Sotolongo-Aguilar, G; Guzmán-Sánchez, MV; Saavedra-Fernández, O., and Carrillo-Calvet, HA. Mining Informetrics Data with Self-organizing Maps. En: Davis, M and Wilson, CS, Ed. Proceedings of the 8 th International Society for for Scientometrics and Informetrics; Sydney, Australia. Sydney, Australia: BIRG; 2001: 665-673. ISBN: 0-7334-18201.

Mining informetric data with self-organizing maps

Sotolongo-Aguilar, Gilberto *; Guzmán-Sánchez, María Victoria *; Saavedra-Fernández, Oscar **; Carrillo-Calvet, Humberto A. ***

* I Finlay, Havana, Cuba: gsotolongo@finlay.edu.cu ; mvguzman@finlay.edu.cu
** EBSCO, Mexico: saavedra@ebsco.com
*** Laboratorio de Dinámica no Lineal, Fac. Ciencias, UNAM, Mexico: carr@servidor.unam.mx

Introduction

Informetric (scientometric o bibliometric) maps are important means for representing qualitative features of data sets. This practice for multivariate data has been available for some decades. Implementing such techniques requires general-purpose statistical software, e.g. STATISTICA, SAS, SPSS, or dedicated software, e.g. Leximapp, Tetralogie, Matrisme, and Vantagepoint. Other approaches use algorithms based on differential equations, i.e. BibTechMon. The application of artificial neural networks (ANN) to informetric data is still at a development stage e.g. NEURODOC; while some scholars consider such applications of difficult implementation. Nevertheless, unsupervised ANN as Kohonen algorithm (self-organizing maps - SOM) seems to be a suiTable approach for informetric "data mining" (DM) seeking "knowledge discovery" (KD).

Map metaphors have been associated with the handling of non-geographic information for a long time. It was probably Paul Otlet, one of the first in making explicit reference to the mapping of intellectual domains, envisioning the use of maps in the exploration of unknown information terrain (Skupin 2000). Otlet was the first in using the word "bibliometrics" in his "Traité de Documentation", 35 year before Alan Pritchard reintroduce and popularized it. Otlet considered that, "... in every knowledge order, measure is a higher form of knowledge ..." (Otlet 1934).

Multidimensional scaling (MDS), principal component analysis (PCA), and self-organizing maps (SOM) are the most commonly used methods for multivariate data analyses. In terms of complexity and computational efficiency, MDS does not scale up very well for large data sets, neither does PCA. Self-organizing maps (SOM), also known as Kohonen maps, are based on an artificial neural network approach. Being computational intensive, they produce meaningful results even for large data sets, receiving great attention in the visualization of text documents (Skupin 2000).

As Usama M. Fayyad states, "... Knowledge Discovery in Databases (KDD) is concerned with extracting useful information from databases... The term data mining has historically been used in the database community and in statistics (often in the latter with negative connotations to indicate improper data analysis). We take the view that any algorithm that enumerates patterns from, or fits models to, data is a data mining algorithm. We further view data mining to be a single step in a larger process that we call the KDD process ... KDD’s goal, as stated above, is very broad, and can describe a multitude of fields of study. Statistics has been preoccupied with this goal for over a century ... (Fayyad 1997)

Our general framework for technology watch is based on the french standard for "Watch services and watch system introduction services" (Std. AFNOR XP 1998), while data mining for that purpose is guided by the CRISP-DM (2000) (CRoss-Industry Standard Process for Data Mining). Our bibliometric work is based on the MOBIS Pro-Soft approach (Sotolongo-Aguilar 2000), where Viscovery® SOMine Enterprise Edition Version 3.0 software is used for creating SOM. In doing so, steps followed according to recommendations of Guido Deboeck (PCAI Magazine 1999, submitted paper)

Viscovery® SOMine is a tool for advanced analysis and monitoring of numerical data sets. Based on the concept of Self-Organizing Maps (SOM), it employs a robust variant of unsupervised neural networks, namely Kohonen’s Batch-SOM, which is further enhanced with a new scaling technique for speeding up the learning process. This tool provides a powerful means to analyze complex data sets without prior statistical knowledge. There is no need to be familiar with the basic algorithm. The user is guided through the training process by an environment of well-defined settings and defaults. Fairly practiced users can, however, tune a number of SOM creation parameters, while focused users may appreciate to be offered an efficient way to control data pre processing. Inside Viscovery® SOMine, a two-dimensional hexagonal grid realizes the SOM. Starting from a set of numerical, multivariate data, the ”nodes” on the grid gradually adapt to the intrinsic shape of the data distribution. Since the order on the grid reflects the neighbourhood within the data, features of the data distribution can be read off from the emerging landscape on the grid. The training process is time- consuming and may be followed by the user on a progress graph. In a second step, the data representation contained in the trained SOM is systematically converted for use in a spectrum of visualization techniques. Evaluating dependencies between components, investigating geometric properties of the data distribution, searching for clusters, monitoring new data—just to mention a few options—thereby become an intuitive and inspiring interactive process. In parallel, numerical information, such as cluster statistics, can be retrieved on demand at any time (Viscovery SOMine 1999)

In this paper informetric - SOM Case Studies are presented. In each case dozens of maps were obtained (some presented in this paper), eliciting features necessary for analysis. These studies could be seen as tasks performed in the framework of technology watch activities using DM & KD approaches.

Case Studies

A. Measuring the scientific production of Latin America in the field of agriculture.: a bibliometric study (Saavedra-Fernández, Sotolongo-Aguilar 1999)

The aim of this study was to explore the performance of Latin American and Caribbean (LA&C) countries in agricultural sciences. With this purpose we mapped a data set of 20 variables (features) from 22 Latin-American countries (title row and second column in Table 1). Variables (indicators) used were: 1) population per million inhabitants; 2) economic-active population per million inhabitants; 3) GIP in million U$D; 4) expenses in R&D in million U$D; References in: 5) SCI; 6) PASCAL; 7) INSPEC; 8) COMPENDEX, 9) CA; 10) BIOSIS; 11) MEDLINE; 12) CAB; 13) ICyT; 14) IME; 15) Journal in LATINDEX Directory; 16) % of World reference in CAB; 17)References in CAB per 100 thousand inhabitants; 18) References in CAB per 1000 millions U$D GIP; 19)References in CAB per R&D expenses in million U$S; 20) References in CAB per 100 researchers (equivalent).

Data was obtained from RICyT, the Iberian-American Network of Science and Technology Indicators (Red Iberoamericana de Indicadores de Ciencia y Tecnología 1999) corresponding to 1996; and data from Latindex, the Directory of Latin-American Journal data (Alonso-Gamboa 2000) from 1997. Data was structured in a 22 X 20 (country X indicators) matrix and used as input for training an ANN based on Kohonen Algorithm.

Summary of map history:

- SOM settings and data pre processing the scaling of data was set to variance; all data with the same priority and a sigmoid transformation was applied. Training was by default set to 36 cycles; tension was set to 0,5 also by default. Training schedule used was Accurate, exact mode for 2000 nodes. Map ratio was set to square.
- Map history description: Training started Thursday, 13 April of 2000 at 12:51:56, while Training ended Thursday, 13 April of 2000 at 12:55:01, lasting 3 minutes. Final errors were: Normalized distortion: 1,108E-016, and Quantization error: 5,046E-018.

The resulting SOM (Figure 1) was further explored for analysis purpose (Figure 2 and Table 1). Figure 1-show clusters formed as result of training a SOM with all data, while Figure 2 shows a SOM resulted by considering only CAB-related data (variables 16 - 20). Explanation of the two SOM statistics for average data are shown in Table 1. Stared (*) clusters in Figure 2 (corresponding to columns 2 and 4 in Table 1 shows the following: in case of C1* there is an increase of all parameter except those of CAB-related indicators, while for C2* members there is a considerable reduction of all parameters except for those CAB-related. Columns 3 and 5 shows the ratio of C1*/C1 and C2*/C2. Results suggest that Caribbean countries (Panama, El Salvador, Trinidad & Tobago, Barbados, Jamaica, Costa Rica and Cuba) shows the best performance in agricultural science production (output) according to input resources, compared to the rest of the countries including the Big Three (Argentina, Brazil and Mexico), that remain unchanged for both clustering.

B. Self-organizing map of Latin America & Caribbean scientific journals (Saavedra-Fernández, Sotolongo-Aguilar 1999)

The purpose here was to identify variables that could improve the visibility of (LA&C) science and technology publications. In this case, we mapped a data set of 19 variables (features) from 22 (LA&C) countries from the same sources of (Red Iberoamericana de Indicadores de Ciencia y Tecnología 1999), corresponding to 1990-1997; and data from Latindex, the Directory of Latin-American Journal data (Alonso-Gamboa 2000) corresponding to year 2000 countries (title row and second column in Table 2). The average of each indicator for the 8 year data set from RICyT for each country was calculated as well as the rank. The rank of data from Latindex was also obtained. All variables (indicators used were ranked data for: 1) Journals included in the LATINDEX Directory; Latin American Journals processed by 2) SCISEARCH; 3) CA;

4) BIOSIS; 5) MEDLINE; 6) CAB; References in 7) SCI; 8) PASCAL; 9) INSPEC; 10) COMPENDEX, 11) CA; 12) BIOSIS; 13) MEDLINE; 14) CAB; 15) ICyT; 16) IME; 17) References in SCI per 100 thousand inhabitants; 18) References in PASCAL per 100 thousand inhabitants; 19) References in SCI per 1000 millions U$D GIP. With all ranked data of 19 indicators from 22 countries, a 19 X 22 (country X indicators-rank) matrix was structure and used as input data for training an ANN based on Kohonen Algorithm.

Summary of map history:

- SOM settings and data pre processing SOM setting and pre processing was similar to the former.
- Map history description: Training lasted 1 minute, and final errors were: Normalized distortion 0, and Quantization error 0.

The resulting SOM (Figure 3e) was further explored for analysis purposes (Figure 3 a, b, c, d and Table 2). Figure 3e was trained withe all data, while the rest, Figure 3 a, b, c, d are the result of only considering: a) data from LATINDEX; b) journals processed by big information services (variables 2-6); c) references of Latin America present in big references and information services (GSIR) (variables 7 -16); d) relative indicators (variables 17 - 19). Results suggest the idea that Latin American effort is not enough to achieve the desired visibility in science and technology.

C. Nonlinear dynamics in biomedicine: bibliometric exploratory data analysis and data mining

(Sotolongo-Aguilar, Carrillo-Calvet ongoing research)

This application shows preliminary results of an ongoing research in the field of nonlinear dynamics in biomedicine. It moves in a much more complex direction than the two previous application, bearing in mind that here we pursue identification of research trend in a scientific field. Pre-processing and processing is much more time consuming and complex. Firstly, the identification of clusters of different features, e.g. keywords, substances, theoretical research, application research. In second place, incorporating data of the obtained clusters (adding value) to original data set; and finally, finding out research trends as a result of mapping the value added data set. In this case the aim was to map trends of research, based on values added to data set (database) extracted from clustering different features previously mapped.

For MESH, Nonlinear Dynamics is considered the study of systems, which respond disproportionately (nonlinearly) to initial conditions or perturbing stimuli. Nonlinear systems may exhibit "chaos" which is classically characterized as sensitive dependence on initial conditions. Chaotic systems, while distinguished from more ordered periodic systems, are not random. When their behaviour over time is appropriately displayed (in "phase space"), constraints are evident which are described by "strange attractors". Phase space representations of chaotic systems, or strange attractors, usually reveal fractal (FRACTALS) self-similarity across time scales. Natural, including biological, systems often display nonlinear dynamics and chaos. This is a fruitful research field and its exploration could lead to a better structuration of research projects in biomedicine.

For this study, a data set of 1884 records was obtained from PubMed 1990-2000, using the search equation: nonlinear dynamic [MESH]. Estimation of parameters for complete bibliography is 92% of articles and 85% of journals (Egghe 1990).

The Domain Features Clustering Schedule defined was the following:

No.	Domain Feature	Methods / Fields Envolved	Graph [Basic-Matrix]	Resulting SOM - Ward Clustering	Cluster Code
1	Basic				
	time development	count / years	graph (number of references per year)	-	-
	team size / kind of research (applied-theoretical)	average / number of signatures	graph (average signatures per year)	-	-
	research complexity		[id, research complexity]	BASIC	B
		years	-	-	-
		count / number of keywords	-	-	-
		count / number of substances	-	-	-
		count / number of document types	-	-	-
		count / number of signatures	-	-	-
2	Subject	correlation of coocurrence / mesh+year	[mesh+year]	MESH	M
3	Substances	correlation of coocurrence / substances+year	[substances+year]	SUBSTANCES	S

No.	Domain Feature	Methods / Fields Envolved	Graph [Basic-Matrix]	Resulting SOM - Ward Clustering	Cluster Code
4	Authors	correlation of coocurrence / authors+year	[authors+year]	AUTHORS	A
5	Document Type	correlation of coocurrence / document type+year	[document type+year]	DOC-TYPE	D
6	Affiliation	correlation of coocurrence / document type+year	[affiliation+year]	AFFILIATION	F
7	Country of Publication	correlation of coocurrence / country of publication+year	[country of publication+year]	COUNTRY	C
8	General Domain Features	correlation of coocurrence / All Clusters	[All Clusters]	GENERAL-DOMAIN-FEATURES	G

Here we only describe the method for Clustering Code M. The method for clustering the rest of the Codes is the same. Considering the year as an extra keyword, we obtained MESH coocurrence. Similarity matrix was obtained with multiple Pearson correlation all for 251 variables representing those MESH terms with a frequency of 9 or more (occurrence over 0,48 % in the whole data set).

Summary of map history:

- SOM settings and data pre processing SOM setting and pre processing were similar to two other studies except that training was by default set to 39 cycles and map ratio was set to principal plane 100:77 ratio.
- Map history description lasted 26 minutes and final errors were: Normalized distortion 0.007716 and Quantization error 0.0003153.

The resulting first generation map Figure 3, was only labelled with stared (*) major MESH. The statistic of average data for each cluster was obtained, and the original database was value-added, classifying documents belonging to each cluster with its Cluster Code. Groups of documents (cluster-coded) were created in the database. Preliminary results at this stage shows how this domain has evolved since 1993 with work mainly in subjects covered by cluster M4, followed in 1994 with cluster M1 and then in 1995 with subjects covered in cluster M2. Finally during last years since 1996 till 2000, main research was conducted in subjects included in cluster M5.

A second generation of maps was obtained for each cluster (not included in this paper). For a second generation of clusters, the method is the following: for each cluster, nodes are filtered and with resulting average data for each node, other ANN is trained in similar conditions as described above. Again members of each cluster are identified and database classified, adding value to it for further analysis.

The procedure will conclude when the whole Domain Features Clustering Schedule is completed. This is an ongoing work in the frame of analysis where database is enhanced. A final map GENERAL-DOMAIN-FEATURES will result by training an ANN with data from a matrix of All Clusters.

The former approach outcomes are two tools. On one hand, a set of maps aiding clients (analysts) to get insight to data, and on the other hand, a searchable value-added database where original data could be found. In a way, this consists on an enhancement of MOBIS-Pro-Soft methodology approach for small and medium size corpora presented during the last 7. ISSI Conference (Sotolongo-Aguilar et al 2000).

Conclusions

Mining informetric data with self-organizing maps is useful for eliciting qualitative features from data. The wealth of tables and matrixes obtained during regular bibliometric or scientometric studies could be used as input for training ANN with the purpose of obtaining a better insight to data. Data analysis is enriched due to the fact that data could be explored to obtain models from it, instead of pre-establishing the models of data. Visual exploration enriches analysis while help end-user to better understand features highlighted. Further research should be conducted for validation of results.

Bibliography

Alonso-Gamboa, Octavio. (2000) Iberoamérica construye LATINDEX: sistema regional de información en línea para revistas científicas. Boletín Infolac (accepted for publication).

CRISP-DM 1.0 Step-by-step data mining guide. (2000) CRISP-DM consortium.

Deboeck, G. Public Domain versus Commercial Tools for creating Neural Self-organizingmaps. (1999) PCAI Magazine.

Deboeck, G. Data Mining with self-organizing maps: Part 1: Main Steps (submitted to PCAI Magazine).

L. Egghe. (1990) Applications of the Theory of Bradford's Law to the Calculations of Leimkuhler's Law and to the Completion of Bibliographies. Journal of the American Society for Information Science, 41(7): 469-492.

Fayyad, Usama M. 1997Editorial . datama mining and Knowledge Discovery., 1 (1):1-6

Otlet, P El Tratado de Documetación. El libro sobre el libro. Teoría y Práctica. M. D. Ayuso (translator) Universidad de Murcia, España, 431 p.

Red Iberoamericana de Indicadores de Ciencia y Tecnología (RICYT). (1999) Indicadores de Ciencias y Tecnología. Iberoamericanos/Interamericanos (1990-1997). Buenos Aires: REDES A.C.

Skupin, A. (2000) From Metaphor to Method: Cartographic Perspectives on Information Visualization. International Symposium on Information Visualization, October 9-10, Salt Lake City.

Saavedra-Fernández, O., Sotolongo-Aguilar, G. (1999) Medición de la producción científica en América Latina en el campo agrícola y afines: un estudio bibliométrico. XII RIBDA, Coronado Costa Rica.

Saavedra-Fernández, O., Sotolongo-Aguilar, G. (in preparation, 2001) Mapeo autoorganizado de las revistas científica de America Latina y el Caribe

Sotolongo-Aguilar, G.R. , Suárez-Balseiro, C. A. , Guzmán-Sánchez, M.V. (2000) Modular Bibliometric Information System with Proprietary Software (MOBIS-ProSoft): a versatile approach to bibliometric research tools. LIBRES: Library and Information Science Electronic Journal;10(2), http://libres.curtin.edu.au/ (September 30)

Skupin, A. (2000) From Metaphor to Method: Cartographic Perspectives on Information Visualization. International Symposium on Information Visualization 2000, October 9-10, Salt Lake City.

Std. AFNOR XP X 50-053-FRE. , Prestation de veille et prestation de mise en place deún systeme de veille. 22 p. 1998m

Vessuri, Hebe M.V. (1987) La revista científica periférica. El caso de Acta Científica Venezolana. Interciencia.;12(3):124-134.

Viscovery SOMine. (1999) Enterprise Edition Version 3.0, Eudaptics Software Gmbh.

Table 1 Cluster´s statistics summary of countries according to general and specific indicators relates to CAB data (Figures 1 and 2)

No.	Indicators	C 1: panama, el salvador, trinidad y tobago, barbados, jamaica, costa rica, guatemala, ecuador, uruguay, rep dominicana, bolivia, paraguay, peru, honduras, nicaragua	C 1*: guatemala, ecuador, uruguay, chile, rep dominicana, bolivia, paraguay, venezuela, peru, honduras, colombia, nicaragua	C1*/ C1	C 2: cuba, chile, venezuela, colombia	C 2*: panama, el salvador, trinidad y tobago, barbados, jamaica, costa rica, cuba	C2*/C2	C 3: argentina, mexico, brasil	C 3*: argentina, mexico, brasil
1	poblacion_millones	6,38	13,79	2,16	21,82	4,34	0,20	94,96	94,96
2	pea_millones	2,44	5,42	2,22	10,52	1,06	0,10	41,45	41,45
3	pbi_millones-U$D	12728,27	31653,58	2,49	60691,50	7692,43	0,13	450044,67	450044,67
4	gasto_CYT_millones-U$D	53,00	282,33	5,33	706,00	362,00	0,51	6828,67	6828,67
5	scicesearch	104,67	318,00	3,04	876,25	179,86	0,21	4968,00	4968,00
6	PASCAL	41,67	128,83	3,09	354,25	70,86	0,20	2250,67	2250,67
7	INSPEC	5,53	48,75	8,81	147,00	12,29	0,08	1451,33	1451,33
8	COMPENDEX	3,60	37,17	10,32	118,25	11,57	0,10	889,33	889,33
9	CA	22,33	123,83	5,54	421,00	76,14	0,18	2534,00	2534,00
10	BIOSIS	58,20	151,08	2,60	427,75	110,14	0,26	3353,67	3353,67
11	MEDLINE	24,60	82,83	3,37	230,75	42,57	0,18	1432,00	1432,00
12	CAB	44,33	109,92	2,48	381,00	124,29	0,33	1794,00	1794,00
13	ICYT	2,47	8,33	3,38	46,00	17,29	0,38	54,67	54,67
14	IME	4,00	8,67	2,17	29,50	10,57	0,36	42,00	42,00
15	Directorio_Latindex	21,50	49,55	2,30	150,75	63,20	0,42	522,33	522,33
16	CAB%mundial	0,03	0,07	2,53	0,25	0,08	0,32	1,19	1,19
17	**CAB_100mil_hab**	**1,73**	**0,78**	**0,45**	**2,40**	**3,74**	**1,56**	**2,07**	**2,07**
18	**CAB_pbi_mil_mill**	**8,08**	**3,45**	**0,43**	**12,15**	**18,34**	**1,51**	**104,17**	**104,17**
19	**CAB__gasto_i+d_mill_U$D**	**2,14**	**2,13**	**0,99**	**2,30**	**2,33**	**1,01**	**0,87**	**0,87**
20	**CAB_100_invest_ejc**	**8,61**	**4,67**	**0,54**	**5,43**	**12,15**	**2,24**	**5,10**	**5,10**

* Clusters with CAB´s data

Table 2 Cluster´s su,,ary statistics of Figure 3e - LA &C - with all indicators

No	Indicators	C 1: Peru	C 2: Costa Rica	C 3: Bolivia	C 4: R. Dominican	C 5: Barbados	C 6: Panama	C 7: Nicaragua,	C 8: Guatemala	C 9: Colombia	C 10: Jamaica	C 11: Brasil	C 12: Ecuador	C 13: Mexico	C 14: Uruguay	C 15: Cuba	C 16: Honduras	C 17: Venezuela	C 18: El Salvador	C 19: Trinidad y	C 20: Chile	C 21: Argentina
	Matching records	1	1	1	1	1	1	2	1	1	1	1	1	1	1	1	1	1	1	1	1	1
	Matching records (%)	4.55	4.55	4.55	4.55	4.55	4.55	9.09	4.55	4.55	4.55	4.55	4.55	4.55	4.55	4.55	4.55	4.55	4.55	4.55	4.55	4.55
1	Revistas (2000)	10	9	15	16	16	13	21	12	6	20	3	11	2	8	4	18	7	14	19	5	1
2	rev_scicesearch	10	6	10	10	10	10	10	10	6	10	3	10	1	10	6	10	3	10	6	2	3
3	rev_CA	9	4	1	9	9	9	9	9	4	9	9	9	4	9	1	9	1	9	9	4	4
4	rev_BIOSIS	6	5	12	12	12	12	12	12	6	12	1	12	2	12	3	6	6	12	6	6	3
5	rev_MEDLINE	6	2	6	6	6	6	6	6	6	6	1	6	2	6	6	6	2	6	6	2	6
6	rev_CAB	7	7	7	7	7	7	7	7	7	7	1	7	2	7	5	7	3	7	5	7	3
7	sci_promedio	11	9	16	18	17	12	20.5	15	7	6	1	14	3	10	8	19	5	22	13	4	2
8	pascal_promedio	9	7	15	21	17	14	18.5	16	4	11	1	13	2	10	8	20	5	22	12	6	3
9	inspec_promedio	10	7	17	14	15	16	20.5	17	4	12	1	13	2	8	11	19	5	22	9	6	3
10	compendex_promedio	10	7	16	21	14	14	18	17	5	13	1	12	2	9	11	19	4	22	8	6	3
11	ca_promedio	9	7	15	20	17	13	18.5	16	4	11	1	14	3	8	10	21	5	22	12	6	2
12	biosis_promedio	9	7	16	21	17	12	16	14	4	10	1	15	3	8	11	18	5	22	20	6	2
13	medline_promedio	9	7	17	21	16	13	18.5	14	4	8	1	15	2	10	11	20	5	22	12	6	3
14	cab_promedio	9	7	16	21	19	14	19	13	5	12	1	15	3	10	8	17	6	22	11	4	2
15	icyt_promedio	8	7	11	14	20	15	17.5	18	4	20	5	12	3	9	10	15	6	13	20	2	1
16	ime_promedio	11	8	15	14	21	16	17.5	12	4	21	5	9	3	7	13	19	6	10	19	2	1
17	sci_100milhab_promedio	6	12	9	10	14	15	20	20	7	18	13	22	8	2	3	11	5	16	1	17	4
18	pascal_100milhab_promedio	6	9	8	10	14	15	20	20	6	17	13	22	12	5	3	11	4	18	1	16	2
19	sci_pbimil_U$Db_promedio	8	4	13	10	12	17	20.5	16	3	11	18	22	6	1	2	14	7	15	5	19	9

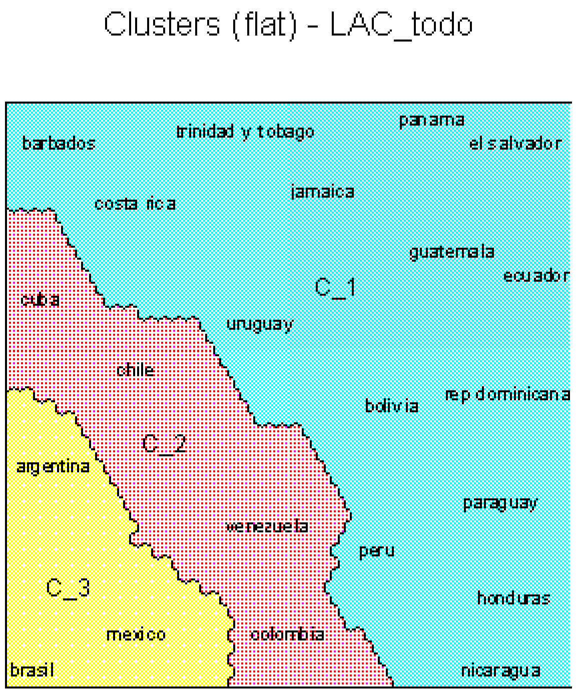


Figure 1 Latin America & Caribbean selected countries SOM according general input-output indicators

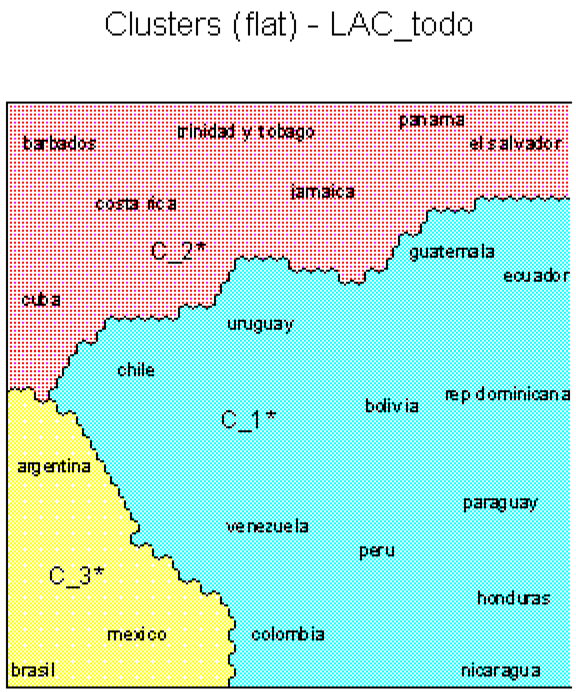
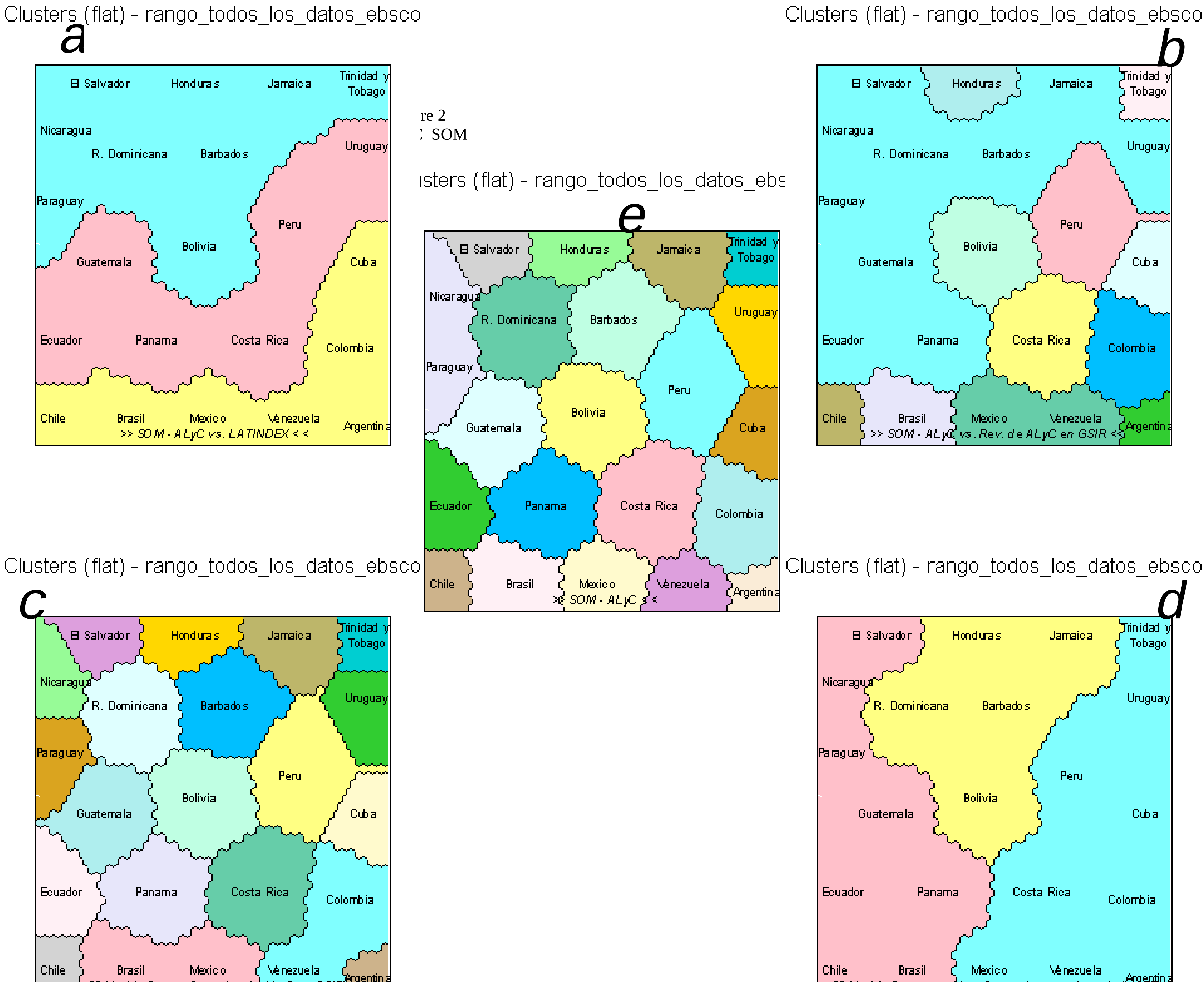


Figure 2 Latin America & Caribbean selected countries SOM according input-output indicators related to CAB data


re 2
SOM
Clusters (flat) - rango_todos_los_datos_ebsco
a
El Salvador
Honduras
Jamaica
Trinidad y Tobago
Nicaragua
R. Dominicana
Barbados
Uruguay
Paraguay
Peru
Bolivia
Guatemala
Cuba
Ecuador
Panama
Costa Rica
Colombia
Chile
Brasil
Mexico
Venezuela
Argentina
>> SOM - ALyC vs. LATINDEX < <
Clusters (flat) - rango_todos_los_datos_ebsco
b
El Salvador
Honduras
Jamaica
Trinidad y Tobago
Nicaragua
R. Dominicana
Barbados
Uruguay
Paraguay
Peru
Bolivia
Guatemala
Cuba
Ecuador
Panama
Costa Rica
Colombia
Chile
Brasil
Mexico
Venezuela
Argentina
>> SOM - ALyC vs. Rev. de ALyC en GSIR <<
e
El Salvador
Honduras
Jamaica
Trinidad y Tobago
Nicaragua
R. Dominicana
Barbados
Uruguay
Paraguay
Peru
Bolivia
Guatemala
Cuba
Ecuador
Panama
Costa Rica
Colombia
Chile
Brasil
Mexico
Venezuela
Argentina
>> SOM - ALyC <<
Clusters (flat) - rango_todos_los_datos_ebsco
c
El Salvador
Honduras
Jamaica
Trinidad y Tobago
Nicaragua
R. Dominicana
Barbados
Uruguay
Paraguay
Peru
Bolivia
Guatemala
Cuba
Ecuador
Panama
Costa Rica
Colombia
Chile
Brasil
Mexico
Venezuela
Argentina
Clusters (flat) - rango_todos_los_datos_ebsco
d
El Salvador
Honduras
Jamaica
Trinidad y Tobago
Nicaragua
R. Dominicana
Barbados
Uruguay
Paraguay
Peru
Bolivia
Guatemala
Cuba
Ecuador
Panama
Costa Rica
Colombia
Chile
Brasil
Mexico
Venezuela
Argentina

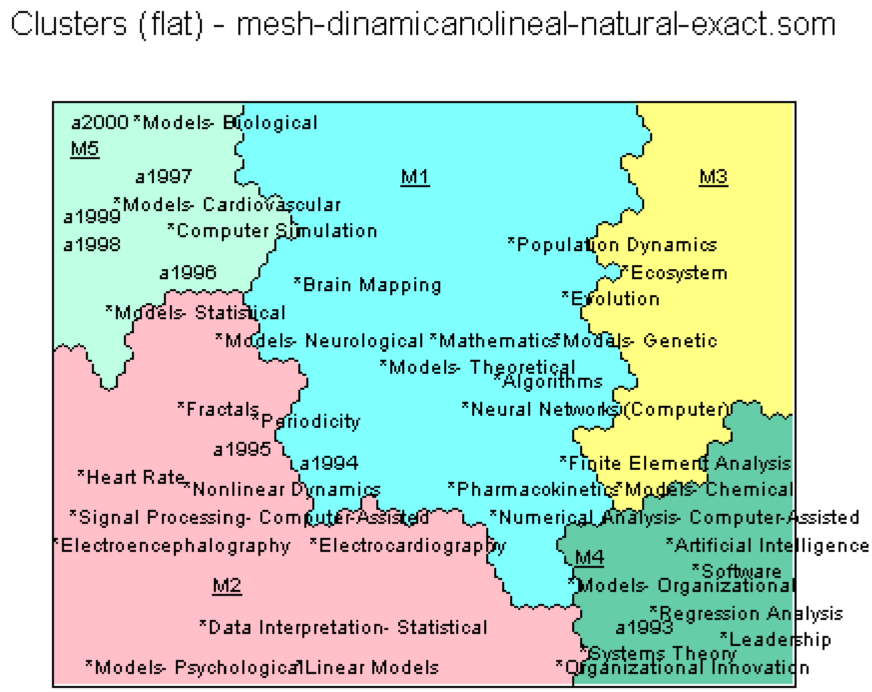


Figure 3 Non Linear Dynamics Domain SOM

Leyend
M 1: (1224 documents)
*Brain Mapping, *Models- Neurological, *Mathematics, *Models- Theoretical, *Algorithms, *Neural Networks (Computer), *Periodicity, a1994, *Pharmacokinetics
M 2: (1374 documents)
*Models- Statistical, *Fractals, a1995, *Nonlinear Dynamics, *Heart Rate, *Signal Processing- Computer-Assisted, *Electroencephalography, *Electrocardiography, *Data Interpretation- Statistical, *Models- Psychological, *Linear Models
M 3: (549 documents)
*Population Dynamics, *Ecosystem, *Evolution, *Models- Genetic, *Finite Element Analysis
M 4: (352 documents)
*Models- Chemical, *Numerical Analysis- Computer-Assisted, *Artificial Intelligence, *Software, *Regression Analysis, a1993, *Leadership, *Models- Organizational, *Systems Theory, *Organizational Innovation
M 5: (1884 documents)
a2000, *Models- Biological, a1997, *Models- Cardiovascular, *Computer Simulation, a1999, a1998, a1996