Towards General Theory of Raster Data Generalization

Paulo Raposo & Tim Samsonov

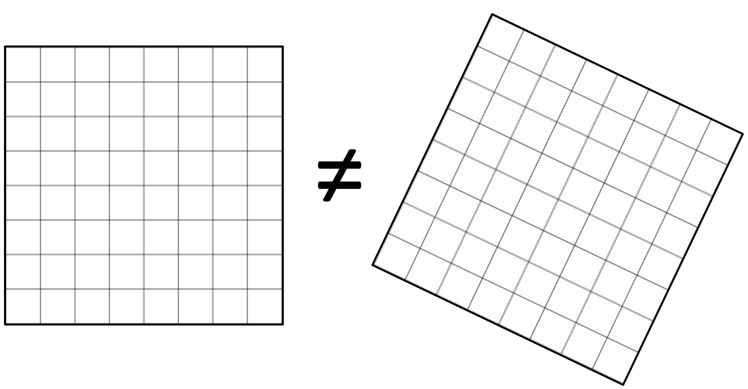
Penn State University & Moscow State University

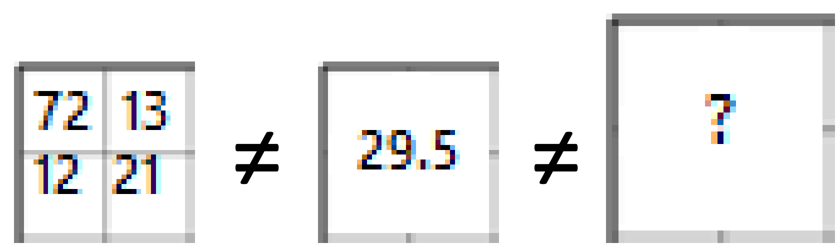
PENNSTATE 1855

FACULTY OF GEOGRAPHY
Lomonosov MSU Moscow

Simple data format, but…

assignment ambiguity & MAUP





2 processes

Cell value recalculation

Cell scaling and resampling

Framework for raster generalization

- Need a useful paradigm for geographical raster generalization
- Specialized thematic maps can’t just use resampled and averaged rasters
  e.g., climatology, oceanography, social science, vector fields, etc.

Frameworks

Cartography:

- McMaster & Monmonier 1989
- Li et al. 2001
- Peuquet 1979
- … also much DEM generalization work

Computer Science:

- Scale-space theory

M&M 1989

1. Structural generalization
   (resolution changes)
2. Numerical generalization
   (kernel-based convolution)
3. Numerical categorization
   (image classification)
4. Categorical generalization
   (kernel-based simplification of categ. data)

Scale-Space Theory

Scale-space theory is a framework for multi-scale signal representation developed by the computer vision, image processing and signal processing communities with complementary motivations from physics and biological vision. It is a formal theory for handling image structures at different scales, by representing an image as a one-parameter family of smoothed images, the scale-space representation, parametrized by the size of the smoothing kernel used for suppressing fine-scale structures.

--Wikipedia

Pointes de généralisation

(Ratajski 1967)

- Scales at which information content cannot be maintained (e.g., quantified by entropy)
- Scales at which patterns cannot be maintained (e.g., quantified by Moran's I, etc.)
- Scales at which features cannot be planimetrically represented (Nyquist-Shannon sampling theorem and resolution)
- Scales at which features must manifest at higher order (e.g., trees to forest, dunes to desert, etc.)

Unexplored raster gen topics

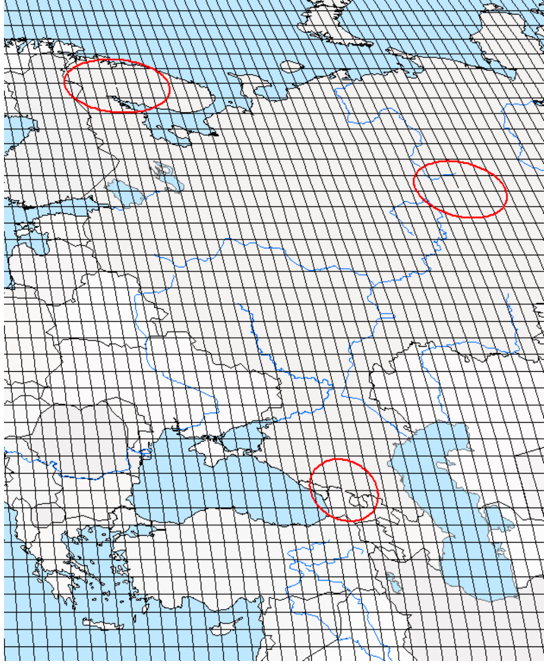
- Entropy (explorations by Bjorke, Li, Knöpfli, etc.)
- Multi-band raster generalization
- Evaluation
- Operator sequences
- Star vs. Ladder
- Projection distortions in data processing

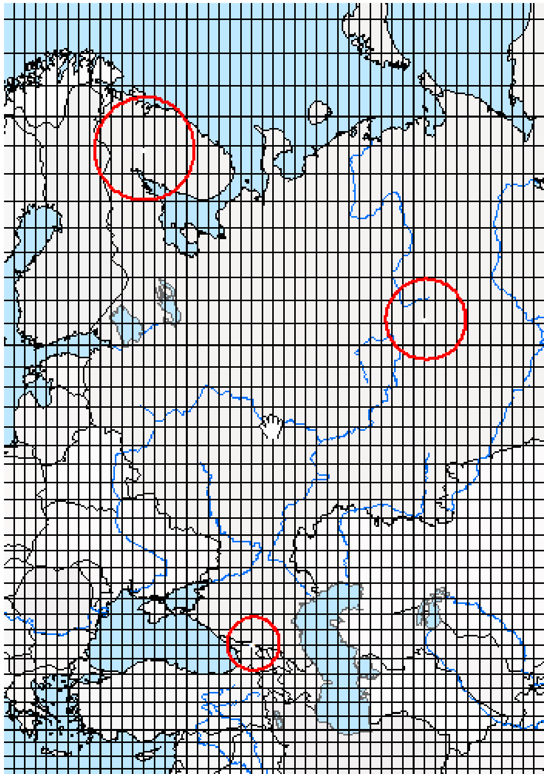
Case study

Raster processing with variable kernel shapes

Map projection distortions and generalization

- Generalization in small scales is highly affected by map projection distortions
- Various measures that are used in generalization as constraints and parameters depend on local distortion ellipse
- Small-scale generalization workflow should be aware of this issue



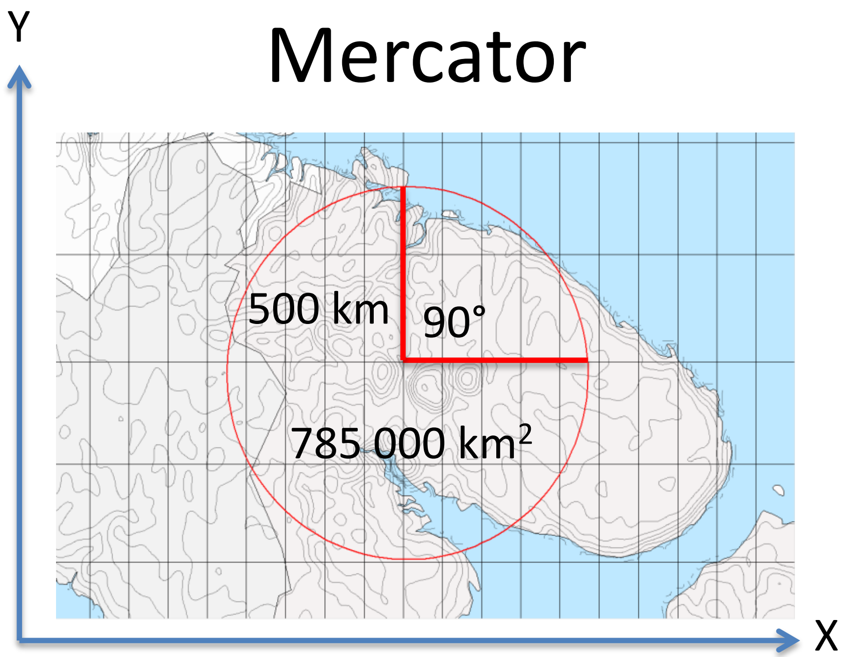


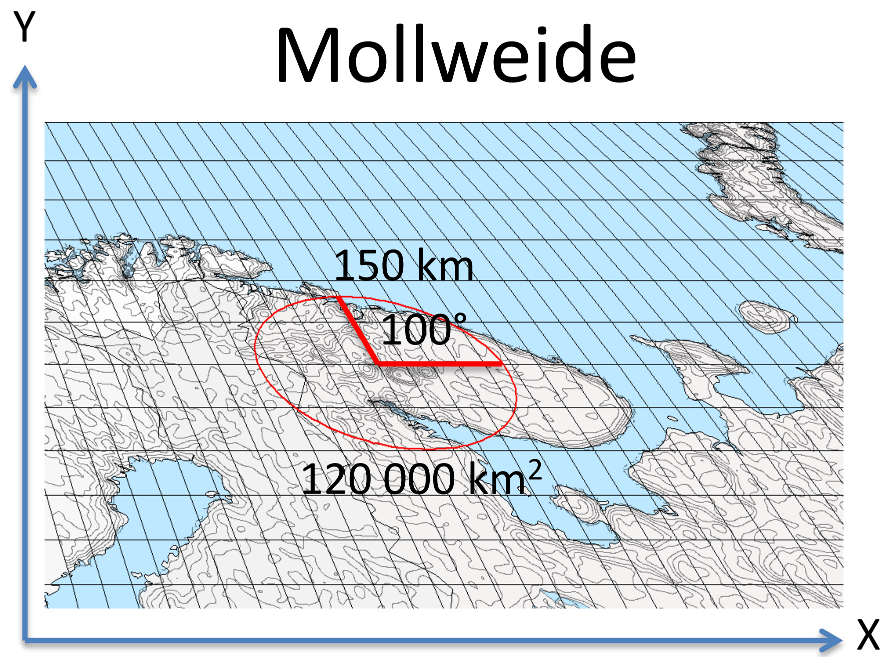
Measures

Rules

Constraints

Evaluation

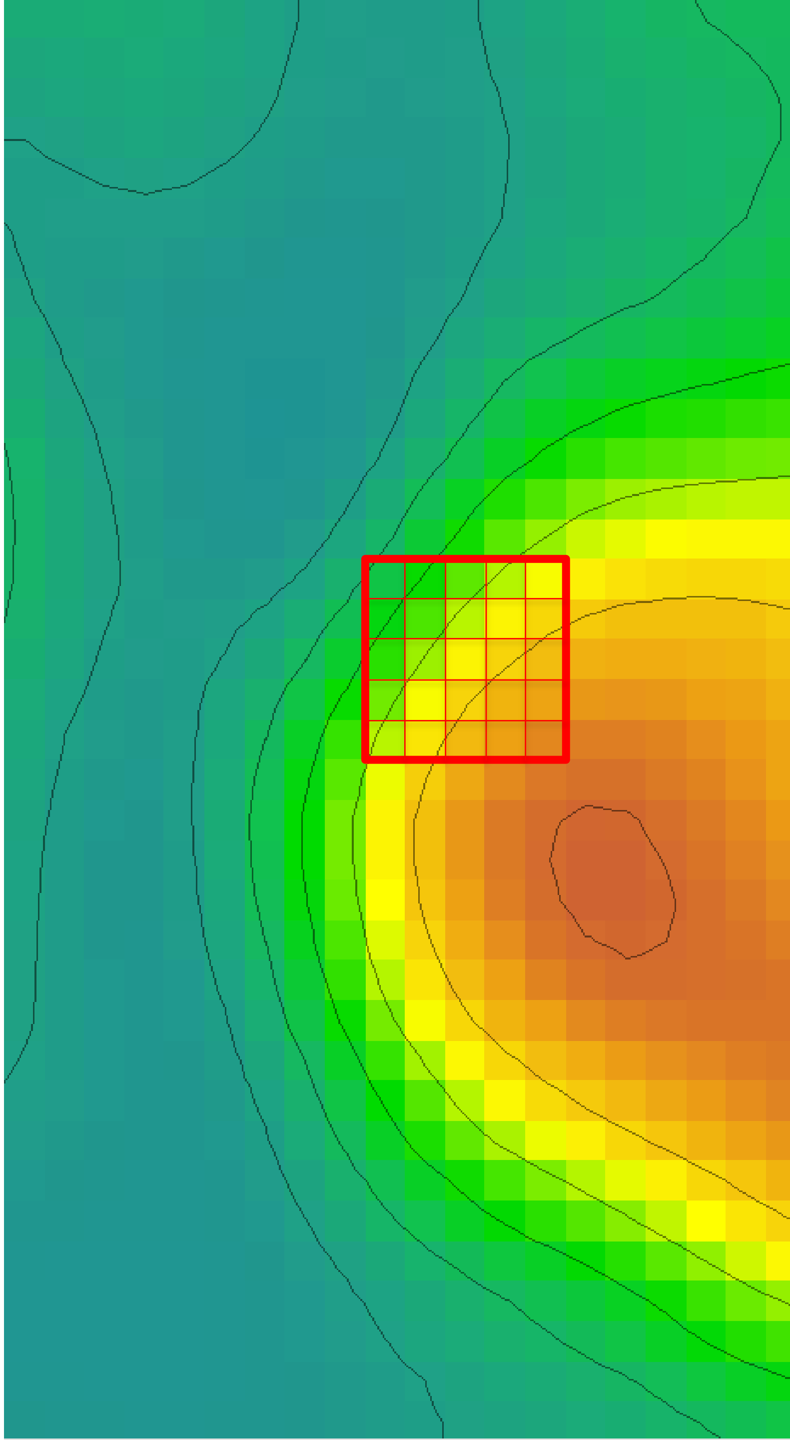




- Distances, areas and polar angles differ greatly
- Results of generalization will depend on projection

Raster processing

- **Floating window** techniques — standard for numerical generalization of rasters
- Standard issue: wrong slope angles, flattened or over-exxagerated hillshading due to map projection distortions.
- Biased raster statistics (mean, standard deviation) due to areal distortions



Floating window processing

Geometric

- Fixed tesselation neighborhood (3x3, 5x5 etc)
- Advantages:
  - Standard technique with fixed kernel can be applied
  - Quick processing
- Disadvantages
  - Incorrect calculation of derivatives

Geographic

- Fixed geographic neighborhood (variable according to local distortions)
- Advantages:
  - The same geographical neighborhood is processed everywhere
  - Correct calculations of derivatives from rasters
- Disadvantages:
  - Slow processing

Geodetic calculations?

$$m_G = \sqrt{\left(\frac{\delta G}{\delta p}\right)_0^2 m_p^2 + \left(\frac{\delta G}{\delta q}\right)_0^2 m_q^2} = \ldots = \frac{0{\cdot}41 m_z}{w(1+p^2+q^2)}, \tag{24}$$

$$m_A = \sqrt{\left(\frac{\delta A}{\delta p}\right)_0^2 m_p^2 + \left(\frac{\delta A}{\delta q}\right)_0^2 m_q^2} = \ldots = \frac{0{\cdot}41 m_z}{w\sqrt{p^2+q^2}}, \tag{25}$$

$$\begin{aligned} m_{k_h} &= \sqrt{\left(\frac{\delta k_h}{\delta r}\right)_0^2 m_r^2 + \left(\frac{\delta k_h}{\delta t}\right)_0^2 m_t^2 + \left(\frac{\delta k_h}{\delta s}\right)_0^2 m_s^2 + \left(\frac{\delta k_h}{\delta p}\right)_0^2 m_p^2 + \left(\frac{\delta k_h}{\delta q}\right)_0^2 m_q^2} = \ldots \\ \ldots &= \frac{1{\cdot}41 m_z}{w(p^2+q^2)} \sqrt{\frac{1}{1+p^2+q^2}\left\{\frac{2q^4+p^2q^2+2p^4}{2w^2} + \frac{1}{3}\left[(qs-pt)^2 + (ps-qr)^2 + \frac{(q^2r-2pqs+p^2t)(p^2+q^2)}{4(1+p^2+q^2)^2} - \frac{(q^2r-2pqs+p^2t)^2}{p^2+q^2}\right]\right\}}, \end{aligned} \tag{26}$$

$$\begin{aligned} m_{k_v} &= \sqrt{\left(\frac{\delta k_v}{\delta r}\right)_0^2 m_r^2 + \left(\frac{\delta k_v}{\delta t}\right)_0^2 m_t^2 + \left(\frac{\delta k_v}{\delta s}\right)_0^2 m_s^2 + \left(\frac{\delta k_v}{\delta p}\right)_0^2 m_p^2 + \left(\frac{\delta k_v}{\delta q}\right)_0^2 m_q^2} = \ldots \\ \ldots &= \frac{1{\cdot}41 m_z}{w(p^2+q^2)} \sqrt{\frac{1}{(1+p^2+q^2)^3}\left\{\frac{2q^4+p^2q^2+2p^4}{2w^2} + \frac{1}{3}\left[(qs+pr)^2 + (ps+qt)^2 + \frac{9(p^2r+2pqs+q^2t)(p^2+q^2)}{4(1+p^2+q^2)^2} - \frac{(p^2r+2pqs+q^2t)^2}{p^2+q^2}\right]\right\}}. \end{aligned} \tag{27}$$

NO

Workflow

1. Define the initial shape of the kernel
2. Sample raster area by the control points which are equally spaced in degrees.
3. Calculate the parameters of distortion ellipse at each point using projection equations.
4. Using distortion ellipse parameters, define the local matrix of affine transformation
5. Transform initial kernel shape and rasterize it. Round the size of the kernel to the odd number if needed.
6. For each pixel in initial raster find the closest control point and assign its number to the pixel.
7. Process the whole raster using kernels from assigned control points.

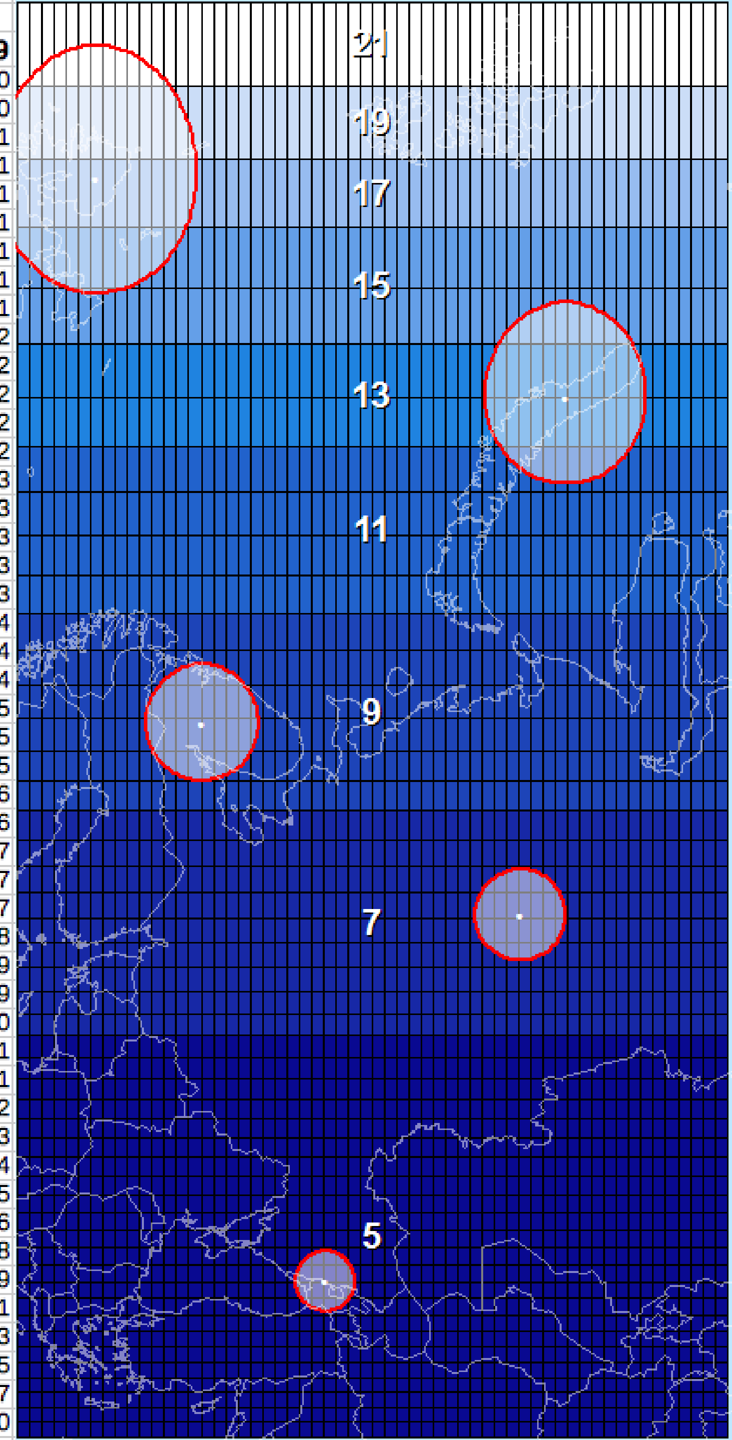
Variable kernels

Mercator Projection

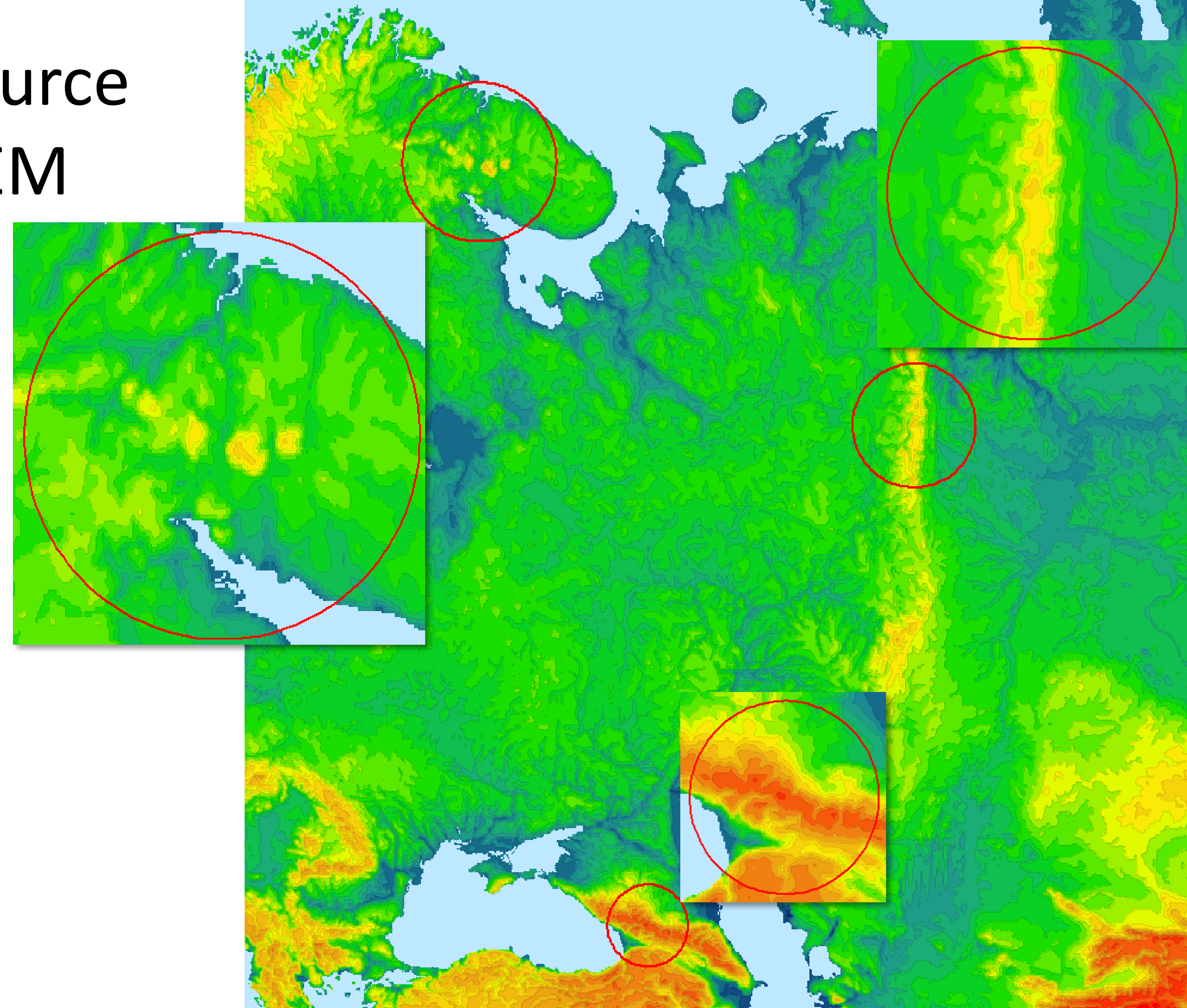
Meridian and parallel scales
$m = n = 1/\cos(B)$

Angle between meridian and parallel
$\Theta = \pi/2$

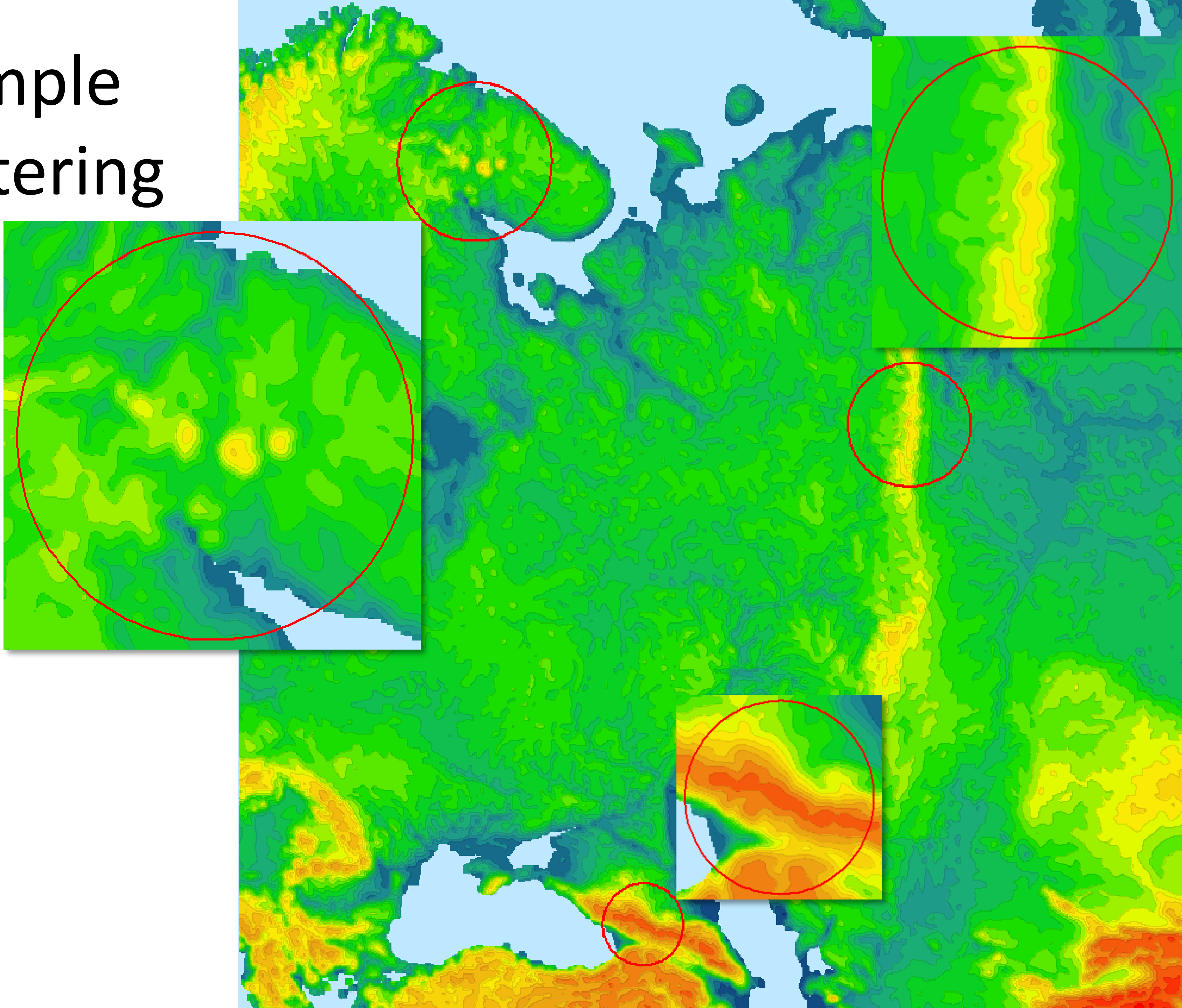
	Window Size			
Latitude	3	5	7	9
30	3	6	8	10
31	3	6	8	10
32	4	6	8	11
33	4	6	8	11
34	4	6	8	11
35	4	6	9	11
36	4	6	9	11
37	4	6	9	11
38	4	6	9	11
39	4	6	9	12
40	4	7	9	12
41	4	7	9	12
42	4	7	9	12
43	4	7	10	12
44	4	7	10	13
45	4	7	10	13
46	4	7	10	13
47	4	7	10	13
48	4	7	10	13
49	5	8	11	14
50	5	8	11	14
51	5	8	11	14
52	5	8	11	15
53	5	8	12	15
54	5	9	12	15
55	5	9	12	16
56	5	9	13	16
57	6	9	13	17
58	6	9	13	17
59	6	10	14	17
60	6	10	14	18
61	6	10	14	19
62	6	11	15	19
63	7	11	15	20
64	7	11	16	21
65	7	12	17	21
66	7	12	17	22
67	8	13	18	23
68	8	13	19	24
69	8	14	20	25
70	9	15	20	26
71	9	15	22	28
72	10	16	23	29
73	10	17	24	31
74	11	18	25	33
75	12	19	27	35
76	12	21	29	37
77	13	22	31	40



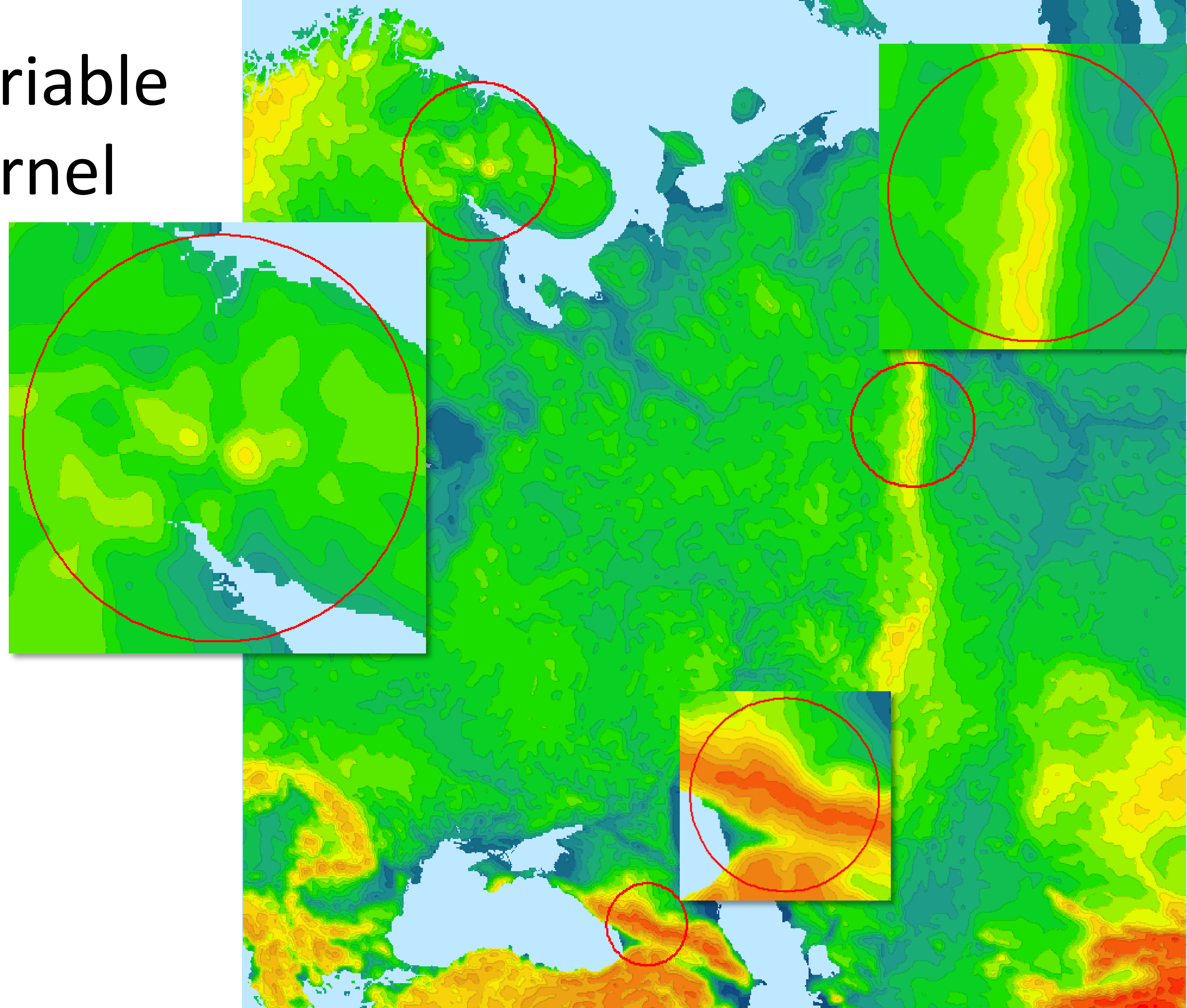
Source DEM



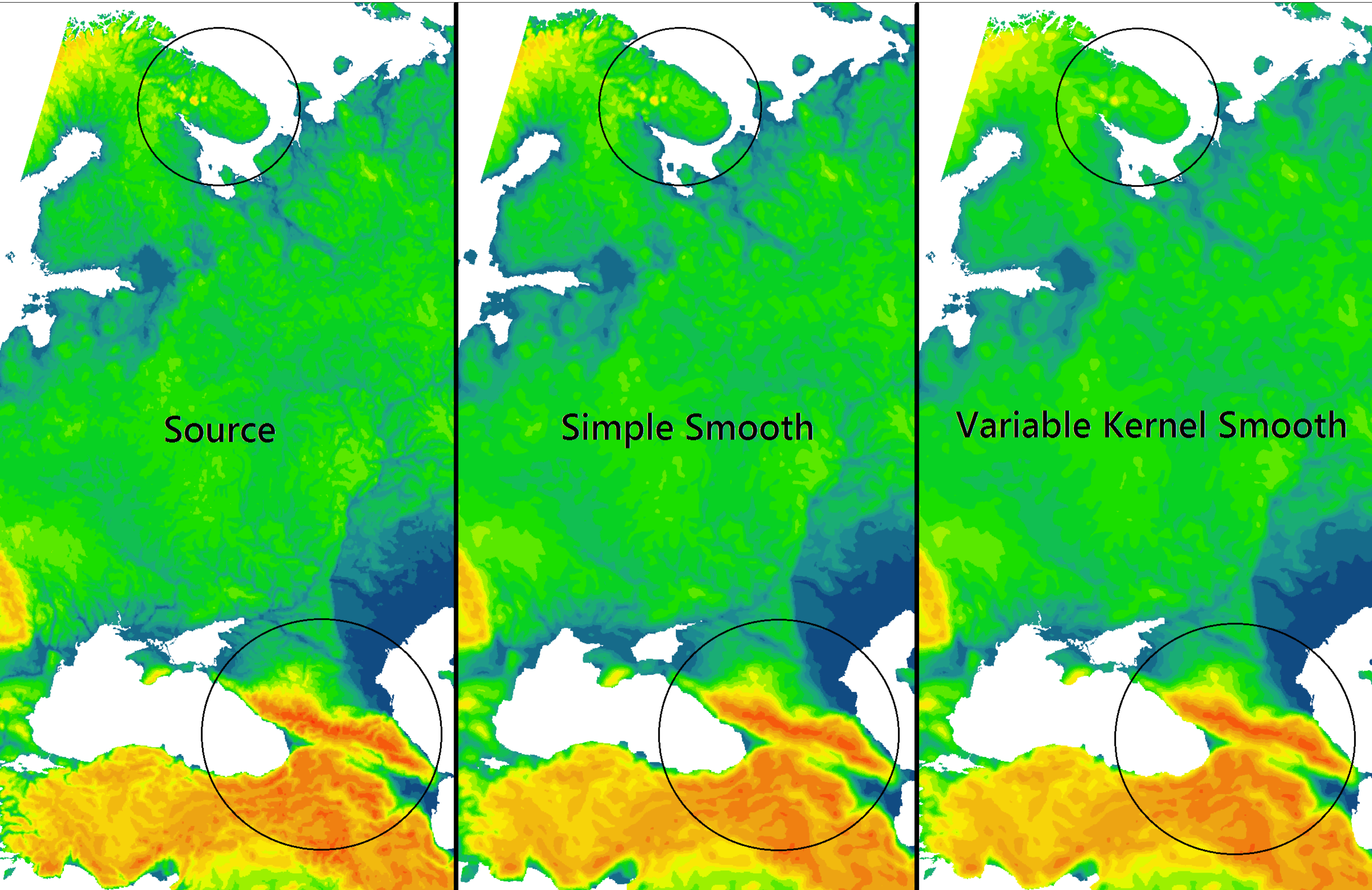
Simple Filtering



Variable Kernel

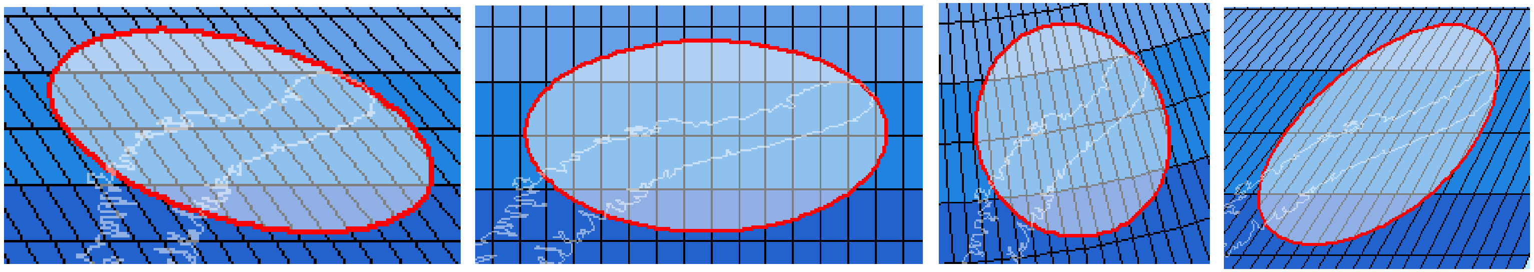


After Processing — Equal Area Conic



Conclusions

- Variable kernel shape is useful for:
  - Geographic averaging and analysis, calculation of derivatives
  - Emphasis on map projection distortions (variable detail on map)
- Future perspectives:
  - General processing framework (all projections)
  - Asymmetric filters (large geographic area)



Paulo Raposo and Tim Samsonov

TOWARDS GENERAL THEORY OF RASTER DATA GENERALIZATION

QUESTIONS?