How to deal with georeferencing phytosociological data?

Remigiusz Pielech, University of Agriculture in Kraków, Poland

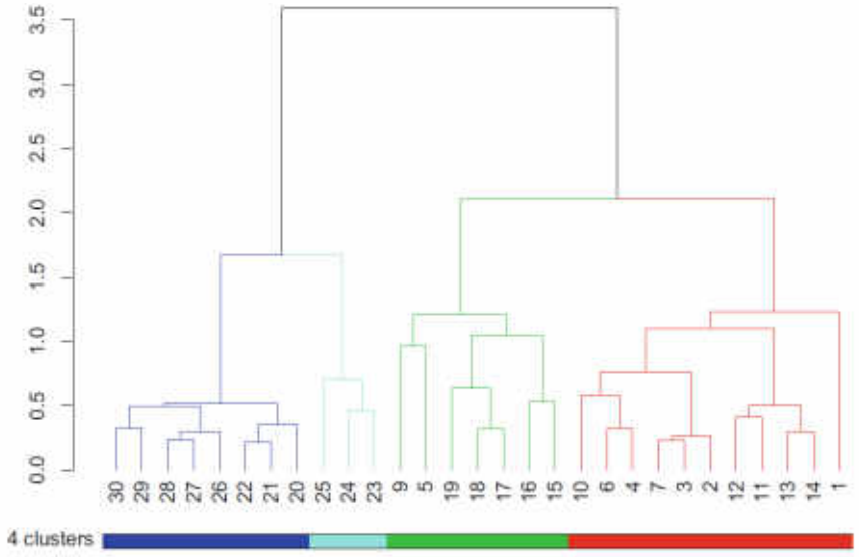
COST CONVERGES (CA16208) workshop in Kraków, Poland, 21-22 May 2018

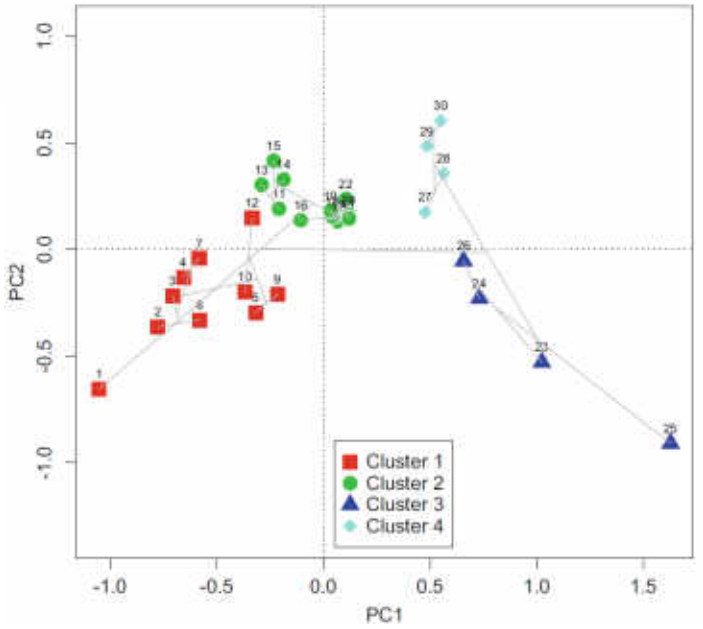
cost
EUROPEAN COOPERATION
IN SCIENCE & TECHNOLOGY

Funded by the Horizon 2020 Framework Programme
of the European Union

Why do we need a precise spatial reference ?

- Not needed for simple vegetation classifications





cost
EUROPEAN COOPERATION
IN SCIENCE & TECHNOLOGY

Going beyond classification

Precise localization can be enable using digital sources as information about the environment (as a proxy for direct measurements), e.g.:

- Digital Elevation Model (and its derivatives)
- CORINE land cover
- Digital river networks
- … and many others

How to understand *”spatial precision”*?

In databases, coordinates are accompanied by additional parameter usually describesd as:

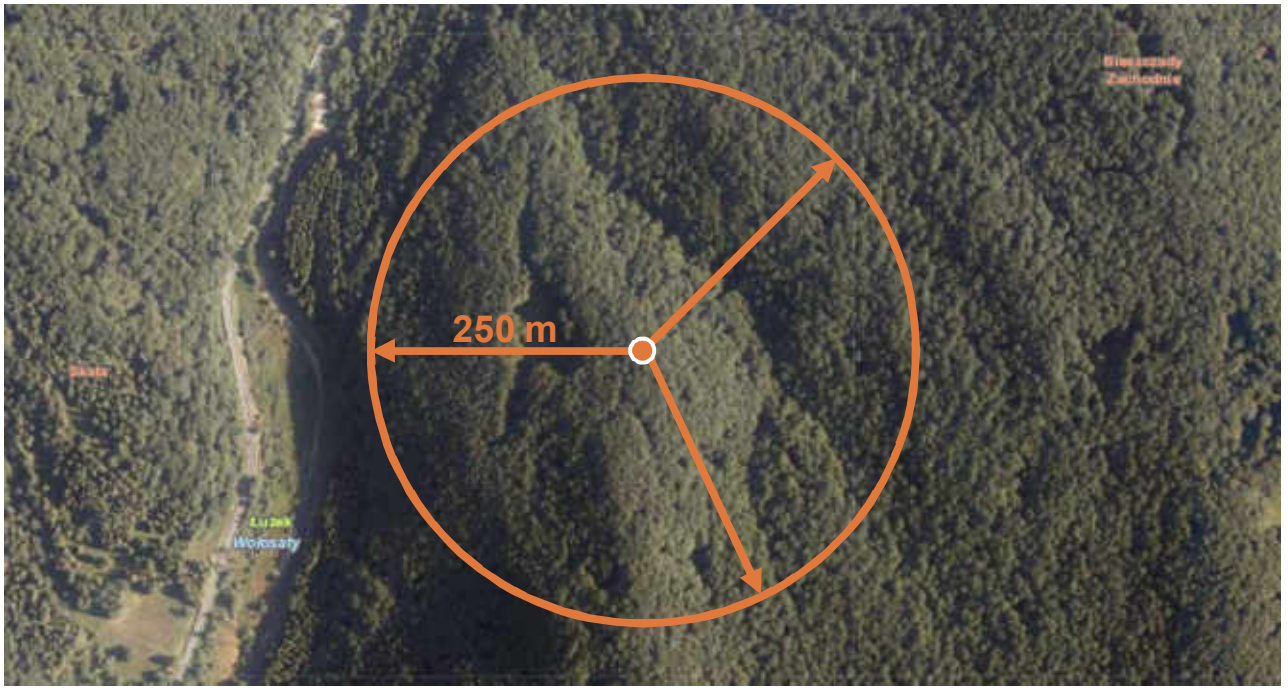
- *Distance error*
- *Locality bias*
- *Spatial uncertainity*

or any combination of the above mentioned terms…

How to understand *"spatial precision"*?

Coordinates:
49° 6' 57.57" N
22° 40' 0.58" E

Locality bias:
250 m



Coordinates: center of the area of possible occurrence

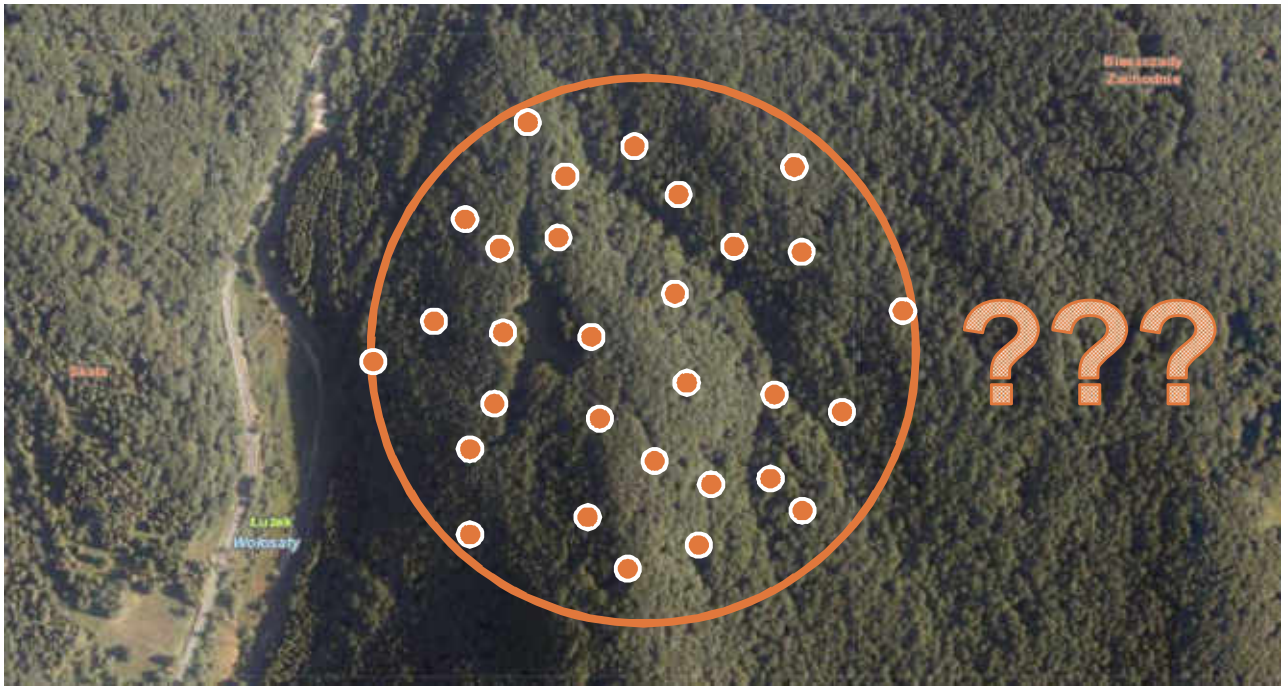
Locality bias: uncertainty radius encompassing the area of possible occurrence

cost
EUROPEAN COOPERATION IN SCIENCE & TECHNOLOGY

How to understand *„spatial precision”*?

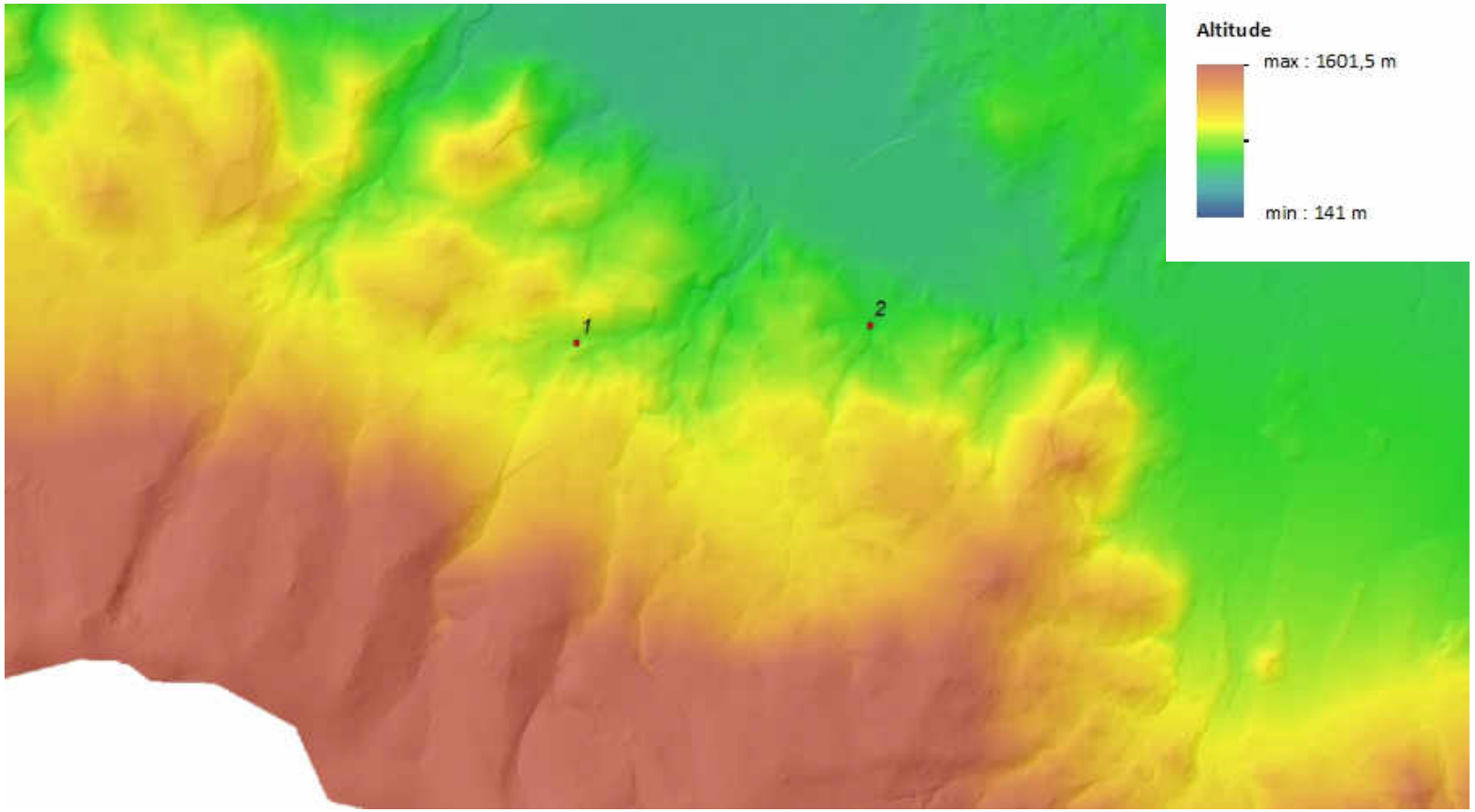
Coordinates:
49° 6’ 57.57” N
22° 40’ 0.58” E

Locality bias:
250 m

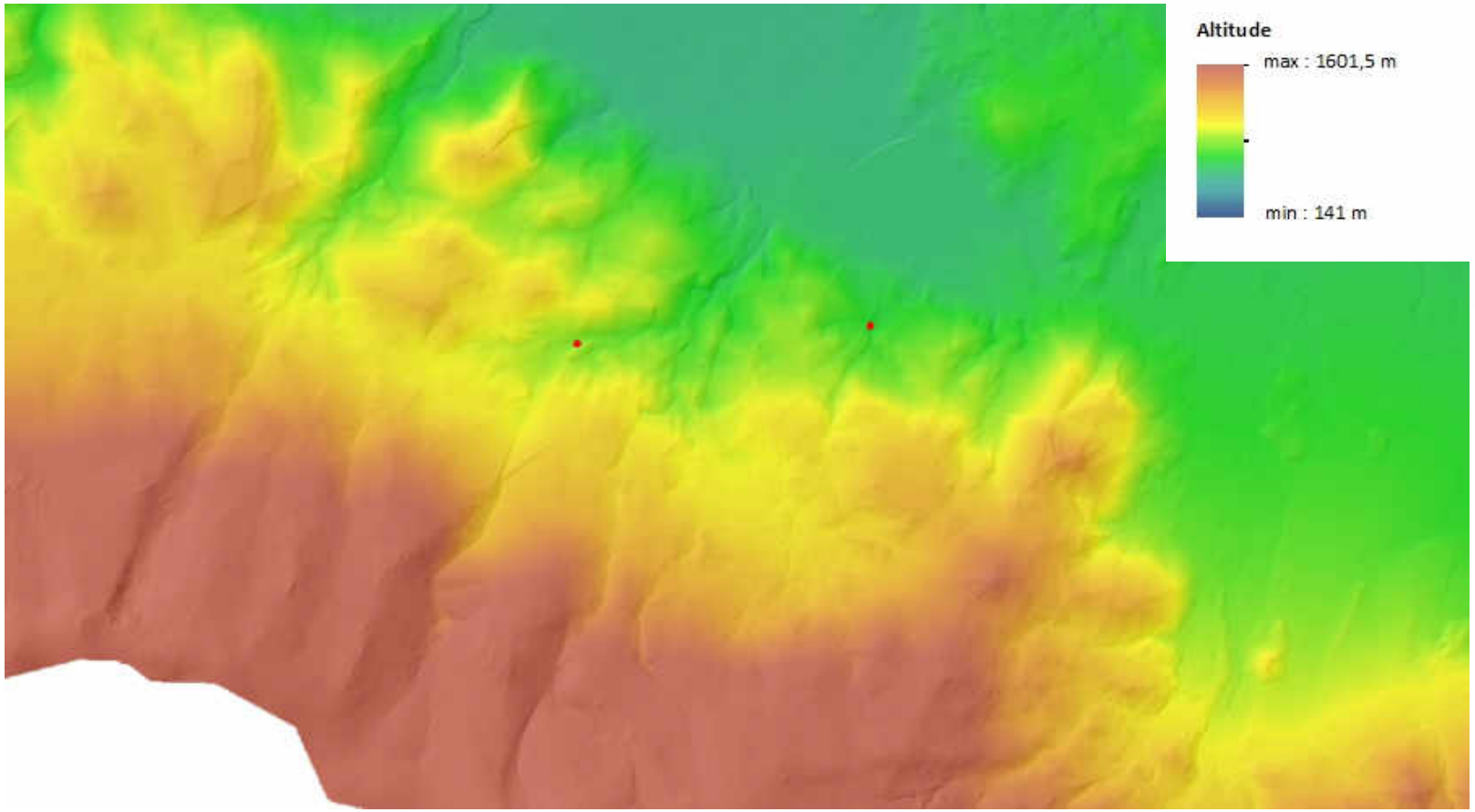


True locality is somewhere in here…

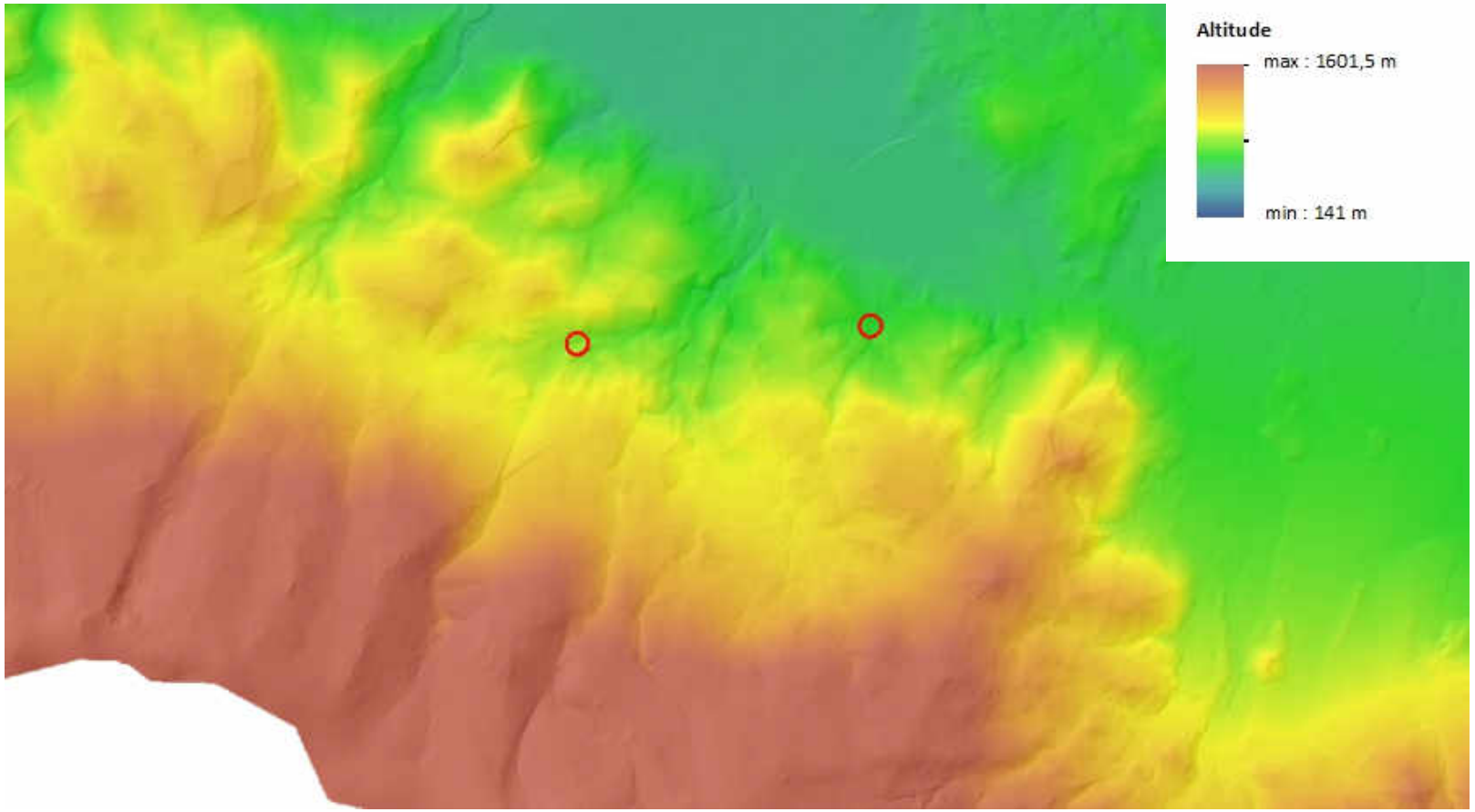
cost
EUROPEAN COOPERATION
IN SCIENCE & TECHNOLOGY



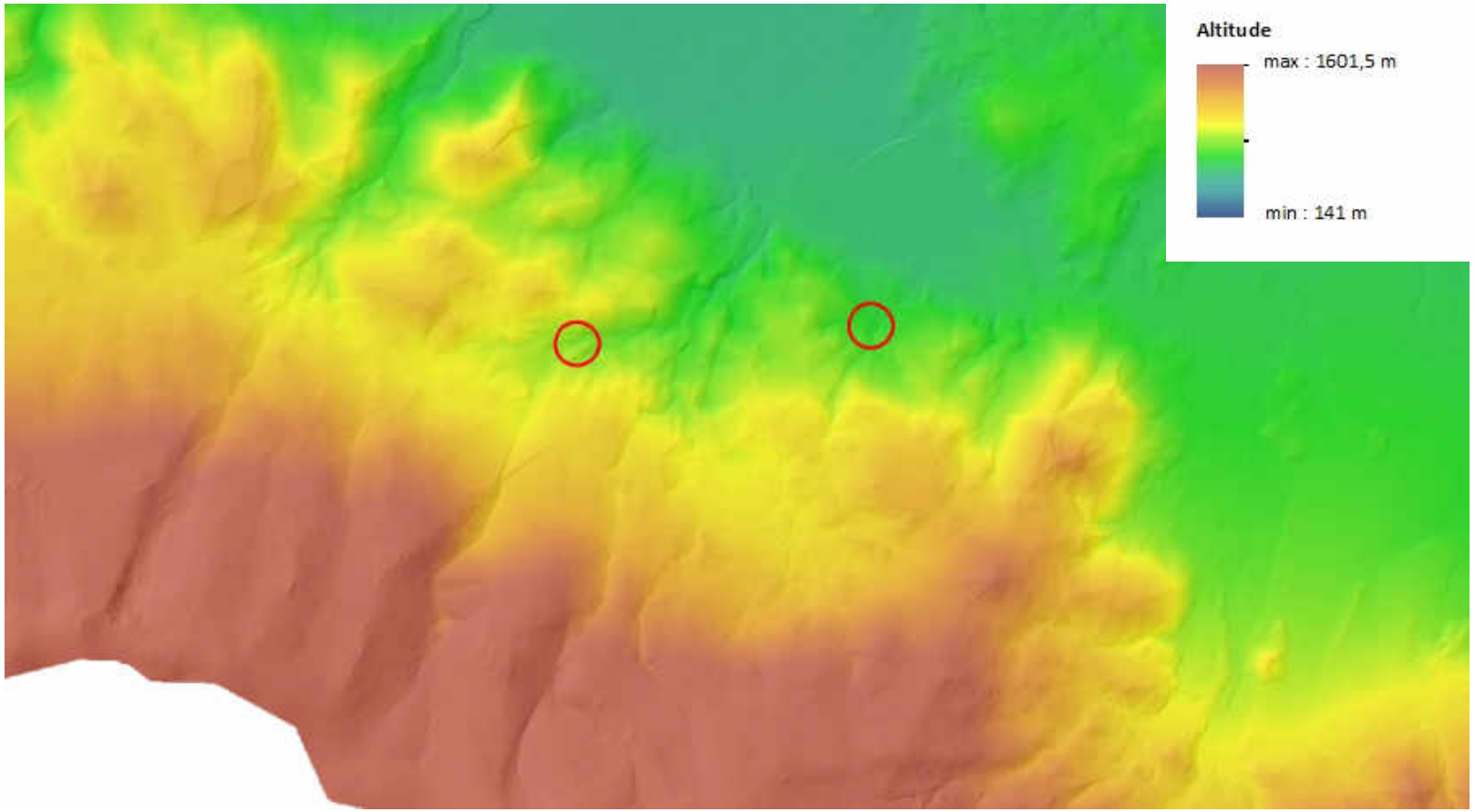
Two localities at the foot of Karkonosze Mts., SW Poland



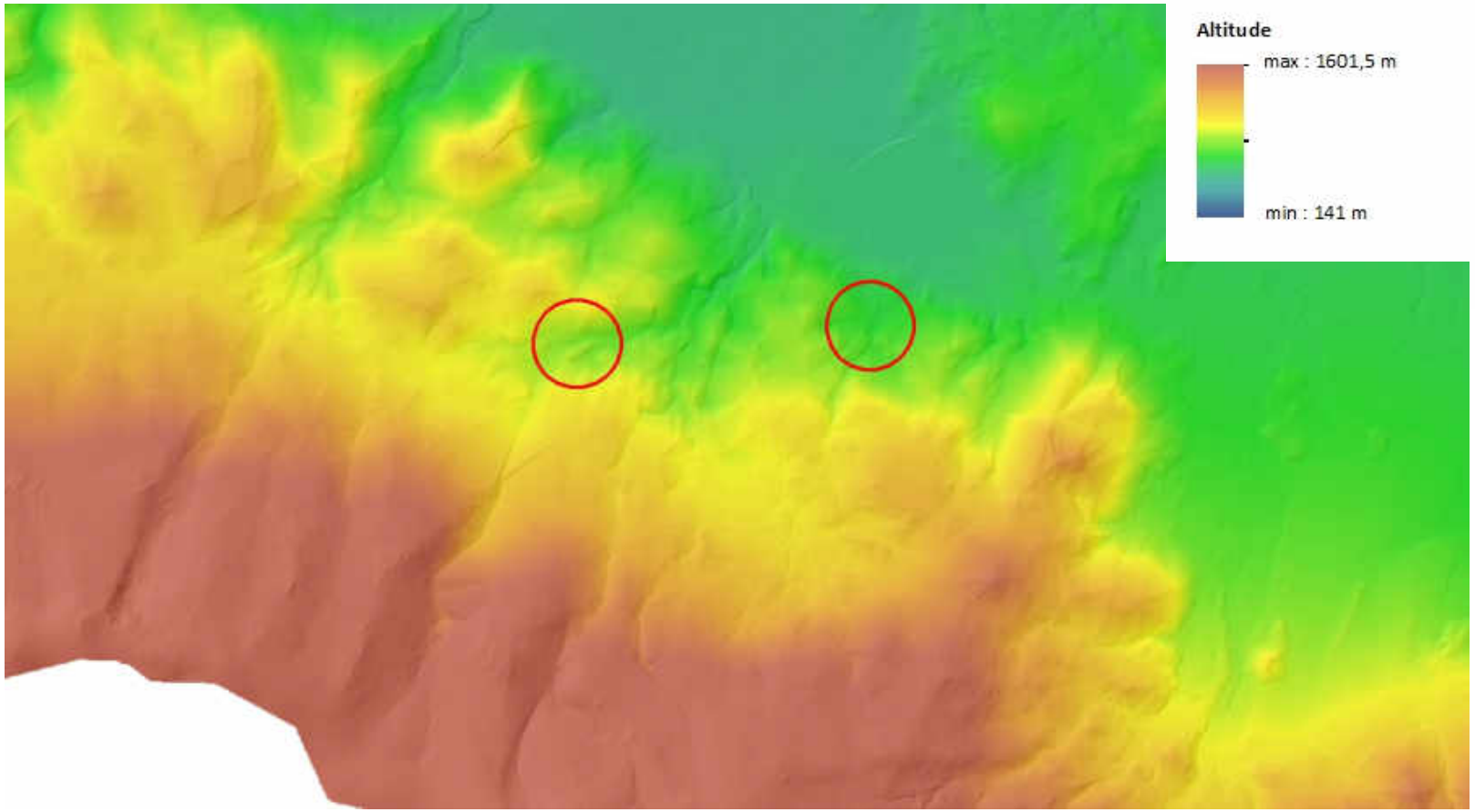
Locality bias: 25 m



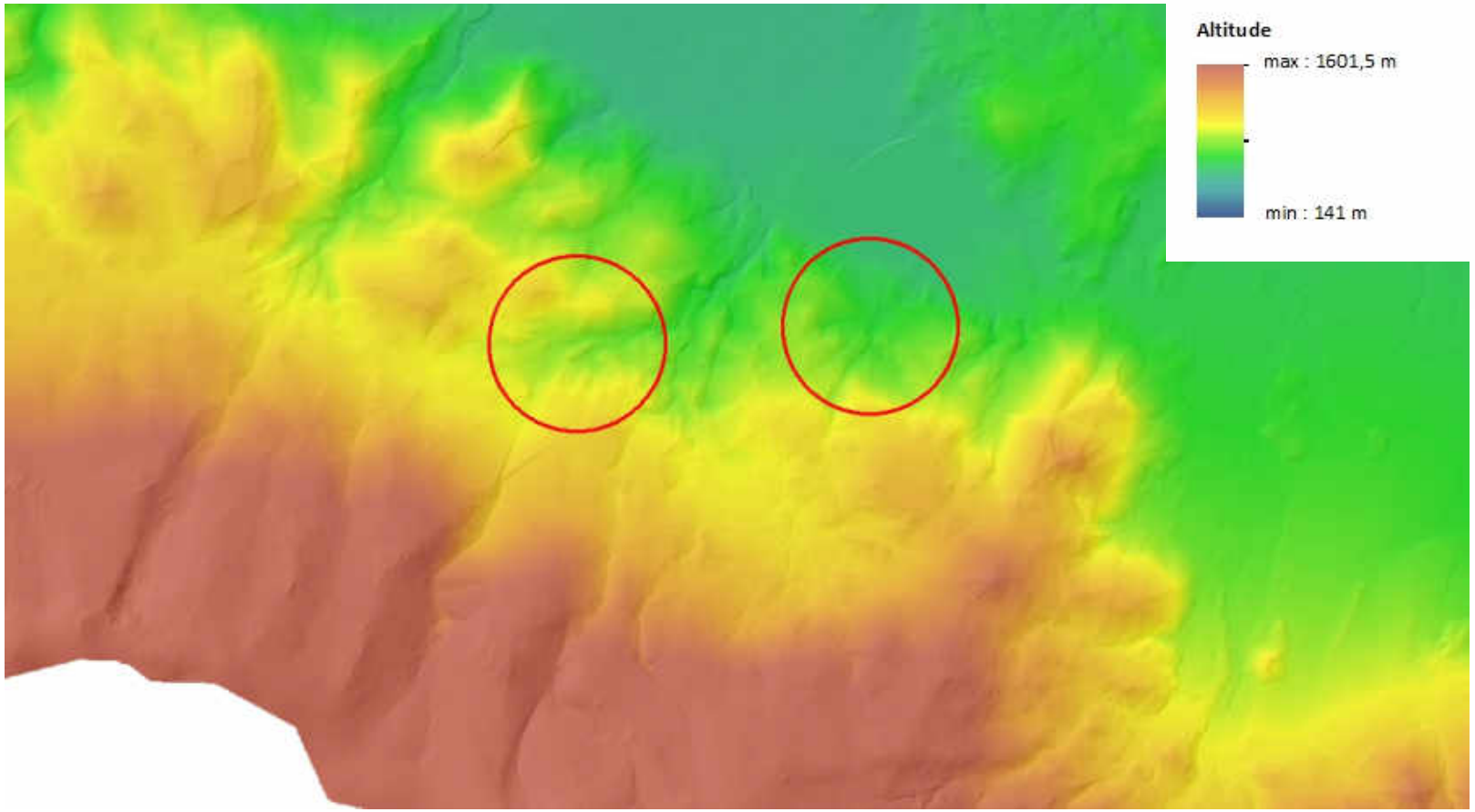
Locality bias: 100 m



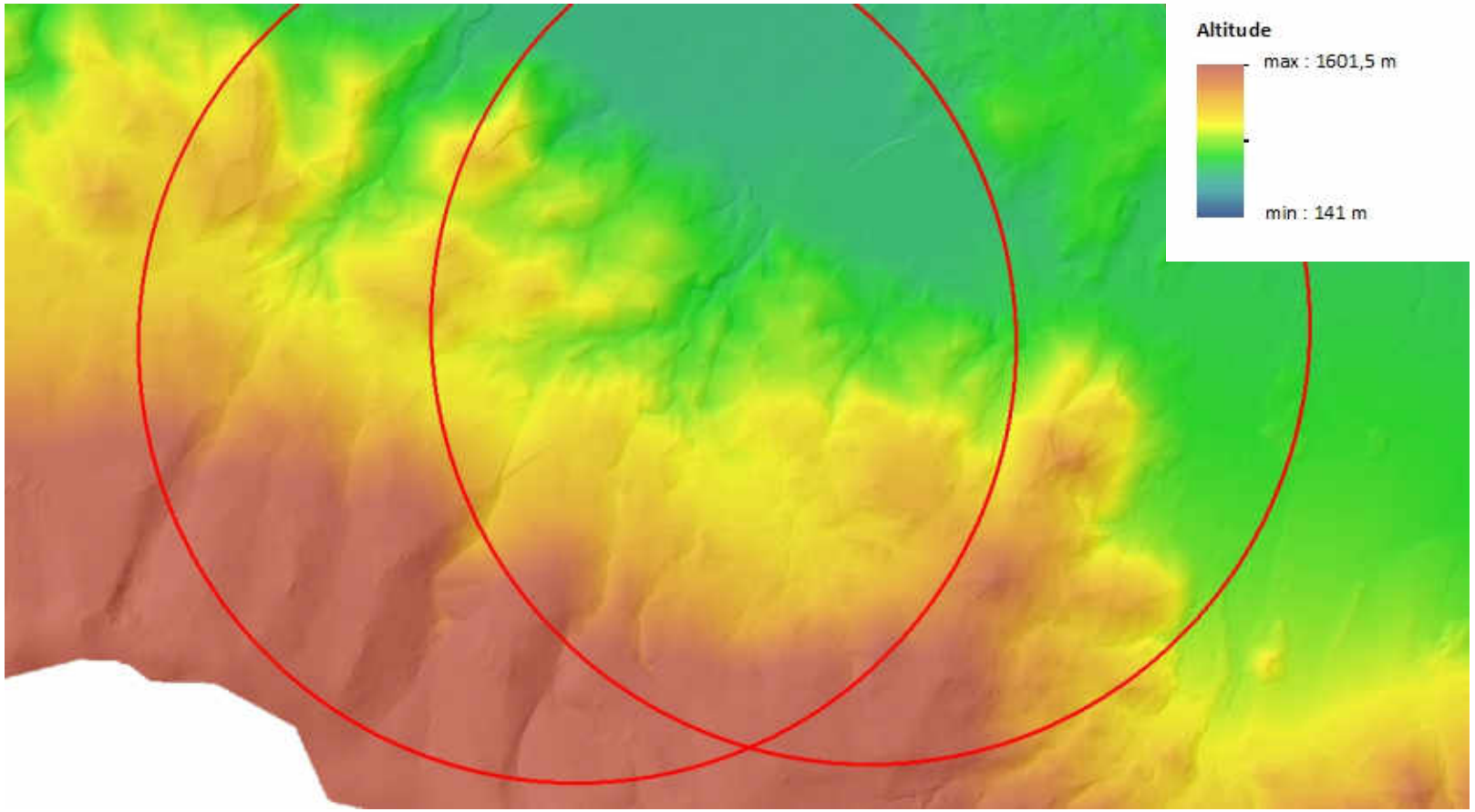
Locality bias: 250 m



Locality bias: 500 m

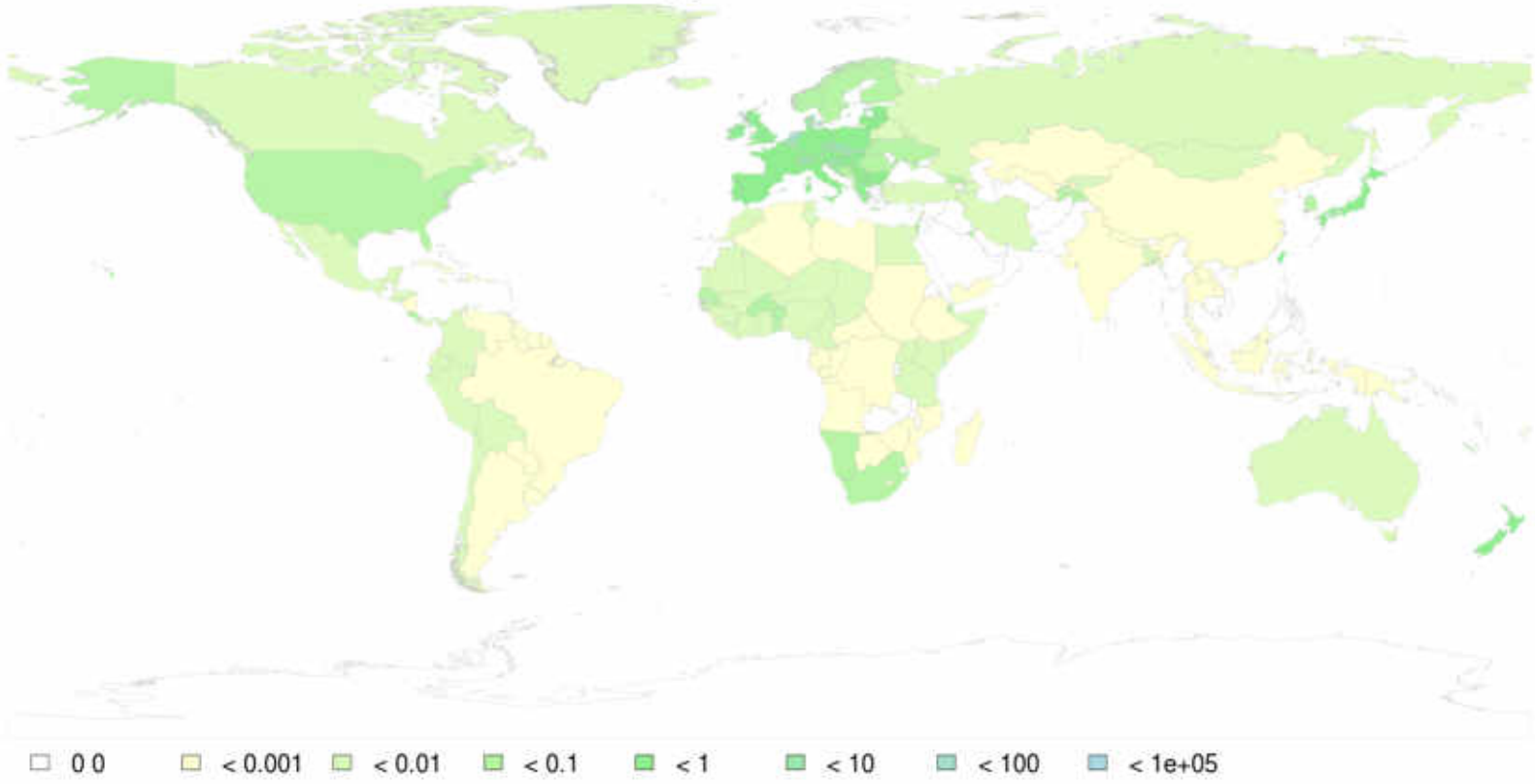


Locality bias: 1000 m



Locality bias: 5000 m

Global Index of Vegetation-Plot Databases



Density of vegetation plots around the world, measured as number of individual vegetation plots divided by country area in square kilometers

EU-NL-001
Dutch National Vegetation Database

The following numbers refer to data you really have, but not to data you are going to collect in the future. So, please no potential vegetation-plots!

Number of veg.-plots with non-overlapping area *(fact sheet)*	600000	Number of veg.-plots observations *(fact sheet)*	650933
Estimate of existing non-overlapping veg.-plots *(fact sheet)*	750000	Completeness	80

Geographic localisation *(fact sheet)*	GPS coordinates (precision 25 m or less)	50	point coordinates less precise than GPS, up to 1 km	in %	
	small grid (not coarser than 10 km)	40	political units or only on a coarser scale (above 10 km)	10	

EU-CZ-001
Czech National Phytosociological Database

The following numbers refer to data you really have, but not to data you are going to collect in the future. So, please no potential vegetation-plots!

Number of veg.-plots with non-overlapping area *(fact sheet)*	108537	Number of veg.-plots observations *(fact sheet)*	108537
Estimate of existing non-overlapping veg.-plots *(fact sheet)*	115000	Completeness	94

Geographic localisation *(fact sheet)*	GPS coordinates (precision 25 m or less	11	point coordinates less precise than GPS, up to 1 km	83.5
	small grid (not coarser than 10 km)	3.5	political units or only on a coarser scale (above 10 km)	2

cost
EUROPEAN COOPERATION IN SCIENCE & TECHNOLOGY

EU-AT-001
Austrian Vegetation Database

The following numbers refer to data you really have, but not to data you are going to collect in the future. So, please no potential vegetation-plots!

Number of veg.-plots with non-overlapping area *(fact sheet)*	53000	Number of veg.-plots observations *(fact sheet)*	53000
Estimate of existing non-overlapping veg.-plots *(fact sheet)*	100000	Completeness	53

Geographic localisation *(fact sheet)*	GPS coordinates (precision 25 m or less)	3	point coordinates less precise than GPS, up to 1 km	30
	small grid (not coarser than 10 km)	50	political units or only on a coarser scale (above 10 km)	99

EU-PL-001
Polish Vegetation Database

The following numbers refer to data you really have, but not to data you are going to collect in the future. So, please no potential vegetation-plots!

Number of veg.-plots with non-overlapping area *(fact sheet)*	84951	Number of veg.-plots observations *(fact sheet)*	84951
Estimate of existing non-overlapping veg.-plots *(fact sheet)*	84951	Completeness	100

Geographic localisation *(fact sheet)*	GPS coordinates (precision 25 m or less	2	point coordinates less precise than GPS, up to 1 km	76
	small grid (not coarser than 10 km)	in %	political units or only on a coarser scale (above 10 km)	22

cost
EUROPEAN COOPERATION IN SCIENCE & TECHNOLOGY

EU-00-004
Iberian and Macaronesian Vegetation Information System (SIVIM)

The following numbers refer to data you really have, but not to data you are going to collect in the future. So, please no potential vegetation-plots!

Number of veg.-plots with non-overlapping area *(fact sheet)*	155000	Number of veg.-plots observations *(fact sheet)*	155000
Estimate of existing non-overlapping veg.-plots *(fact sheet)*	175000	Completeness	88

Geographic localisation *(fact sheet)*	GPS coordinates (precision 25 m or less)	in %	point coordinates less precise than GPS, up to 1 km	5
	small grid (not coarser than 10 km)	92	political units or only on a coarser scale (above 10 km)	3

cost
EUROPEAN COOPERATION IN SCIENCE & TECHNOLOGY

What can we do to increase the spatial accuracy of our data?

Follow best practices for georeferencing

cost
EUROPEAN COOPERATION IN SCIENCE & TECHNOLOGY

What is georeferencing?

Georeferencing (=geocoding)

Converting locality description into mappable representation of a feature

What is georeferencing?

Locality description

Best possible accuracy

Georeferencing procedure

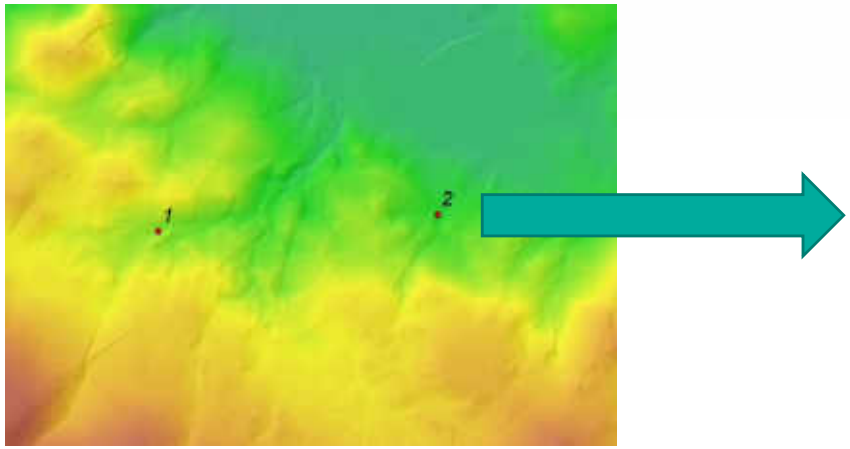
Once for a particular database:

- Selecting geographical reference system
- Georeferencing fields in database

For each record:

- Localize the area of possible occurrence
- Recording the coordinates of the central point
- Estimation of spatial accuracy (*very important!*)

Geographical reference system



Coordinates for point no 2:

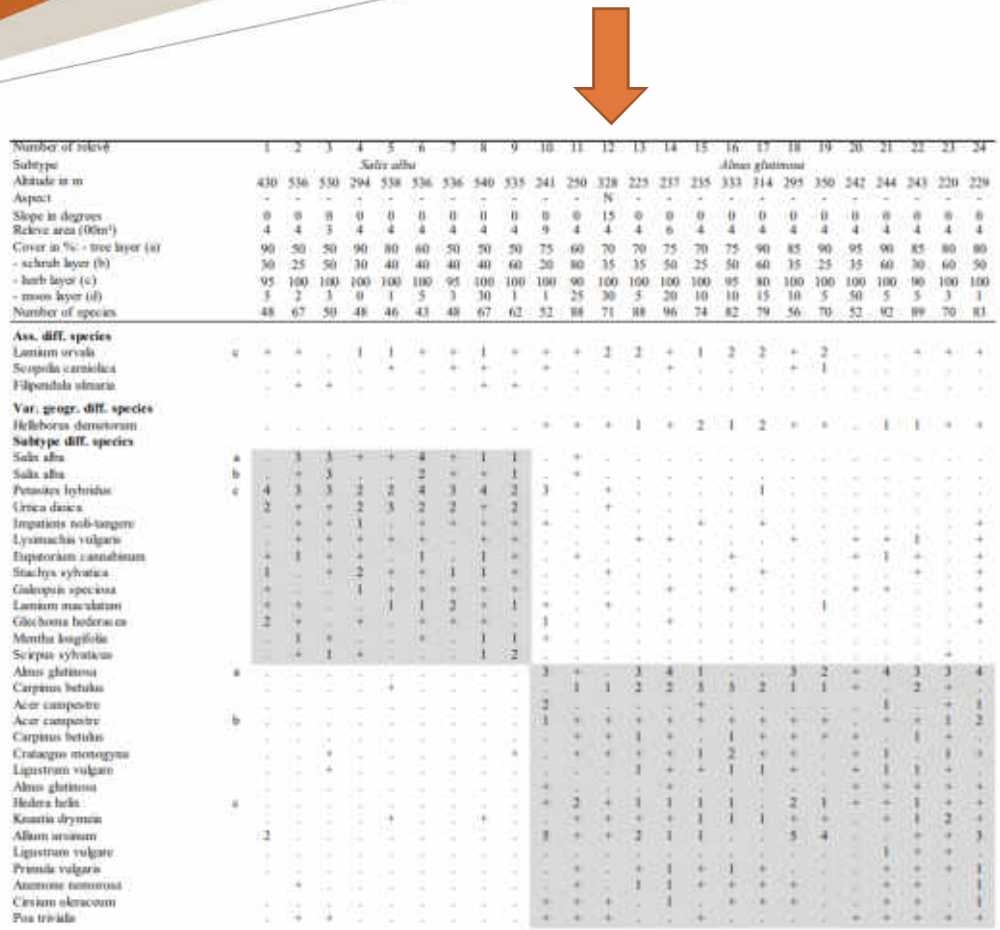
	X	Y
Układ 1992	267875	332672
Układ 1965	3634752	5532182
Układ 2000	6338185	5633516
WGS 84	15° 42' 13.36" E	50° 48' 51.75" N
WGS 84	15° 42.223' E	50° 48.863' N
WGS 84	15.703712	50.814375

cost
EUROPEAN COOPERATION
IN SCIENCE & TECHNOLOGY

Georeferencing fields

<u>The following are a minimum:</u>

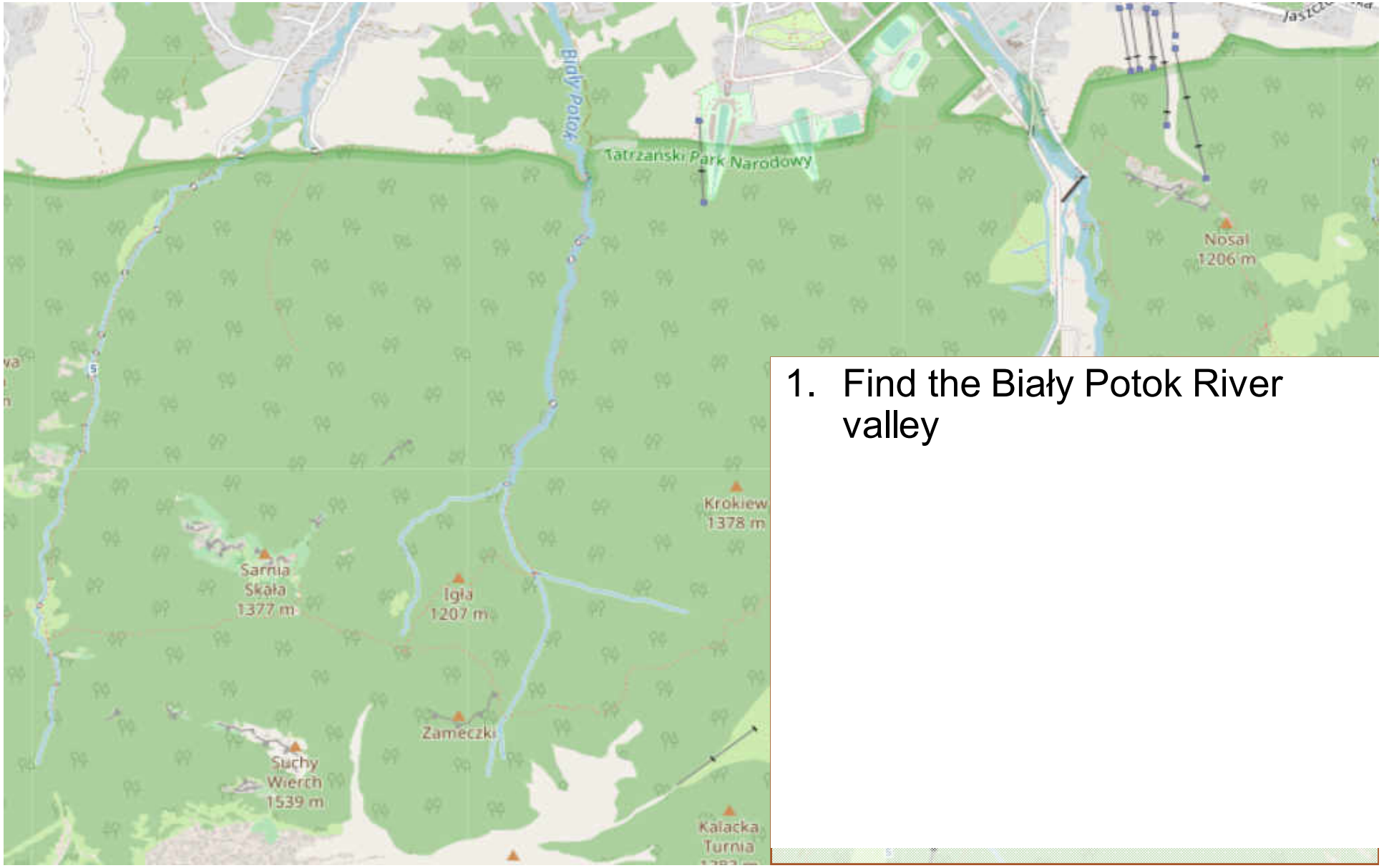
- Longitude
- Latitude
- Spatial accuracy/distance error/locality bias
- Remarks (usually omitted in phytosociological dbs)



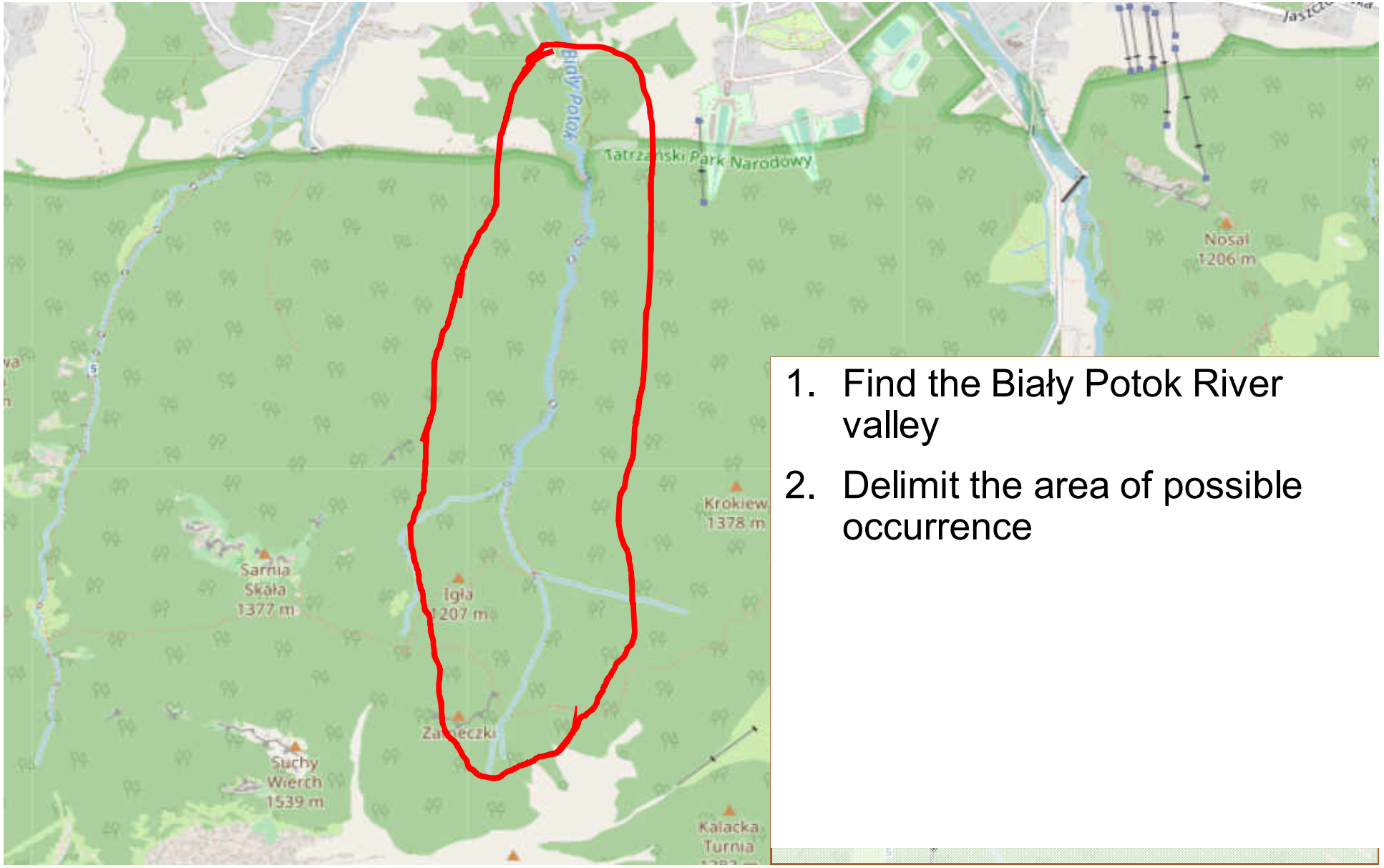
Number of relevé		1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24
Subtype						*Salix alba*												*Alnus glutinosa*							
Altitude in m		430	536	550	294	538	536	536	540	535	241	250	328	225	237	235	333	314	295	350	242	244	243	220	229
Aspect		-	-	-	-	-	-	-	-	-	-	-	N	-	-	-	-	-	-	-	-	-	-	-	-
Slope in degrees		0	0	0	0	0	0	0	0	0	0	0	15	0	0	0	0	0	0	0	0	0	0	0	0
Relevé area (00m²)		4	4	3	4	4	4	4	4	4	9	4	4	4	6	4	4	4	4	4	4	4	4	4	4
Cover in %: - tree layer (a)		90	50	50	90	80	60	50	50	50	75	60	70	70	75	70	75	90	85	90	95	90	85	80	80
- schrub layer (b)		30	25	50	30	40	40	40	40	60	20	80	35	35	50	25	50	60	35	25	35	60	30	60	50
- herb layer (c)		95	100	100	100	100	100	95	100	100	100	90	100	100	100	100	95	80	100	100	100	100	90	100	100
- moss layer (d)		5	2	3	0	1	5	3	30	1	1	25	30	5	20	10	10	15	10	5	50	5	5	3	1
Number of species		48	67	50	48	46	43	48	67	62	52	88	71	88	96	74	82	79	56	70	52	92	89	70	83
Ass. diff. species																									
Lamium orvala	c	+	+	.	1	1	+	+	1	+	+	+	2	2	+	1	2	2	+	2	.	.	+	+	+
Scopolia carniolica		.	.	.	.	+	.	+	+	.	+	.	.	.	+	.	.	.	+	1	.	.	.	.	.
Filipendula ulmaria		.	+	+	.	.	.	.	+	+	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
Var. geogr. diff. species																									
Helleborus dumetorum		.	.	.	.	.	.	.	.	.	+	+	+	1	+	2	1	2	+	+	.	1	1	+	+
Subtype diff. species																									
Salix alba	a	.	3	3	+	+	4	+	1	1	.	+	.	.	.	.	.	.	.	.	.	.	.	.	.
Salix alba	b	.	+	3	.	.	2	+	+	1	.	+	.	.	.	.	.	.	.	.	.	.	.	.	.
Petasites hybridus	c	4	3	3	2	2	4	3	4	2	3	.	+	.	.	.	.	1	.	.	.	.	.	.	.
Urtica dioica		2	+	+	2	3	2	2	+	2	.	.	+	.	.	.	.	.	.	.	.	.	.	.	.
Impatiens noli-tangere		.	+	+	1	.	+	+	+	+	+	.	.	.	.	+	.	+	.	.	.	.	1	.	+
Lysimachia vulgaris		.	+	+	+	+	+	.	+	+	.	.	.	+	+	.	.	.	+	.	+	+	1	.	+
Eupatorium cannabinum		+	1	+	+	.	1	.	1	+	.	+	.	.	.	.	+	.	.	.	+	1	+	.	+
Stachys sylvatica		1	.	+	2	+	+	1	1	+	.	.	+	.	.	.	.	+	.	.	.	.	+	.	+
Galeopsis speciosa		+	.	.	1	+	+	+	+	+	.	.	.	.	+	.	+	.	.	.	+	+	.	.	+
Lamium maculatum		+	+	.	.	1	1	2	+	1	+	.	+	.	.	.	.	.	.	1	.	.	.	.	+
Glechoma hederacea		2	+	.	+	.	+	+	+	.	1	.	.	.	+	.	.	.	.	.	.	.	.	.	+
Mentha longifolia		.	1	+	.	.	+	.	1	1	+	.	.	.	.	.	.	.	.	.	.	.	.	.	.
Scirpus sylvaticus		.	+	1	+	.	.	.	1	2	.	.	.	.	.	.	.	.	.	.	.	.	.	+	.
Alnus glutinosa	a	.	.	.	.	.	.	.	.	.	3	+	.	3	4	1	.	.	3	2	+	4	3	3	4
Carpinus betulus		.	.	.	.	+	.	.	.	.	.	1	1	2	2	3	3	2	1	1	+	.	2	+	.
Acer campestre		.	.	.	.	.	.	.	.	.	2	.	.	.	.	+	.	.	.	.	.	1	.	+	1
Acer campestre	b	.	.	.	.	.	.	.	.	.	1	+	+	+	+	+	+	+	+	+	.	+	.	1	2
Carpinus betulus		.	.	.	.	.	.	.	.	.	.	+	+	1	+	.	1	+	+	+	+	.	1	+	.
Crataegus monogyna		.	.	+	.	.	.	.	.	+	.	+	+	+	+	1	2	+	+	.	+	1	.	1	+
Ligustrum vulgare		.	.	+	.	.	.	.	.	.	.	.	.	1	+	+	1	1	+	.	+	1	1	+	.
Alnus glutinosa		.	.	.	.	.	.	.	.	.	+	.	.	.	+	.	.	.	.	.	+	+	+	+	+
Hedera helix	a	.	.	.	.	.	.	.	.	.	+	2	+	1	1	1	1	.	2	1	+	+	1	+	+
Knautia drymeia		.	.	.	.	+	.	.	+	.	.	+	+	+	+	1	1	1	+	+	.	+	1	2	+
Allium ursinum		2	.	.	.	.	.	.	.	.	1	+	+	2	1	1	.	.	5	4	.	.	+	+	1
Ligustrum vulgare		.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	1	+	+	.
Primula vulgaris		.	.	.	.	.	.	.	.	.	.	+	.	+	1	+	1	+	.	.	.	+	+	+	1
Anemone nemorosa		.	+	.	.	.	.	.	.	.	.	+	.	1	1	+	+	+	+	.	.	+	+	.	1
Cirsium oleraceum		.	.	.	.	.	.	.	.	.	+	+	+	.	1	.	+	+	+	.	.	+	+	.	1
Poa trivialis		.	+	+	.	.	.	.	.	.	+	+	+	.	.	+	.	.	.	.	+	+	+	+	+

Locality description from publication:

Tatra Mts., Biały Potok river valley

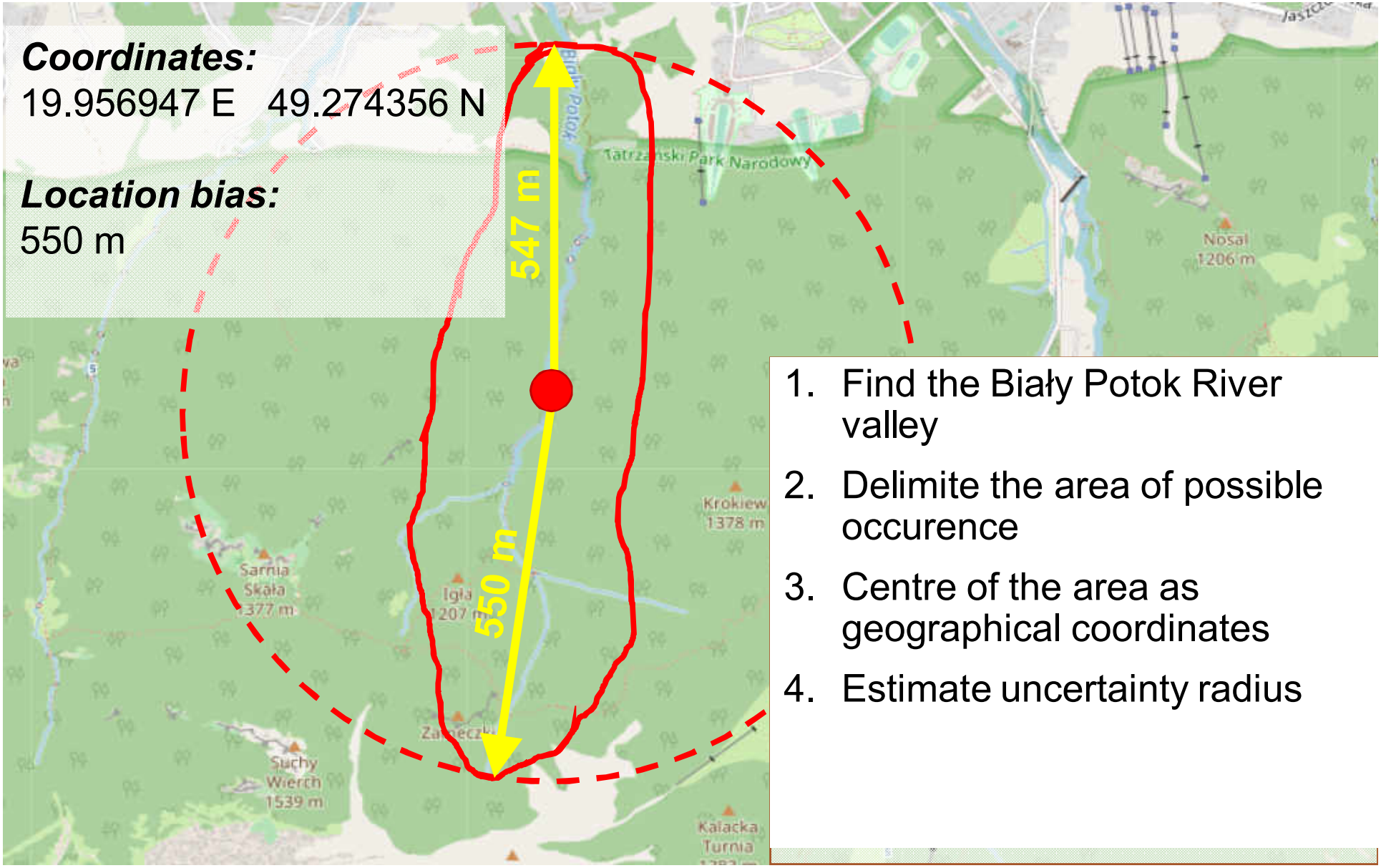


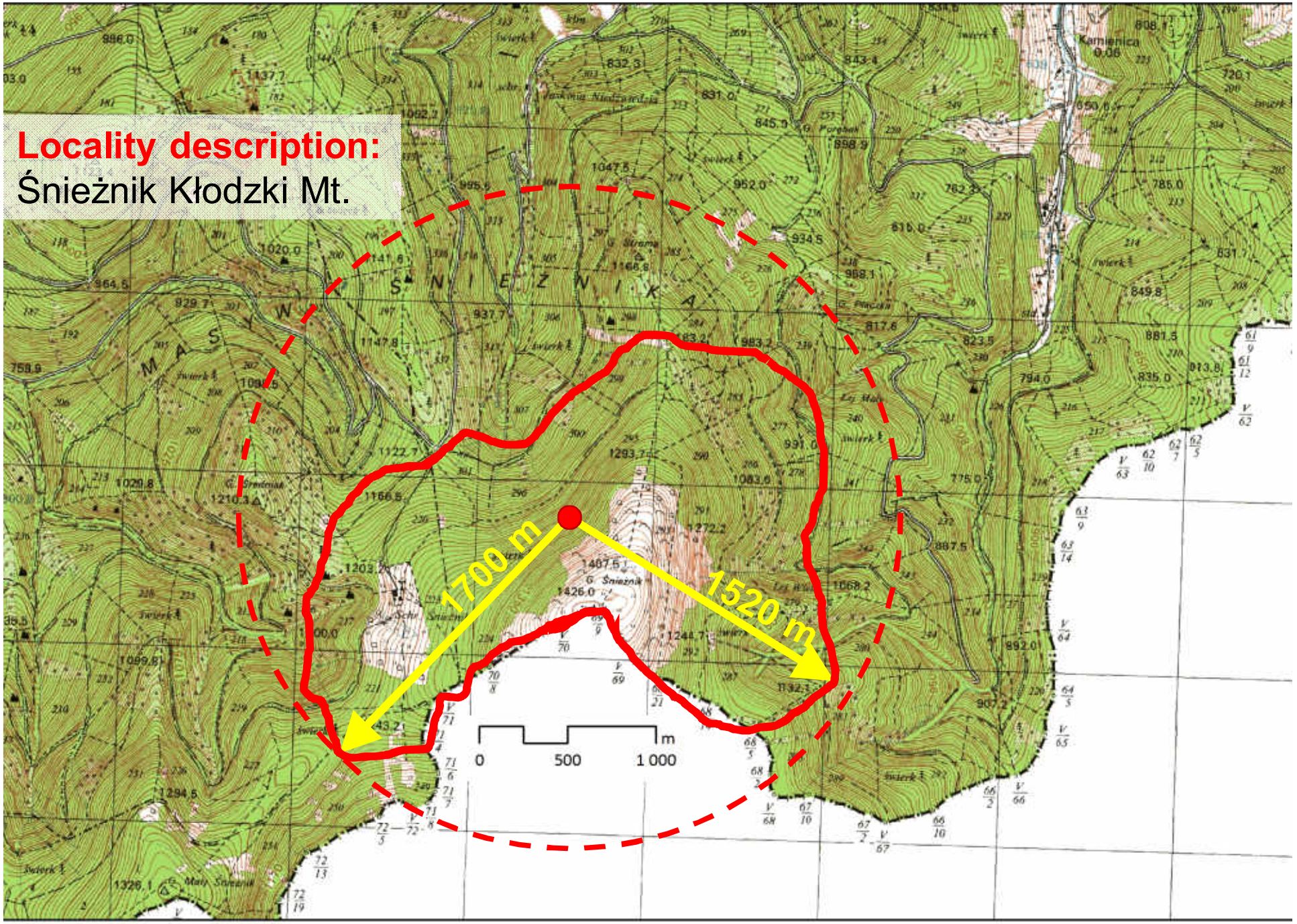
1. Find the Biały Potok River valley

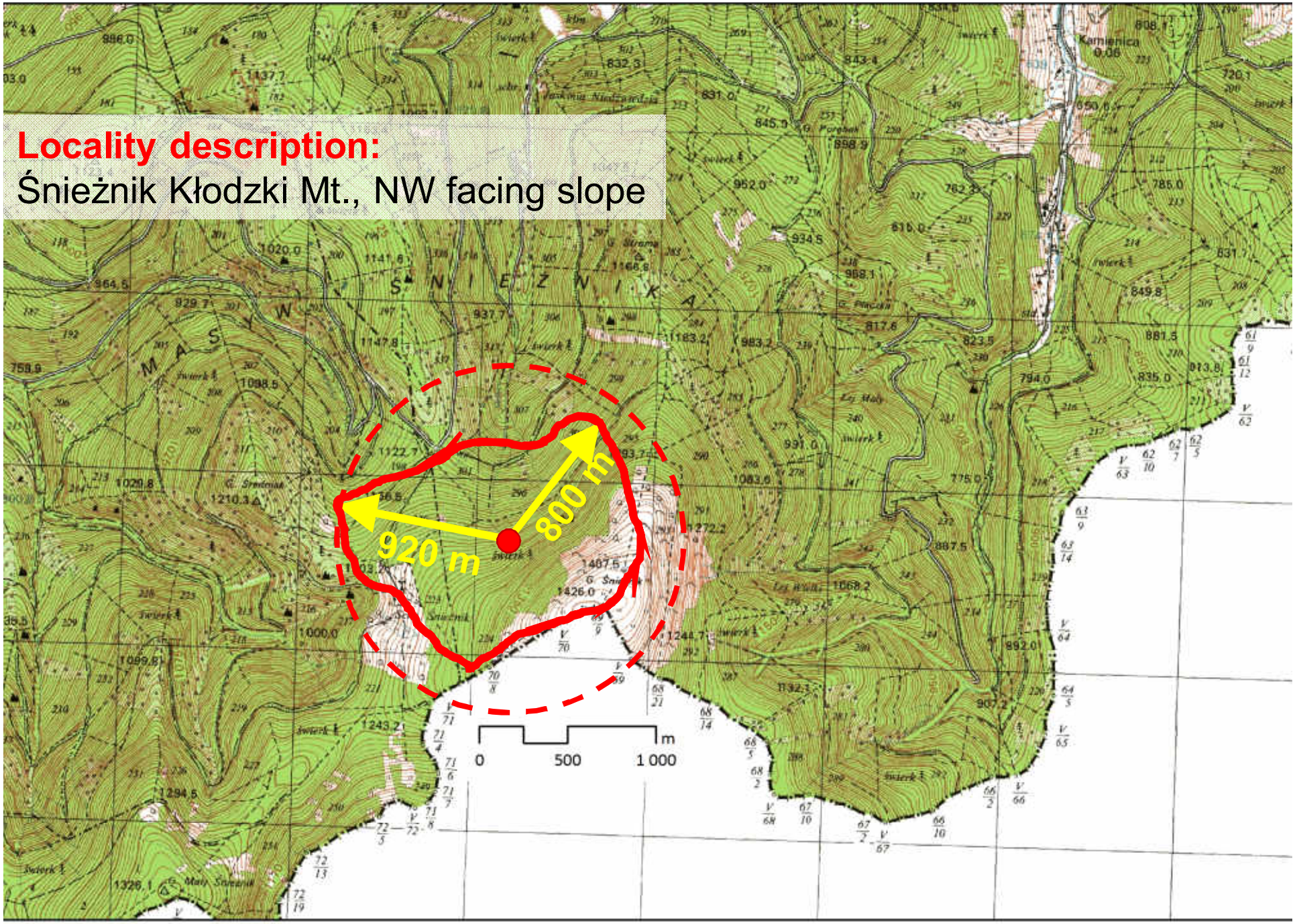


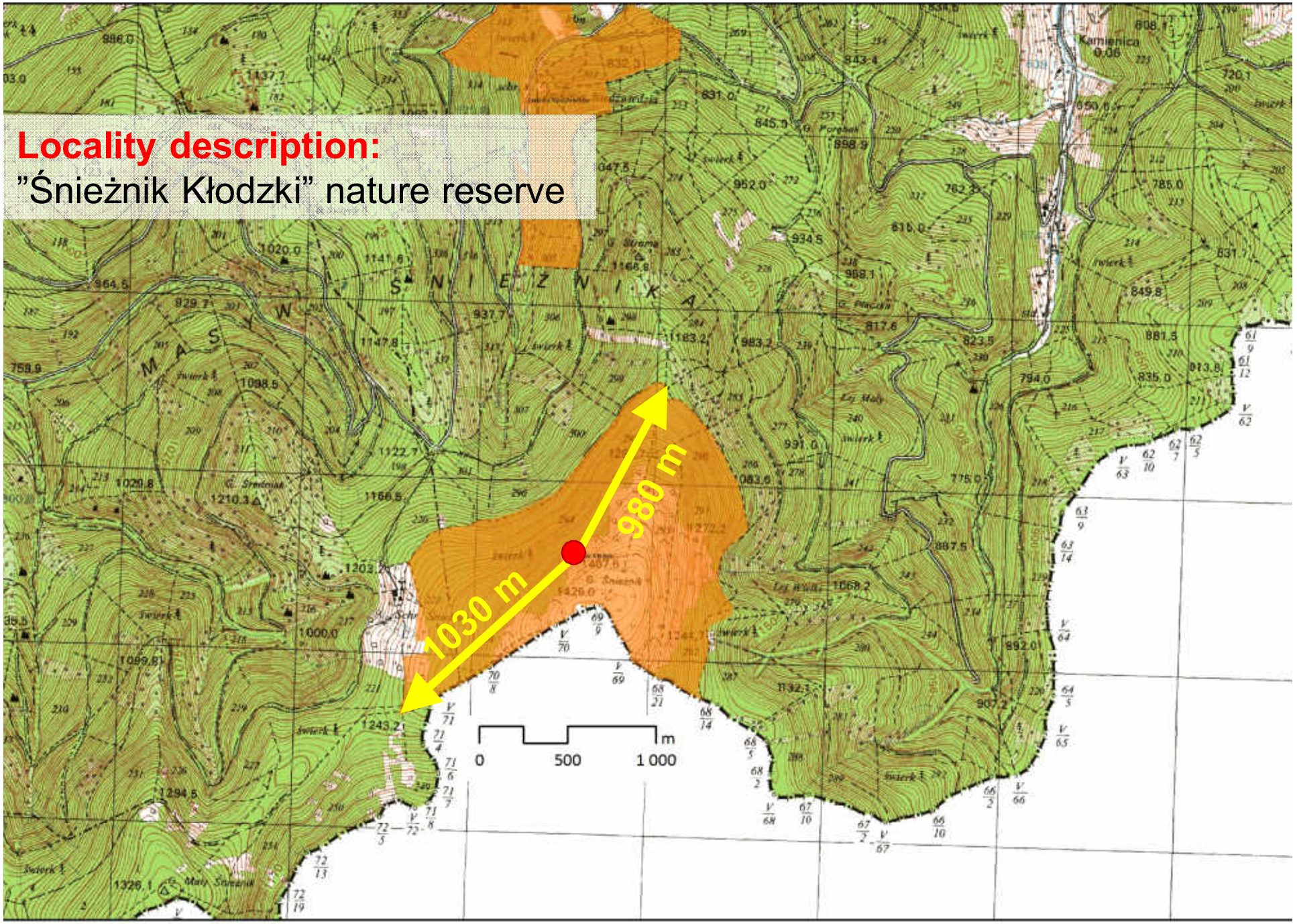
1. Find the Biały Potok River valley
2. Delimit the area of possible occurrence


Coordinates:
19.956947 E 49.274356 N
Biały Potok
Tatrzański Park Narodowy
Nosal
1206 m
Krokiew
1378 m
Sarnia
Skała
1377 m
Igła
1207 m
Suchy
Wierch
1539 m
Kalacka
Turnia
1. Find the Biały Potok River valley
2. Delimite the area of possible occurence
3. Centre of the area as geographical coordinates


Coordinates:
19.956947 E 49.274356 N
Location bias:
550 m
547 m
550 m
Biały Potok
Tatrzański Park Narodowy
Nosal
1206 m
Krokiew
1378 m
Sarnia
Skała
1377 m
Igła
1207 m
Suchy
Wierch
1539 m
Kalacka
Turnia
1. Find the Biały Potok River valley
2. Delimite the area of possible occurence
3. Centre of the area as geographical coordinates
4. Estimate uncertainty radius

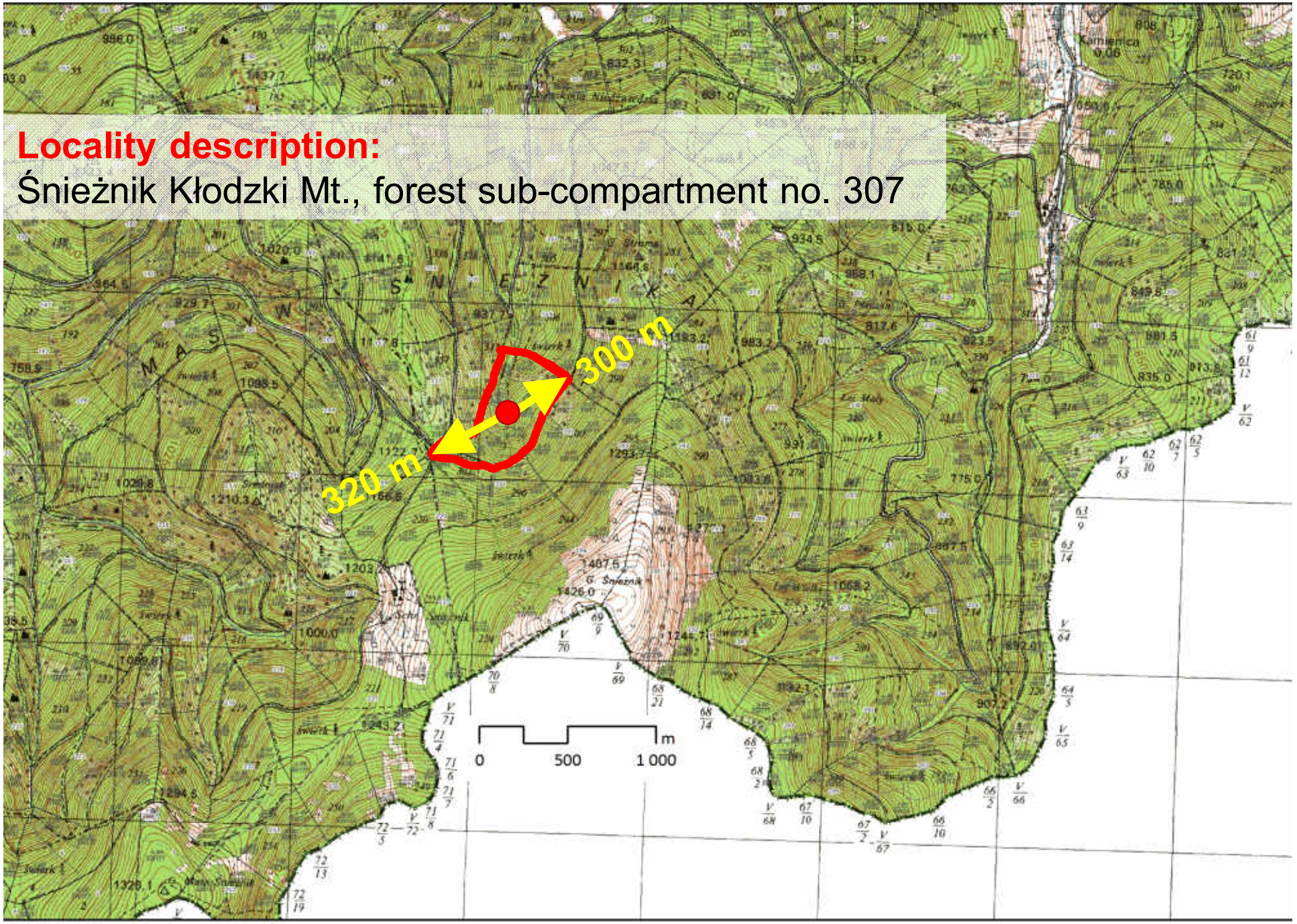

Locality description:
Śnieżnik Kłodzki Mt.
1700 m
1520 m
0
500
1 000
m


Locality description:
Śnieżnik Kłodzki Mt., NW facing slope
S N I E Ż N I K A
M A S
W
920 m
800 m
0
500
1 000
m


Locality description:
”Śnieżnik Kłodzki” nature reserve
980 m
1030 m
0
500
1 000
m

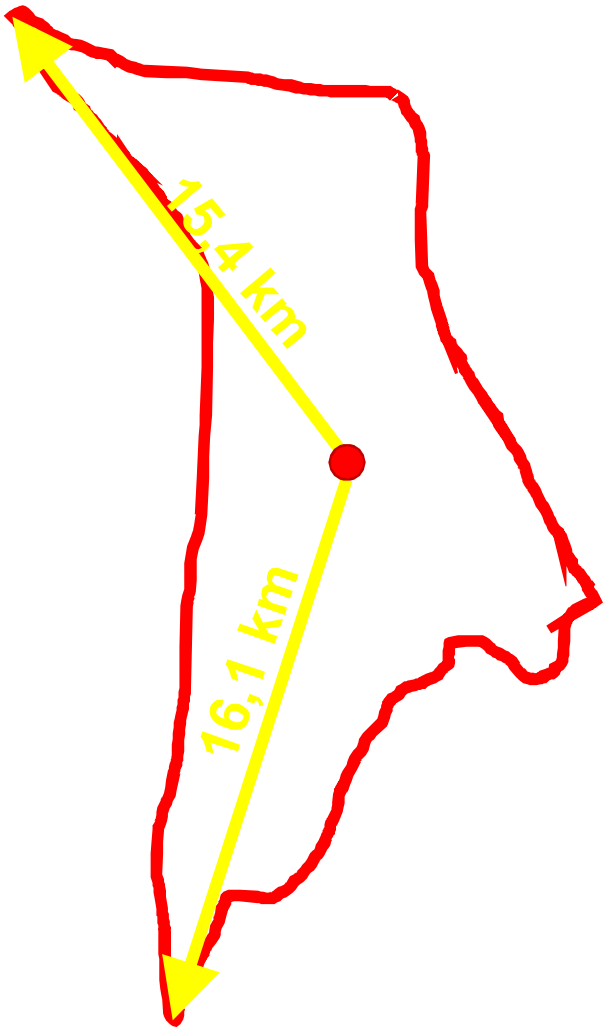
Locality description:

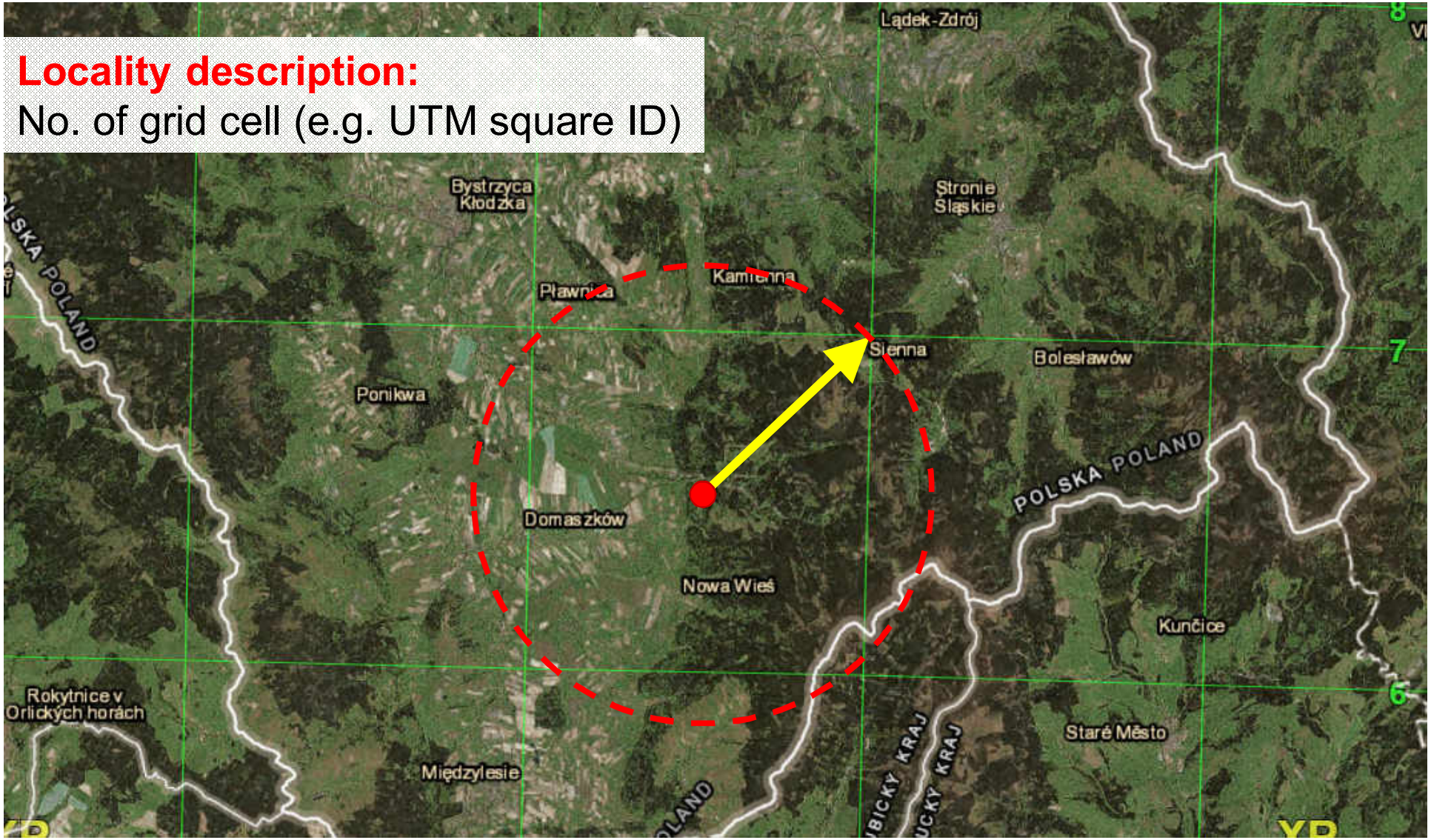
Śnieżnik Kłodzki Mt., forest sub-compartment no. 307



Locality description:

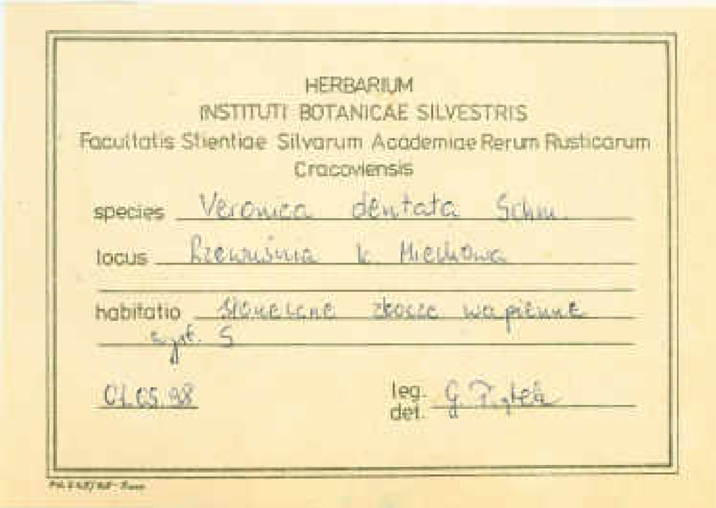
Śnieżnik Massiff




Locality description:
No. of grid cell (e.g. UTM square ID)
Lądek-Zdrój
Bystrzyca Kłodzka
Stronie Śląskie
Pławnica
Kamienna
Sienna
Bolesławów
Ponikwa
Domaszków
Nowa Wieś
Kunčice
Rokytnice v Orlických horách
Międzylesie
Staré Město
POLSKA POLAND
8
7
6

Why do we need ’Remarks’ field?

Why do we need 'Remarks' field?



HERBARIUM
INSTITUTI BOTANICAE SILVESTRIS
Facultatis Stientiae Silvarum Academiae Rerum Rusticarum
Cracoviensis

species Veronica dentata Schm.

locus Rzeczyca k. Miechowa

habitatio Słoneczne zbocze wapienne [illegible]

01.05.98

leg. [illegible]
det.

Pros and cons:

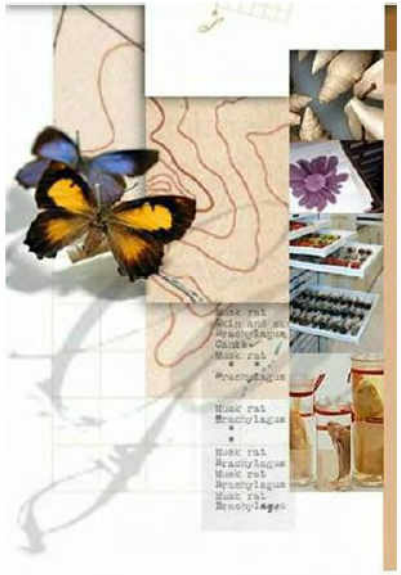
Pros:

- Best possible precision for our data
- Possibility to filter data according to spatial precision (depending on the scale of the study)

Cons:

- Time-consuming

Want more about georeferencing?



Chapman, A.D. and J. Wieczorek (eds). 2006. *Guide to Best Practices for Georeferencing*. Copenhagen: Global Biodiversity Information Facility.

cost
EUROPEAN COOPERATION IN SCIENCE & TECHNOLOGY