PINV18-661

Parallel Evolutionary Biclustering of Short-Term Electric Energy Consumption

Diego P. Pinto-Roa, Hernán Medina, Federico Román, Miguel García-Torres, Federico Divina, Francisco Gómez-Vela, Félix Morales, Gustavo Velázquez, Federico Daumas, José L. Vázquez Noguera, Carlos Sauer, Pedro Gardel
diego.pinto@ua.edu.py/dpinto@pol.una.py

CONSEJO NACIONAL DE CIENCIA Y TECNOLOGÍA
CONACYT

PROCIENCIA
PROGRAMA PARAGUAYO PARA EL DESARROLLO DE LA CIENCIA Y TECNOLOGÍA

UNIVERSIDAD AMERICANA

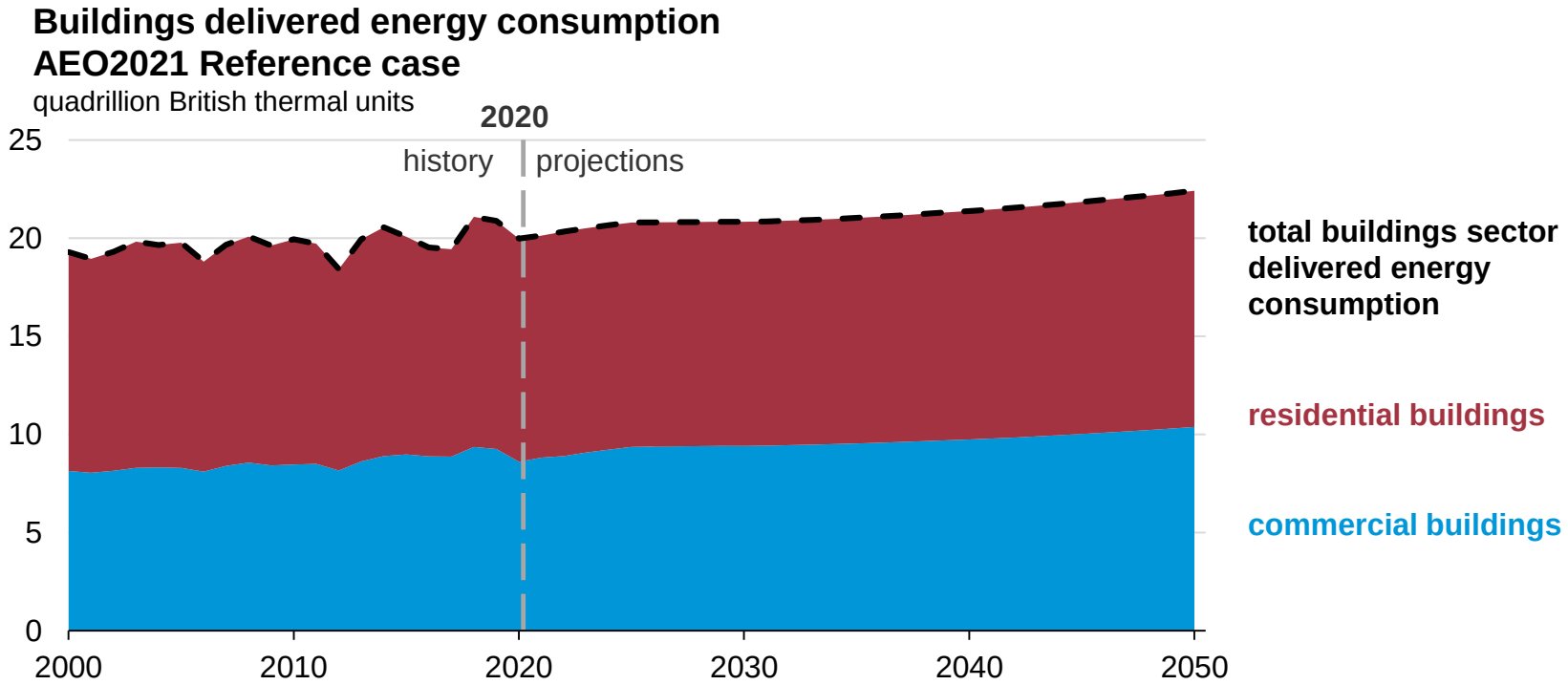
Con el apoyo de:
Feei
Fondo para la Excelencia de la Educación y la Investigación

Agenda

- Short-term electric energy consumption
- Bi-clustering
- Evolutionary Bi-clustering
- Parallel Evolutionary Bi-clustering
- Simulation
- Conclusion and Futures Works

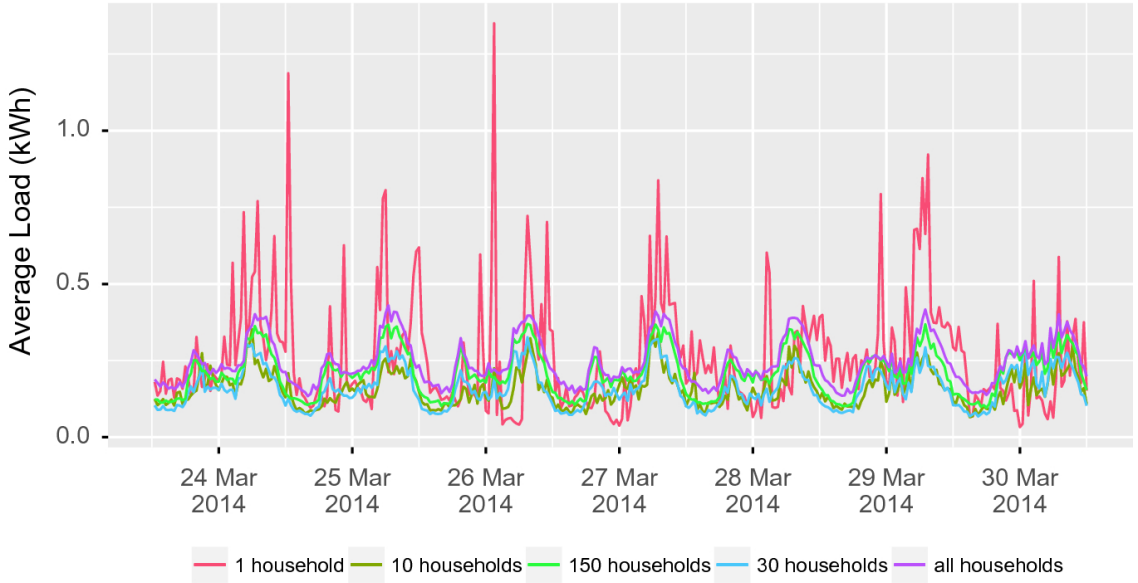
Short-term electric energy consumption

Energy consumption of buildings is receiving more and more attention in today's economies.



Anual Energy Outlook 2021

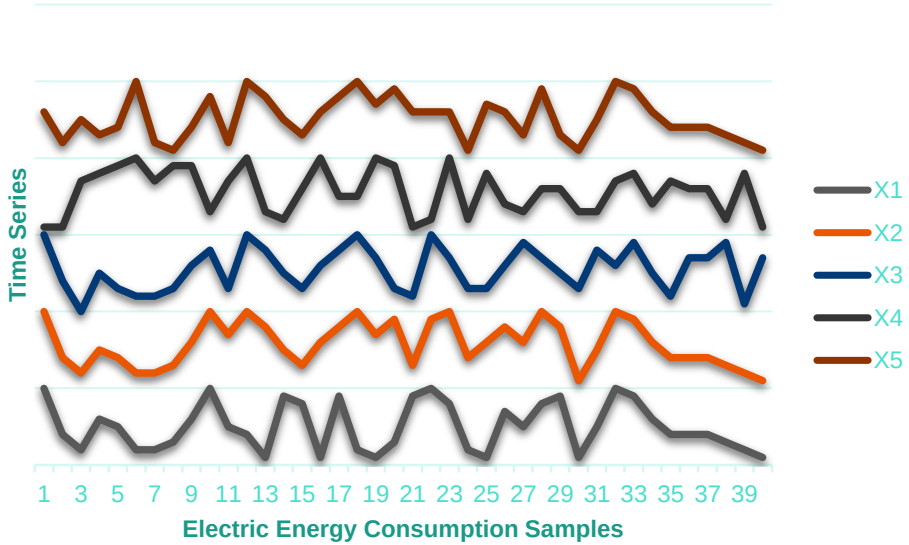
Short-term Electrical Energy Consumption



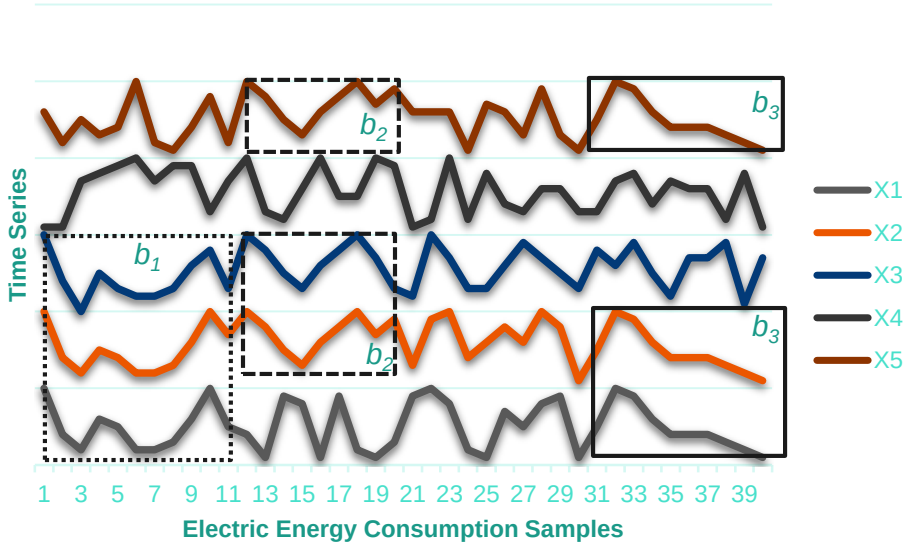
- The short-term load forecast is for a period of one day to one week.
- Short term load forecasting can help in estimating load flows and making decisions that can prevent over loading.

Jacob, M., Neves, C., & Greetham, D. V. (2020). Short term load forecasting. In *Forecasting and Assessing Risk of Individual Electricity Peaks* (pp. 15-37). Springer, Cham.

Short-term Electrical Energy Consumption

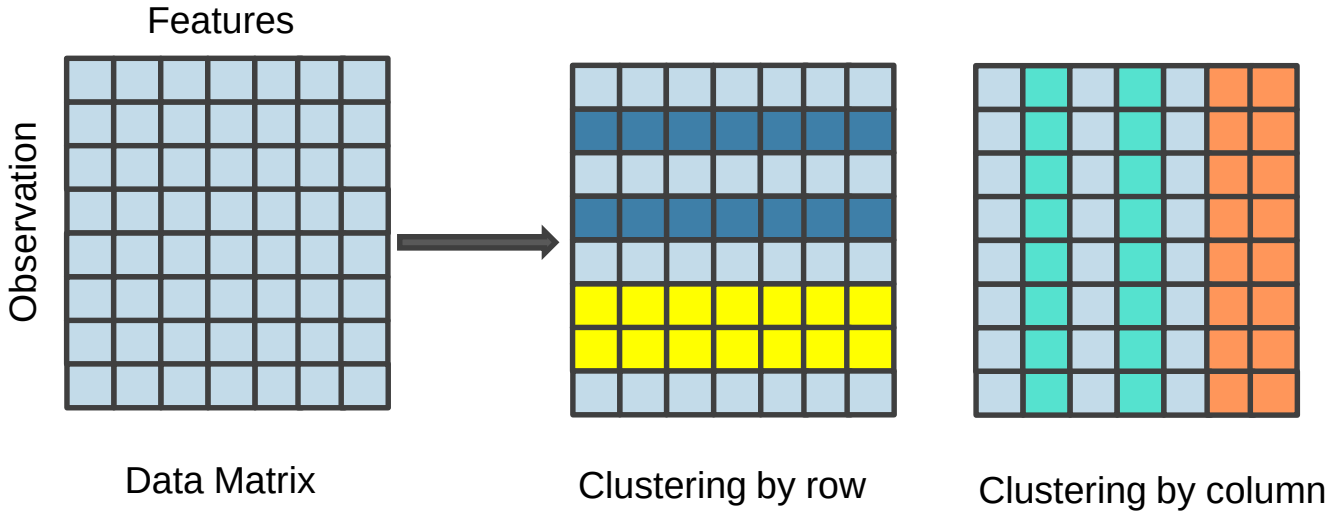


Short-term Electrical Energy Consumption



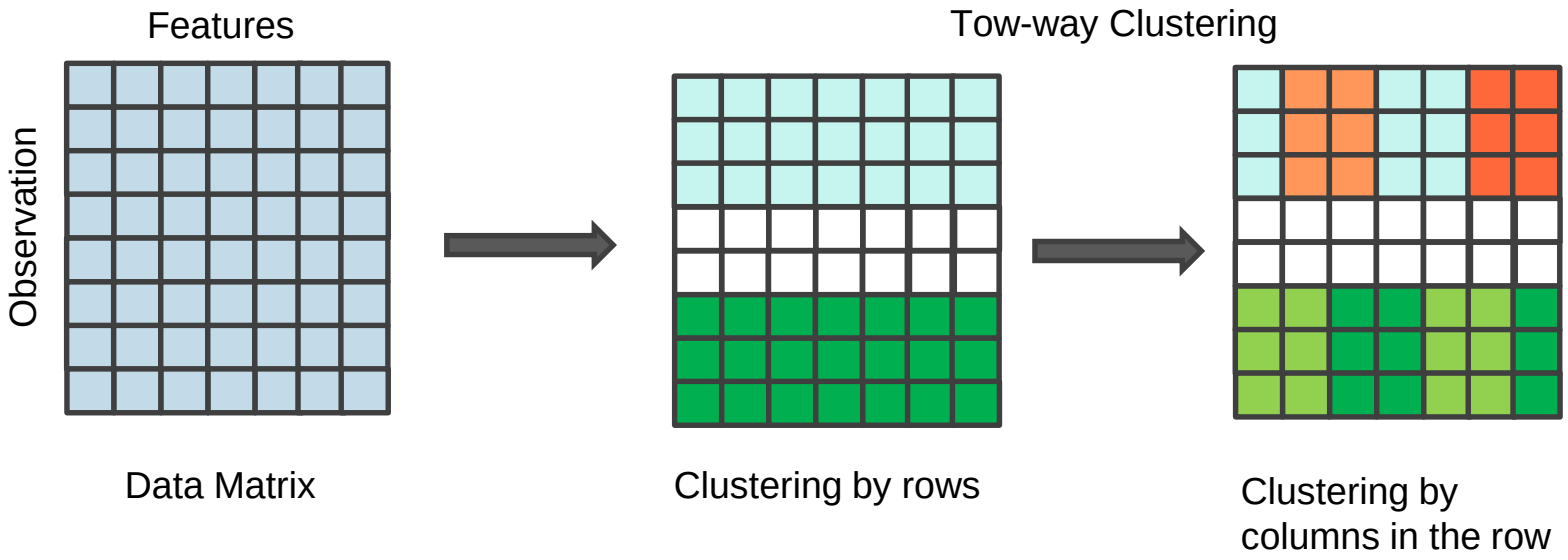
Bi-clustering
is a complex task of Machine Learning

Bi-Clustering vs. Clustering



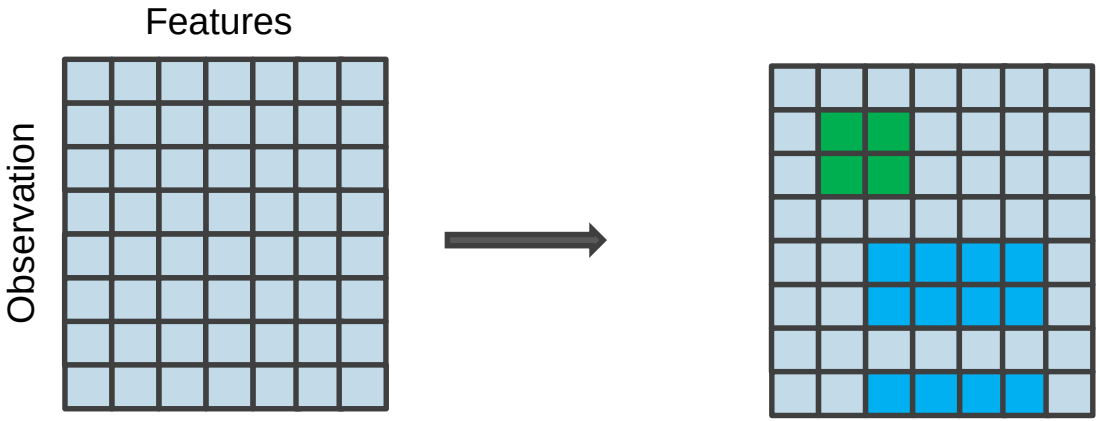
Kameshwaran, K., & Malarvizhi, K. (2014). Survey on clustering techniques in data mining. *International Journal of Computer Science and Information Technologies*, *5*(2), 2272-2276.

Bi-Clustering vs. Clustering



Kameshwaran, K., & Malarvizhi, K. (2014). Survey on clustering techniques in data mining. *International Journal of Computer Science and Information Technologies*, *5*(2), 2272-2276.

Bi-Clustering vs. Clustering



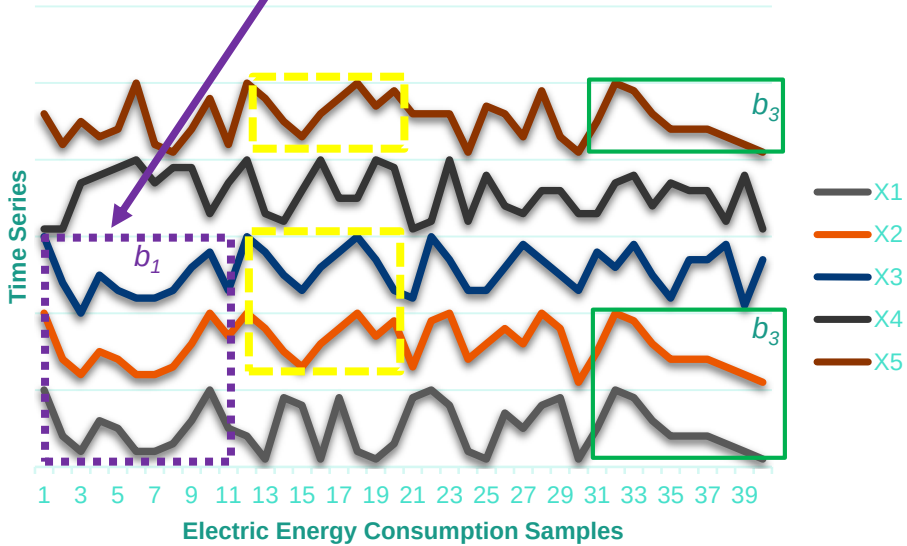
Goal of bi-clustering is Identify homogeneous submatrices

Hartigan, J. A. (1972). Direct clustering of a data matrix. *Journal of the american statistical association*, *67*(337), 123-129

Bi-cluster of a Short-term Electrical Energy Consumption

Data Matrix

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32	33	34	35	36	37	38	39	40
X5	46	42	45	43	44	50	42	41	44	48	42	50	48	45	43	46	48	50	47	49	46	46	46	41	47	46	43	49	43	41	45	50	49	46	44	44	44	43	42	41
X4	31	31	37	38	39	40	37	39	39	33	37	40	33	32	36	40	35	35	40	39	31	32	40	32	38	34	33	36	36	33	33	37	38	34	37	36	36	32	38	31
X3	30	24	20	25	23	22	22	23	26	28	23	30	28	25	23	26	28	30	27	23	22	30	27	23	23	26	29	27	25	23	28	26	29	25	22	27	27	29	21	27
X2	20	14	12	15	14	12	12	13	16	20	17	20	18	15	13	16	18	20	17	19	13	19	20	14	16	18	16	20	18	11	15	20	19	16	14	14	14	13	12	11
X1	10	4	2	6	5	2	2	3	6	10	5	4	1	9	8	1	9	2	1	3	9	10	8	2	1	7	5	8	9	1	5	10	9	6	4	4	4	3	2	1

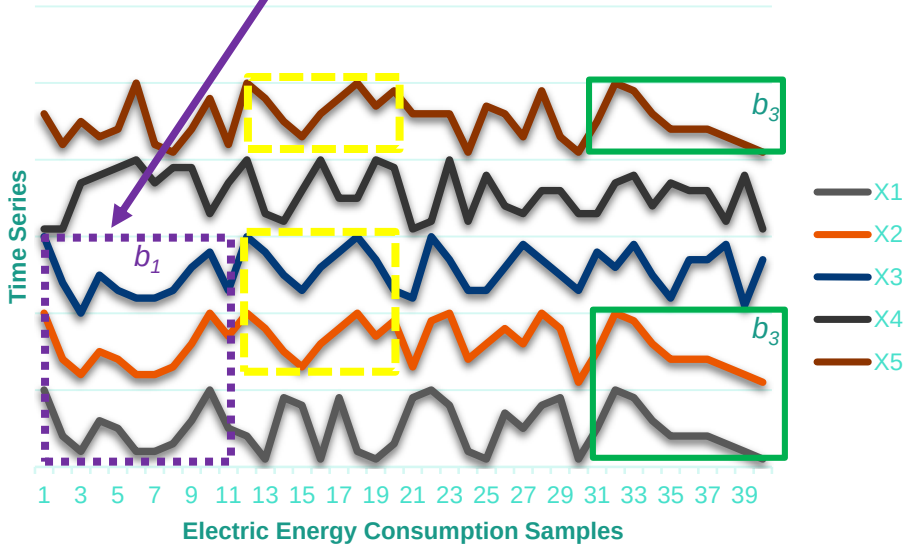


Bi-cluster = Sub-Matrix with contiguous columns that enclosed a consumption patter

Bi-cluster of a Short-term Electrical Energy Consumption

Data Matrix

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32	33	34	35	36	37	38	39	40
X5	46	42	45	43	44	50	42	41	44	48	42	50	48	45	43	46	48	50	47	49	46	46	46	41	47	46	43	49	43	41	45	50	49	46	44	44	44	43	42	41
X4	31	31	37	38	39	40	37	39	39	33	37	40	33	32	36	40	35	35	40	39	31	32	40	32	38	34	33	36	36	33	33	37	38	34	37	36	36	32	38	31
X3	30	24	20	25	23	22	22	23	26	28	23	30	28	25	23	26	28	30	27	23	22	30	27	23	23	26	29	27	25	23	28	26	29	25	22	27	27	29	21	27
X2	20	14	12	15	14	12	12	13	16	20	17	20	18	15	13	16	18	20	17	19	13	19	20	14	16	18	16	20	18	11	15	20	19	16	14	14	14	13	12	11
X1	10	4	2	6	5	2	2	3	6	10	5	4	1	9	8	1	9	2	1	3	9	10	8	2	1	7	5	8	9	1	5	10	9	6	4	4	4	3	2	1



b_1

	1	2	3	4	5	6	7	8	9	10
X1	10	4	2	6	5	2	2	3	6	10
X2	20	14	12	15	14	12	12	13	16	20
X3	30	24	20	25	23	22	22	23	26	28

Virtual Error= 0.305
Row variance= 8.323
Volume= 30

b_2

	12	13	14	15	16	17	18	19
X2	20	18	15	13	16	18	20	17
X3	30	28	25	23	26	28	30	27
X5	50	48	45	43	46	48	50	47

Virtual Error= 0.555
Row variance= 5.1
Volume= 24

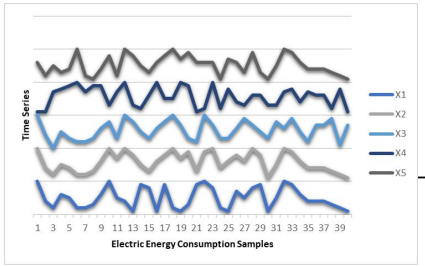
b_3

	30	31	32	33	34	35	36	37	38	39	40
X1	1	5	10	9	6	4	4	4	3	2	1
X2	11	15	20	19	16	14	14	14	13	12	11
X5	41	45	50	49	46	44	44	44	43	42	41

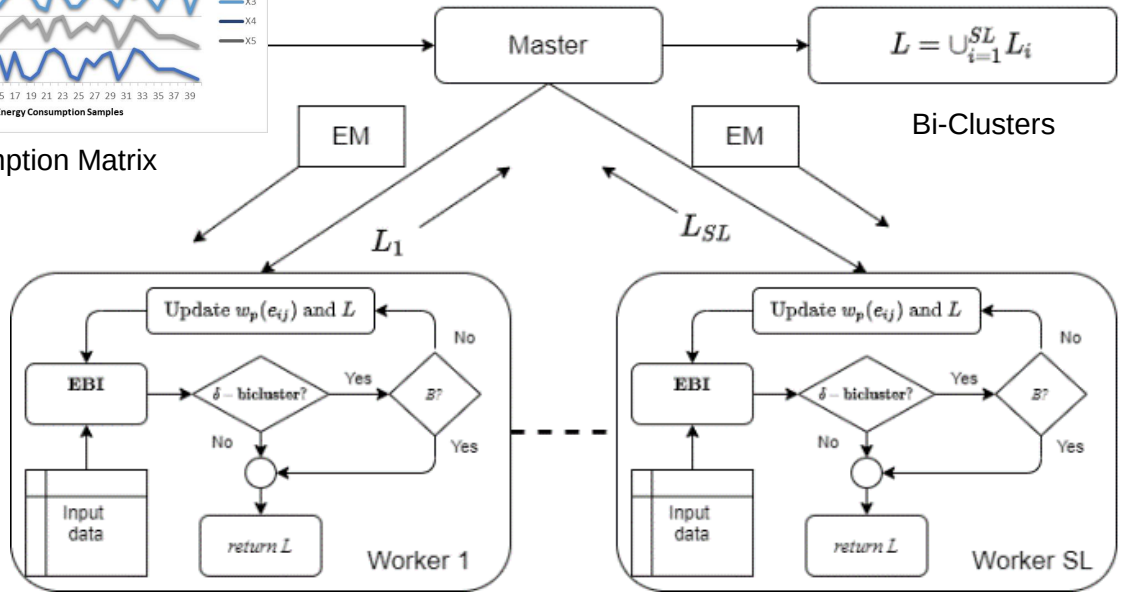
Virtual Error= 0.107
Row variance= 7.884
Volume= 33

Parallel Evolutionary Bi-cluster

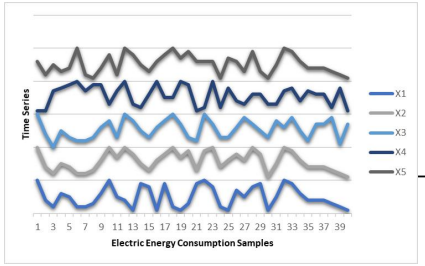
Parallel Search Covering – Master Slave Paradigm



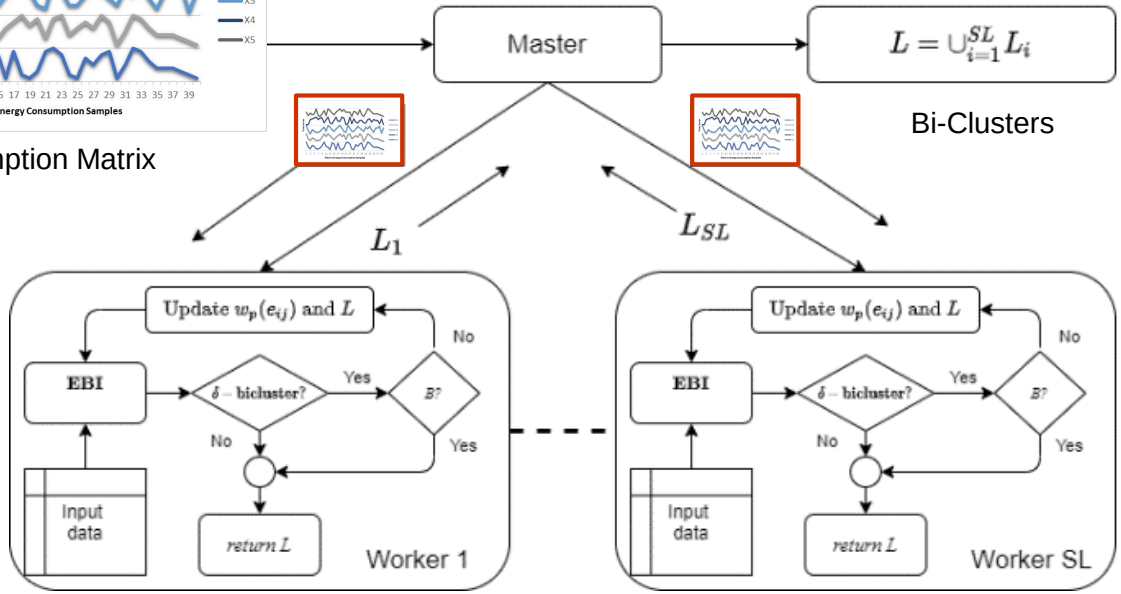
Consumption Matrix



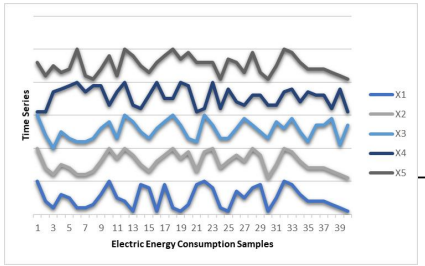
Parallel Search Covering



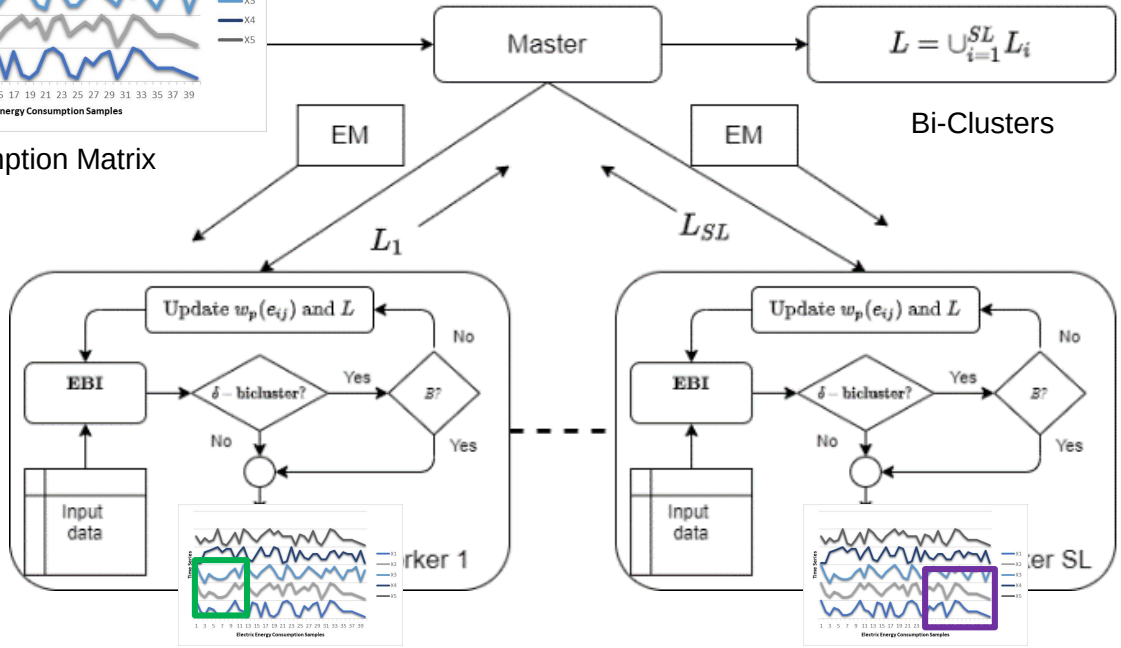
Consumption Matrix



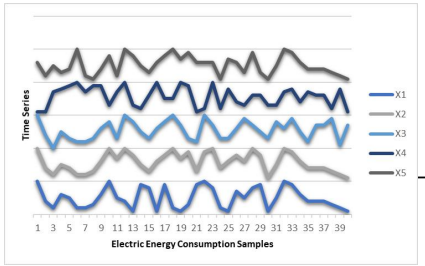
Parallel Search Covering



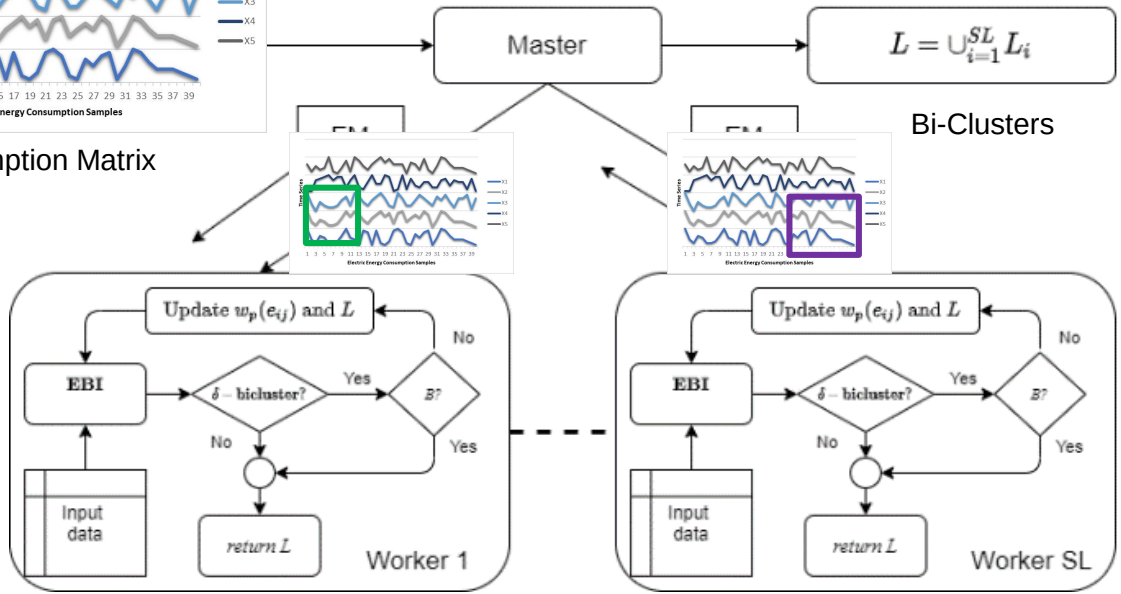
Consumption Matrix



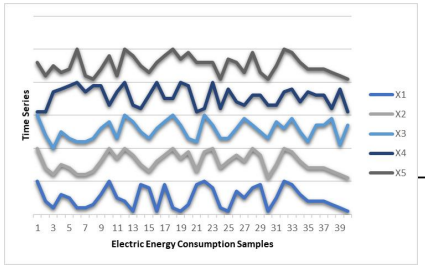
Parallel Search Covering



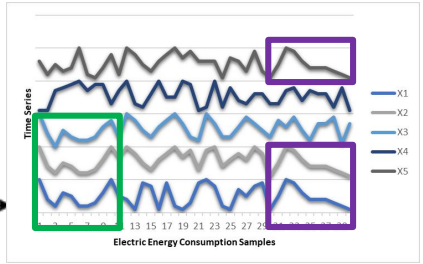
Consumption Matrix



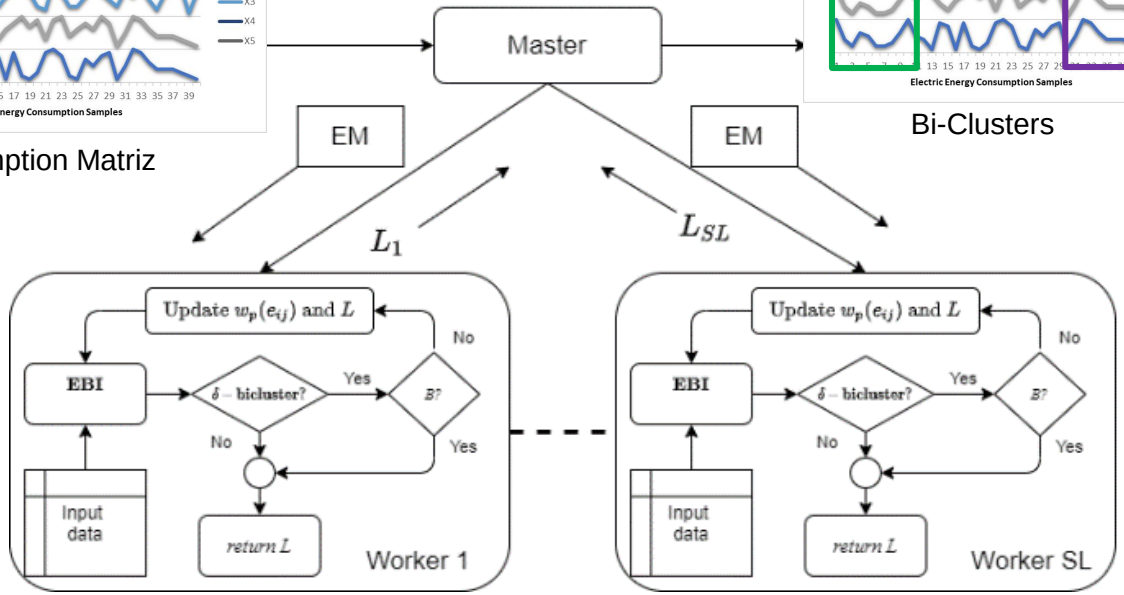
Parallel Search Covering



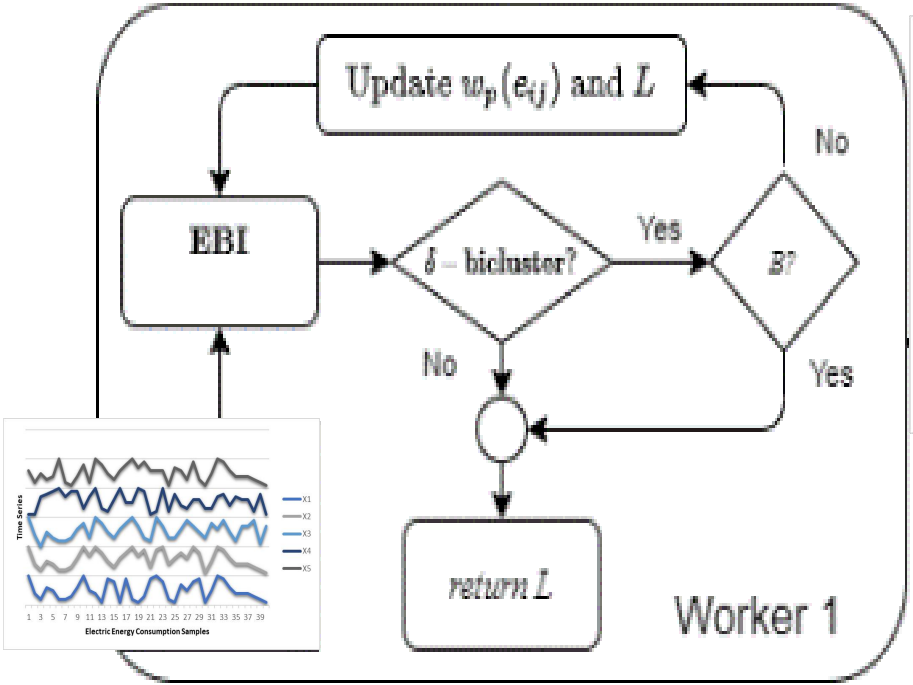
Consumption Matriz

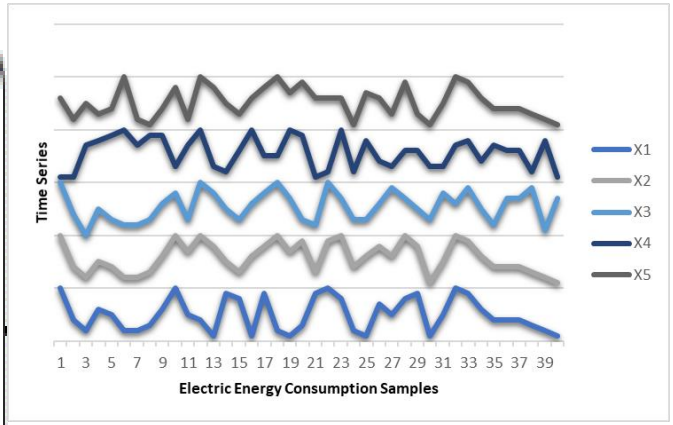


Bi-Clusters



Worker node
Sequential Evolutionary Bi-clustering (SEBI)

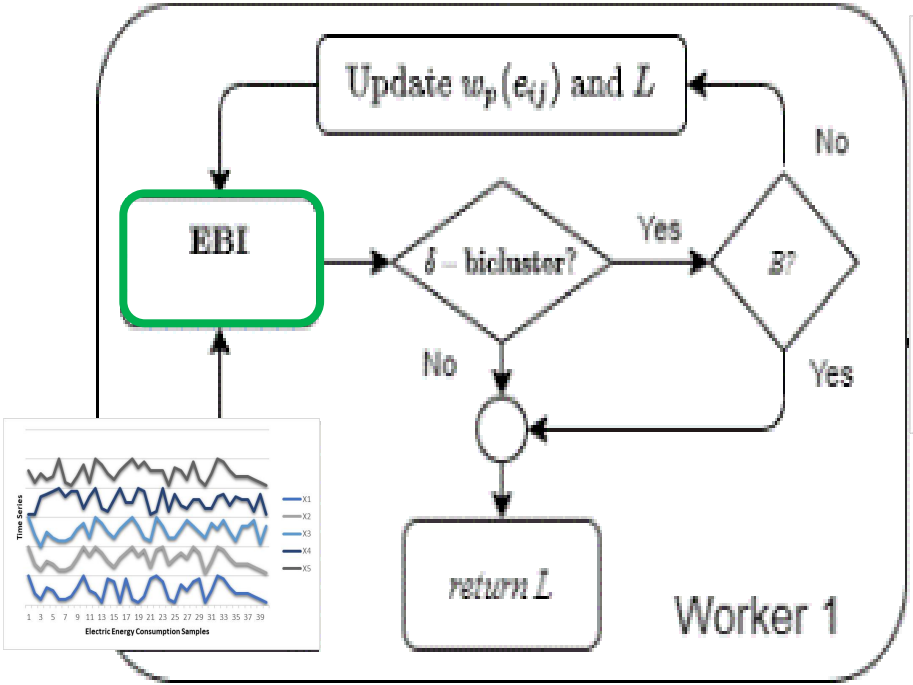


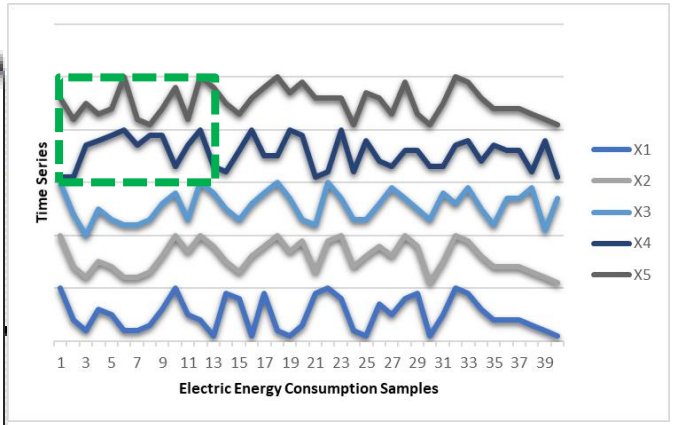


SEBI searches bi-clusters one-by-one
EBI calculates just one bi-cluster in each request.

F. Divina and J. S. Aguilar-Ruiz, "Biclustering of expression data with evolutionary computation," in *IEEE Transactions on Knowledge and Data Engineering*, vol. 18, no. 5, pp. 590-602, May 2006, doi: 10.1109/TKDE.2006.74.

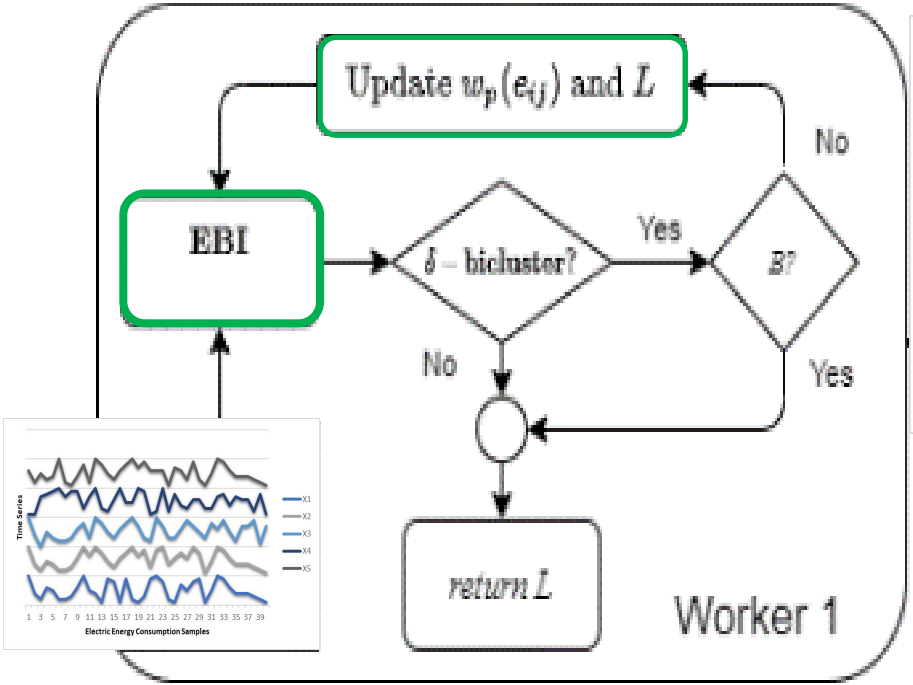
Worker node
Sequential Evolutionary Bi-clustering (SEBI)

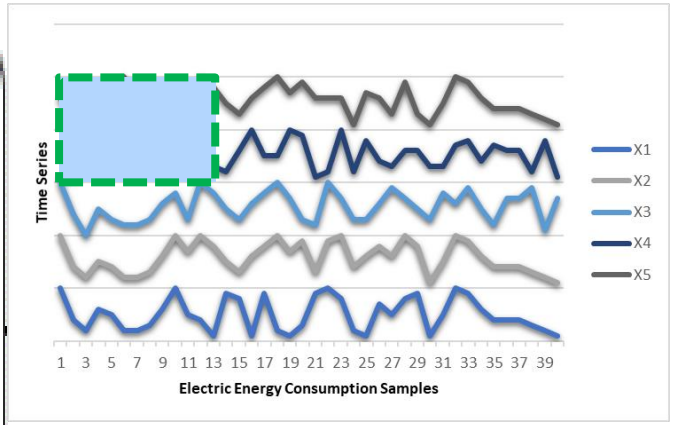




SEBI searches bi-clusters one-by-one
EBI calculates just one bi-cluster in each request.

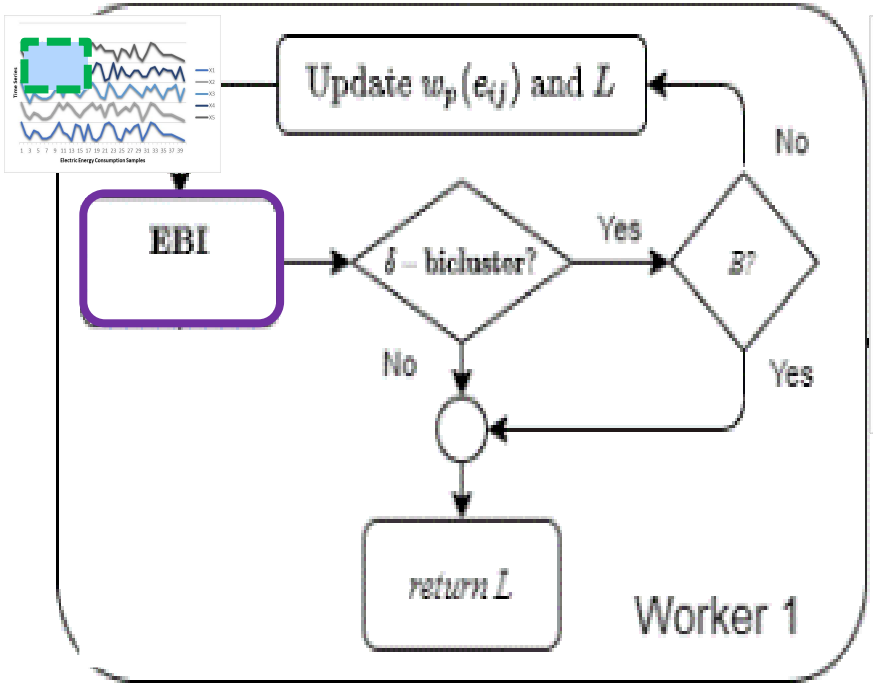
Worker node
Sequential Evolutionary Bi-clustering (SEBI)

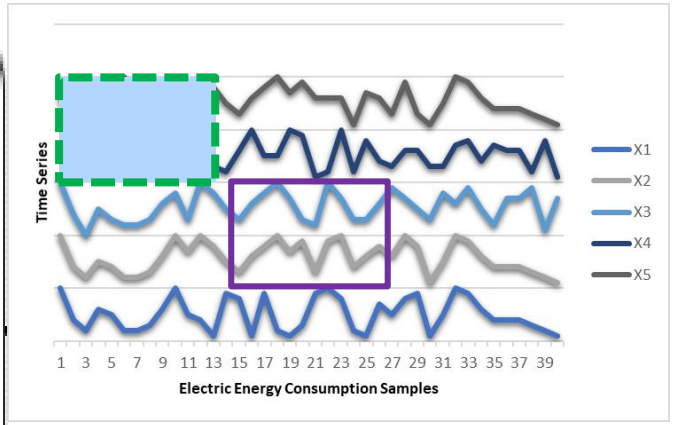




SEBI searches bi-clusters one-by-one
EBI calculates just one bi-cluster in each request.

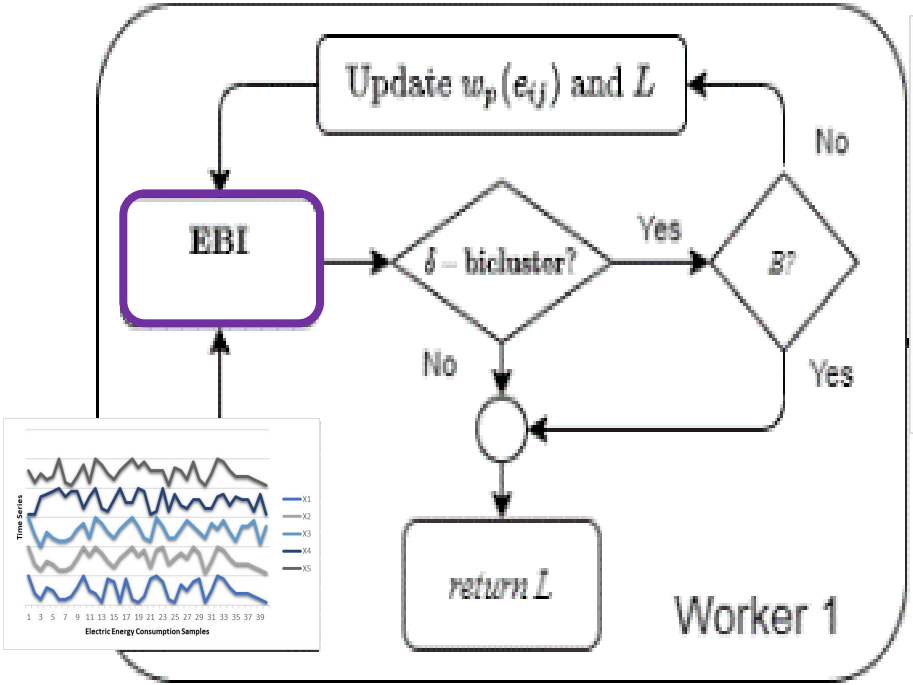
Worker node
Sequential Evolutionary Bi-clustering (SEBI)

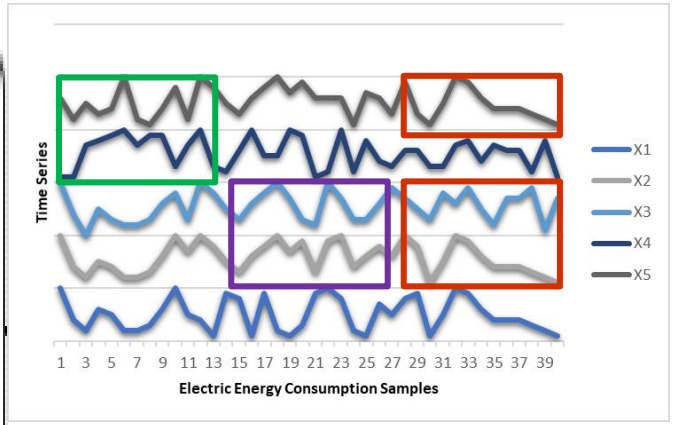




SEBI searches bi-clusters one-by-one
EBI calculates just one bi-cluster in each request.

Worker node
Sequential Evolutionary Bi-clustering (SEBI)



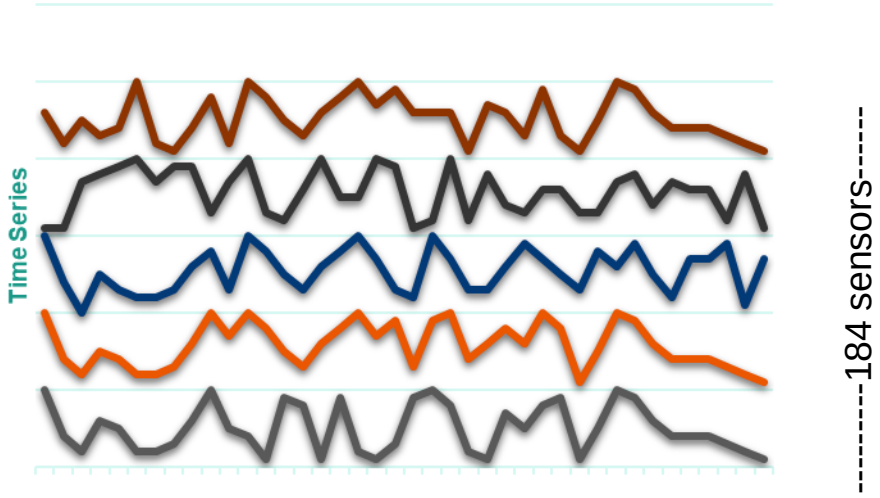


SEBI searches bi-clusters one-by-one
EBI calculates just one bi-cluster in each request.

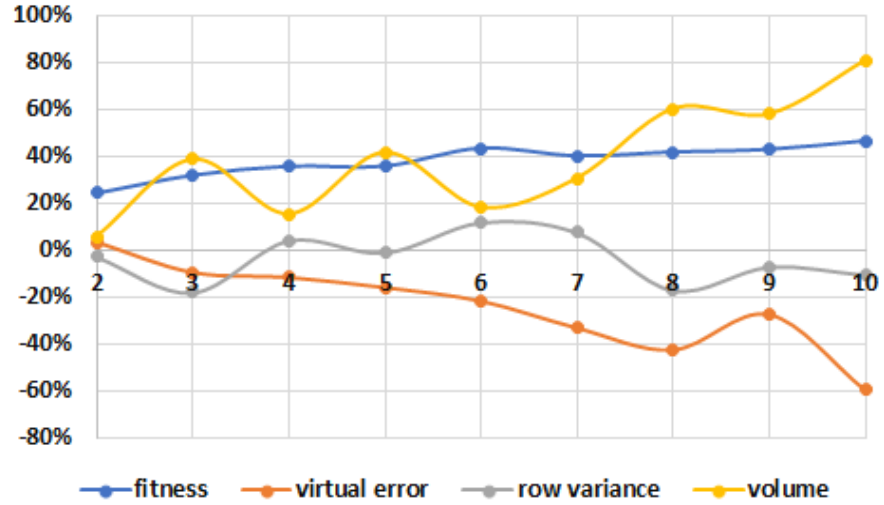
Simulations

Data Set

- This dataset contains 744 samples of electric energy consumption for 184 sensors gathered in December of 2013.



Performance vs Number of Workers/PC

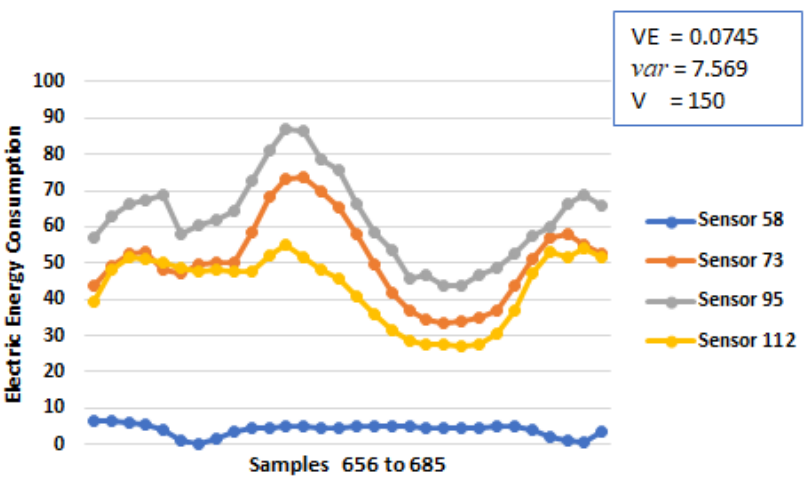


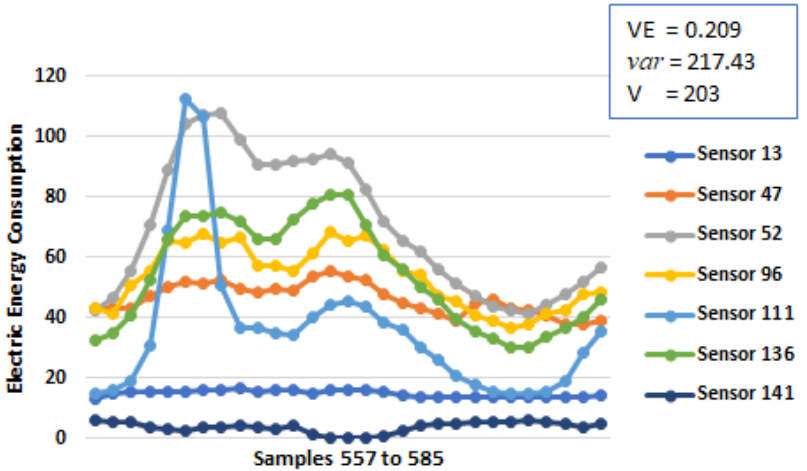
Relative performance of PSC over SC

	VE	Row Var	Volume
fitness	-0.72	0.07	-0.69
VE		-0.25	0.86
Row Var			-0.59

- Fitness improves when VE and Volume improve.
- Virtual error and Volume improves when K increases.

Some real bi-clusters calculated





The patterns show abrupt changes of sensor groups at different times of the month. Such descriptions of electricity consumption are crucial to energy policy or decision-making.

Conclusion and future works

Summary

- We have developed a parallel evolutionary algorithm for a biclustering problem.
- Biclustering helps to discover patterns of power consumption by describing which consumption points and time periods make up a pattern.
- Quality measures, virtual error, row variance and volume, help to find coherent and highly variable patterns of power consumption.

- We propose to extend to the multi-objective optimization approach
- We want to consider other parallelization paradigms such as island parallelization.
- We are interested in studying consumption patterns that are displaced in time.

¡Thank for your attention!

CONSEJO NACIONAL DE CIENCIA Y TECNOLOGÍA
CONACYT

PROCIENCIA
PROGRAMA PARAGUAYO PARA EL DESARROLLO DE LA CIENCIA Y TECNOLOGÍA

UNIVERSIDAD AMERICANA

Con el apoyo de:
Feei
Fondo para la Excelencia de la Educación y la Investigación