A machine learning approach to the assessment of the vulnerability of *Posidonia oceanica* meadows

Catucci E., Scardi M.

University of Rome 'Tor Vergata', Italy
CoNISMa

catucci.elena@gmail.com – mscardi@mclink.it

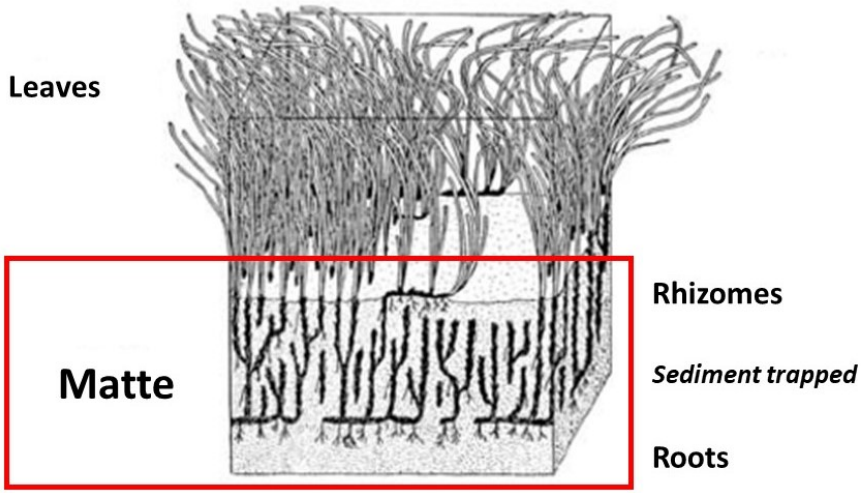
10th International Conference on Ecological Informatics

September, 24-28, Jena – Germany

Posidonia oceanica

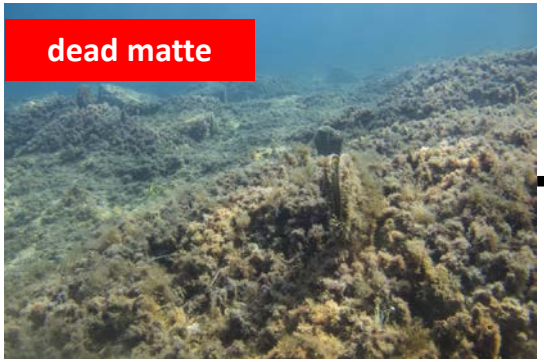
- endemic Mediterranean seagrass
- habitat-forming species
- one of the most productive ecosystems on Earth
- biodiversity hotspot
- …

Posidonia oceanica: matte



...perturbations

Living meadows



Regressed meadows

Dead matte *to identify areas in which the environmental conditions no longer allow the presence of living meadows*

Crucial in habitat suitability modeling!

State of the art on P. oceanica: *Italian vs. Mediterranean*

Italian seas

- ✓ spatial distribution entirely known
- ✓ meadows status: living meadows *vs.* regressed meadows
- ✓ *P. oceanica* regression: ~25%

Mediterranean basin

- ✓ *P. oceanica* distribution known in the North and West of the basin
- x No data for the South and East
- ✓ *P. oceanica* regression: ~15%

Scardi et al., 2013

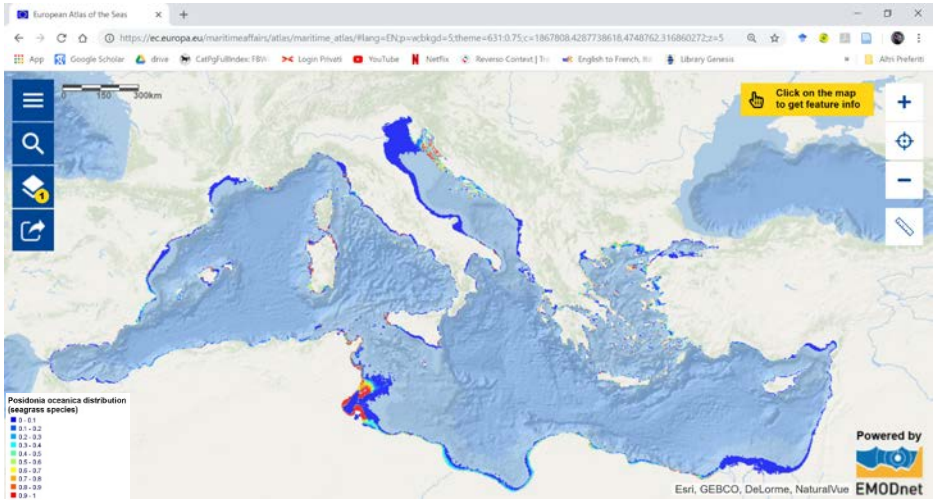
From Scardi et al., 2013
I analyzed…

…what I did

SDM using RF to define the potential spatial distribution

RF output not independent to the meadows status

HSM *aimed at assessing the vulnerability of meadows in the Italian seas*



……

Italian seas

Aim:

to assess habitat suitability
for P. oceanica *meadows*

↓

entire data set
used for
the RF training

↓

data preprocessing*:*
conservative
rasterization

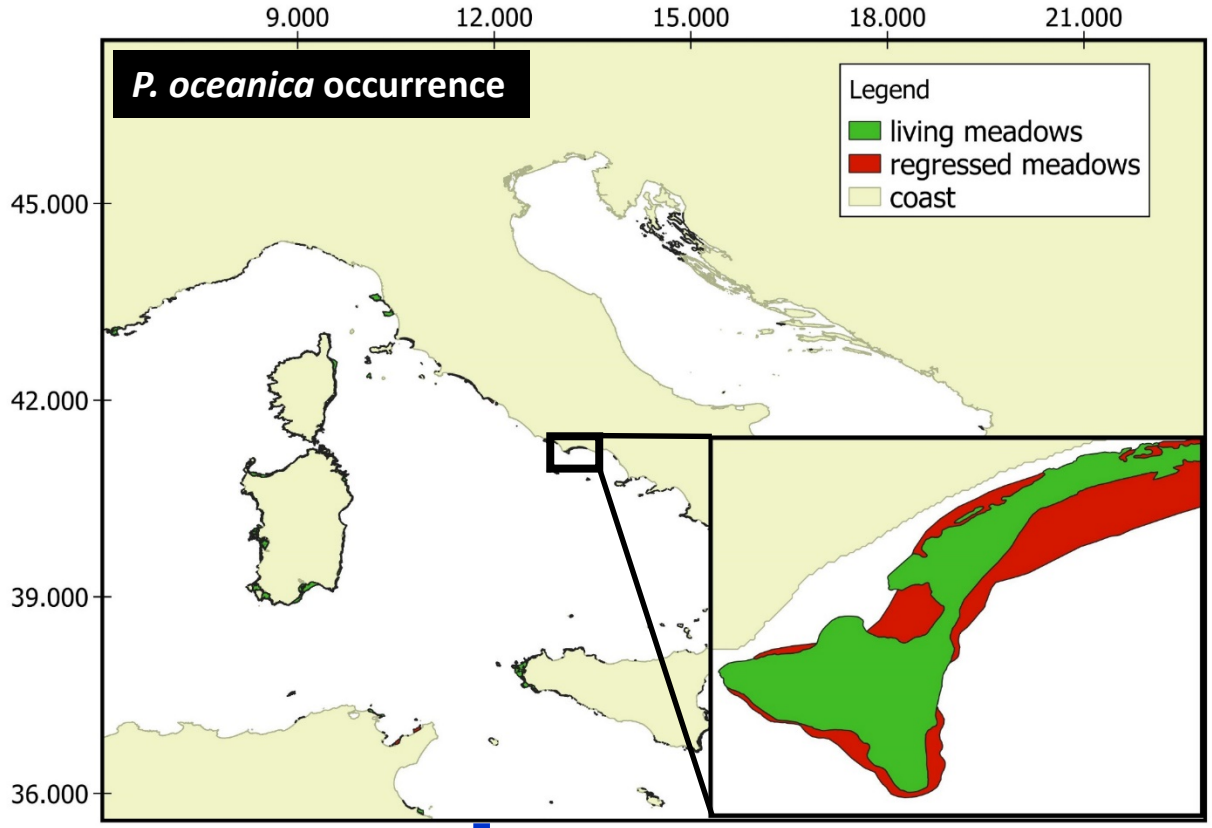
Mediterranean basin

Aim:

to predict spatial distribution
of P. oceanica *meadows*

↓

data set splitted into
training set
and test set

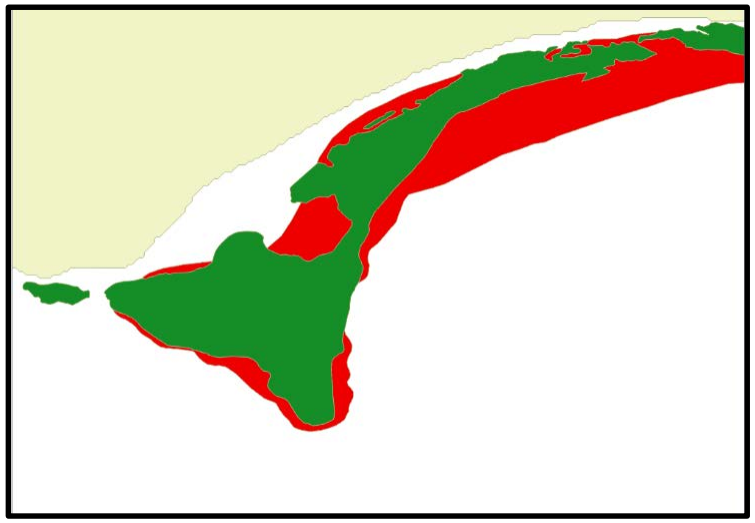


SHAPEFILE *vs.* **RASTER**

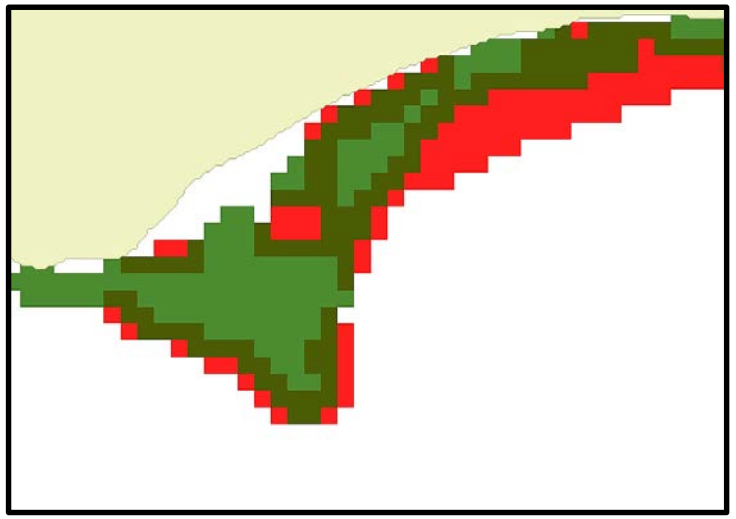
PREDICTORS
Artisanal fishing
Aspect of the seafloor
Bathymetry
Bottom salinity
Bottom temperature
Bottom type
Calcite concentration
Chlorophyll a concentration (a. r.)
Chlorophyll a concentration (mean)
Climate change (sea surface temperature)
Climate change (UV)
Diffuse attenuation coefficient
Dissolved oxygen concentration
Distance to 200 m isobath
Distance to coast
Distance to ports
Distance to river mouths
Euphotic depth
Human impact to marine ecosystems
Nitrate concentration
Nutrient input (fertilizers)
Ocean acidification
pH
Phosphate concentration
Photosynthetically available radiation
Pollutants (inorganic)
Pollutants (organic)
Pollution (ocean-based)
Population pressure
Salinity
Sea surface temperature (ann. range)
Sea surface temperature (mean)
Shipping intensity
Silicate concentration
Slope of the seafloor

Data preprocessessing

Shapefile



Raster



Conservative rasterization → **3 meadows status**

any pixel that intersect a
P. oceanica *polygon*
is turned on

→ Living meadows
→ Mixed conditions
→ Regressed meadows

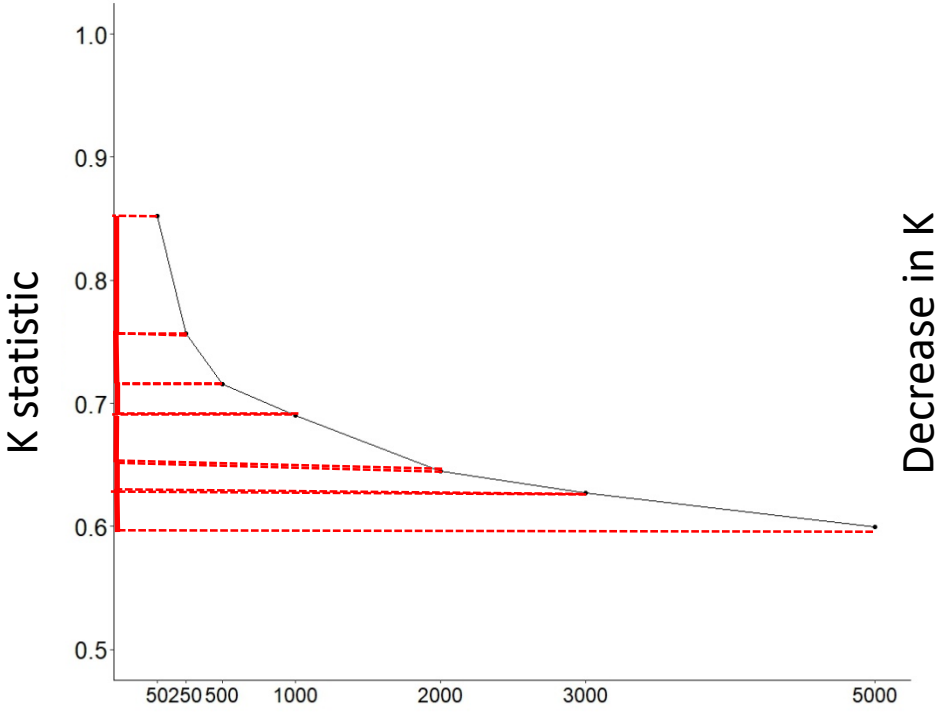
Presence records: suitable habitat

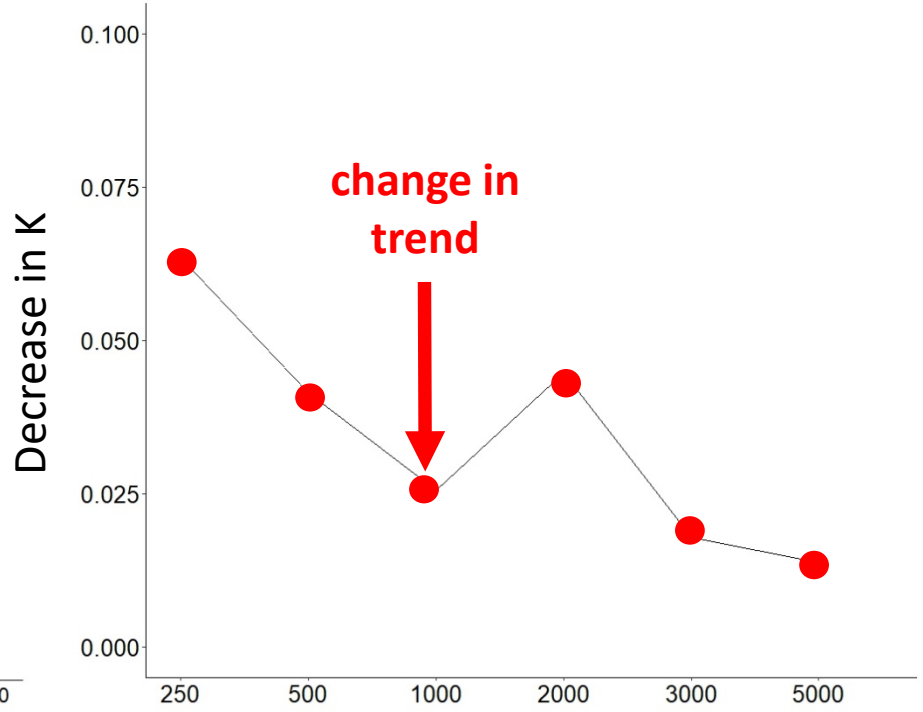
Random forest: *tuning parameters*

Italian data set

→ *266,634 records (87.5% absence – 12.5% presence)*

→ *35 predictors*



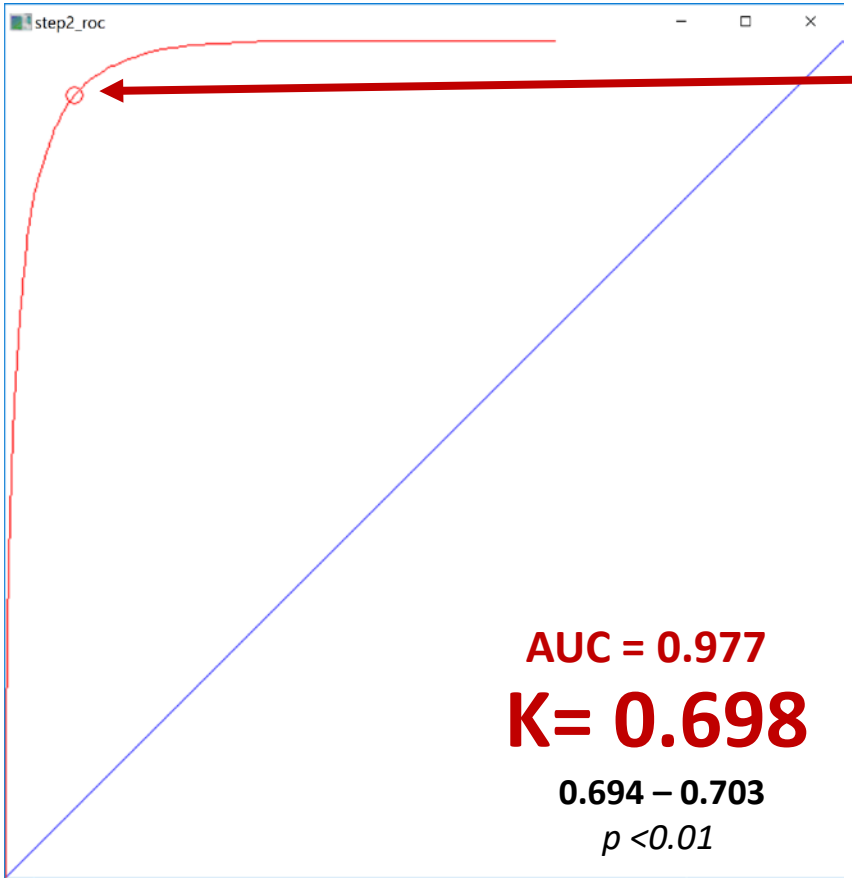


tradeoff between overfitting and accuracy

HSM → nodesize 1000 and mtry 4

HSM: *ROC CURVE & KAPPA STATISTIC*

Data set → *87.5% absence* vs. *12.5% presence*



Optimal threshold= 0.104

AUC = 0.977

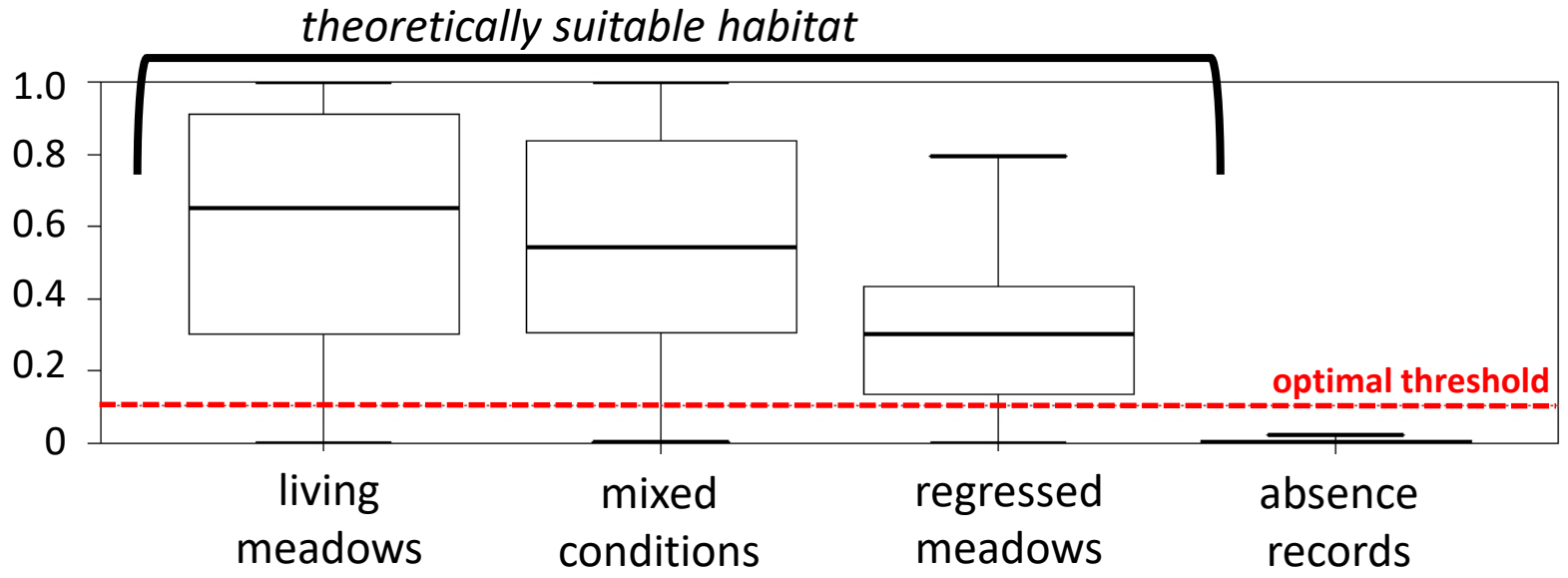
K= 0.698

0.694 – 0.703

p <0.01

✓ **HSM with good level of accuracy**

observed vs. predicted



***KS test**: null hypothesis of equal distribution*

Living vs. Mixed	Mixed vs. Regressed	Living vs. Regressed
D_{max}= 0.124	D_{max}= 0.359	D_{max}= 0.441
p << 0.001	p << 0.001	p << 0.001

✓ **Relationship between RF output & observed meadows status**

large values

intermediate values

low values

Vulnerability assessment

Therefore…

- ***Dead matte*** *is found where environmental conditions no longer allow the existence of living meadows*

- ***RF*** *aims at assessing the habitat suitability for* P. oceanica *meadows*

Areas *within which…*

HSM provided low values of probability of presence

…are the ***most vulnerable*** *ones!*

Main conclusions

- ✓ **HSM with good level of accuracy**
- ✓ **relationship between predicted and observed data:**

HSM provides quantitative estimates of meadows status

to identify the areas which are currently marginally suitable for the living meadows

TAKE AT HOME

HSM as a supporting tool to assess the vulnerability of *P. oceanica* meadows

→ *… practically*

HSM to identify potential regression risk

Thank you for the attention!

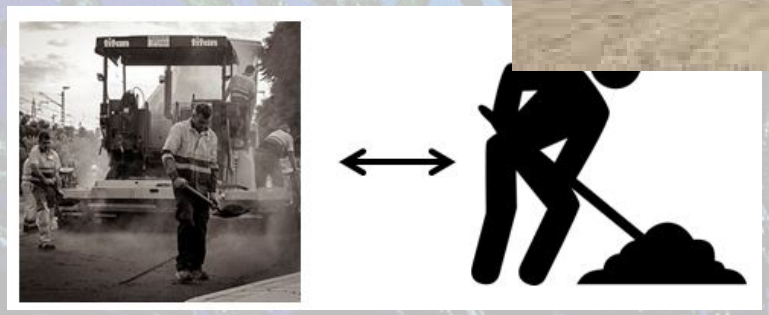
Catucci E., Scardi M.

catucci.elena@gmail.com – mscardi@mclink.it

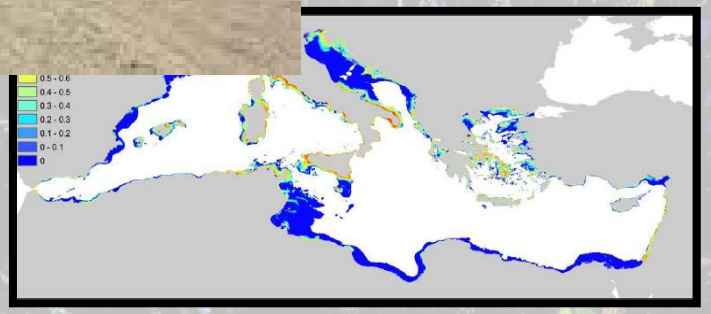
Posidonia oceanica

Modeling

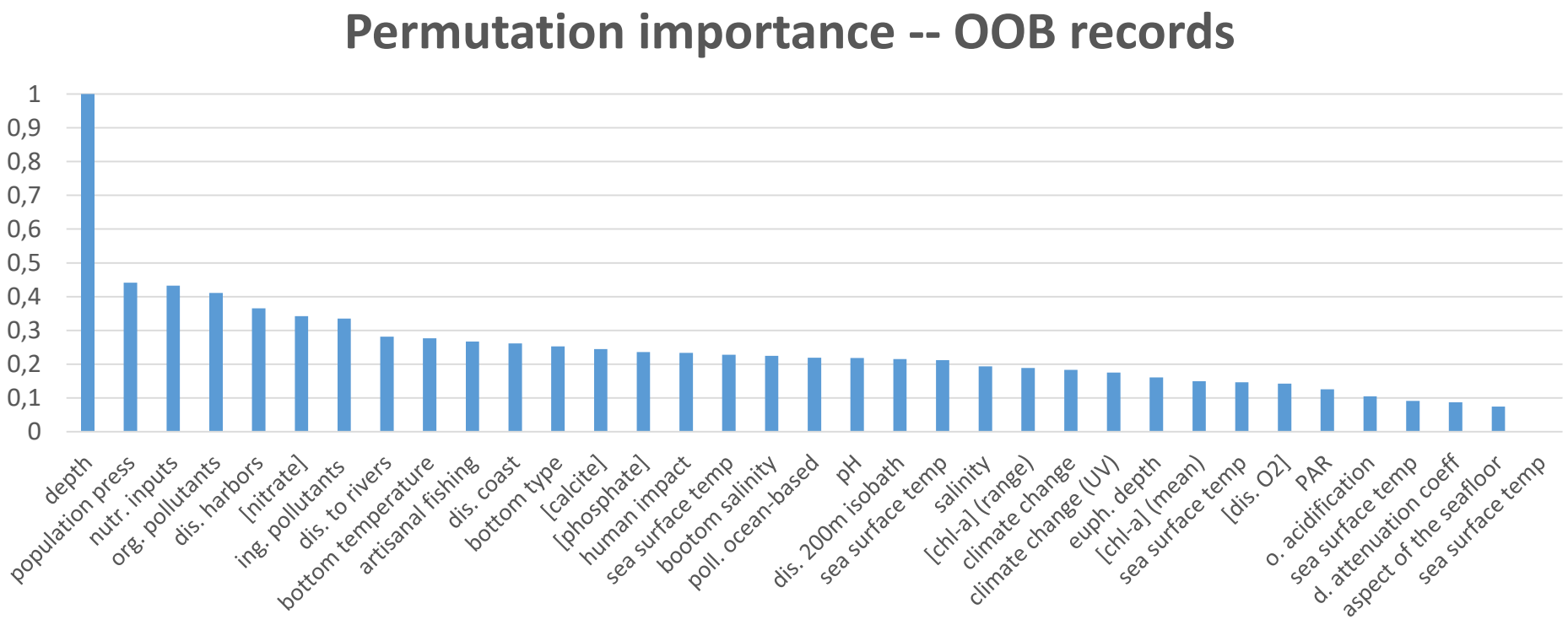
Modeling



HSM



RF: *predictors relative importance*



→ handle high-order relationships between predictors and *P. oceanica*

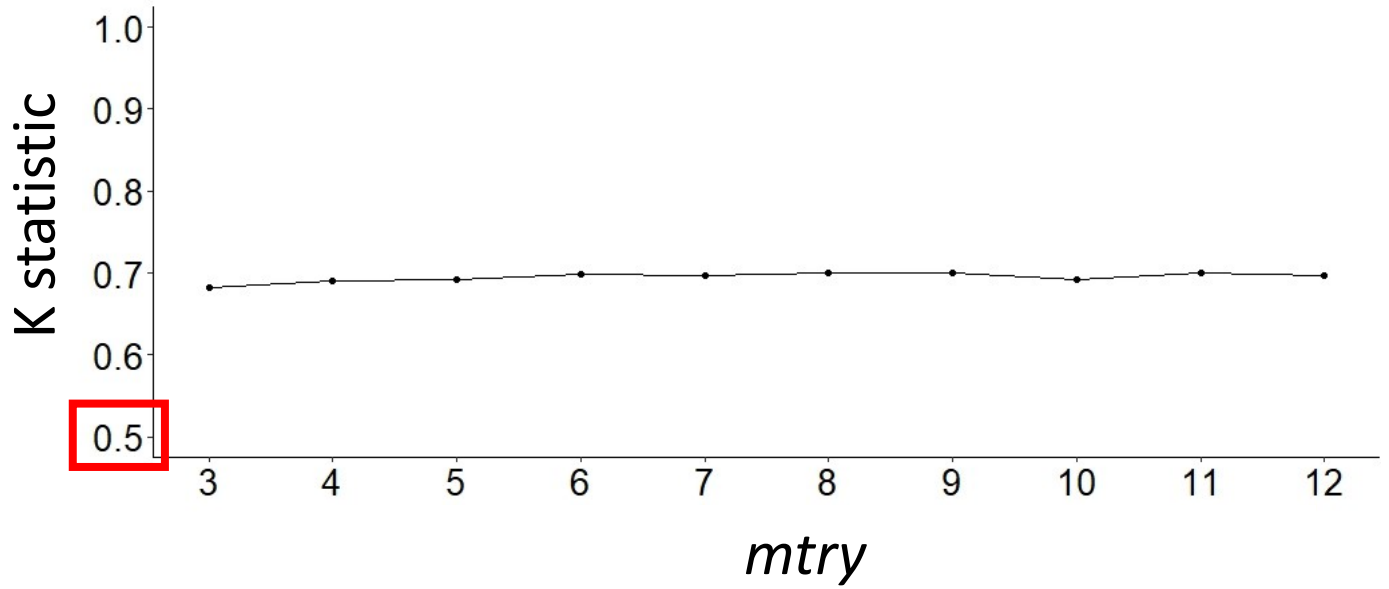
→ resemblance between the spatial distribution of predictors and *P. oceanica* meadows

Random forest: *tuning parameters*

Italian data set

→ *266,634 records (87.5% absence – 12.5% presence)*

→ *35 predictors*



tradeoff between overfitting and accuracy