modextreme

agriculture facing extreme climatic events

European Project n° 613817

2nd Annual Meeting

Yield forecast at EU level

Valentina Pagani, Tommaso Guarneri, Roberto Confalonieri

(Cassandra lab, University of Milan)

Tommy Klein, Pierluigi Calanca

(WBF-Agroscope)

Cassandra models for change

November 3-4, 2015

WP 5 Objectives

- Quantifying the reliability of crop yield forecasts produced using
  - Modelling solutions currently implemented in CGMS
  - Improved (for extreme events) solutions based on
    - process based models
    - (agro-)climatic indicators
- Define country- and crop-specific workflows for yield forecast

Introduction

- There's an **increasing demand** for crop yield forecasting system in both developed and developing countries
- **Early warnings** in case of **poor crop harvests** allow indeed governments and other stakeholders to
  - assure food imports
  - regulate agricultural markets

Introduction

- In the last years, a variety of forecasting systems were proposed, based on:
  - Field surveys
  - Crop models
  - Remote sensing
  - Integrated approaches
- However, in specific contexts where **yield fluctuations** are driven by **few main factors**, systems based on **agro-climatic indicators** proved their reliability (e.g. Balaghi et al., 2012)

Introduction

- Objectives of Task 5.2:
  - Evaluate the potential of forecasting systems based on **temperature** and **drought indicators** for **winter and summer crops** and for **grasslands** in Europe
  - Identify **combinations crop × country** where systems based on agro-climatic indicators present sufficient reliability

Methods

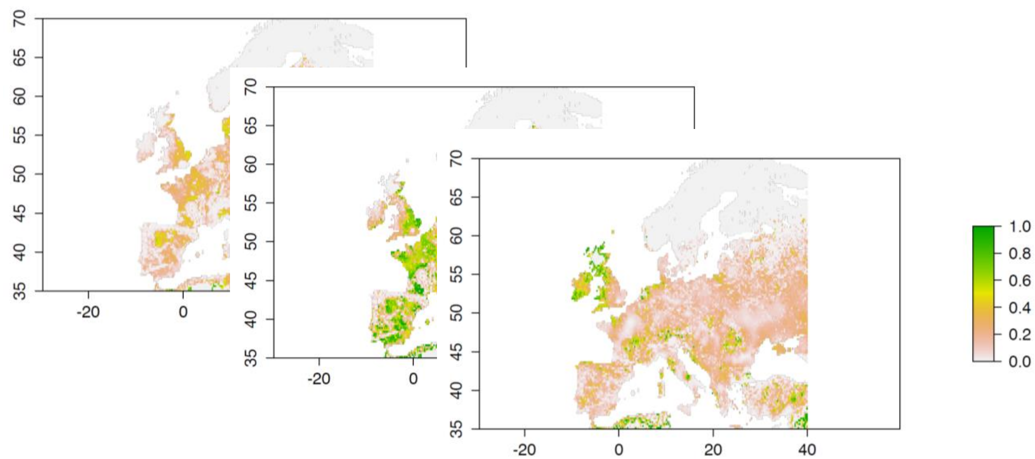
- Agro-climatic indicators (WP1):
  - Heat
    - Tmaxcr (# days with Tmax > threshold)
  - Frost
    - Tmincr (# days with Tmin < threshold)
  - Drought
    - ARIDmean (mean Agricultural Reference Index [ARID; Woli et al., 2012] for drought) $ARID = 1 - \frac{T_a}{T_p}$
    - ARIDcr (# days with ARID higher than a threshold)
    - Fu (Fu drought index; Fu, 1981) $FU = -\frac{\sum P}{\sum ET_0} + \left[1 + \left(\frac{\sum P}{\sum ET_0}\right)^w\right]^{(1/w)}$

Methods

- Crops (CGMS):
  - Maize
  - Rice
  - Sunflower
  - Potatoes
  - Sugar beet
  - Wheat
  - Barley
  - Rye
  - Triticale
  - Rapeseed
  - Grassland
- Countries:
  - Whole Europe
- Time series:
  - 1990-2013 (3 forecasting events per season)

Methods

- Input data:
  - Weather data: E-OBS database (Haylock et al., 2008)
  - Official yield statistics: FAOSTAT database
  - Flowering-maturity: according to the MARS crop calendars
  - Crop masks: Monfreda et al. (2008)
    - cereals
    - other crops
    - grasslands

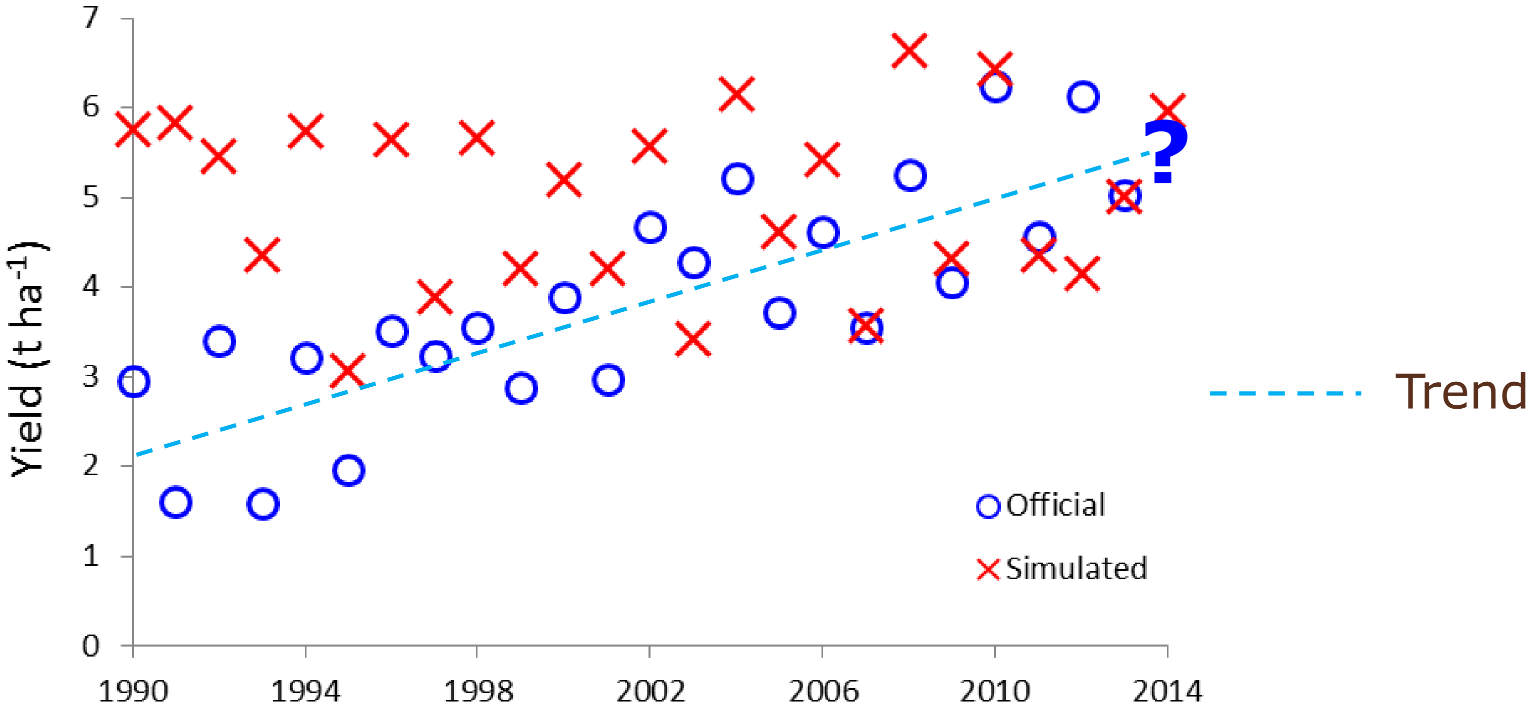


Methods

- Spatial level for analysis:
  - NUTS0
- Post processing (JRC-MARS operational chain):
  - Something is **not explicitly considered** by the system
  - This generates **uncertainty** and **specific system behaviours** under specific conditions (assumption)
  - The **same behaviour** will be reproduced **under similar conditions**

Methods

- De-trending historical yield statistics



Methods

- De-trending historical yield statistics
  - Are there discontinuities in the series of official yields?

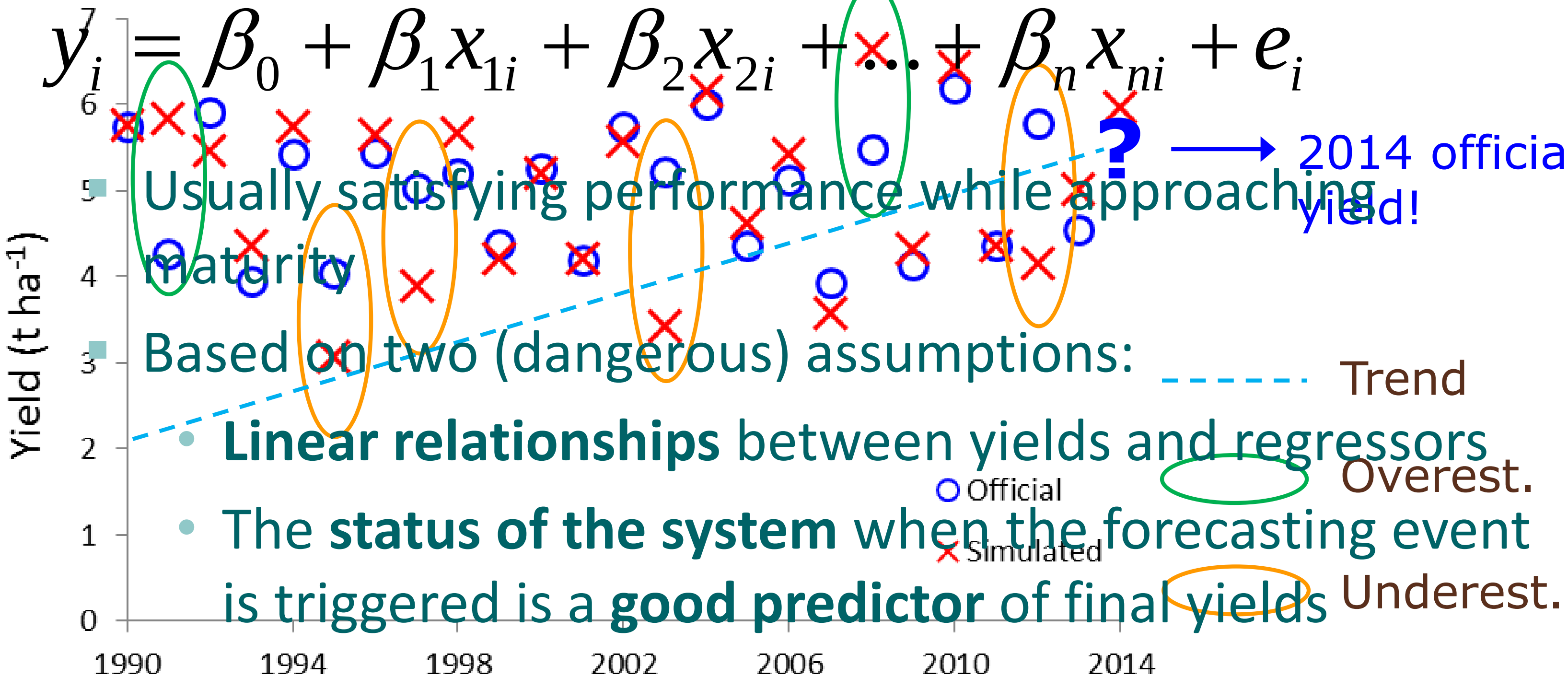
    (e.g. the dissolution of the Soviet Union let to the abandonment of irrigation systems in many areas)
  - No, but there’s a trend
    - Technological (e.g. new varieties, machines)
    - Due to something the system should capture (e.g. climate change)
  - The **analyst** should understand **case-by-case** how the trend should be managed
  - If due to something that cannot be reproduced by the system, it should be eliminated

Methods

- Multiple linear regression

$$y_i = \beta_0 + \beta_1 x_{1i} + \beta_2 x_{2i} + \ldots + \beta_n x_{ni} + e_i$$

- Usually satisfying performance while approaching maturity
- Based on two (dangerous) assumptions:
  - **Linear relationships** between yields and regressors
  - The **status of the system** when the forecasting event is triggered is a **good predictor** of final yields



2014 official yield = *f*(official[1990-2013], simulated[1990-2013, 2014])

Methods

- Analysis performed in **two steps**:
  - **Screening** all the combinations crop × country × forecasting moment
  - **Percentage of variance** in yields explained by:
    - Trend
    - Regression model
  - **Complete cross-validation** where
    - Regression model was able to explain "*in a reasonable way*" the **inter-annual variability** in yields
  - Completely unsupervised

Results

- Sample results for the regression models
- General considerations:
  - **Forecasting capability** of agro-climatic indicators was **higher in central and southern Europe**
  - **Exceptions** were grasslands in Germany and potatoes in Poland
  - **Indicators selected** by the **step-wise** procedure were mostly related with **warm/dry conditions**
    - Often Mediterranean countries
    - Dry summers and mild/rainy winters

Maize

Country	Period	Trend	27th ten-day period Model	27th ten-day period Total	24th ten-day period Model	24th ten-day period Total	20th ten-day period Model	20th ten-day period Total
Austria	1990-2013	0.72	0.18	0.9	0.18	0.9	0.05	0.77
Belgium	2000-2013	0	0.74	0.74	0.76	0.76	0.7	0.7
Bulgaria	1990-2013	0	0.49	0.49	0.58	0.58	0.49	0.49
Croatia	**1992-2013**	**0.25**	0.6	0.85	0.61	0.86	**0.64**	**0.89**
Czech R.	1993-2013	0.49	0.22	0.71	0.23	0.72	0.21	0.7
France	1990-2013	0.51	0.37	0.88	0.38	0.89	0.16	0.67
Greece	1990-2013	0.5	0.17	0.67	0.15	0.65	0.13	0.63
Germany	1990-2013	0.71	0.21	0.92	0.2	0.91	0.13	0.84
Hungary	1990-2013	0	0.57	0.57	0.69	0.69	0.41	0.41
Italy	**1990-2013**	**0.38**	**0.43**	**0.81**	**0.44**	**0.82**	0.44	0.82
Netherlands	1992-2013	0.54	0.07	0.61	0.06	0.6	0.08	0.62
Poland	1990-2013	0.48	0.32	0.8	0.34	0.82	0.39	0.87
Portugal	1990-2013	0.78	0.07	0.85	0.06	0.84	0.05	0.83
Romania	1990-2013	0	0.53	0.53	0.67	0.67	0.45	0.45
Slovakia	1993-2013	0	0.44	0.44	0.58	0.58	0.6	0.6
Slovenia	1992-2013	0.46	0.38	0.84	0.38	0.84	0.2	0.66
Spain	1990-2013	0.88	0.02	0.9	0.03	0.91	0.01	0.89

Wheat

Country	Period	Trend	15th 10-day period Model	15th 10-day period Total	13th 10-day period Model	13th 10-day period Total	11th 10-day period Model	11th 10-day period Total
Austria								
Belgium								
Bulgaria								
Croatia								
Denmark								
Czech R.								
Estonia								
Finland								
France								
Greece								
Germany								
Hungary								
Ireland								
Italy								
Latvia								
Lithuania								
Netherlands								
Poland								
Portugal								
Romania	1990-2013	0	0.34	0.34	0.27	0.27	0.09	0.09
Slovakia	1993-2013	0	0.56	0.56	0.44	0.44	0.05	0.05
Slovenia	**1992-2013**	**0.24**	**0.57**	**0.81**	0.24	0.48	0.12	0.36
Spain	**1990-2013**	**0.31**	**0.58**	**0.89**	0.5	0.81	0.37	0.68
Sweden	1990-2013	0	0.18	0.18	0.21	0.21	0.29	0.29
United Kingdom	1990-2013	0.38	0.05	0.43	0.05	0.43	0.01	0.39

Results

- Results for the **step 2** (**complete cross validation** on combinations crop × country **screened**)
  - In general:
  - **Good results** achieved for each of the 11 crops **in one or two countries**
  - **Exception was rice**, for which no satisfactory results were achieved
    - Flooded
    - Macrothermal crop originated in the tropics

    No conditions characterized by **heat and drought** (differently from other crops)

Results

- Results for the **step 2** (**complete cross validation** on combinations crop × country **screened**)
  - **ARIDcr** (# days with ARID higher than a threshold) and **Tmaxcr** (# days with Tmax > threshold) were the indicators **selected more frequently** as regressors
  - **No relevant impact** of **low temperatures**: Tmincr (# days with Tmin < threshold) was rarely selected
  - In most cases **2 or 3 indicators were selected**: the fourth one usually did not increase the forecasting capability

    (low risk of overfitting!)

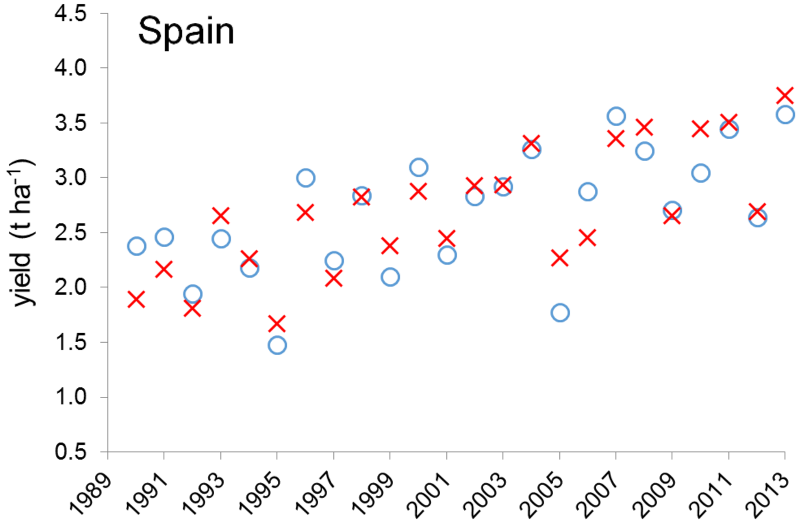
	Crop	Country	Regression model	RRMSE	EF	CRM	R^2
Winter crops	Wheat	Spain	$ARID_{mean}$, $ARID_{cr}$, T_{maxcr}	0.24	0.80	0.00	0.82
		Slovenia	$ARID_{mean}$, $ARID_{cr}$, Fu, T_{mincr}	6.19	0.66	-0.01	0.67
	Barley	Italy	$ARID_{cr}$, Fu, T_{maxcr}, T_{mincr}	3.42	0.48	0.00	0.58
		Spain	$ARID_{cr}$, T_{maxcr}	9.84	0.84	0.01	0.85
	Rye	Spain	$ARID_{cr}$, T_{maxcr}, T_{mincr}	0.21	0.76	0.00	0.76
	Triticale	Romania	$ARID_{cr}$, Fu, T_{maxcr}	15.96	0.07	0.00	0.33
	Rapeseed	Romania	$ARID_{mean}$, T_{maxcr}	20.52	0.65	-0.03	0.66
Summer crops	Maize	Italy	$ARID_{mean}$, T_{maxcr}, T_{mincr}	4.86	0.62	0.00	0.65
		Croatia	$ARID_{mean}$, T_{maxcr}	11.35	0.68	0.00	0.71
	Sunflower	Bulgaria	$ARID_{mean}$, Fu, T_{maxcr}	12.97	0.80	0.01	0.81
		Italy	T_{maxcr}, T_{mincr}	6.87	0.43	0.00	0.45
	Potato	Romania	$ARID_{cr}$, Fu, T_{maxcr}	9.82	0.52	-0.01	0.53
		Poland	$ARID_{mean}$, $ARID_{cr}$, T_{maxcr}	8.04	0.66	0.00	0.68
	Sugar beet	Italy	$ARID_{cr}$, T_{mincr}	7.02	0.69	0.00	0.69
		Croatia	$ARID_{cr}$, Fu	10.78	0.73	0.00	0.74
	Mown Grasslands	Germany	$ARID_{mean}$, $ARID_{cr}$, T_{mincr}	4.49	0.47	0.00	0.59

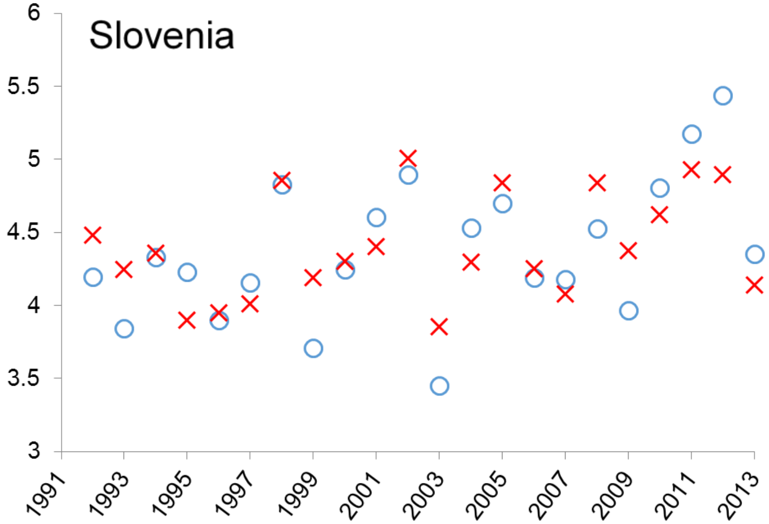
Results

- Results for the **step 2** (**complete cross validation** on combinations crop × country **screened**)
  - **Best results** achieved for **wheat**, **barley** and **rye** in **Spain** (82%, 85% and 76% of yields variability explained)
  - Good results also for **sunflower in Bulgaria** and **sugar beet in Croatia** (81% and 74%), the latter mainly because of technological innovations
  - **Poor results** for **triticale in Romania**, mainly because of the **short historical series** (regression model not robust)

Results

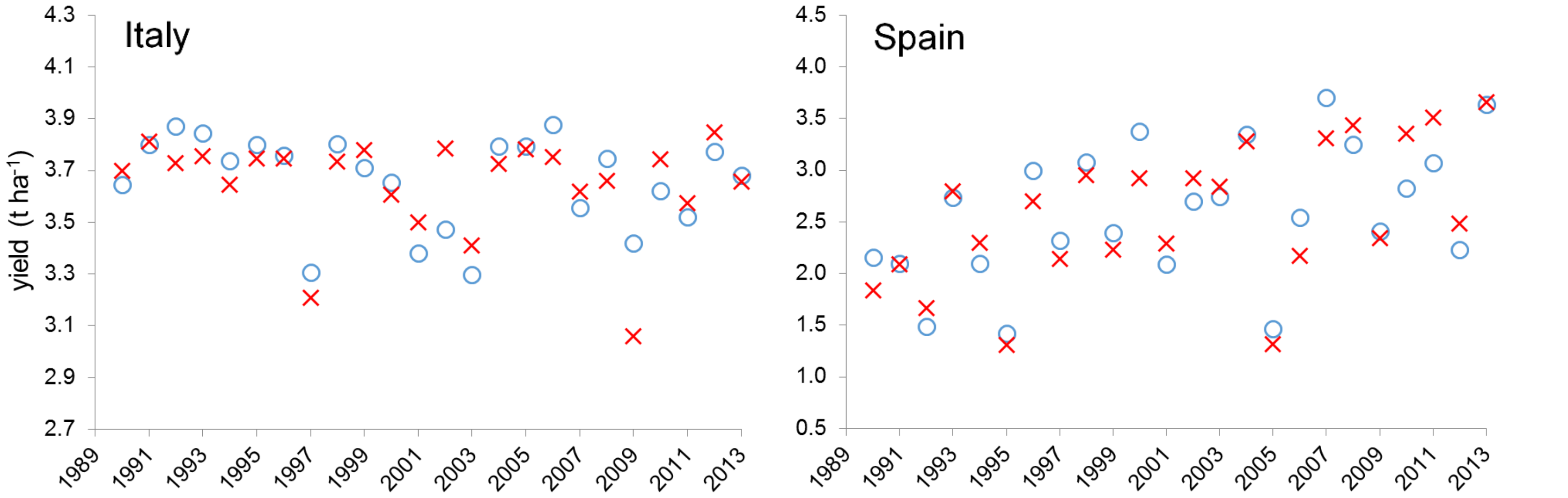
- Results for the **step 2** (**complete cross validation** on combinations crop × country **screened**)
  - Examples for:
    - Wheat





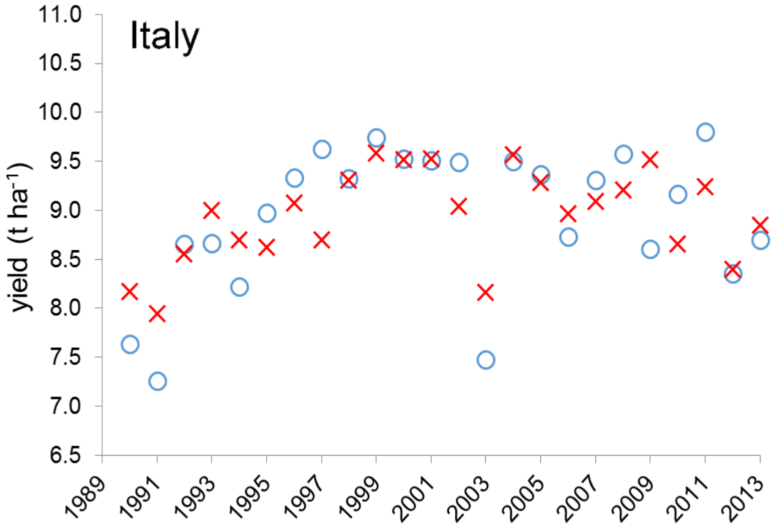
Results

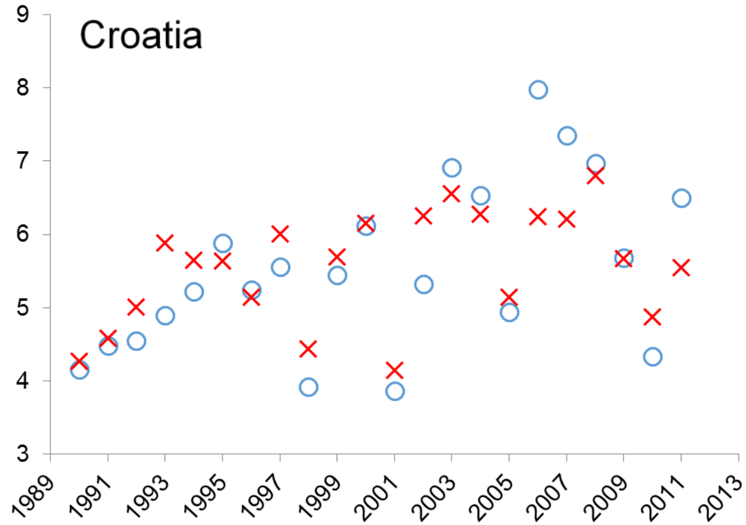
- Results for the **step 2** (**complete cross validation** on combinations crop × country **screened**)
  - Examples for:
    - Barley



Results

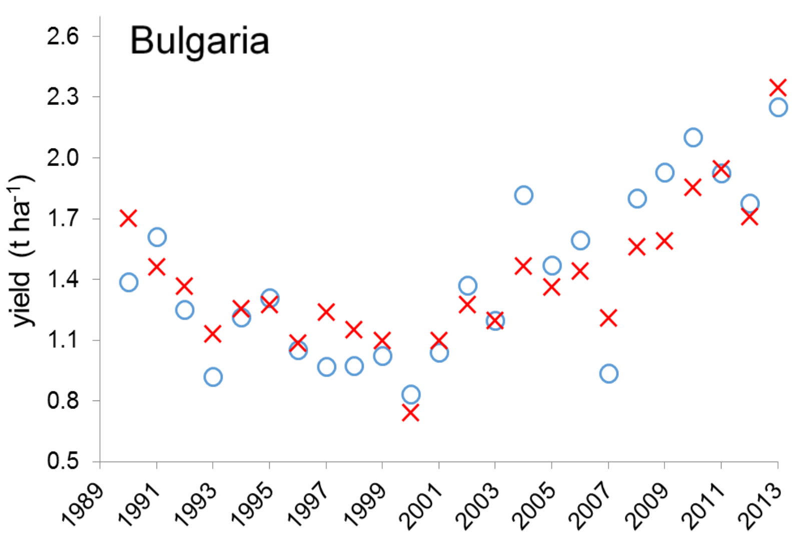
- Results for the **step 2** (**complete cross validation** on combinations crop × country **screened**)
  - Examples for:
    - Maize

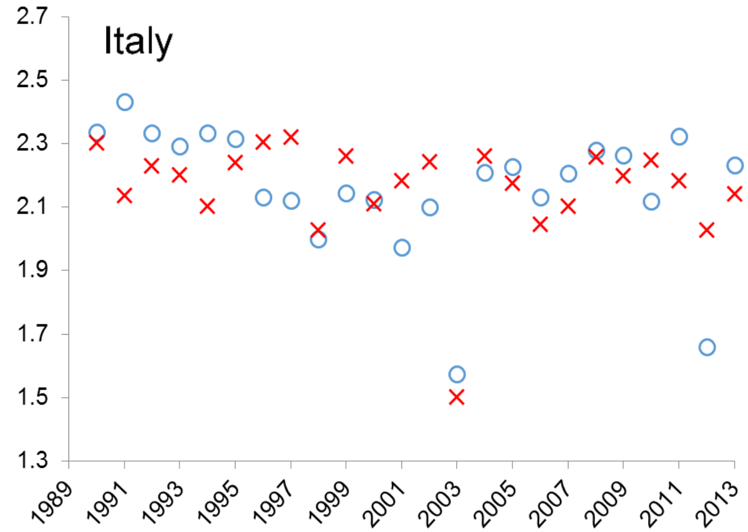




Results

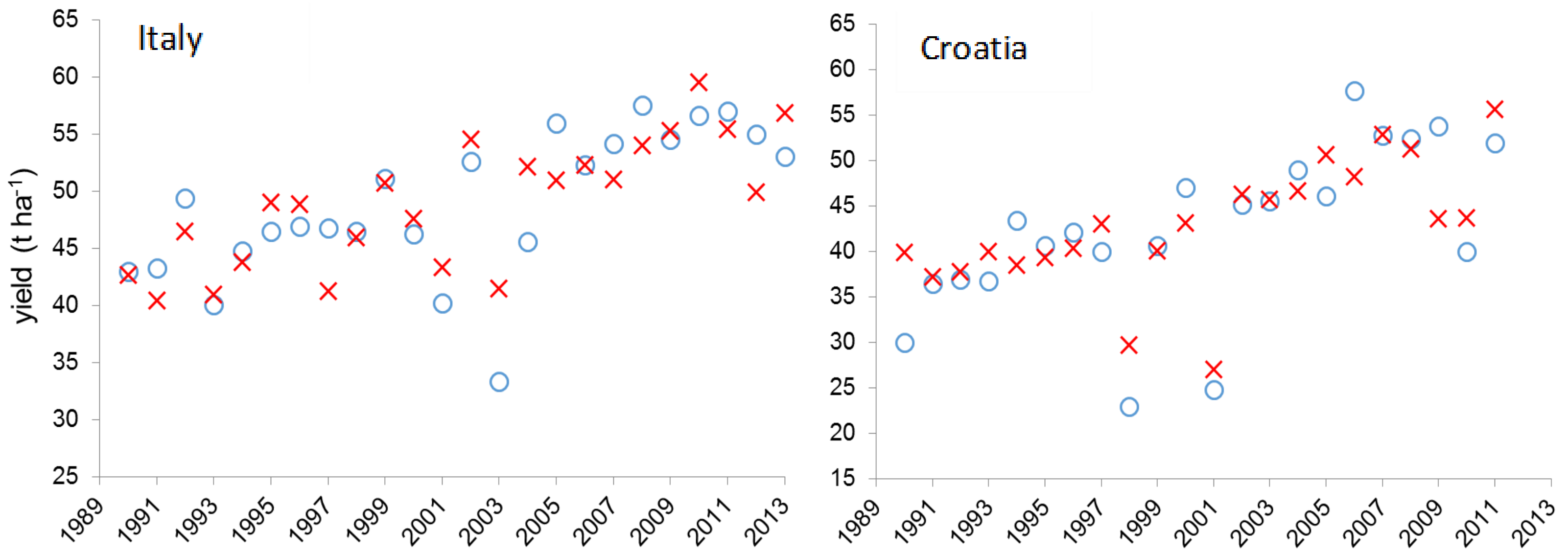
- Results for the **step 2** (**complete cross validation** on combinations crop × country **screened**)
  - Examples for:
    - Sunflower





Results

- Results for the **step 2 (complete cross validation** on combinations crop × country **screened)**
  - Examples for:
    - Sugar beet



Conclusions

- **Good results** in countries where **production is mainly driven** by the occurrence of **extreme conditions**, especially **drought** and **heat**
- It will be interesting to test **hybrid approaches** (dynamic models + agro-climatic indicators)
- Ideas for the future:
  - Perform the **test at NUTS2** (or NUTS3…) level
    (**Extreme events can be local**)
  - **Crop masks changing** during the time series (?)

Acknowledgement

"The research leading to these results has received funding from the European Community's Seventh Framework Programme – FP7 (KBBE.2013.1.4-09) under Grant Agreement No. 613817, 2013-2016"

Chinese Academy of Agricultural Sciences

Dmi Vejr, klima og hav

INRA Science & Impact

FAO Fiat Panis

Softeco Sismat information technology

UHMI

Embrapa

Universidad de Córdoba

Δημοκρίτειο Πανεπιστήμιο Θράκης

crea Consiglio per la ricerca in agricoltura e l'analisi dell'economia agraria

Universitas Studiorum Mediolanensis

Federal Department of Economic Affairs, Education and Research EAER
Agroscope

Universidade Técnica de Lisboa

INTA

University of Pretoria

INRA Transfert

UEA University of East Anglia