Universitat Politècnica de València
Departamento de Ingeniería Hidráulica y Medio Ambiente
Programa de Doctorado en Ingeniería del Agua y Medioambiental

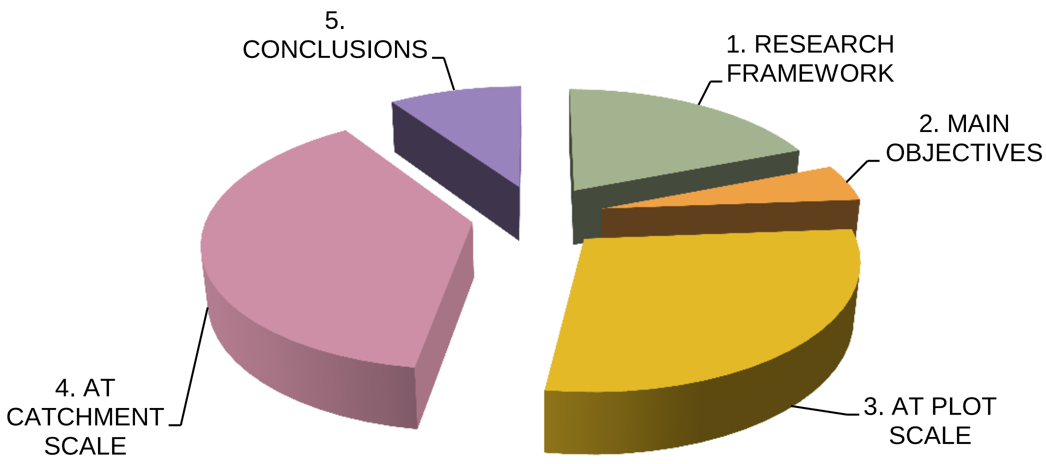
On the Use of Satellite Data to Calibrate a Parsimonious Ecohydrological Model in Ungauged Basins

Author: Guiomar Ruiz Pérez
Supervisor: Félix Francés García

September, 2016

Presentation Outline

- Research Framework
- Main objectives
- At plot scale: testing the model
- At catchment scale: spatio-temporal modelling
- Conclusions



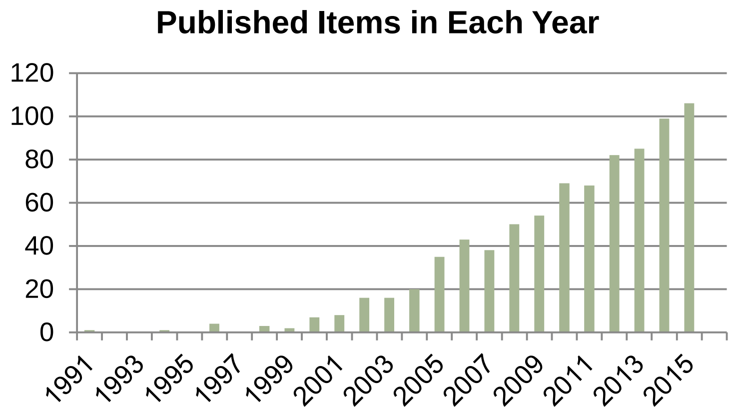
1. Research Framework

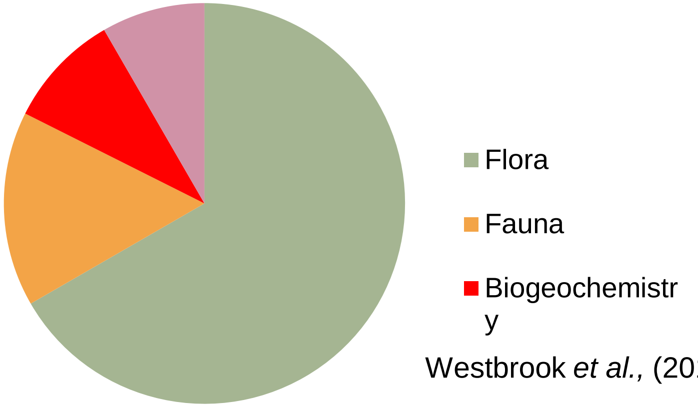
'Science, my lad, is made up of mistakes, but they are mistakes which it is useful to make, because they lead little by little to the truth'

-Jules Verne-

Ecohydrology: a new discipline

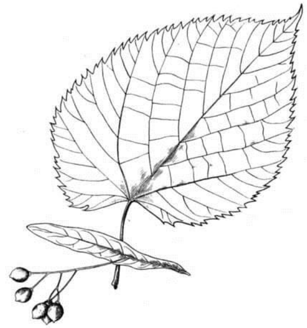
- This discipline seeks:
  - how hydrological processes influence the distribution, structure, function and dynamics of biological communities
  - how feedbacks from biological communities affect the water cycle
- In continuous growth
- Main topics:
  - Flora, Fauna, Biogeochemistry, Human impact





Westbrook *et al.*, (2013

Ecohydrological Modelling



- Water to survive
  - Affected by the spatio-temporal distribution of water
  - Strategies to cope with water restriction (stomatal closure, small leaves, rooting system)

- Primary conduit for returning terrestrial water
  - Transpiration is the largest water flux from Earth's continent
  - ET is ~67% of mean annual P
  - T is ~90% of ET

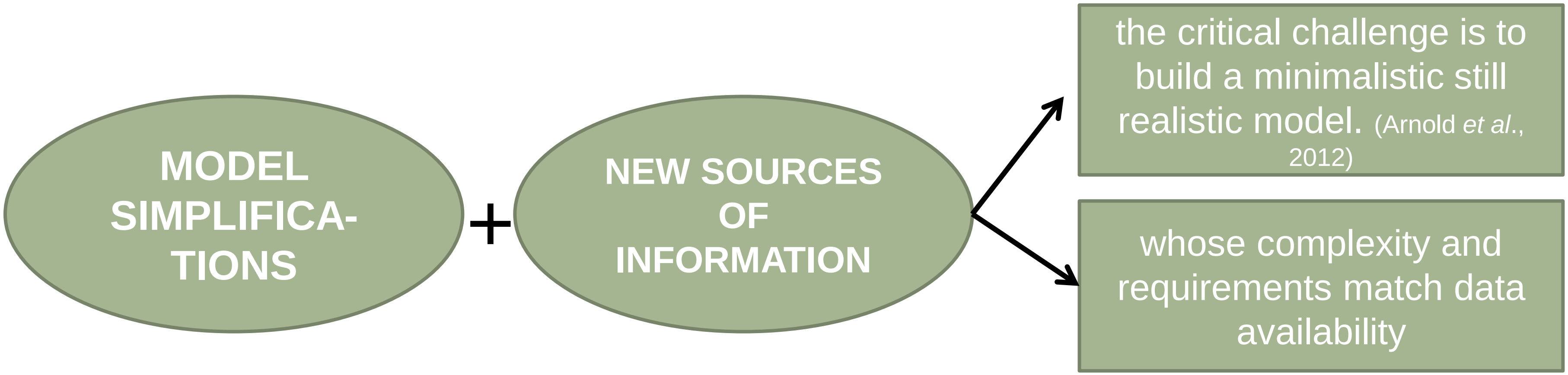
Zhang, Y. *et al.* (2016) and Jasechko, S., *et al.* (2013)

PLANT'S PIVOTAL ROLE → ECOHYDROLOGICAL MODELLING

Ecohydrological Modelling

- Traditionally very few hydrological models have included the vegetation as state variable
- Nowadays, they have increased substantially

High parametrical requirement – Data scarcity



Satellite Data

- Applicability of remote sensing data
  - Models forced by remote sensing data
  - Proxy of some parameters
  - Calibration and validation → challenging task
- Bibliographic survey of the Web of Knowledge:

Satellite | Calibration | Implementation | Modelling | Ecohydrology

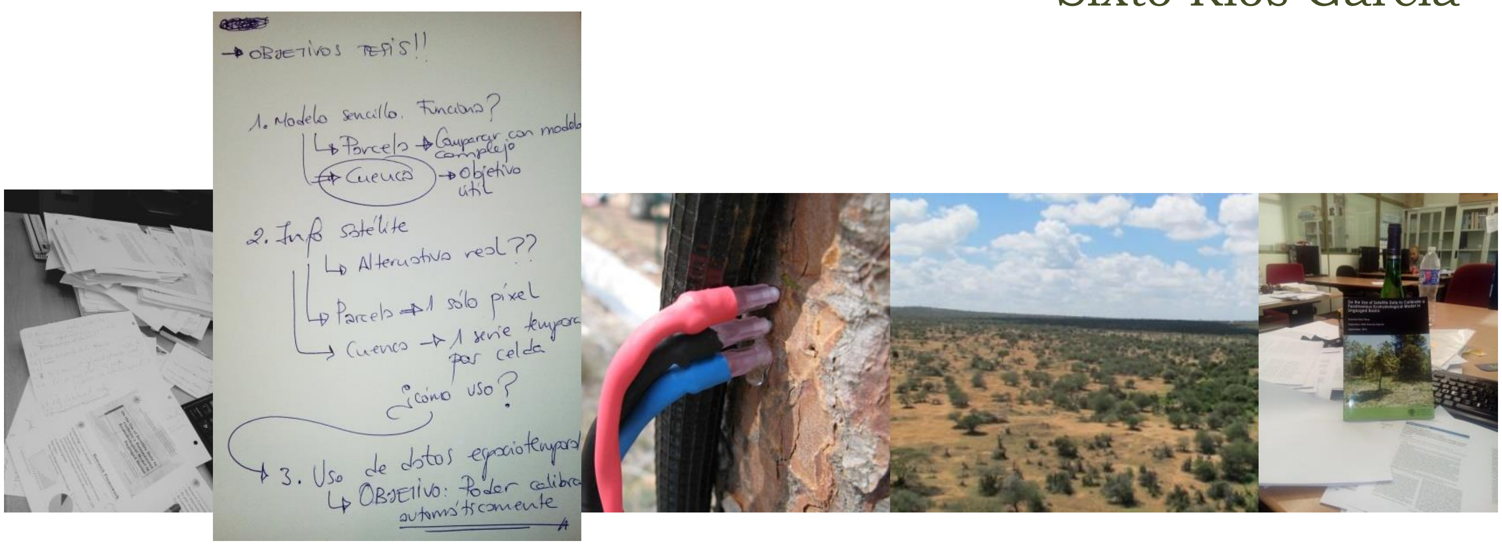
  - Lumped or semi-distributed models: 76.5%
  - Distributed models: 23.5% → Multi-objective approach

SPATIO-TEMPORAL DATA

2. Main Objectives

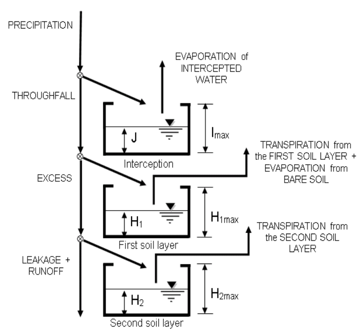
'[…] la cualidad más importante del matemático no es la memoria para retener mil fórmulas conocidas y emplearlas oportunamente, sino crear ideas fecundas y modelos sencillos bien adaptados a las situaciones reales. La simplificación es el objetivo.'

-Sixto Ríos García-



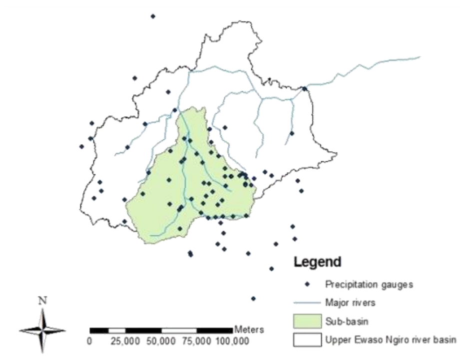
Main Objectives

- To test a parsimonious ecohydrological model in different places and at different working scales
- To explore the applicability of satellite data in ecohydrological modelling
- To develop a methodology to tackle the spatio-temporal data provided by satellite



The LUE-Model

At plot scale



At catchment scale

3. At plot scale: Testing the model

'There is always an open book for all eyes: Nature'

-Jean Jacques Rousseau-

Research Questions

- Is the proposed parsimonious model capable to satisfactory simulate vegetation and hydrological dynamics or is a more complex model needed?
- Could satellite products be used to implement a dynamic vegetation model or are field measurements totally necessary?

ELSEVIER

Ecological Modelling

Volume 324, 24 March 2016, Pages 45–53

Can a parsimonious model implemented with satellite data be used for modelling the vegetation dynamics and water cycle in water-controlled environments?

G. Ruiz-Pérez[a], M. González-Sanchis[b], A.D. Del Campo[b], F. Francés[a]

[a] Research Group of Hydrological and Environmental Modelling (GIHMA), Research Institute of Water and Environmental Engineering, Universitat Politècnica de València, Valencia, Spain

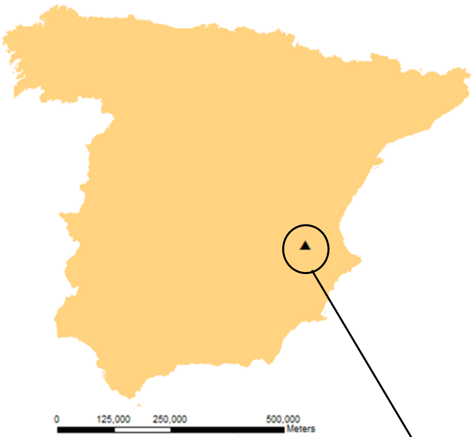
[b] Research Group in Forest Science and Technology (Re-ForeST), Research Institute of Water and Environmental Engineering, Universitat Politècnica de València, Valencia, Spain

Methodology/outline

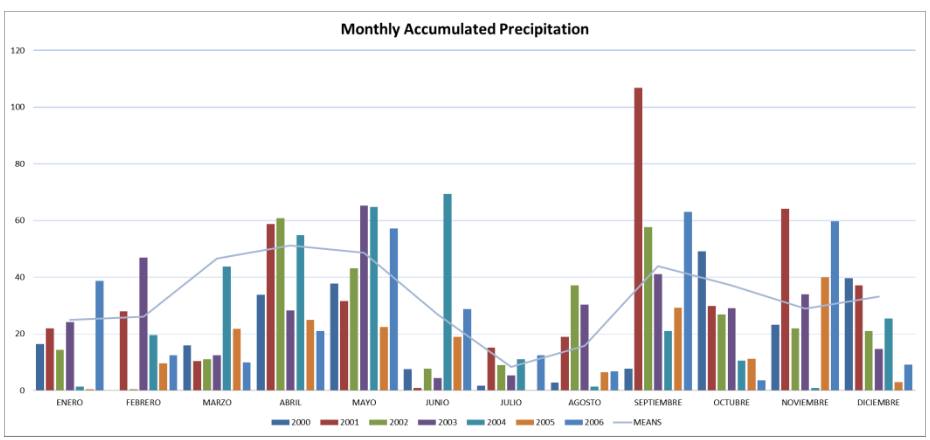
- Study area and models:
  - **Study area**: Aleppo pine experimental plot in La Hunde forest (East Spain)
  - Proposed **parsimonious dynamic vegetation model** (LUE-Model)
  - Selected **complex dynamic vegetation model** with successful results in the study area (Biome-BGC)
- Implementation of both models:
  - LUE-Model: with only NDVI (**satellite information**)
  - Biome-BGC: with **field data**
- Analysis of **results and conclusions**

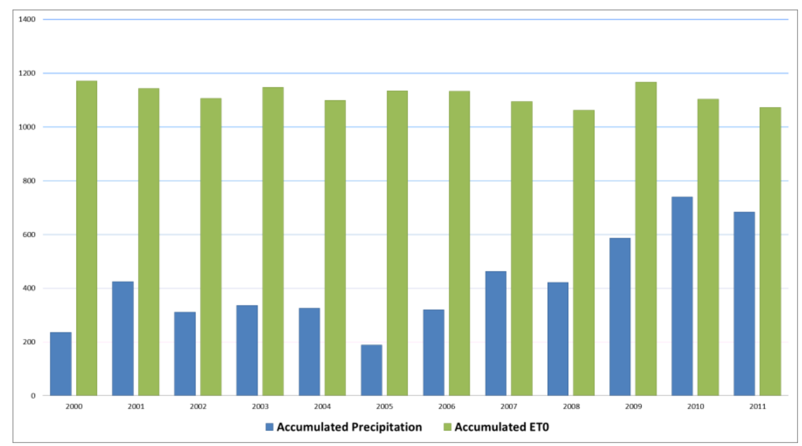
Study Area

- Experimental plot
  - Mediterranean semiarid climate:
    - Seasonality
    - Water-controlled area
  - Aleppo pine

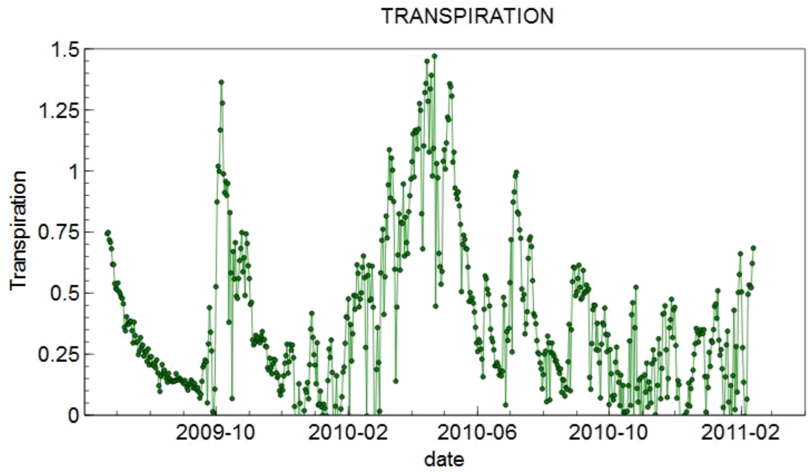


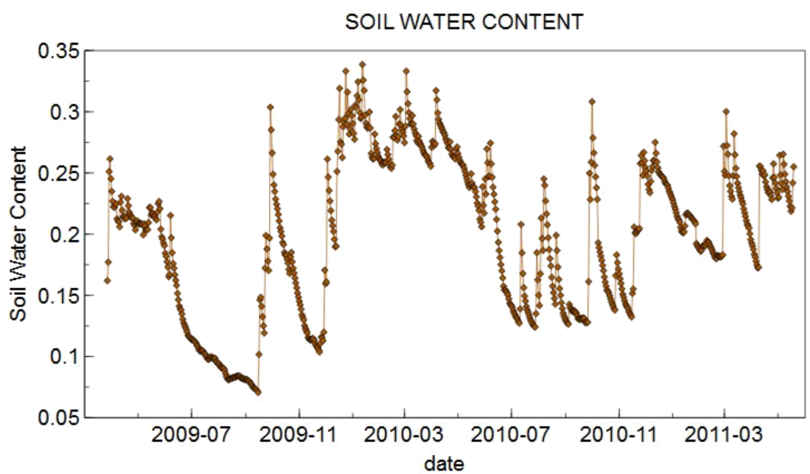
Experimental plot location





- Field Data
  - Transpiration
    - Sap flow sensors → Heat-Ratio Method
    - Three theoretical diameter classes
  - Soil Water Content
    - Soil Moisture sensors
    - 30 cm depth
    - 9 sensors:
      - 6 with tree's direct influence
      - 3 without tree's direct influence



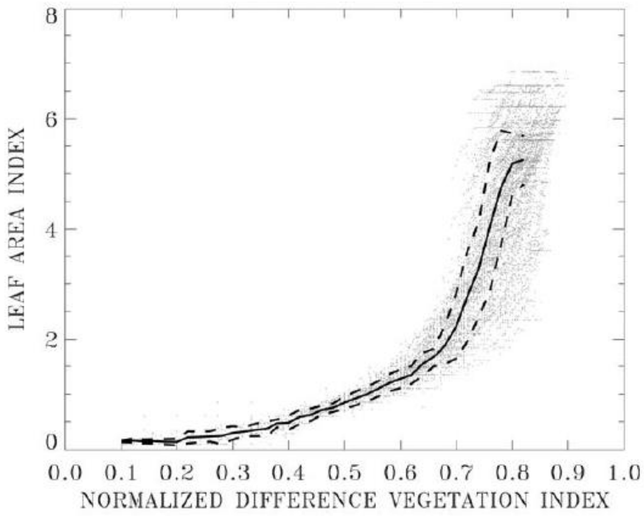


Data

- Satellite Data:
  - Vegetation Indices: NDVI and EVI

$$NDVI = \frac{NIR - RED}{NIR + RED} \qquad EVI = G * \frac{NIR - RED}{NIR + C_1 * RED - C_2 * Blue + L}$$

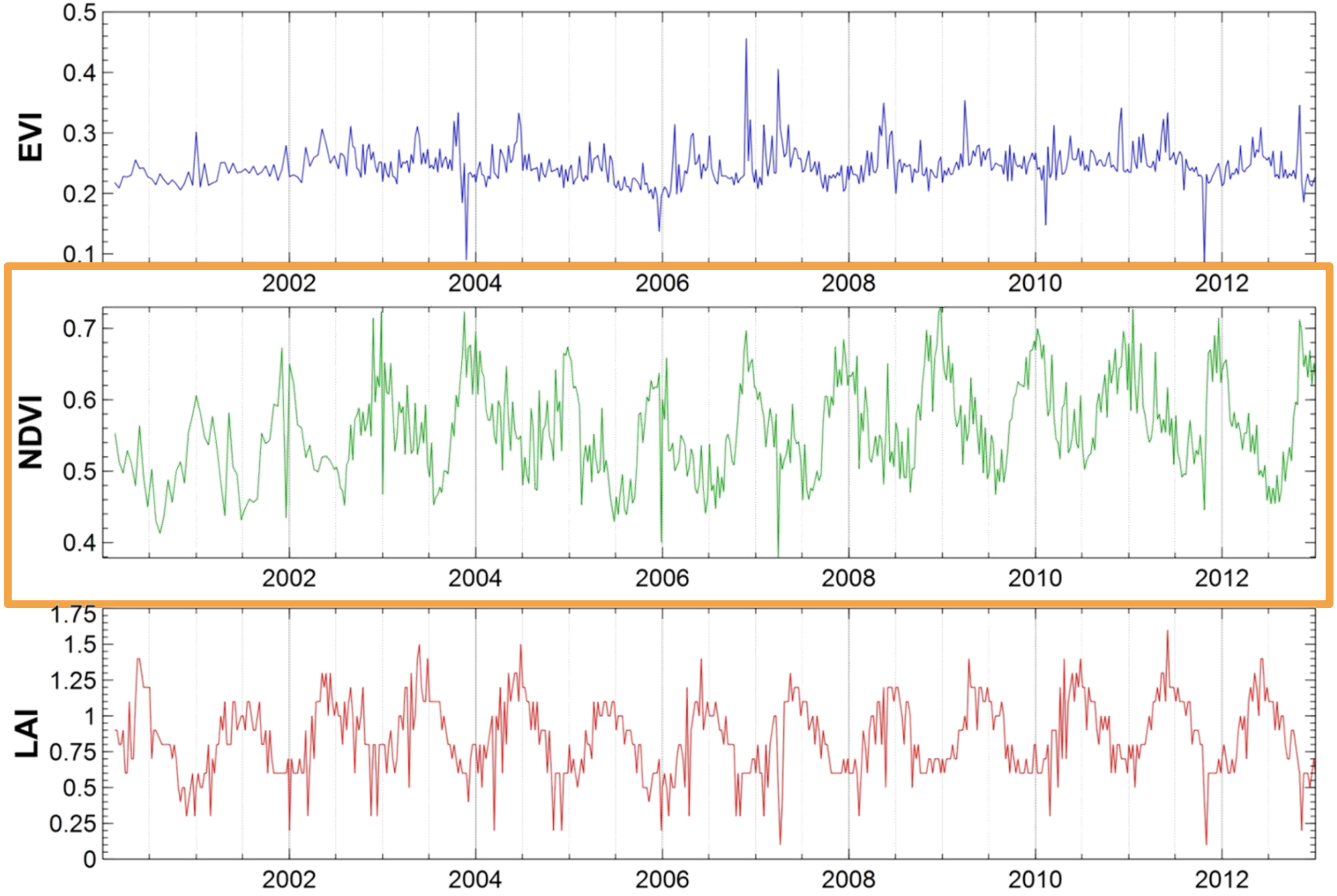
  - Other products: LAI (Leaf Area Index)



Extracted from Bruemann *et al.*, 2002

Data

- Satellite Data: Modis Vegetation Indices



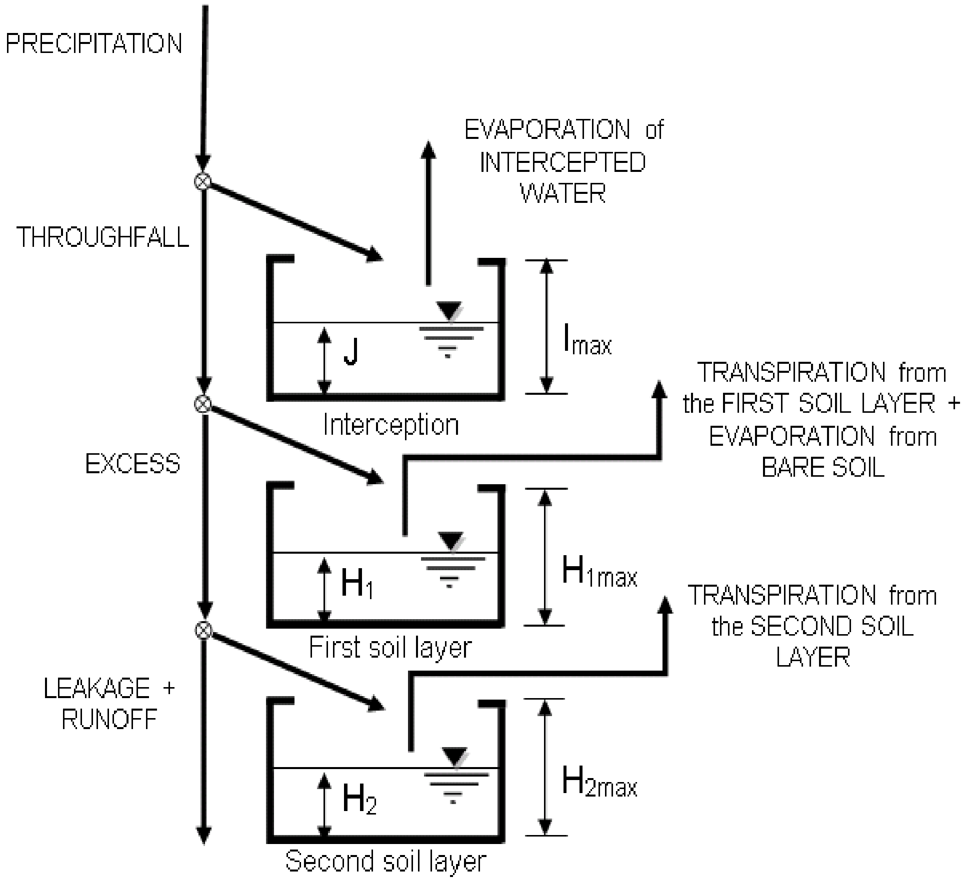
EVI
250m; 16days
No sense!

NDVI
250m; 16days
max_1: Nov/December
max_2: April/May
min: July/August

LAI
1km; 16days
max: March/May
min: Nov/January
Inconsistent with field data!

Description of the Models

- LUE-Model
  - Hydrological sub-model



Water balance

First soil layer

$$\frac{dH_1}{dt} = (P - I) - D - E - T_1$$

Second soil layer

$$\frac{dH_2}{dt} = D - L - T_2$$

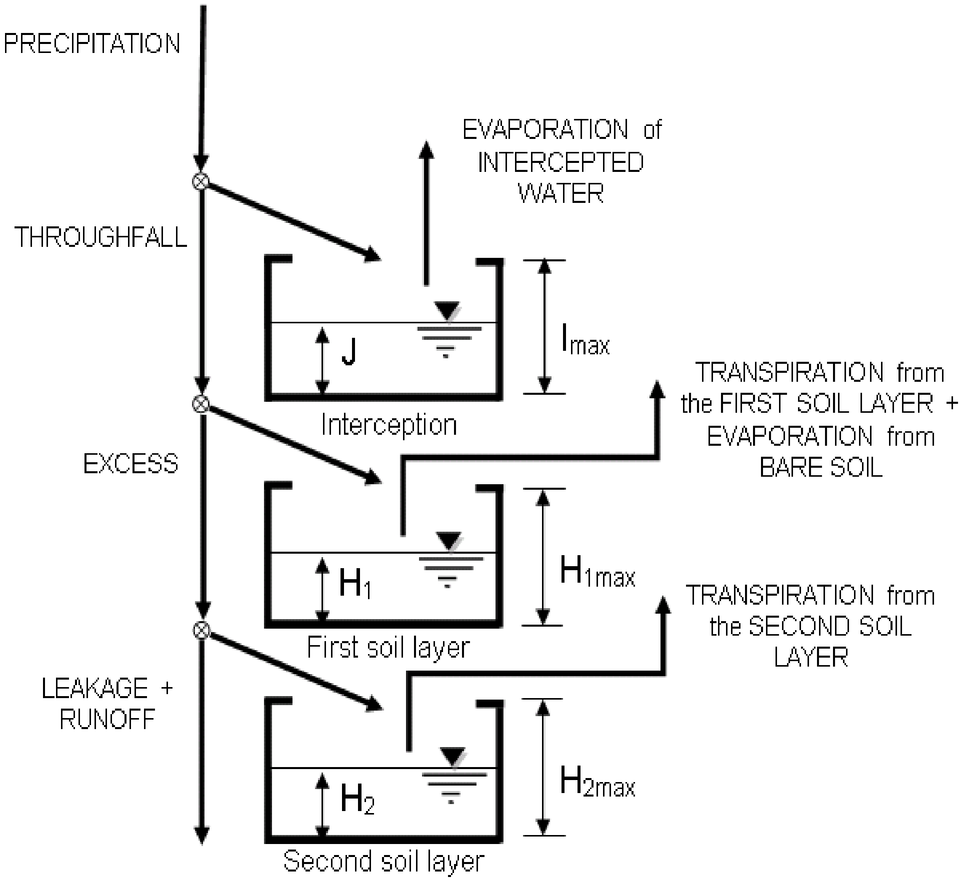
Interception storage

$$\frac{dJ}{dt} = I - \min\left(ET_o \cdot f_t, J\right)$$

Pasquato *et al.*, 2015

Description of the Models

- LUE-Model
  - Hydrological sub-model



Transpiration

FAO:

$$T = ET_o \cdot \lambda_v \cdot \lambda_s$$

$$T_1 = ET_o \cdot f_t \cdot \min(LAI,1) \cdot \beta_t(H_1) \cdot r_1$$

$$T_2 = ET_o \cdot f_t \cdot \min(LAI,1) \cdot \beta_t(H_2) \cdot (1 - r_1)$$

Bare Soil Evaporation

$$E = ET_o \cdot f_b \cdot \beta_b(H_1)$$

Description of the Models

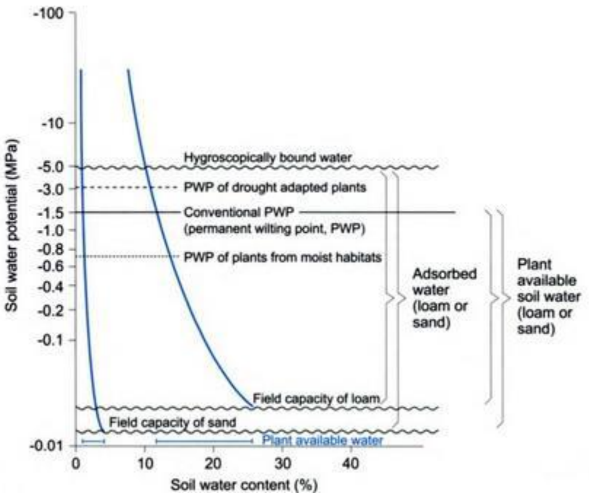
- LUE-Model
  - Hydrological sub-model

Transpiration

$$T_1 = ET_o \cdot f_t \cdot \min(LAI,1) \cdot \beta_t(H_1) \cdot r_1$$

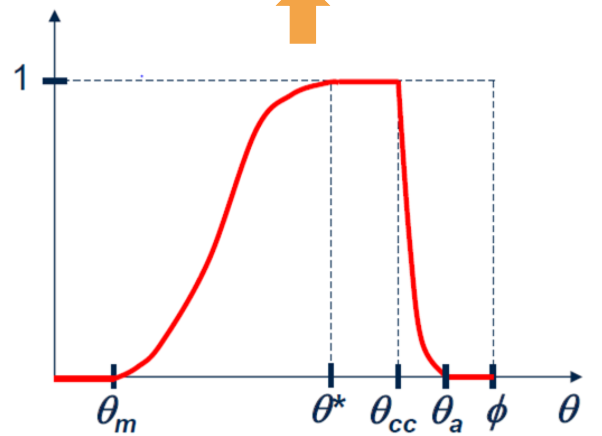
$$T_2 = ET_o \cdot f_t \cdot \min(LAI,1) \cdot \beta_t(H_2) \cdot (1-r_1)$$

$$\beta_j(H_i) = \begin{cases} 1 & for\ H_i \geq H_{i,cr} \\ \left(\frac{H_i - H_{i,lim}}{H_{i,cr} - H_{i,lim}}\right)^q & for\ H_{i,lim} < H_i < H_{i,cr} \\ 0 & for\ H_i \leq H_{i,lim} \end{cases}$$



$$\psi = \psi_{ae} * \left(\frac{n}{H}\right)^b$$

Clapp and Hornberger



Description of the Models

- LUE-Model
  - Dynamic vegetation sub-model

$$\frac{dB_l}{dt} = \left(LUE \cdot \varepsilon \cdot APAR - \mathrm{Re} \right) \cdot \varphi_l - \kappa_l \cdot B_l$$

$$\varphi = 1 - \frac{LAI}{LAI_{\max}}$$

LEAF BIOMASS

B_l [kg DM m^{-2} veg cover]

LIGHT USE EFFICIENCY

LUE [kg DM m^{-2} MJ^{-1}]

ε depends on:

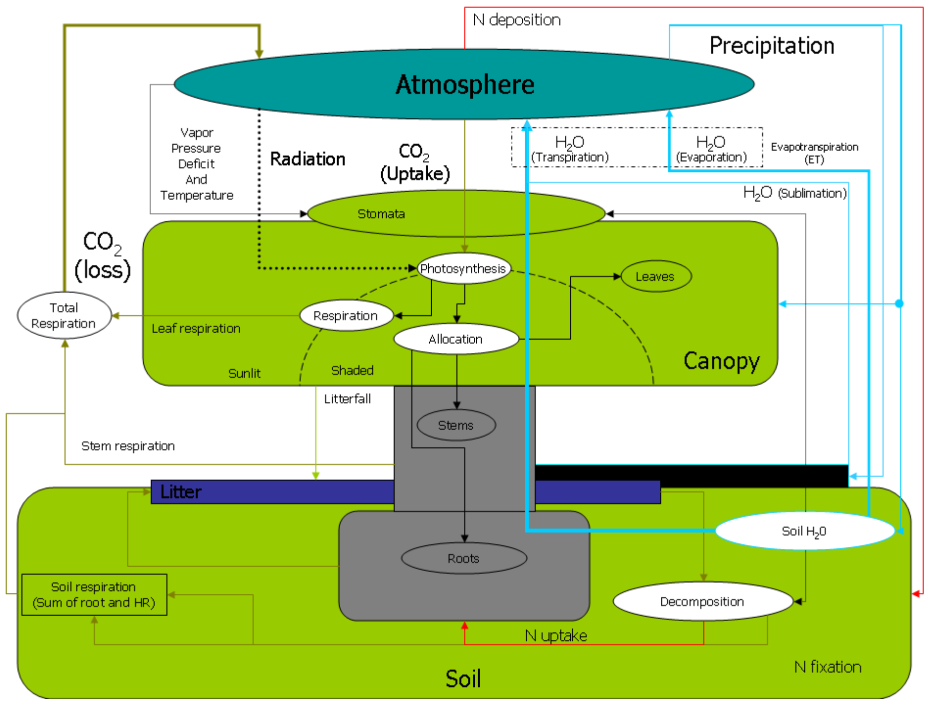
- Water Stress → connection with hydrological model
- Temperature

$$LAI = B \cdot SLA \cdot f_t$$

William and Albertson (2005)

Description of the Models

- Biome-BGC Model
  - Complex Physically-based model
  - Source: Numerical Terradynamic Simulation Group. Montana University



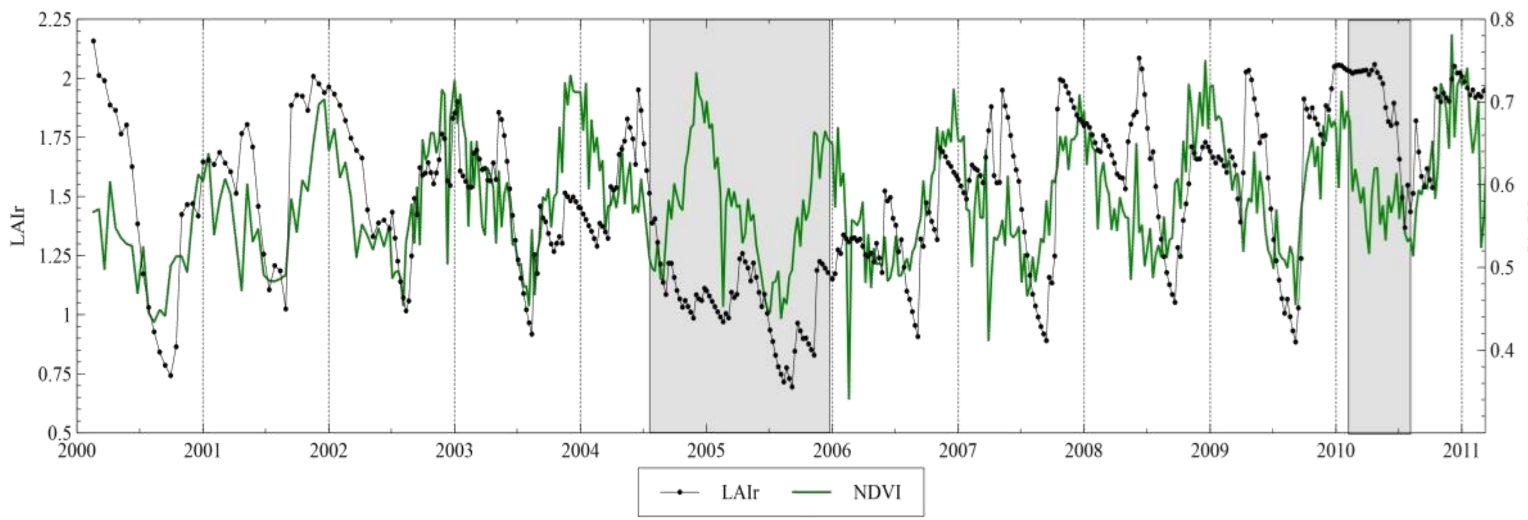
extracted from the Theoretically Framework of Biome-BGC, 2010

Calibration and Validation

- Implementation of the models:
  - LUE-Model:
    - Thirteen parameters
    - Calibration: NDVI data provided by Modis
    - Validation: Field measurements
  - Biome-BGC:
    - More than thirty parameters
    - Calibration: 70% of the field measurements
    - Validation: 30% of the field measurements
  - Goodness-of-fit indexes:
    - Nash and Sutcliffe efficiency index
    - RMSE
    - Pearson Correlation coefficient

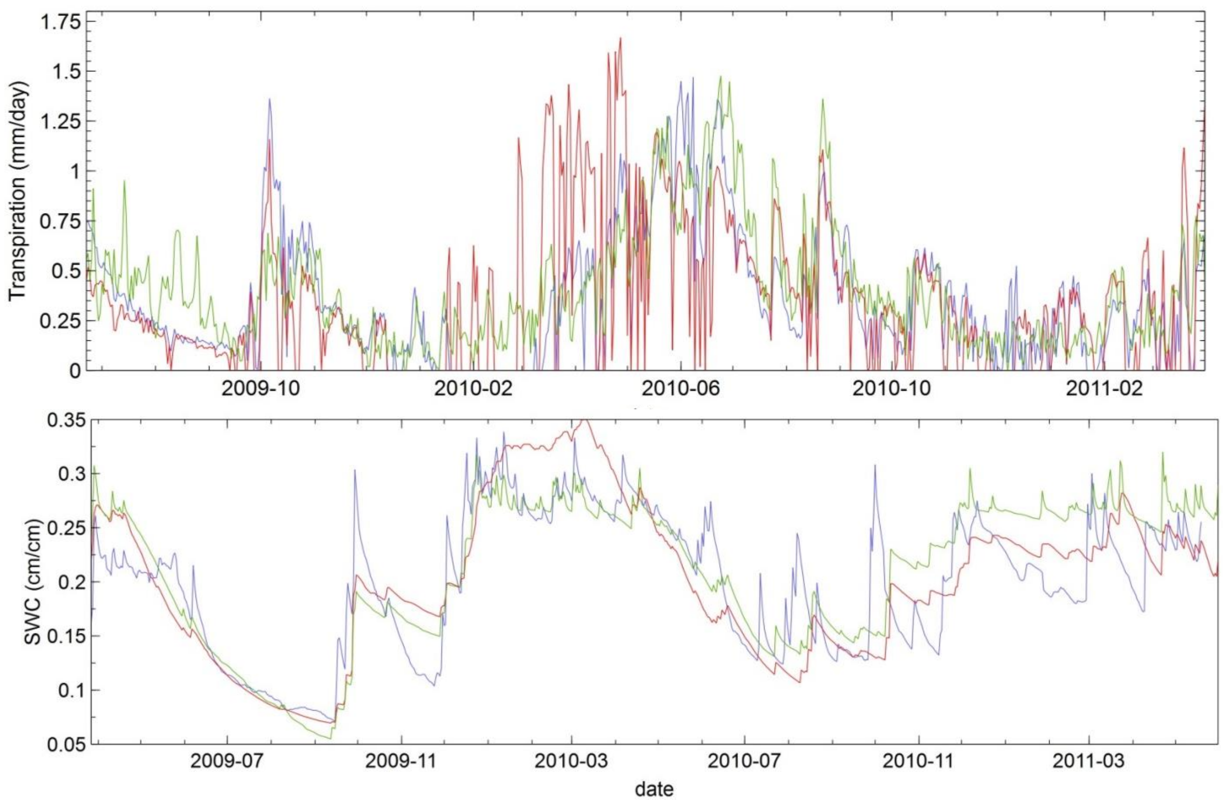
Results and Discussion

- Comparison between LAI simulated by the model and the NDVI provided by satellite
  - Strong correspondence
  - Two exceptions:
    - From July 2004 to December 2005 → extremely dry period
    - Spring 2010→ extremely wet period



Results and Discussion

- Comparison between the models



FIELD | BIOME | LUE

LUE-MODEL
RMSE= 0.360 E=0.34

BIOME-BGC
RMSE= 0.282 E=0.64

LUE-MODEL
RMSE= 0.06 E=0.42

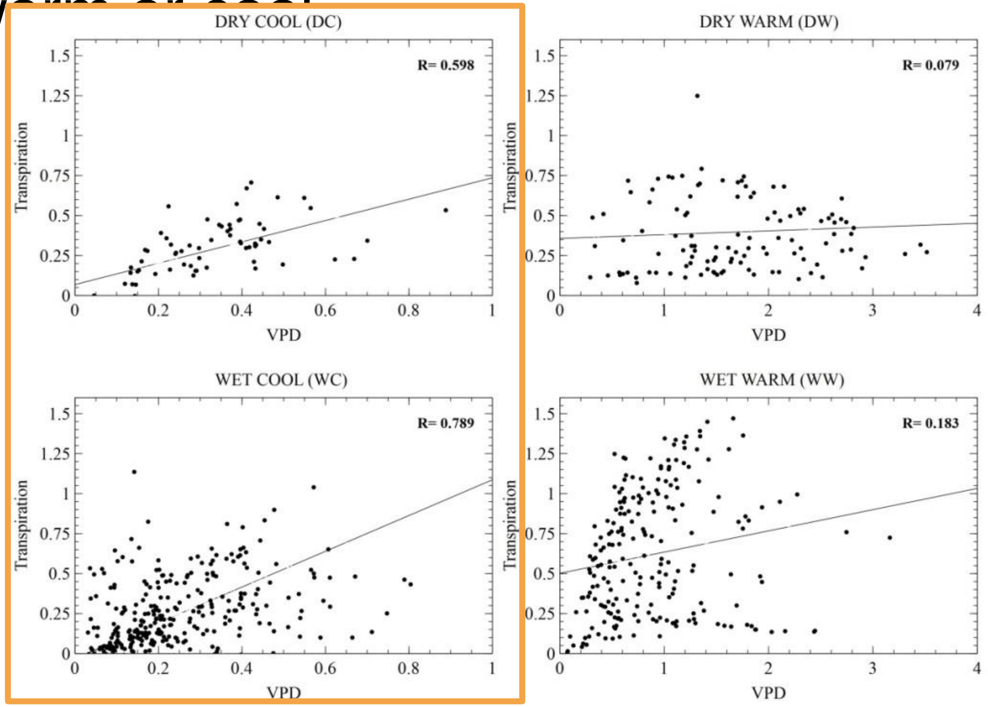
BIOME-BGC
RMSE= 0.05 E=0.517

Results and Discussion

- Comparison between the models
  - Goal: identify the main reasons and understand why the model performances were different
  - Classification: Four spells according to ppt and temperature
    - Dry Cool (DC)
    - Dry Warm (DW)
    - Wet Cool (WC)
    - Wet Warm (WW)
  - Criteria:
    - Dry: None of the previous consecutive 14 days registered daily ppt > 5mm
    - Warm: average daily temperature ≥ 13.2 ºC

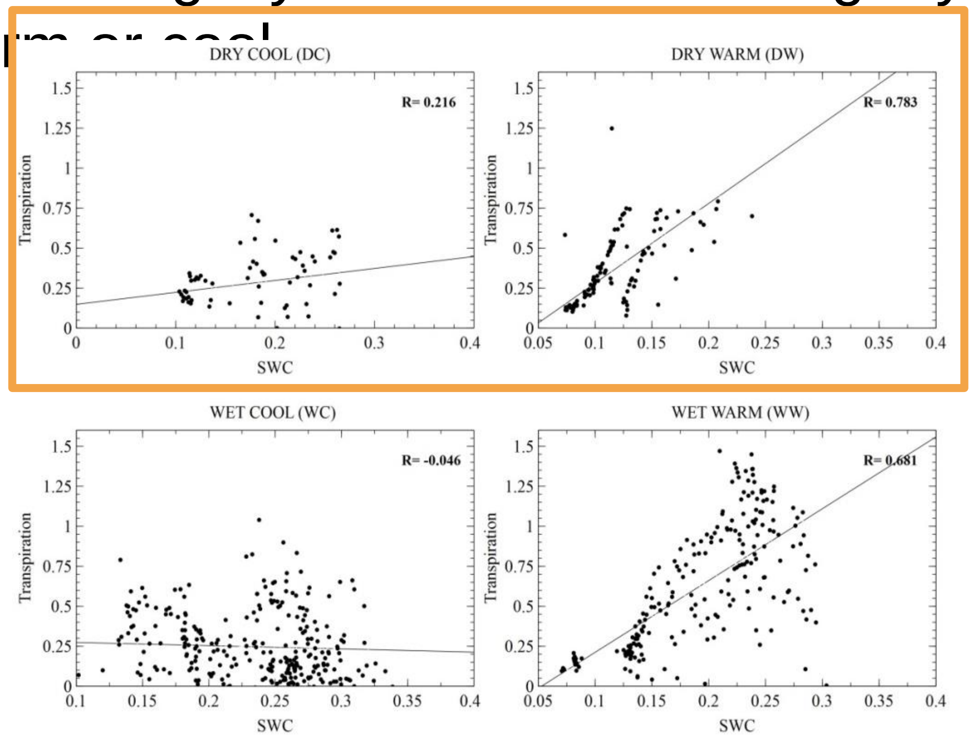
Results and Discussion

- Comparison between the models
  - LUE-Model: bad results during cool spells, either dry or wet
  - Biome-BGC: slightly less accurate during dry spells, either warm or cool
  - Why?



Results and Discussion

- Comparison between the models
  - LUE-Model: bad results during cool spells, either dry or wet
  - Biome-BGC: slightly less accurate during dry spells, either warm or cool!
  - Why?

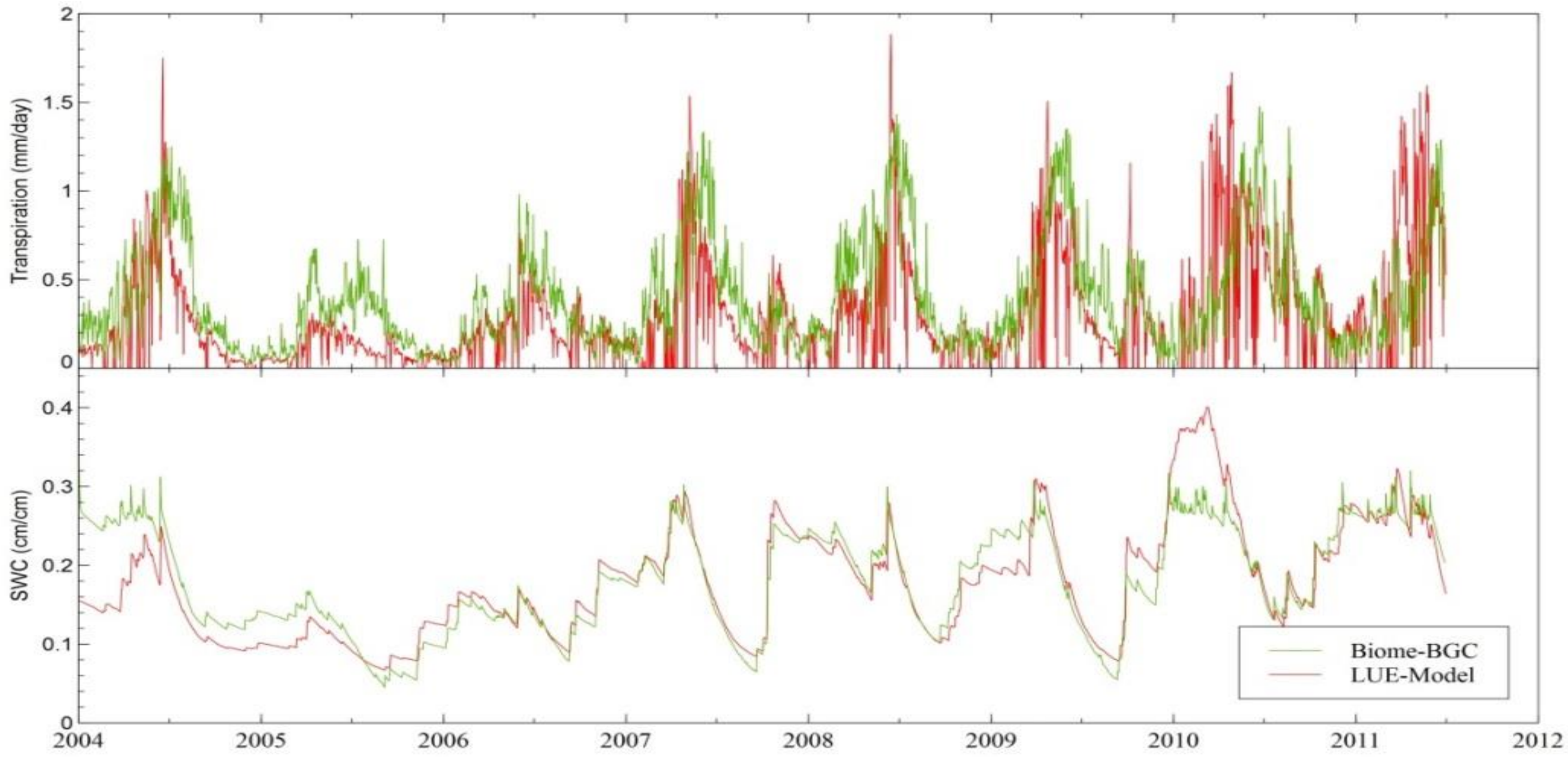


Results and Discussion

- Comparison between the models
  - LUE-Model: bad results during cool spells, either dry or wet
  - Biome-BGC: slightly less accurate during dry spells, either warm or cool
  - Why?
    - LUE-Model: ET_0 calculated using Hargreaves methodology which does not take into account the actual atmospheric evaporative demands in its formulation.
    - Biome-BGC: not designed for arid and semi-arid environments.

Results and Discussion

- Comparison between the models



BIOME	LUE

Results and Discussion

- Comparison between the models

LUE-MODEL

Flows	Dry year (2005)		Wet year (2010)	
	mm	%	mm	%
Ppt	188		739	
ET (EI+T+Es)	165.2	87.9	431.9	58.4
Excedence	16.3	8.7	326.9	44.2
Blue/Green	**0.098**		**0.757**	

BIOME-BGC

Flows	Dry year (2005)		Wet year (2010)	
	mm	%	mm	%
Ppt	188		739	
ET (EI+T+Es)	156.3	83.1	408.8	55.3
Excedence	16.3	8.7	330.1	44.7
Blue/Green	**0.104**		**0.807**	

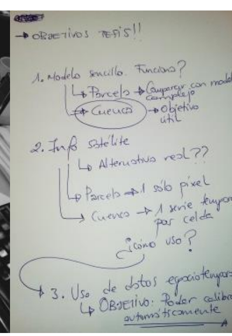
Conclusions at plot scale

- A parsimonious model with simple equations can achieve good results in general terms (at least, comparable with a physically-based model).
- LUE-Model's accuracy is worse when the transpiration is limited by the atmospheric demands but, the atmospheric demands should/could be taken into account in ET_0 calculation.
- Satellite data was a very useful source of information and its combination with the LUE-Model demonstrated to be an accurate tool capable to predict the role of the vegetation in the water cycle with no field data.
- Limitation: this application was at plot scale. **What about at catchment scale?**

4. At catchment scale: spatio-temporal modelling

'… then the plastic container appeared. A miracle! A revolution! What a relief this is for the exhausted African woman! What a transformation in her life! In Africa, water is a treasure.'

-Ryszard Kapuściński -



Research Goals

- To determine if the satellite data can be used as an alternative in ungauged or scarce data basins
- To propose a methodology in order to calibrate the model using the spatio-temporal information provided by satellite data

UNIVERSITAT POLITÈCNICA DE VALÈNCIA

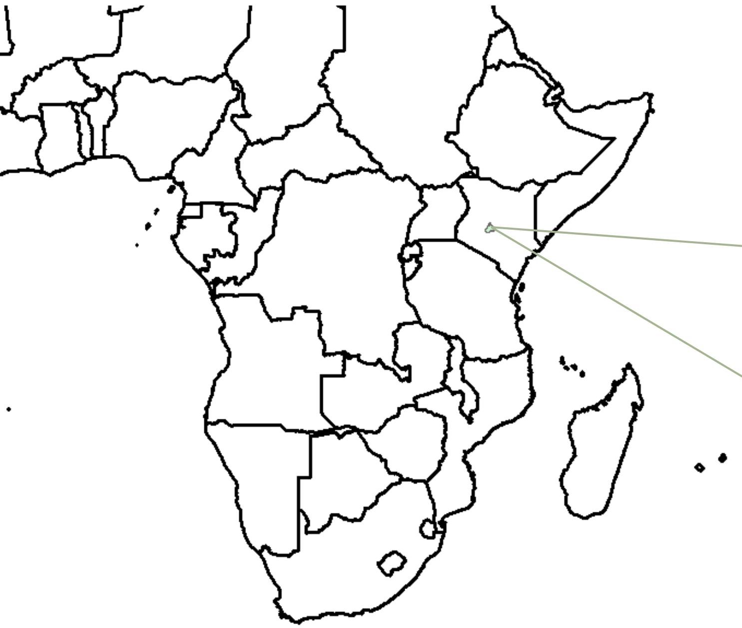
Università degli Studi Basilicata

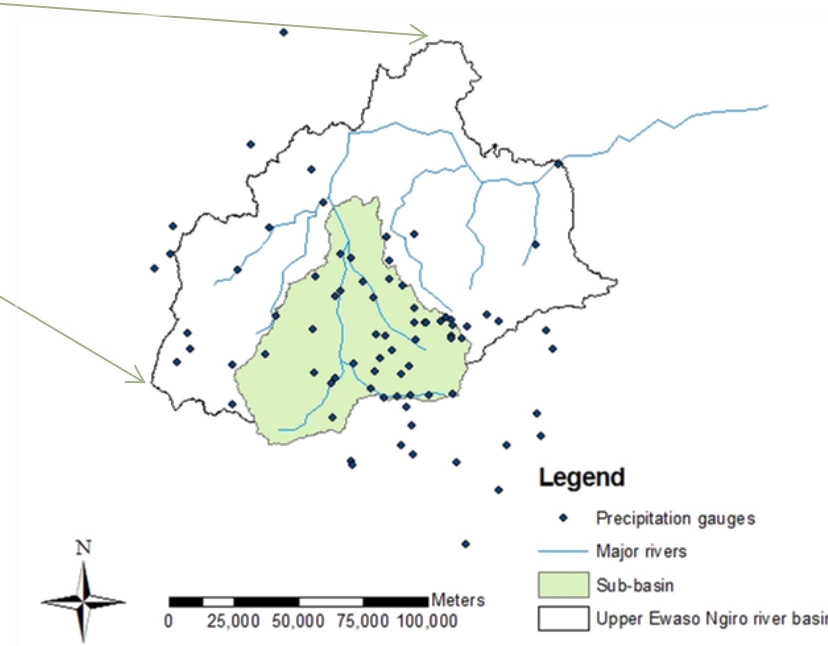
PRINCETON UNIVERSITY

Calibration of a parsimonious distributed ecohydrological daily model in a data scarce basin using exclusively the spatio-temporal variation of NDVI. Not published yet.

Study Area and Data

- Area: 4,605 km2
- Mostly water-controlled
- Sensitive to global change





- Rainfall: 1950 - 2003
- Temperature: 1950 - Present
- Observed discharge: 1980 - 2002
- NDVI: 2000 - Present

The TETIS-VEG Model

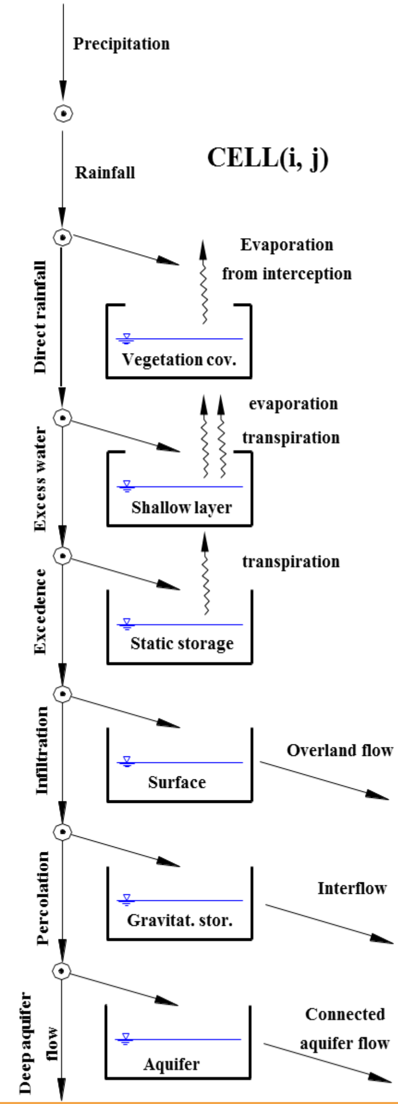
- Hydrological sub-model
  - Developed in the UPV since 1994
  - Distributed and conceptual (tank structure) model, with physically based parameters
  - Automatic calibration algorithm

- Vegetation sub-model
  - Based on the LUE-Model

$$T_1 = ET_o \cdot f_t \cdot \min(LAI,1) \cdot \beta_t(H_1) \cdot r_1$$

$$\frac{dB_l}{dt} = (LUE \cdot \varepsilon \cdot APAR - \mathrm{Re}) \cdot \varphi_l - \kappa_l \cdot B_l$$

$$LAI = B \cdot SLA \cdot f_t$$



Methodology

- Manual calibration
  - EOF
  - Confusion matrix
- Automatic calibration
  - Only using NDVI data
  - Spatio-temporal data → EOF
  - Period: year 2003
- Validation
  - NDVI data
  - Spatio-temporal data → EOF
  - Historical streamflow at the outlet point
  - Period: years 2000, 2001 and 2002

Empirical Orthogonal Functions (EOF)

Empirical Orthogonal Functions

- Methodology: EOF analysis
  - 1. Matrix configuration
  - II. Covariance matrix
  - III. To solve the eigenvalue problem
  - IV. Results:
    - Eigenvectors → They can be regarded as maps (patterns)
    - Eigenvalues → variance explained by each eigenvector
    - Loadings → Time evolution of each eigenvector

$$F = \begin{bmatrix} x_{11} & x_{12} & \cdots & x_{1p} \\ \vdots & \vdots & \ddots & \vdots \\ \vdots & \vdots & & \vdots \\ x_{n1} & x_{n2} & \cdots & x_{np} \end{bmatrix}$$

A map for time t_1

A time series for location x_1

$$R = F^T * F$$

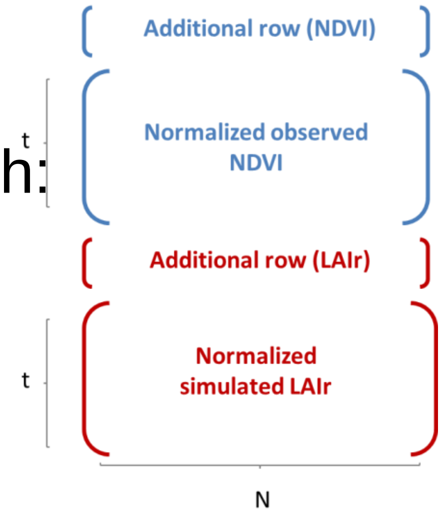
$$R * C = C * \Lambda$$

Manual Calibration

- Goals:
  - To test the applicability of the proposed model in the study basin
  - To obtain a first approximation for the parameters
- Procedure:
  - Manual adjustment of parameters values in 32 different cells
  - Objective function: Pearson correlation coefficient between LAI and NDVI
  - Selection of these 32 different cells:
    - Identification of the NDVI main patterns → EOF maps
    - Link between these main patterns and the available spatial maps (e.g. land use map) → confusion matrices

Automatic Calibration

- Goals:
  - To develop a strategy in order to incorporate spatio-temporal data during the model calibration
- Procedure:
  - Construction of one integral matrix with both:
    - Normalized observed NDVI
    - Normalized simulated LAI
    - Two additional rows



  - EOF methodology in order to obtain:
    - Common main patterns, loadings and portion of variance explained by each one
  - Objective function:

$$Error = \sum_{i=1}^{k} w_i * \sum_{j=1}^{t} \left| load\_sim_{i,j} - load\_obs_{i,j} \right|$$

Koch *et al.*, 2015

Automatic Calibration

- Goals:
  - To develop a strategy in order to incorporate spatio-temporal data during the model calibration
- Additional metrics:
  - Temporal Pearson correlation coefficient in each cell → maps
  - Spatial correlation coefficient distinguishing between:
    - Trees
    - Shrubs
    - Grasses

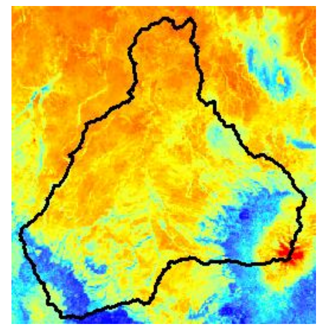
Validation

- Goals:
  - To test the accuracy of the model in a period not used during the model calibration
  - To check if the model is capable to simulate satisfactorily the streamflow being calibrated only with satellite data
- Period: year 2000, 2001 and 2002
- Data: Satellite NDVI and Historical streamflow
- Procedure:
  - EOF methodology as used during the model calibration
  - Temporal and spatial Pearson correlation coefficients
  - Comparison between the observed and simulated hydrograph

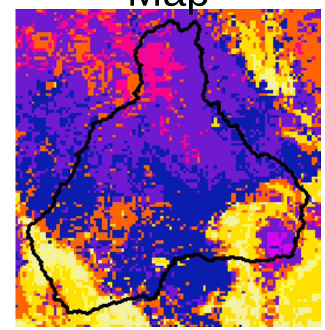
Results and Discussion

- Manual calibration:
  - EOF results:
    - EOF_1:
      - Explained variance: 61.5%
      - Linked with: Land Use map
    - EOF_2:
      - Explained variance: 10.5%
      - Linked with: No links
    - EOF_3:
      - Explained variance: 5.5%
      - Linked with: Soil maps

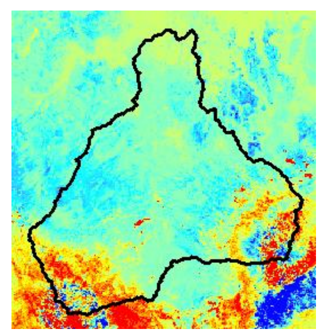
EOF_1



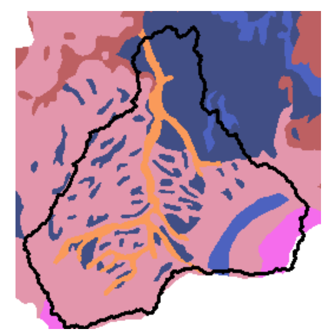
Land Use Map



EOF_3

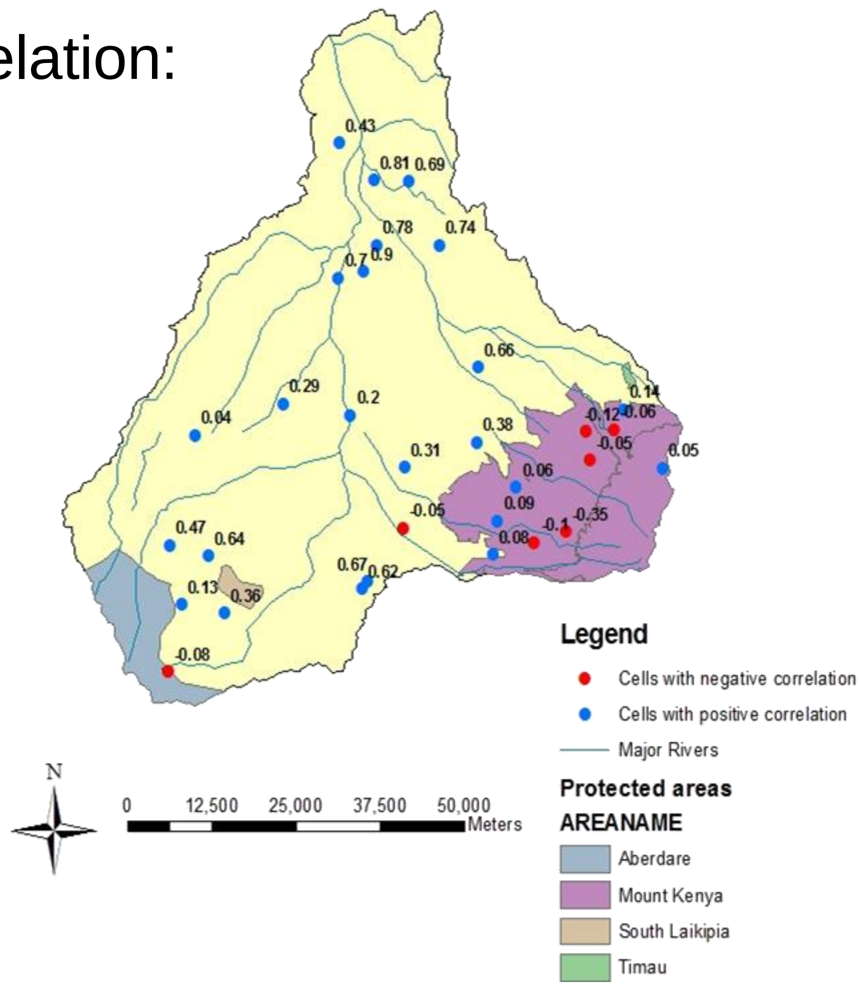


Soil Map



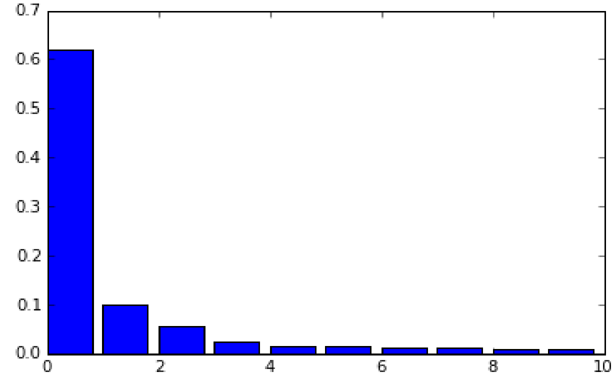
Results and Discussion

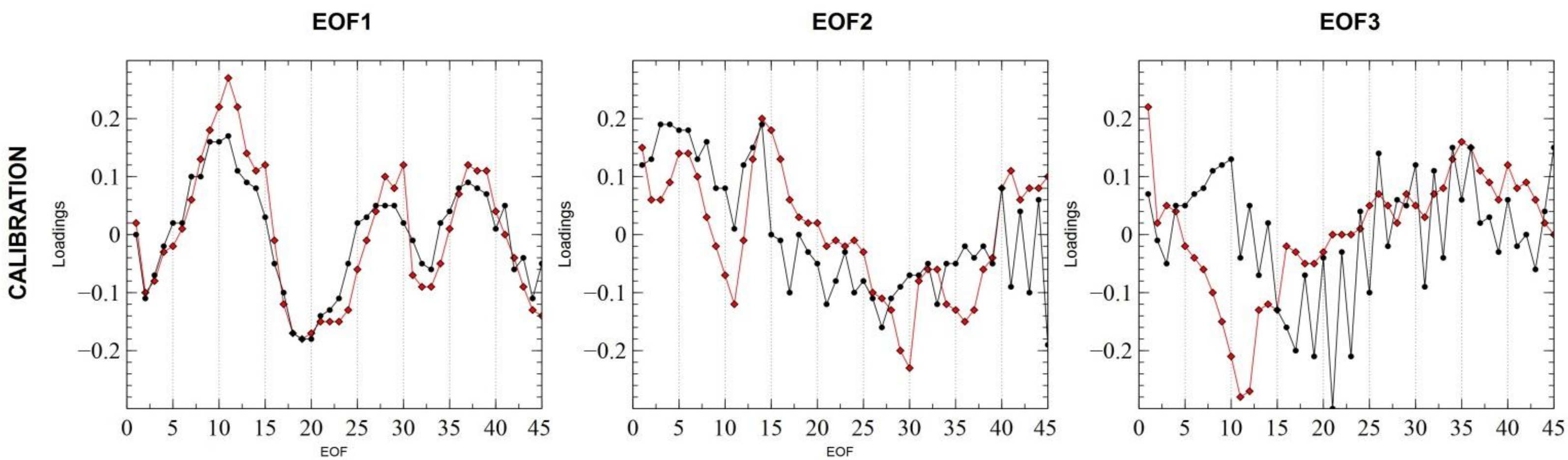
- Manual calibration:
  - Pearson correlation:



Results and discussion

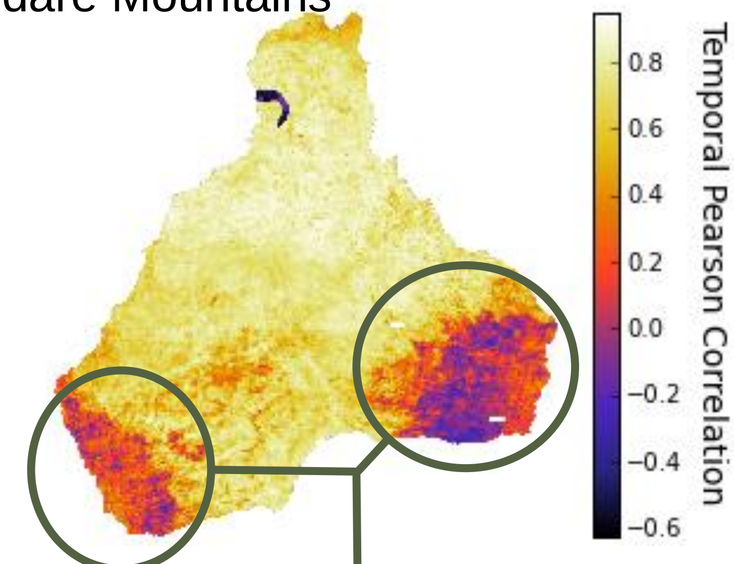
- Automatic calibration:
  - EOF results:
    - EOF_1 → two growing season
    - EOF_2 → Sensitive to initial conditions
    - EOF_3 → Sensitive to initial conditions





Results and Discussion

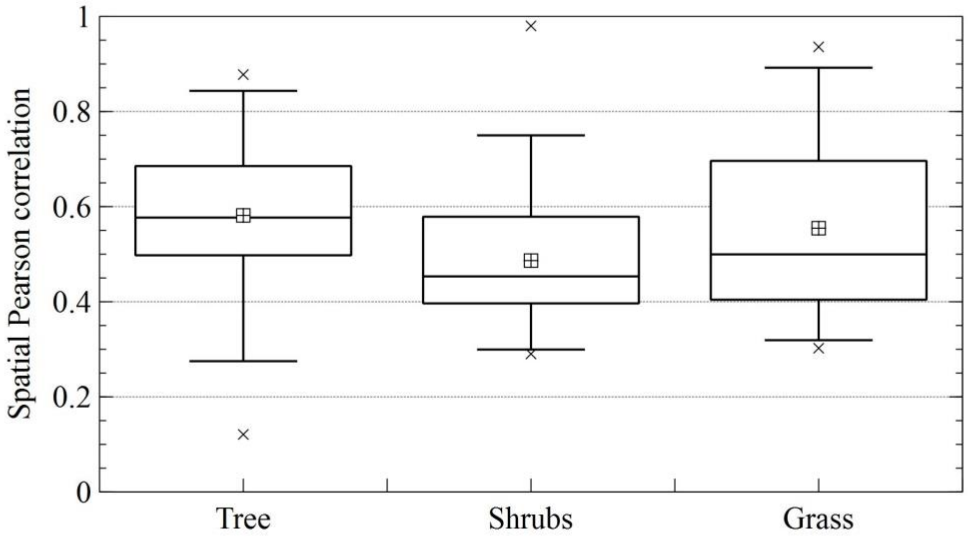
- Automatic calibration:
  - Temporal Pearson correlation:
    - Higher than 0.4 in almost the whole catchment
    - Negative in two small areas:
      - Mount Kenya
      - Aberdare Mountains



- Nutrient-limited areas
- High NDVI oscillations

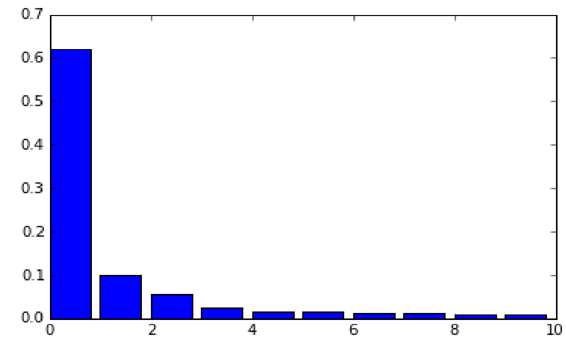
Results and Discussion

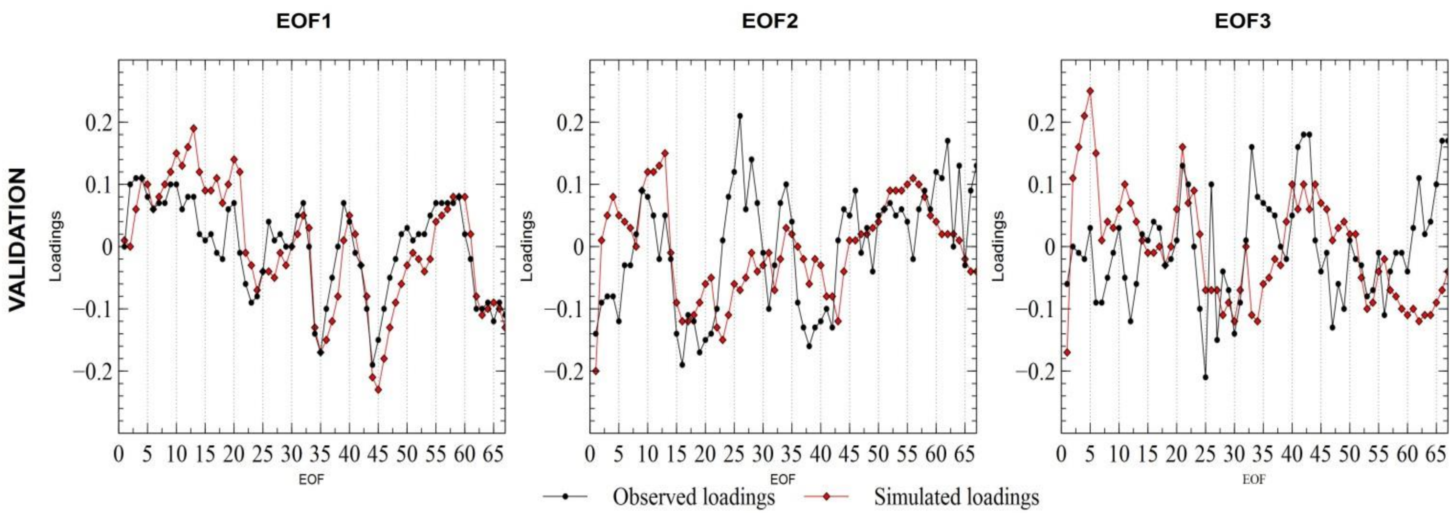
- Automatic calibration:
  - Spatial Pearson correlation:
    - The mean spatial correlations were higher than 0.45 for all main covers
    - Best results → cells classified as trees
    - Worst results → Shrubs



Results and Discussion

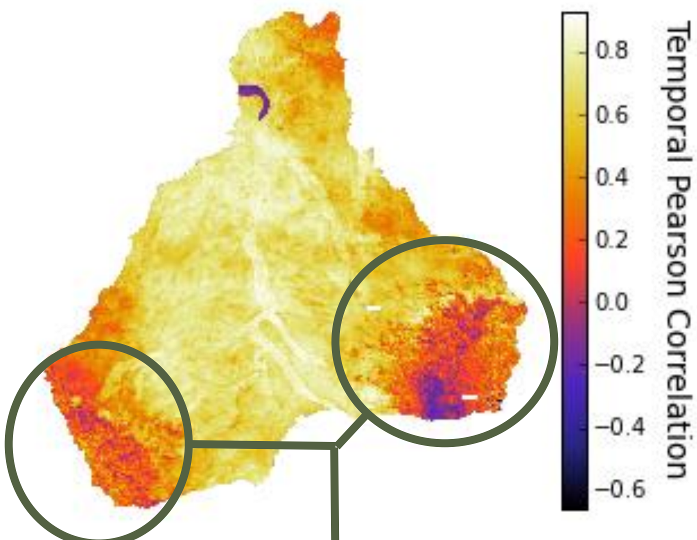
- Validation:
  - Satellite validation. EOF results
    - EOF_1: completely captured
    - EOF_2: worse results
    - EOF_3: worse results





Results and Discussion

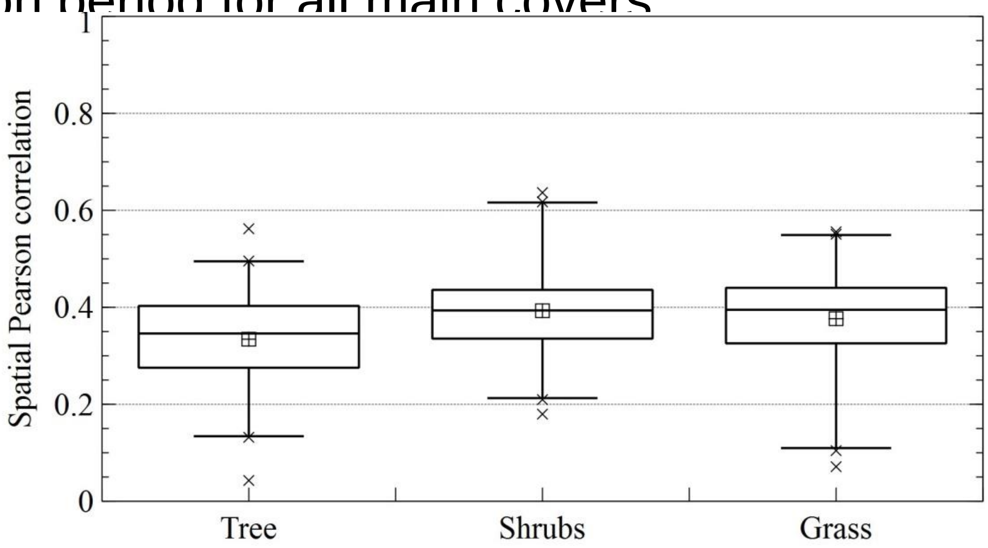
- Validation:
  - Temporal Pearson correlation:
    - Between 0.3 and 0.9 in more than 80% of the catchment
    - Negative in two small areas:
      - Mount Kenya
      - Aberdare Mountains



Expected bad results.

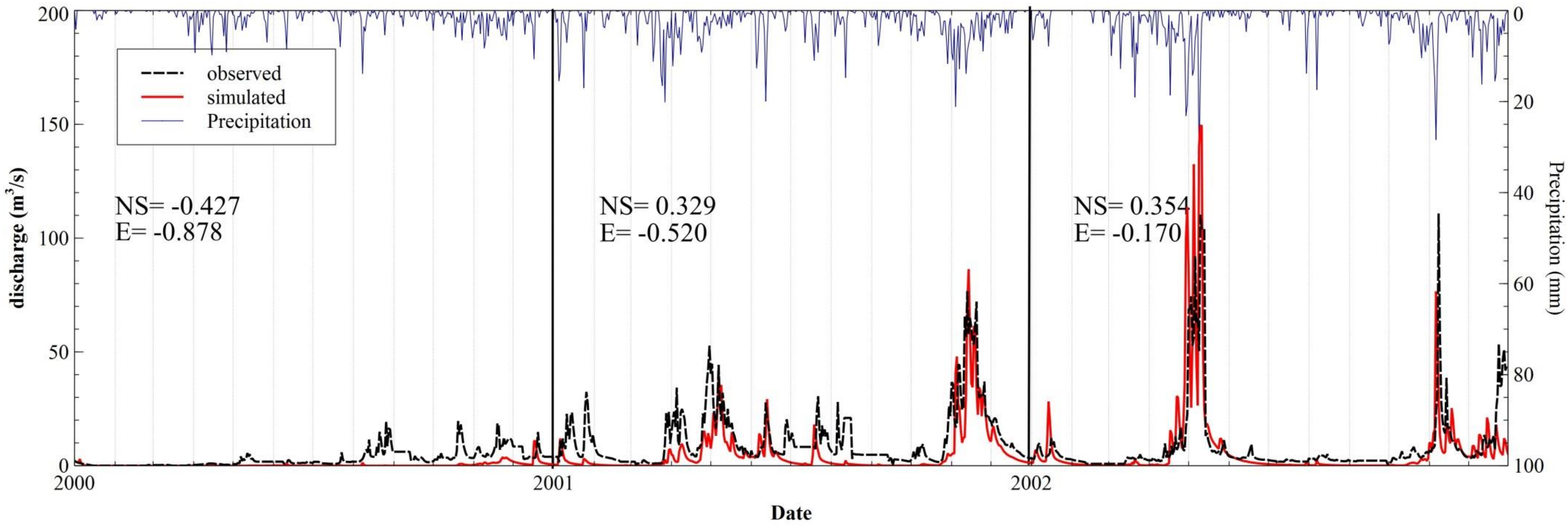
Results and Discussion

- Validation:
  - Spatial Pearson correlation:
    - The mean spatial correlations were higher than 0.35 for all main covers
    - The variance was narrower than the one obtained during the calibration period for all main covers



Results and Discussion

- Discharge Validation:
  - Comparison between streamflows at outlet point:
    - NS→ Nash and Sutcliffe efficiency index
    - E→ Volume error



Conclusions at catchment scale

- Simple models together with remote sensing data could be a potential alternative in places with no data or scarce available data.
- Some limitations:
  - Related to the model (nutrient-limited areas)
  - Related to the quality of the data
- The proposed methodology is an innovative option in order to include spatio-temporal data for conceptual models calibration
- Satellite data could be key in some places but it has to be properly used extracting both temporal and spatial information.
- In arid and semi-arid areas, the vegetation dynamics has been key in the water cycle's modelling (even for discharge simulation)

5. Conclusions

‘The important thing is not to stop questioning. Curiosity has its own reason for existing’

-Albert Einstein-

Concluding Remarks

- Vegetation plays a key role → Ecohydrological modelling
  - Problem: Data scarcity + high parametrical requirement
  - Potential solution: Parsimonious models + remote sensing data
- First step: at plot scale
  - Results comparable with the ones obtained by a physically based model
  - The model was able to reproduce vegetation and water dynamics
- Second step: at catchment scale
  - A methodology based on the use of EOF analysis was proposed and successfully applied.
  - A spatio-temporal state variable was used as target (LAI)
  - The model was capable to produce daily LAI maps and observed discharge
  - Promising results particularly in ungauged basins

Future Lines

- Use of other satellite products
- Development of new objective functions related to the EOF analysis → Selection rules:
  - Dominant variance rules
  - Time-history rules
  - Space-map rules
- Use of a multi-objective approach in which different sources of information could be mixed → Mixing different uncertainty levels
- Improvement of the TETIS-VEG model → nitrogen and carbon
- Analysis of the impact of the variability of hydrologic drivers on the vegetation patterns formation → seasonality, spatial patterns, extreme weather conditions…
- Parallel programming

Publications and Merits

- Publications in indexed journals
  - **Ruiz-Pérez G.,** Medici C., Latron J., Llorens P., Gallart F., Francés F., 2016a. Investigating the behaviour of a small Mediterranean catchment using three different hydrological models as hypotheses. *Hydrological Processes*, online version, DOI: 10.1002/hyp.10738
  - **Ruiz-Pérez, G**; González-Sanchis, M; del Campo, A.; Francés, F. Can a parsimonious model implemented with satellite data be used for modelling the vegetation dynamics and water cycle in water-controlled environments? *Ecological Modelling*, 324, 2016, 45-53.
  - **Ruiz-Pérez, G**; Koch, J.; Manfreda, S.; Caylor, K.; Francés, F. Calibration of a parsimonious distributed ecohydrological daily model in a data scarce basin using exclusively the spatio-temporal variation of NDVI. Under review
  - **Ruiz-Pérez, G**; González-Sanchis, M; del Campo, A.; Francés, F. Ecohydrological-based forest management in semiarid environment: adaptative sylviculture strategies and consequences. In preparation

Publications and Merits

- Conference contributions:
  - 16 International conferences (12 oral presentations + 4 posters)
  - 1 National conference (Oral presentation)
- Other merits
  - Teaching experience:
    - Course for the TETIS model's users
    - Course about Python applied in Science
    - Together with my supervisor → dynamic vegetation modelling part
  - Stay abroad experiences:
    - Università Degli Studi della Basilicata
      - Three months, Prof. Salvatore Manfreda, Hydrolab
    - Princeton University
      - Four months, Prof. Kelly Caylor, Civil and Environmental Engineering Dept.

Thanks for your attention!

ACKNOWLEDGEMENTS

AEMET, CEDEX, NASA Land Processes Distributed Active Archive Center (LP DAAC), USGS/Earth Resources Observation and Science (EROS) Center, Spanish Ministry of Economy and Competitiveness

PROJECTS

INTEGRA ECOTETIS (CGL2011-28776-C02-01), TETISMED (CGL2014-58127-C3-3-R) and EEBB-I-15-10262

RESEARCH GROUPS

HYDROLAB (Università degli studi della Basilicata), Princeton Ecohydrology Lab (Princeton University), ReForest (Universitat Politècnica de València), GIMHA (Universitat Politècnica de València)