



MINISTERIO
DE ECONOMÍA
Y COMPETITIVIDAD

CSIC
CONSEJO SUPERIOR DE INVESTIGACIONES CIENTÍFICAS



*Relevant data and parameters for understanding and
validating models and comparing observations about
Water Reservoirs and Lakes,
with specific emphasis on addressing **Algae Bloom***

Presented by A.Monteoliva (ECOHYDROS SL, Spain)

J.Marco, F.Aguilar (IFCA, CSIC-UC)

@ Global Water Information IG,

7th RDA plenary, 1st March 2016, Tokyo

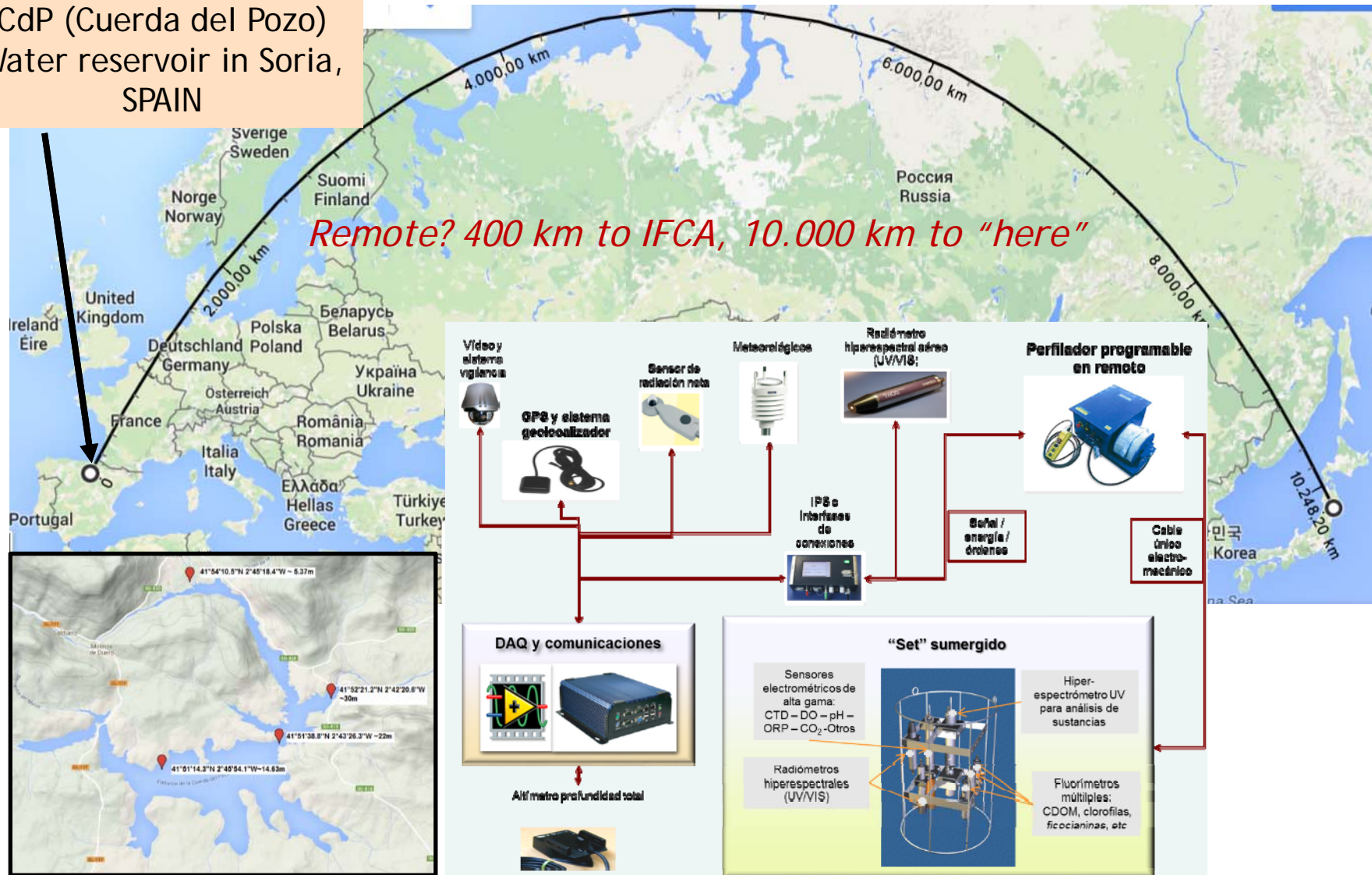


Introducing our collaboration on Water Reservoirs

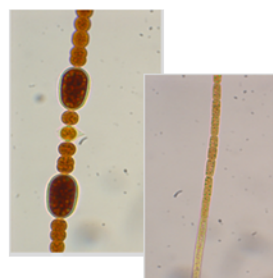
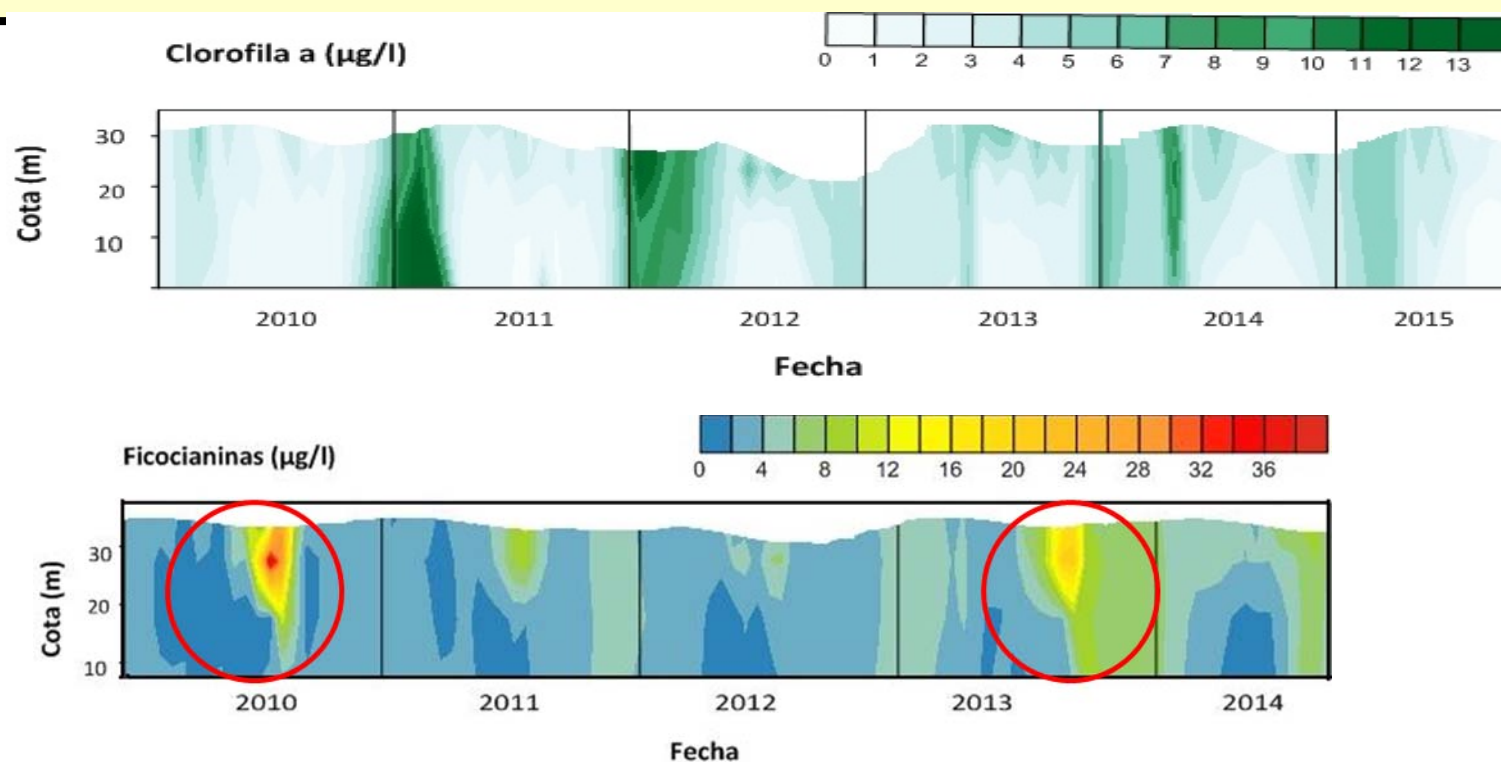
- ✿ *IFCA = basic research center (CSIC=National Research Council in Spain, + University of Cantabria), Advanced Computing & e-Science group*
- ✿ *ECOHYDROS SL = SME on Advanced Water Ecology*
- ✿ **2005: Start of IFCA-ECOHYDROS (SME) collaboration**
 - Modeling for Itoiz water reservoir, CWM models (ELCOM, DYRESM+ CAEDYM)
Modelling of A Watershed: A Distributed Parallel Application in a Grid Framework
- ✿ **2008: DORII EU FP7 project**
 - **Monitoring platform** @CdP, Instrumentation, Labview, Web Services, Grid Instrumentation
- ✿ **2012: ROEM+ (LIFE+ project)**
 - Extension of monitoring system
 - **Start of use of Delft3D for hydrological and water quality (algae bloom)**
- ✿ **2013: Connecting with LifeWatch (EU ESFRI) initiative**
 - Presentation to EGI.eu
 - SCARCE Int.Conference, NETLAKE
- ✿ **2014: Extension of monitoring network**
 - Cogotas water reservoir: new monitoring station connected via 3G
- ✿ **2015: Case Studies for EGI LW Competence Center and INDIGO-DataCloud**
- ✿ **2016: New initiative at Sanabria (Alpin) Lake in Spain**

Monitoring a Remote Water Reservoir

CdP (Cuerda del Pozo)
Water reservoir in Soria,
SPAIN

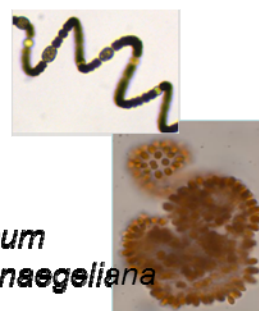


From monitoring to a CHAB warning system



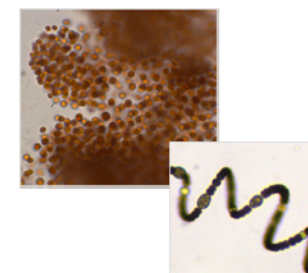
2010 - 2011

Dolichospermum planctonicum
Aphanizomenon flos-aquae



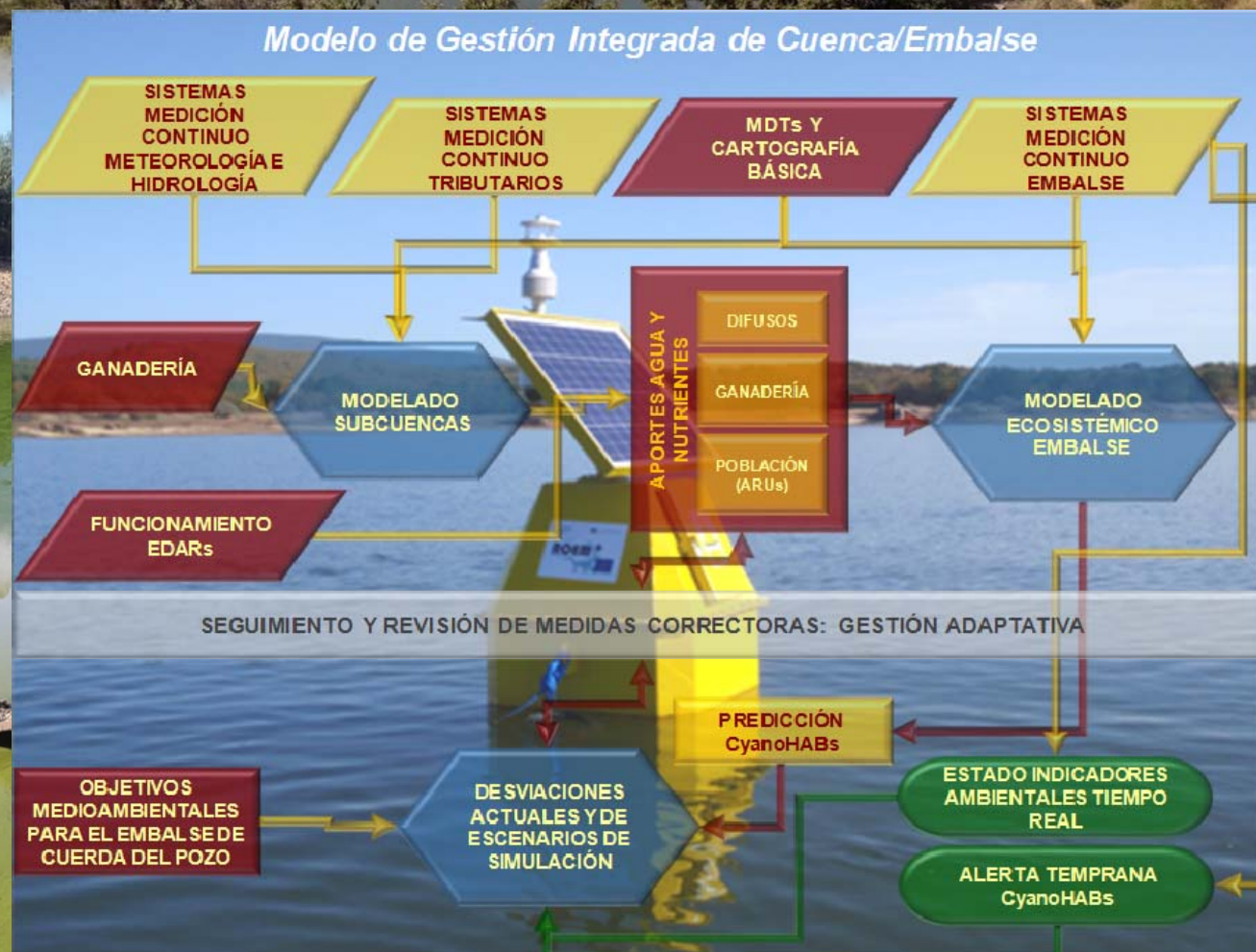
2013

Dolichospermum crassum
Colonias Woronichinia naegeliana



2014

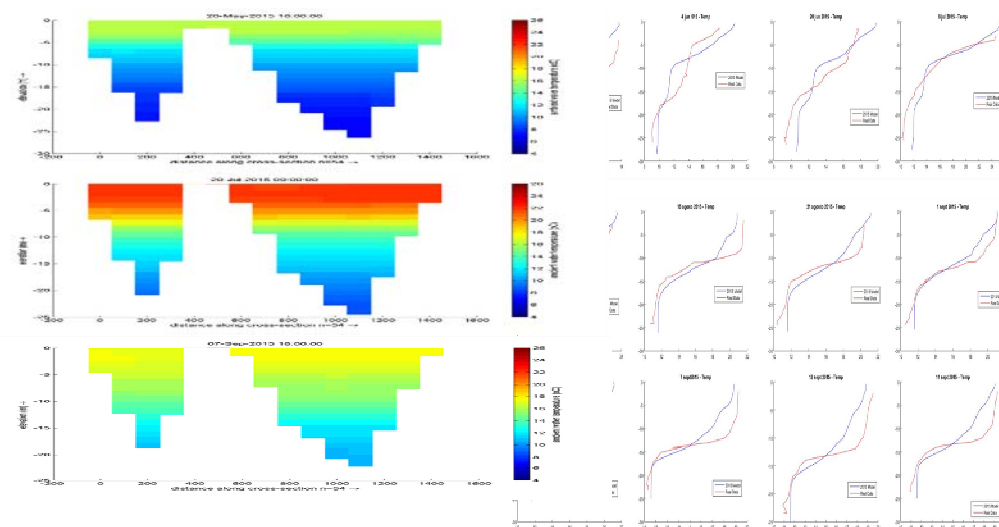
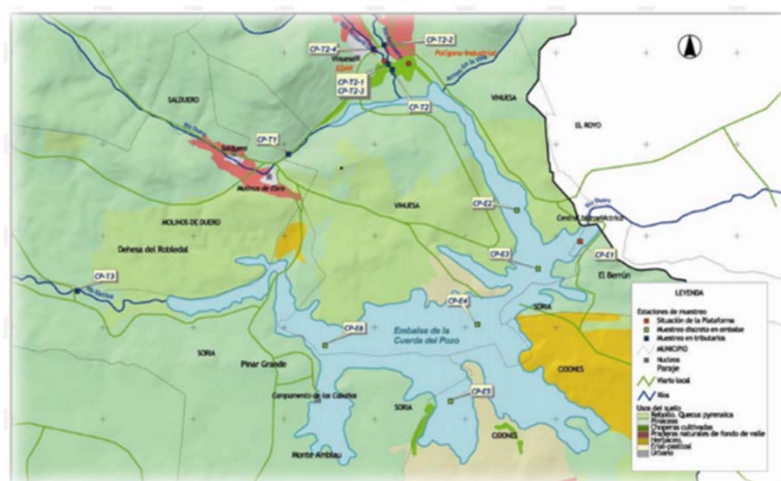
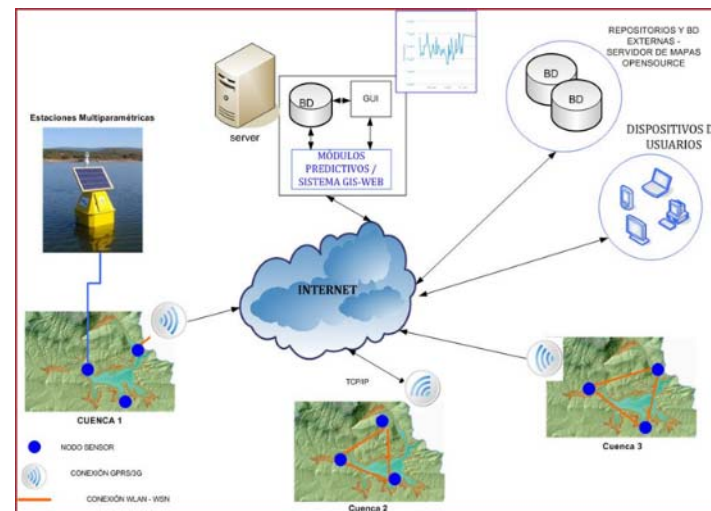
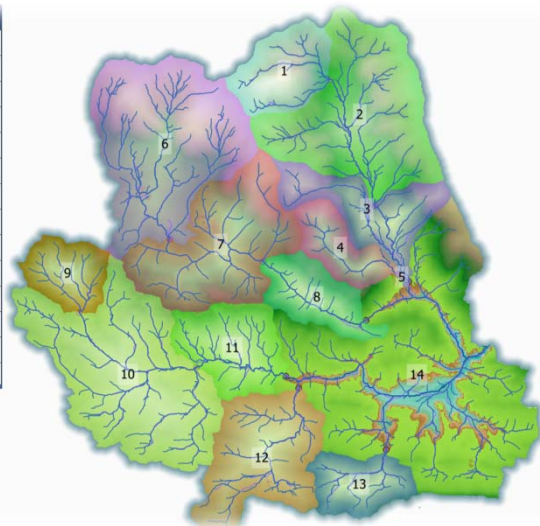
Colonias Microcystis novacekii
Dolichospermum crassum



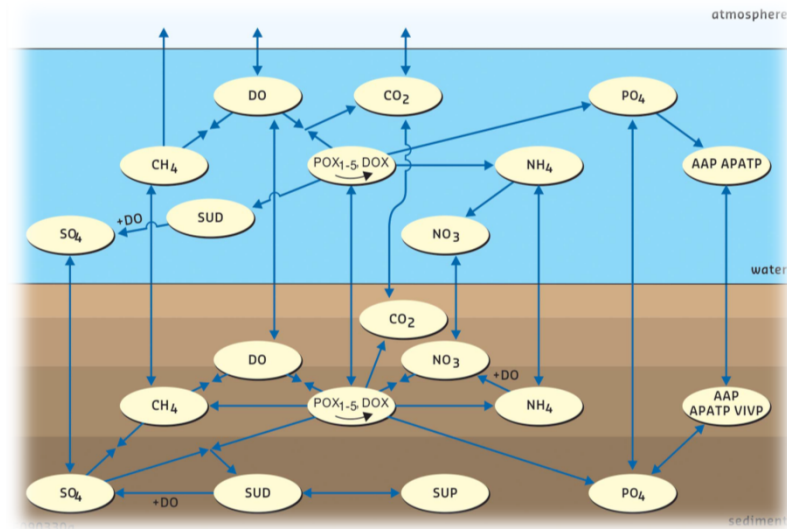
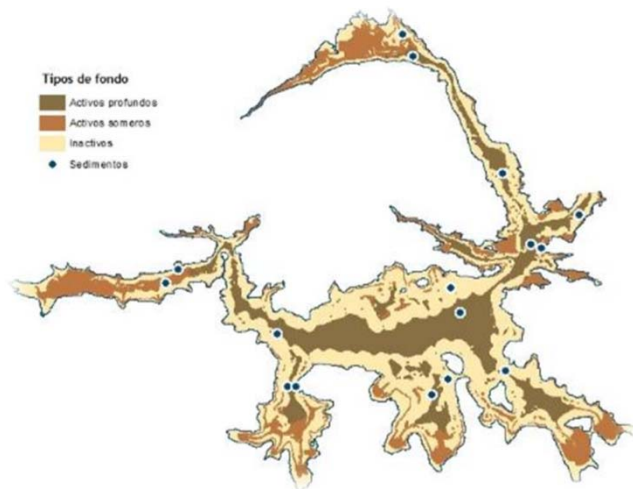
FROM A.MONTEOLIVA PhD THESIS, 2016

(Validated) Hydrological Model: Delft3D

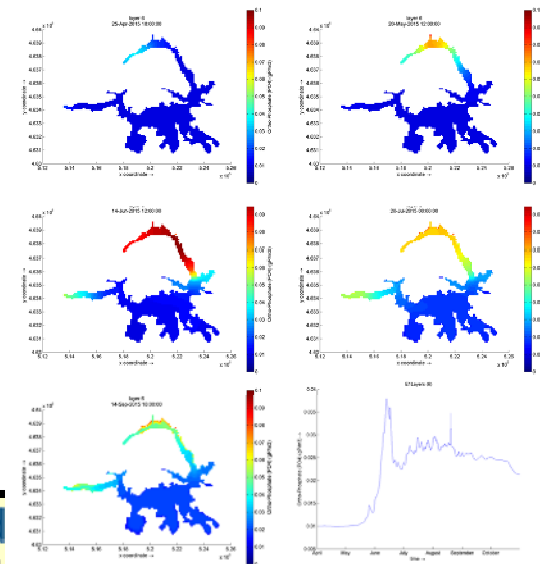
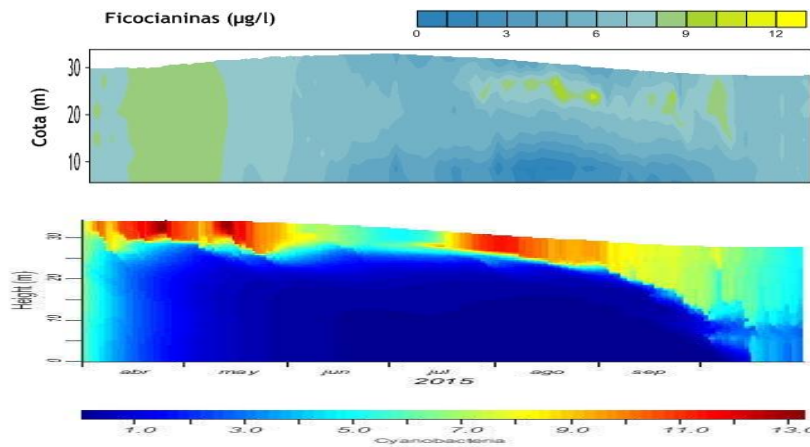
ID	Nombre	AREA (Ha)
1	Revinuesa	2.142
2	Revinuesa	6.003
3	Revinuesa	2.530
4	Remoncio	1.397
5	Remoncio	13
6	Duero	6.720
7	Duero	4.878
8	Duero	1.810
9	Ebrillos	1.569
10	Ebrillos	7.249
11	Ebrillos	2.495
12	Dehesa	3.812
13	Bajero	1.615
14	CED	12.390



BioGeoChemical Models: large number of processes and parameters !!!



1-D models are not enough, **we need to reproduce the evolution in 3D**



Our problem, our aim...

- ✚ We would like to have a well defined set of parameters/variables allowing us to explore the complete modeling of a water reservoir, describing algae blooms, and its validation against monitoring data
- ✚ We would like to integrate this set (as a vocabulary? As an ontology?) into our OPEN DATA PORTAL, being organized around the FULL DATA LIFE-CYCLE
- ✚ **So we fully support Ilya suggestion (email) to consider:**

*-How to setup a formal vocabulary of such basic parameters, and how to agree on the meaning, definitions and data for these parameters;
-how to manage and reference time series and possibly real time data feeds so that to ensure reproducibility of model results and comparison;
-how to maintain identifiers of hydrologic features across various model runs and throughout data lifecycle;
-how to publish such basic parameter data as part of model validation workflows;
-how to deal with large volumes of model data*

OUR e-INFRASTRUCTURE FRAMEWORK

- **LifeWatch (lifewatch.eu) is an ESFRI** (*EU Research Infrastructure*)
 - Addressing Biodiversity & Ecosystems
 - An e-Infrastructure to build Virtual Research Environments (VRE)
 - Integrating **OPEN DATA** information
 - GBIF, LTER, GENBANK, SATELLITE IMAGES, TERRESTRIAL MAPS...
- **EGI-LifeWatch Competence Center**
 - Framework: EGI FedCloud
 - Dedicated Resources (~5000 cores + PB, new node in Seville, Andalusia, SPAIN)
- **Support LW VRE**
 - Marine VRE (marine.lifewatch.eu)
 - Terrestrial + FreshWater VRE
- **Pilot projects**
 - **Ecological Observatories Data Flow and “Big Data” analysis**
 - Workflows: Galaxy and TRUFA; Network of Life
 - Citizen Science: Assisted Pattern Recognition

Monitoring & Modeling ALGAE BLOOM in a Water Reservoir

LIFE+ Project lead by a SME, collecting monitoring data (environmental station+ water quality and chloro-cyano profiler), and modeling hydro+bio

– INTEGRATE EXISTING OPEN RESEARCH DATA

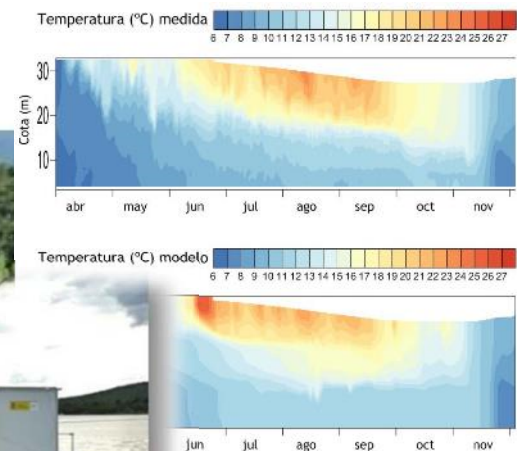
- USE **METEO**, TERRAIN, BATHIMETRY, LAND USE
- **HYDROLOGICAL** INPUT

**STORAGE+COMPUTING
NEED HPC FOR DELFT-3D**

– PRESERVE NEW OPEN RESEARCH DATA & ANALYSIS

- REMOTELY COLLECTED DATA INTO REPLICATED DB
- COMPLEX MODEL OUTPUTS
- MULTIPARAMETRIC ANALYSIS

*Model already running in FedCloud
Adapt Multiparametric scan
Preserve Thermoclines analysis (in R)*



SOLUTIONS EXPLORED

See OPEN DATA Commons session @ RDA Paris

- Support external resources (data, tools): **VRE**
- Enable a “/lifewatch/home” for each researcher/each community, accessible with ID via a **preservation portal**
- Users will define the “openness” of their
 - DATA (private/**embargo**/open/published-DOI)
 - ANALYSIS (R/python, via github)
 - WORKFLOWS at SaaS level (R,python)
- Support it with a global (federated) distributed storage
 - OneData (Data Commons basic component)
- Integrated also with FedCloud computing resources
 - We will rely on INDIGO project developments to optimize!
- Enforce DMP (Data Management Plan)



*If it needs to be preserved => **DMP** & **OPEN** (after embargo)*

Summary/ next steps


- ✚ Eutrophication is an important (increasing?) problem
- ✚ We aim to model it in detail!
- ✚ Along next weeks we will be sweeping the space of bio-geo-chemical parameters in ECO-DELFT
- ✚ We will be also populating our new OPEN DATA PRESERVATION portal
- ✚ We would like to use STANDARD PARAMETERS/VARIABLES TO ASSURE A GOOD DESCRIPTION USING 3D MODELS
- ✚ We will also employ the scheme to enable the processing/validation of simulations



elroto.elpais@gmail.com

Join us at Indigo WP2 & EGI-LW CC @ EGI Conference in AMSTERDAM, 4-8 April

Key question: Incorporate Digital Knowledge

- ✚ Software (VM) + Data preservation is not enough
 - Ideas explored under CMS preservation
 - Validation
 - Analysis Description
 - CHEP 2014 discussion  Knowledge Preservation
- ✚ Incorporate Digital Knowledge from start
 - Understand the use of “ontologies” / “semantics”
 - Ontologies are not taxonomies
 - Ontologies are not metadata
 - Ontologies are not (restricted) vocabularies
- ✚ *Under analysis for Fresh Water VRE*
 - *Ontological Framework explored: SWEET*
 - *Ontologies: EML and WaterEML*
 - *Consider report at RDA 2015 (Paris) devoted session*
 - *What about INSPIRE directive?*
- ✚ *First try:*
 - *Build on SWEET*
 - *Start from vocabularies used in CdP*
- ✚ Yet, how to integrate into DMP? Ideal (unify) is obvious...

Incorporating Digital Knowledge...

CuerdaDelPozo (http://www.semanticweb.org/guadalupecanasherrera/ontologies/2015/10/CuerdaDelPozo) : [/Users/guadalupecanasherrera/CuerdaDelPozo]

Active Ontology x Entities x Object Properties x Data Properties x Individuals by class x OntoGraf x SPARQL Query x

SPARQL query:

```
PREFIX rdf: <http://www.w3.org/1999/02/22-rdf-syntax-ns#>
PREFIX owl: <http://www.w3.org/2002/07/owl#>
PREFIX rdfs: <http://www.w3.org/2000/01/rdf-schema#>
PREFIX xsd: <http://www.w3.org/2001/XMLSchema#>
PREFIX onto: <http://www.semanticweb.org/guadalupecanasherrera/ontologies/2015/10/CuerdaDelPozo#>

SELECT ?instruments ?magnitudes
WHERE {
  ?instruments rdf:type onto:ServoirPlatform .
  ?instruments onto:is_type onto:Water .
  ?instruments onto:measures_magnitude ?magnitudes
}
```

OWL via PORTEGE
(to be integrated !!!)

BUT.

We NEED key questions
to understand what we can/want to do

BACKUP SLIDE

instruments	magnitudes
radiometer	incident_upcoming_irradiance
radiometer	incident_downcoming_irradiance
fluorometer	CDOM
fluorometer	fycocynine
fluorometer	Chlorophyll_a
UV_Spectrometer	Nitrate
UV_Spectrometer	chemical_oxygen_demand
UV_Spectrometer	suspended_solids
UV_Spectrometer	Nitrite
UV_Spectrometer	Carbonate
UV_Spectrometer	spectrum
multiparametric_probe	depth_at_measurement
multiparametric_probe	water_temperature
multiparametric_probe	ORP
multiparametric_probe	Salinity
multiparametric_probe	dissolved_oxygen_concentration

Incorporating Digital Knowledge...

CuerdaDelPozo (http://www.semanticweb.org/guadalupecanasherrera/ontologies/2015/10/CuerdaDelPozo) : [/Users/guadalupecanasherrera/CuerdaDelPozo/CuerdaDelPozo.t

CuerdaDelPozo (http://www.semanticweb.org/guadalupecanasherrera/ontologies/2015/10/CuerdaDelPozo)

Active Ontology x Entities x Object Properties x Data Properties x Individuals by class x OntoGraf x SPARQL Query x

SPARQL query:

```
PREFIX rdf: <http://www.w3.org/1999/02/22-rdf-syntax-ns#>
PREFIX owl: <http://www.w3.org/2002/07/owl#>
PREFIX rdfs: <http://www.w3.org/2000/01/rdf-schema#>
PREFIX xsd: <http://www.w3.org/2001/XMLSchema#>
PREFIX onto: <http://www.semanticweb.org/guadalupecanasherrera/ontologies/2015/10/CuerdaDelPozo#>

SELECT ?magnitudes
WHERE { onto:Cyanobacteria onto:is_detected_by ?magnitudes }
```

magnitudes
fycocynine
spectrum
ORP
pH
suspended_solids
dissolved_oxygen_concentration
Carbonate
air_temperature
Chlorophyll_a
Nitrate
dissolved_oxygen_saturation
barimetric_pressure
chemical_oxygen_demand
relative_humidity
CDOM
Nitrite
Salinity

SPARQL via PORTEGE
(to be integrated !!!)

BUT.

We NEED key questions
to understand what we can/want to do

BACKUP SLIDE