR(1)

3.1.5 Estimativa dos parâmetros para monitoração e identificação do estado da máquina

A monitoração e a identificação do estado de funcionamento da máquina são suportadas por técnicas paramétricas. Nesta fase estimam-se os parâmetros que vão permitir efectuar estas tarefas. A estimativa destes parâmetros, baseada na informação disponibilizada pelas fases anteriores, segue uma aproximação análoga à utilizada na escolha e estimativa dos coeficientes dos modelos. Neste caso, utiliza-se também um conjunto de sucessões recolhidas em situações idênticas às iniciais, noutra janela temporal. A partir da análise dessas sucessões, calcula-se o valor dos parâmetros que melhor se adequam ao método de monitoração e identificação escolhidos.

3.1.6 Identificação

Utilizando os modelos determinados anteriormente, pode-se identificar de uma forma probabilística o estado de funcionamento da máquina. Esta identificação é feita confrontando os dados que estão a ser recolhidos no momento, com os previstos pelo modelo escolhido para representar essa situação. A análise estatística deste confronto permite que se apresente, de uma forma probabilística, qual o estado de funcionamento da máquina. A identificação do modelo, confrontada com o estado em que se pressupõe que a máquina esteja a funcionar, ajuda a detectar situações de anomalia.

3.1.7 Monitoração

Baseado no modelo do estado de funcionamento da máquina, obtido nas fases anteriores para cada um dos sensores, pode-se efectuar a monitoração do seu funcionamento. O objectivo da monitoração será o de acompanhar o funcionamento de um determinado estado da máquina, reconhecendo situações de tendência para defeito ou falha, como por exemplo, o desgaste ou a quebra de uma ferramenta. A monitoração é feita através do confronto entre os valores sensoriais recolhidos e os previstos. O resultado do processo de monitoração resulta da análise desse confronto.

4. Apresentação e análise dos resultados obtidos

Este capítulo apresenta e analisa os resultados mais importantes obtidos através da experimentação apresentada.

4.1 Sucessões recolhidas

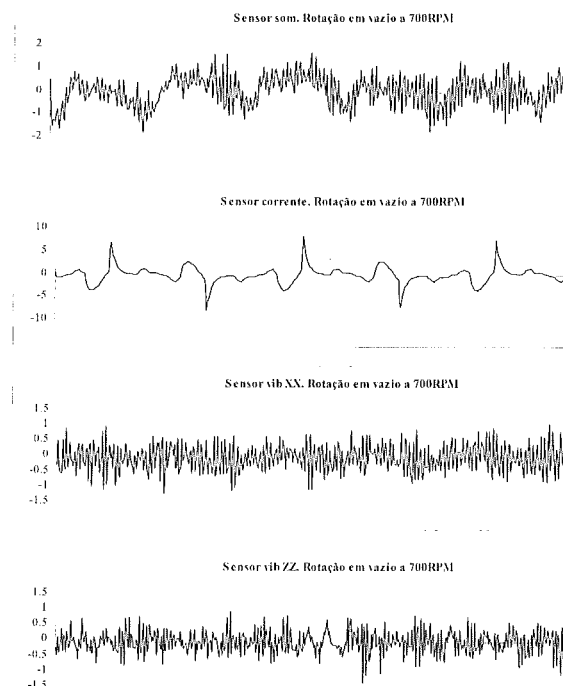


Figura 6 - Visualização de sucessões cronológicas características da máquina (*rotação a 700RPM em vazio*). Dados recolhidos através dos sensores instalados na máquina

A figura 6 apresenta um exemplo da visualização das sucessões cronológicas amostradas através de cada sensor, características para cada estado de funcionamento considerado. Cada sucessão foi recolhida durante 62.47 ms, com um tempo entre recolha de dois pontos consecutivos de 122 μ s. Esta frequência de amostragem esteve condicionada às características do "hardware" utilizado.

Da análise das sucessões recolhidas pelo sensor de som verifica-se que as sucessões apresentam um padrão semelhante nas baixas frequências. Na situação de *rotação em vazio* nota-se uma ligeira variação nas altas frequências, existindo contudo uma variação significativa na amplitude do sinal, crescente com o aumento da velocidade de rotação da máquina. Na situação de *corte*, as altas frequências apresentam um padrão significativamente distinto dos encontrados nas situações de *rotação em vazio*. O motivo porque o padrão nas baixas frequências se mantém constante, é que o microfone também capta o ruído de fundo existente no laboratório, que, neste caso, é dominado pelo ruído emitido (som grave) pelos dispositivos de ventilação. Quando a máquina é posta a rodar em vazio, ela produz, para todas as velocidades de rotação, um som agudo de frequência semelhante, que aumenta da intensidade com a velocidade. Na situação de *corte*, o som agudo muda de tom, tornando-se mais grave.

As sucessões recolhidas pelo sensor de corrente apresentam um padrão semelhante, existindo no entanto uma ligeira variação na amplitude do sinal, crescente com o aumento da velocidade de rotação da máquina e com a entrada em corte. O padrão mantém-se semelhante, uma vez que a forma de onda do sinal de corrente que alimenta a máquina é sempre o mesmo para todas as situações. Com o aumento de velocidade de rotação da máquina, o motor vai consumir mais energia, provocando um aumento na amplitude do sinal. O mesmo acontece quando a máquina entra em corte.

Cada uma das sucessões recolhidas pelo sensor de vibração segundo o eixo dos XX apresenta um padrão distinto, existindo igualmente uma variação significativa na amplitude do sinal, crescente com o aumento da velocidade de rotação da máquina e com a entrada em corte. Este facto verifica-se porque as vibrações segundo o eixo dos XX, resultantes da execução de um comando, são distintas, aumentando de intensidade com o aumento de velocidade e com a entrada em corte.

As conclusões que se podem retirar, através da análise das sucessões recolhidas através do sensor de vibração segundo o eixo dos ZZ, são semelhantes às da análise das sucessões recolhidas pelo outro sensor de vibração.

4.2 Modelos seleccionados

Os modelos seleccionados para representar cada uma das situações consideradas para estudo do funcionamento da máquina ferramenta, encontram-se descritos no quadro 1. Para cada situação, apresenta-se a ordem dos parâmetros de cada modelo ARIMA(p,d,q), a transformação aplicada, assim como o valor estimado dos respectivos coeficientes, após ter sido retirada a periodicidade, quando significativa (e.g., som de fundo de baixa frequência

resultante do funcionamento do sistema de ventilação do laboratório, a funcionar em situação estável; ou do ciclo de 20ms correspondente à corrente alterna de alimentação do motor da máquina, reflectido no sensor de corrente).

Analisando os resultados obtidos, verifica-se que os modelos correspondentes ao sensor de corrente não têm nenhuma transformação aplicada, necessitando todos de uma integração de primeiro grau e sendo a ordem dos parâmetros p e q inferiores ou iguais a um.

Os modelos correspondentes aos outros três sensores têm a transformação Box-Cox aplicada, com valor de $\lambda = 0$ (equivalente à transformação log) e $\lambda = -1$. Embora na descrição destes modelos não esteja explicitamente referida a integração das sucessões, ela existe. É a proximidade de 1 do primeiro coeficiente autoregressivo, AR1, que mostra que um modelo equivalente seria o dado pelo $AR1 = 0$ e $d = 1$, mantendo-se os outros coeficientes com os mesmos valores.

Sensor	CORRENTE						SOM					
	100RPM	300RPM	500RPM	700RPM	900RPM	Corte	100RPM	300RPM	500RPM	700RPM	900RPM	Corte
Comando CNC												
ARIMA(p,d,q)	(1,1,0)	(0,1,1)	(0,1,1)	(0,1,1)	(1,1,0)	(0,1,1)	(1,0,1)	(0,0,1)	(1,0,1)	(1,0,3)	(1,0,3)	(1,0,1)
Transformação	-	-	-	-	-	-	LOG	LOG	LOG	LOG	BC(l=-1)	BC(l=-1)
AR1	0.107	0	0	0	0.165	0	1	1	1	1	1	0.998
AR2	0	0	0	0	0	0	0	0	0	0	0	0
AR3	0	0	0	0	0	0	0	0	0	0	0	0
MA1	0	-0.104	-0.121	-0.128	0	-0.115	-0.331	-0.688	-0.737	0.528	0.782	-0.54
MA2	0	0	0	0	0	0	0	0	0	0.643	0.545	0
MA3	0	0	0	0	0	0	0	0	0	-0.537	-0.658	0

Sensor	Vibração X						Vibração Y					
	100RPM	300RPM	500RPM	700RPM	900RPM	Corte	100RPM	300RPM	500RPM	700RPM	900RPM	Corte
Comando CNC												
ARIMA(p,d,q)	(1,0,1)	(1,0,3)	(1,0,3)	(1,0,3)	(1,0,3)	(1,0,0)	(1,0,1)	(1,0,1)	(1,0,1)	(1,0,0)	(1,0,0)	(1,0,3)
Transformação	LOG	LOG	LOG	LOG	LOG	LOG	LOG	LOG	LOG	BC(l=-1)	BC(l=-1)	BC(l=-1)
AR1	1	1	1	1	1	1	1	1	1	1	1	1
AR2	0	0	0	0	0	0	0	0	0	0	0	0
AR3	0	0	0	0	0	0	0	0	0	0	0	0
MA1	0.242	0.62	0.63	0.387	0.681	0	0.334	0.449	0.808	0	0	0.696
MA2	0	0.309	0.581	0.753	0.64	0	0	0	0	0	0	0.501
MA3	0	-0.247	-0.454	-0.514	-0.608	0	0	0	0	0	0	-0.53

Quadro 1 - Modelos ARIMA(p,d,q) seleccionados. Apresentação por sensor e comando executado

4.3 Verificação dos modelos seleccionados

Para verificação dos modelos seleccionados, efectuou-se um estudo aplicando-os a outras sucessões recolhidas através dos mesmos sensores em condições semelhantes, mas noutra janela temporal.

Sensor	CORRENTE						SOM					
Comando CNC	100RPM	300RPM	500RPM	700RPM	900RPM	Corte	100RPM	300RPM	500RPM	700RPM	900RPM	Corte
NãoRejeitados(%)	93	95	100	97	99	98	100	100	99	100	97	95

Sensor	Vibração X						Vibração Y					
Comando CNC	100RPM	300RPM	500RPM	700RPM	900RPM	Corte	100RPM	300RPM	500RPM	700RPM	900RPM	Corte
NãoRejeitados(%)	93	100	100	98	95	100	99	100	91	100	100	100

Quadro 2 - Percentagem de sucessões não rejeitadas por um modelo ARIMA, correspondente a um sensor e um comando CN

Os resultados obtidos, e que se apresentam no quadro 2, mostram que, com um nível de significância de 90%, se pode afirmar que estes modelos são adequados para modelar cada um dos estados de funcionamento da máquina a partir da informação recolhida por esses sensores.

4.4 Monitoração

A monitoração do estado de funcionamento da máquina é efectuada tendo por base os modelos seleccionados. Para os casos estudados, o sistema de monitoração apresentou um nível de significância da ordem dos 97% na indicação de bom funcionamento quando a máquina estava a funcionar correctamente, e de 90% na indicação de anomalia, quando a máquina se encontrava com problemas (ferramenta de corte com o gume fracturado).

4.5 Identificação automática de comandos CN

Um novo aspecto, estudado no decurso deste trabalho, foi o da identificação automática do comando CN que está a ser executado pela máquina, a partir da recolha sensorial.

4.5.1 Identificação aplicando os critérios A e B

Foram utilizados os seguintes dois critérios no processo de identificação: A) menor erro absoluto médio, e B) não rejeição do desvio como ruído branco.

O critério A identifica um comando aplicando cada um dos modelos seleccionados sobre a sucessão recolhida para cada um dos sensores. A identificação é feita considerando o modelo que apresentou menor erro absoluto médio.

Para cada um dos sensores, o critério B identifica um comando aplicando cada um dos modelos seleccionados sobre a sucessão recolhida. A identificação é feita considerando o modelo que apresentou maior significância em não rejeitar o desvio como ruído branco.

Para cada um dos critérios A e B, as probabilidades de se acertar na identificação de um comando, dado que se está na presença de uma sucessão pré determinada, são apresentadas no quadro 3.

Probabilidade[Acertar utilizando sensor|presença da serie] Probabilidade[Acertar utilizando sensor|presença da serie]

Sensor	Comando					
	100RPM	300RPM	500RPM	700RPM	900RPM	Corte
Corrente	20	36	12	49	74	7
Som	0	0	0	0	100	58
VibraçãoX	0	0	0	0	0	100
VibraçãoZ	0	0	0	0	0	100

Sensor	Comando					
	100RPM	300RPM	500RPM	700RPM	900RPM	Corte
Corrente	0	0	0	0	100	0
Som	12	0	0	42	97	3
VibraçãoX	44	79	67	3	100	0
VibraçãoZ	6	48	69	0	0	100

Quadro 3 - Probabilidade de identificação correcta por um sensor, estando na presença de uma sucessão pré determinada. Critério: A) Menor erro absoluto médio, e B) não rejeição do desvio como ruído branco.

Esta análise permite chegar aos resultados apresentados no quadro 4 que indicam a probabilidade de se acertar na identificação de um comando, segundo cada um dos critérios A e B. Examinando este quadro, observa-se que o desempenho dos sensores de corrente e de som na identificação é melhor quando utilizado o critério do menor erro absoluto médio, enquanto que o desempenho dos sensores de vibração é melhor quando utilizado o critério de não rejeição do resíduo como ruído branco.

Sensor	Critério: Menor Erro Absoluto Médio	Critério: Resíduo como ruído branco
	Probabilidade [acertar utilizando o sensor]	Probabilidade [acertar utilizando o sensor]
Corrente	33.00%	16.67%
Som	26.33%	25.67%
VibraçãoX	16.67%	48.83%
VibraçãoZ	16.67%	32.08%

Quadro 4 - Probabilidade de identificação correcta por um sensor, segundo os critérios A e B

Por esse motivo, para efeitos de identificação utilizou-se um novo critério baseado nos resultados da aplicação dos critérios A e B. Assim, aplicou-se o critério A sobre os dados recolhidos pelos sensores de corrente e som, com um desempenho de 33% e 26.33% respectivamente, e o critério B sobre os dados recolhidos pelos sensores de vibração segundo o eixo dos XX e dos ZZ, com um desempenho de 48.83% e 32.08% respectivamente. A apreciação destes resultados, leva a que se possa associar um nível de credibilidade a cada sensor, com base na identificação de comandos utilizando o melhor critério. O quadro 5 apresenta, em forma de quadro, estes resultados.

Sensor	Probabilidade [acertar utilizando o sensor]	Critério	Credibilidade
Corrente	33.00%	Menor Erro Absoluto Médio	0.22
Som	26.33%	Menor Erro Absoluto Médio	0.18
VibraçãoX	48.83%	Resíduo como ruído branco	0.34
VibraçãoZ	32.08%	Resíduo como ruído branco	0.26
			Total = 1

Quadro 5 - Credibilidade de cada sensor, quando lhe é aplicado o melhor critério

4.5.2 Identificação aplicando o critério C

Utilizando a credibilidade de cada sensor, como descrito anteriormente, definiu-se um novo critério (critério C) que permite a identificação de um comando recorrendo às conclusões

apresentadas pelo quadro 5. Este critério consiste em somar as credibilidades de cada sensor para cada comando identificado, escolhendo-se aquele que corresponde a um maior total.

O quadro 6 exemplifica a aplicação do critério C. Nele, o sensor de corrente, utilizando o critério de menor erro absoluto médio, identificou o comando 100RPM, o sensor de som, utilizando o mesmo critério, identificou o comando 300RPM, o sensor de vibração segundo o eixo dos XX, utilizando o critério de não rejeição do resíduo como ruído branco, identificou o comando 100RPM e o sensor de vibração segundo o eixo dos ZZ, utilizando o mesmo critério, identificou o comando 700RPM. Como totais, o comando 100RPM tem uma credibilidade de 56%, o comando 300RPM de 18%, o comando 700RPM de 26%, e todos os outros comando têm credibilidade nula. Assim, segundo este critério, seria identificado o comando 100RPM (comando com maior credibilidade).

Credibilidade	Sensor	Comando					
		100RPM	300RPM	500RPM	700RPM	900RPM	corte
0.22	Corrente	0.22					
0.18	Som		0.18				
0.34	VibX	0.34					
0.26	VibZ				0.26		
	Total	0.56	0.18	0	0.26	0	0

Quadro 6 - Exemplificação da aplicação do critério C

4.5.3 Propriedade de absorção detectada entre modelos

Um exame aos resultados obtidos referentes à identificação pelos diferentes sensores permite concluir que os modelos correspondentes ao comando 900RPM têm fortes características de absorção, nomeadamente sobre os do comando 700RPM. A mesma situação acontece, embora de uma forma não tão significativa, para os pares de modelos (100RPM, 300RPM), (500RPM, 900RPM), e (corte, 300RPM). Uma forma de melhorar o processo de identificação será encontrar outros modelos para estes comandos, para o qual esta propriedade não se verifique.

4.5.4 Estudo de diferentes cenários

Para teste do identificador, utilizando o critério C, fez-se um estudo considerando os seguintes cenários:

1. Utilização de todos os comandos CN, e todos os sensores.
2. Utilização de todos os comandos CN excepto o 700RPM, e todos os sensores.
3. Utilização de todos os comandos CN excepto o 900RPM, e todos os sensores.
4. Utilização de todos os comandos CN, e todos os sensores excepto o sensor do som.
5. Utilização de todos os comandos CN excepto o 700RPM, e todos os sensores excepto o sensor do som.

6. Utilização de todos os comandos CN excepto o 900RPM, e todos os sensores excepto o sensor do som.

De notar que se escolheram os comandos 700RPM e 900RPM, por serem aqueles que apresentavam uma maior característica de absorção, e o sensor de som por ser o que manifesta uma menor credibilidade.

O quadro 7 apresenta os resultados obtidos.

A análise deste quadro permite concluir que, embora, à primeira vista, o sensor de som manifeste menor credibilidade, a sua utilização é importante uma vez que aumenta a probabilidade de acertar na identificação de um comando. Ela permite ainda concluir que deixar de considerar um dos comandos que está em conflito pela propriedade da absorção (700RPM e 900RPM), aumenta a eficácia do identificador.

	Comandos		
	Todos	Sem 700RPM	Sem 900RPM
P[Identificação correcta, considerando todos os sensores]	56.00%	66.00%	69.00%
P[Identificação correcta, não considerando o sensor de som]	49.00%	58.00%	56.00%

Quadro 7 - Probabilidade de identificação correcta de um comando, considerando diferentes cenários

Assim, para efeitos de identificação, devem ser utilizados todos os sensores e deixar de considerar um dos comandos 700RPM ou 900RPM ou, por forma a tentar eliminar a propriedade de absorção existente, procurar encontrar outros modelos para estes comandos.

5. Conclusões e perspectivas de trabalho futuro

Este capítulo apresenta as conclusões mais relevantes, obtidas durante a realização do trabalho exposto, bem como as perspectivas de trabalho futuro que lhe dará continuidade.

5.1 Conclusões

Os resultados obtidos pela utilização do protótipo desenvolvido e apresentado nesta comunicação permitem concluir que, para os cenários considerados, a sua utilização como sistema para monitoração de máquinas ferramenta CN, utilizando processos estocásticos no apoio à manutenção preditiva para ambientes industriais, se apresenta adequada.

O método utilizado no desenvolvimento da análise experimental apresentou resultados animadores, embora a sua utilização seja morosa, mesmo após automatizada a maior parte das suas tarefas. A aplicação da metodologia Box-Jenkins, considerando somente a utilização de modelos ARIMA, conjuntamente com as técnicas para monitoração apresentadas, permitiram chegar a resultados encorajadores.

O facto do identificador apresentar índices de reconhecimento da ordem dos 60% pode ser melhorado se, em tempo real, como é o caso, for possível efectuar várias classificações intermédias e decidir em função da mais significativa. O resultado obtido será de maior confiança quanto maior for o número de classificações intermédias.

É importante realçar ainda que o trabalho apresentado não se encontra completamente terminado, considerando o desenvolvimento que será necessário efectuar com o propósito de transformar o protótipo apresentado num sistema para ser utilizado em ambientes industriais reais, assim como o trabalho de investigação a desenvolver por forma à obtenção de sistemas para monitoração e supervisão mais inteligentes e adequados ao apoio à manutenção preditiva para esses ambientes.

5.2 Perspectivas de trabalho futuro

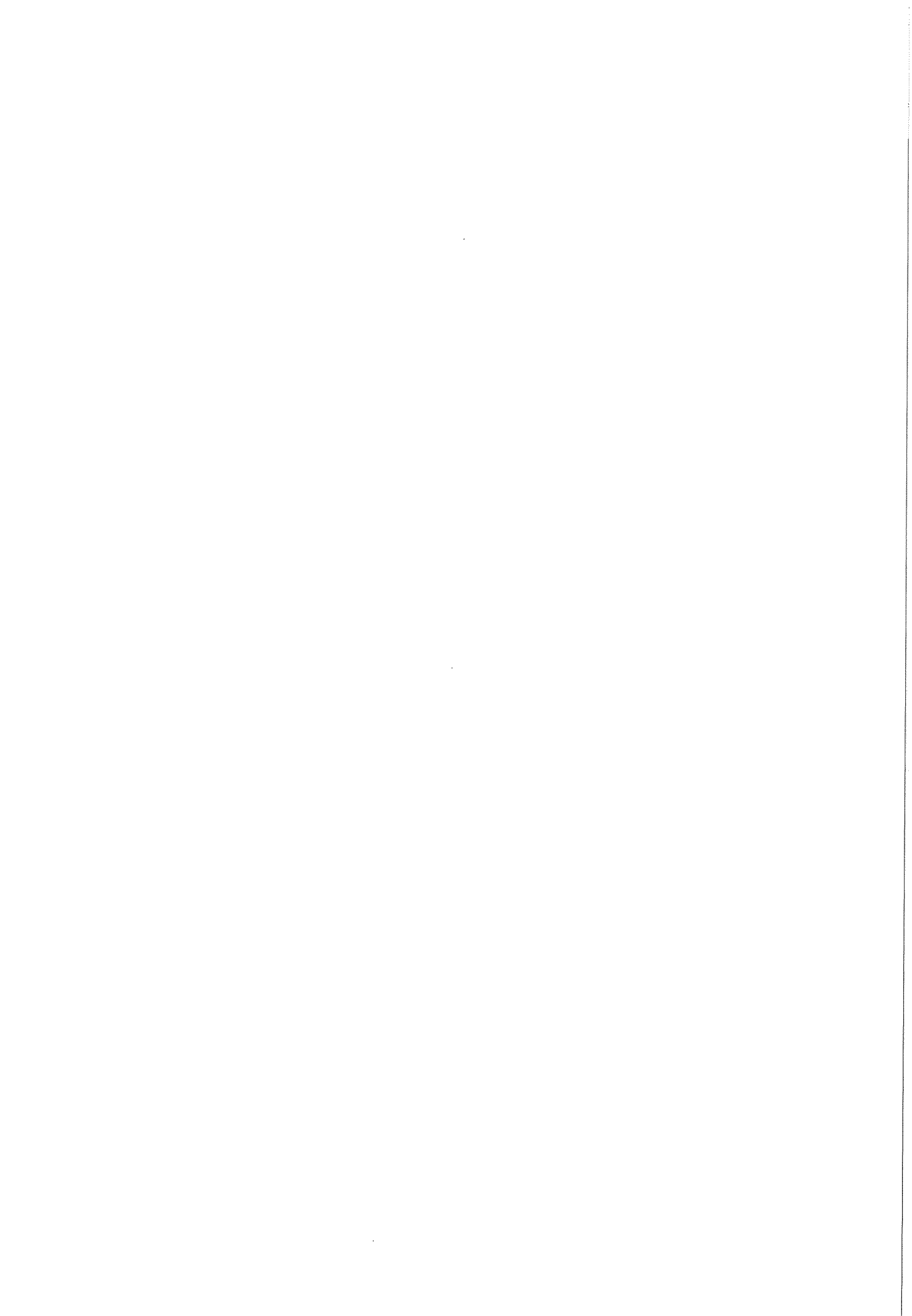
Perspectiva-se que o trabalho futuro se desenvolva segundo as seguintes linhas gerais:

- Refinar o método utilizado durante a análise experimental, na tentativa de se obterem melhores resultados de uma forma mais eficiente.
- Integrar outros sensores e proceder à aquisição de mais dados, utilizando não só a máquina ferramenta CN - torno STARTURN 4, mas também outros, como por exemplo a fresa que existe no laboratório, com o objectivo de verificar a consistência do método utilizado.
- Aplicar outras técnicas e metodologias, para além da modelação ARIMA seguindo a metodologia Box-Jenkins, como base para o desenvolvimento do sistema de monitoração. Neste contexto, a utilização de algoritmos genéticos [6], redes neuronais [9], análise espectral e diferentes técnicas de modelação baseada em métodos estatísticos [11] [16], entre outros a considerar, poderá ser avaliada, constituindo, sem dúvida, um importante trabalho de comparação de mecanismos para monitoração.
- Integrar o trabalho realizado na área de análise e concepção de monitores específicos, como por exemplo os de Dvijver [5], Pun [14] e Szélig [21] e no campo da inteligência artificial aplicada à aprendizagem em monitoração, diagnóstico e prognóstico, como por exemplo o de Wright [23], com o método apresentado, de maneira a poder reavaliar o protótipo do sistema de monitoração desenvolvido.
- Como perspectiva mais ambiciosa, mas talvez mais aliciante, aperfeiçoar o protótipo desenvolvido no sentido de se obter um sistema de monitoração apto a ser incorporado numa fábrica com produção integrada por computador, de preferência utilizando sistemas flexíveis de produção, permitindo a relevância do sistema proposto sobre um ambiente real.

6. Referências

- [1] Barata, M. and Steiger-Garção, A., Predictive Environment for Prognosis and Monitoring CNC Systems Support, 23rd International Symposium on Automotive Technology & Automation, Vienna Hofburg, Austria, 3-7 December (1990).
- [2] Barata, M., Rauber, T., et al., Caracterização Automática de Padrões em Sensores de Máquinas CNC, Workshop Iberoamericana de Sistemas Autónomos em Robótica e CIM, Lisboa 2-4 Novembro (1992); Pub. in Revista Robotica e Automatização 10 (Nov.92) pp.95-105.
- [3] Barata, M., Jardim-Gonçalves, R., Steiger-Garção, A. and Assis-Lpes, J., CNC machines monitoring using sensorial stochastic models, Balanced Automation Systems - Architectures and design methods (1995) pp.439-450, Chapman & Hall.

- [4] Box, G., and Jenkins, G., *Time Series Control. Forecasting and control*. Holden-Day, Library of Congress Catalog (1970).
- [5] Devijver, P. and Kittler, J., *Pattern Recognition: A Statistical Approach*, Prentice/Hall Int., London (1992).
- [6] Goldberg, D., *Genetic algorithms in search, optimization and machine learning*, Reading, MA. Addison Wesley (1989).
- [7] Gonçalves, R., *Processos estocásticos de apoio à manutenção preditiva para ambientes industriais*, Dissertação apresentada para obtenção de grau de mestre em Investigação Operacional e Engenharia de Sistemas pela UTL/IST (1996).
- [8] Harvey, A., *Forecasting, structural time series models and the Kalman filter*, Cambridge University Press (1992).
- [9] Lippmann, R., *An introduction to computing with neural nets*, IEEE ASSP Magazine (1987) pp. 4-22.
- [10] Makridakis, S. and Wheel Wright, S., *Forecasting methods for management*, Fifth edition, John Wiley & Sons, Inc. (1989).
- [11] Michie, D. et al., *Machine Learning, neural and statistical classification*, Ellis Horwood Limited (1994).
- [12] Montgomery, D., *Introduction to Statistical Quality Control*, Second edition, John Wiley & Sons (1990).
- [13] Murteira, B., Müller, D. e Turkman, K., *Análise de Sucessões Cronológicas*, Editora McGraw-Hill de Portugal, Lda. (1993a).
- [14] Pun, L., Archimeda, B. et al., *Intelligent learning aid for an intelligent FMS-monitoring process*, in *Computers in industry* 17 (1991) pp. 225-235.
- [15] Rauber, T., Barata, M. e Steiger-Garção, A., *Ambiente para Análise e Visualização de Informação Sensorial em Supervisão de Sistemas*, Actas das 4 Jornadas Nacionais de PPP/AC, Portugal-Lisboa, Maio 17-19 (1993).
- [16] Rauber, T., *Inductive pattern classification - methods - features - sensors*, Dissertação para obtenção do grau de doutor em Engenharia Electrotécnica (1995).
- [17] Ryabinin, I., *Reliability of engineering systems - Principles and analysis*, MIR Publishers, Moscow (1976).
- [18] Sammon, Jr., *A nonlinear mapping for data structure analysis*, IEEE Trans. on Computers 18 (1969) pp.401-409.
- [19] Staugard, A., *Robotics and AI: An Introduction to Applied Machine Intelligence*, Prentice Hall (1987) pp.96-112.
- [20] Steiger-Garção, A., Barata, M., et al., *Integrated Environment for Prognosis and Monitoring Systems Support*, 1st UNIDO workshop on Robotics and Computer Integrated Manufacturing, Lisbon-Portugal (1989).
- [21] Szélig, K., Alpek, F. et al., *Automatic inspection in a CIM system*, in *Computers in Industry* 17 (1991) pp. 101-119.
- [22] Whitehall, B. and Lu, S., *Machine Learning in engineering automation-The present and the future*, in *Computers in Industry* 17 (1991) pp. 91-100.
- [23] Wright, P. and Bourne, D., *Manufacturing Intelligence*, Addison-Wesley Publishing Company (1988).



O SISTEMA DE INFORMAÇÃO PARA GESTÃO DO INEGI

António Amador

Instituto de Eng. Mecânica e Gestão Industrial
Rua dos Bragas
4099 Porto Codex - Portugal

João Falcão e Cunha

Faculdade de Eng. da Univ. do Porto
Rua dos Bragas
4099 Porto Codex - Portugal

José António Cabral

Faculdade de Eng. da Univ. do Porto
Rua dos Bragas
4099 Porto Codex - Portugal

Rui Sá

Instituto de Eng. Mecânica e Gestão Industrial
Rua dos Bragas
4099 Porto Codex - Portugal

Vasco Sanches e Sá

Faculdade de Eng. da Univ. do Porto
Rua dos Bragas
4099 Porto Codex - Portugal

Abstract

This paper presents SIGEI, the Management Information System of INEGI. SIGEI supports users in managing the main information processes at INEGI, in an integrated and distributed way. SIGEI has been designed and developed using robust methods and tools, in particular OMT and Access. Particular care was taken in ensuring that the system would be easily adapted to new technologies, and to the skill of its users.

SIGEI has been identified by a strategic study carried out in INEGI. It has been closely followed by a working group including senior INEGI management. The development process followed the waterfall and the prototyping software engineering models. The article described SIGEI's sub-systems, namely Contact Management, Proposal Management, Project Management, it describes the software development method, it provides examples of the conceptual models and user interfaces, and it finishes by indicating expected developments of SIGEI.

Resumo

Apresenta-se neste trabalho o SIGEI, Sistema de Informação para Gestão do INEGI. O SIGEI oferece apoio aos utilizadores nas principais funções de gestão de informação existente no INEGI, de uma forma integrada e distribuída. O SIGEI foi concebido e desenvolvido com o auxílio de metodologias e ferramentas informáticas robustas, em particular o OMT e o Access. Houve um cuidado particular em assegurar a adaptação a evoluções de tecnologia e a adequação aos utilizadores.

O SIGEI surgiu de uma necessidade identificada no estudo estratégico do INEGI, tendo sido a sua construção e implantação acompanhada por um grupo de trabalho ao mais alto nível do INEGI. O processo de

desenvolvimento da aplicação faseou-se pelo modelo cascata e utilizou prototipificação na construção dos vários módulos. O artigo descreve globalmente os sub-sistemas do SIGEI, gestão de contactos, gestão de propostas e gestão de projectos, o método de desenvolvimento utilizado, apresenta exemplos dos modelos conceptuais e das interfaces gráficas com os utilizadores e conclui com as necessidades e possibilidades actuais de evolução.

Keywords

Information system, OMT, Access.

1. Introdução

O INEGI, Instituto de Engenharia Mecânica e Gestão Industrial, é um Instituto de Inovação e Transferência de Tecnologia, constituindo a interface de ligação entre o Departamento de Engenharia Mecânica e Gestão Industrial da Faculdade de Engenharia da Universidade do Porto e o meio empresarial. Foi criado em 1986, tendo o estatuto de associação sem fins lucrativos e de utilidade pública.

O INEGI é constituído por unidades operacionais, designadas por secções e centros, que são grupos de investigação, desenvolvimento e consultoria, com grande autonomia de gestão e dispersos geograficamente por instalações na Faculdade de Engenharia da Universidade do Porto e no Complexo Empresarial do Barroco (CEB).

A gestão destas unidades, e do INEGI globalmente, depende de um sistema de informação que foi evoluindo a partir de uma estrutura assente em documentos em papel e em procedimentos de contacto pessoal directo (Sá, 1993).

O apoio informático às actividades do INEGI é actualmente realizado por três aplicações mono-utilizador independentes entre si: contabilidade geral, contabilidade analítica por projecto e facturação. Esta situação deverá alterar-se com a introdução de um sistema integrado de apoio informático, o SIGEI, para as várias funções de apoio à gestão do INEGI.

1.1 História do SIGEI

O projecto de Acreditação dos Laboratórios do INEGI e a sua integração na Rede Nacional de Laboratórios Acreditados levou à realização do estudo do Sistema da Qualidade do INEGI (Cabral, 1994a). Na mesma altura realizou-se um Estudo de Diagnóstico e Análise Estratégica (EDAE) pela empresa Challenge (Challenge, 1994). Para que a acreditação fosse efectuada era necessário reconhecer ao INEGI "adequada capacidade técnica e organizativa" (Cabral, 1994a). Estes estudos tornaram evidente as limitações do sistema de informação existente.

Foi então proposta a definição de um Plano Estratégico para o INEGI (Cabral, 1994b) no qual se previa a necessidade de criar um sistema de informação de apoio à gestão que viabilizasse um melhor acompanhamento das actividades realizadas, disponibilizando indicadores de gestão, numa lógica distribuída de utilização pelas várias unidades operacionais.

Para tal foi constituído um grupo de trabalho, ao mais alto nível da instituição. Deste grupo faziam parte elementos da direcção do INEGI, da coordenação de algumas das suas unidades operacionais, bem como técnicos dos serviços centrais. Além disso o grupo integrava os

responsáveis pela concepção e futuro desenvolvimento do sistema. Este grupo efectuou uma análise detalhada dos processos de gestão de informação no INEGI e elaborou os requisitos para o Sistema de Informação para Gestão do INEGI. Embora não participando directamente nas fases seguintes, o grupo de trabalho mantém-se activamente empenhado no sucesso do SIGEI e no seu bom funcionamento.

Uma vez acordados em pormenor os requisitos, o SIGEI passou à fase de concepção, implementação e teste, como veremos de seguida na secção 1.3 e com mais detalhe na secção 3. Após os primeiros testes em situação real simulada bem sucedidos, foram realizadas acções de formação para todos os funcionários envolvidos na sua utilização. Simultaneamente o SIGEI foi instalado em rede e disponibilizado durante algum tempo em paralelo com o sistema existente. Actualmente, o SIGEI encontra-se em utilização em tempo real por todas as unidades do INEGI sediadas no Complexo Empresarial do Barroco e em utilização em tempo diferido pelas sediadas na FEUP.

1.2 Requisitos gerais do SIGEI e seus utilizadores

O estudo da estrutura organizacional, dos canais de informação e dos processos de gestão do INEGI levou o grupo de trabalho a elaborar um conjunto de requisitos para o seu Sistema de Informação.

Da análise dos requisitos a que o SIGEI teria de dar resposta, chegou-se à conclusão que deveria ser um sistema que oferecesse:

1. Integração da informação gerada pelas várias Unidades num repositório comum;
2. Disponibilização em tempo real de toda a informação relativa aos vários projectos das várias unidades do INEGI, esta disponibilização pode ser condicionada;
3. Integração ou interligação com as aplicações informáticas existentes por forma a minimizar a duplicação de procedimentos de introdução de dados;
4. Autonomia de gestão da informação local de cada unidade, para o caso de dificuldades de interligação informática pela rede; no entanto o sistema global do INEGI deveria conter informação actualizada em tempo real sempre que possível e sempre que as ligações informáticas fossem re-estabelecidas;
5. Garantia de integridade e de correcção da informação existente em base informática, nomeadamente através da definição de procedimentos de validação e verificação;
6. Segurança da informação existente em base informática, nomeadamente através da definição de procedimentos de salvaguarda e de níveis de acesso.

Como se pode constatar, algumas destas características terão que ser necessariamente garantidas também através de procedimentos externos ao SIGEI, o que não invalida que o sistema ofereça um bom apoio à sua realização.

Colocou-se em determinado momento do estudo inicial a possibilidade de permitir o acesso ao SIGEI através da Internet¹. O receio existente por parte das direcções das instituições num possível acesso a informação interna levou o INEGI a não enveredar por essa via. Ficou sempre em aberto a possibilidade futura de tal vir a suceder, se os mecanismos de protecção se tornassem robustos e convincentes.

Um aspecto muito importante do estudo prévio refere-se à caracterização dos utilizadores, no que diz respeito aos seus conhecimentos. Essencialmente o SIGEI tem dois tipos de utilizadores:

- os coordenadores (de projectos e das unidades), e técnicos superiores dos serviços centrais;
- os funcionários dos secretariados (das unidades e dos serviços centrais).

Do ponto de vista dos conhecimentos, qualquer dos grupos tem bons conhecimentos de informática ao nível da utilização de ambientes gráficos, tipo escritório electrónico. De qualquer forma o sistema terá grande utilização sobretudo pelos secretariados, uma vez que se pensa dar-lhes maiores responsabilidades de futuro. Desta forma, haverá particular cuidado, nomeadamente nos testes e na formação.

1.3 Estratégia de Construção do SIGEI

O grupo de trabalho que foi constituído para a análise dos requisitos do SIGEI, e que continuou envolvido no processo de construção, decidiu que a melhor estratégia para o desenvolvimento do sistema seria a especificação, implementação e teste de sub-sistemas para diferentes funções. Foi ainda considerada a opção de selecção, aquisição e adaptação de um sistema existente no mercado. No entanto, a percepção do custo e do esforço de adaptação levou a não enveredar por esta solução. Com base nos requisitos funcionais, foram identificados e desenvolvidos os seguintes sub-sistemas:

- Sub-sistema de informação para gestão de contactos externos ou sub-sistema de correio ('mailing');
- Sub-sistema de informação para gestão de propostas de projectos, desde a preparação de propostas à abertura dos projectos;
- Sub-sistema de informação para gestão de projectos: gestão de toda a informação relacionada com um projecto.

Após o desenvolvimento dos três protótipos destes sub-sistemas houve uma primeira fase em que foram testados localmente, avançando-se depois para o seu teste em rede. Em todos estes testes foram ouvidas as críticas e sugestões dos utilizadores, tendo os protótipos sofrido

¹ O estudo inicial, que balizou o desenvolvimento futuro do SIGEI, foi efectuado em finais de 1994. Nessa altura a Internet dava os últimos passos para se tornar acessível a qualquer utilizador. O navegador WWW NCSA MosaicTM surge em 1993, e em 1995 realiza-se em Bolsa a OPV da empresa Netscape, que comercializa o primeiro produto do tipo, baseado no Mosaic. Apenas em 1996 a Microsoft está em condições de lhe fazer frente, com o Internet Explorer, também baseado no Mosaic, depois de alguma hesitação aparente face ao futuro da Internet (cf. R. H. Zakon, ©1993-7, 'Hobbes' Internet Timeline', v3.1, <http://info.isoc.org/zakon/Internet/History/HIT.html>).

melhorias, ao nível das interfaces gráficas, da base de informação e das restantes funcionalidades.

2. Descrição Global do SIGEI

Sendo o SIGEI um sistema estruturante de toda a área administrativa de apoio à gestão, o seu estudo abrangeu necessariamente todos os procedimentos administrativos. Sempre que possível foram implementadas formas simplificadas dos processos com vista a facilitar a sua utilização futura.

O desenvolvimento de um sistema deste tipo deve ter em atenção que a componente mais estável é a estrutura da informação base por si manipulada, e não os procedimentos adoptados para manipulação da mesma. Este facto justifica que o sistema tenha sido desenvolvido tendo como suporte uma base de dados relacional cuidadosamente normalizada e implementada com o apoio de um SGBDR², que alia as capacidades de gestão da informação à facilidade de desenvolvimento e manutenção das interfaces gráficas com os utilizadores que se tornam necessárias.

2.1 Sub-sistemas actuais do SIGEI

Como vimos, a análise de requisitos permitiu identificar 3 grandes sub-sistemas principais a construir no âmbito do SIGEI, nomeadamente, *Gestão de Contactos*, *Gestão de Propostas* e *Gestão de Projectos*.

O sub-sistema de *Gestão de Contactos* foi identificado porque se considerou essencial ao sucesso da instituição uma correcta e eficaz gestão da informação referente aos seus clientes e fornecedores. A informação presente neste sistema vai bastante para além da que se encontra num sistema de organização pessoal (tipo agenda), pois permite, por exemplo, manter informação parcialmente estruturada relativa a uma instituição cliente, tais como os vários responsáveis departamentais, o resumo de contactos já efectuados e a marcação de alertas para contactos futuros (tipo 'workflow').

O sub-sistema de *Gestão de Propostas* passou a ser responsável por todo o ciclo de processamento de informação que vai desde o momento em que o contacto com um cliente torna necessária uma resposta (em termos de proposta de fornecimento de serviço ou produto) até à formalização de um contrato. Este sub-sistema suporta, por exemplo, informação relativa ao historial das versões de uma proposta.

O sub-sistema de *Gestão de Projectos* entra em acção após a recepção de uma confirmação de encomenda, sendo desencadeado todo o processo que levará à sua execução. Este sub-sistema é responsável pelo tratamento de informação ao nível da gestão das actividades produtivas do projecto, apoiando a satisfação dos prazos da encomenda (essencialmente quando se trata de projectos que envolvem entregas de produtos faseadas no tempo) e ao nível da gestão contabilística do projecto como centro de custo.

² Sistema de Gestão de Bases de Dados Relacionais.

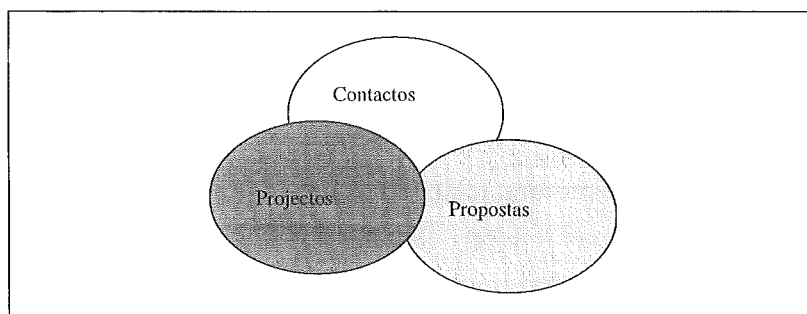


Figura 1 - Sub-Sistemas do SIGEI

O SIGEI permite assim criar, alterar e consultar todo o tipo de informação relacionada com as actividades das várias unidades do INEGI (figura 1).

2.2 Arquitectura Distribuída

Como já foi anteriormente referido, o SIGEI é um sistema distribuído, actuando na informação das várias unidades do INEGI (ver figura 2). Consequentemente, teve que ser pensado tendo em atenção que não seria para utilização apenas local mas também para permitir que utilizadores fisicamente afastados tivessem a possibilidade de aceder a toda a informação.

A opção tomada (ver figura 2) foi a de centralizar a base de dados e distribuir pelas diferentes unidades as aplicações necessárias para interagir com ela, ou seja, adoptar uma arquitectura cliente-servidor.

Como a informação é relativamente independente entre as diferentes unidades, poder-se-iam criar bases de dados locais mantendo-se a possibilidade de os utilizadores acederem à totalidade da informação. Assim, se um utilizador desejasse obter informação sobre todos os projectos existentes no INEGI teria de desencadear uma operação que iria a todas as bases de dados locais ver os projectos existentes.

Esta opção não foi a seguida, uma vez que o número de operações a fazer sobre todo o conjunto de informação (listagens e consultas) provocaria um fluxo de informação na rede superior ao fluxo gerado pela solução centralizada.

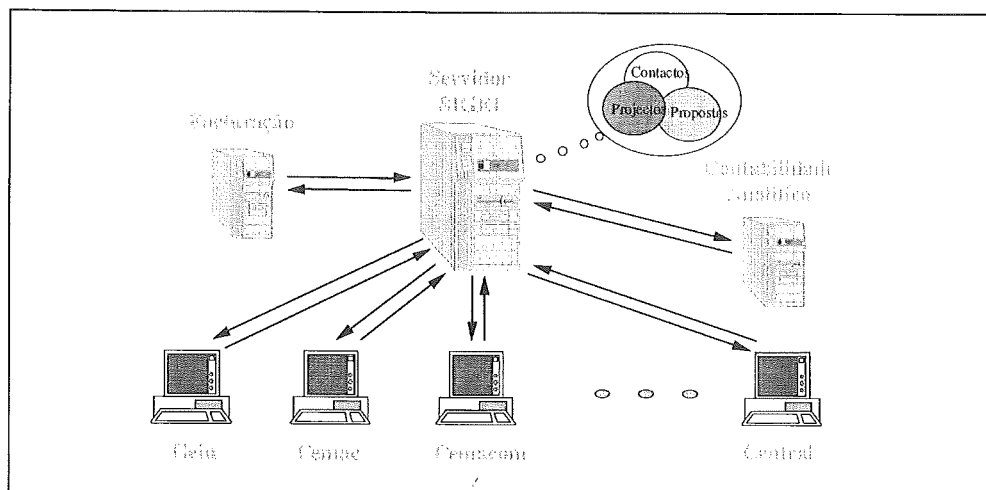


Figura 2 - Arquitectura cliente-servidor do SIGEI

2.3 Interface com Outros Sistemas

Como já foi referido, o SIGEI irá interagir com dois outros sistemas já existentes (ver figura 2), Facturação e Contabilidade Analítica. Prevê-se ainda a ligação à Contabilidade Geral.

A colaboração entre o SIGEI e estes sistemas resulta do facto de existir informação que está presente simultaneamente nesses sistemas e no SIGEI. Por exemplo, quando existe um pedido de emissão de factura por uma unidade, a informação vai ser guardada no SIGEI mas também é necessária ao sistema de facturação.

Numa primeira fase, essa ligação é indirecta, isto é, está dependente da confirmação de um utilizador específico. Por exemplo, no caso de ser solicitada a emissão de uma factura, esse pedido fica numa lista de espera até que o responsável dos serviços administrativos e financeiros indique ao SIGEI que pode transferir essa solicitação para o sistema de facturação (possivelmente introduzindo mais alguns dados, que ao nível da unidade não eram necessários, como o valor do IVA).

Numa fase posterior de evolução do SIGEI, estas interfaces entre sistemas existentes poderão ser melhoradas, diminuindo ao mínimo a necessidade de um utilizador validar as transferências.

2.4 Caracterização dos Tipos de Acesso dos Utilizadores

Como já foi referido, um dos requisitos do SIGEI era o de permitir definir utilizadores com diferentes níveis de acesso ao sistema.

Efectivamente, devido à quantidade e natureza da informação existente, foi necessário implementar um controlo de acesso à informação disponibilizada. Como o SIGEI é um sistema que congrega toda a informação das várias unidades, não haverá a necessidade de qualquer utilizador aceder a toda a informação. Sendo assim, existem utilizadores que poderão manipular a informação ao nível da unidade e utilizadores que o poderão fazer a nível global. Ainda dentro destes dois grupos, há que distinguir os utilizadores que poderão alterar dados dos que poderão apenas ler informação (referente apenas a um projecto ou a todos os projectos da unidade).

3. Método de desenvolvimento Utilizado

Esta secção tem por objectivo descrever com algum detalhe as fases (ver figura 3) que levaram ao desenvolvimento do protótipo do SIGEI.

A abordagem seguida utilizou em simultâneo o processo de desenvolvimento em cascata (Boehm, 1976) e o processo de prototipificação (Pressman, 1980). A ênfase no primeiro processo foi devida à clara necessidade de se pensar primeiro no que se vai fazer a seguir, tentando abordar os problemas do geral para o particular (abordagem 'top-down'). Note-se que, ao contrário das aplicações reais do modelo clássico cascata, não houve, por exemplo, uma separação entre a pessoa que concebeu e a que construiu o SIGEI. O técnico principal do projecto participou em todas as fases do mesmo, tendo sido responsável pela concepção,

construção, teste e formação dos utilizadores. A ênfase na prototipificação permitiu testar as maquetes do SIGEI, que incluíam funcionalidades rudimentares, com o objectivo particular de avaliar a receptividade e a ergonomia das interfaces gráficas com o utilizador (estruturas de menus, formulários, gráficos e outros relatórios).

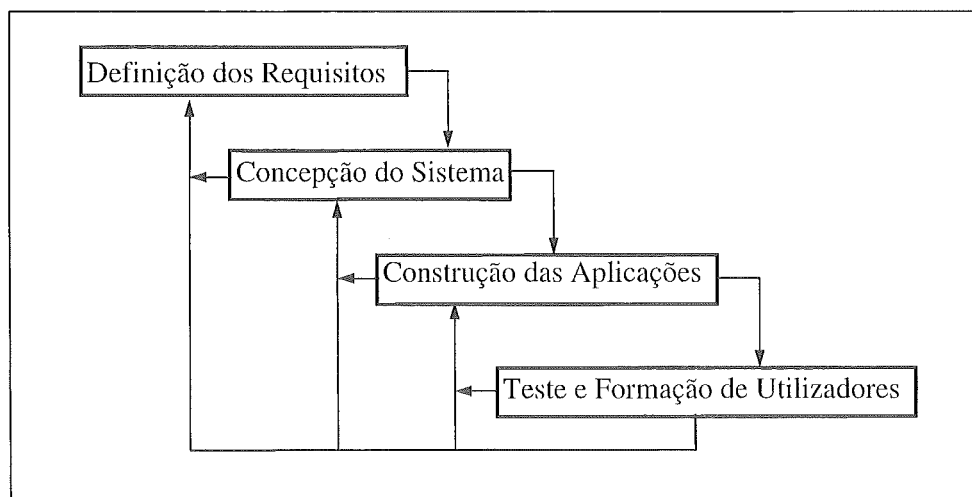


Figura 3 - Fases de desenvolvimento dos protótipos do SIGEI

3.1 Definição dos Requisitos

A partir do momento em que ficaram definidos os principais requisitos do sistema de informação para gestão do INEGI e se verificou a sua inexistência no mercado, partiu-se para o seu desenvolvimento. A equipa de trabalho responsável pelo projecto não só identificou um conjunto global de requisitos, como produziu documentos detalhando minuciosamente os requisitos e fluxos de informação dos processos.

O sistema deveria integrar a informação gerada pelas várias unidades, disponibilizando-a em tempo real. Esta disponibilização deveria ser efectuada nos moldes especificados pelos utilizadores. Deveria ainda interagir com os sistemas existentes, nomeadamente o de facturação e o de contabilidade analítica, e futuramente com o de contabilidade geral, e, finalmente, rever possíveis problemas de ligação à base central (o que implica uma gestão local autónoma, actualizando logo que possível a base central).

Teria ainda que haver um especial cuidado com procedimentos para assegurar e garantir a integridade, segurança e correcção da informação existente, o que levou por exemplo à definição de diferentes tipos de utilizadores.

3.2 Concepção do Sistema

Após a definição dos requisitos avançou-se para a concepção do sistema. Como já referimos, uma das primeiras opções tomadas, devido à complexidade do problema, foi a de dividir a sua concepção em sub-sistemas. Desta divisão resultaram os três sub-sistemas intimamente relacionados para Gestão de Contactos, Gestão de Propostas e Gestão de Projectos.

O parte mais relevante do processo de concepção foi a construção do modelo da informação, tarefa essencial no desenvolvimento de qualquer sistema de dimensão razoável. Esta modelação foi feita recorrendo a OMT, "Object Modeling Technique" (Blaha et al 1988, Rumbaugh et al 1991), em particular utilizando o seu modelo estático de classes. O modelo estático de classes do OMT baseia-se no modelo entidade-associação (Chen, 1976) com extensões que permitem modelação orientada por objectos. O OMT inclui, para além da componente estática utilizada neste trabalho, componentes dinâmicas e funcionais que não serão aqui referidas³.

Em conjunto com o modelo de informação dos sub-sistemas do SIGEI foram sendo concebidas, desenhadas e avaliadas maquetes das interfaces gráficas.

3.3 Definição da Plataforma de Desenvolvimento

O grupo de trabalho decidiu que a plataforma de desenvolvimento seria baseada no Sistema Access da Microsoft, devido às suas potencialidades, nomeadamente na execução das operações de gestão da base de dados, na flexibilidade da prototipificação das interfaces gráficas, ao nível da interacção com o utilizador e dos relatórios, e da sua capacidade de interligação com outras aplicações (via OLE, ODBC e SQL). Esta solução permitiria também uma futura utilização do SQL Server, se se viesse a verificar a necessidade de um sistema de maior dimensão.

Além disso, o facto de a Microsoft oferecer fortes garantias de permanente actualização e integração das suas plataformas, contribuiu para esta decisão que, até ao momento, parece ter sido feliz.

3.4 Construção das Aplicações

A construção das aplicações foi feita com base numa abordagem do particular para o geral (abordagem 'bottom-up') em contrapartida à abordagem inversa seguida nas fases anteriores de definição do SIGEI. Esta abordagem foi adoptada por se ter considerado que permitiria obter um sistema mais flexível e adaptado às necessidades.

Foram identificadas as operações aplicadas a cada uma das entidades da base de dados e especificadas as suas funcionalidades. Foi possível assim obter um sistema que pode ser facilmente expandido e cuja arquitectura se encontra baseada na sua componente mais estável, que é a estrutura de informação.

Foi também dado relevo à capacidade oferecida pelo Microsoft Access para o desenvolvimento de aplicações em ambiente gráfico, que se tornam agradáveis do ponto de vista do utilizador, nomeadamente para os que não têm formação específica em informática.

³ O principal autor do OMT, James Rumbaugh, conjuntamente com Grady Booch e Ivar Jacobson estão envolvidos desde há algum tempo na normalização das linguagens de modelação conceptual, tendo proposto a UML, 'Unified Modelling Language' (ver exemplo <http://www.rational.com/uml/index.html>, ou [12]).

3.5 Teste e Formação de Utilizadores

Uma vez criado o protótipo, houve a necessidade de o testar. Numa primeira fase os testes foram efectuados localmente, existindo somente uma base de dados e apenas um utilizador. Apesar da utilidade deste primeiro teste que veio permitir alguns ajustes, é de realçar o teste seguinte que foi efectuado em rede com os utilizadores finais do sistema. Este teste veio permitir adaptar o sistema às necessidades expressas pelos utilizadores (principalmente ao nível da inserção e consulta de informação).

Depois de desenvolvido, testado e instalado, o SIGEI necessitará sempre de manutenção, quer para o manter de acordo com as alterações da realidade, quer para satisfazer as expectativas dos utilizadores, sempre mais exigentes.

4. Modelos Conceptuais OMT e Exemplos de Interfaces do SIGEI

De seguida, descrevem-se sucintamente os três sub-sistemas, implementados como módulos do SIGEI, e as suas interligações. Como exemplo, apresenta-se o modelo conceptual estático OMT do sub-sistema de gestão de propostas e das suas interligações com os outros dois sub-sistemas.

4.1 Gestão de contactos

Este foi o primeiro módulo a ser desenvolvido, tendo como finalidade a criação de uma base de dados de clientes e fornecedores útil a todas as unidades do INEGI.

Basicamente, a informação que é necessário guardar nesta base de dados é a relacionada com clientes e fornecedores. Essa informação inclui nomes, endereços e números de contribuinte das empresas, assim como contactos individuais nessas empresas. Também são guardados os aspectos mais relevantes referentes à caracterização da empresa, ao tipo de produtos fabricados, aos interesses manifestados e aos interesses eventuais. Desta forma, possibilita-se a manutenção de um conhecimento mínimo sobre as empresas existentes na base de dados.

Por último, é também feita uma classificação das diferentes empresas de acordo com os aspectos mais relevantes das suas actividades, para facilitar operações de pesquisa e selecção à base de dados. Esta classificação é feita pela atribuição de várias palavras-chave a cada empresa.

4.2 Gestão de propostas

Após o desenvolvimento do módulo anterior passou a existir uma base de dados para contacto com diferentes empresas. Esta base de dados permitiu avançar para a próxima etapa, que consistia no desenvolvimento de um módulo para a gestão das propostas. Este módulo engloba a estrutura de informação necessária para a criação e acompanhamento de uma proposta até uma possível situação de encomenda de um serviço ou produto ou até ao seu cancelamento.

O ciclo de vida do processo abordado por este módulo parte de um contacto inicial, solicitando um serviço ou produto, e prossegue com o envio de uma proposta por parte do INEGI. A proposta inclui, por exemplo, a informação que descreve o trabalho a desenvolver, as condições de pagamento, o seu valor total e o prazo de conclusão. É esta proposta que o INEGI envia ao cliente, podendo haver um acompanhamento da sua evolução. A informação relativa a este acompanhamento vai sendo armazenada, para facilitar que pessoas não directamente envolvidas na proposta acedam ao seu estado ou à situação das propostas pendentes.

Chegando-se a acordo com o cliente, é então emitida uma confirmação de encomenda, passando a proposta a estar disponível no sub-sistema de gestão de projectos. Caso não haja acordo a proposta é anulada, indicando-se o motivo porque tal aconteceu.

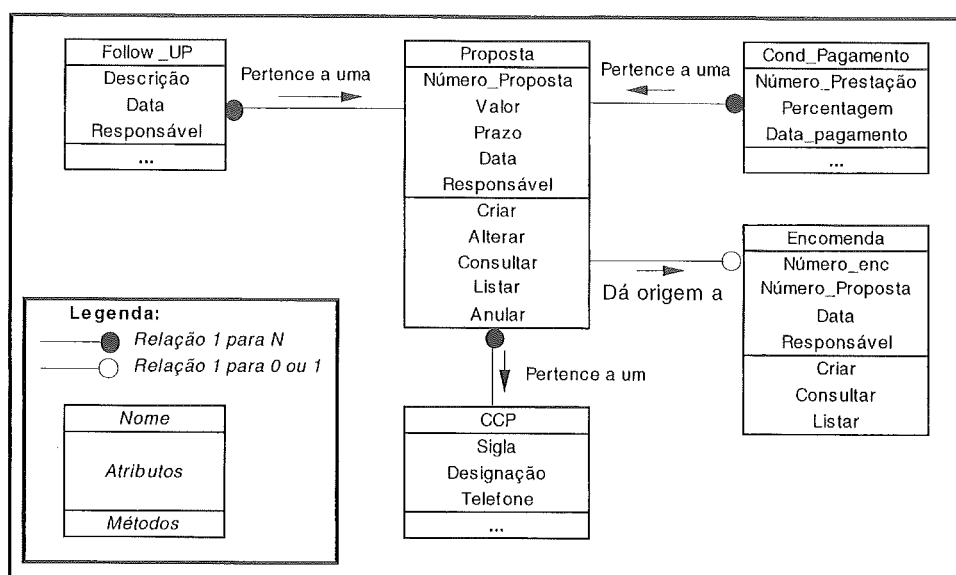


Figura 4 - Gestão de propostas

4.3 Gestão de projectos

Quando uma proposta dá origem a uma encomenda, é aberto um projecto que terá a seu cargo a satisfação dessa encomenda. É toda a estrutura de informação que envolve a satisfação de uma encomenda que é tratada neste módulo.

O primeiro documento a ser criado é o referente à abertura do projecto. Neste documento guarda-se vária informação, da qual se poderá destacar a que diz respeito ao cliente ou clientes, nome e designação do projecto, tipo e entidades que o irão financiar, responsável pelo projecto e despesas e receitas previstas.

A partir do momento em que um projecto é aberto e se inicia todo o ciclo para a resolução da tarefa a que terá de dar resposta, gera-se um conjunto de informações que terá de ser tratado e guardado. Esta informação relaciona-se com o tratamento dos processos de requisição a fornecedores, de facturação aos clientes, facturação de fornecedores, justificação de despesas e de receitas e emissão de guias de remessa.

4.4 Ligação entre os módulos do SIGEI

A ligação entre os diferentes módulos do SIGEI é assegurada ao nível da informação pelas respectivas entidades principais. Por exemplo, a entidade Empresa, do módulo de Gestão de Contactos, é utilizado pelos outros dois módulos, uma vez que fornece toda a informação necessária para o contacto com clientes e fornecedores. Esse contacto tem de existir tanto ao nível da constituição de uma proposta e do seu envio (módulo Propostas) como ao nível das guias de remessa e requisições que são feitas a fornecedores (módulo Projectos).

Note-se que para um projecto ser aberto terá que existir pelos menos uma proposta correspondente ou associada a esse projecto, daí o serem necessárias ligações entre estes dois módulos (ver figura 5).

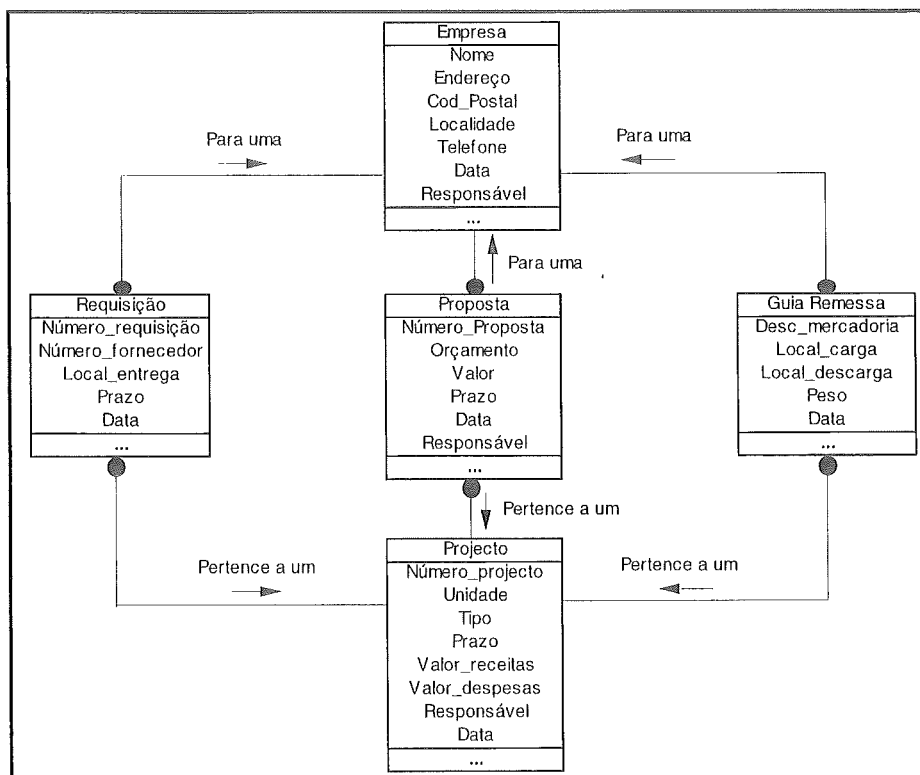


Figura 5 - Ligação dos módulos

4.5 Exemplos da interface com os utilizadores

Após a definição do modelo de classes do sistema, exemplificado na secção anterior com o sub-sistema de gestão de propostas, apresentam-se em seguida, também recorrendo ao mesmo sub-sistema, exemplos das interfaces gráficas. A interface inclui a estrutura de diálogo ou menus (figuras 6 e 8), cujas opções principais são os métodos da classe Proposta (figura 4), inclui as janelas de interacção ou formulários (figura 7) e os relatórios a imprimir (figura 9).

4.6 Interfaces com o utilizador por formulários

O sistema apresenta um conjunto variado, mas uniforme e consistente, de interfaces com os utilizadores. Apresenta-se a seguir um exemplo de como criar uma proposta usando o sub-

sistema de gestão de propostas. De notar que todas as interfaces foram definidas tendo em vista um sistema com características típicas do ambiente Windows.

Depois de o utilizador se identificar e lhe ser permitido o acesso, fica perante o menu inicial do SIGEI. Pode assim seleccionar a opção que lhe permite executar a operação desejada, seja por exemplo, criar uma proposta.

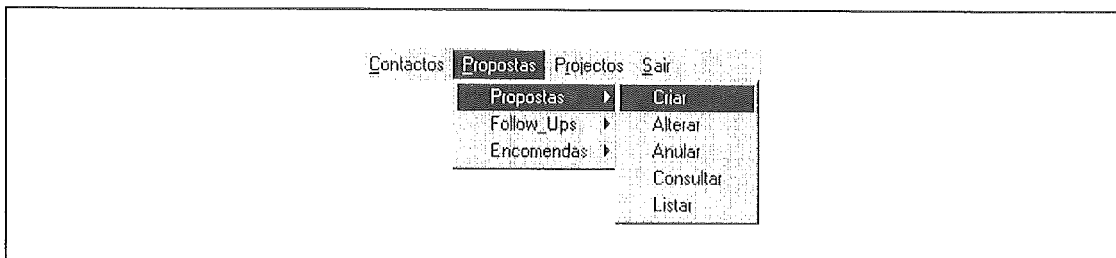


Figura 6 - Opção para criar uma proposta

Após a selecção, o utilizador fica perante o formulário adequado, tendo apenas de preencher os campos aí indicados (figura 7).

 A screenshot of a Windows-style window titled 'Criar Proposta'. The form contains several sections:

- Designação CCP:** GEIN
- Número Proposta:** PE96.04.1
- Designação do Trabalho:** Desenvolvimento de um sistema para gestão administrativa da Universidade de Aveiro
- Cliente:** Universidade de Aveiro
- Nome:** Universidade de Aveiro
- Morada:** Aveiro
- Código Postal:** 3800
- Localidade:** Aveiro
- Telefone:** [empty]
- Telefax:** [empty]
- Número de Contribuinte:** [empty]
- Orçamento Previsional:** Orçamento dos custos para este sistema ...
- Condições de Pagamento:** A table with columns 'Prestação', 'Valor', and 'Data'.

Prestação	Valor	Data
1	30,00%	96.04.30
2	60,00%	96.12.01
0	0,00%	
- Valor:** 1.000\$00
- Prazo:** 96.12.01
- Responsável Follow-UP:** António Amador
- Responsável:** João Falcão e Cunha
- Data Próximo Contacto:** 96.04.10
- Data:** 96.04.03
- Fechar** button

Figura 7 - Formulário para criar uma proposta

De salientar a facilidade de preenchimento destes formulários, havendo, sempre que possível, indicações de ajuda aos utilizadores.

4.7 Interfaces com o utilizador por relatórios

Outro dos aspectos relevantes num sistema deste tipo é a resposta que ele tem de dar ao nível da geração de relatórios. O SIGEI permite produzir relatórios de diferentes tipos, desde as etiquetas para envio de correspondência, aos documentos pré-formatados prontos a enviar para o exterior, bem como documentos internos dos projectos ou mapas de situação de gestão.

Por exemplo, para preparar o documento de Proposta que é enviado para um cliente a confirmar uma Encomenda, o utilizador selecciona a opção correspondente a Listar (figura 8).

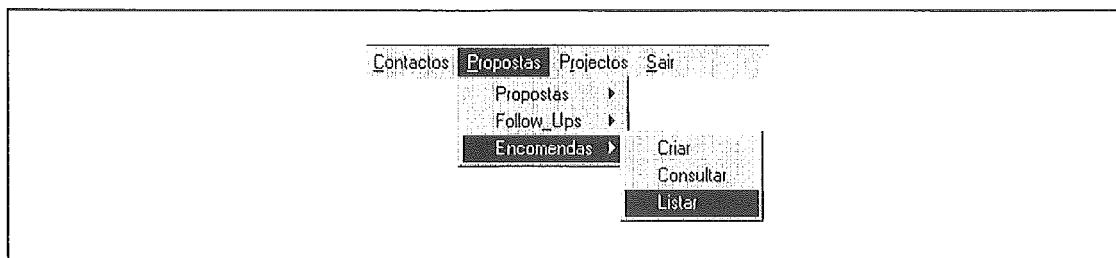


Figura 8 - Opção para listar uma confirmação de encomenda

Aparece então um formulário com as propostas existentes para o utilizador poder seleccionar qual a que deseja. Ao efectuar esta selecção o documento é visualizado, podendo então ser impresso (figura 9).

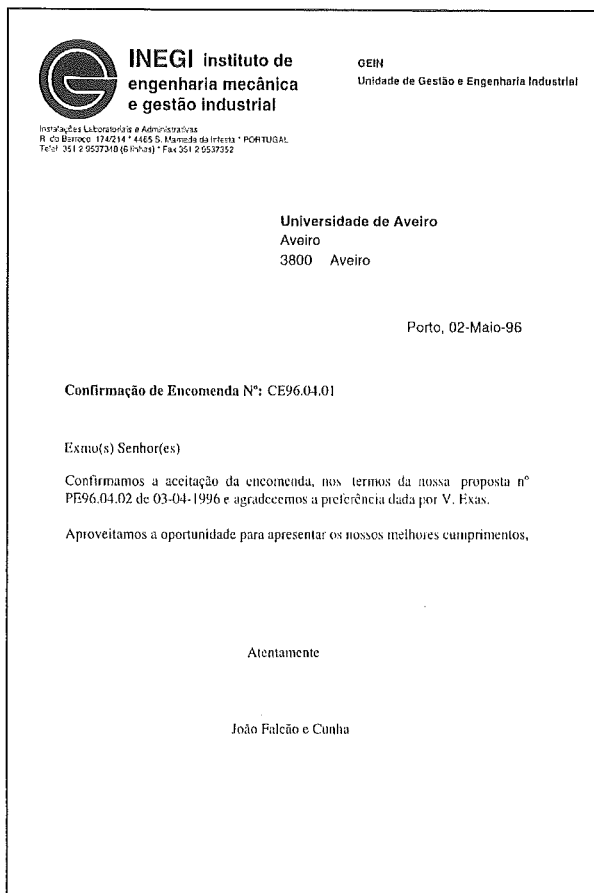


Figura 9 - Impresso de confirmação de encomenda

5. Conclusões e perspectivas de evolução

O presente trabalho descreve o desenvolvimento de um sistema de informação adequado à estrutura administrativa do INEGI e responsável pela gestão dos procedimentos existentes. O sistema responde globalmente aos requisitos exigidos, quer do ponto de vista da gestão do INEGI, quer do ponto de vista dos seus utilizadores.

Como é natural, o SIGEI tem vindo a sofrer alterações para se adaptar a alterações da realidade e para cumprir expectativas acrescidas dos seus utilizadores e dos gestores do INEGI.

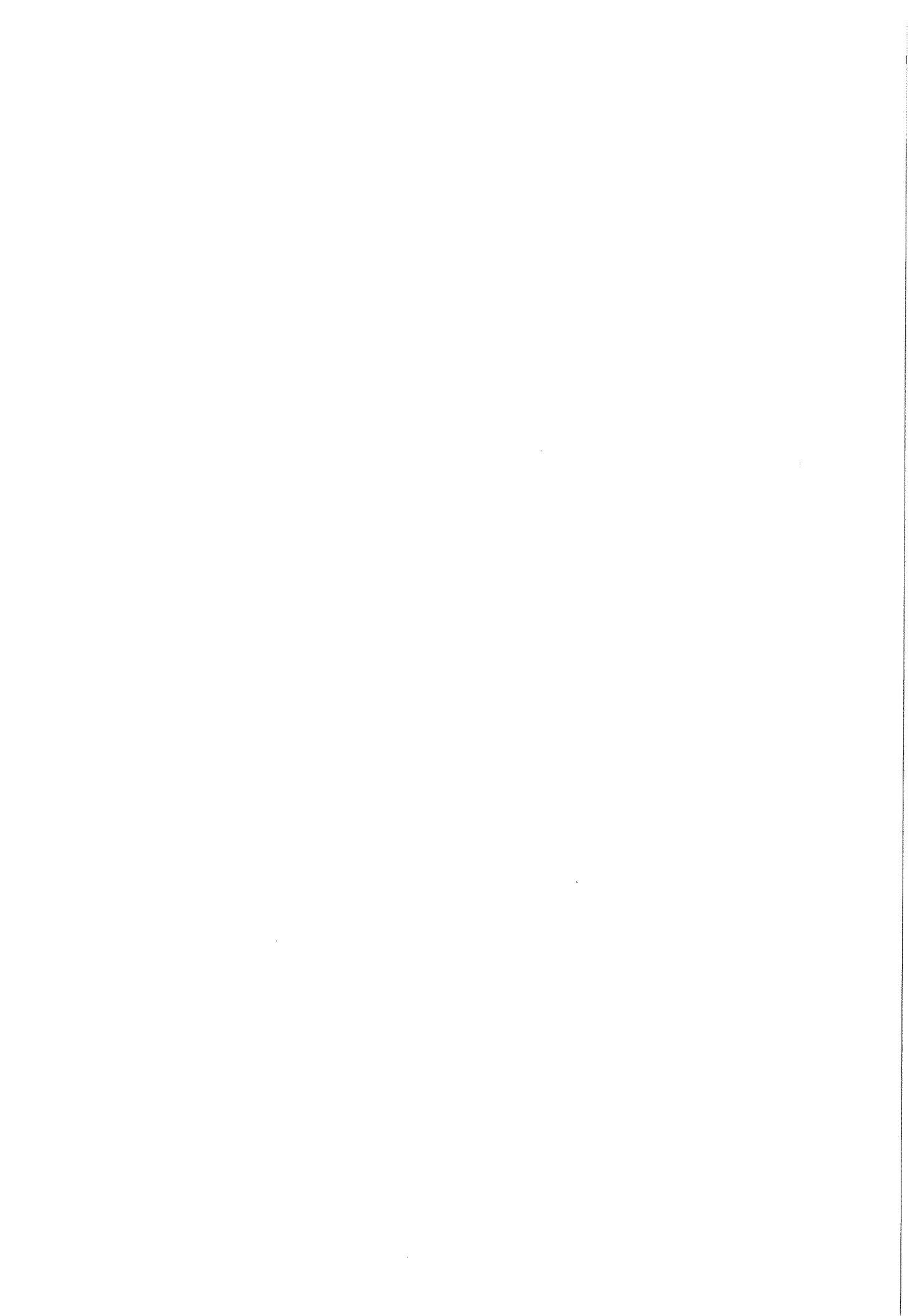
Em particular, estão em curso intervenções para facilitar as ligações com os sistemas de Facturação e de Contabilidade Analítica.

Está neste momento também em estudo a disponibilização de interfaces com o SIGEI na Internet, com o objectivo de permitir aos responsáveis dos projectos e das unidades o acesso de qualquer local à informação relevante, sobretudo com canais de comunicação lentos (tipo modem). O actual estado da tecnologia torna esta tarefa extremamente simples, devido às ferramentas existentes. Seria também simples, utilizando por exemplo bibliotecas CGI (Tittel, 1996) para interface com o gestor da base de dados ACCESS, disponibilizar na Internet uma versão idêntica ao SIGEI, quer do ponto de vista funcional, quer do ponto de vista da interface gráfica, partilhando a mesma base de informação, em tempo real. A utilização da Internet, como infraestrutura de desenvolvimento de um sistema de gestão de informação é hoje em dia fundamental, sobretudo para possibilitar a abertura do sistema aos parceiros. Para proteger a informação mais crítica de acessos indevidos, existem tecnologias de segurança (exemplo das 'firewall') que permitem a constituição de uma Intranet bem protegida (Bremner et al., 1997).

Tendo surgido de um projecto de certificação do sistema da qualidade do INEGI, o projecto de desenvolvimento do SIGEI pautou-se por uma preocupação constante e invulgar com a qualidade do próprio processo da sua construção, desde a constituição de um grupo de trabalho ao mais alto nível do INEGI até à selecção de métodos e ferramentas robustas. O resultado é assim um sistema que pode ser facilmente adaptado a novas tecnologias e que está sem dúvida adaptado ao INEGI e às pessoas que com ele trabalham.

6. Referências

- [1] Blaha, Michael R., William J. Premerlani and James E. Rumbaugh, *Relational Database Design Using an Object-Oriented Methodology*, Comm. of the ACM, 31 (1988) pp. 414-427.
- [2] Boehm, Barry W., *Software Engineering Economics*, Englewood Cliffs, (1976) xxiv+767pp.
- [3] Cabral, J. A., *Acreditação dos Laboratórios do INEGI*, Relatório interno (1994a).
- [4] Cabral, J. A., *A Gestão da Qualidade no INEGI e a Acreditação dos seus Laboratórios: Algumas notas para reflexão*, Relatório interno (1994b).
- [5] Chen, Peter Pin-Shan, *The Entity-Relationship Model - Toward a Unified View of Data*, ACM Trans. on Database Systems, 1 (1976) pp. 9-36.
- [6] Challenge, *Novo Rumo*, Consolidação do INEGI (1994).
- [7] Codd, E. F., *A Relational Model of Data for Large Shared Data Banks*, Comm. of the ACM, 13 (1970) pp. 377-387.
- [8] Date, C. J., *An Introduction to Database Systems*, Vol.2, Addison-Wesley (1983) xv+383 pp.
- [9] Date, C. J., *Relational Database: Selected Writings*, Addison-Wesley (1986) xiv+497 pp.
- [10] Date, C. J., *An Introduction to Database Systems*, Vol.1 (4th Ed.), Addison-Wesley (1986) xx+639 pp.
- [11] Tittel, Ed, Mark Gaither, Sebastian Hassinger, *Cgi Bible*, IDG Books Worldwide (1996) 672pp.
- [12] Eriksson, Hans-Erik, Magnus Penker, *UML Toolkit*, John Wiley & Sons, (1997) xvii+397pp.
- [13] Bremner, Lynn M., Anthony F. Iasi, Al Servati, Anghony F. Iasi, *Intranet Bible*, Jamsa Pr (1997) 464pp
- [14] NCC, *Data Protection Codes of Practice*, NCC Blackwell Limited, Oxford-UK, (1990) 126 pp.
- [15] Pressman, Roger S., *Software Engineering (2nd Ed.)*, McGraw-Hill (1988) xx+567 pp.
- [16] Rumbaugh, James E., Michael R. Blaha, William J. Premerlani, Frederick Eddy e William Lorensen, *Object-Oriented Modeling and Design*, Prentice-Hall (1991) xii+500pp.
- [17] Sá, Rui, *Propostas de Alteração ao Sistema Informático de Contabilidade Analítica do INEGI*, Relatório interno (1993).
- [18] Teorey, Toby J., Dongqing Yang e James P. Fry, *A Logical Design Methodology for Relational Databases Using the Extended Entity-Relationship Model*, ACM Comp.Surveys 18 (1986) pp.197-222.



INSTRUÇÕES AOS AUTORES

Os autores que desejam submeter um artigo à Investigação Operacional devem enviar três cópias desse trabalho para:

Prof. Joaquim J. Júdice
Departamento de Matemática
Universidade de Coimbra
3000 Coimbra, Portugal

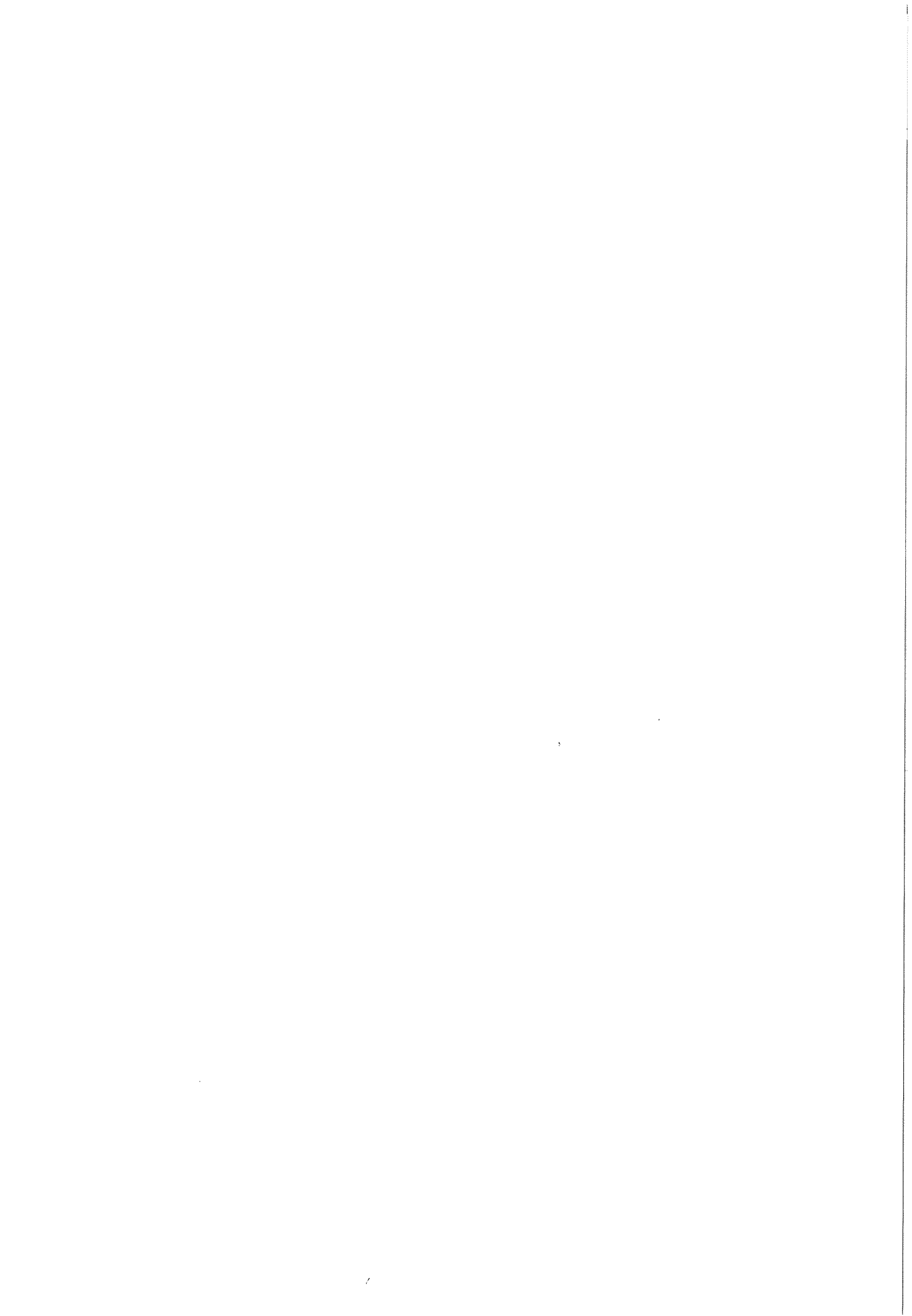
Os artigos devem ser escritos em Português ou Inglês. A primeira página deve conter a seguinte informação:

- Título do artigo
 - Autor(es) e instituição(ões) a que pertence(em)
 - Abstract (em inglês)
 - Resumo
 - Keywords (em inglês)
 - Título abreviado

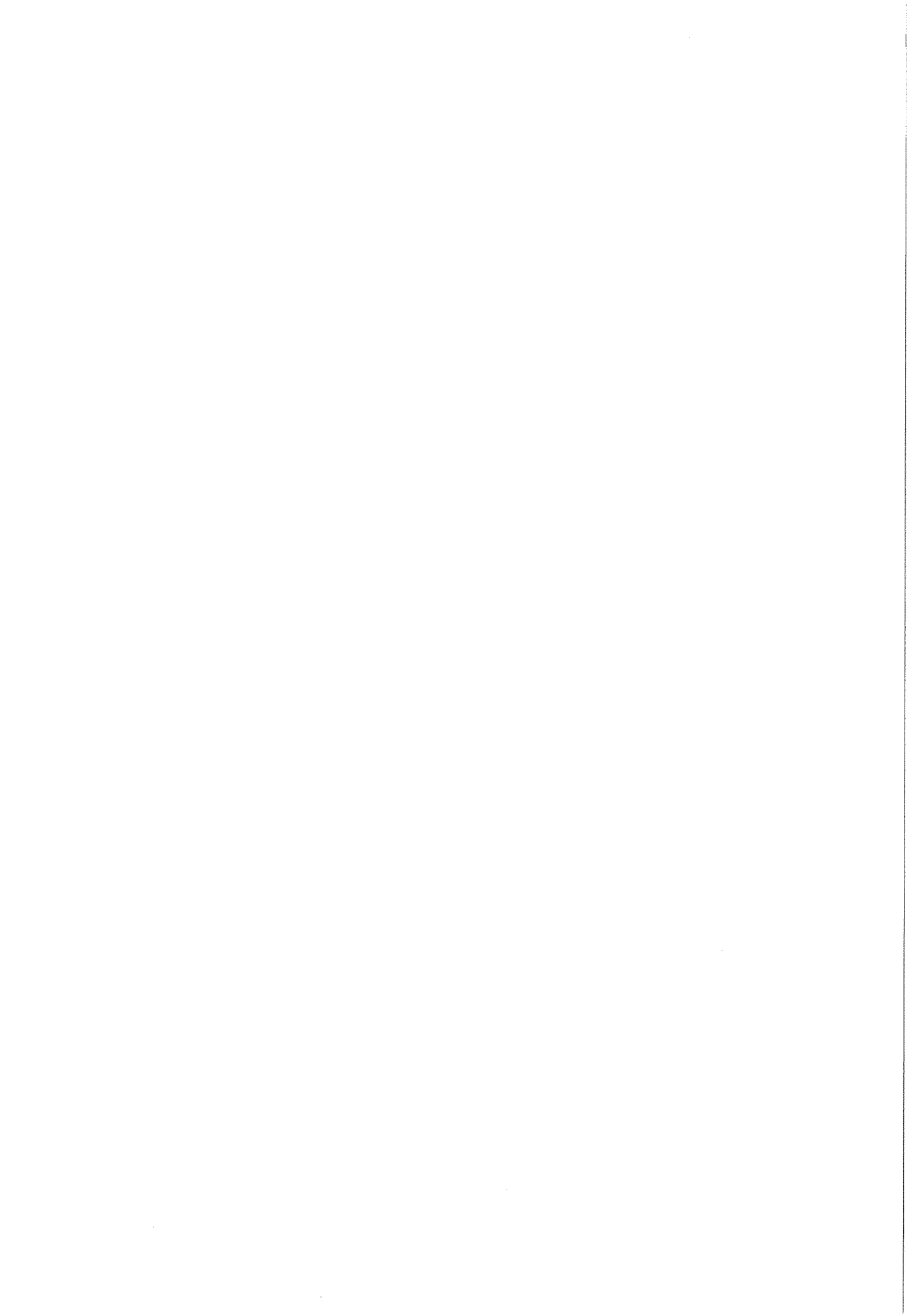
As figuras devem aparecer em separado de modo a poderem ser reduzidas e fotocopiadas. As referências devem ser numeradas consecutivamente e aparecer por ordem alfabética de acordo com os seguintes formatos:

Artigos: autor(es), título, título e número da revista (livro com indicação dos editores), ano, páginas.

Livros: autor(es), título, editorial, local de edição, ano.



Execução Gráfica:
Joaquim F. Macedo
S^{to}. Ant. Cavaleiros



ÍNDICE

P. M. Bruno, H. Ramalinho Lourenço, Branch-and-Bound para o problema de Job-Shop: Novas regras de ramificação	3
J. J. Júdice, M. C. Soares, Solution of some linear complementarity problems arising in variational models of mechanics	17
J. Filipe, I. H. Themido, Um modelo causal de previsão de vendas para um produto de grande consumo	33
R. M. Abreu, R. A. Costa, Modelação de um sistema de comercialização de especialidades farmacêuticas no distrito de Castelo Branco	49
R. Jardim Gonçalves, M. Martins Barata, A. Steiger Garção, J. Assis Lopes, Processos estocásticos de apoio à manutenção preditiva para ambientes industriais	65
A. Amador, J. Falcão e Cunha, J. A. Cabral, R. Sá, V. Sanches e Sá, O sistema de informação para gestão do INEGI	85

