

Una herramienta de Machine Learning para la actualización del Directorio Empresas y Establecimientos con Actividad Económica en Andalucía

Rafael Blanquero¹, Elisa Isabel Caballero Ruiz², Emilio Carrizosa¹,
Marina Enguidanos Weyler², Ana Gema Galera Pozo², Nuria Gómez-Vargas¹,
Jasone Ramírez-Ayerbe¹

¹Instituto de Matemáticas de la Universidad de Sevilla (IMUS), Seville, Spain

²Instituto de Estadística y Cartografía de Andalucía (IECA), Seville, Spain

Lleidatech 2023, Lleida, 27/10/23

1 Approaching the problem

2 Text Mining

- Web Scraping
- Preprocessing

3 Grouping and feature selection techniques

4 Classification model

- Random Forest
- Interpretation of results and definition of innovation

5 Conclusions and further work

Approaching the problem

Characterizing a company



(a) Official surveys, General Business Register, Chamber of Commerce, etc.



(b) Data providing companies

Figure: Traditional methodology

Approaching the problem

Characterizing a company



(a) Official surveys, General Business Register, Chamber of Commerce, etc.



(b) Data providing companies

Figure: Traditional methodology

Proposed new methodology: **Web scraping** [ten Bosch et al., 2018]

- **Field of growing importance** for Official Statistics Institutes.
- Valuable way to **supplement** administrative sources and metadata systems.
- Web data are more **volatile** and usually **unstructured** but in many cases also **richer** and **more frequently updated**.

- 1 Approaching the problem
- 2 **Text Mining**
 - Web Scraping
 - Preprocessing
- 3 Grouping and feature selection techniques
- 4 Classification model
 - Random Forest
 - Interpretation of results and definition of innovation
- 5 Conclusions and further work

State-of-the-Art

- Linkage between business websites addresses (URLs) and a business population frame [Van Delden et al., 2019].
- Supervised Learning model to determine if a company is innovative by studying the text on its website [Daas and van der Doef, 2020].

State-of-the-Art

- Linkage between business websites addresses (URLs) and a business population frame [Van Delden et al., 2019].
- Supervised Learning model to determine if a company is innovative by studying the text on its website [Daas and van der Doef, 2020].

State-of-the-Art

- Linkage between business websites addresses (URLs) and a business population frame [Van Delden et al., 2019].
- Supervised Learning model to determine if a company is innovative by studying the text on its website [Daas and van der Doef, 2020].



Figure: Scraped variables

- Removal of characters: numbers and punctuation marks.
- Converting words to lower case.
- Words cleaning by language detection:
 - Detection and removal of stop words.
 - Stemming (mapping of the different morphological variants to their base form).

```
def limpiar(lista_palabras):  
    stemmer = SnowballStemmer("spanish")  
    stop_words = stopwords.words('spanish')  
    lista_filtrada = [p.lower() for p in lista_palabras if(p.lower() not in stop_words and p.isalpha())]  
    lista_final = [stemmer.stem(p) for p in lista_filtrada]  
  
    return lista_final
```

```
print(limpiar(pag_parsed))
```

```
['accesibil', 'corre', 'rankings', 'iguald', 'emprend', 'histori', 'eleccion', 'conveni', 'normat', 'bous', 'contact', 'estud  
i', 'grad', 'master', 'doctor', 'investig', 'investig', 'crai', 'fius', 'convocatori', 'doctor', 'bibliotec', 'cultur', 'depor  
t', 'agend', 'campus', 'empres', 'emprend', 'ebc', 'patent', 'catedr', 'internacional', 'alianz', 'conveni', 'ogpi', 'rankin  
g', 'cooper', 'profesor', 'pas', 'directori', 'estudi', 'profesor', 'pas', 'alumni', 'search', 'rankings', 'iguald', 'emprend',  
'histori', 'eleccion', 'conveni', 'normat', 'bous', 'contact', 'estudi', 'grad', 'master', 'doctor', 'investig', 'investig', 'c  
rai', 'fius', 'convocatori', 'doctor', 'bibliotec', 'cultur', 'deport', 'agend', 'campus', 'empres', 'emprend', 'ebc', 'paten  
t', 'catedr', 'internacional', 'alianz', 'conveni', 'ogpi', 'rankings', 'cooper', 'profesor', 'pas', 'directori', 'estudi', 'pr  
ofesor', 'pas', 'alumni', 'aup', 'estudi', 'actual', 'univers', 'estudi', 'investig', 'cultur', 'deport', 'vist', 'vist', 'vis  
t', 'vist', 'vist', 'vist', 'vist', 'estudi', 'investig', 'estudi', 'transparent', 'destac', 'estudi', 'investig', 'emp  
res', 'directori', 'editorial', 'editorial', 'search', 'encuentran', 'encuentran', 'rankings', 'iguald', 'emprend', 'histori',  
'eleccion', 'conveni', 'normat', 'bous', 'contact', 'estudi', 'grad', 'master', 'doctor', 'investig', 'investig', 'crai', 'fius  
's', 'convocatori', 'doctor', 'bibliotec', 'cultur', 'deport', 'agend', 'campus', 'empres', 'emprend', 'ebc', 'patent', 'cated  
r', 'internacional', 'alianz', 'conveni', 'ogpi', 'rankings', 'cooper', 'profesor', 'pas', 'directori']
```

Figure: Processed text

- 1 Approaching the problem
- 2 Text Mining
 - Web Scraping
 - Preprocessing
- 3 Grouping and feature selection techniques
- 4 Classification model
 - Random Forest
 - Interpretation of results and definition of innovation
- 5 Conclusions and further work

Clustering

- Grouping words by their weights β_j in a classification model.
 - Penalized logistic regression: ridge regression

$$\min_{\beta} \sum_{i=1}^n (y_i - \beta_0 - \sum_{j=1}^p \beta_j x_{ij})^2 + \lambda \cdot \sum_{j=1}^p \beta_j^2, \lambda > 0$$

- Density-based clustering: DBSCAN [Ester et al., 1996].
 - Designed for the clustering of large noisy datasets.
 - Clusters with arbitrary shape.
 - Detection of outliers.

Clustering

- Grouping words by their weights β_j in a classification model.
 - Penalized logistic regression: ridge regression

$$\min_{\beta} \sum_{i=1}^n (y_i - \beta_0 - \sum_{j=1}^p \beta_j x_{ij})^2 + \lambda \cdot \sum_{j=1}^p \beta_j^2, \lambda > 0$$

- Density-based clustering: DBSCAN [Ester et al., 1996].
 - Designed for the clustering of large noisy datasets.
 - Clusters with arbitrary shape.
 - Detection of outliers.

Clustering

- Grouping words by their weights β_j in a classification model.
 - Penalized logistic regression: ridge regression

$$\min_{\beta} \sum_{i=1}^n (y_i - \beta_0 - \sum_{j=1}^p \beta_j x_{ij})^2 + \lambda \cdot \sum_{j=1}^p \beta_j^2, \lambda > 0$$

- Density-based clustering: DBSCAN [Ester et al., 1996].
 - Designed for the clustering of large noisy datasets.
 - Clusters with arbitrary shape.
 - Detection of outliers.

Clustering

- Grouping words by their weights β_j in a classification model.
 - Penalized logistic regression: ridge regression

$$\min_{\beta} \sum_{i=1}^n (y_i - \beta_0 - \sum_{j=1}^p \beta_j x_{ij})^2 + \lambda \cdot \sum_{j=1}^p \beta_j^2, \lambda > 0$$

- Density-based clustering: DBSCAN [Ester et al., 1996].
 - Designed for the clustering of large noisy datasets.
 - Clusters with arbitrary shape.
 - Detection of outliers.

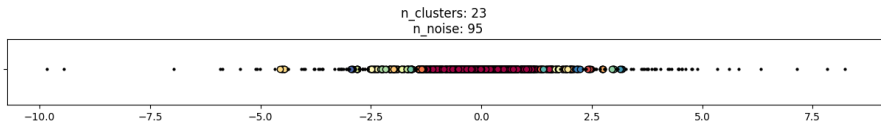


Figure: Clustering according to the weights

Fisher's Exact Test

Determining the dependence between two categorical variables: X : the variable appears **yes/no** in the company's website; Y : the company is **innovative yes/no**.

	Appearance	Non-appearance
Innovative	a	b
Non-innovative	c	d

Table: Contingency table

p-value \leq threshold \rightarrow we assume that there is a significant dependence between the appearance and the innovative character.

Grouping and feature selection techniques

Fisher's Exact Test

Determining the dependence between two categorical variables: X : the variable appears **yes/no** in the company's website; Y : the company is **innovative yes/no**.

	Appearance	Non-appearance
Innovative	a	b
Non-innovative	c	d

Table: Contingency table

$p\text{-value} \leq \text{threshold} \rightarrow$ we assume that there is a significant dependence between the appearance and the innovative character.

```
Tabla de contingencia para la palabra: adult
adult      False True
INNOVACION
0           515   13
1           797    3
p_valor: 0.001104618269509042
```

(a) Rejecting independence

```
Tabla de contingencia para la palabra: afric
afric      False True
INNOVACION
0           525    3
1           793    7
p_valor: 0.7482567547446752
```

(b) Assuming independence

Figure: Examples of Fisher's Test

Grouping and feature selection techniques

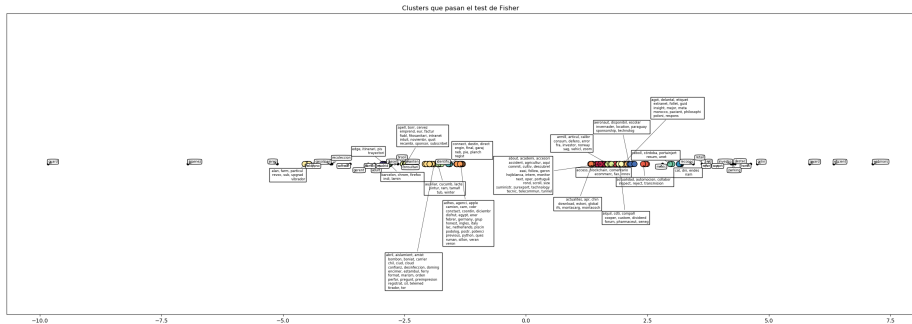


Figure: Selected clusters for the features of the model

- 1 Approaching the problem
- 2 Text Mining
 - Web Scraping
 - Preprocessing
- 3 Grouping and feature selection techniques
- 4 Classification model
 - Random Forest
 - Interpretation of results and definition of innovation
- 5 Conclusions and further work

Classification model

Random Forest

Model variables

- 1310 instances (companies).
- 4336 features 87 features:
 - 4204 words 48 clusters (30 noise + 18 groups).
 - 132 html tags 39 html tags.
- Response: Innovative YES/NO.

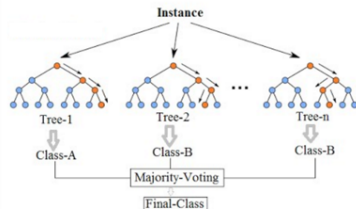


Figure: Machine Learning model: Random Forest

Classification model

Random Forest

Model variables

- 1310 instances (companies).
- 4336 features 87 features:
 - 4204 words 48 clusters (30 noise + 18 groups).
 - 132 html tags 39 html tags.
- Response: Innovative YES/NO.

		Prediction label	
		1	0
True label	1	228	34
	0	92	79

Table: Confusion matrix over test set

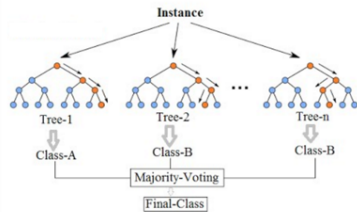


Figure: Machine Learning model: Random Forest

Effectiveness measures

- 0.7 CV accuracy
- 0.87 sensitivity
- 0.46 specificity

Classification model

Interpretation of results and definition of innovation

PDP interact for "etiqueta_long_head" and "etiqueta_long_body"

Number of unique grid points: (etiqueta_long_head: 10, etiqueta_long_body: 10)

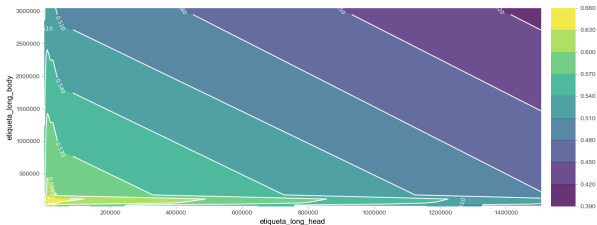


Figure: Partial dependence plot (PDP): *long_head* and *long_body*

PDP for feature "n_imagenes"

Number of unique grid points: 10



Figure: Individual Conditional Expectation (ICE) plot: *n_imagenes*

- 1 Approaching the problem
- 2 Text Mining
 - Web Scraping
 - Preprocessing
- 3 Grouping and feature selection techniques
- 4 Classification model
 - Random Forest
 - Interpretation of results and definition of innovation
- 5 Conclusions and further work

Conclusions

- The webpage content of innovative companies in Andalusia is more homogeneous than that of non-innovative ones.
- How the webpage is constructed prevails over its content.
- It is possible to carry out a study in time of the concept of innovation through the importance of the variables.

Conclusions

- The webpage content of innovative companies in Andalusia is more homogeneous than that of non-innovative ones.
- How the webpage is constructed prevails over its content.
- It is possible to carry out a study in time of the concept of innovation through the importance of the variables.

Further work

- Extrapolate the model to small companies.
- Enhance interpretability:
 - Exploiting background knowledge, e.g., semantic meaning clustering: Must-Link and Cannot-Link constraints (C-DBSCAN [Ruiz et al., 2010]).

- Piet JH Daas and Suzanne van der Doef. Detecting innovative companies via their website. *Statistical Journal of the IAOS*, 2020.
- Martin Ester, Hans-Peter Kriegel, Jörg Sander, Xiaowei Xu, et al. A density-based algorithm for discovering clusters in large spatial databases with noise. In *kdd*, volume 96, 1996.
- Carlos Ruiz, Myra Spiliopoulou, and Ernestina Menasalvas. Density-based semi-supervised clustering. *Data mining and knowledge discovery*, 21(3), 2010.
- Olav ten Bosch, Dick Windmeijer, Arnout van Delden, and Guido van den Heuvel. Web scraping meets survey design: combining forces. In *Big Data Meets Survey Science Conference, Barcelona, Spain*, 2018.
- Arnout Van Delden, Dick Windmeijer, and Olav Ten Bosch. Finding enterprise websites. In *European Establishment Statistics Workshop, Bilbao, Spain*, 2019.

Thank you for your attention!

`ecarrizosa@us.es`

This work has been carried out within the framework of the research project “Una herramienta de Machine Learning para la actualización y el desarrollo del Directorio de Empresas y Establecimientos con actividad en Andalucía” CEI-23-FQM329, which has received funding from the Junta de Andalucía (Consejería de Economía, Conocimiento, Empresas y Universidad).

Transferencia de conocimiento matemático a la industria y las organizaciones

Emilio Carrizosa,
Presidente de math-in

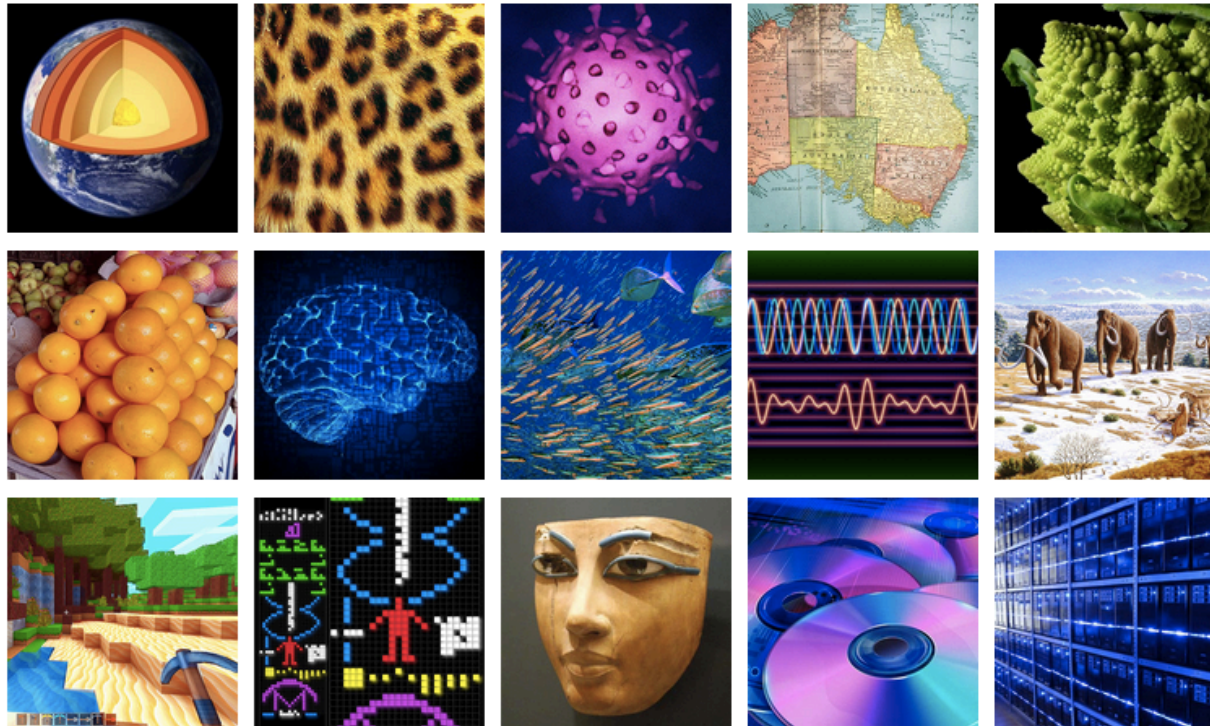
Lleida, 27 de octubre de 2023



LAS MATEMÁTICAS ESTÁN EN TODAS PARTES

Más sobre el Día Internacional
de las Matemáticas

A Language ▾





DÍA INTERNACIONAL DE LAS
MATEMÁTICAS
14 DE MARZO

**MATEMÁTICAS PARA UN
MUNDO MEJOR**

Más sobre el Día Internacional
de las Matemáticas

A Language ▾





Spanish Network for Mathematics in Industry

Descubre el poder de las matemáticas.

La Red Española de Matemática-Industria pone todo el poder de las matemáticas al servicio de tu organización.

Consúltanos



La mayor red de profesionales
especializados en
matemáticas

Members



math-in research groups

- 31 associated members:
 - 27 research groups belonging to 18 universities and research centres.
 - 1 institutional member
 - 3 sponsor members.
- +400 researchers



Universidad de Valladolid

Universida de Vigo



POLITÉCNICA



UNIVERSIDAD DE ALMERÍA



Universidad del País Vasco Euskal Herriko Unibertsitatea



27
Research
Groups



+400
Researchers in
Applied &
Industrial
Mathematics

Modelling Weeks



IX IMW
IBERIAN MODELLING WEEK
JULY 17-21 2023 , PALMA DE MALLORCA

[math-in] PT MATHS IN UTB Universitat de les Illes Balears IAC3

Cofinanciado por la Unión Europea

G CONSELLERIA
O FONS EUROPEUS,
I UNIVERSITAT I CULTURA
B DIRECCIÓ GENERAL
POLÍTICA UNIVERSITÀRIA
I RECERCA



Modelling Weeks

Problems

5 problems are presented, 2 of which are proposed by the PT-maths -in (Portuguese network) and the other 3 by the Spanish math-in.

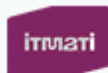
1. Route Optimization for a set of vendors.
2. Group testing of large populations.
3. Diversity in Cluster Analysis.
4. Development of an algorithm for automatic segmentation and guidance of cryoablation intervention for desmoid tumors.
5. Towards more efficient Lithium metal batteries.

Study Groups

163 ESGI

EUROPEAN STUDY GROUP WITH INDUSTRY

JULY 12-16, 2021, ONLINE, Santiago de Compostela



www.mathin.net/163esgi



Study Groups

Problema 2: "Análisis de medidas de temperatura para cocinado inteligente"

B/S/H/

BSH Electrodomésticos España, S.A.

Coordinadores Académicos | Samuele Rubino, Profesor Contratado Doctor Interino. Departamento de Ecuaciones Diferenciales y Análisis Numérico, Universidad de Sevilla, e Isabel Serra Mochales, Doctora en Matemáticas, Grupo de Sistemas Complejos del Centre de Recerca Matemàtica (CRM) y grupo CAOS del BSC.

Coordinador Empresarial | Emilio Placer Maruri head of global gas pre-development (research lines & innovation).

Especialista | Ángel Gerardo Velázquez López, Catedrático de Ingeniería Aeroespacial, Universidad Politécnica de Madrid (UPM).

Descripción | Se trata de crear soluciones sencillas que mejoren la calidad de vida de los usuarios, ya que es una obligación, sobre todo para las sociedades envejecidas. Para ello y basándose en los sensores disponibles, que recogen la información mediante una lectura continua cada 100 ms de voltaje de una serie de funciones con duración en torno a los 10 minutos, se busca analizar dicha información y tratar de evitar eventos que puedan causar problemas a los usuarios. En la mayor parte de los casos, una mirada profunda permite identificar tendencias que indiquen lo que va a suceder, en este caso, mediante una correlación entre los cambios de comportamiento de las variables estudiadas con los eventos, es decir, se conoce el segundo en el que se produce el evento y se busca hacer un análisis previo de las variables capturadas para intentar predecir en las siguientes repeticiones en qué momento hay riesgo que se produzca el evento. Desde el punto de vista matemático, habría que integrar problemas de simulación numérica con el análisis de datos.

Our aims

- To provide solutions and to transfer mathematical technology to industry and organizations.
- To introduce innovations and improvements through the application of today's latest and most demanded mathematical technologies.

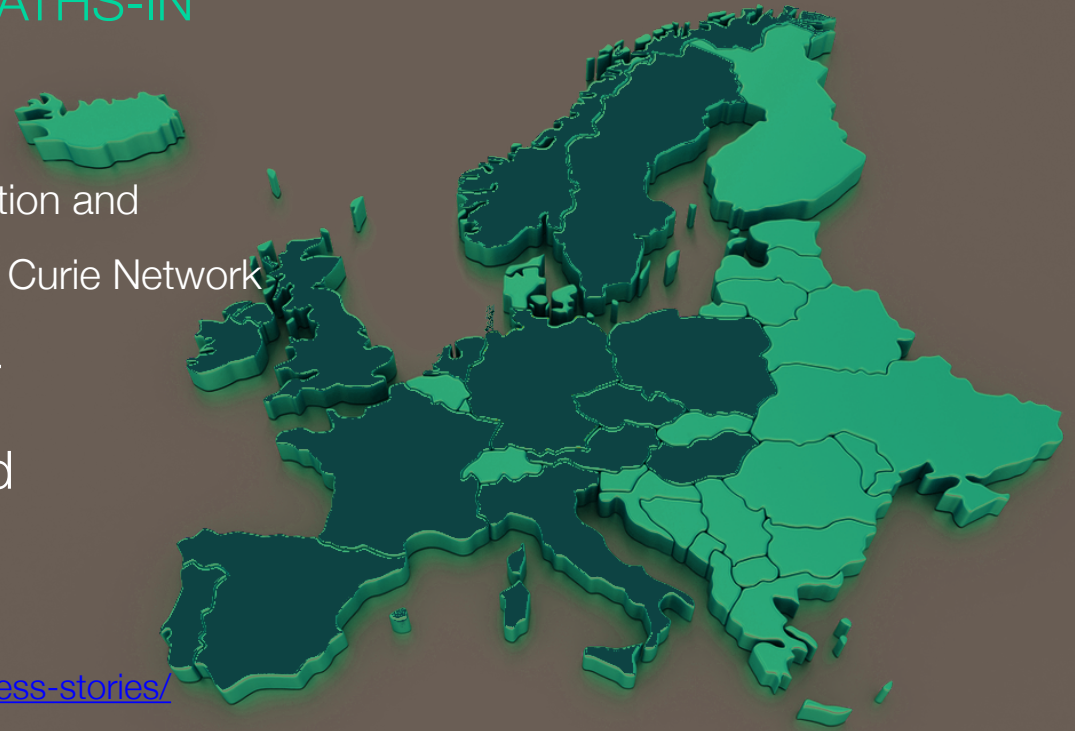


The screenshot shows the homepage of the math-in.net website. At the top, there is a navigation bar with the logo, language options (Español, English), a login field, and a search icon. Below the navigation bar is a main menu with categories: MATH-IN, OFERTA, SECTORES, EXPERIENCIA, INVESTIGACIÓN, ACTUALIDAD, and +INFORMACIÓN. The main content area features a large banner with the text "Las principales empresas y administraciones son nuestros clientes" and a sub-header "Nuestra experiencia nos avala 60 grupos y más de 600 investigadores para solucionar su problema". Below the banner are three featured sections: "MATH-IN Red Española Matemática-Industria" with a map of Spain, "SECTORES Farmacia y biomedicina" with a microscope image, and "CONFÍAN EN NOSOTROS VICUSdt Desarrollos tecnológicos" with the VICUSdt logo. To the right of these sections is a "Siguenos en Twitter" button. Below the featured sections are two columns: "NOTICIAS" (News) and "AGENDA" (Calendar). The "NOTICIAS" column lists several news items with dates and titles, such as "02/05 ESGI 131" and "01/05 Publicada la Memoria de Actividades 2016". The "AGENDA" column lists upcoming events with dates and titles, such as "15 May ESGI 131" and "15 May Semana de la Innovación Open EUREKA". On the right side of the page, there are several promotional banners for events, including "ESGI {131}", "IV IBERIAN MODELLING WEEK VIGO 12-16 JUNIO 2017", and "1ª Instituto verano internacional sobre 'Network Physiology'".

math-in, Spanish node of EU-MATHS-IN



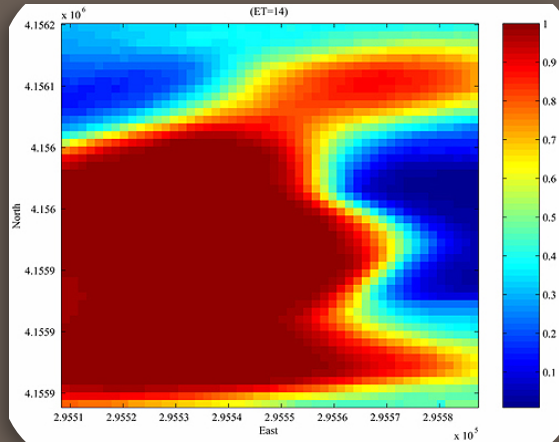
- European project promotion from EU-MATHS-IN
 - Cloud HPC
 - ROMSOC (Reduced Order Modelling, Simulation and Optimization of Coupled Systems): EID Marie Curie Network
- European Technology Platform (ETP) for Mathematical Modelling, Simulation, and Optimisation (MSO): ETP4MSO
- Success Stories: <https://www.eu-maths-in.eu/success-stories/>



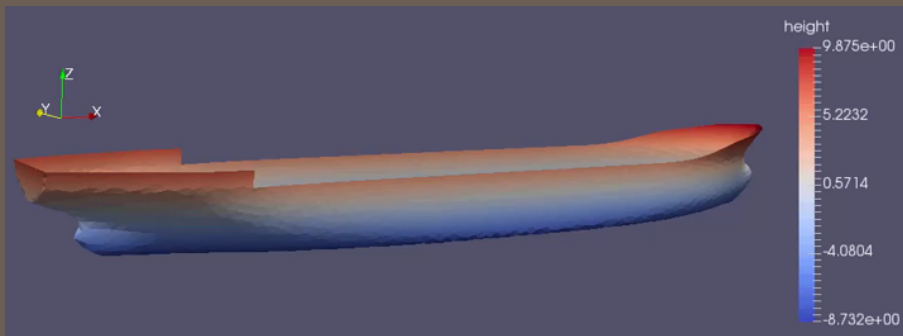
- ABB CORDOBA
- ABENGOA SOLAR
- ACSUG
- ADDLINK RESEARCH
- ADIF
- ADVANCED DYNAMICS
- AIMEN
- AFI
- ALCOA – INESPAL
- AMADIX
- BALIÑO
- BRICOKING
- BSH Electrodomésticos
- CAIXA CATALUNYA
- CAIXA GALICIA
- CASTROSÚA
- CNES - Centre Nacional Etudes Spatiales
- CESGA
- CIS MADEIRA
- CITRÖEN HISPANIA
- CONSELLO SOCIAL. USC.
- CSIC
- DALPHIMETAL
- DASSAULT AVIATION
- DETEGASA
- ENDESA
- EUROPEAN RAILWAYS
- EUSKOTREN
- EUROPIZARRAS
- FERROATLÁNTICA
- FERROCARRILES DE VÍA ESTRECHA (FEVE)
- FINSA
- FUJITSU
- GAMELSA
- GRAN CASINO DE BARCELONA
- GRANITOS MONTE FARO
- IBERDROLA
- IDOM
- IKERLAN
- INDITEX
- INGETEAM
- INDRA
- INDUSTRIAS GONZÁLEZ
- INSTITUT ESTADÍSTICA DE CATALUNYA
- INSTITÜT FÜR UNDWIRSTCHAFT MATHEMATIK
- INSTITUTO GALEGO DE ESTATÍSTICA
- INSTITUTO NACIONAL DE ESTATÍSTICA
- JUNTA DE ANDALUCÍA
- JUNTA DE CASTILLA Y LEÓN
- LABORATORIOS ISDIN
- LIGNITOS DE MEIRAMA
- MAPAL
- METRO BILBAO
- MICRONICS
- NAVANTIA
- NORCONTROL
- NOVOTEC CONSULTORES
- PERAMA INGENIERÍA
- PROCEDIMIENTOS UNO
- RENFE
- REPSOL EXPLORACIÓN
- SGL CARBON
- SIDENOR
- TECNOLOGÍAS AVANZADAS INSPIRALIA (ITAV)
- TOTEMA ENGINEERING
- TURGALICIA
- TRANSPORTES URBANOS DE SEVILLA
- TSK FLAGSOL
- UDX
- UFISA SOLUZIONA
- UNILEVER FOOD ESPAÑA
- UNIÓN ESPAÑOLA DE EXPLOSIVOS
- UNION – FENOSA GENERACIÓN
- SCHNEIDER ELECTRIC
- XUNTA DE GALICIA



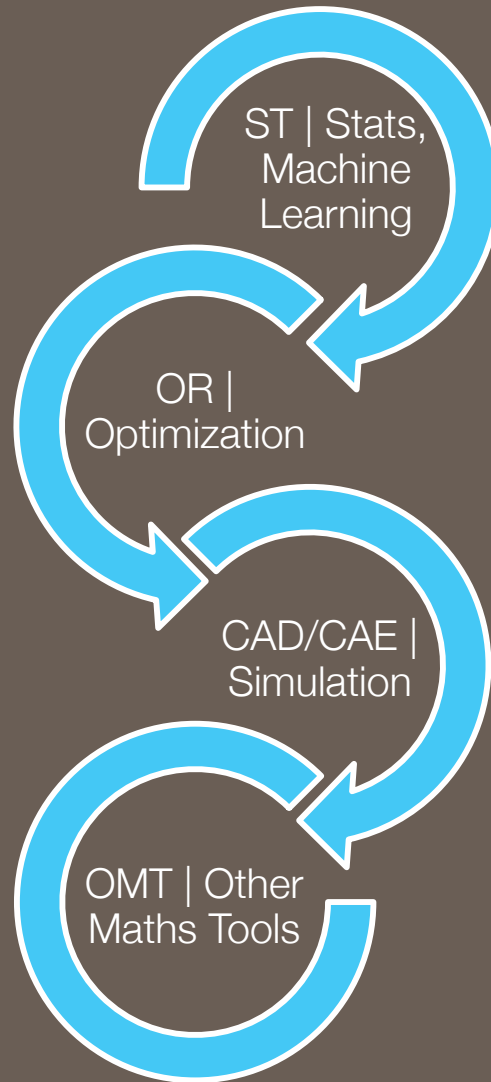
Mathematical technologies



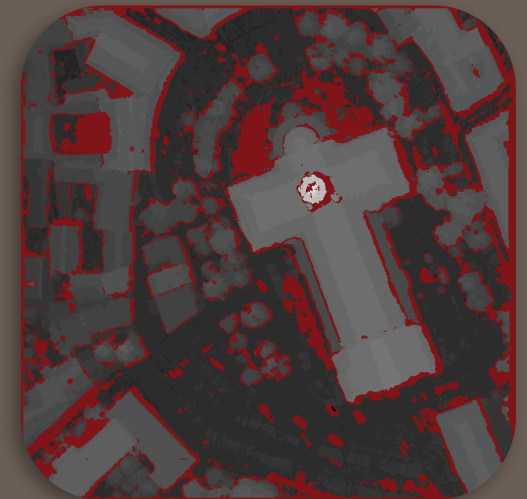
Maps risk of weed infestation
MODES | ITMATI, UDC



Dynamic draft study with constant environmental variables.
mat+i, M2NICA | ITMATI, USC, UDC
i4sea.



Modelling and installation of geothermal systems
InterTech | UPV



Resolve the depth of satellite images
TAMI | UIB

Mathematical technologies



REAL ACADEMIA ESPAÑOLA



Diccionario de la lengua española Edición del Tricentenario Actualización 2020

Consulta posible gracias al compromiso con la cultura de la  Fundación "la Caixa"

por palabras Consultar

algoritmo

Quizá del lat. tardío **algobarismus*, y este abrev. del ár. clás. *ḥisābu lġubār* 'cálculo mediante cifras arábigas'.

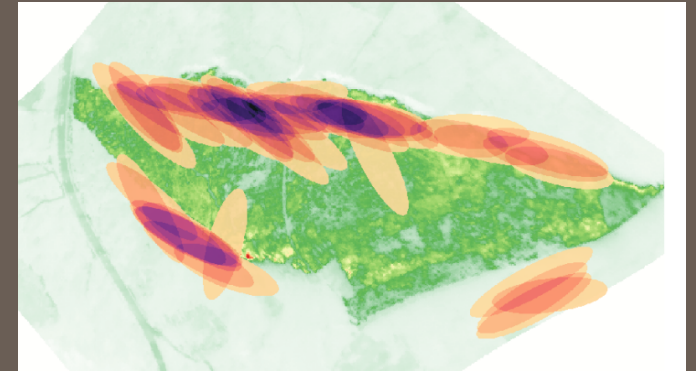
1. **m.** Conjunto ordenado y finito de operaciones que permite hallar la solución de un problema.
2. **m.** Método y notación en las distintas formas del cálculo.

Mathematical technologies

$$\begin{aligned}
 r : \rho \left(\frac{\partial u_r}{\partial t} + u_r \frac{\partial u_r}{\partial r} + \frac{u_\phi}{r} \frac{\partial u_r}{\partial \phi} + u_z \frac{\partial u_r}{\partial z} - \frac{u_\phi^2}{r} \right) = \\
 - \frac{\partial p}{\partial r} + \mu \left[\frac{1}{r} \frac{\partial}{\partial r} \left(r \frac{\partial u_r}{\partial r} \right) + \frac{1}{r^2} \frac{\partial^2 u_r}{\partial \phi^2} + \frac{\partial^2 u_r}{\partial z^2} - \frac{u_r}{r^2} - \frac{2}{r^2} \frac{\partial u_\phi}{\partial \phi} \right] + \rho g_r \\
 \phi : \rho \left(\frac{\partial u_\phi}{\partial t} + u_r \frac{\partial u_\phi}{\partial r} + \frac{u_\phi}{r} \frac{\partial u_\phi}{\partial \phi} + u_z \frac{\partial u_\phi}{\partial z} + \frac{u_r u_\phi}{r} \right) = \\
 - \frac{1}{r} \frac{\partial p}{\partial \phi} + \mu \left[\frac{1}{r} \frac{\partial}{\partial r} \left(r \frac{\partial u_\phi}{\partial r} \right) + \frac{1}{r^2} \frac{\partial^2 u_\phi}{\partial \phi^2} + \frac{\partial^2 u_\phi}{\partial z^2} + \frac{2}{r^2} \frac{\partial u_r}{\partial \phi} - \frac{u_\phi}{r^2} \right] + \rho g_\phi \\
 z : \rho \left(\frac{\partial u_z}{\partial t} + u_r \frac{\partial u_z}{\partial r} + \frac{u_\phi}{r} \frac{\partial u_z}{\partial \phi} + u_z \frac{\partial u_z}{\partial z} \right) = \\
 - \frac{\partial p}{\partial z} + \mu \left[\frac{1}{r} \frac{\partial}{\partial r} \left(r \frac{\partial u_z}{\partial r} \right) + \frac{1}{r^2} \frac{\partial^2 u_z}{\partial \phi^2} + \frac{\partial^2 u_z}{\partial z^2} \right] + \rho g_z.
 \end{aligned}$$

16084244485963766983895
 22868478312355265821314
 49576857262433441893039
 68642624341077322697802
 80731891544110104468232
 52716201052652272111660
 39666557309254711055785
 37634668206531098965269
 18620564769312570586356
 62018558100729360659876
 48611791045334885034611
 36576867532494416680396
 26579787718556084552965
 41266540853061434443185
 86769751456614068007002
 37877659134401712749470
 42056223053899456131407
 11270004078547332699390
 81454664645880797270826
 68306343285878569830523
 58089330657574067954571
 63775254202114955761581
 40025012622859413021647
 15509792592309907965473
 76125517656751357517829
 66645477017450112006148

- Classification.
- Forecasting.
- Quality control.
- Risk analysis. Credit scoring.
- Customer analysis. Market/product studies.
- Experimental design.
- Clinical analyses. Biostatistics



Watermark of the discharges made by five aircraft that participated in the extinction of a fire that occurred in Galicia in the summer of 2017.
MODES | ITMATI, USC, Coremain



Information Systems for Health Planning
US
CARTO

- Off the shelf Machine Learning tools, e.g. (deep) neural networks

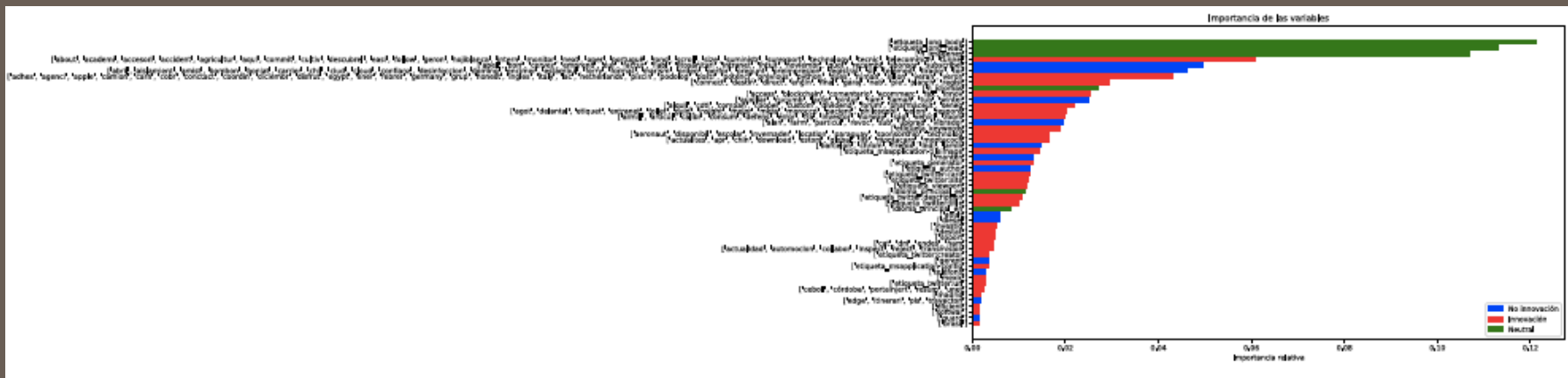


CNN
with trained parameters

INFERENCE



- From back to gray boxes. The search of explainability.



Detecting innovative firms through their web pages

US

IECA

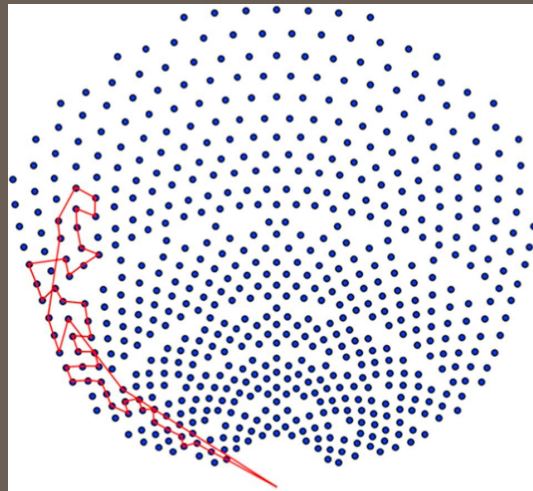
Mathematical technologies

Operations Research (OR)

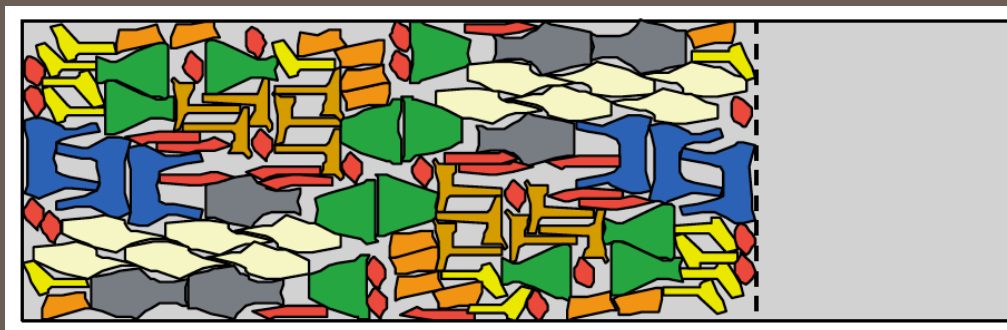
- Supply Chain Management.
- Production.
- Routing.
- Strategic decision-making.



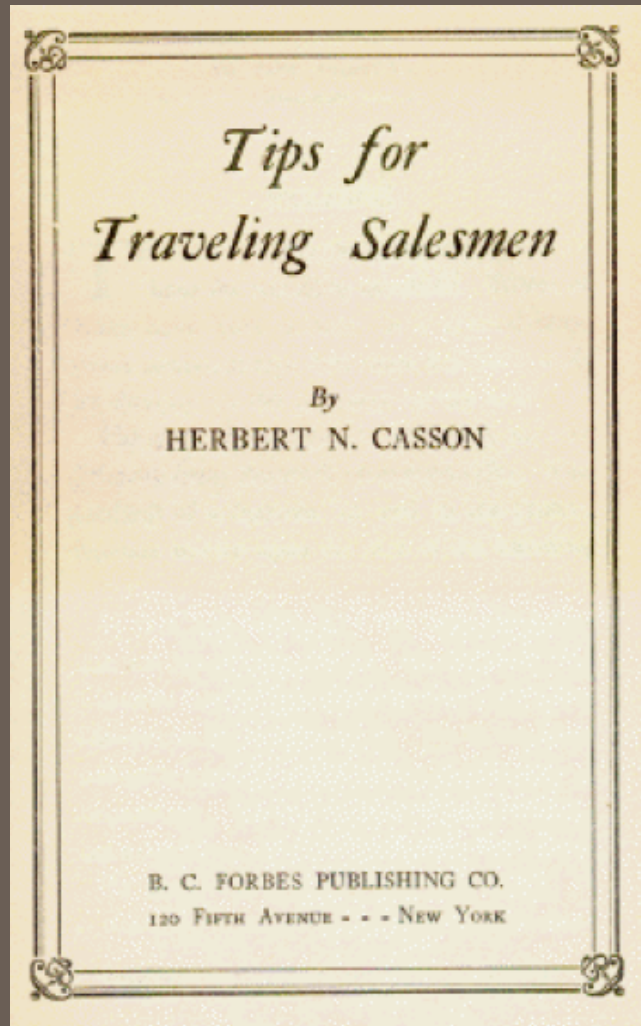
Optimal vectorization of airspace.
US
CRIDA, ENAIRE



Solar Tower Plants Design and Maintenance.
US
TSK, Abengoa Solar



Efficient implementation of 2D nesting problem to minimize waste of metal.
BCAM
Lantek



*Tips for
Traveling Salesmen*

By
HERBERT N. CASSON

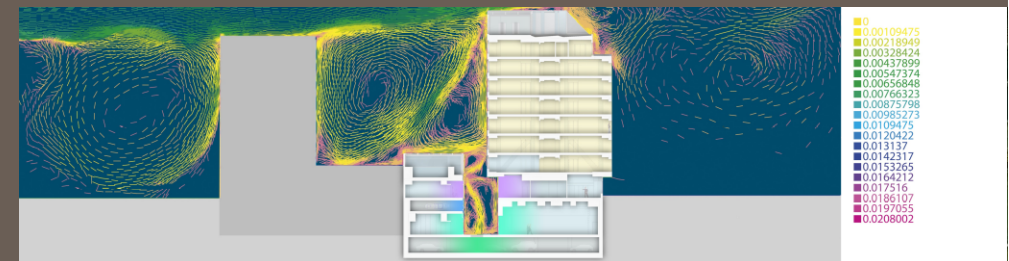
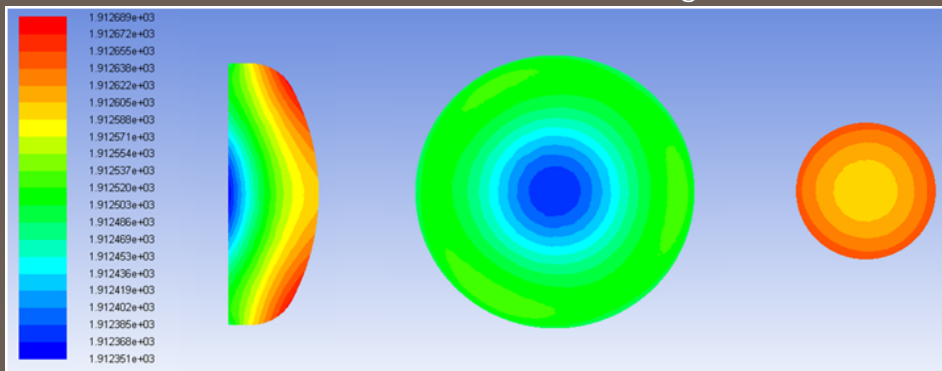
B. C. FORBES PUBLISHING CO.
120 FIFTH AVENUE - - - NEW YORK

Mathematical technologies

CAD/CAE

- Mechanical or structural.
- Thermal or thermodynamic.
- Manufacturing processes.
- Electronic and/or electromagnetic.
- Fluids.
- Acoustic or vibro-acoustic.
- Environmental.
- Other, such as multi-physics.

Efficient Metal Purification Strategy
mat+i | ITMATI, USC
Ferroglobe

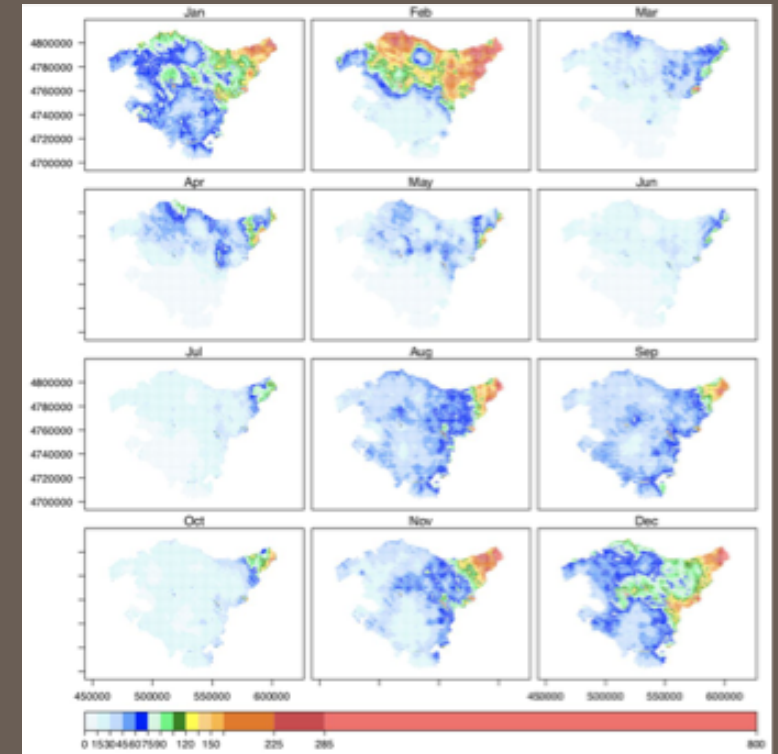


Intelligent Buildings
US
ILUNION Hotels

Mathematical technologies

Other maths techniques (OMT)

- Digital images.
- Geometric analysis.
- Digital signal processing. Search and processing of information
- Design of geographical tracking systems.
- Communication networks.
- Codification of information.
- Computation, computational algebra.
- Bio-informatics, Bio-mathematics, genomics and proteomics.



Mapping high-resolution soil properties with geoaddivitive models. Predicted R-factor Map (left), R-factor by month (above). BCAM, Neiker

Take-home messages

- Maths is even more everywhere than ever.
- Modern paradigms (data driven decision making) call for more intensive (and new) maths to perform **accurate and safe**.

Take-home messages

- Maths is even more everywhere than ever.
- Modern paradigms (data driven decision making) call for more intensive (and new) maths to perform **accurate and safe**.
- Applied mathematicians are **ready to expand collaborations** with industry and organizations (in Spain and abroad!).

Take-home messages

- Maths is even more everywhere than ever.
- Modern paradigms (data driven decision making) call for more intensive (and new) maths to perform **accurate and safe**.
- Applied mathematicians are **ready to expand collaborations** with industry and organizations (in Spain and abroad!).
- The network math-in is a reliable and efficient bridge between industry and mathematicians in universities.

Take-home messages

- Maths is even more everywhere than ever.
- Modern paradigms (data driven decision making) call for more intensive (and new) maths to perform **accurate and safe**.
- Applied mathematicians are **ready to expand collaborations** with industry and organizations (in Spain and abroad!).
- The network math-in is a reliable and efficient bridge between industry and mathematicians in universities.

Big thanks!!!

emilio.carrizosa@math-in.net

<http://math-in.net/>