A predictive occupancy model for vascular plants in temporary freshwater rockpools

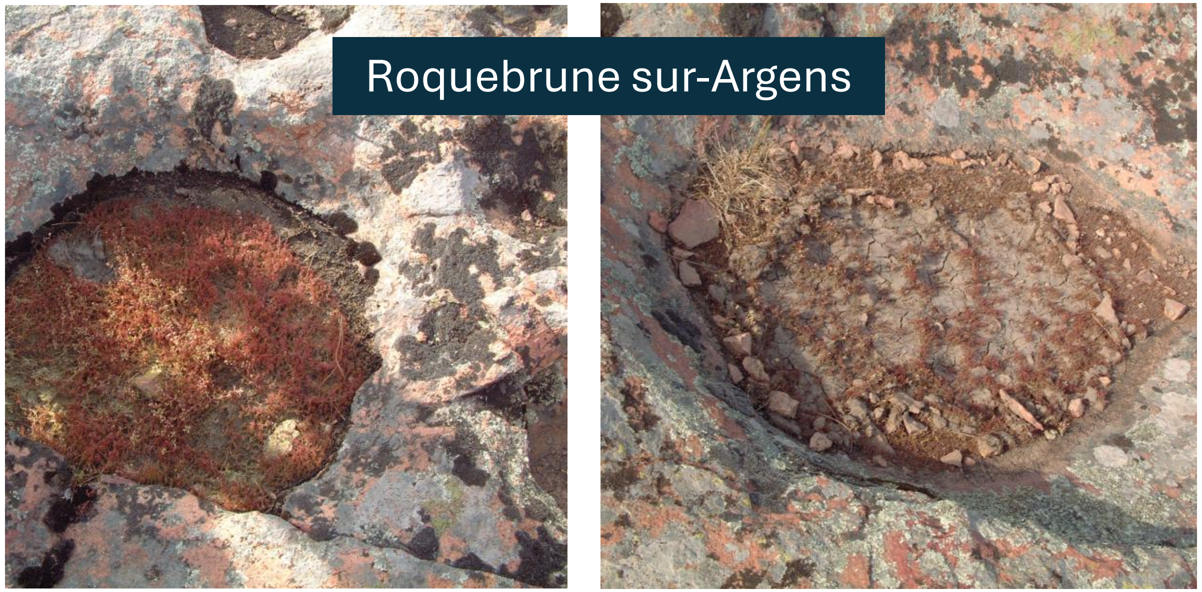
Sandro Lanfranco & Leanne Camilleri

Department of Biology

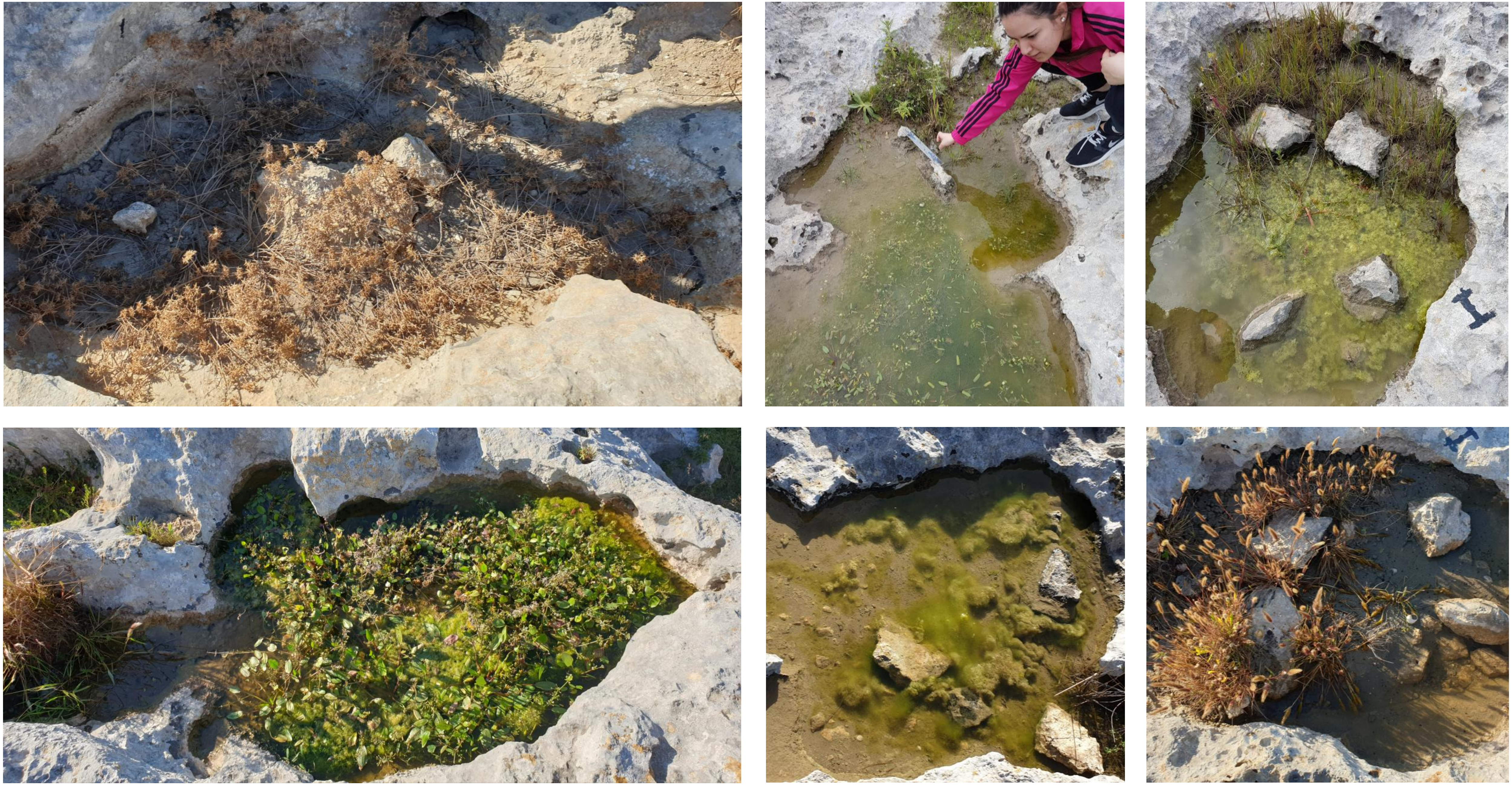
University of Malta

The Habitat


Malta


Roquebrune sur-Argens


Lagoa, Algarve




Lagoa, Algarve, May 2017

Principal characteristics

Solution hollows in karst landscapes

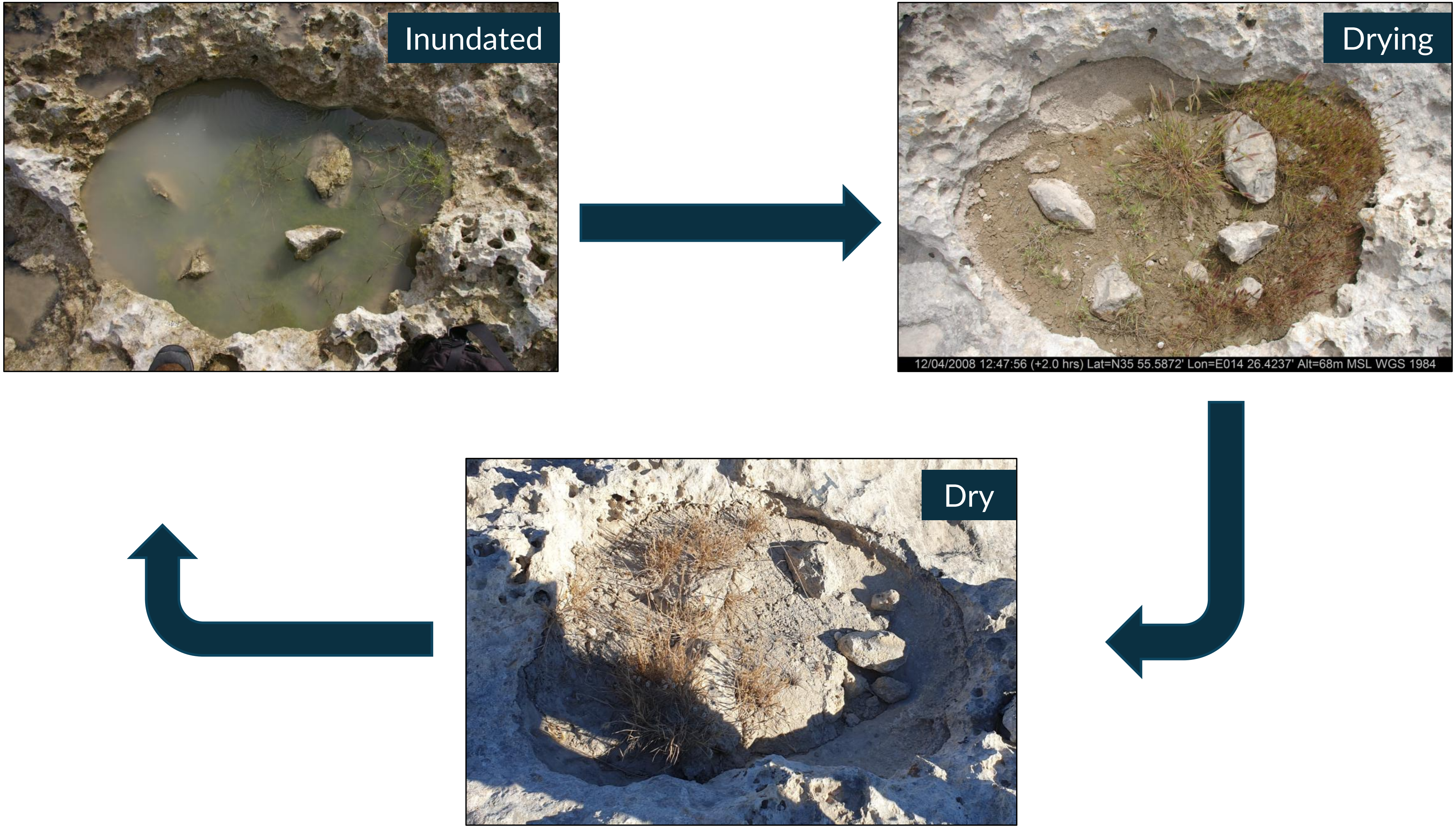
Input from precipitation and local surface runoff

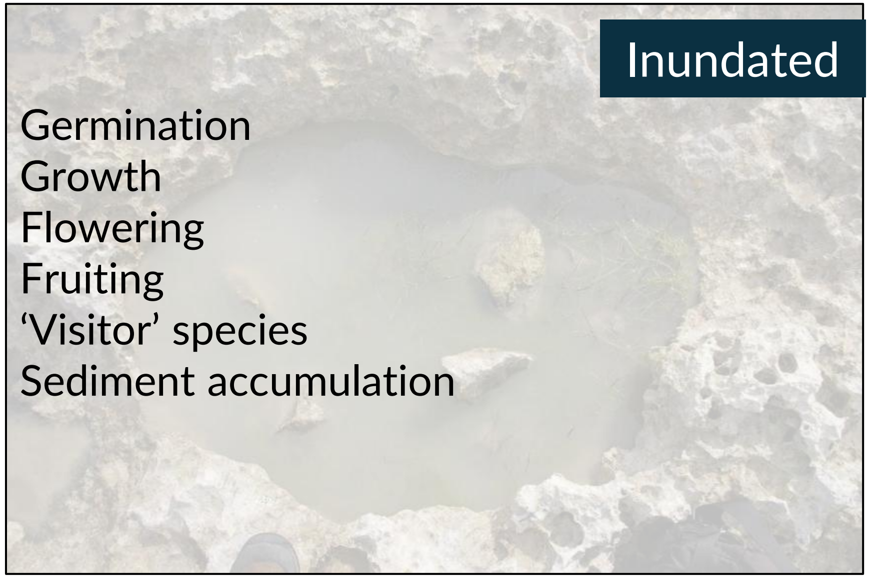
Ecological phases completely linked to climatic patterns – no immediate buffering

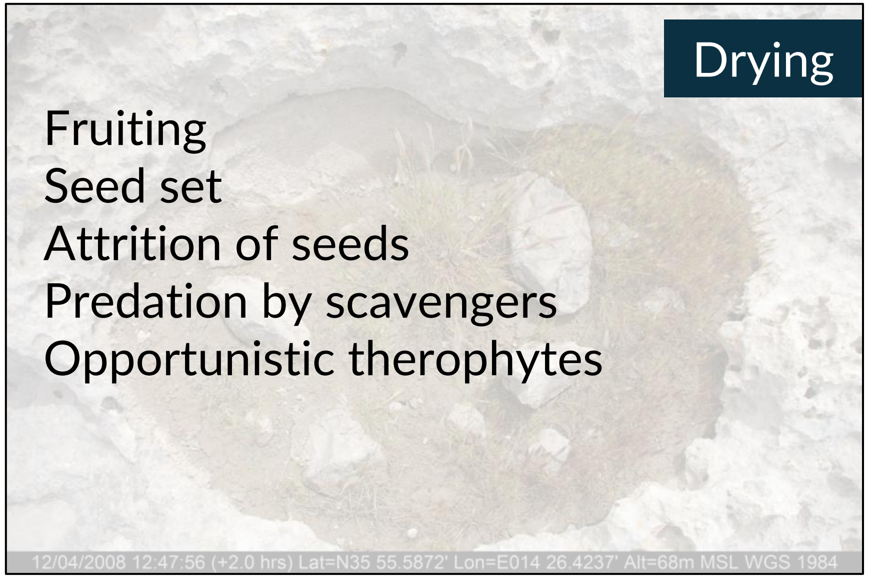

Lythrum hyssopifolia

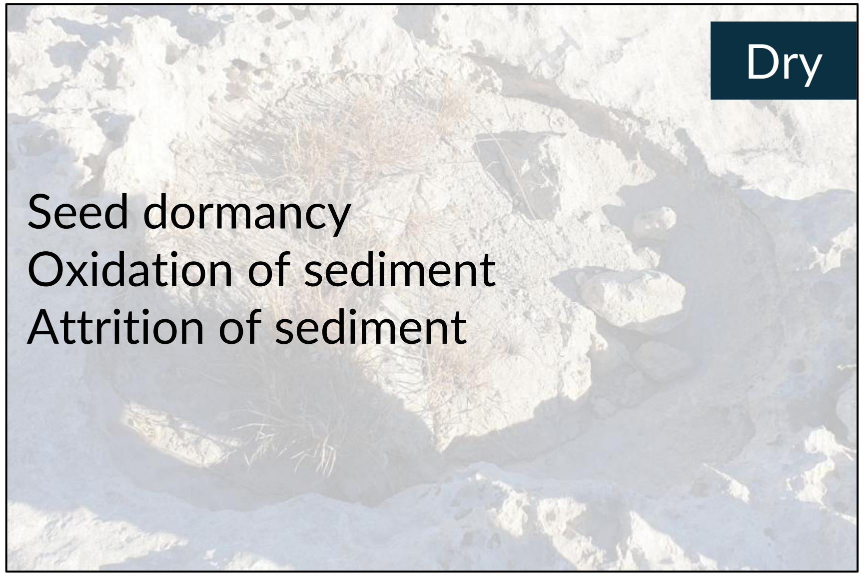

Ranunculus saniculifolius


Elatine gussonei


Inundated
Drying
12/04/2008 12:47:56 (+2.0 hrs) Lat=N35 55.5872' Lon=E014 26.4237' Alt=68m MSL WGS 1984
Dry


Inundated
Germination
Growth
Flowering
Fruiting
'Visitor' species
Sediment accumulation


Drying
Fruiting
Seed set
Attrition of seeds
Predation by scavengers
Opportunistic therophytes
12/04/2008 12:47:56 (+2.0 hrs) Lat=N35 55.5872' Lon=E014 26.4237' Alt=68m MSL WGS 1984


Dry
Seed dormancy
Oxidation of sediment
Attrition of sediment

The Challenge

The habitat: constraints and importance

- The ‘Inundated’ phase and ‘Dry’ phase are Alternate Stable States
- Flora of conservation significance is mainly present during the ‘Inundated’ phase

The challenge

- Assessing the conservation status of pools is problematic outside the ‘Inundated’ phase as the significant species are not vegetative at that time

Our Aim

- Develop a fine-grained occupancy model according to the known occurrences of each species of interest
- Would allow the ‘potential phytocoenosis’ of each pool to be estimated
- Should be simple and straightforward; can be used to inform non-specialists

Target Species

Species	Family	Life-cycle strategy
Callitriche truncata	Plantaginaceae	Aquatic
Crassula vaillantii	Crassulaceae	Amphibious-aquatic
Damasonium bourgaei	Alismataceae	Amphibious-aquatic
Elatine gussonei	Elatinaceae	Amphibious-aquatic
Lythrum hyssopifolia	Lythraceae	Amphibious-terrestrial
Mentha pulegium	Lamiaceae	Amphibious-terrestrial
Ranunculus saniculifolius	Ranunculaceae	Amphibious-aquatic
Zannichellia melitensis	Potamogetonaceae	Aquatic

Rationale behind the model

- Basin dimensions determine hydroperiod characteristics
- Hydroperiod characteristics act as an ecological filter
- It is theoretically possible to predict occupancy of a pool based on its dimensions, if ecological niche characteristics of a species are known or inferred

Collection of field data

Pool basin data

Primary pool dimensions: r_1, r_2, Z_{max}, S_{max}

Derived properties: volume (V), surface area (SA), SA:V

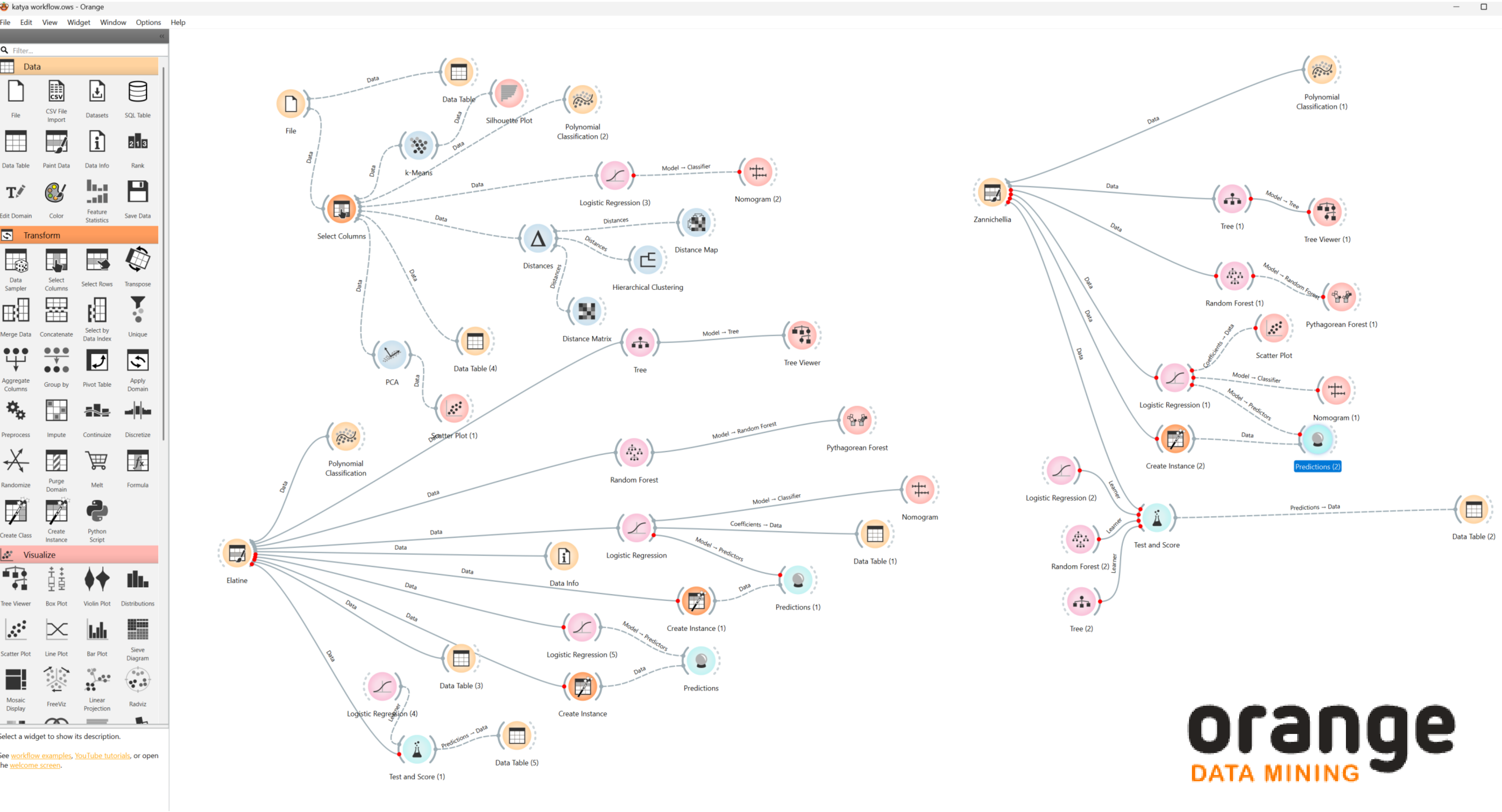
Species data

Depth of occurrence of each target species: Z_i

Sample size

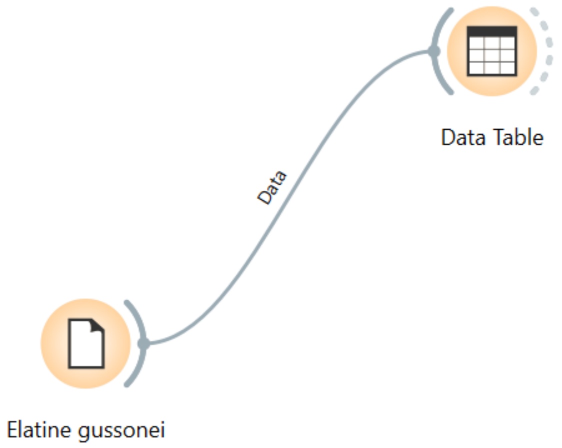
110 individual pools; October 2023 - March 2025

The Model



katya workflow.ows - Orange
File
Edit
View
Widget
Window
Options
Help
Filter...
Data
File
CSV File Import
Datasets
SQL Table
Data Table
Paint Data
Data Info
Rank
Edit Domain
Color
Feature Statistics
Save Data
Transform
Data Sampler
Select Columns
Select Rows
Transpose
Merge Data
Concatenate
Select by Data Index
Unique
Aggregate Columns
Group by
Pivot Table
Apply Domain
Preprocess
Impute
Continuize
Discretize
Randomize
Purge Domain
Melt
Formula
Create Class
Create Instance
Python Script
Visualize
Tree Viewer
Box Plot
Violin Plot
Distributions
Scatter Plot
Line Plot
Bar Plot
Sieve Diagram
Mosaic Display
FreeViz
Linear Projection
Radviz
Select a widget to show its description.
See workflow examples, YouTube tutorials, or open the welcome screen.
File
Data Table
Silhouette Plot
Polynomial Classification (2)
k-Means
Select Columns
Logistic Regression (3)
Model → Classifier
Nomogram (2)
Distances
Distance Map
Hierarchical Clustering
Distance Matrix
Data Table (4)
PCA
Scatter Plot (1)
Tree
Model → Tree
Tree Viewer
Polynomial Classification
Random Forest
Model → Random Forest
Pythagorean Forest
Elatine
Logistic Regression
Coefficients → Data
Model → Predictors
Nomogram
Data Table (1)
Data Info
Create Instance (1)
Predictions (1)
Logistic Regression (5)
Predictions
Data Table (3)
Create Instance
Logistic Regression (4)
Test and Score (1)
Predictions → Data
Data Table (5)
Zannichellia
Polynomial Classification (1)
Tree (1)
Tree Viewer (1)
Random Forest (1)
Pythagorean Forest (1)
Scatter Plot
Logistic Regression (1)
Nomogram (1)
Create Instance (2)
Predictions (2)
Logistic Regression (2)
Random Forest (2)
Tree (2)
Learner
Test and Score
Data Table (2)
Data
Distances
orange
DATA MINING

Step 1: formatting the data



Data Table - Orange

File Edit View Window Help

Info
24 instances
9 features (6.9 % missing data)
Target with 2 values
1 meta attribute

Variables
- [x] Show variable labels (if present)
- [ ] Visualize numeric values
- [x] Color by instance classes

Selection
- [ ] Select full rows

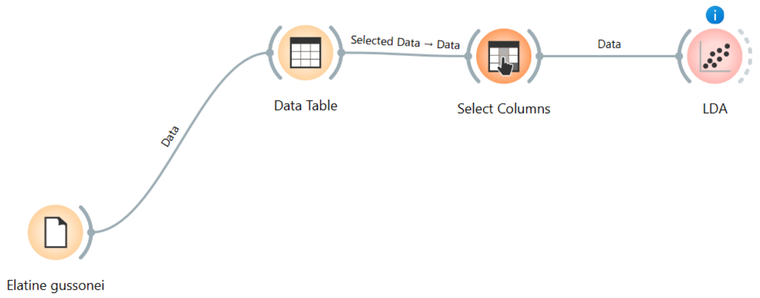
Restore Original Order

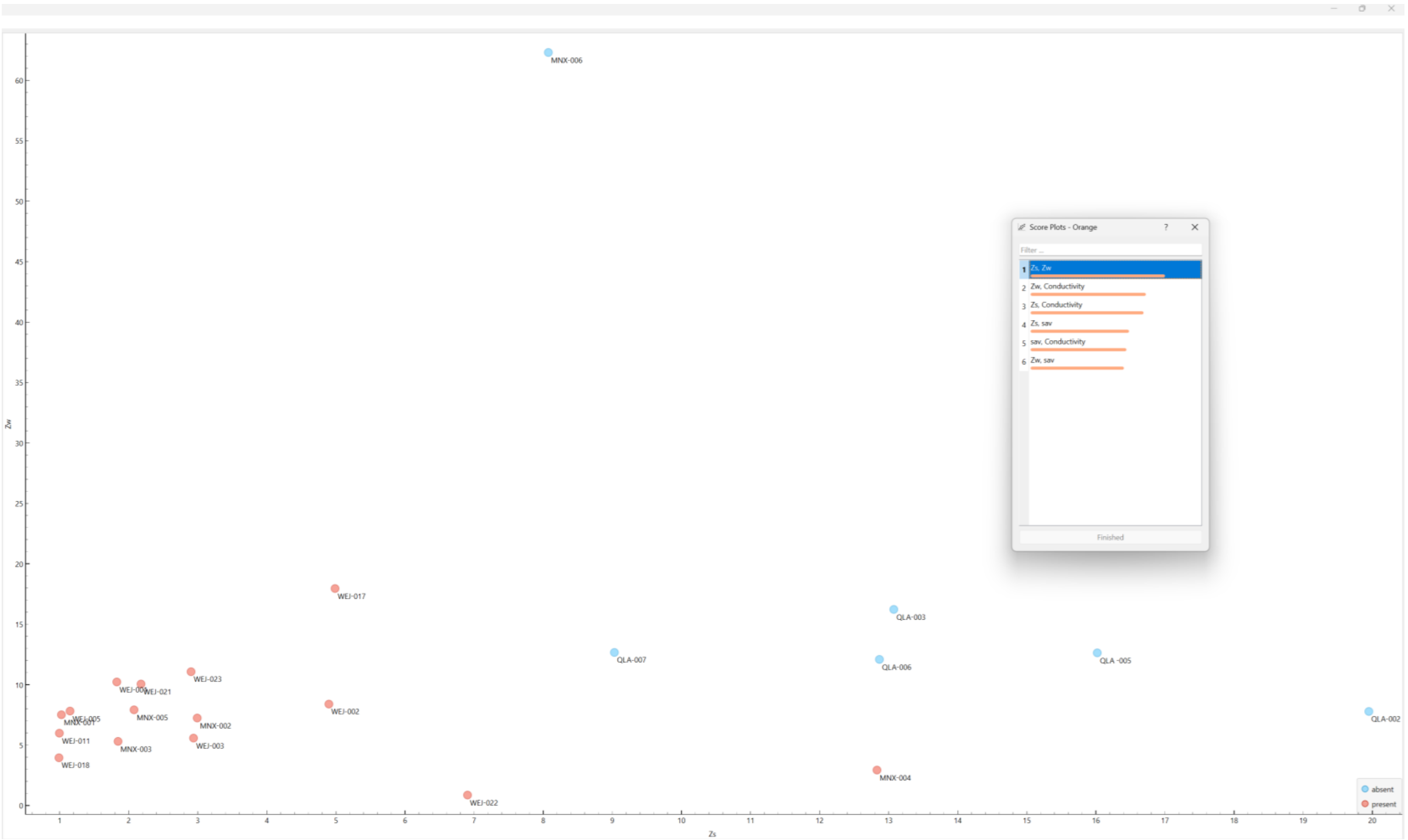
- [x] Send Automatically

	occurrence	pool_code	r1	r2	z	sav	Sed_max	Zw	Zs
1	absent	QLA-007	225	83	13	4513	9.0	13	9
2	absent	QLA-006	128	84	12	2815	13.0	12	13
3	absent	QLA -005	160	115	13	4624	16.0	13	16
4	absent	QLA-004	200	170	12	8901	?	12	?
5	absent	QLA-003	286	236	16	13680	13.0	16	13
6	absent	QLA-002	117	116	8	5685	20.0	8	20
7	absent	QLA-001	183	80	16	2875	?	16	?
8	absent	MNX-006	290	220	62	3233	8.0	62	8
9	present	MNX-005	174	114	13	4794	?	8	2
10	present	MNX-004	88	58	22	746	13.0	3	13
11	present	MNX-003	70	60	9	1466	5.0	5	2
12	present	MNX-002	224	171	14	8595	17.5	7	3
13	present	MNX-001	181	124	11	6410	8.0	8	1
14	present	WEJ-023	930	320	18	53425	16.0	11	3
15	present	WEJ-022	409	370	12	39618	8.5	1	7
16	present	WEJ-021	670	300	11	57405	14.0	10	2
17	absent	WEJ-020	160	44	16	1382	7.0	?	?
18	present	WEJ-018	440	290	15	26724	9.5	4	1
19	present	WEJ-017	188	140	14	5906	8.0	18	5
20	present	WEJ-011	240	167	?	?	12.2	6	1
21	present	WEJ-005	327	230	12	19690	8.5	8	1
22	present	WEJ-003	608	280	9	59425	20.0	6	3
23	present	WEJ-002	290	270	?	?	20.0	8	5
24	present	WEJ-001	655	542	14	79664	20.0	10	2

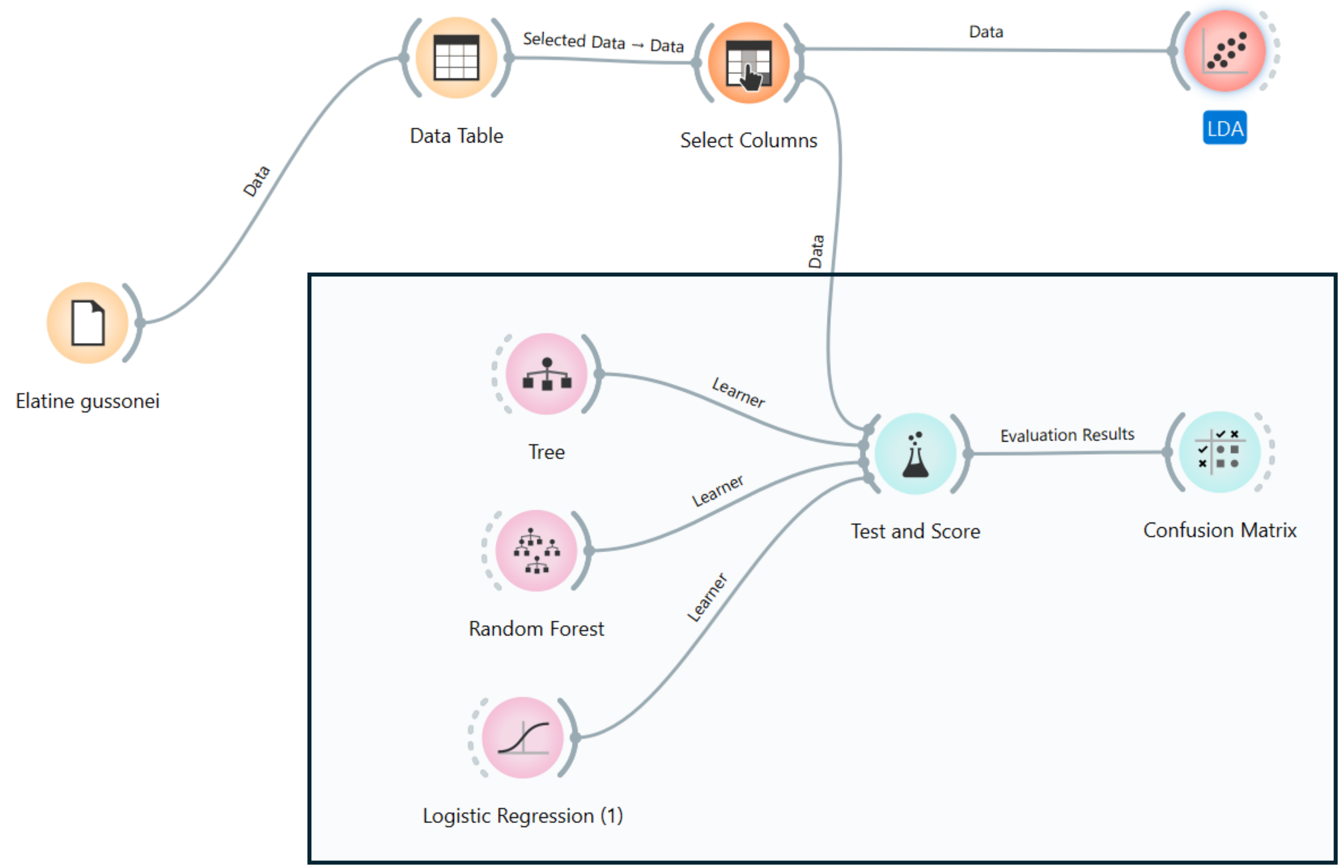
24 | 24 | 24

Step 2: finding informative projections using LDA





Step 3: Selection and testing of classifiers



Testing segment of the model

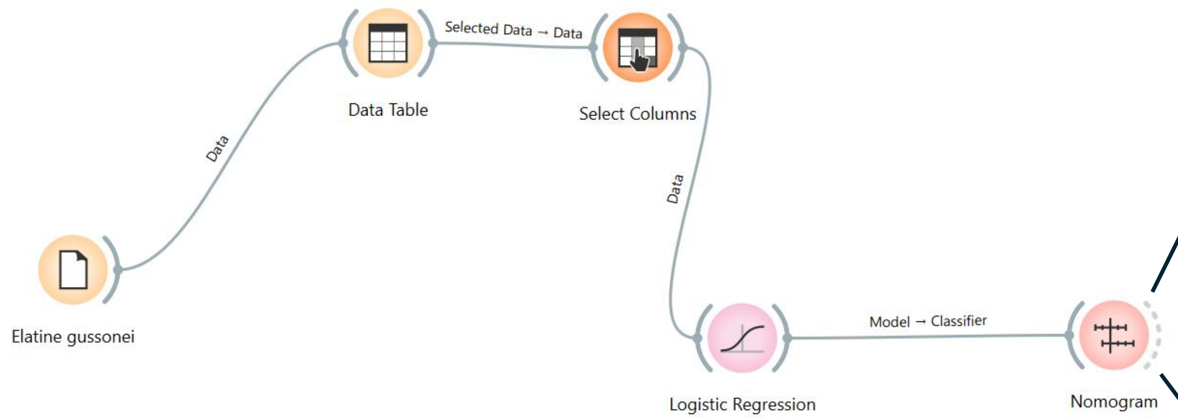
Parameters: Training set 80% Testing set 20%

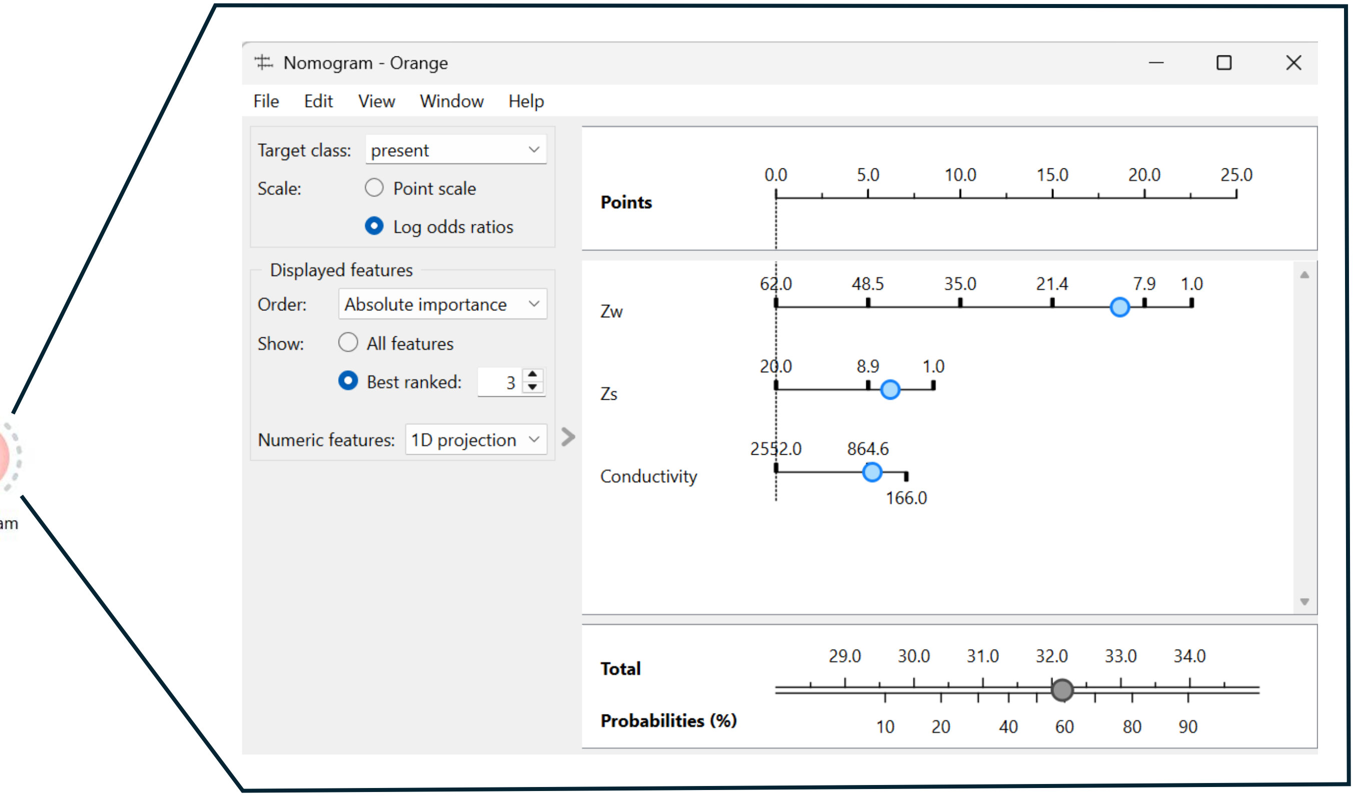
Model	AUC	MCC	F1
Logistic Regression	0.967	0.669	0.857
Random Forest	0.950	0.713	0.846
Classification Tree	0.817	0.289	0.733

Actual \ Predicted	absent	present	Σ
absent	72.7 %	7.7 %	9
present	27.3 %	92.3 %	15
Σ	11	13	24

Confusion matrix

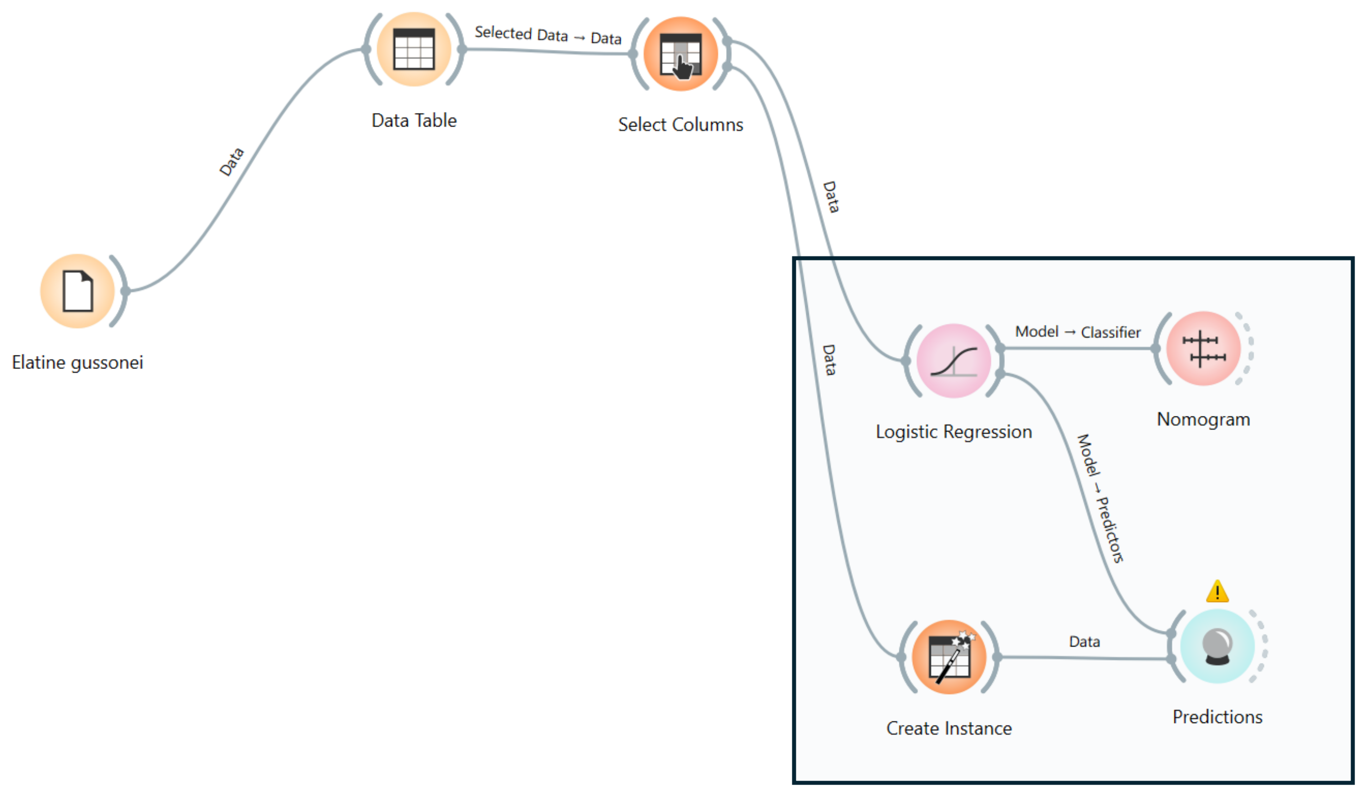
Step 4: Decision-making – using a nomogram

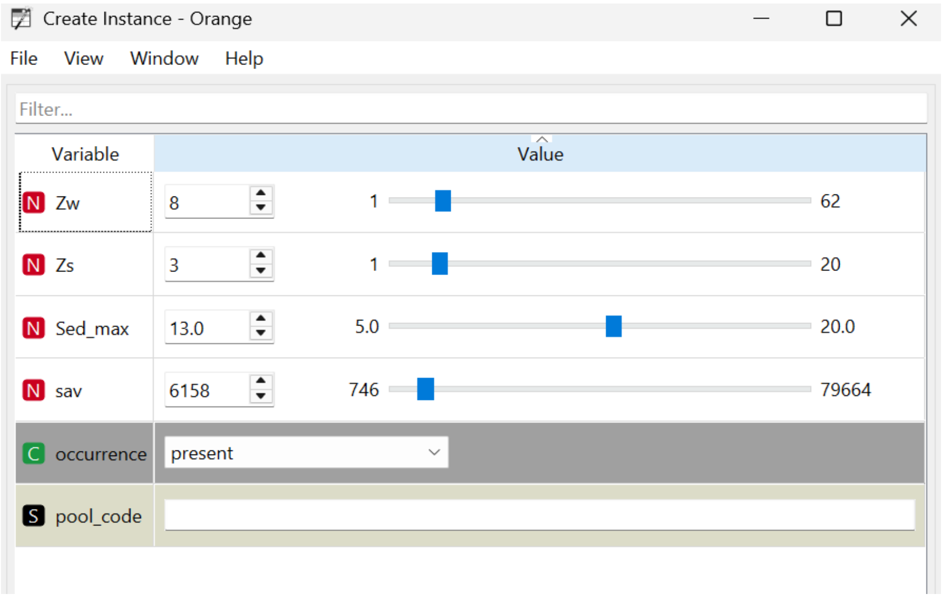


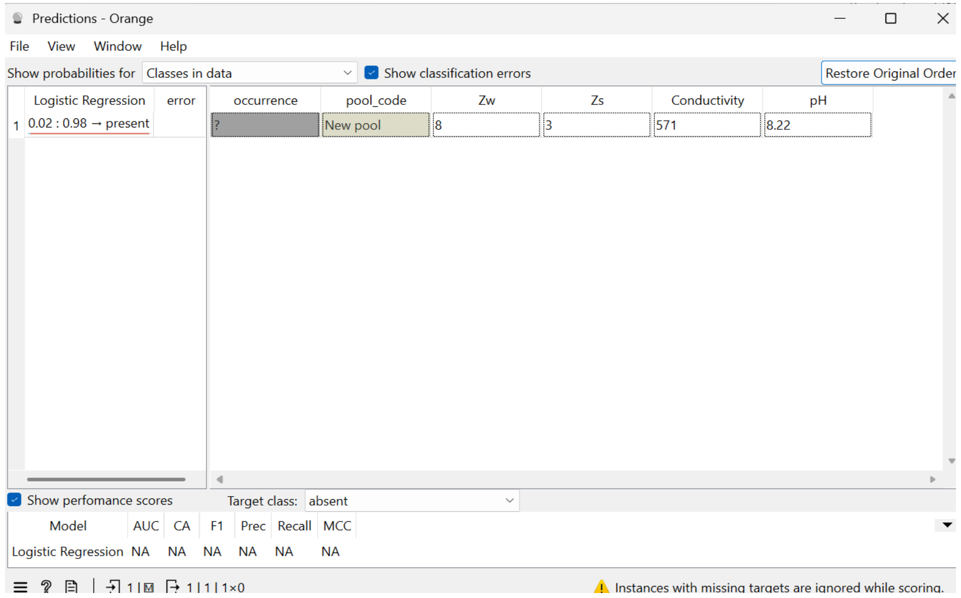


Output from the nomogram can be used to simulate different conditions

Step 4: Decision-making: creating an ‘instance’







Step 5: Phytocoenosis

The model is run repeatedly for every species of interest with this result:

Parameters

Z_{max} 8 cm

S_{max} 3 cm

r_1 184 cm

r_2 102 cm

Species	**Probability of occurrence**
Damasonium bourgaei	0.29
Elatine gussonei	0.98
Lythrum hyssopifolia	0.74
Mentha pulegium	0.55
Ranunculus saniculifolius	0.82
Zannichellia melitensis	0.07

Step 6: Conservation and Protection

Species	Probability of occurrence
Damasonium bourgaei	0.29
Elatine gussonei	0.98
Lythrum hyssopifolia	0.74
Mentha pulegium	0.55
Ranunculus saniculifolius	0.82
Zannichellia melitensis	0.07

These values can be used to estimate a protection rating for an individual pool or for an entire pool landscape

Conclusions

Pros

- The model is visual, simple, and can be used by non-specialists
- Performs relatively well even with a very small number of descriptors

Cons

- Oversimplifies the natural context
- Assumes linear combination of parameters
- Does not account for interspecific interactions

Next steps…

Future research

- Adapt model to both alternate stable states
- Adding more parameters would turn it into an explanatory model

Thank you!
From the whole team

Matthew Calleja
Leanne Camilleri
Katya Debono
Nina Gatt
Greta Micallef
Lena Schmidbauer