azti tecnalia
Transforming Science into Business

Universidad del País Vasco / Euskal Herriko Unibertsitatea

Intelligent Systems Group

PML Plymouth Marine Laboratory

www.azti.es

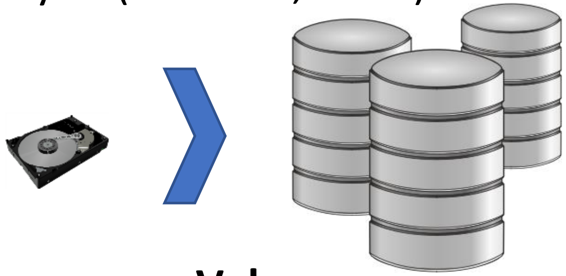
Big Data in Marine Science: A few modelling examples

Jose A. Fernandes
jfs@pml.ac.uk
Senior Scientist in Big Data for Marine Research and Innovation (AZTI)
PML fellow

Ibon Galparsoro (AZTI)
Iñaki Quincoces (AZTI)
Iñaki Inza (ISG, UPV)
Xabier Irigoien (AZTI)

What is the meaning of Big Data? The 3 Vs, the 5 Vs, the 7Vs

In addition to its **sheer volume**, big data also exhibits other unique characteristics as compared with traditional data. It is commonly **unstructured** and require more **real-time** analysis (Hu et al., 2014).



Volume

Variety

Velocity

Variability: E.g. temporal and spatial

Veracity: data with errors

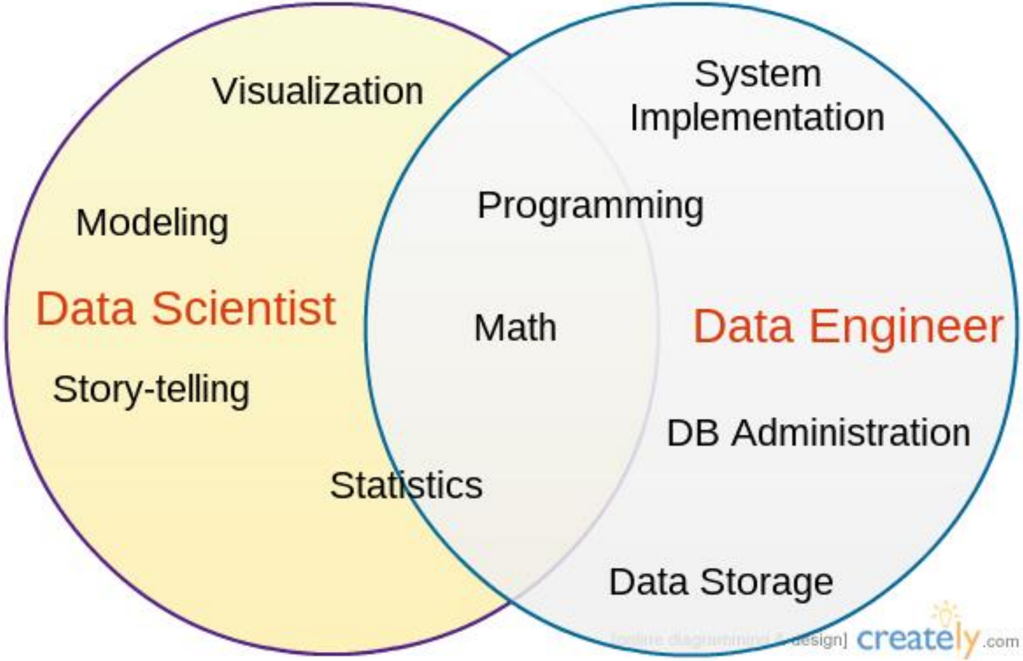
Viability: quickly and cost-effectively test

Value: More for less -> blue economy

Hu, H., Wen, Y., Chua, T. S., & Li, X. (2014). Toward scalable systems for big data analytics: A technology tutorial. IEEE Access, 2, 652-687.

Main roles in the Big Data process

Value:
Other scientists
Industry
Managers
General public



http://101.datascience.community/2014/07/08/data-scientist-vs-data-engineer/

Avoiding the **DARK side**
of the **Big Data** requires
the right combination of **expertise** and **responsible** application.

Image analysis, machine learning and field expert

Details

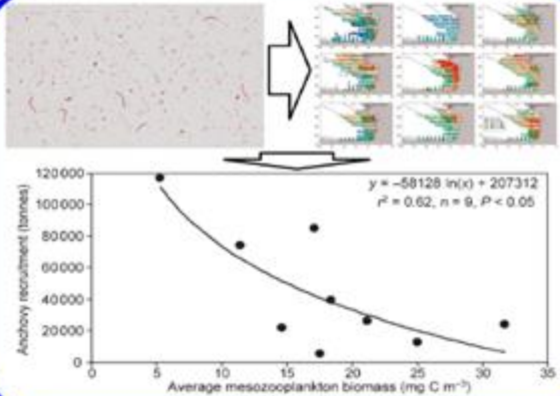
9 years of data
4000 samples
2,000,000 particles

From 24 to 67 taxa
Break 0.5 mm limit

Methods advance

Training-set elaboration

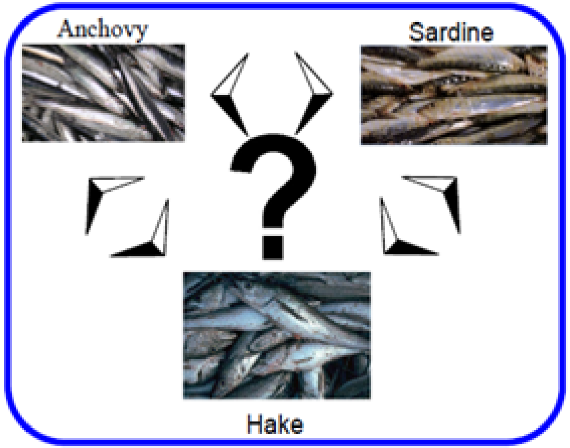
Application



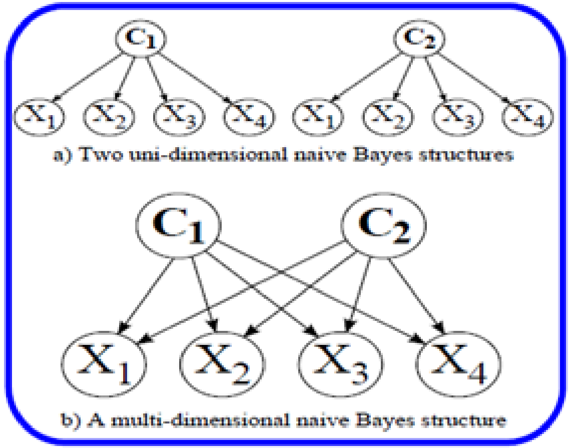
(2012) Improving semi-automated zooplankton classification using an internal control and different imaging devices. Bachiller E., Fernandes J.A., Irigoien X. Limnol. Oceanogr.: Methods 10: 1-9.
(2011) Zooplankton Image Analysis Manual: automated identification by means of scanner and digital camera as imaging devices. Bachiller E. & Fernandes J.A.. Revista Investigación Marina 18(2) :16-37.
(2009) Spring zooplankton distribution in the Bay of Biscay from 1998 to 2006 in relation with anchovy recruitment.Irigoien X., Fernandes J.A., Grosjean P, Denis K., Albaina A, Santos M. J. Plankton Res. 31(1): 1-17.
(2009) Optimizing the number of classes in automated zooplankton classification. Fernandes J.A., Irigoien X., Boyra G., Lozano J.A., Inza I. J. Plankton Res. 31(1): 19-29.
(2009) Changes in plankton size structure and composition, during the generation of a phytoplankton bloom, in the central Cantabrian sea. Zarauz L, Irigoien X., Fernandes J.A. J. Plankton Res. 31(2): 193-207. [Journal link] [ResearchGate]
(2008) Modelling the influence of abiotic and biotic factors on plankton distribution in the Bay of Biscay, during three consecutive years (2004-06). Zarauz L., Irigoien X., Fernandes J.A. J. Plankton Res. 30(8): 857-872.

Fish recruitment forecasting

Details



Methods advance

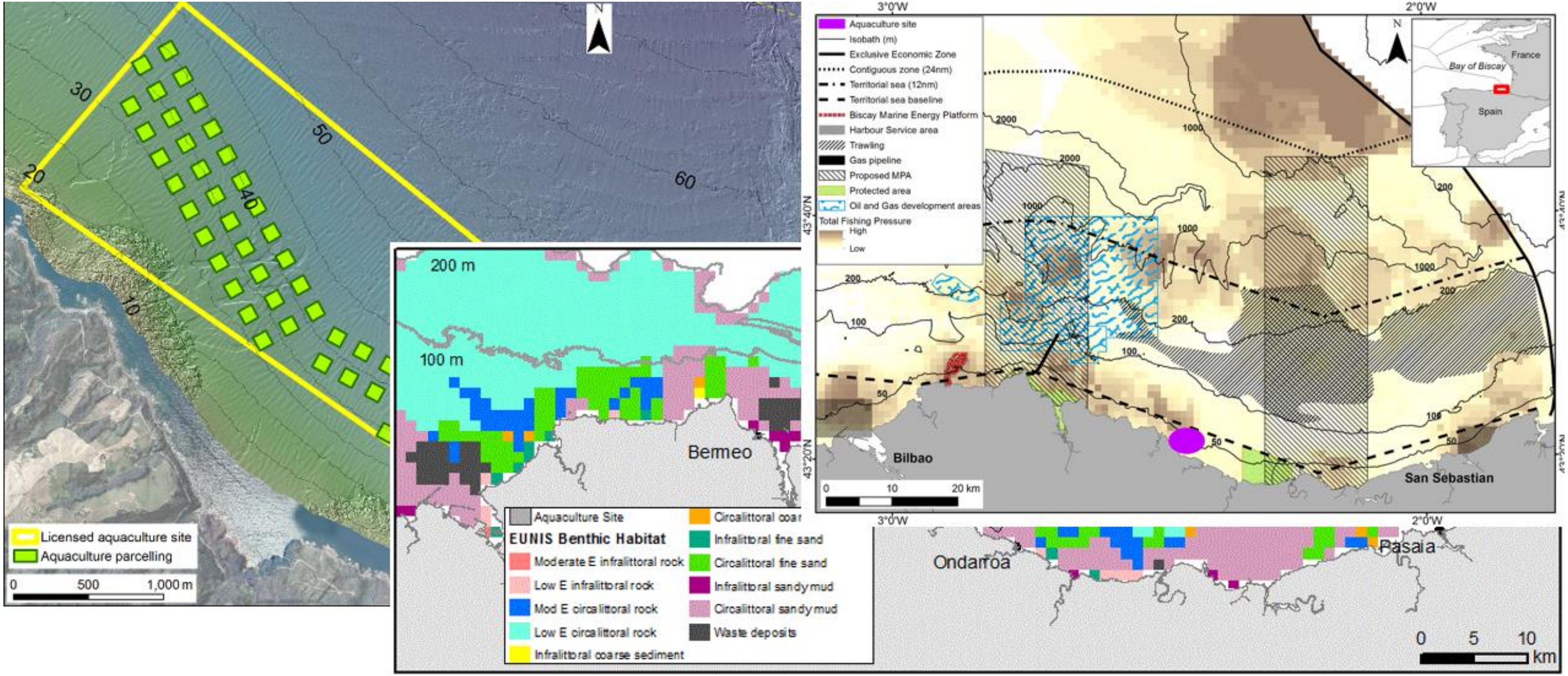


Application

Anchovy BS	Sardine BS	Hake BS	Joint Acc.
0.36	0.34	0.27	17.3 ± 4.8
0.35	0.27	0.21	28.9 ± 4.5
0.32	**0.24**	0.19	22.6 ± 4.3
0.32	0.25	**0.18**	19.7 ± 5.5
0.30	0.27	0.21	$\mathbf{29.5 \pm 4}$
0.32	0.27	**0.18**	28.5 ± 4.7

(2015) Evaluating machine-learning techniques for recruitment forecasting of seven North East Atlantic fish species. Fernandes J.A., Irigoien X, Lozano J.A., Inza I., Goikoetxea N., Pérez A. Ecol. Inform. 25, 35-42
(2015) Spatio-Temporal Bayesian Network Models with Latent Variables for Revealing Trophic Dynamics and Functional Networks in Fisheries Ecology. Trifonova N., Kenny A., Maxwell D., Duplisea D., Fernandes J.A., Tucker A. Ecol. Inf. 30: 142-158.
(2013) Supervised pre-processing approaches in multiple class-variables classification for fish recruitment forecasting. Fernandes J.A., Lozano J.A., Inza I., Irigoien X., Rodríguez J.D., Pérez A. Environ. Modell. Softw. 40, 245-254.
(2011) The potential use of a Gadget model to predict stock responses to climate change in combination with Bayesian Networks: the case of the Bay of Biscay anchovy. Andonegi E., Fernandes J.A., Quincoces I., Uriarte A., Pérez A., Howell D., Stefansson G. ICES J. Mar. Sci. 68(6): 1257-1269.
(2010) Fish recruitment prediction, using robust supervised classification methods. Fernandes J.A., Irigoien X., Goikoetxea N., Lozano J.A., Inza I., Pérez A, Bode A. Ecol. Model. 221(2): 338-352.
(2009) Anchovy Recruitment Mixed Long Series prediction using supervised classification. Fernandes J.A., Irigoien X., Uriarte A., Ibaibarriaga L., Lozano J.A., Inza I. Working document to the ICES benchmark workshop on short lived species (WKSHORT) Bergen (Norway), 31 August-4 September.

A local aquaculture spatial planning example



Science hot topic:
Hake as a choke-species (mackerel and horse mackerel before)

FISH and FISHERIES

FISH and FISHERIES, 2015, 16, 563–575

Adverse consequences of stock recovery: European hake, a new “choke” species under a discard ban?

Alan R Baudron & Paul G Fernandes

Zoology Department, Institute of Biological and Environmental Sciences, University of Aberdeen, Aberdeen, AB24 2TZ, UK

Becoming a policy hot topic:
ICES workshop in answer to DGmare request

ICES WKFISHDISH REPORT 2016

ICES ADVISORY COMMITTEE

ICES CM 2016/ ACOM: 55

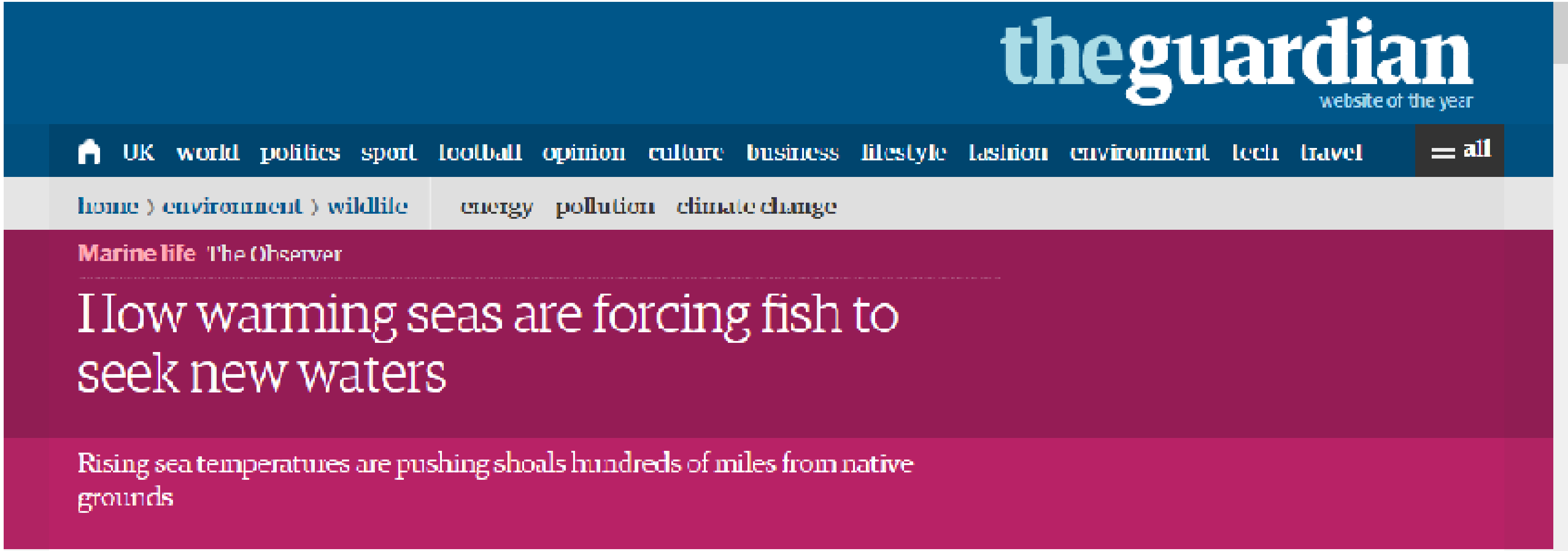
REF. ACOM

Report of the Working Group on Fish Distribution Shifts (WKFISHDISH)

22–25 November 2016

- 10 key species identified as big movers out of 19 widely distributed species: anchovy, anglerfish, blue whiting, cod, hake, herring, mackerel, horse mackerel, megrims, and plaice.
- Key areas identified: The North Sea; Bristol and English Channel; West of Scottland and Ireland; and, North Scotland, Norwegian Sea and Skagerrak-Kattegat.

Public hot topic: reaching the general press

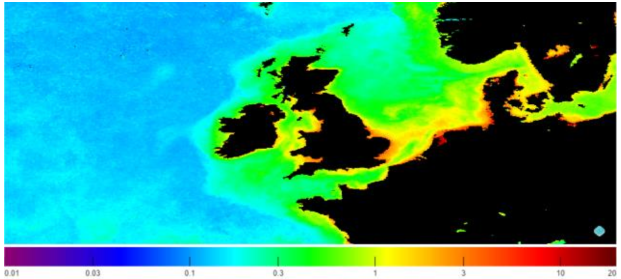


“Then stocks began to head north, most probably because sea temperatures were rising. Eventually, mackerel reached Iceland – at which point Iceland asked to be included in fishing quotas. This request was rejected – so Iceland went ahead and started catching mackerel in any case.”

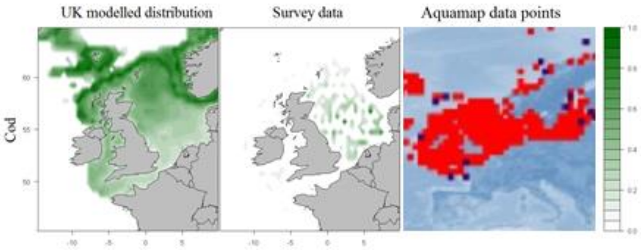
“Unless we find ways to adapt quota agreements speedily and efficiently, we are going to see a lot more disputes like this one in future,” Roberts said.

Bringing data and technologies together

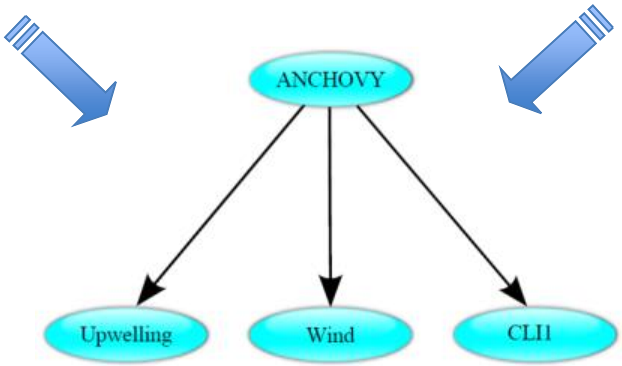
Satellite data



Survey data



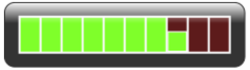
Probabilistic forecasting



Results of the proof-of-concept

Cod / Hake in the North Sea

Presence / absence performance


Cod



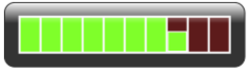
Hake



Both simultaneously

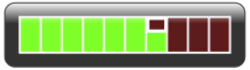
Cod / Haddock in the South UK

Presence / absence performance


Cod



Haddock



Both simultaneously

Hake: high biomass accuracy 88%
Cod: high biomass accuracy 74%
Haddock high biomass accuracy 65%

Pelagic species

- Presence / absence
  - Sardine: 97 %
  - Mackerel: 77 %
  - Herring: 71.2 %
  - Sprat: 79 %

- High biomass:
  - Sardine: 73 %
  - Mackerel: 74 %
  - Herring: 73 %
  - Sprat: 61 %

Workforward (Big Data)

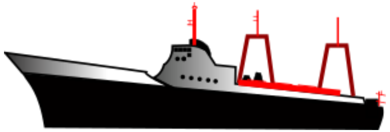
Socio-economics?

Making it operational and funding its use?

Funding further research & related topics?

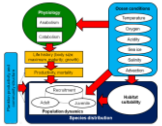
- Industry efficiency and sustainability
- Avoiding bycatch species
- Ecosystem impact evaluation
- Assessment of cumulative pressures
- Support to stock assessments and management
- Mapping and Assessment of Marine Ecosystem Services
- Maritime Spatial Planning

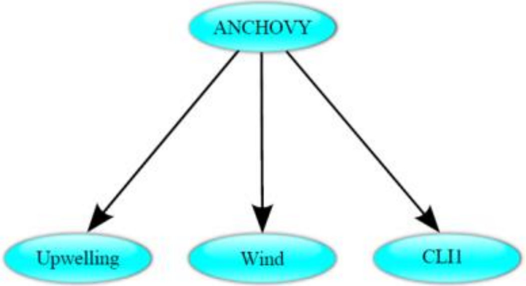
Fishing vessels capturing scientific data for improving management and help industry



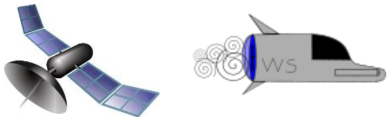
Uranga, J., et al. (2017). Detecting the presence-absence of bluefin tuna by automated analysis of medium-range sonars on fishing vessels. *PloS one*, *12*(2), e0171382.

Other sources of knowledge





Other sources of data



Big Data in Marine Science: A few modelling examples

Jose A. Fernandes
jfernandes@azti.es
Senior Scientist in Big Data for Marine Research and Innovation (AZTI)
PML fellow (PML)

azti tecnalia

Transforming Science into Business

Txatxarramendi ugarte z/g
48395 Sukarrieta, Bizkaia

Herrera Kaia. Portualdea z/g
20110 Pasaia, Gipuzkoa

Astondo Bidea, Edificio 609
Parque Tecnológico de Bizkaia
48160 Derio, Bizkaia