

# Development, deployment and validation of an oceanographic virtual laboratory based on Grid computing

David Mera Pérez

Santiago de Compostela, Feb. 15<sup>th</sup> 2013



# Index

- 1 Context and Motivation
- 2 Objectives
- 3 Virtual laboratory development
- 4 Virtual laboratory validation
- 5 Results and conclusions



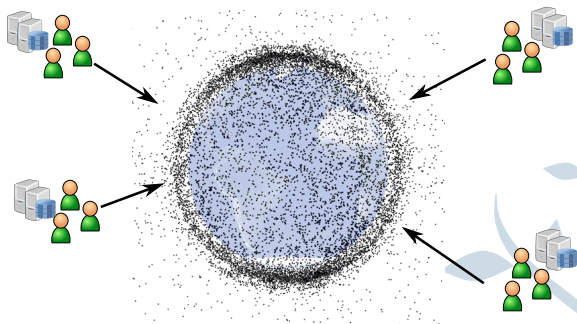
# Index

- 1 Context and Motivation
- 2 Objectives
- 3 Virtual laboratory development
- 4 Virtual laboratory validation
- 5 Results and conclusions



# Context and Motivation

- ▷ Satellite missions for Earth observation increase every year.
- ▷ The study of the ocean requires multidisciplinary teams
- ▷ Distributed computing paradigm.



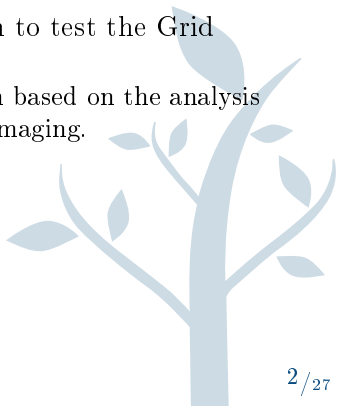
# Index

- 1 Context and Motivation
- 2 Objectives**
- 3 Virtual laboratory development
- 4 Virtual laboratory validation
- 5 Results and conclusions



# Objectives

1. To develop a user-friendly distributed computational environment based on Grid computing.
2. To develop an oceanographic application to test the Grid environment.
  - An oil spill automatic detection system based on the analysis of satellite Synthetic Aperture Radar imaging.



# Index

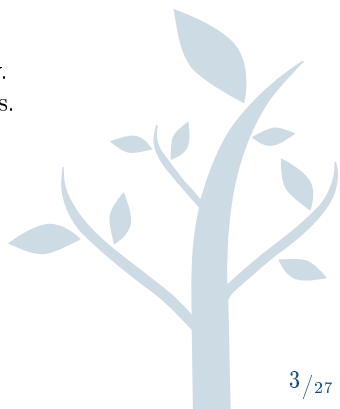
- 1 Context and Motivation
- 2 Objectives
- 3 Virtual laboratory development**
- 4 Virtual laboratory validation
- 5 Results and conclusions



# Retelab

## User access and registration

- ▷ The access to most of the Grids is not intuitive.
  - Command line interface.
  - Digital certificates.
  - The computer knowledge is mandatory.
  - The users management is based on files.

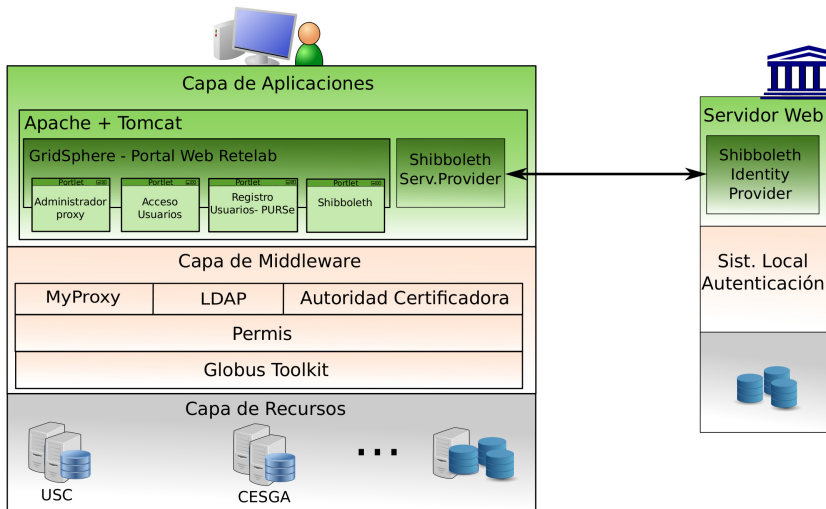




# Retelab

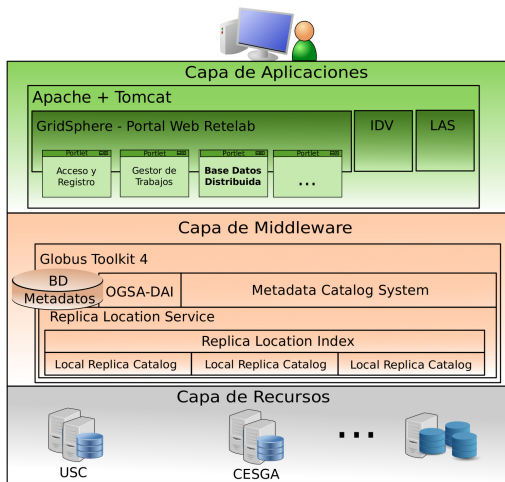
## User access and registration

### ▷ Retelab approach



# Retelab Distributed storage system

- ▷ Based on Metadata - ISO 19115.
- ▷ Integration of visualization tools.
  - Live Access Server.
  - Integrated Data Viewer.



# Retelab

## Job submission and monitoring system

- ▷ Previous job submission systems:
  - They need interaction with the users.
  - The interaction decreases simplicity and transparency.
- ▷ Retelab approach:
  - Grid metascheduler.
    - > To make decisions on behalf of users.
    - > To facilitate the optimal utilization of the Grid resources.
    - > It undertakes the tasks for resource discovery, job scheduling, executing, monitoring and output retrieval.
  - It was mainly developed by a CESSGA researcher.

# Retelab Integration

The image displays two screenshots of the Retelab web interface, showing the configuration and execution of a virtual experiment.

**Top Screenshot: Configuration**

- Page Title:** Produccion primaria
- Section:** IDL - PRIMARY PRODUCTION
- Form Fields:**
  - Title: \_\_\_\_\_
  - Main: \_\_\_\_\_
  - Input files: \_\_\_\_\_
  - Result files: \_\_\_\_\_
  - Remote: \_\_\_\_\_
  - Local: \_\_\_\_\_
- Virtual Data Base:**
  - Region: Canary
  - Sensor: \_\_\_\_\_
  - Parameter: \_\_\_\_\_
  - Search: \_\_\_\_\_
- File Browser:**
  - Path: /home/retelab/david\_mera
  - Select input files
  - loadct.pro
  - pp\_out\_nc.pro
  - aware.pro
  - fotoeer.pro
- Button:** Enviar

**Bottom Screenshot: Execution Results**

- Page Title:** Produccion primaria
- Section:** IDL - PRIMARY PRODUCTION
- Form Fields:**
  - Title: virtual9
  - Main: aware
  - Input files:
 

```
gsiftp://ui.retelab.cesga.es/tmp/david_mera
/AV_SST_2006314_6_600.N7.newmedvirado
gsiftp://ui.retelab.cesga.es/tmp/david_mera
/SW_CHL_2006314_6_600.14.newmedvirado
gsiftp://ui.retelab.cesga.es/tmp/david_mera
/S20063132006320.L3m_8D_PAR_9_file.///home/r
```
  - Result files: \_\_\_\_\_
  - Remote: results.tar.gz
  - Local: results.tar.gz
- Virtual Data Base:**
  - AV\_SST\_2006314\_6\_600.N7.newmedvirado [+]
  - SW\_CHL\_2006314\_6\_600.14.newmedvirado [+]
  - S20063132006320.L3m\_8D\_PAR\_9\_2 [+]
  - sw\_ppo\_2006314\_6\_600.14.newmedvirado.nc [+]
  - Select
  - Show Search
- File Browser:**
  - Path: /home/retelab/david\_mera
  - Select input files
  - 1243337736781
  - 1245862964637
  - loadct.pro
- Button:** Enviar

# Retelab Integration

The screenshot displays the Retelab virtual laboratory interface. At the top, the Retelab logo is visible, along with the text "LABORATORIO VIRTUAL PARA LA RED NACIONAL DE TELEDETECCION OCEANOGRÁFICA". The main interface is divided into several sections:

- Job Monitor:** A table showing the status of a job. The job ID is 4, titled "virtual19", with a status of "Finished". The user is "testuser9" and the results are located at "/testuser9/1259935615423".
- File Monitor:** A table listing files generated by the job. The files are "stderr-4.txt" (657 bytes), "stdout-4.txt" (759 bytes), and "sw\_ppo\_2006314\_6\_600.14.newmedivradio.nc" (4771292 bytes).
- Attributes:** A section for defining attributes, showing "sw\_ppo\_2006314\_6\_600.14" with a value of "AVVHR".
- Unidata IDV - Map View - One Pane:** A map view showing a geographical area with a contour plot overlay. The map is titled "grub\_grad - Contour: Plan View".

The Job Monitor table data is as follows:

GWID	Title	Status	DN proxy	Results	Action
4	virtual19	Finished	testuser9	/testuser9/1259935615423	Delete

The File Monitor table data is as follows:

Type	Name	Size	Last Modification	Actions
File	stderr-4.txt	657	20100119124826	add DataGrid
File	stdout-4.txt	759	20100119124826	add DataGrid
File	sw_ppo_2006314_6_600.14.newmedivradio.nc	4771292	20100119124826	add DataGrid

The Attributes section shows:

Attribute:  Type:  Value:

The Unidata IDV map view shows a contour plot of a geographical area, with a legend on the right side indicating values from 1.31 to 1.32.

# Index

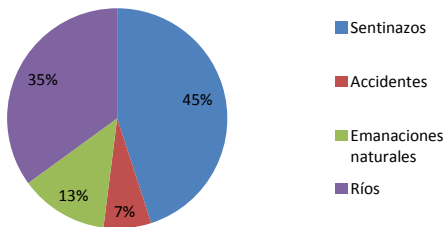
- 1 Context and Motivation
- 2 Objectives
- 3 Virtual laboratory development
- 4 Virtual laboratory validation**
- 5 Results and conclusions



# Sentinazos

## Introduction

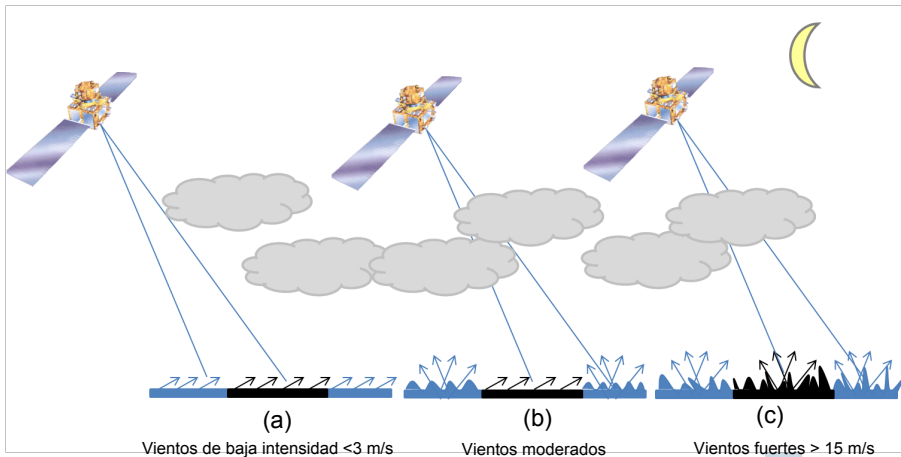
- ▷ The international trade is mainly supported by maritime transport.
- ▷ The intensive traffic sails along the Exclusive Economic Zones (EEZ) of the countries and generates important pollution problems.
- ▷ Only the 7% of oil spills come from catastrophes like tanker and oil platform accidents.



# Sentinazos

## Introduction

### ▷ Synthetic Aperture Radar

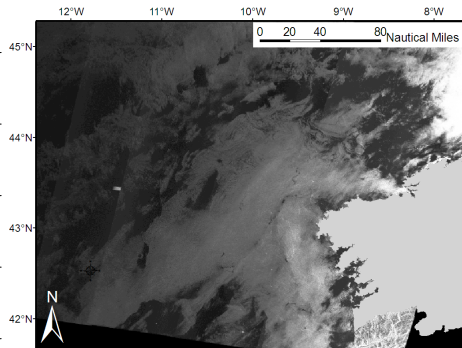
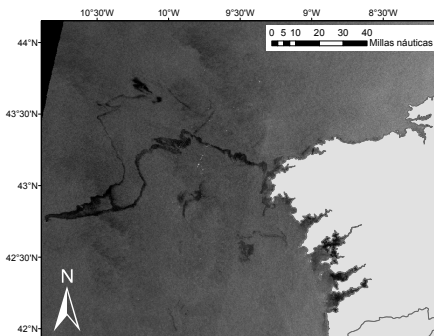




# Sentinazos

## Introduction

### ▷ Synthetic Aperture Radar - Examples



# Sentinazos

## Introduction

### ▷ Synthetic Aperture Radar - Oil Spills

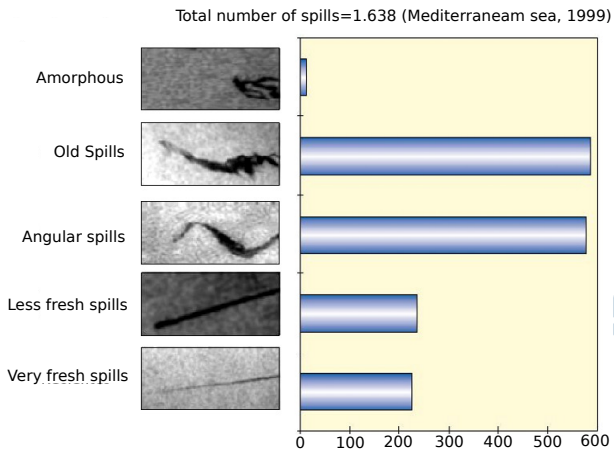


Figura: Classification of detected spills in terms of their shapes.

# Sentinazos

## Goals

### Hypothesis

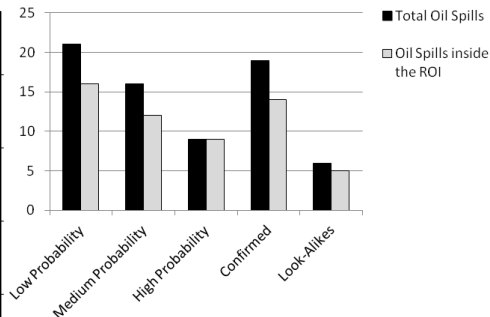
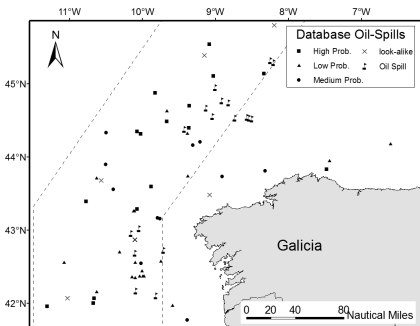
1. Is possible to use wind information to segment oil candidates from SAR images?
2. Is the shape analysis relevant to classify the oil candidates?

### Goal

- ▷ To develop an oil spill automatic detection system focused on the galician coast and based on SAR images.

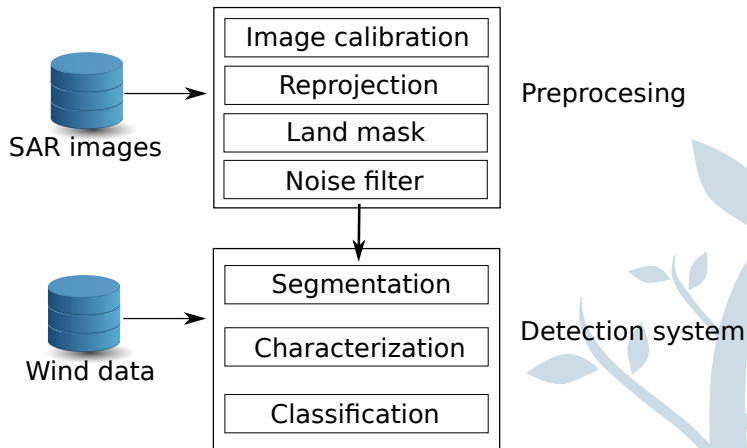
# Sentinazos Methodology

- ▷ Dataset: a collection of 47 SAR images from the Envisat.
  - Galician coast. Finisterre Traffic Separation Scheme (2007-2011)
  - Wide Swath Mode
  - Polarization: vertical-vertical
  - Coverage: 400 km x 400 km



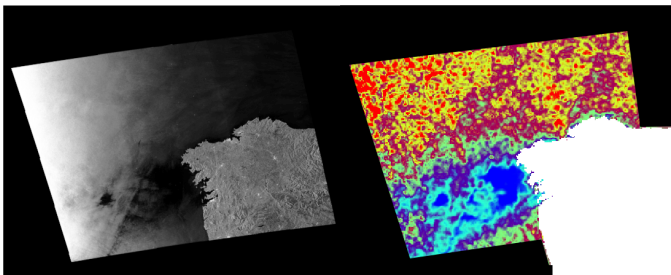
# Methodology Sentinazos

## ▷ Oil Spill detection system architecture

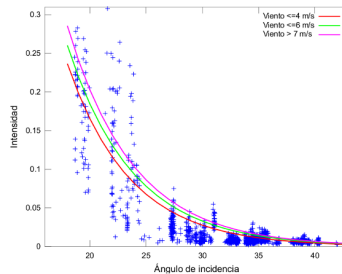


# Sentinazos Methodology

## ▷ Segmentation - Establishing the Adaptive Threshold.

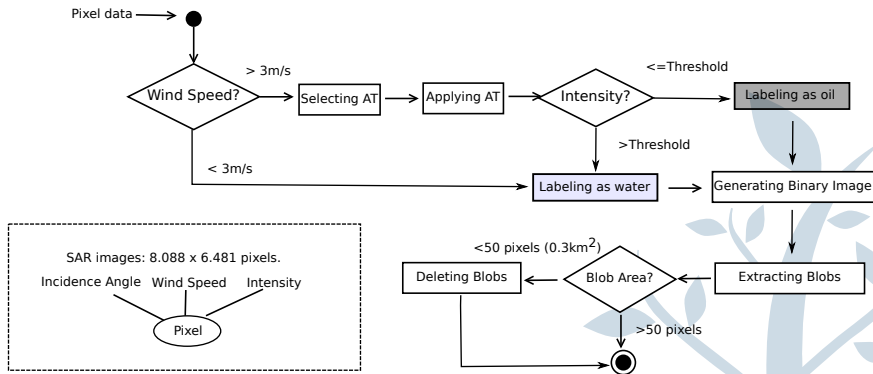


X	Y	Wind speed	Incidence Angle	Intensity
200	567	5.4	37,28°	0.021
203	577	3.2	36,46°	0.003
300	367	3.4	30,23°	0.0285
320	467	4.1	20,67°	0.0423
...	...	...	...	...



# Sentinazos Methodology

## ▷ Segmentation - Applying the Adaptive Threshold

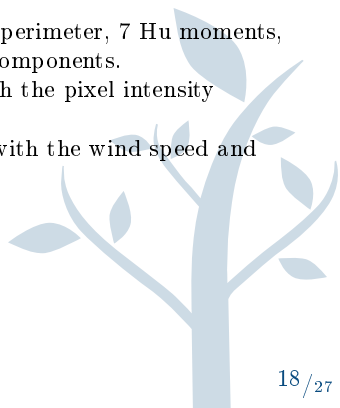


# Sentinazos

## Methodology

### ▷ Characterization

- The segmented areas are analyzed to get a characteristic vector:
  - > 17 shape characteristics (Ratio area perimeter, 7 Hu moments, Thickness, etc) → PCA → 5 main components.
  - > 2 physical characteristics related with the pixel intensity values.
  - > 2 contextual characteristics related with the wind speed and the incidence angle.





# Sentinazos

## Methodology

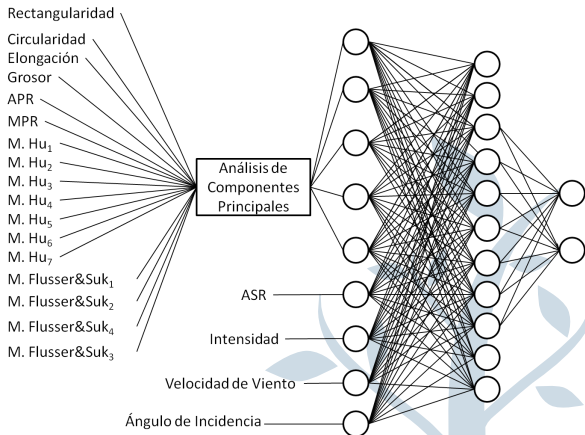
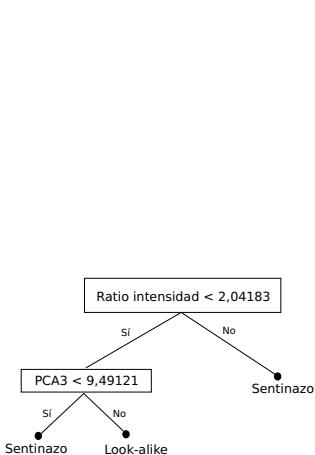
### ▷ Classification

- Clustering of oil spills and look alike.
- Evaluation of the characteristics vector.
- Machine learning classifiers.
  - > Artificial Neural Network
  - > Decision Tree



# Sentinazos Methodology

## ▷ Classification

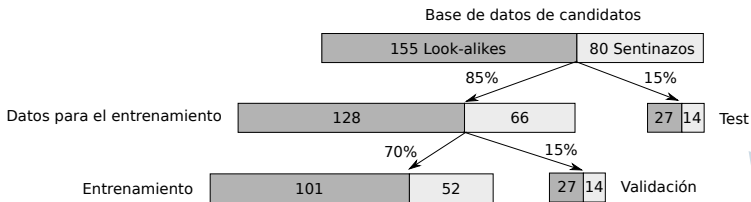


# Index

- 1 Context and Motivation
- 2 Objectives
- 3 Virtual laboratory development
- 4 Virtual laboratory validation
- 5 Results and conclusions

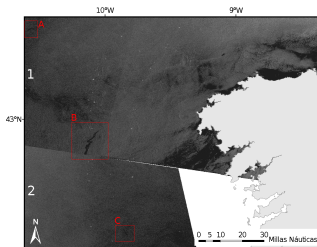


# Results and conclusions

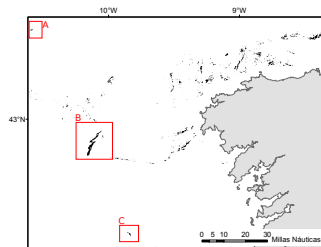


	Validation set		Test set	
	Sentinazos	Look-alikes	Sentinazos	Look-alikes
ANN	85,7 %	85,2 %	92,9 %	96,3 %
Pruned decision tree	92,9 %	85,2 %	92,9 %	92,6 %

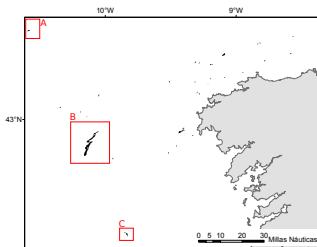
# Results and conclusions



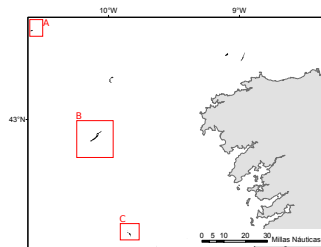
a



b

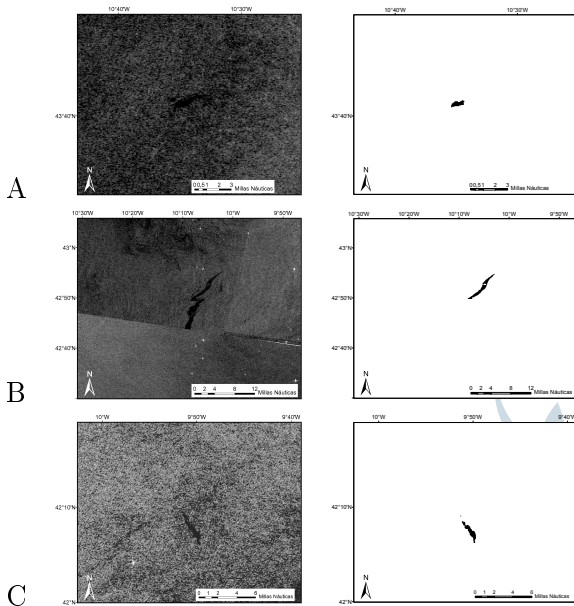


c



d

# Results and conclusions



# Results and conclusions

MainWindow

File Noise Filters Tools Oil Spills ML

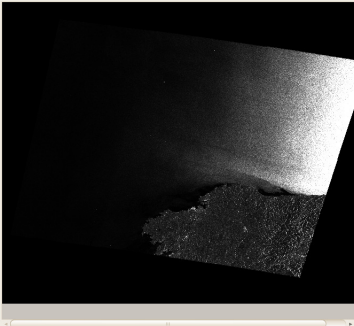
Image World Map Edited

Image Name: `josgalician@ffs/2011/14-10-2011/ASA_WSM_1PNIPA20111014_105250_000000733107_00310_50324_4066.N1_Crv.tif`

Geotiff Information:  
Version: 1  
Key\_Revision: 1.2  
Tagged Information:  
ModelTransformationTag (4,4):  


0,0007793291006237270				
0	-0,0007793291006237270			
0	0	0	0	
0	0	0	0	1

  
End\_Of\_Tags.  
Keyed Information:  
GTMModelTypeGeoKey (Short,1): Mod  
GTRasterTypeGeoKey (Short,1): Rast  
GTCitationGeoKey (Ascii,7): "WGS 84  
GeographicTypeGeoKey (Short,1): G  
PCSCitationGeoKey (Ascii,7): "WGS 84  
End\_Of\_Keys.  
End\_Of\_Geotiff:  
3CS: 4326/WGS 84  
Datum: 6326/World Geodetic System 1984  
Spheroid: 7030/WGS 84 (6378137,00,6378137,00)  
Prime Meridian: 8901/Greenwich (0,000,0)  
Prime Meridian: 8901/Greenwich (0,000,0)



ZoomIn

ZoomOut



# Results and conclusions

Firefox RETELAB Grid Portal

Logout  
Bienvenido, David Mera

**RETELAB**  
LABORATORIO VIRTUAL PARA LA RED NACIONAL DE TELEDETECCIÓN OCEANOGRÁFICA

Welcome Administration file-manager Primary Production Jobssubmit-portlet Proxymanager Portlet **sentinazos** gridway-datagrid-portlet

Proyecto Sentinazos

Sentinazos

**PROYECTO SENTINAZOS**

Título del trabajo: galicial2-11-2012

Acción a realizar: Segmentar y Clasificar

Clasificador utilizado: Redes Neuronales

Imagen SAR: \_\_\_\_\_

IA: \_\_\_\_\_

Vientos: \_\_\_\_\_

Remoto: \_\_\_\_\_

Local: \_\_\_\_\_

Enviar

**Virtual Data Base**

Region: \_\_\_\_\_

Sensor: \_\_\_\_\_


Parameter: \_\_\_\_\_

Search

📁 /home/retelab/david\_mera

**select\_input\_files**

- 📁 SARImage.tif
- 📁 SARImageWind.nc
- 📁 SARImageIA.tif



12 de noviembre de 2012



# Results and conclusions

## ▷ Conclusions

- Processing time.
- The AT could be improved using other wind models.
- The oil spills inside of low wind areas are not discovered.
- Is the shape relevant?

## ▷ Ongoing work

- New wind speed models.
- New satellites (Sentinel).
- New classifiers.
- Add contextual data (ships, FTSS, etc).



# Collaborations

## ▷ Retelab

- Marine and Food Technological Centre (AZTI Tecnalia)
- The Centre of Supercomputing of Galicia (CESGA)
- Canarian Institute of Marine Sciences (ICCM)

## ▷ Sentinazos

- University of Coruña
- Median Engeniering Group (GIM), University of Extremadura
- MacDonald Image Lab, Department of Earth, Ocean and Atmospheric Sciences, Florida State University
- Spanish Maritime Safety Agency (SASEMAR)

Thank you!!

