ingeniería isa de sistemas y automática

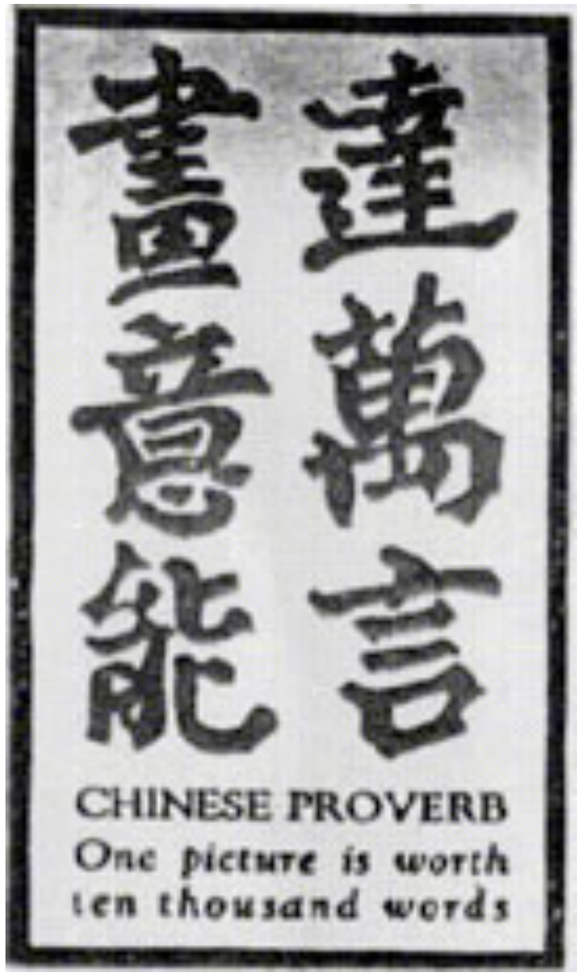
Universidad de Oviedo

A Visual Analytics approach for process analysis

Why visualization?

Why Visualization?

"A picture is worth a thousand words"



Fred R. Barnard

Printers' Ink. December 8, 1921

Why Visualization?

"Un bon croquis vaut mieux qu'un long discours"

Napoleon Bonaparte

"A picture shows me at a glance
what it takes dozens of pages of a book to expound."

Russian writer Ivan Turgenev
(in *Fathers and Sons* in 1862)

Why Visualization?

ingeniería isa de sistemas y automática

Universidad de Oviedo

I *see* it...

It is very *clear*

Show me why...

She *unveiled* the mystery

This *sheds* some *light* to the problem

This author is a *visionary*

Their explanation was rather *obscure*

He always says the truth: he is very *transparent* kid

Why Visualization?

ingeniería isa de sistemas y automática

Universidad de Oviedo

It is very *clear*

I *see* it...

Show me why...

She *unveiled* the mystery

IN-SIGHT

This *sheds* some *light* to the problem

This author is a *visionary*

Their explanation was rather *obscure*

He always says the truth: he is very *transparent* kid

Why Visualization?

40+ % of the brain dedicated to visual tasks

ingeniería isa de sistemas y automática

Why Visualization?

Universidad de Oviedo

How many “sevens” do you find?

1093840192384901834093281481083459830985560984095809456894
0680946804986049860458094802943209385029348509898690285096
7890956890456809348560394563956034560584560935050358609358
6904690586034870909289091001012948019238401983401923109234
0480923489043859023850923896508309683098670395860394586039
4860934869034586093860938093468093458038034680934568039458
6034560386034563945839586358568039539568309458038505609568
0349568345809358045860958603758609458603583405684305680349
5680395680395840958034860394680938609348609685309456039465
0439580349758603496093458620943580281039803481894104391092

Please, count...

Why Visualization?

ingeniería isa de sistemas y automática

Universidad de Oviedo

How many “sevens” do you find?

1093840192384901834093281481083459830985560984095809456894
0680946804986049860458094802943209385029348509898690285096
7890956890456809348560394563956034560584560935050358609358
6904690586034870909289091001012948019238401983401923109234
0480923489043859023850923896508309683098670395860394586039
4860934869034586093860938093468093458038034680934568039458
6034560386034563945839586358568039539568309458038505609568
0349568345809358045860958603758609458603583405684305680349
5680395680395840958034860394680938609348609685309456039465
0439580349758603496093458620943580281039803481894104391092

Why Visualization?

ingeniería isa de sistemas y automática

Universidad de Oviedo

How many “sevens” do you find?

1093840192384901834093281481083459830985560984095809456894
0680946804986049860458094802943209385029348509898690285096
7890956890456809348560394563956034560584560935050358609358
6904690586034870909289091001012948019238401983401923109234
0480923489043859023850923896508309683098670395860394586039
4860934869034586093860938093468093458038034680934568039458
6034560386034563945839586358568039539568309458038505609568
0349568345809358045860958603758609458603583405684305680349
5680395680395840958034860394680938609348609685309456039465
0439580349758603496093458620943580281039803481894104391092

Visualization provides “*context*”

location, cadence, patterns, proportion, ...
subtle forms of knowledge are combined
connected knowledge = more knowledge !!

ingeniería isa de sistemas y automática

Why Visualization?

Universidad de Oviedo

are these data random?

what are the most repeated “xy” patterns?

1093840192384901834093281481083459830985560984095809456894
0680946804986049860458094802943209385029348509898690285096
7890956890456809348560394563956034560584560935050358609358
6904690586034870909289091001012948019238401983401923109234
0480923489043859023850923896508309683098670395860394586039
4860934869034586093860938093468093458038034680934568039458
6034560386034563945839586358568039539568309458038505609568
0349568345809358045860958603758609458603583405684305680349
5680395680395840958034860394680938609348609685309456039465
0439580349758603496093458620943580281039803481894104391092

Why Visualization?

most repeated “xy” patterns?

1093840192384901834093281481083459830985560984095809456894
0680946804986049860458094802943209385029348509898690285096
7890956890456809348560394563956034560584560935050358609358
6904690586034870909289091001012948019238401983401923109234
0480923489043859023850923896508309683098670395860394586039
4860934869034586093860938093468093458038034680934568039458
6034560386034563945839586358568039539568309458038505609568
0349568345809358045860958603758609458603583405684305680349
5680395680395840958034860394680938609348609685309456039465
0439580349758603496093458620943580281039803481894104391092

Why Visualization?

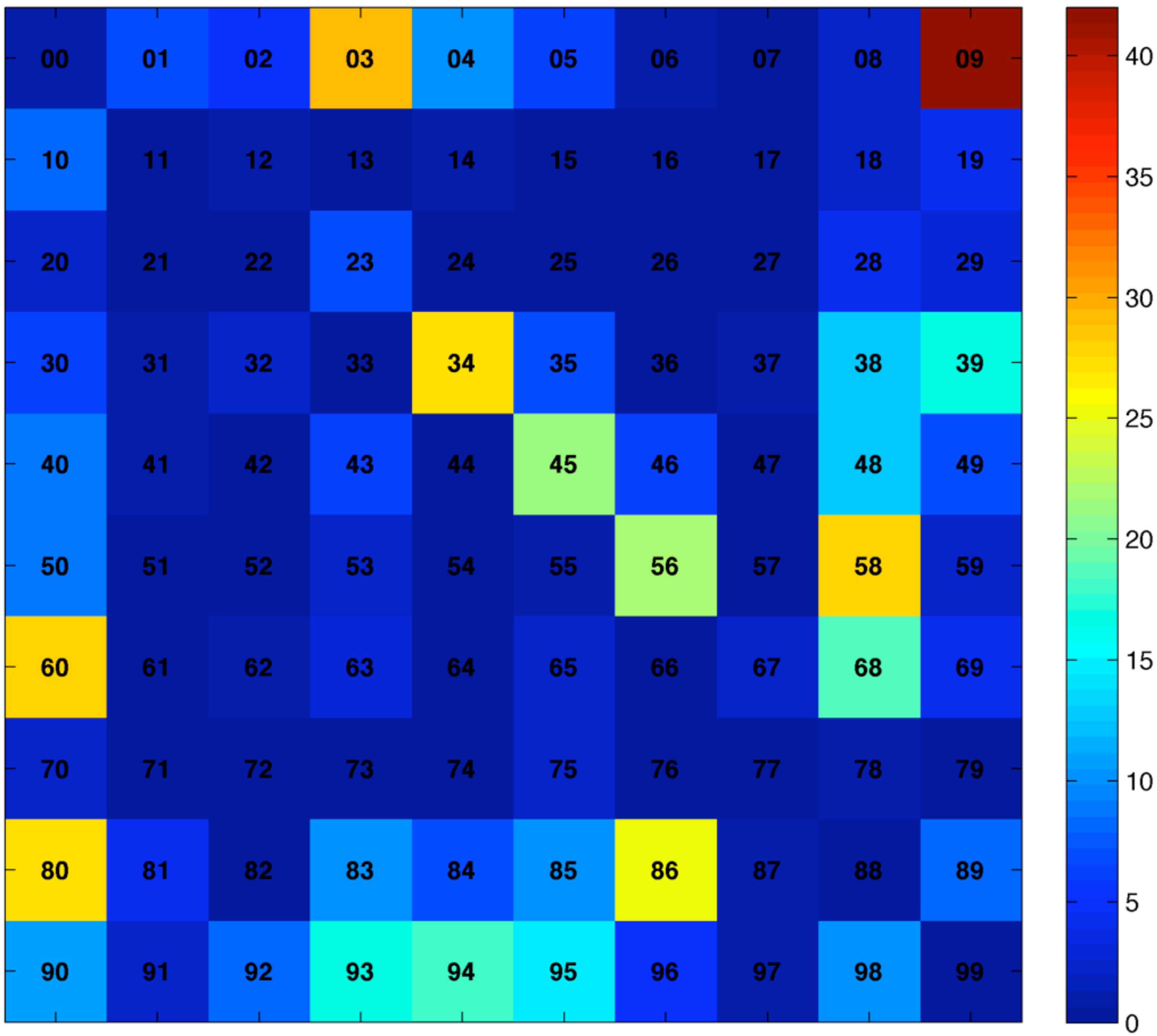
second most repeated “xy” patterns

1093840192384901834093281481083459830985560984095809456894
0680946804986049860458094802943209385029348509898690285096
7890956890456809348560394563956034560584560935050358609358
6904690586034870909289091001012948019238401983401923109234
0480923489043859023850923896508309683098670395860394586039
4860934869034586093860938093468093458038034680934568039458
6034560386034563945839586358568039539568309458038505609568
0349568345809358045860958603758609458603583405684305680349
5680395680395840958034860394680938609348609685309456039465
0439580349758603496093458620943580281039803481894104391092

Why Visualization?

Universidad de Oviedo

visual histogram of “xy” patterns?



Statistics sometimes mislead

Ancombe's quartet

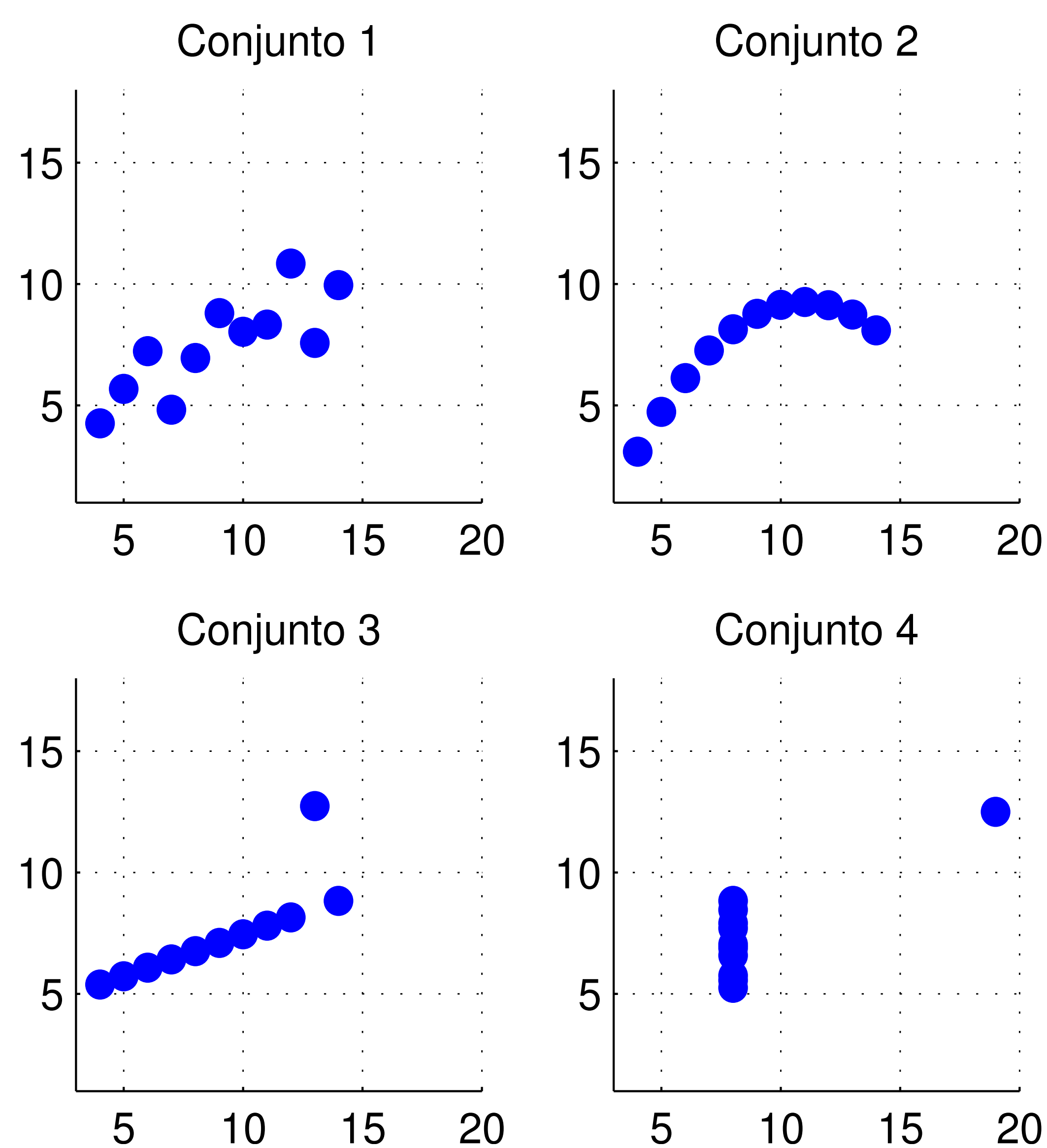
```
clear;
close all;

% Datos del cuarteto Anscombe (http://en.wikipedia.org/wiki/Anscombe's_quartet)
p  = [10.0  8.04    10.0    9.14    10.0    7.46    8.0 6.58
8.0 6.95    8.0 8.14    8.0 6.77    8.0 5.76
13.0    7.58    13.0    8.74    13.0    12.74   8.0 7.71
9.0 8.81    9.0 8.77    9.0 7.11    8.0 8.84
11.0    8.33    11.0    9.26    11.0    7.81    8.0 8.47
14.0    9.96    14.0    8.10    14.0    8.84    8.0 7.04
6.0 7.24    6.0 6.13    6.0 6.08    8.0 5.25
4.0 4.26    4.0 3.10    4.0 5.39    19.0    12.50
12.0    10.84   12.0    9.13    12.0    8.15    8.0 5.56
7.0 4.82    7.0 7.26    7.0 6.42    8.0 7.91
5.0 5.68    5.0 4.74    5.0 5.73    8.0 6.89];

% Separamos los cuatro conjuntos
x{1} = p(:,1:2);
x{2} = p(:,3:4);
x{3} = p(:,5:6);
x{4} = p(:,7:8);

figure(1);
clf;

for k = 1:4,
    subplot(2,2,k);
    scatter(x{k}(:,1),x{k}(:,2),50,'filled');
    axis equal;
    axis([3 20 1 18]);
    grid on;
    C{k} = cov(x{k}),
    M{k} = mean(x{k}),
    R{k} = corrcoef(x{k}),
    title(sprintf('Conjunto %d',k));
end
```



¡The 4 sets have the same statistic descriptors!

Covariance Matrices

```
>> for k = 1:4, C{k}, end;
ans =

   11.0000    5.5010
    5.5010    4.1273


ans =

   11.0000    5.5000
    5.5000    4.1276


ans =

   11.0000    5.4970
    5.4970    4.1226


ans =

   11.0000    5.4990
    5.4990    4.1232
```

Mean vectors

```
>> for k = 1:4, M{k}, end;
ans =

    9.0000    7.5009


ans =

    9.0000    7.5009


ans =

    9.0000    7.5000


ans =

    9.0000    7.5009
```

Big Data

Industrial systems also pose a big data problem

- Huge amounts of data
  - Many sensors
  - Ubiquitous information
  - Heterogeneous information

- Complex systems
  - Dynamic behavior
  - Connected, interacting, coupled…
  - Many variables, multiway problems
  - Human factors
  - Interaction with other factories

Problem:

Get insight
from an ocean of data

Detect structure in data
and model it

References:
Big data: The next frontier for innovation, competition, and productivity.
http://www.mckinsey.com/insights/business_technology/big_data_the_next_frontier_for_innovation.

A few Big Data numbers

Visual Exploration *of* Large Data Sets

Computer systems today store vast amounts of data. Researchers, including those working on the "How Much Information?" project at the University of California, Berkeley, recently estimated, about 1 exabyte (1 million terabytes) of data is generated annually worldwide, including 99.997% available only in digital form. This worldwide data deluge means that in the next three years, more data will be generated than during all previous human history.

In the rising tide of business transaction data, these tools help distinguish which are strategic assets and which are not worth collecting in the first place.

Daniel A. Keim

Data is often recorded, captured, and stored automatically via sensors and monitoring systems. Many of the simple transactions now part of our everyday lives, such as paying for food and clothes by credit card or using the telephone, are typically recorded for future reference by computers. Many parameters of each transaction are routinely captured, resulting in highly dimensional data. The data is collected because companies, including those engaged in some kind of e-commerce, view it as a source of potentially valuable information that, as a strategic asset, could provide a competitive advantage. But actually finding this valuable information is difficult. Today's data management systems make it possible to view only small portions of it. If the data is presented in text form, the amount that can be displayed amounts to only about 100 data items—a drop in the ocean when dealing with data sets containing millions of data items. Lacking the ability to adequately explore the large amounts being collected, and despite its potential usefulness, the data becomes useless and the databases data dumps. Visual data exploration, which aims to provide insight by visualizing the data, and information visualization techniques (such as distorted overview displays and dense pixel displays) can help solve this problem.

Effective data mining depends on having a human in the data exploration process while combining this person's flexibility, creativity, and general knowledge with the enormous storage capacity and computational power of today's computers. Visual data exploration seeks to integrate humans in the data exploration process, applying their perceptual abilities to the large data sets now available. The basic idea is to present the data in some visual form, allowing data analysts to gain insight into it and draw conclusions, as well as interact with it. The visual representation of the data reduces the cognitive work needed to perform certain tasks.

Visual data mining techniques have proved their value in exploratory data analysis; they also have great

Computer systems today store vast amounts of data. Researchers, including those working on the "How Much Information?" project at the University of California, Berkeley, recently estimated, about 1 exabyte (1 million terabytes) of data is generated annually worldwide, including 99.997% available only in digital form. This worldwide data deluge means that in the next three years, more data will be generated than during all previous human history.

A few Big Data numbers

VISUAL EXPLORATION *of* LARGE DATA SETS

Computer systems today store vast amounts of data. Researchers, including those working on the "How Much Information?" project at the University of California, Berkeley, recently estimated, about 1 exabyte (1 million terabytes) of data is generated annually worldwide, including 99.997% available only in digital form. This worldwide data deluge means that in the next three years, more data will be generated than during all previous human history.

In the rising tide of business transaction data, these tools help distinguish which are strategic assets and which are not worth collecting in the first place.

Daniel A. Keim

Data is often recorded, captured, and stored automatically via sensors and monitoring systems. Many of the simple transactions now part of our everyday lives, such as paying for food and clothes by credit card or using the telephone, are typically recorded for future reference by computers. Many parameters of each transaction are routinely captured, resulting in highly dimensional data. The data is collected because companies, including those engaged in some kind of e-commerce, view it as a source of potentially valuable information that, as a strategic asset, could provide a competitive advantage. But actually finding this valuable information is difficult. Today's data management systems make it possible to view only small portions of it. If the data is presented in text form, the amount that can be displayed amounts to only about 100 data items—a drop in the ocean when dealing with data sets containing millions of data items. Lacking the ability to adequately explore the large amounts being collected, and despite its potential usefulness, the data becomes useless and the databases data dumps. Visual data exploration, which aims to provide insight by visualizing the data, and information visualization techniques (such as distorted overview displays and dense pixel displays) can help solve this problem.

Effective data mining depends on having a human in the data exploration process while combining this person's flexibility, creativity, and general knowledge with the enormous storage capacity and computational power of today's computers. Visual data exploration seeks to integrate humans in the data exploration process, applying their perceptual abilities to the large data sets now available. The basic idea is to present the data in some visual form, allowing data analysts to gain insight into it and draw conclusions, as well as interact with it. The visual representation of the data reduces the cognitive work needed to perform certain tasks.

Visual data mining techniques have proved their value in exploratory data analysis; they also have great

Computer systems today store vast amounts of data. Researchers, including those working on the "How Much Information?" project at the University of California, Berkeley, recently estimated, about 1 exabyte (1 million terabytes) of data is generated annually worldwide, including 99.997% available only in digital form. This worldwide data deluge means that in the next three years, more data will be generated than during all previous human history.

COMMUNICATIONS OF THE ACM August 2001/Vol. 44, No. 8

!

ingeniería isa de sistemas y automática

A few Big Data numbers
(according to Mc Kinsey Global Institute report)

Universidad de Oviedo

600 $ cost of a hard disk having space to store all the music composed ever

40% projected growth in global data generated per year

5% estimated growth of IT spending

http://www.mckinsey.com/insights/business_technology/big_data_the_next_frontier_for_innovation

J. Manyika, M. Chui, B. Brown, J. Bughin, R. Dobbs, C. Roxburgh, and A. H. Byers. Big data: The next frontier for innovation, competition, and productivity. Technical report, McKinsey Global Institute, May 2011.

A few Big Data numbers
(according to Mc Kinsey Global Institute report)

ingeniería isa de sistemas y automática
Universidad de Oviedo

60% potential increase of benefits in retailers' operating margins possible with big data

300 G$ potential annual value to USA health care (more than double the total annual health care spending in Spain)

250 G€ estimated value of big data from EU public administration (more than the greek GDP)

190k estimation of highly qualified data analysis experts needed in USA

+

1.5M estimation of data-savvy managers needed in USA to fully harness big data.

Visual Analytics

Process analysis workflow

Universidad de Oviedo

Data	Information	Knowledge

Process	Data acquisition	Feature extraction	Intelligent Data Analysis (IDA)	Visualization
	Advanced sensors Artificial vision Exploitation data Web data (XML, CSV, GeoJSON, etc.)	Frequency analysis (FFT, etc) System Identification Statistical descriptors Geometrical descriptors	Modeling Prediction Clustering and classification Dimensionality reduction Change/Novelty detection	Data visualization Interaction mechanisms Interface design Visualization of IDA algorithms

Process analysis workflow

Universidad de Oviedo

VISUAL ANALYTICS

Data — Information — Knowledge

Process

Data acquisition

Advanced sensors
Artificial vision
Exploitation data
Web data (XML, CSV, GeoJSON, etc.)

Feature extraction

Frequency analysis (FFT, etc)
System Identification
Statistical descriptors
Geometrical descriptors

Intelligent Data Analysis (IDA)

Modeling
Prediction
Clustering and classification
Dimensionality reduction
Change/Novelty detection

Visualization

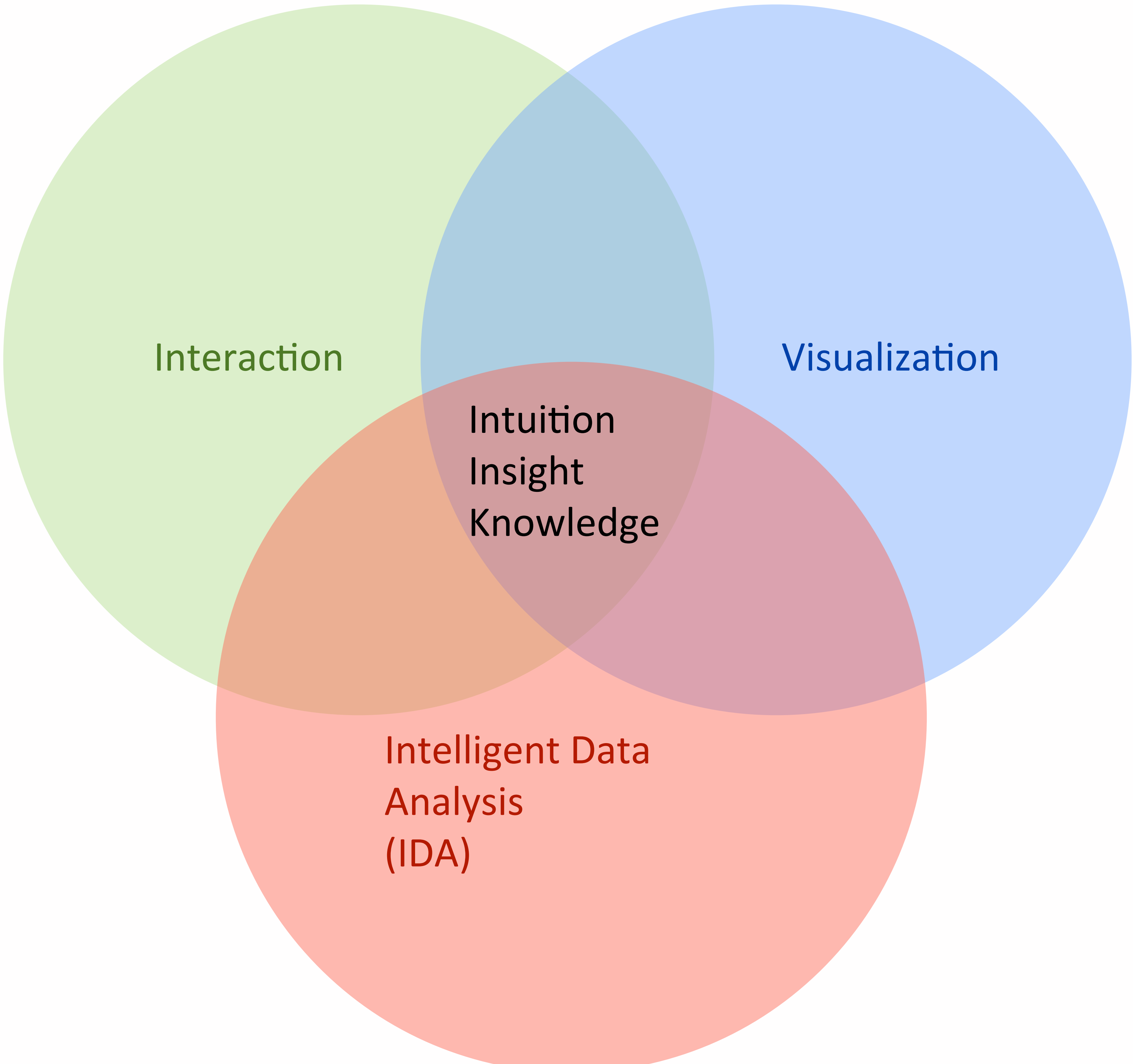
Data visualization
Interaction mechanisms
Interface design
Visualization of IDA algorithms

ingeniería isa de sistemas y automática

Visual Analytics approach

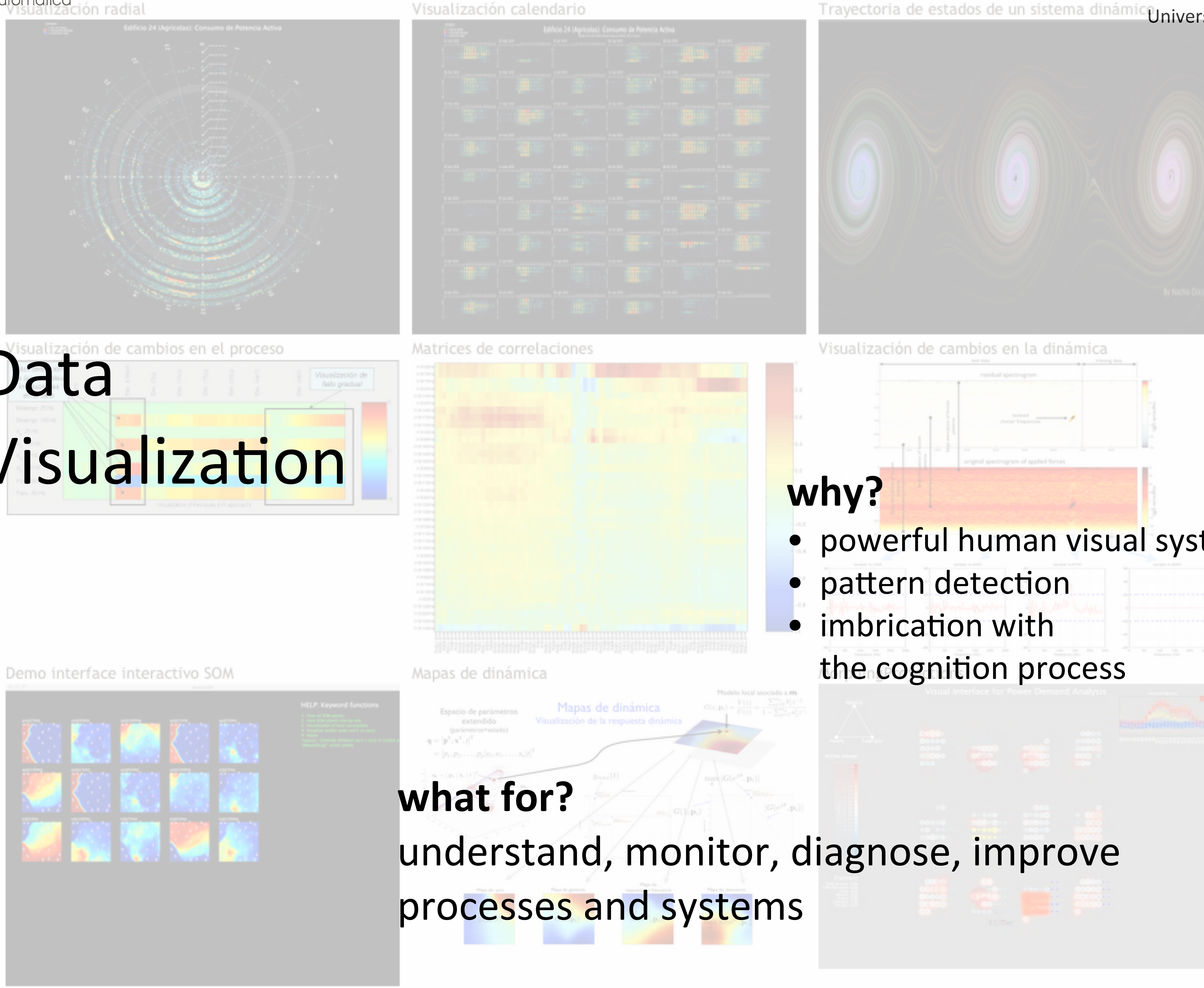
Universidad de Oviedo

Human Learning + Machine Learning Synergy



Process data analysis

Universidad de Oviedo



Data Visualization

why?

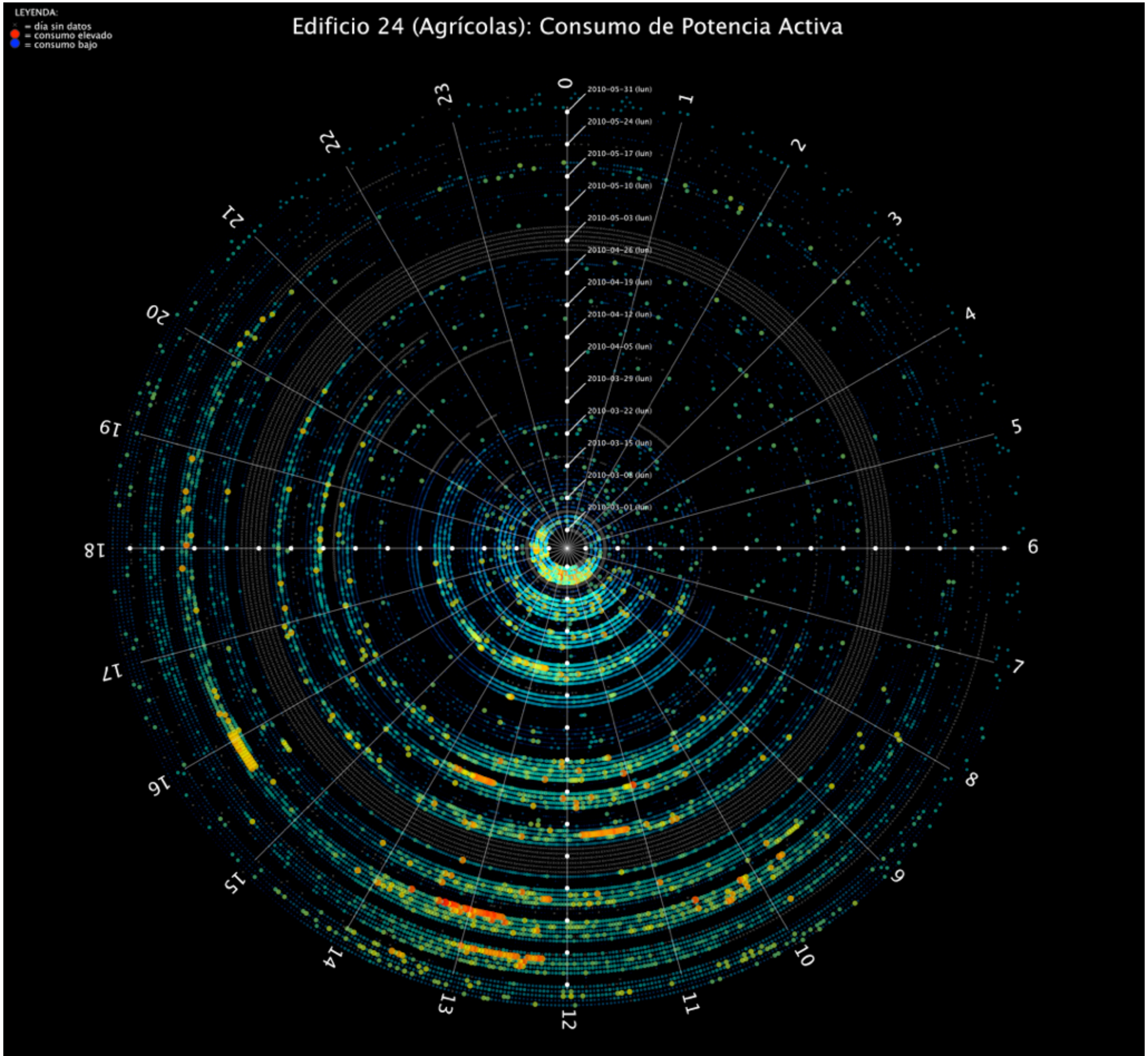
- powerful human visual system,
- pattern detection
- imbrication with the cognition process

what for?

understand, monitor, diagnose, improve processes and systems

Data visualization

power demand data visualization (radial)

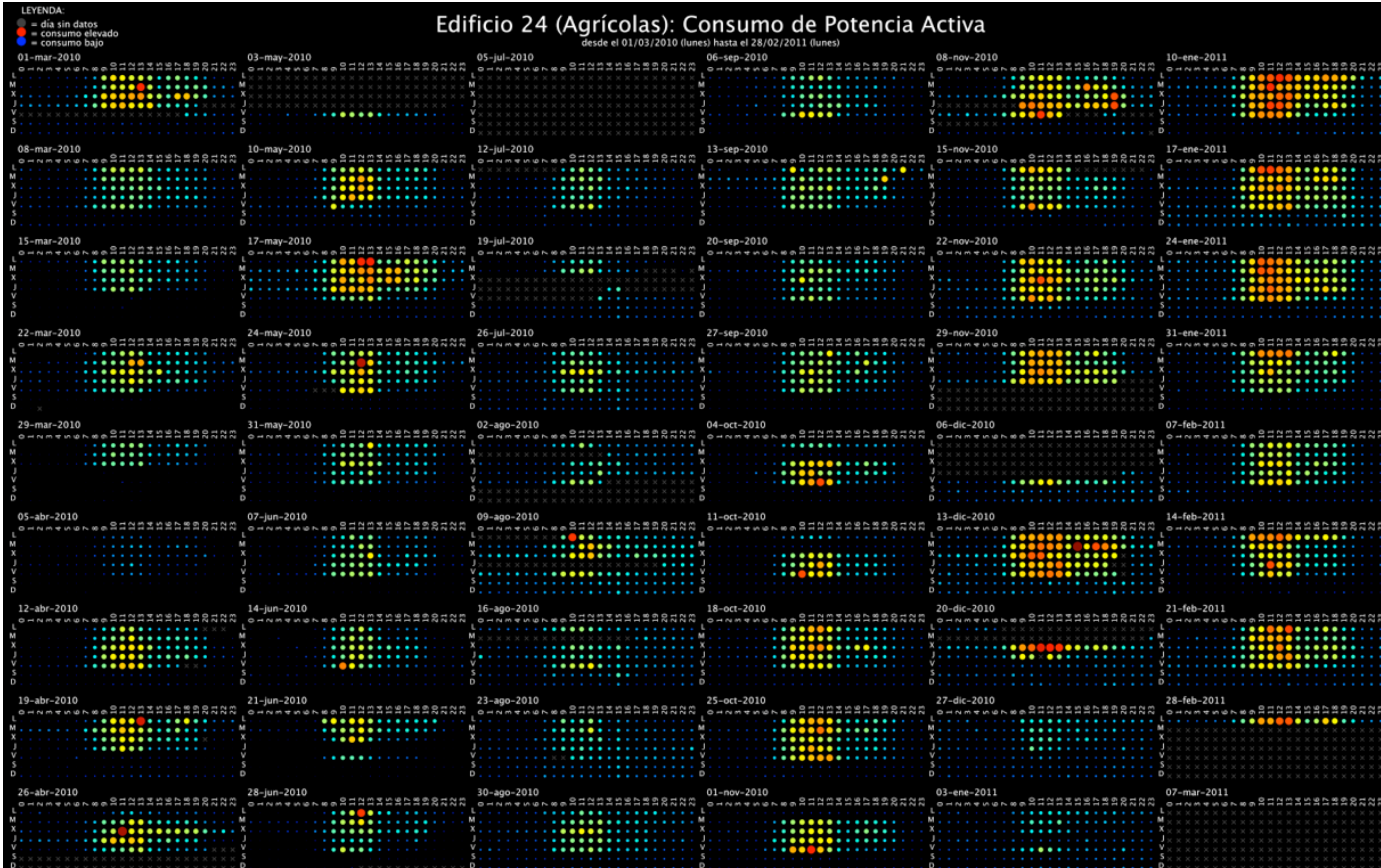


ingeniería isa de sistemas y automática

Data visualization

power demand data visualization (small-multiples)

Universidad de Oviedo



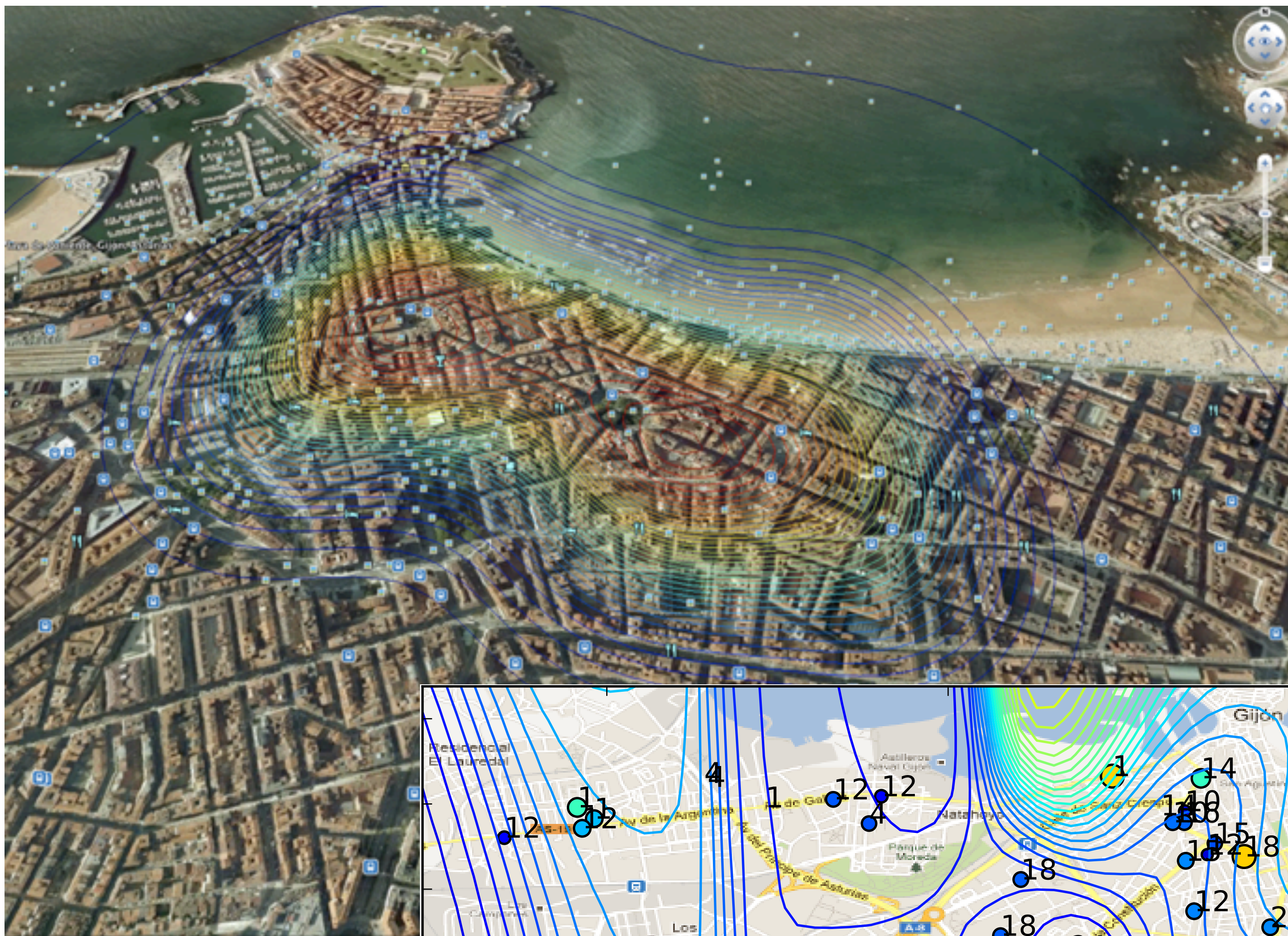
ingeniería isa de sistemas y automática

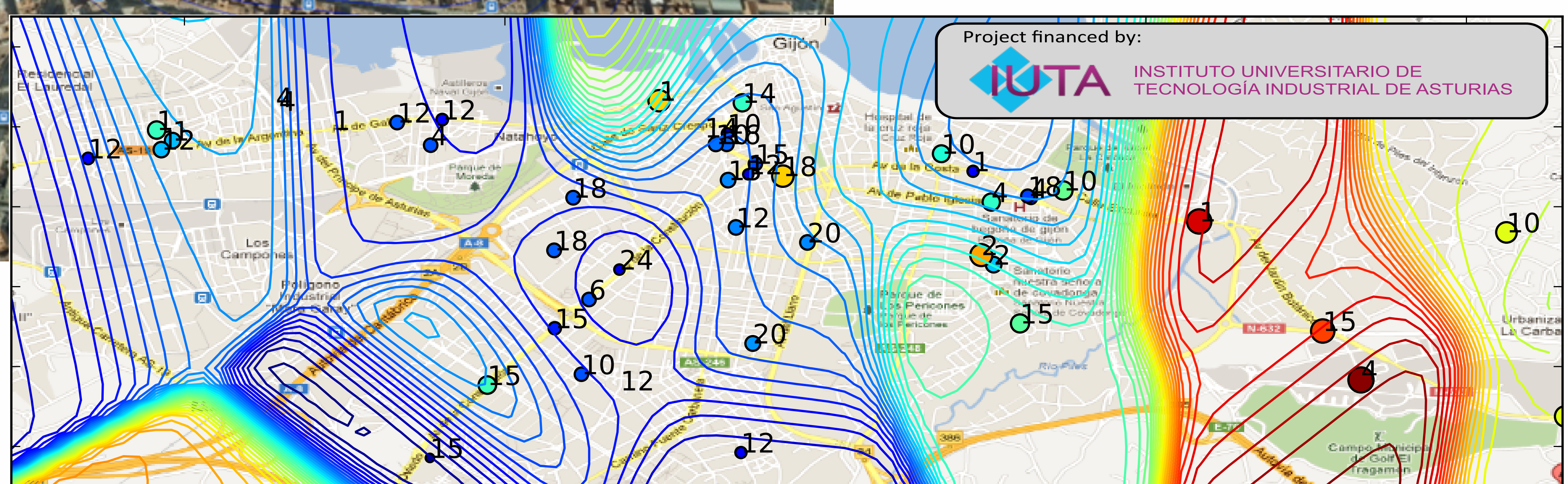
Data visualization

Geo + data = insight

Universidad de Oviedo

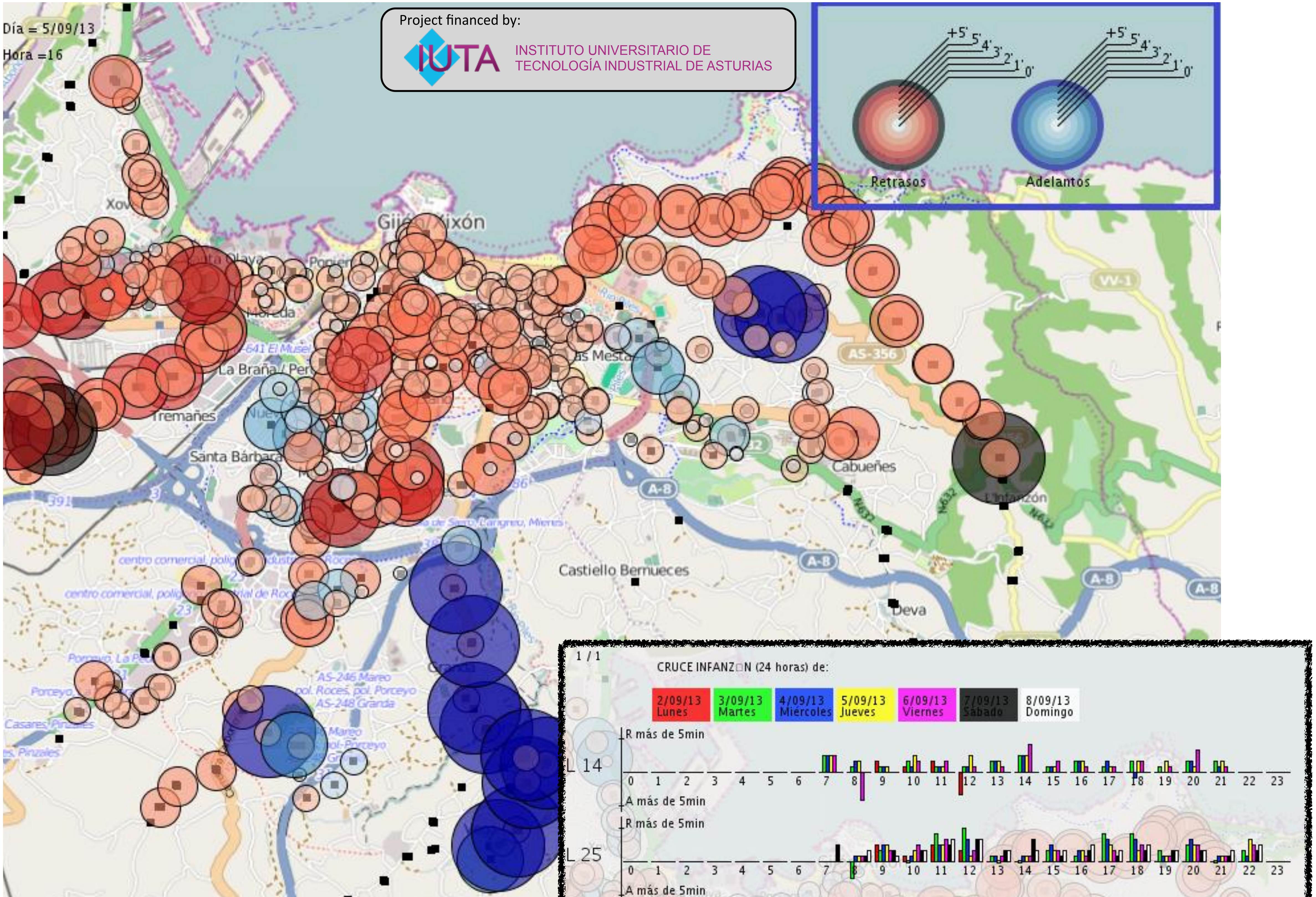
“Hot spot” visualization from geotagged data





Data visualization

Spatio-temporal analysis: urban transport system in Gijón



Manifold learning methods appeared in Science, dec. 2000

[1] S. T. Roweis and L. K. Saul. Nonlinear dimensionality reduction by locally linear embedding. *Science*, 290:2323–2326, Dec., 22 2000.

[2] J. B. Tenenbaum, V. de Silva, and J. C. Langford. A global geometric framework for nonlinear dimensionality reduction. *Science*, 290:2319–2323, Dec, 22 2000.

Nonlinear Dimensionality Reduction by Locally Linear Embedding

Sam T. Roweis[1] and Lawrence K. Saul[2]

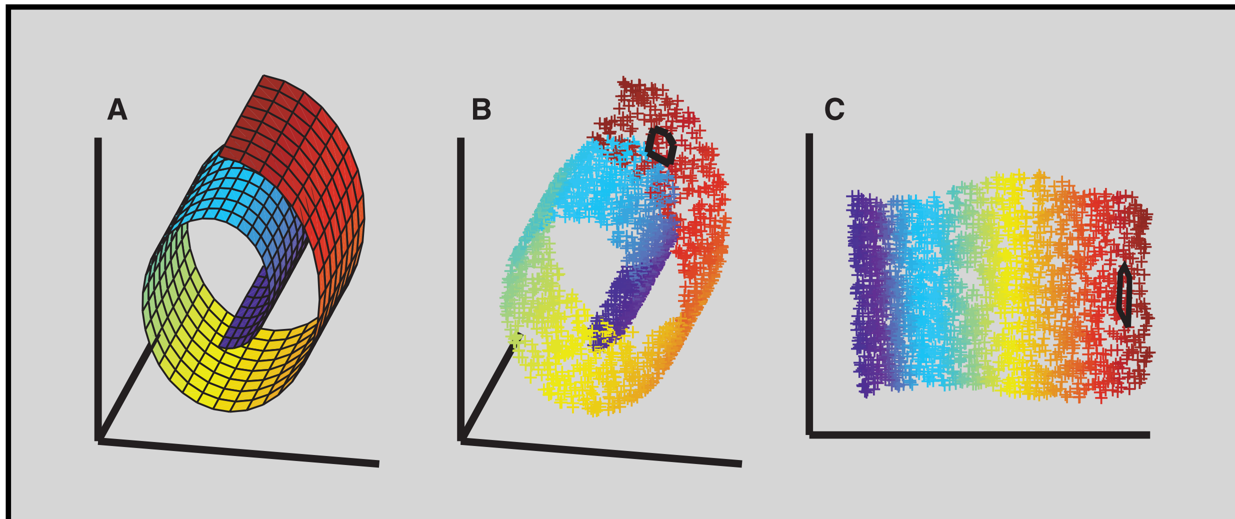
Many areas of science depend on exploratory data analysis and visualization. The need to analyze large amounts of multivariate data raises the fundamental problem of dimensionality reduction: how to discover compact representations of high-dimensional data. Here, we introduce locally linear embedding (LLE), an unsupervised learning algorithm that computes low-dimensional, neighborhood-preserving embeddings of high-dimensional inputs. Unlike clustering methods for local dimensionality reduction, LLE maps its inputs into a single global coordinate system of lower dimensionality, and its optimizations do not involve local minima. By exploiting the local symmetries of linear reconstructions, LLE is able to learn the global structure of nonlinear manifolds, such as those generated by images of faces or documents of text.

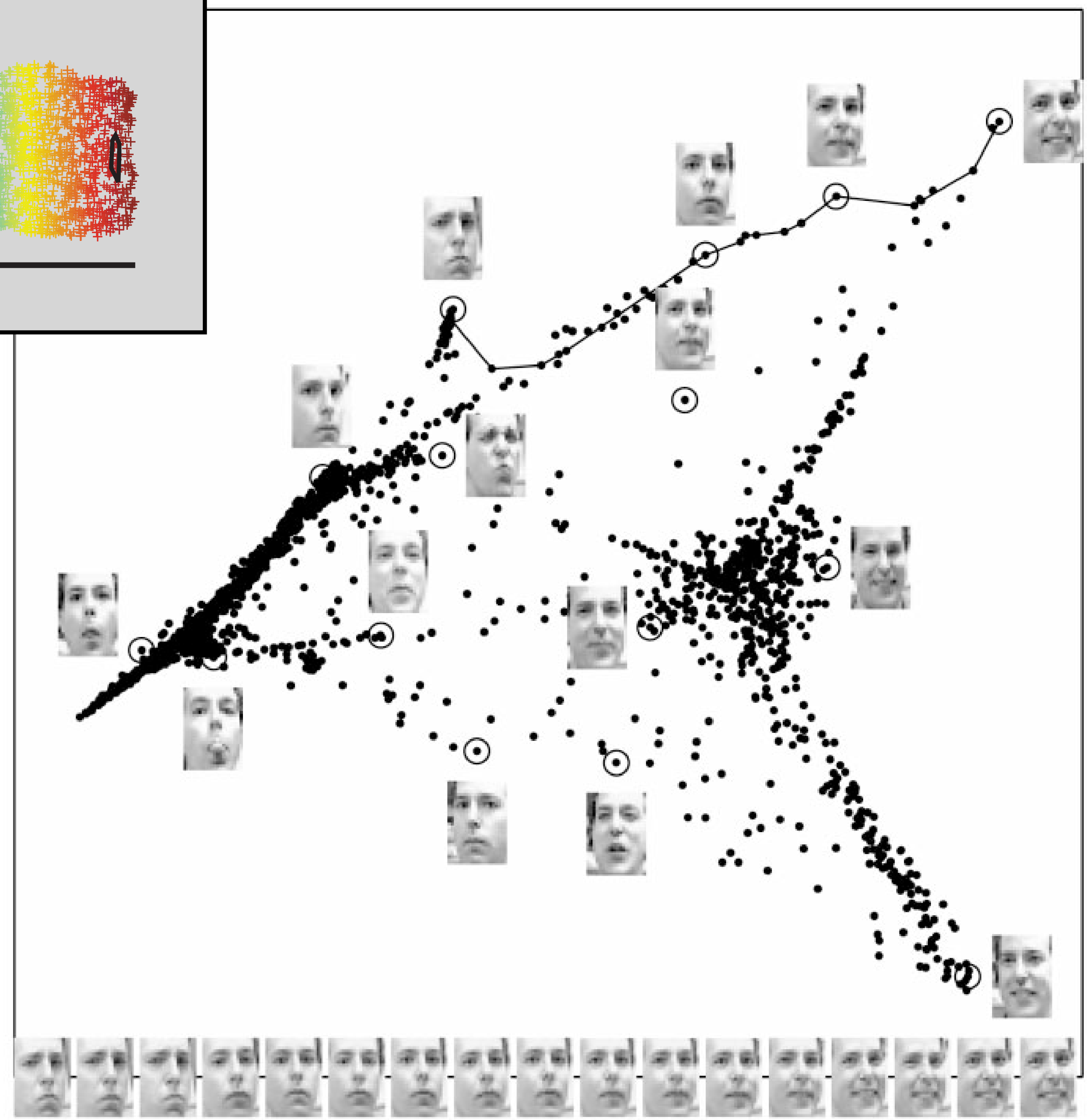
A Global Geometric Framework for Nonlinear Dimensionality Reduction

Joshua B. Tenenbaum,[1]* Vin de Silva,[2] John C. Langford[3]

Scientists working with large volumes of high-dimensional data, such as global climate patterns, stellar spectra, or human gene distributions, regularly confront the problem of dimensionality reduction: finding meaningful low-dimensional structures hidden in their high-dimensional observations. The human brain confronts the same problem in everyday perception, extracting from its high-dimensional sensory inputs—30,000 auditory nerve fibers or 10^6 optic nerve fibers—a manageably small number of perceptually relevant features. Here we describe an approach to solving dimensionality reduction problems that uses easily measured local metric information to learn the underlying global geometry of a data set. Unlike classical techniques such as principal component analysis (PCA) and multidimensional scaling (MDS), our approach is capable of discovering the nonlinear degrees of freedom that underlie complex natural observations, such as human handwriting or images of a face under different viewing conditions. In contrast to previous algorithms for nonlinear dimensionality reduction, ours efficiently computes a globally optimal solution, and, for an important class of data manifolds, is guaranteed to converge asymptotically to the true structure.

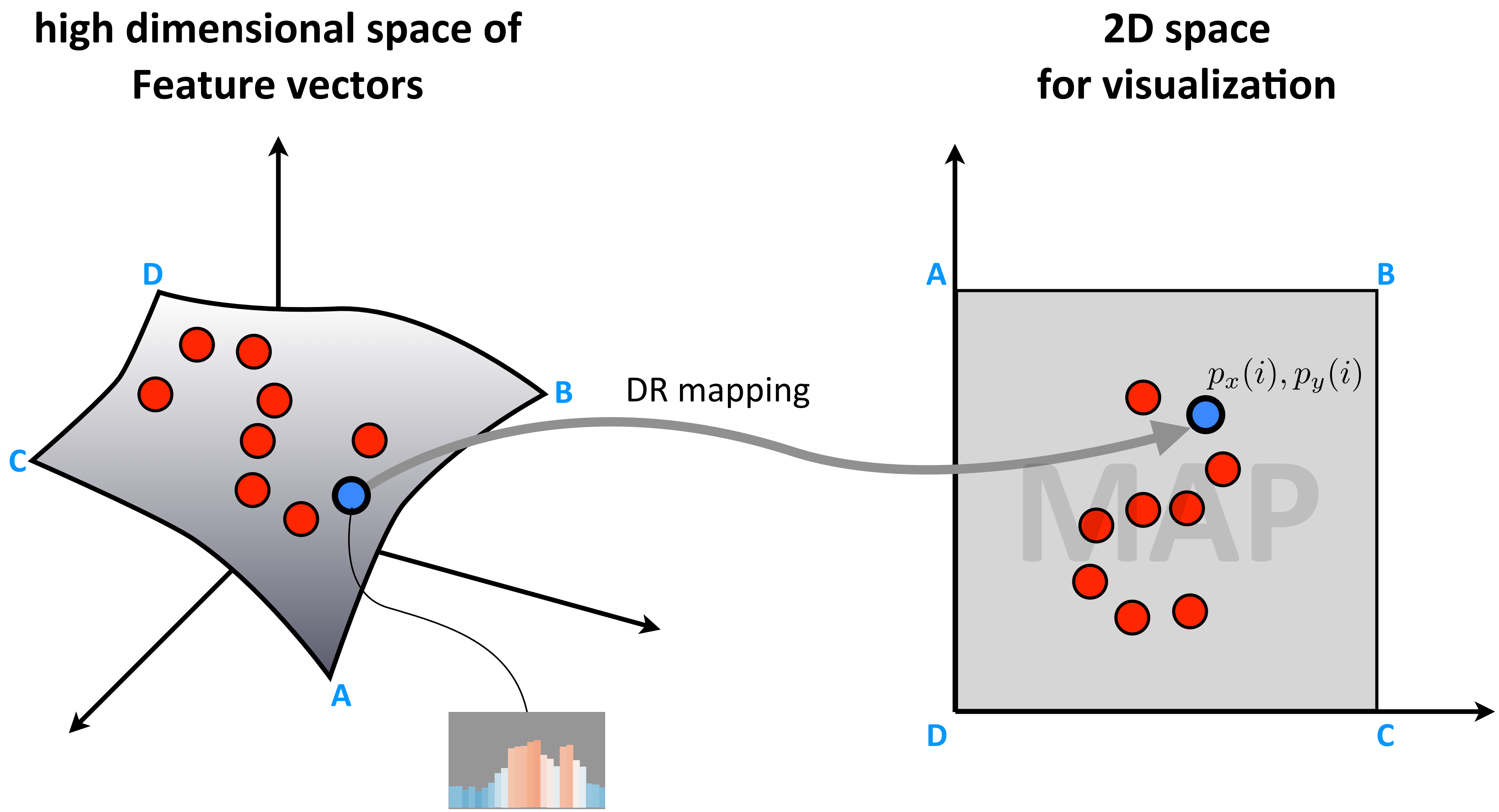
Dimensionality Reduction





Source:
(Roweis et al., 2000)

Dimensionality Reduction



Feature vectors can be:

raw process variables
harmonics,
statistical or geometrical descriptors
model parameters,
etc.

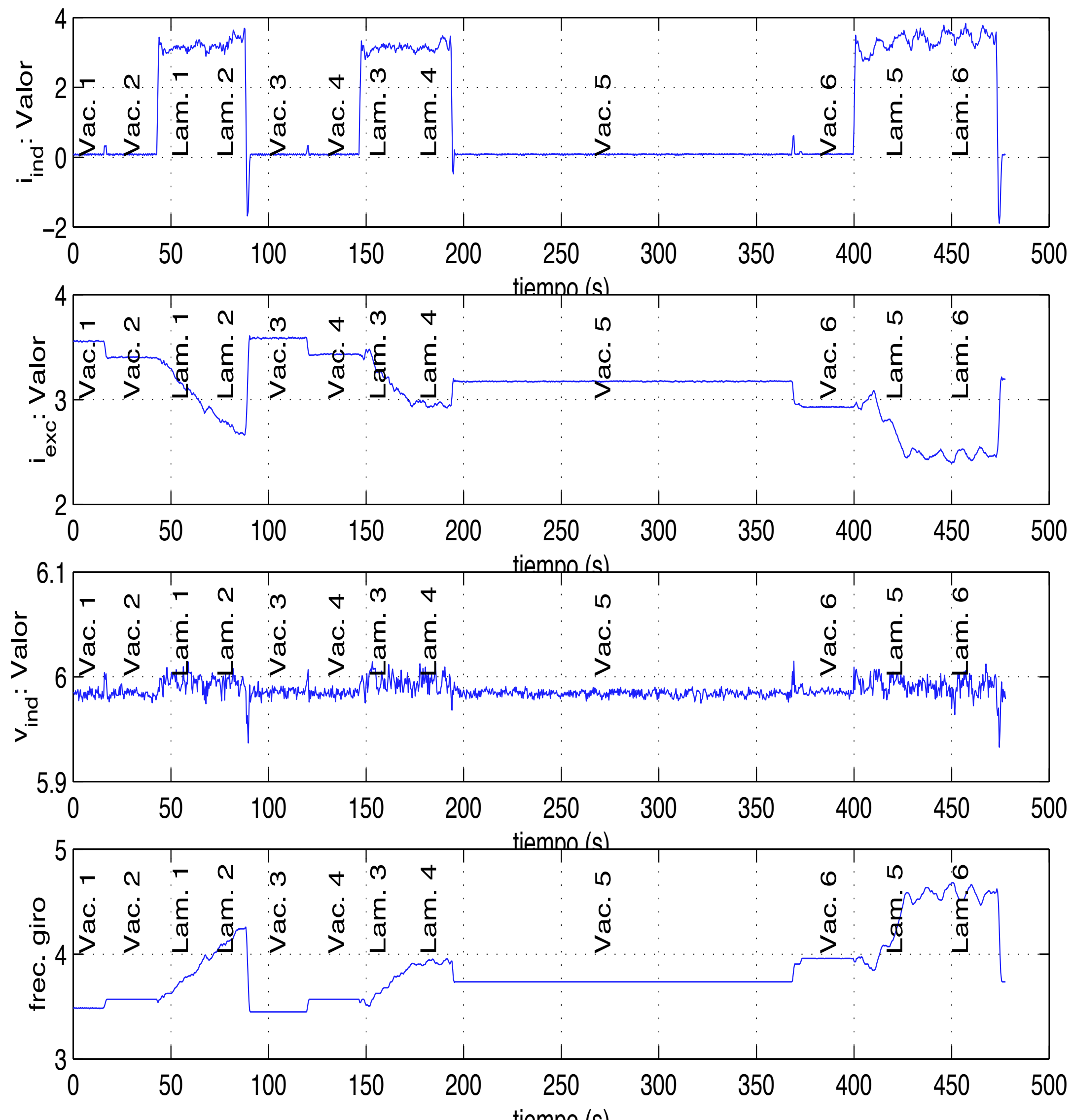
The visualization space:

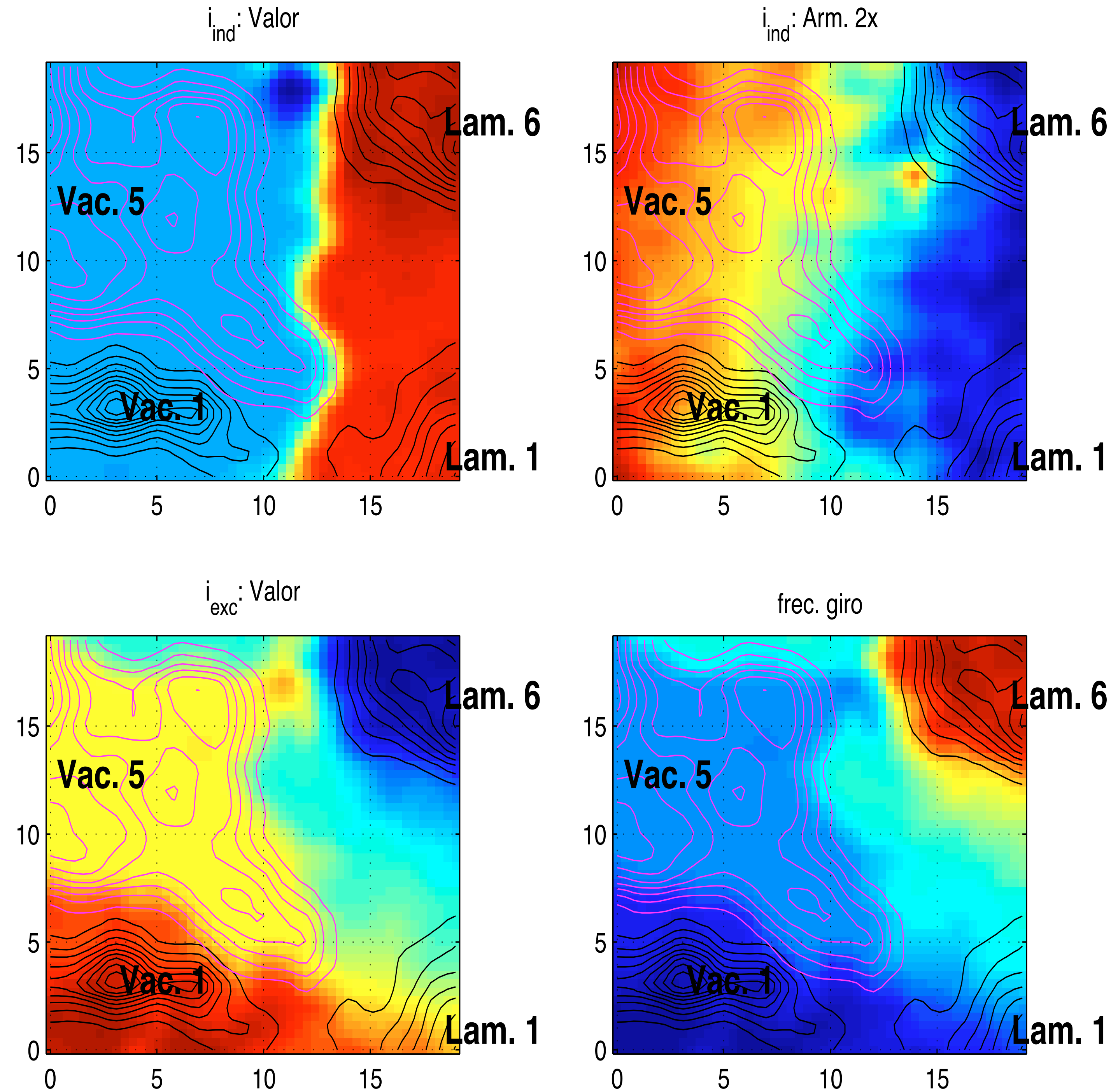
shows a “map” of the process
nearby points => similar feature vectors
Regions in the VS represent regions in FS

Dimensionality Reduction

Process data analysis

Universidad de Oviedo



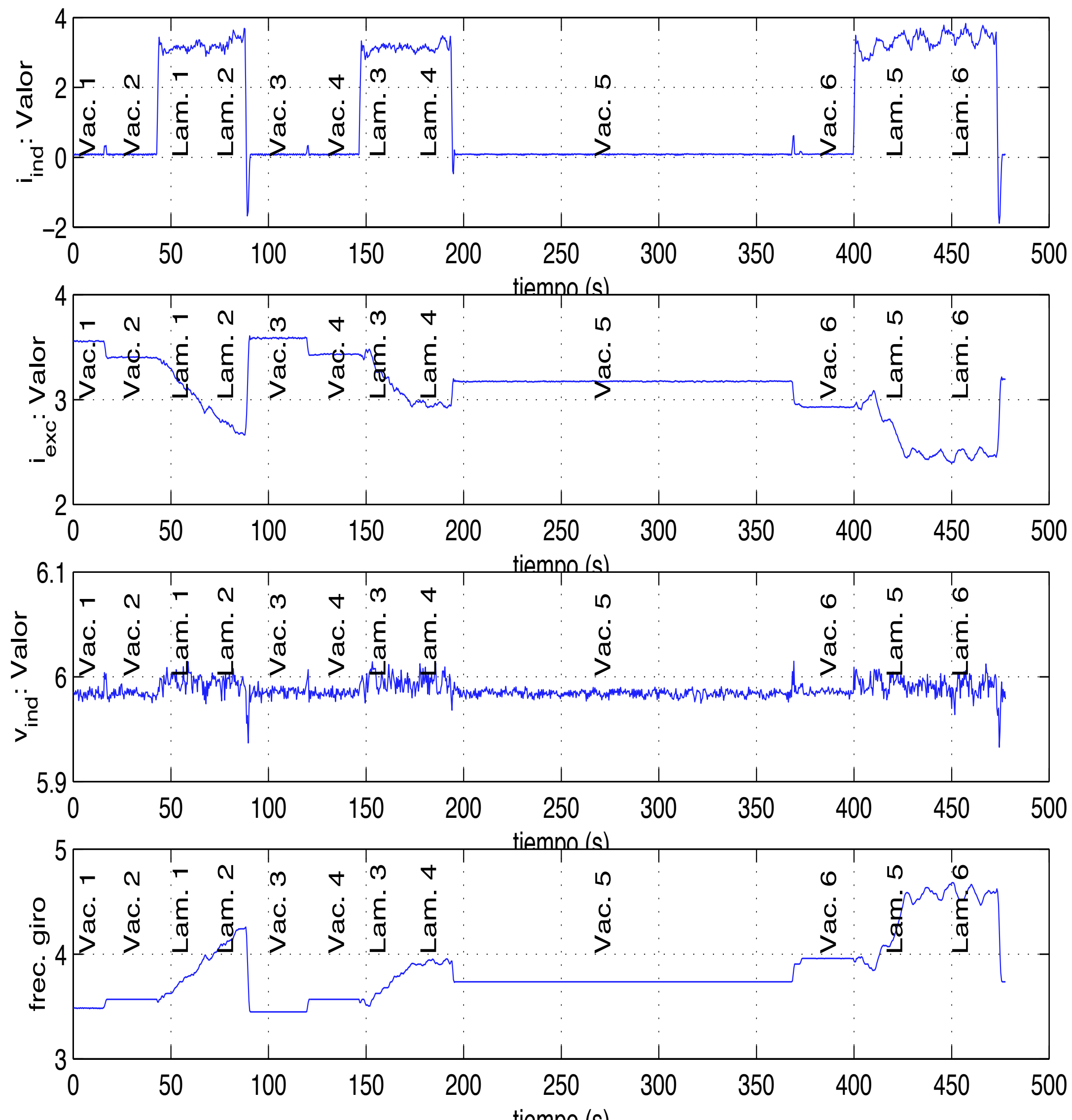


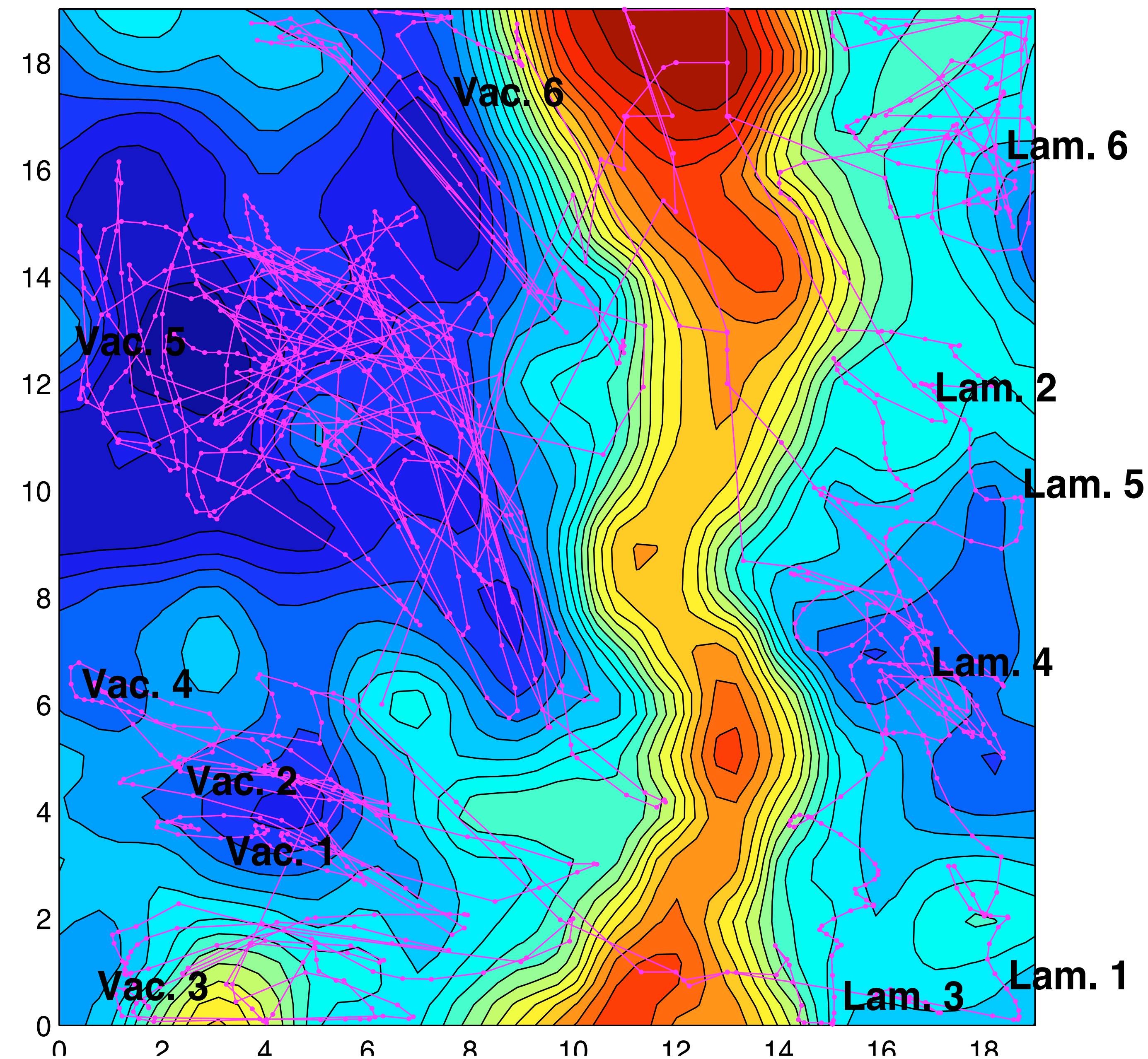
ingeniería isa de sistemas y automática

Dimensionality Reduction

condition monitoring

Universidad de Oviedo



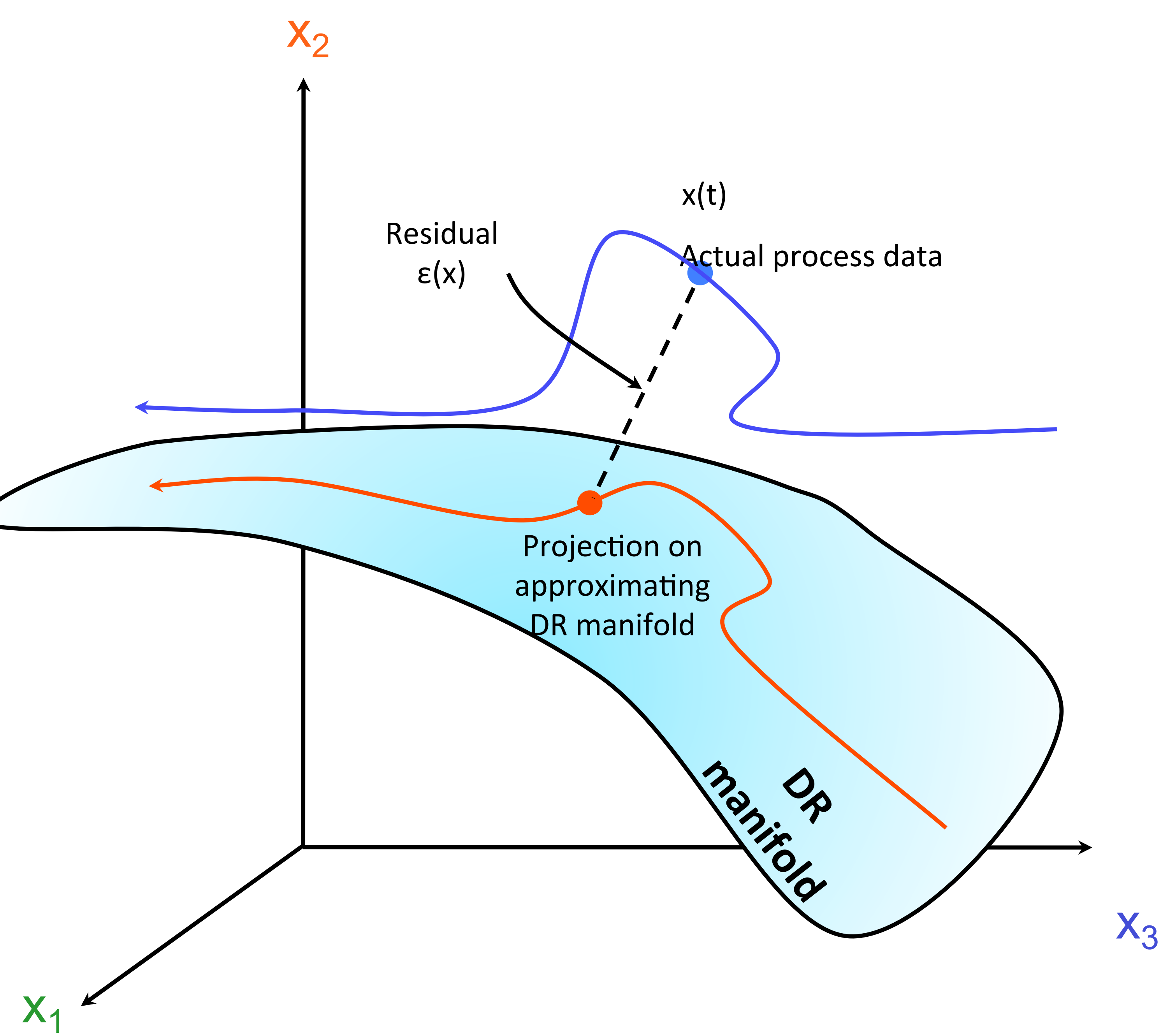


Dimensionality Reduction

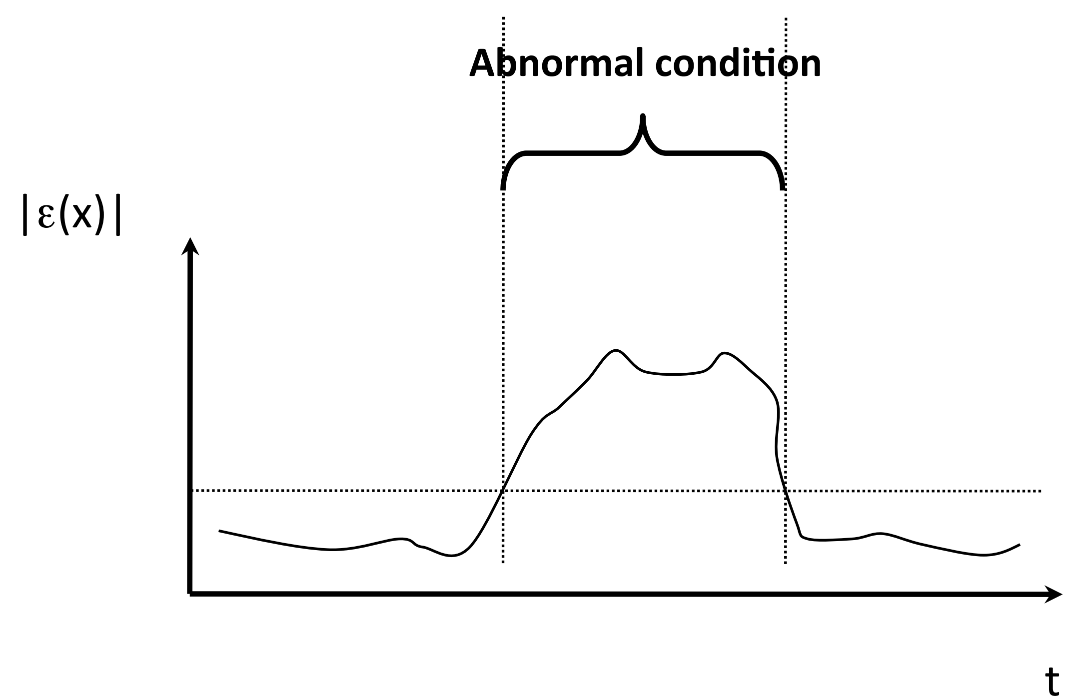
Novelty detection (residuals)

Process data space

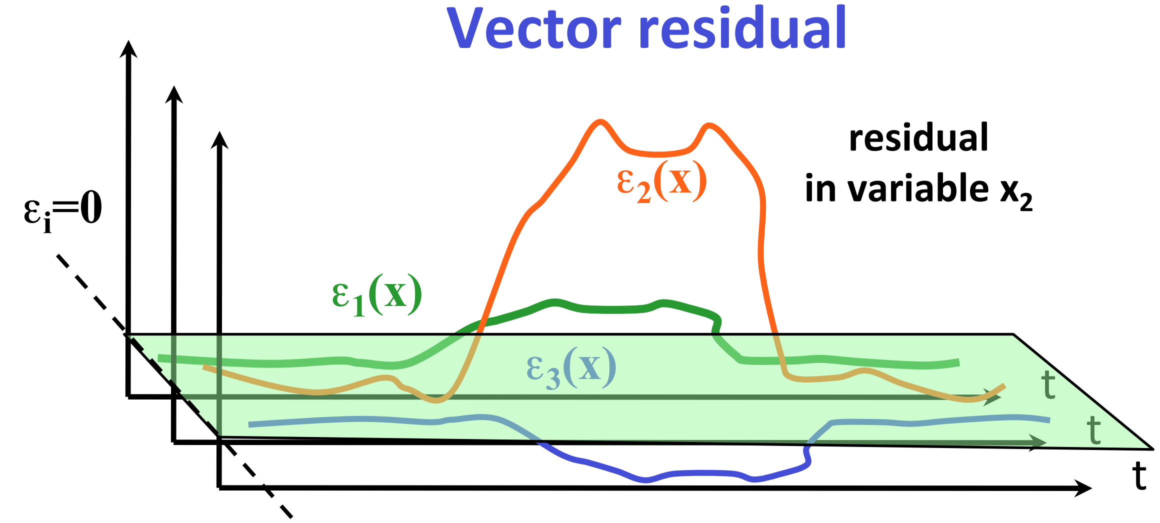
(x_1, x_2, ... are process variables or features)



Global residual

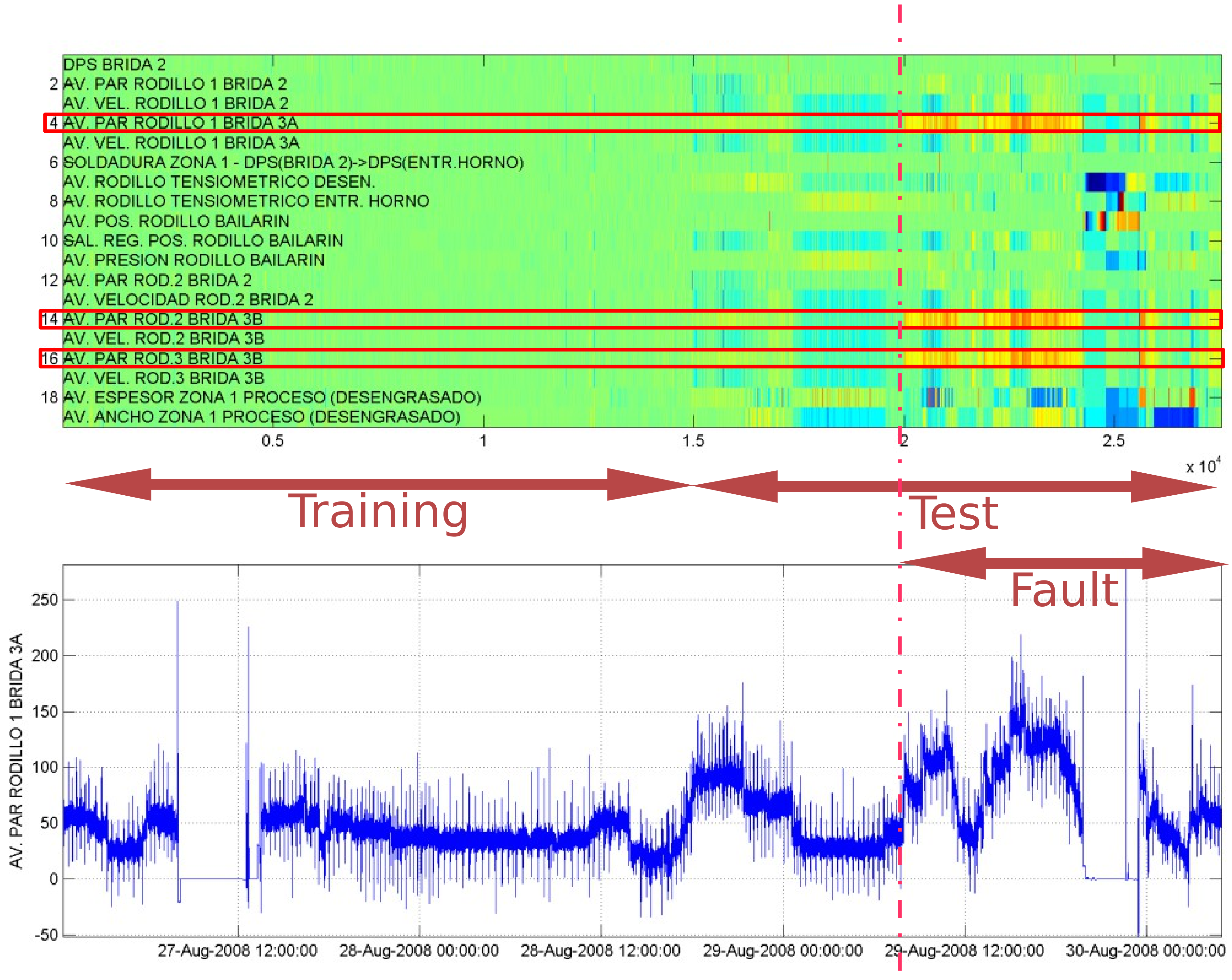


Vector residual



Dimensionality Reduction

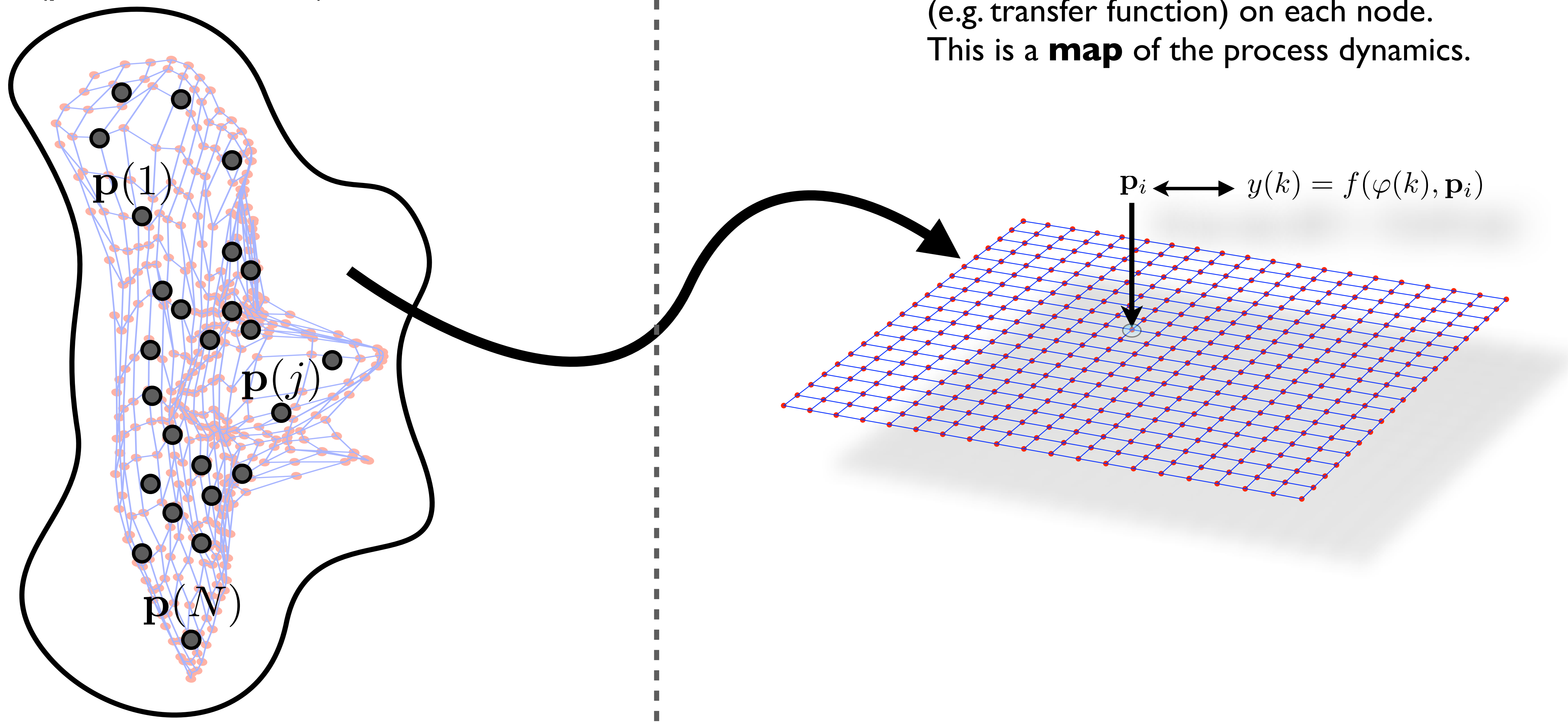
Novelty detection based on residuals



Basic idea...

input space
(parameter set *P*)

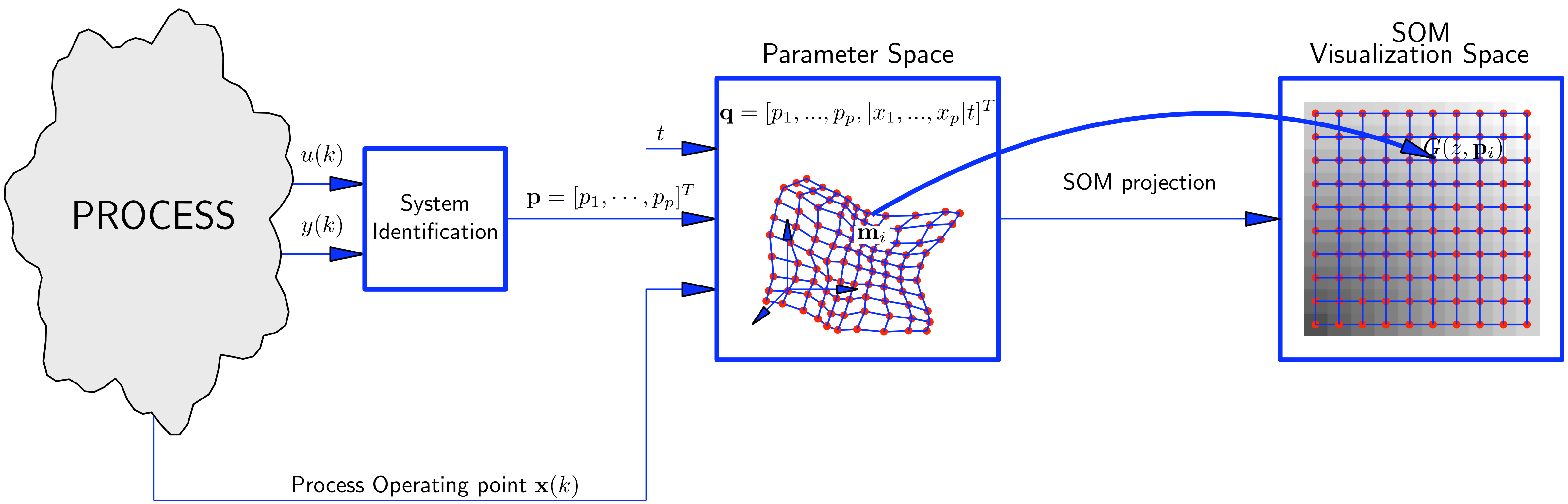
visualization space
The SOM stores a dynamic model (e.g. transfer function) on each node. This is a **map** of the process dynamics.



Dimensionality Reduction

Visualizing process dynamics through maps

… a bit more detailed approach



Dimensionality Reduction

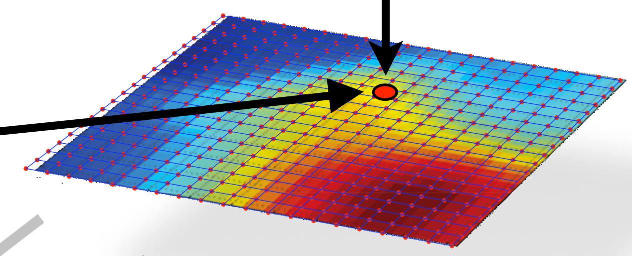
Visualizing process dynamics through maps

Universidad de Oviedo

ingeniería isa de sistemas y automática

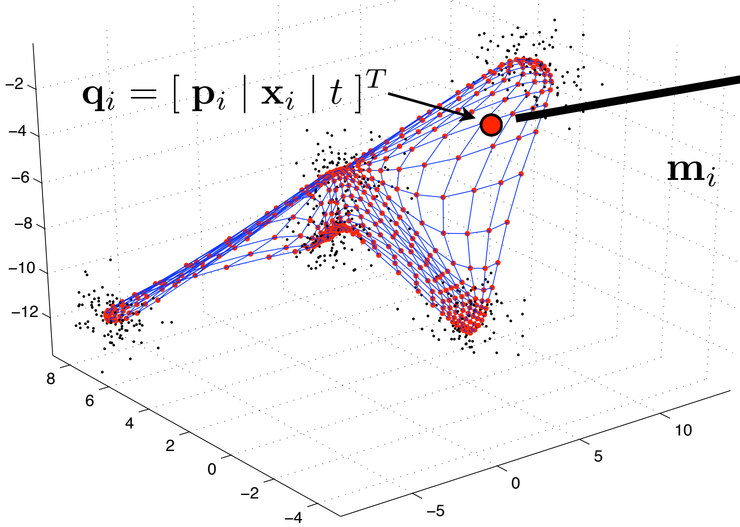
Local model bound to $\mathbf{m}_i$

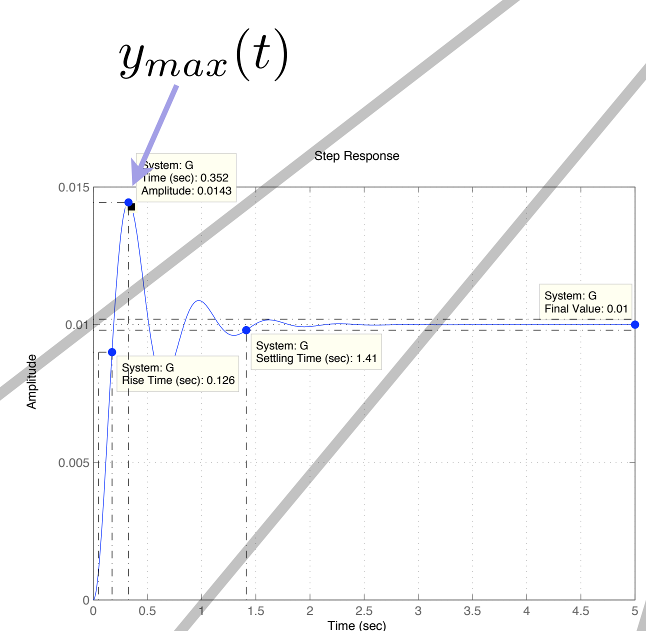
$$G(z, \mathbf{p}_i) = \frac{Y(z)}{U(z)} = \frac{\sum_{j=0}^{n_u} b_j^i z^{-j}}{1 - \sum_{j=1}^{n_y} a_j^i z^{-j}}$$

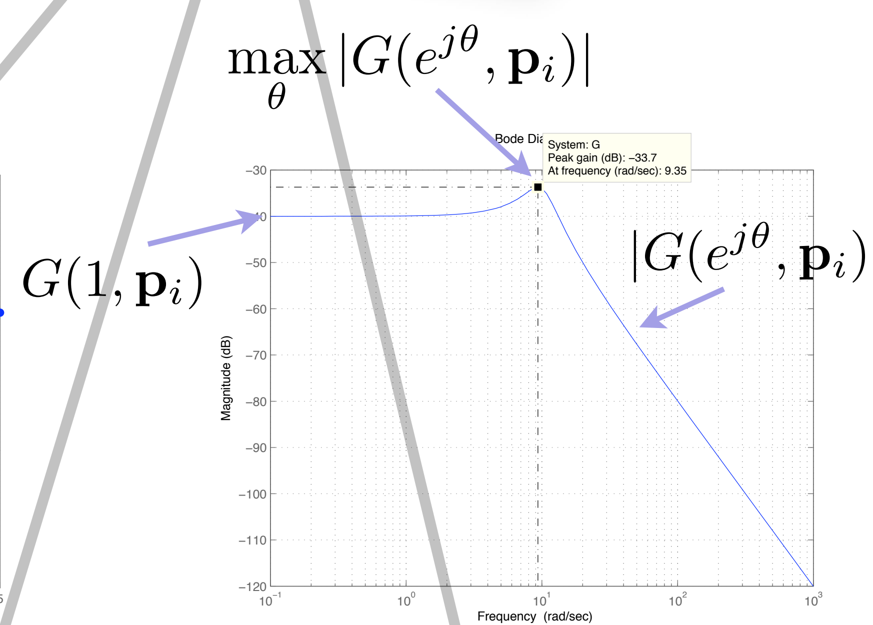


Extended parameter space
(parámeters+state)

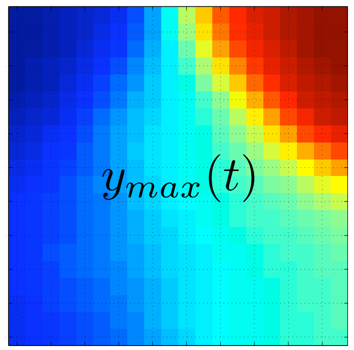
$$\mathbf{q} = [\mathbf{p}^{\mathrm{T}}, \mathbf{x}^{\mathrm{T}}, t]^{\mathrm{T}}$$
$$= [p_1, p_2, \ldots, p_p | x_1, x_2, \ldots, x_n | t]^{\mathrm{T}}$$



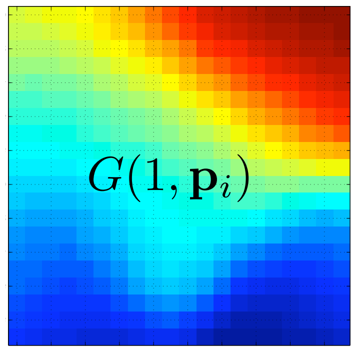




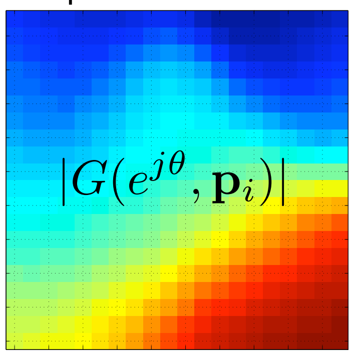
Mapa de pico



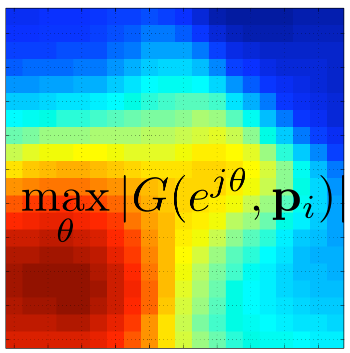
Mapa de ganancia



Mapa de
respuesta en frecuencia

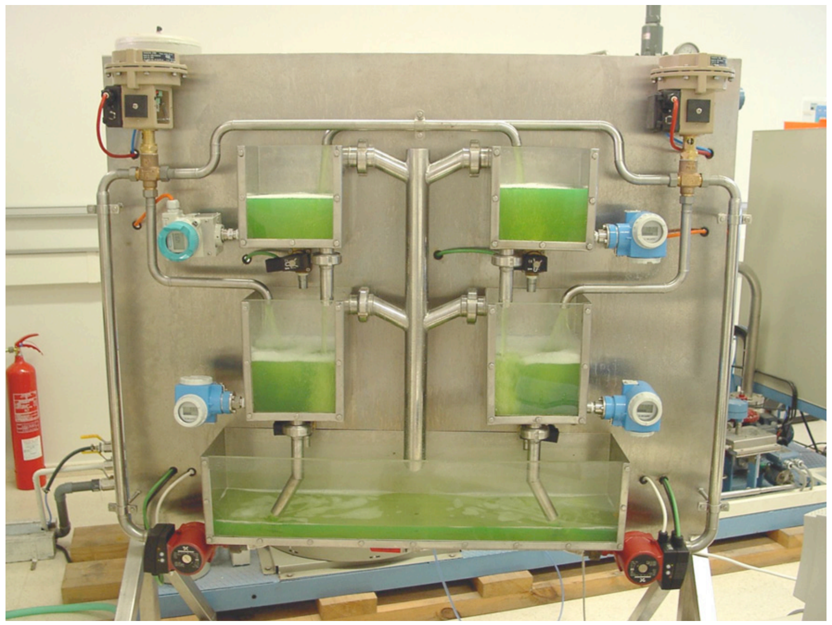


Mapa de resonancia



Dimensionality Reduction

Example: visualizing local dynamics



Dynamic model

$$y(k) = a_1 y(k-1) + a_2 y(k-2) + b_0 u(k) + b_1 u(k-1) + b_2 u(k-2)$$

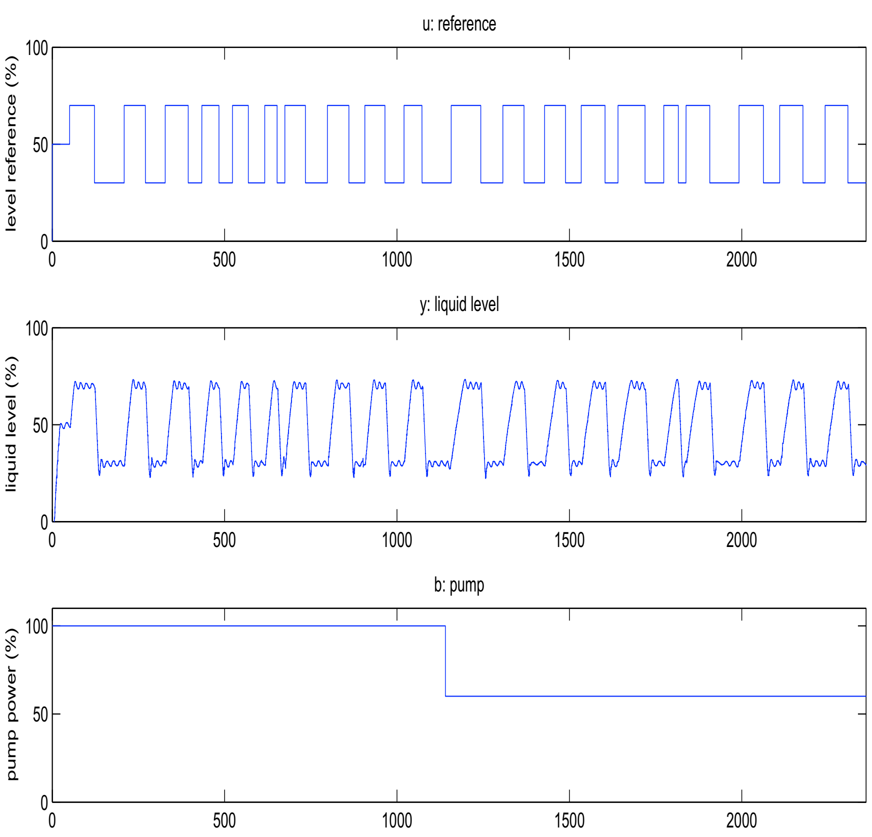
$$G(z, \mathbf{p}_i) = \frac{b_o^i + b_1^i z^{-1} + b_2^i z^{-2}}{1 + a_1^i z^{-1} + a_2^i z^{-2}}$$

State

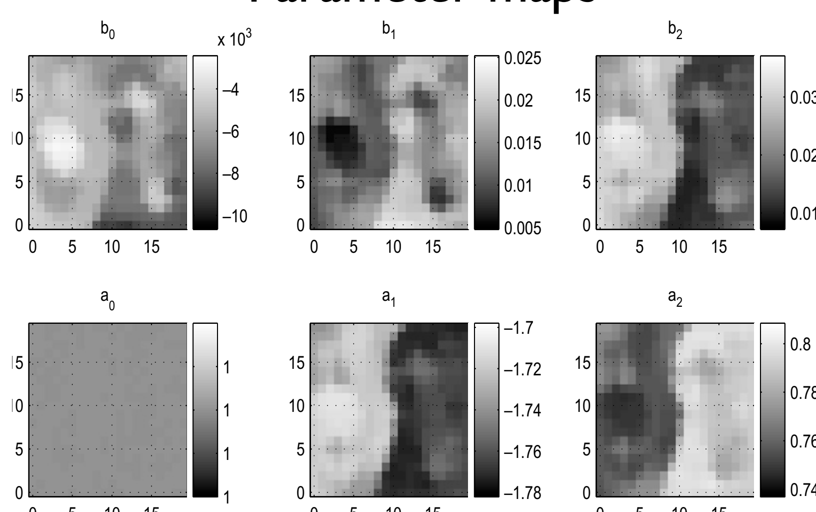
(working point)

$$\mathbf{x} = [u, y, b]^T$$

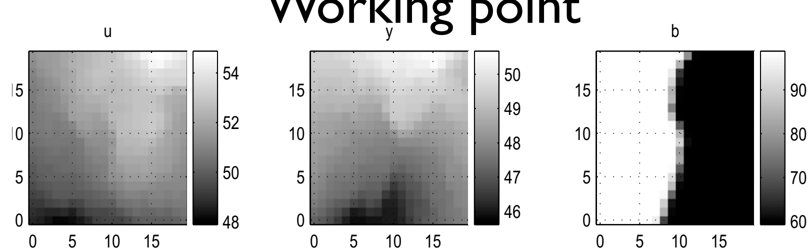
$$\boxed{\mathbf{q} = [a_1, a_2, b_0, b_1, b_2, |u, y, b]}$$



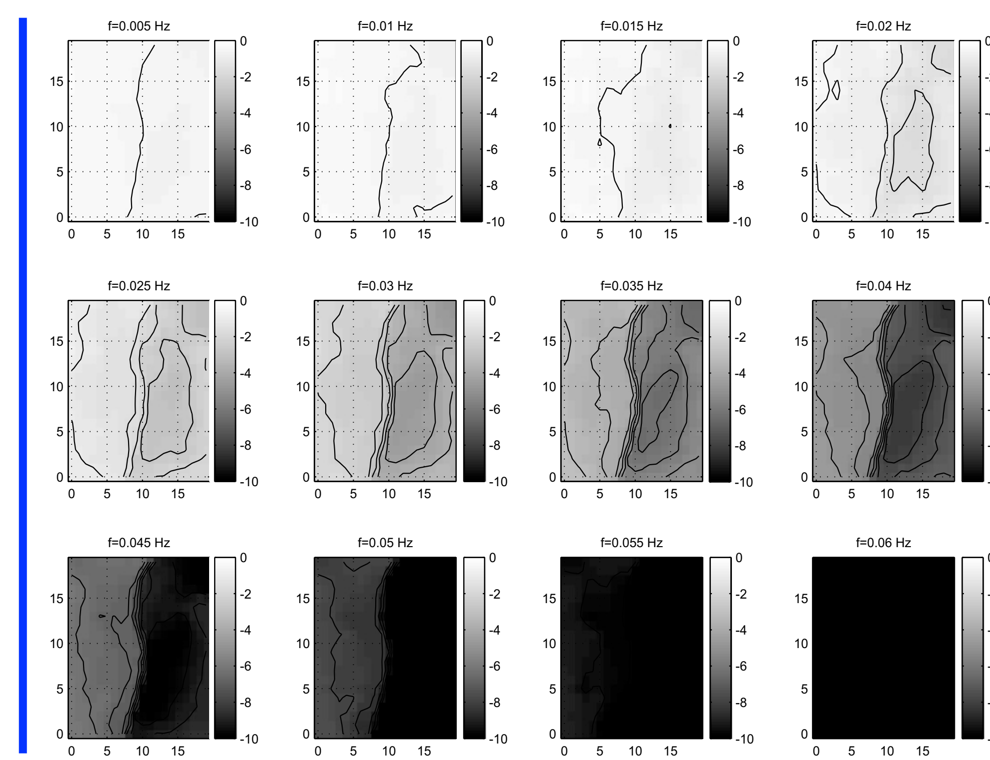
Parameter maps



Working point

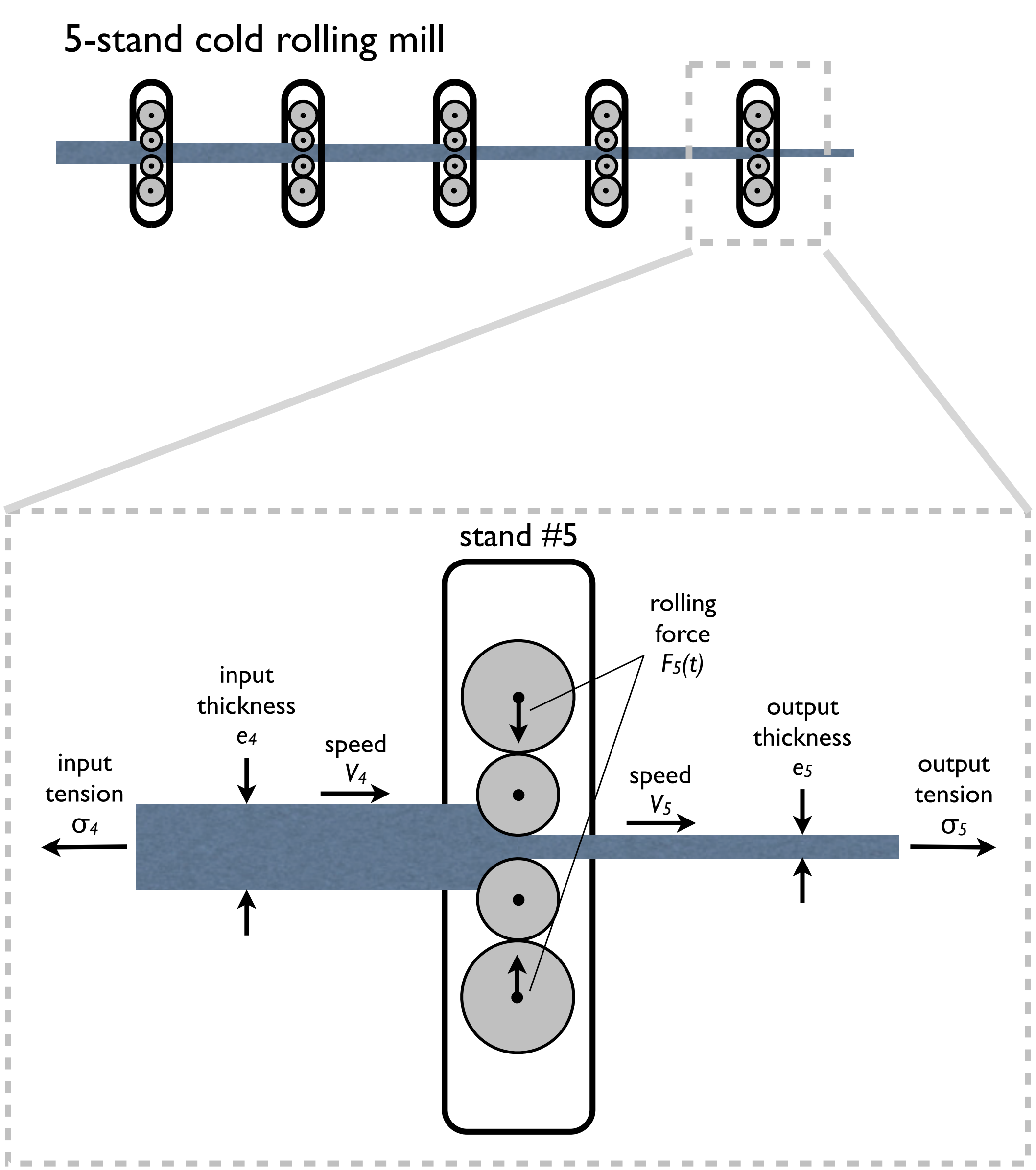


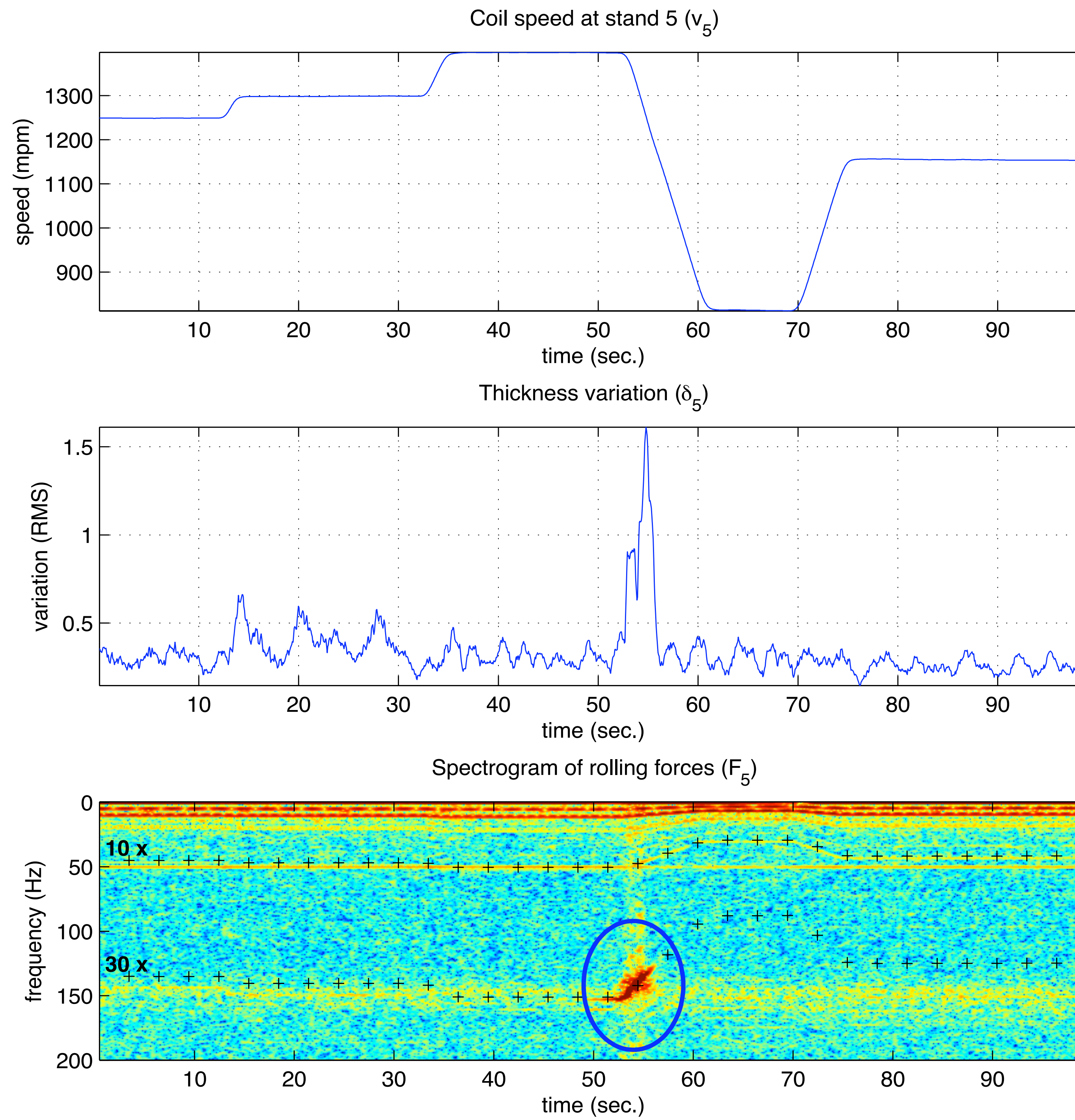
Frequency response maps



Dimensionality Reduction

Building maps of vibrational states (chatter)





Dimensionality Reduction

Building maps of vibrational states (chatter)

ingeniería isa de sistemas y automática

Universidad de Oviedo

Dynamic models (AR)

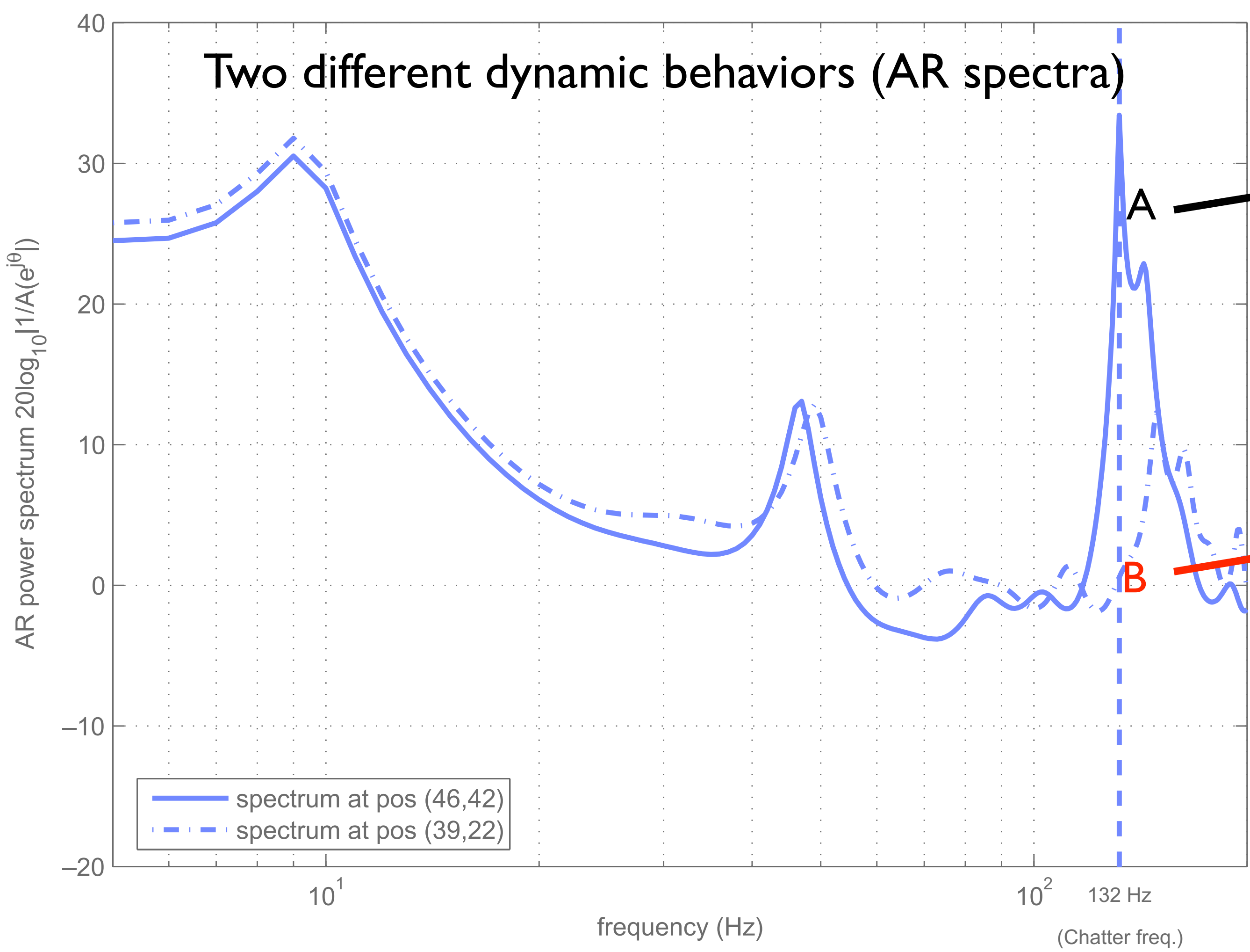
$$F_5(k) = a_1 F_5(k-1) + \cdots + a_{29} F_5(k-29) + \epsilon$$

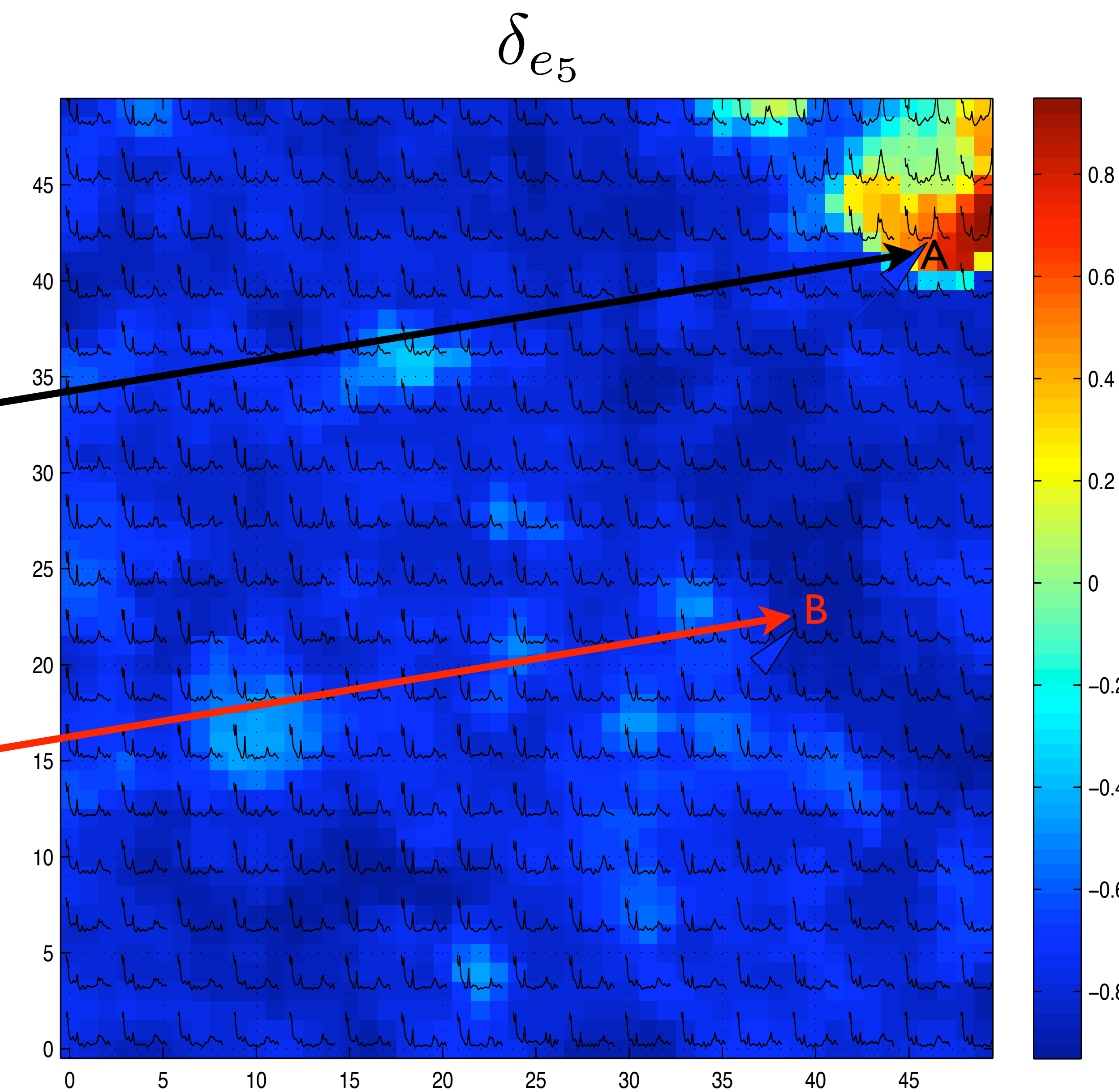
$$\mathbf{p} = [a_1, a_2, \ldots, a_{29}]^{\mathrm{T}}$$

Working point

$$\mathbf{x}_i = [\delta e_5, v_5]^T$$

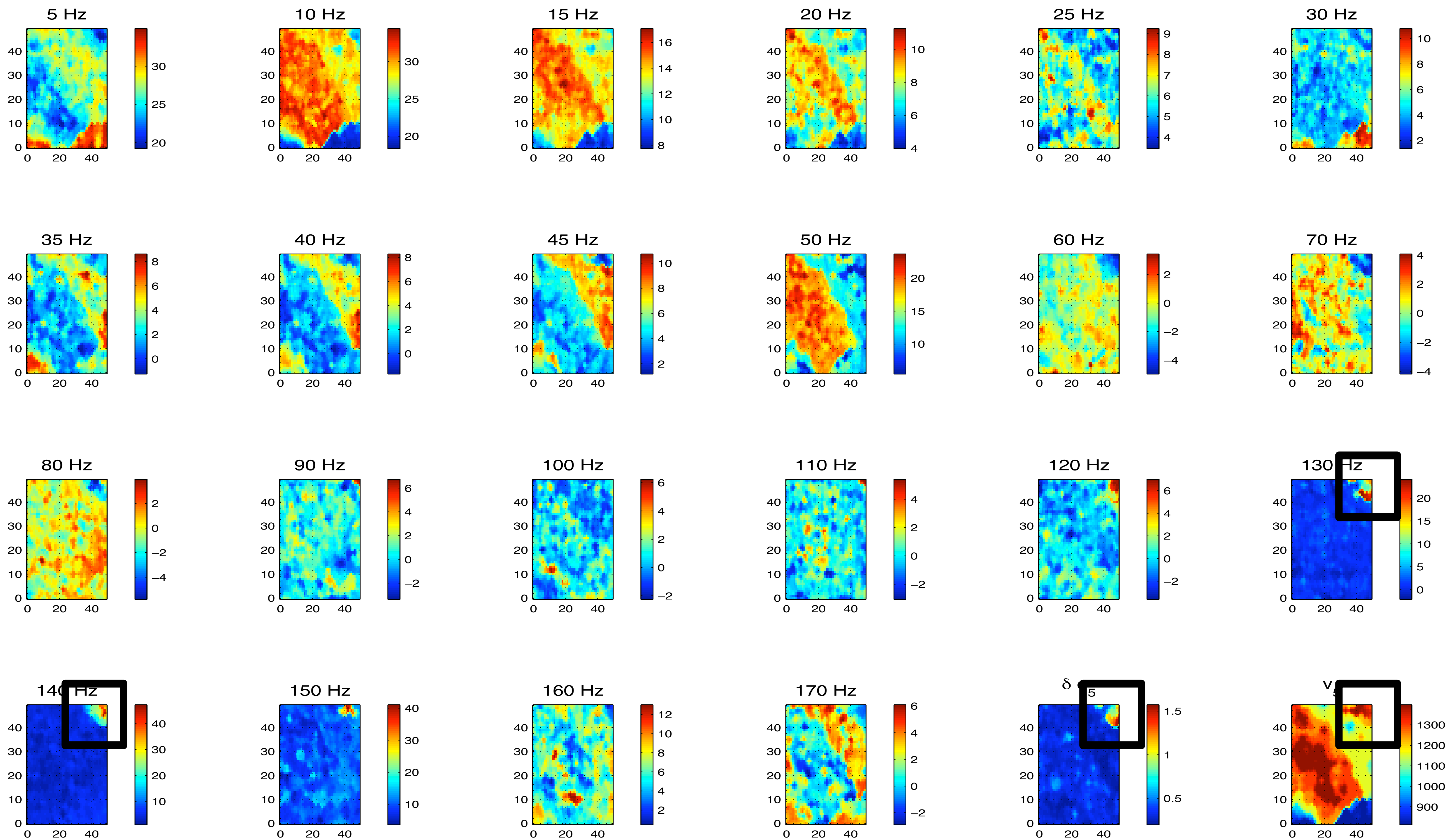
$$\mathbf{q} = [a_1, a_2, \ldots, a_{29}, |\delta e_5, v_5]^{\mathrm{T}}$$





Dimensionality Reduction

Building maps of vibrational states (chatter)



Dimensionality Reduction

vibration analysis in rotating machines

Universidad de Oviedo

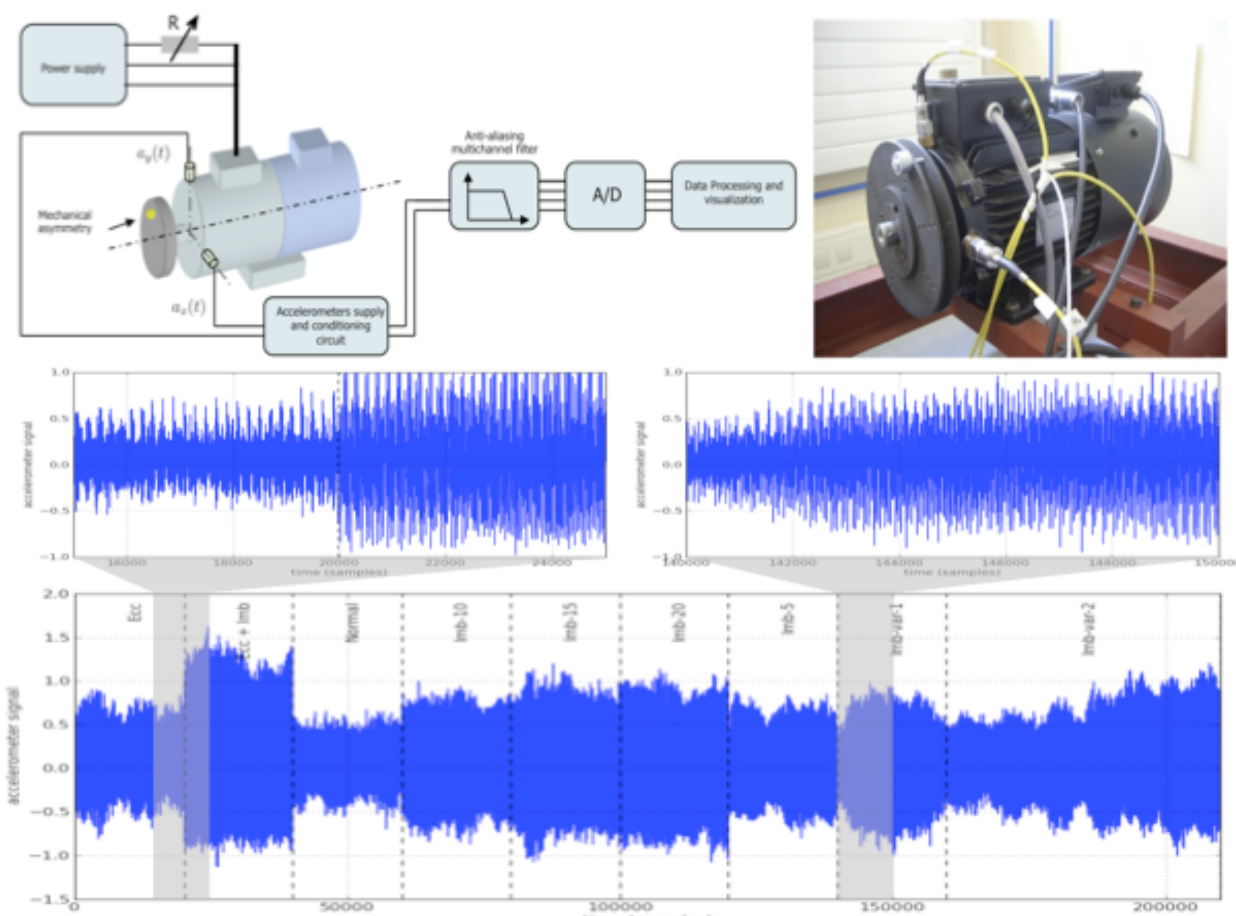
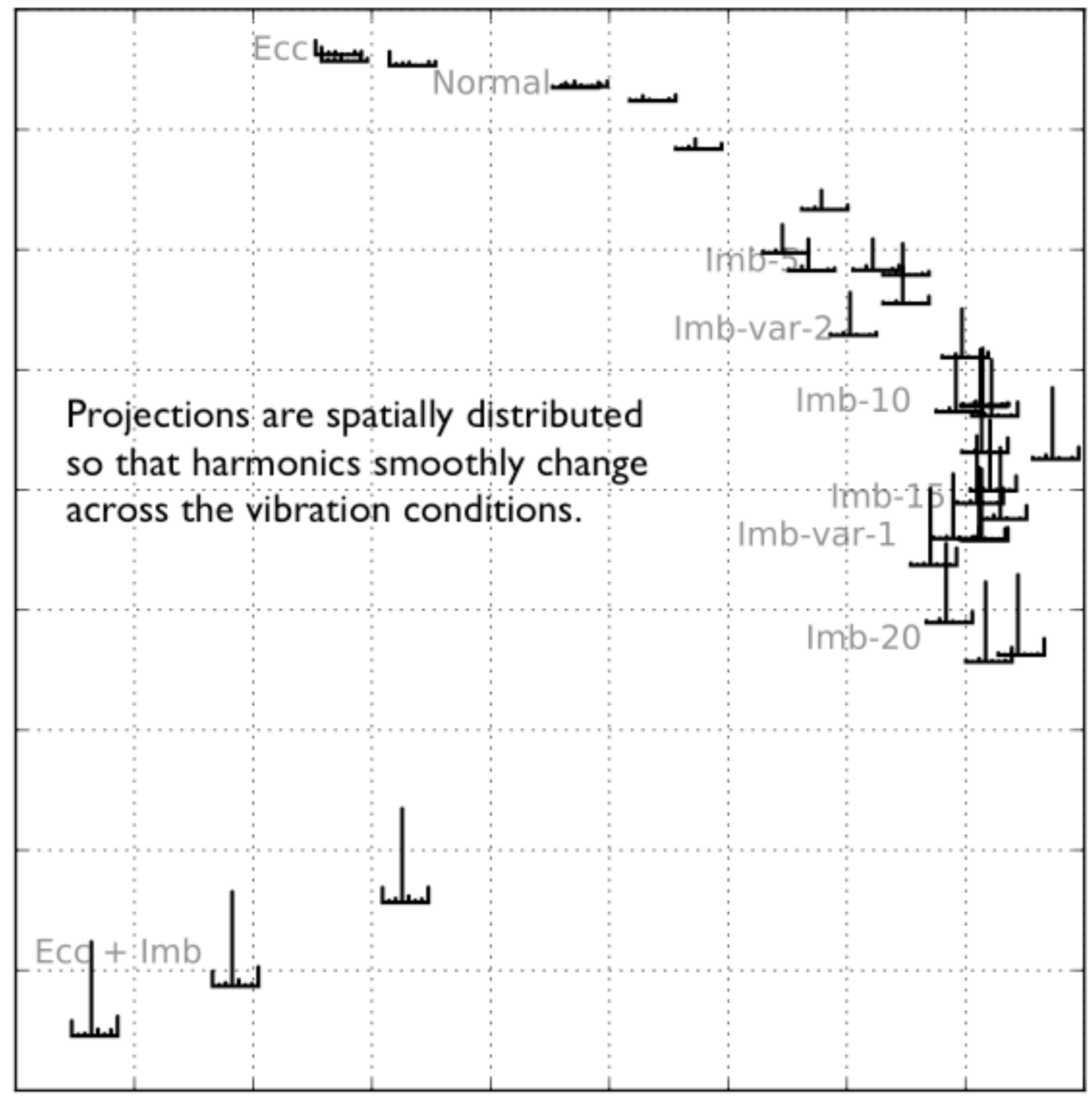


Table 1. Description of the 9 experiments.

Experiment No.	Label	Descripction
#1	Ecc	Mechanical eccentricity
#2	'Ecc+Imb'	Mechanical eccentricity + Electrical imbalance
#3	'Normal'	No eccentricity, no electrical imbalance.
#4	'Imb-10'	Electrical imbalance (fixed load, 10 Ω)
#5	'Imb-15'	Electrical imbalance (fixed load, 15 Ω)
#6	'Imb-20'	Electrical imbalance (fixed load, 20 Ω)
#7	'Imb-5'	Electrical imbalance (fixed load, 5 Ω)
#8	'Imb-var-1'	Electrical imbalance (variable load, low→high→low)
#9	'Imb-var-2'	Electrical imbalance (variable load, low→high)

map of vibrational states



Projections are spatially distributed so that harmonics smoothly change across the vibration conditions.

Dimensionality Reduction

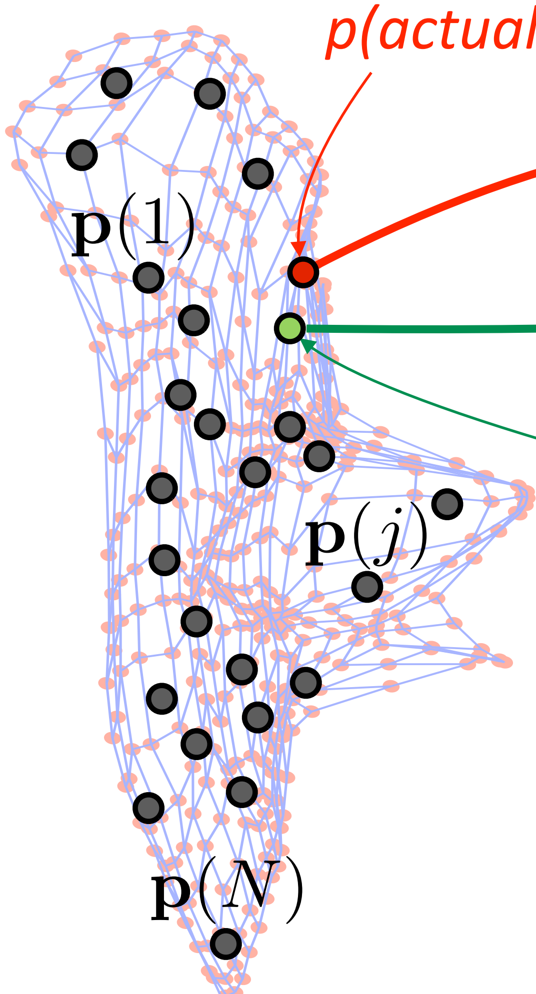
Novelty detection in dynamic behavior

Basic idea

actual process
inputs and outputs

system identification

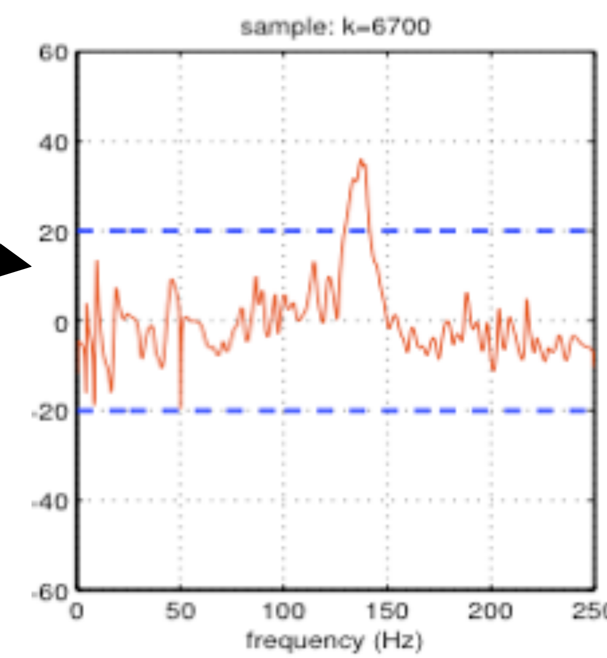
Actual dynamics
p(actual)



Best matching
dynamics *p(best)*

diff. operator
A - B

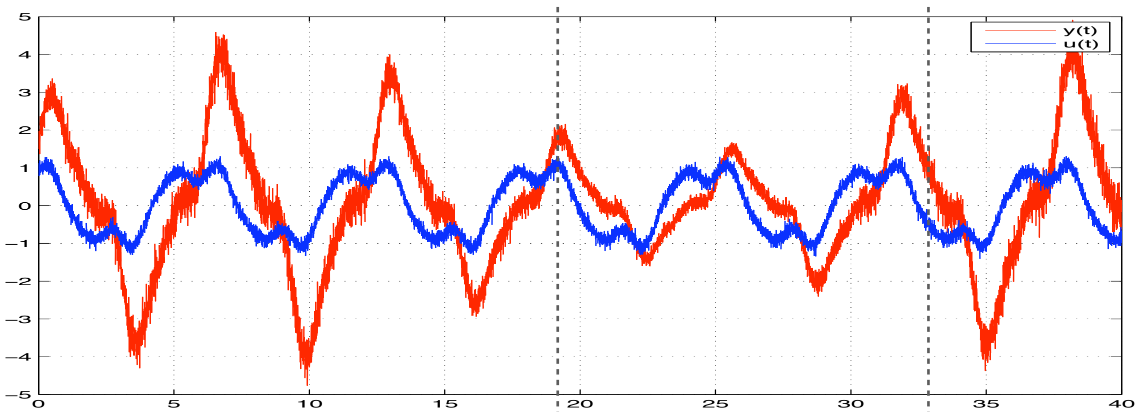
Residual dynamics



Dimensionality Reduction

Novelty detection in dynamic behavior

Example: input (blue) and output (red) from a process



$W_k = \{k - n + 1, \cdots, k\}$

Get data from k_{th} window

$$\{y(k), \varphi(k)\}_{k \in W_k}$$

minimize:

$$J = \sum_{k \in I_j} \|y(k) - f(\varphi(k), \mathbf{p}(k))\|^2$$

SOM bmu

$$c(k) = \arg\min_i \{\|\mathbf{p}(k) - \mathbf{m}_i\|\}$$

$$\mathbf{m}_{c(k)} \leftarrow \mathbf{p}(k)$$

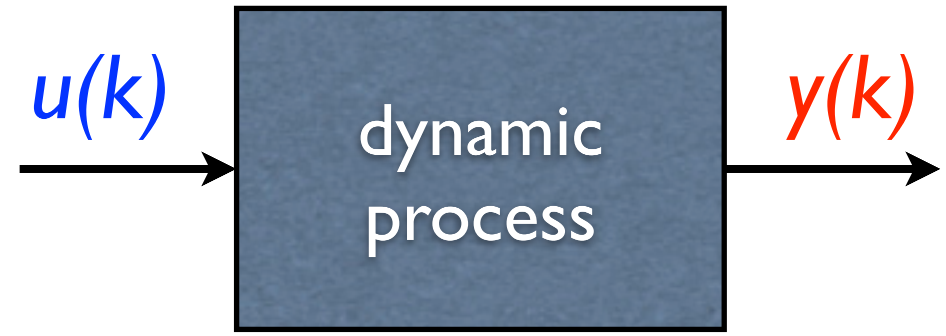
$G(e^{j\theta}, \mathbf{m}_{c(k)})$ (closest model)

$G(e^{j\theta}, \mathbf{p}(k))$ (actual model)

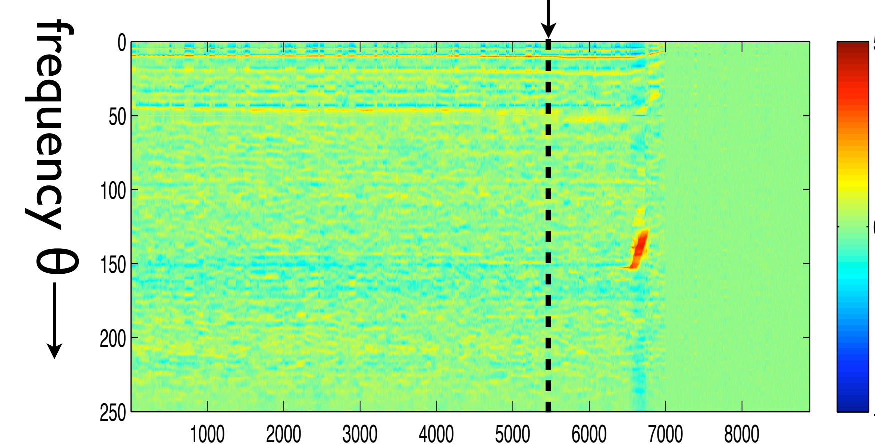
Compare models (actual vs closest)

$$\mathbf{R}(e^{j\theta}, k) = \frac{G(e^{j\theta}, \mathbf{p}(k))}{G(e^{j\theta}, \mathbf{m}_{c(k)})}$$

visualize



k_{th} window

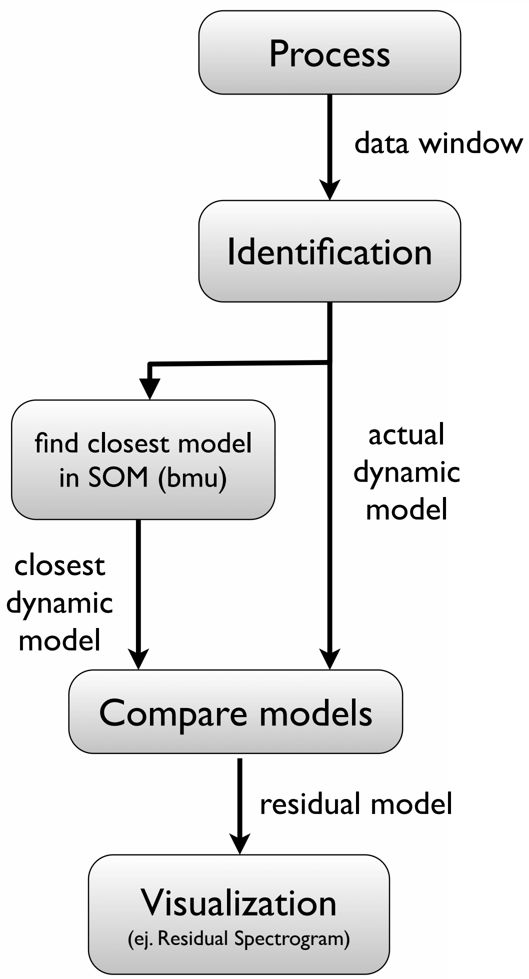


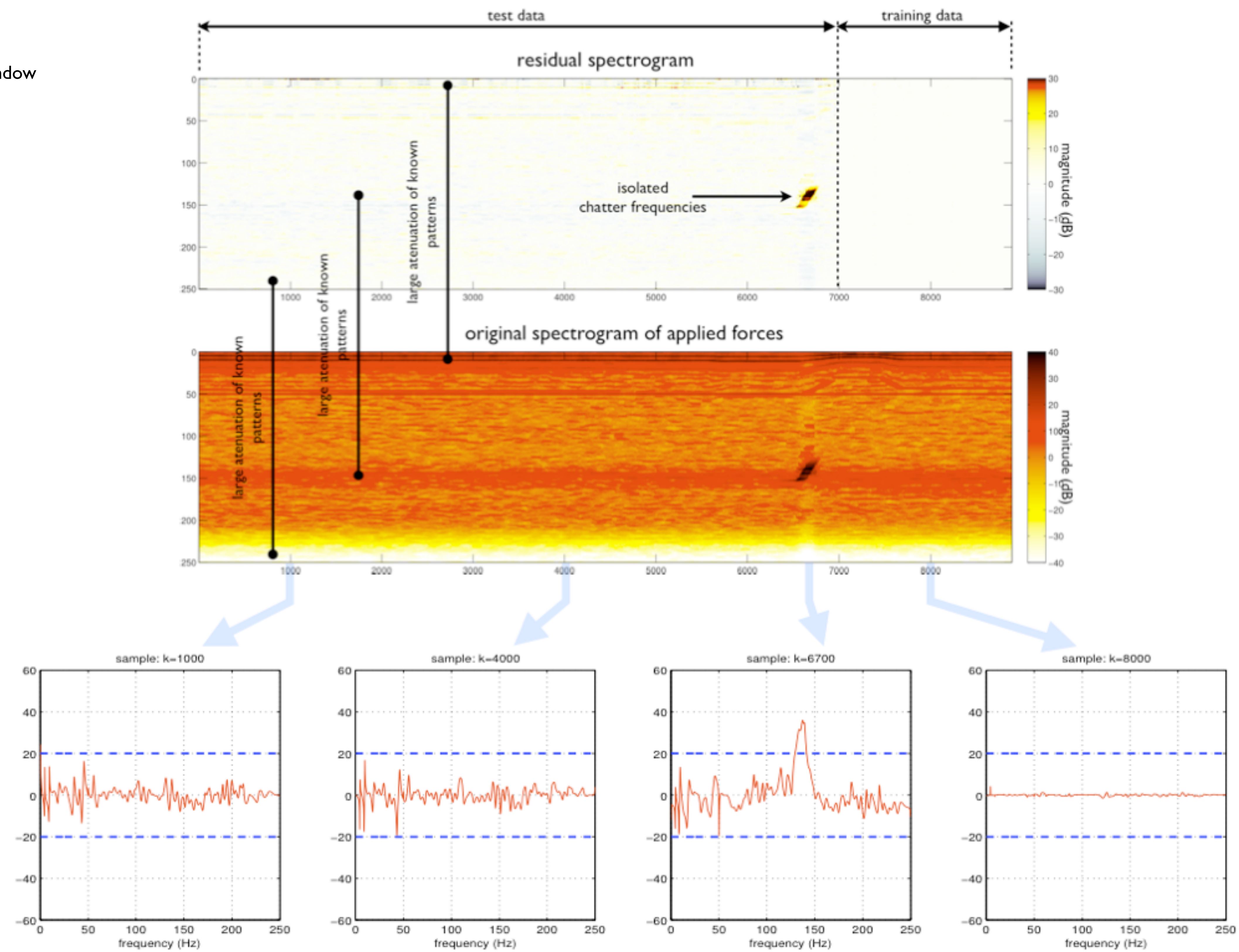
Dimensionality Reduction

Example: Isolation of harmful dynamics

ingeniería isa de sistemas y automática

Universidad de Oviedo



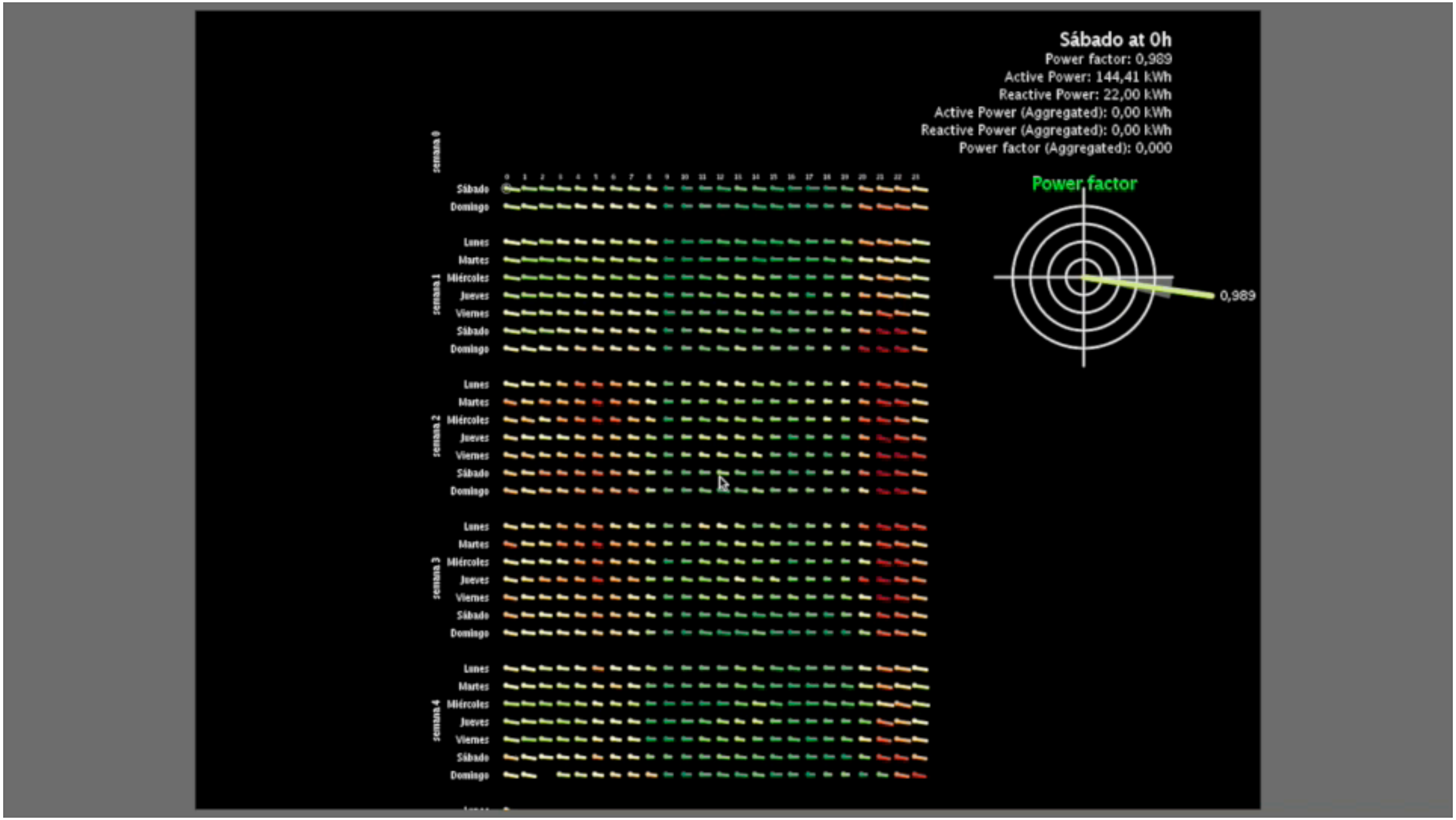


Interaction

a new dimension for data understanding

Basic interaction

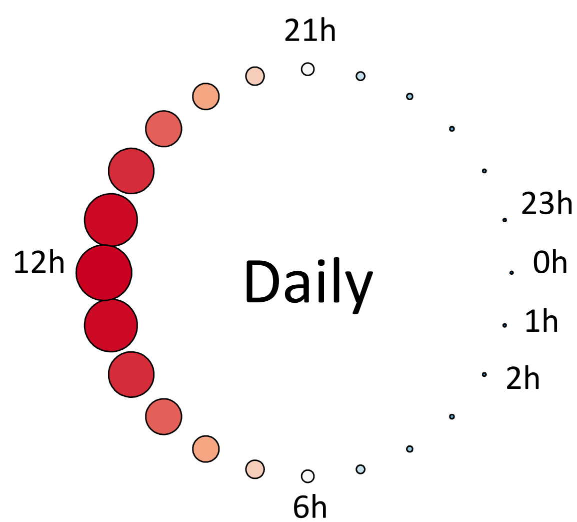
zoom, pan, hover...

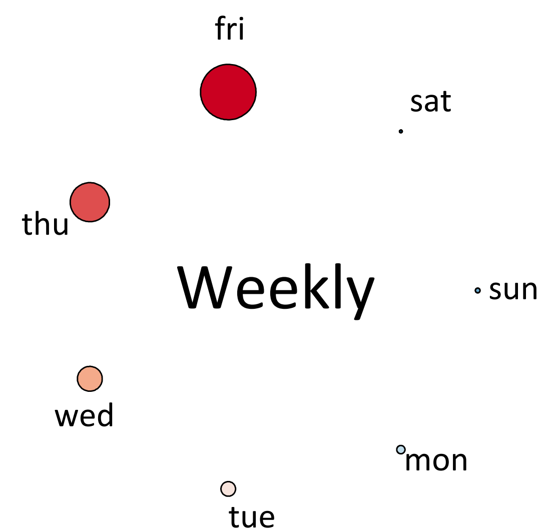


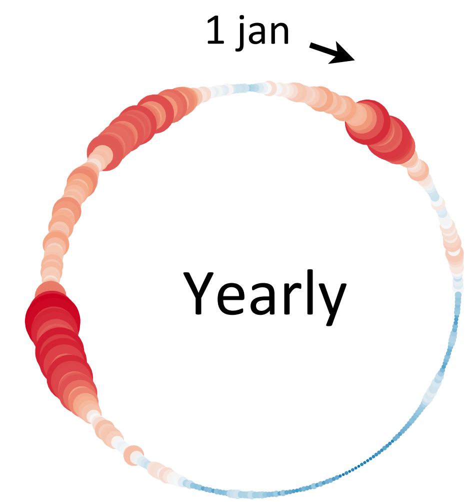
Animated Transitions

example: electric power demand analysis

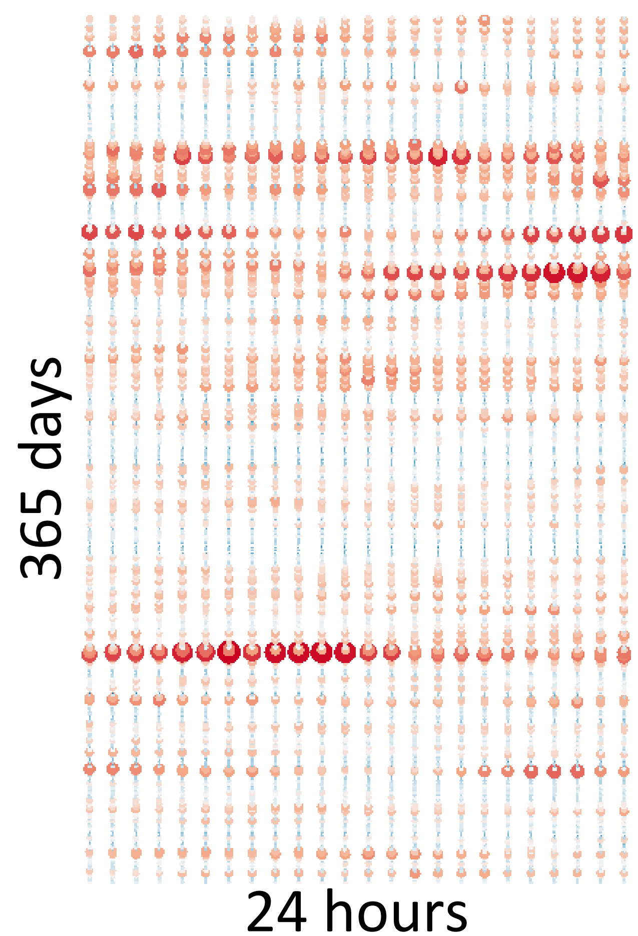
Universidad de Oviedo



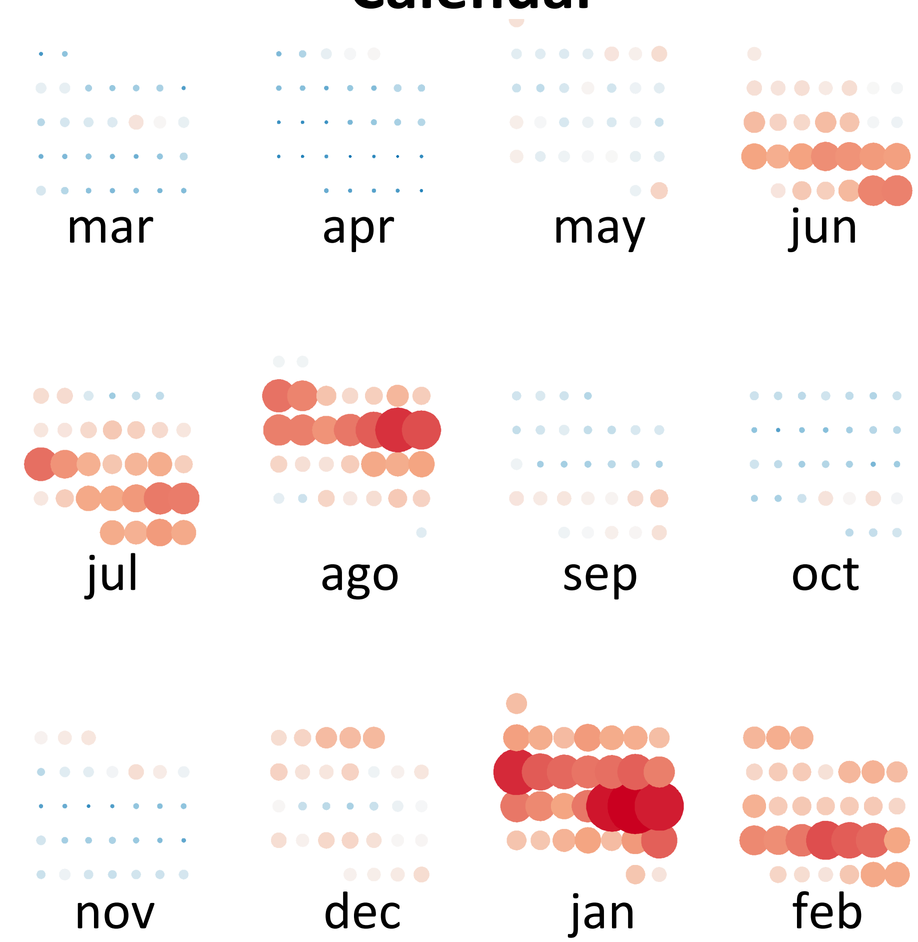




365x24



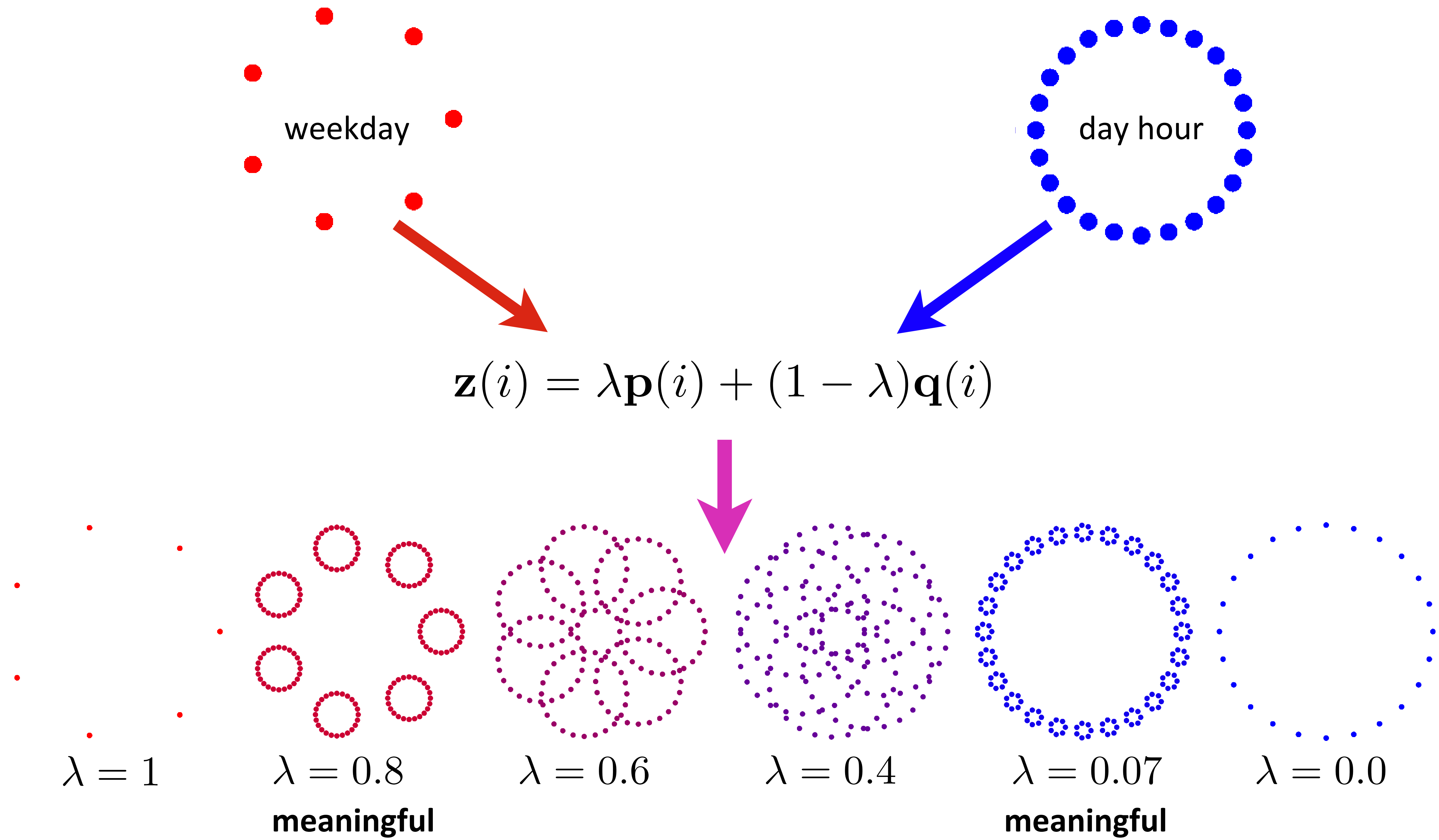
Calendar



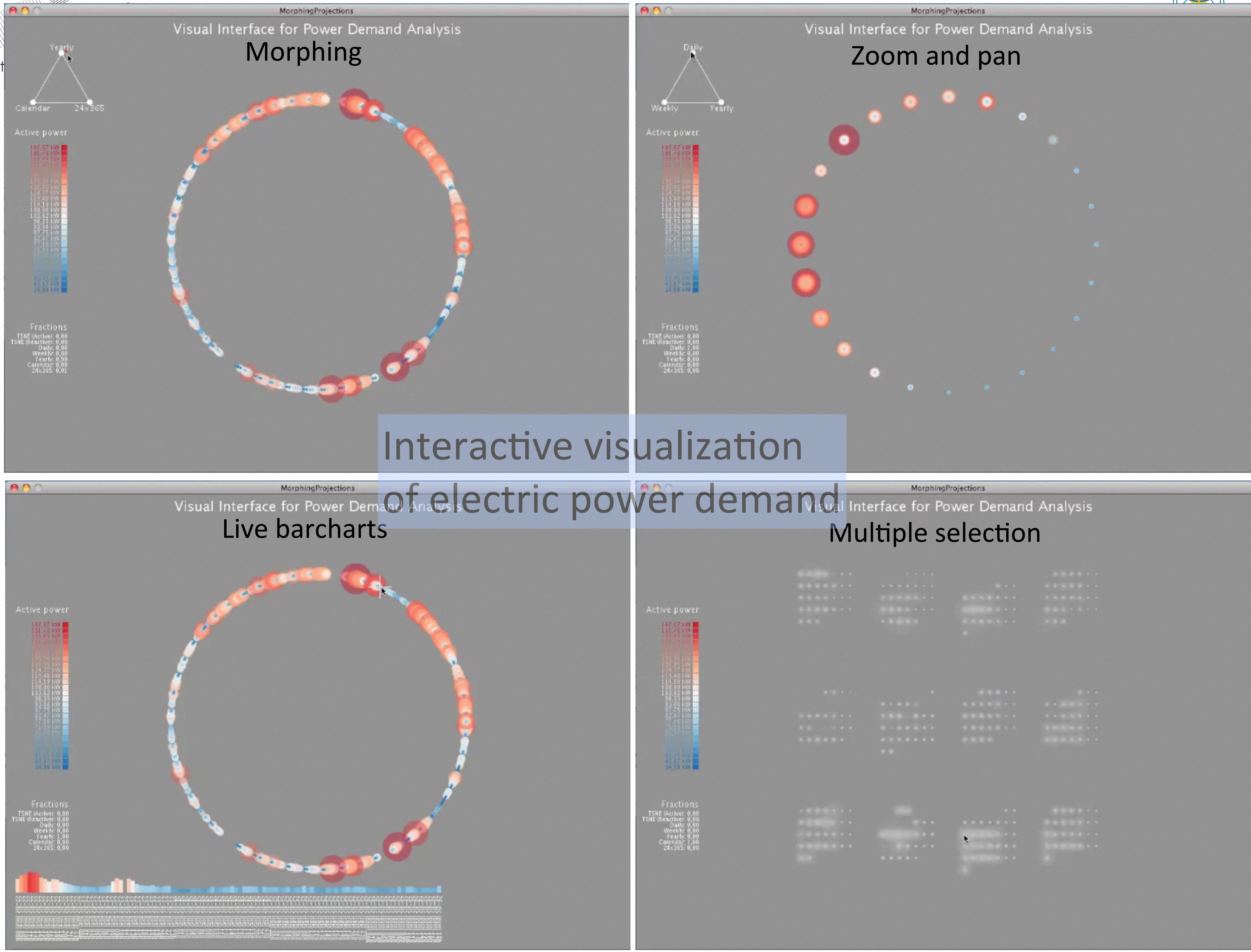
Animated Transitions

morphing operation

Universidad de Oviedo



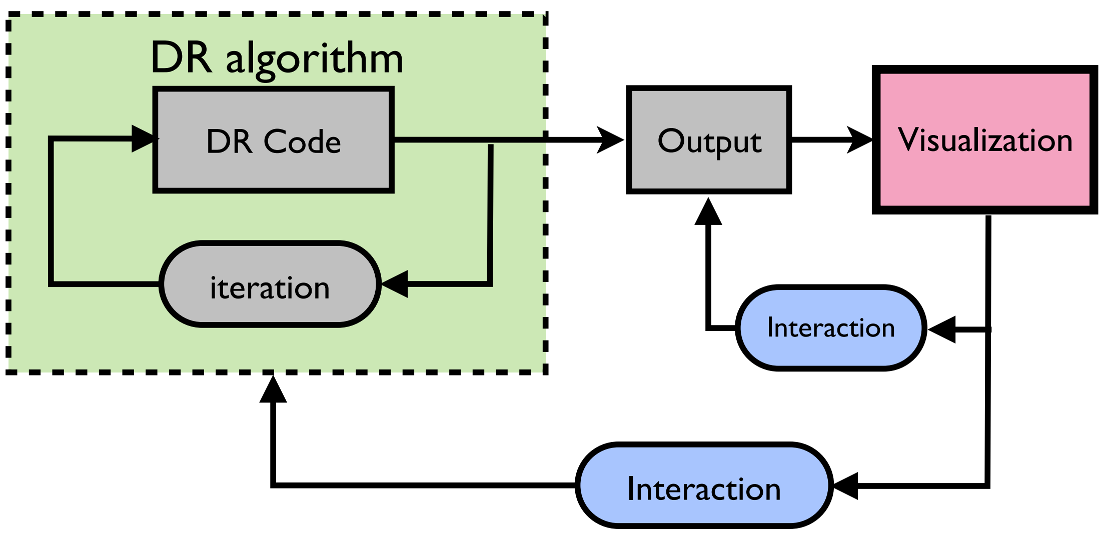
Interactive visualization of electric power demand



Interactive Dimensionality Reduction (iDR)

Universidad de Oviedo

plain DR

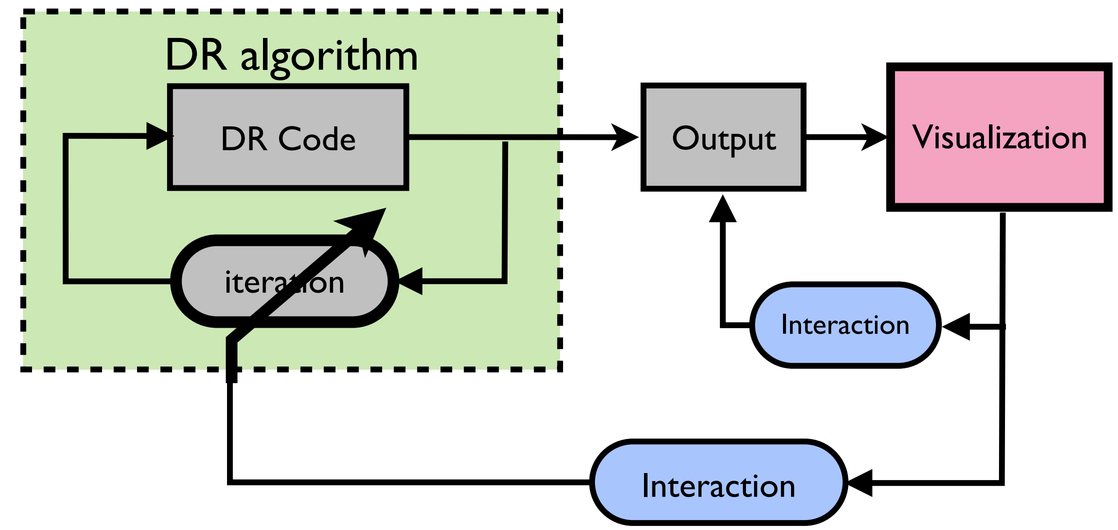


batch mode interaction scheme

set initial configuration
run until convergence (N steps)
visualize the results
fine-tune the DR algorithm

slow!
$\sim 10^{1}$s

interactive DR



iDR interaction scheme

set initial configuration

run 1step
visualize the results

fine-tune the DR algorithm

quick!
$\sim 10^{-1}$s

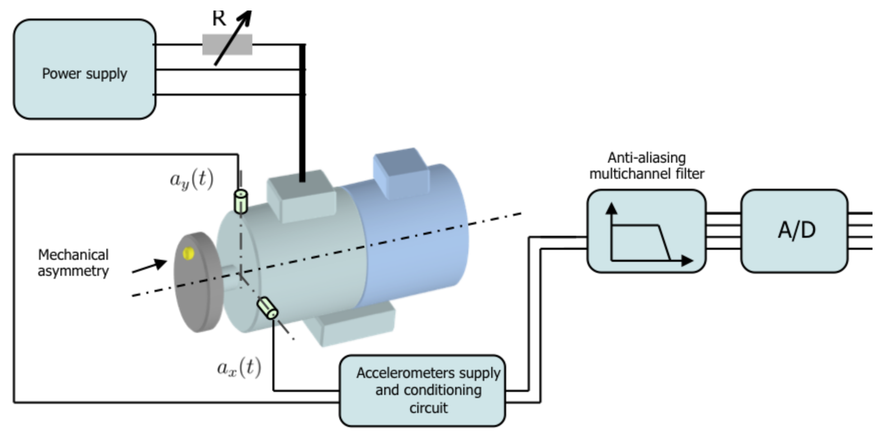
enables a quick feedback to the user and hence a much better user integration in the exploration process

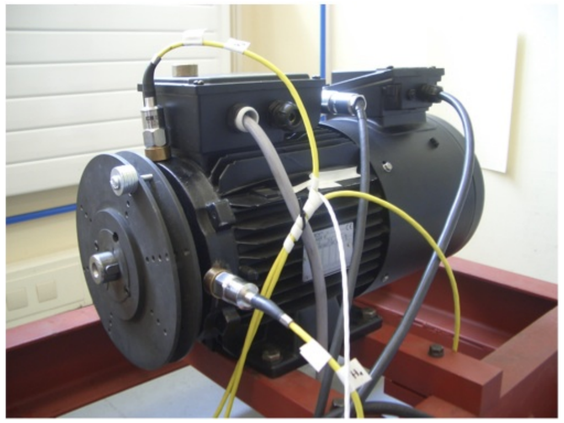
ingeniería isa de sistemas y automática

Interactive Dimensionality Reduction

fault analysis of AC motor

Universidad de Oviedo



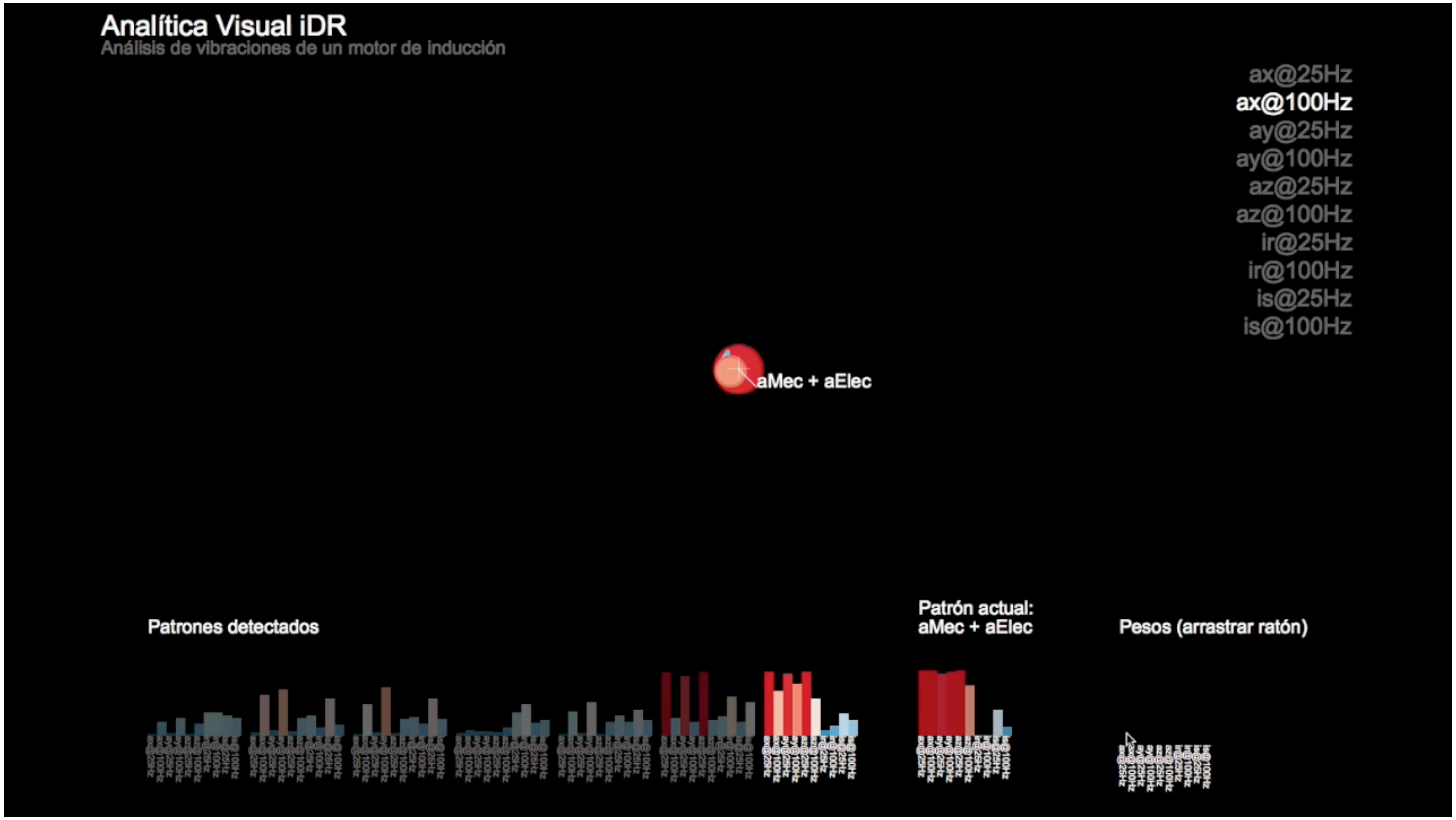


Application interface with iDR user-driven modification of the input metric space

Javascript application using processingjs (http://processingjs.org)

Analysis of three vibration signals $a_x(t)$ $a_y(t)$ $a_z(t)$ and two phase currents $i_R(t)$, $i_S(t)$

of a 4kW 2 pole-pair asynchronous motor http://isa.uniovi.es/~idiaz/demos/iDR-vibracionesMotor/



Thank you for your attention!