European Network on New Sensing Technologies for Air Pollution Control and Environmental Sustainability - *EuNetAir*
COST Action TD1105

Comparing Sensor Data with the Aid of Self Organizing Maps

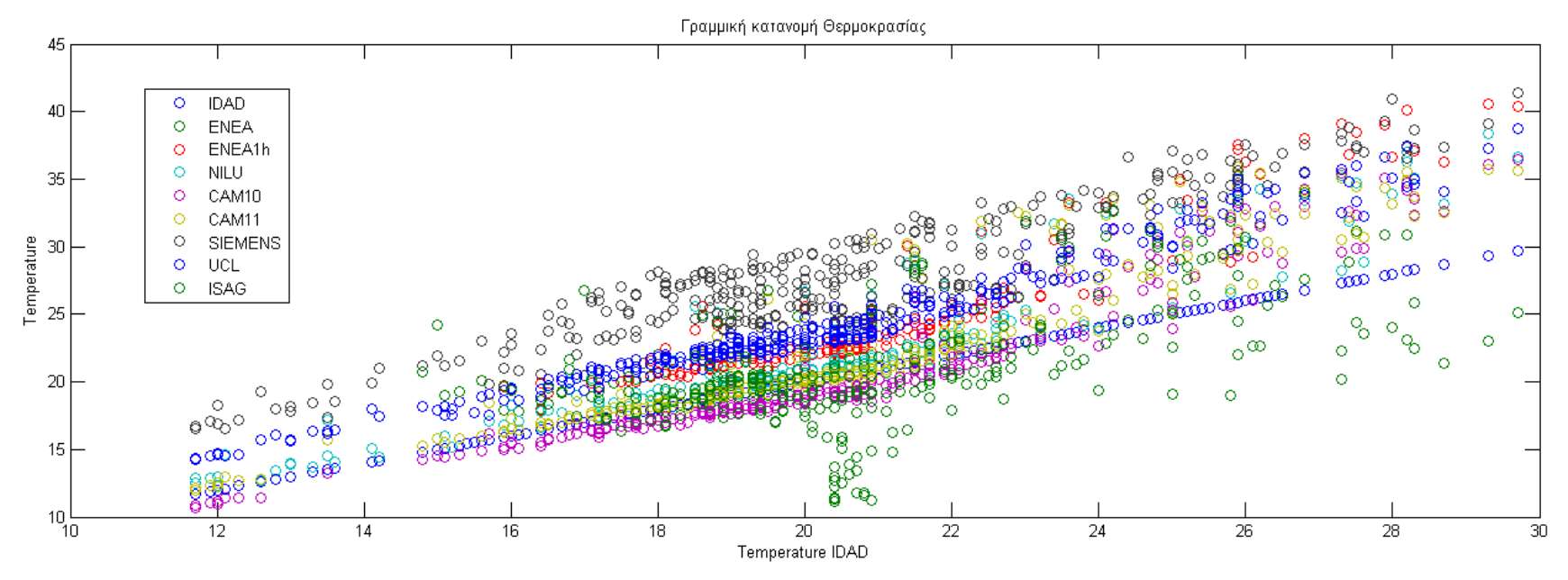
Kostas Karatzas
Informatics Systems and Applications – Environmental Informatics Research Group
Dept. of Mechanical Engineering
Aristotle University of Thessaloniki, Greece
kkara@eng.auth.gr, http://isag.meng.auth.gr

COST is supported by the EU Framework Programme
ESF provides the COST Office through a European Commission contract

Contents

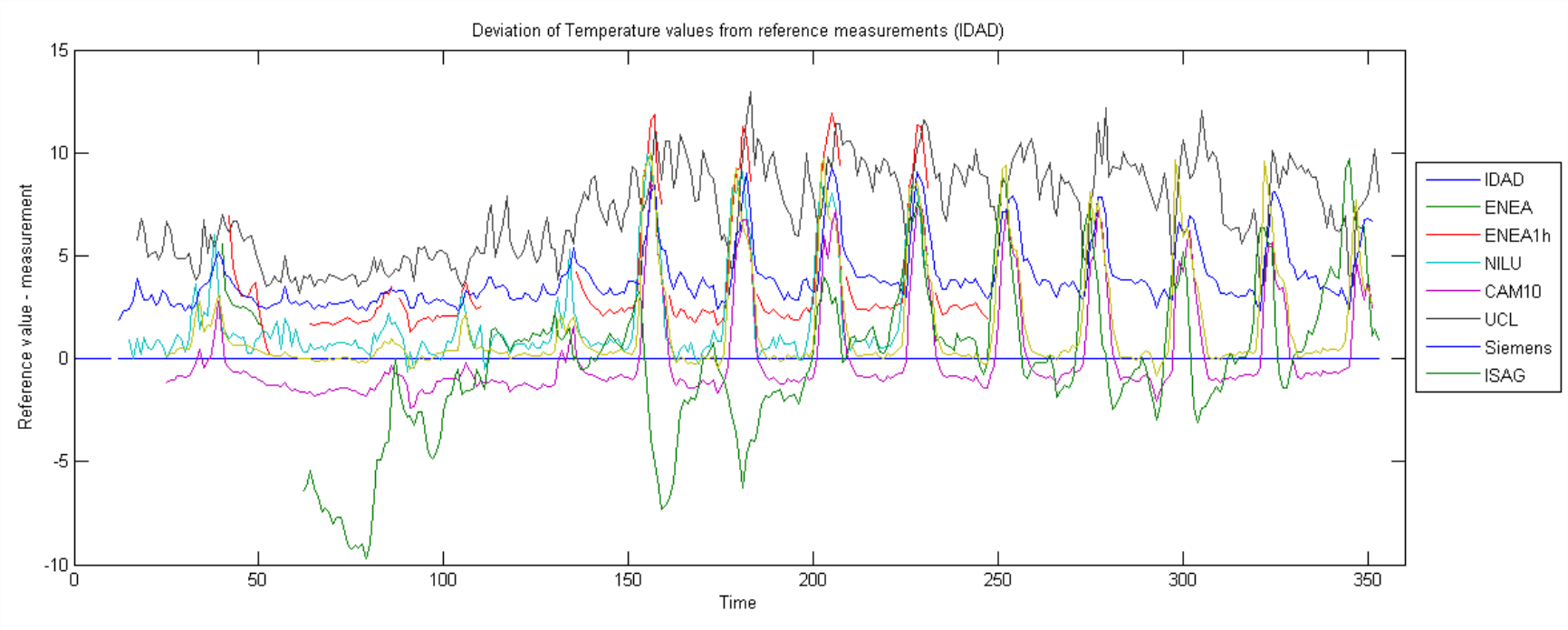
1. The need to compare sensor behavior
2. The SOM method
3. Indicative results from Aveiro DB analysis
4. First conclusions

1. The need to compare sensor behavior



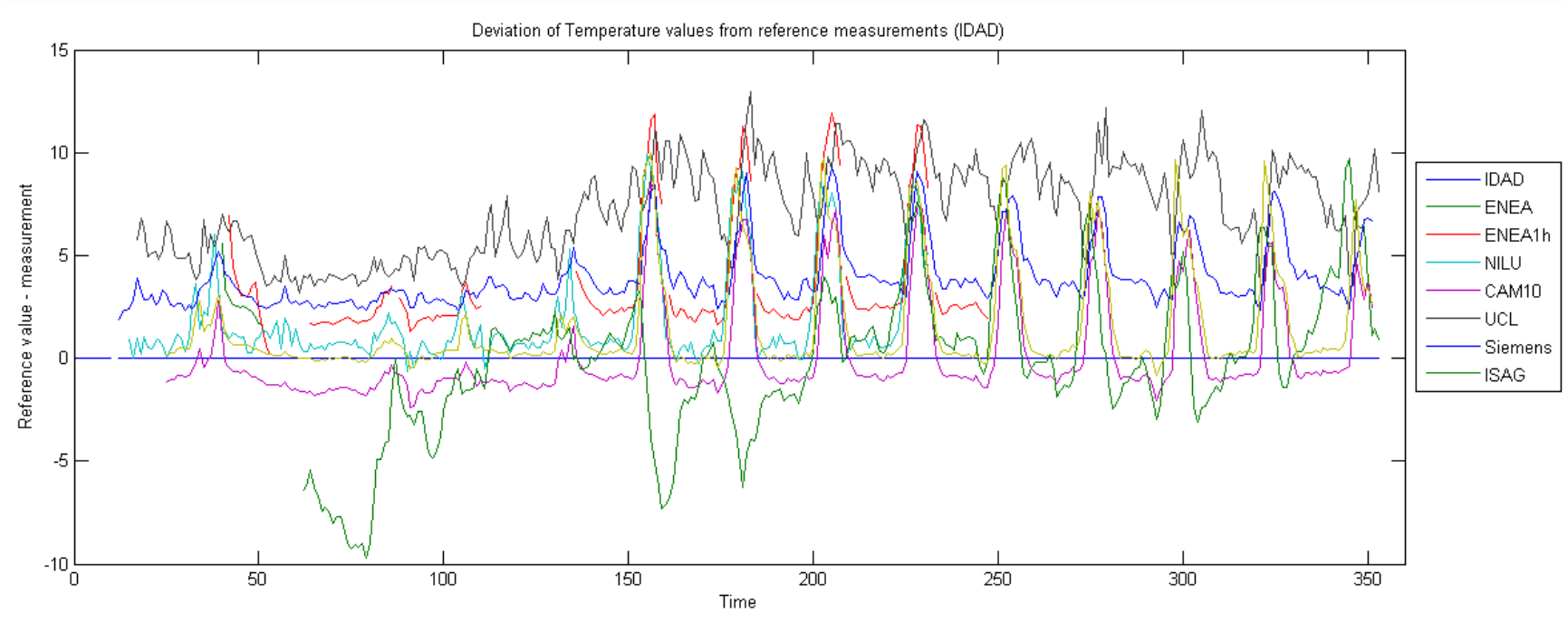
- Putting it all together is always a challenge

1. The need to compare sensor behavior



Our common starting point: differences and trends may become obvious but..

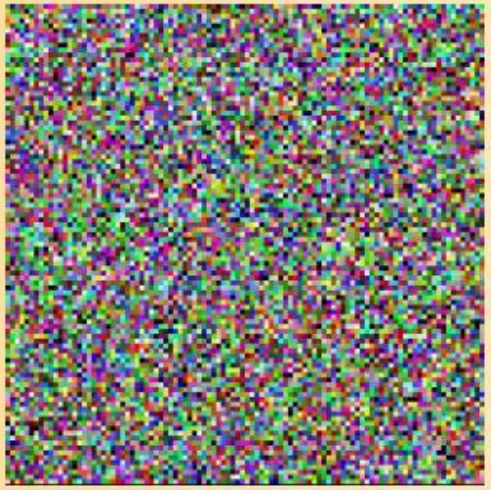
1. The need to compare sensor behavior



Our common starting point: differences and trends may become obvious but..

What about simultaneous, cross-parameter, behavior analysis?

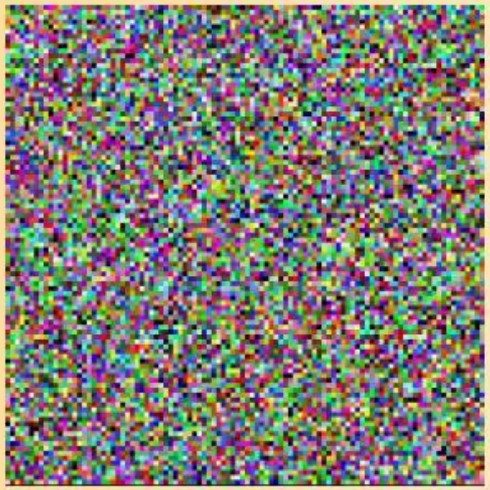
2. The SOM method



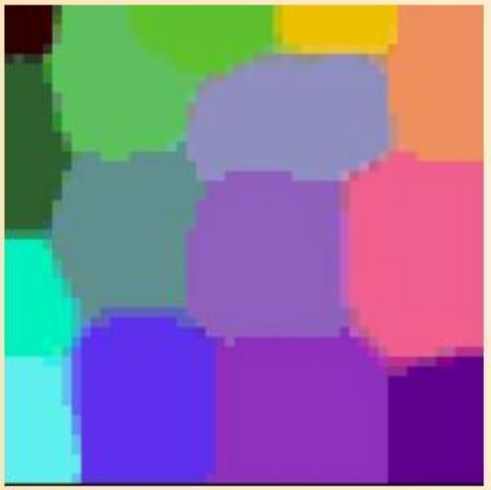
(Left) The initial stage of a Self Organization Map.

If there are relationships ("similarities" or "dependencies" within the data)

2. The SOM method



(Left) The initial stage of a Self Organization Map.

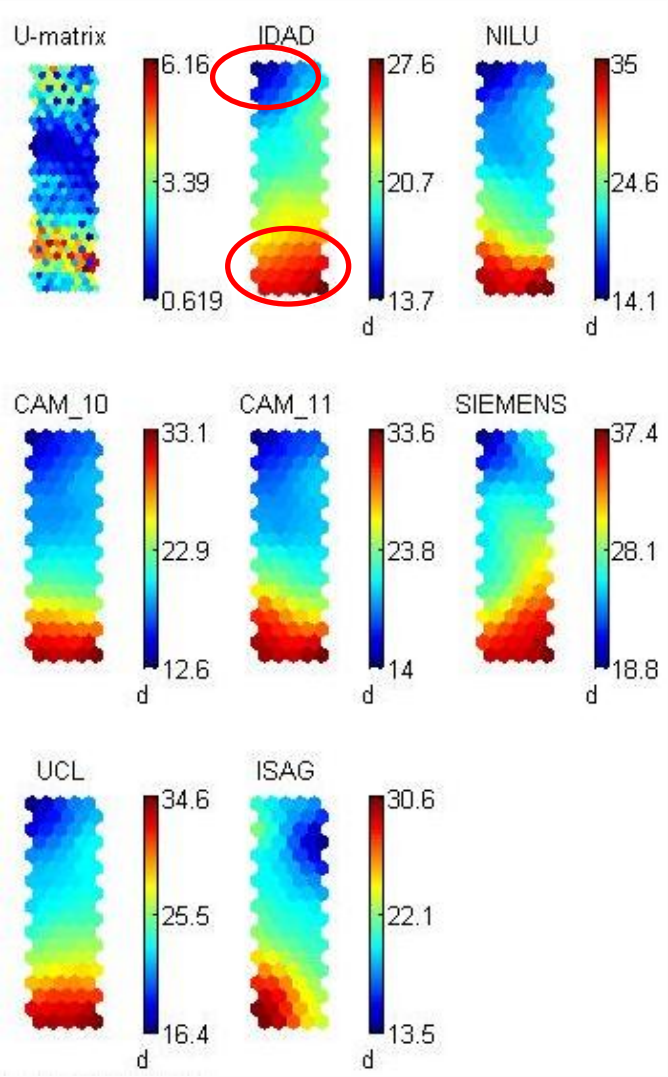


(Right) The final stage. Notice how similar colors are clustered together.

If there are relationships ("similarities" or "dependencies" within the data)

Can we identify them and create a profile?

3. Indicative results from Aveiro DB analysis

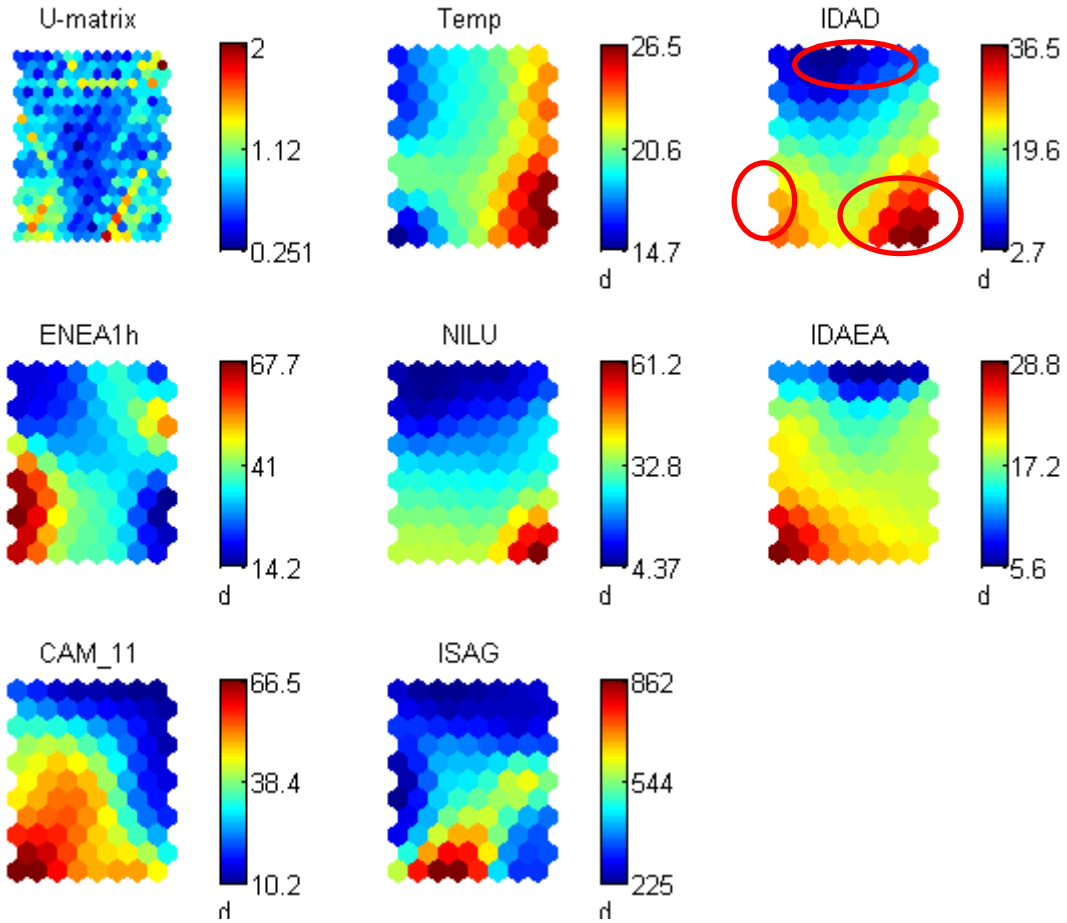


Temperature

Same areas in the SOM correspond to "similar" or "neighboring" measurement points

- Overall SOM pattern represents "real' pattern.
- All measurements in quite good agreement
- UCL and Siemens closer to IDAD

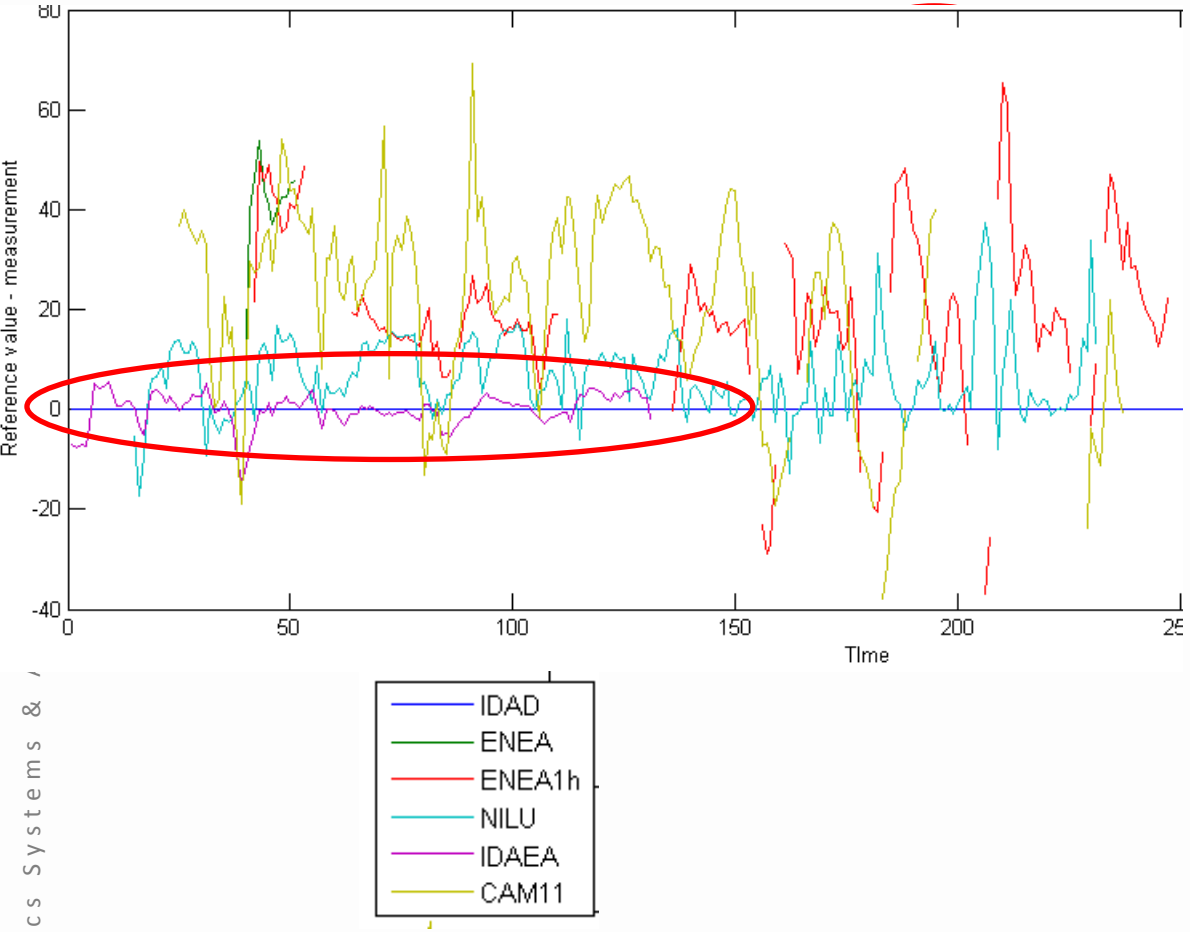
3. Indicative results from Aveiro DB analysis



O3

- Dependency from Temperature
- ENEA: high values in "wrong place", indicate sensor problems
- ISAG: wrong profile, major sensor problems.
- CAM_11: overestimations, lower sensitivity in lower values?
- IDAEA: lower sensitivity in lower values
- NILU: unbalanced behavior results in differences between higher and lower values. Overall best profile.

3. Indicative results from Aveiro DB analysis



O3

- Dependency from Temperature
- ENEA: high values in "wrong place", indicate sensor problems, incomplete dataset.
- ISAG: wrong profile, major sensor problems.
- CAM_11: overestimations, lower sensitivity in lower values~~?~~
- IDAEA: lower sensitivity in lower values due to incomplete dataset. **Overall best profile.**
- NILU: unbalanced behavior results in differences between higher and lower values. ~~Overall best profile.~~

4. First conclusions

- Profile analysis important forAQ sensor data
- An unsupervised, “inbiased” method necessary to be employed
- SOM-based analysis reveals patterns in behavior and relationships-interdependencies
- In combination with other graphs and math analysis may constitute a reference methodology for sensor data investigations

QoL information services!