

Analysis, Buffers and Map Algebra [2020]

GEOG 176B: Technical Issues in GIS / Spatial Data

Krzysztof Janowicz

What You Will Learn in This Lecture

- Understand **buffers** over raster data and different vector primitives.
- Understand design decisions when selecting buffers, e.g., when to **nest buffers** or **dissolve their boundaries**.
- Learn about examples of **proximity analysis**.
- Review terms such as the **scope of operations** or moving windows.
- Understand the basics of **map algebra** and how to apply it to several use cases.

Some Common Types of Overlays (Recap)

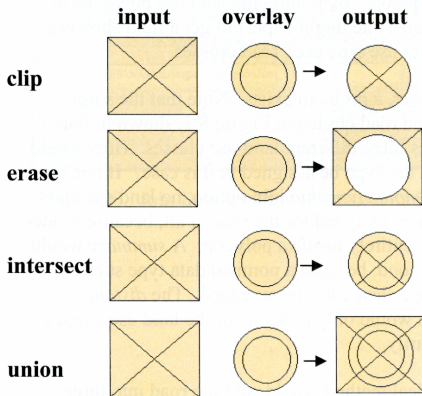
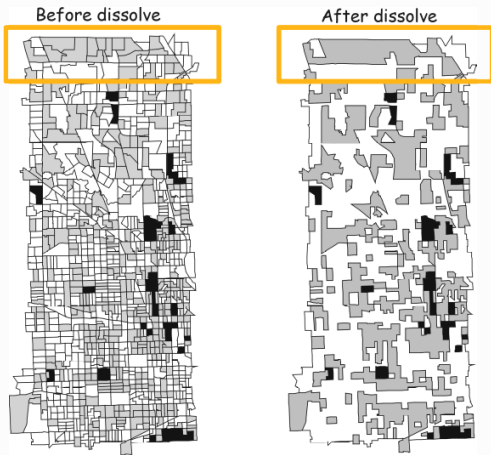


Fig. 8.4. Polygon overlay operations and results. Union and intersect join the attributes of the two input layers. Clip and erase do not.

Dissolve Operation (Recap)

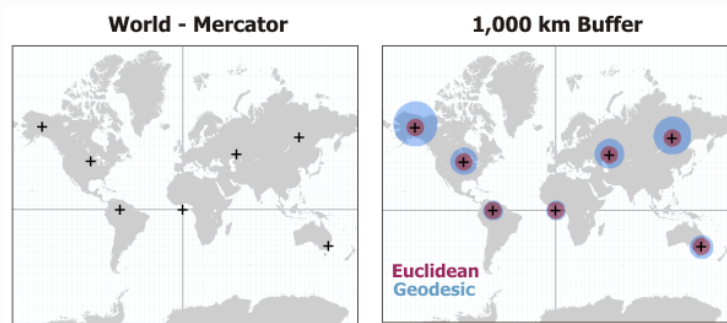


- Creation of **aggregated polygons** based on **merging adjacent polygons** that **share a common attribute value**
- In this neighborhood example, adjacent polygons are merged when they belong to the **same class**.

Buffers in GIS

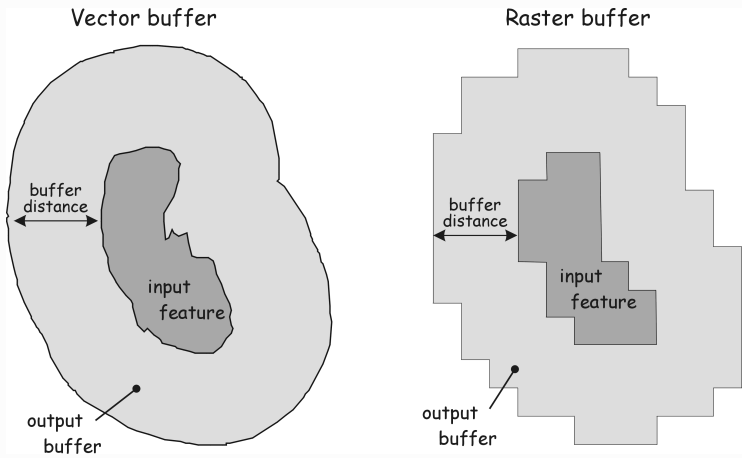
- A **buffer is a feature or zone** surrounding a input feature (or zone) that is generated using some **distance** (and distance measure).
- This distance can also be measured in **time** (or other costs).
- Buffers around **vector** geometries, e.g., points and polygons, are very common, but buffers can also be computed for **raster** datasets.
- **Buffers are areas.**
- Buffers can be **combined** when their borders are adjacent or when the buffers overlap.
- Buffers can be **nested** based on different distance bins.
- A common example for the usage of buffers is excluding zones, e.g., around a nature preserve.

Buffers Distance Methods

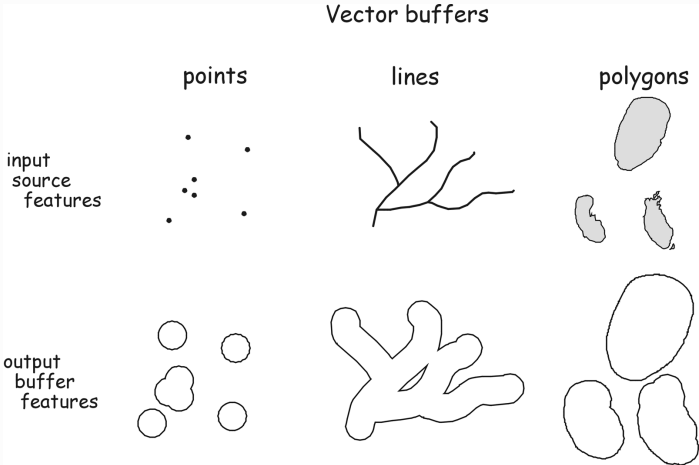


The **Geodesic buffers** (using a geographic coordinate system) the **Euclidean buffers** (using a projected coordinate system) are computed around the same cities; notice how they differ.

Vector and Raster Buffers



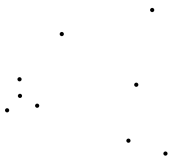
Vector Buffers



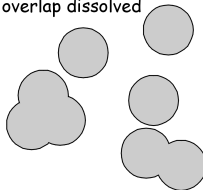
Note how all vector buffers are **polygons**.

Point Buffers

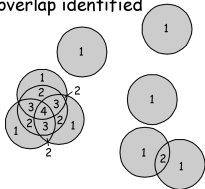
a) point layer



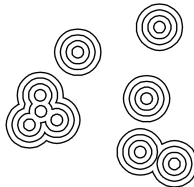
b) simple buffer, overlap dissolved



c) compound buffer, overlap identified



d) nested buffers



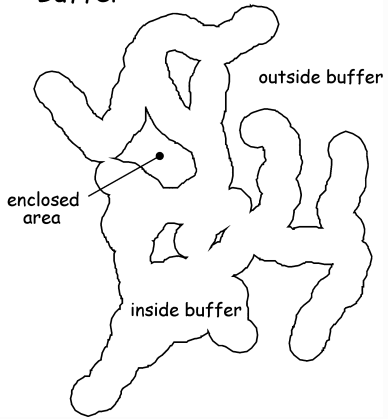
Note how c) translates to densities.

Line Buffers

Line features

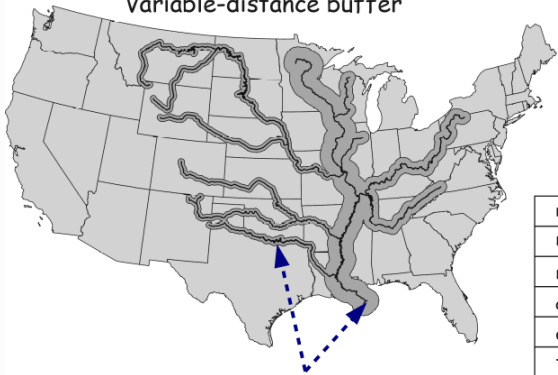


Buffer



Variable-Distance Buffers

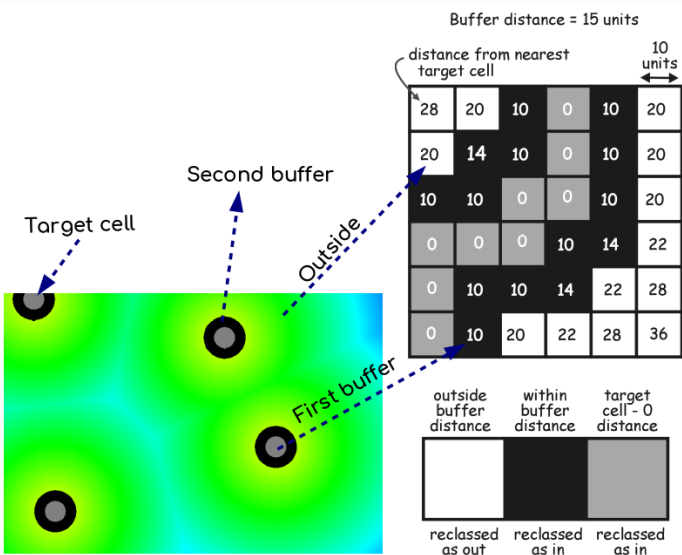
Variable-distance buffer



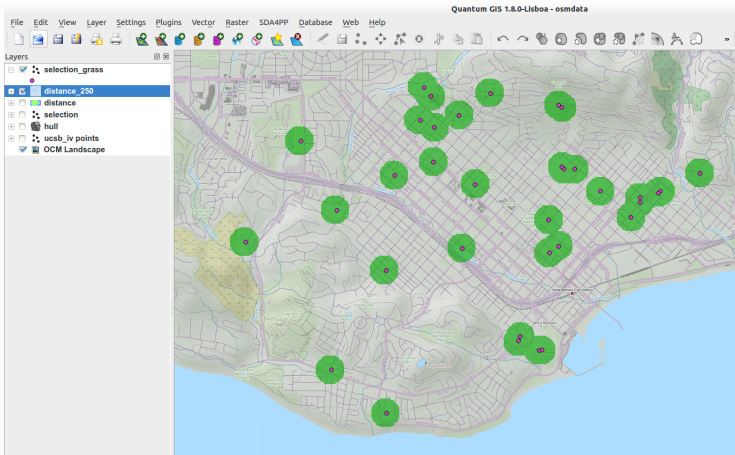
Variable buffer size based on river size

river_identifier	buffdist
mississippi	100
missouri	50
arkansas	50
ohio	75
tennessee	75
st. croix	75
illinois	75
wisconsin	75

Raster Buffers

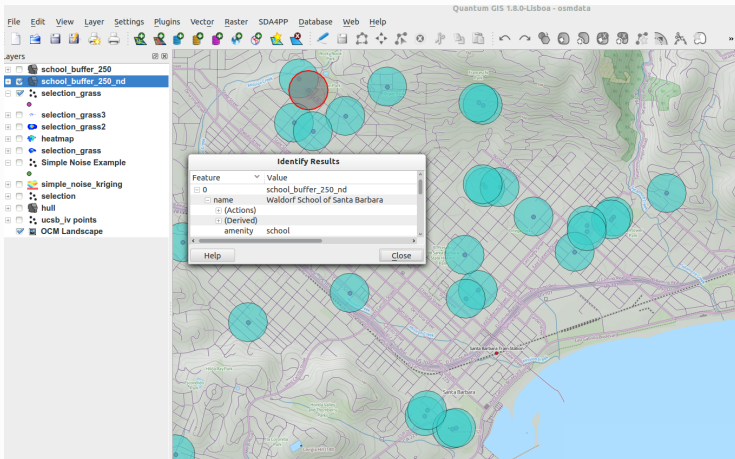


Raster Buffer Example



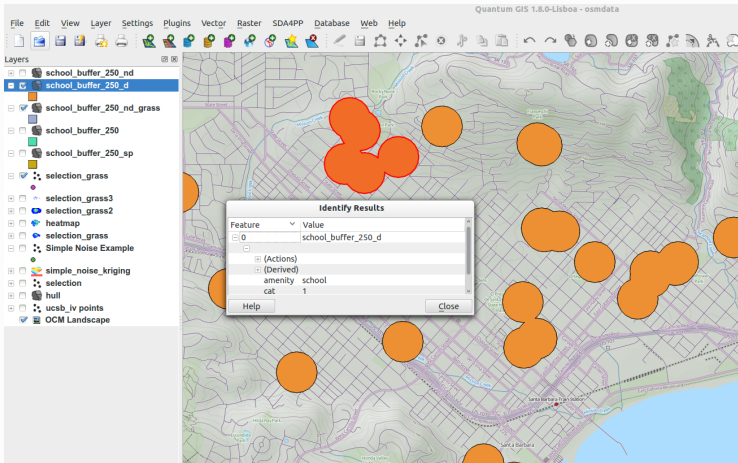
What can be asked of such raster buffer layer; when to use them?

Point-Based Vector Buffer Example



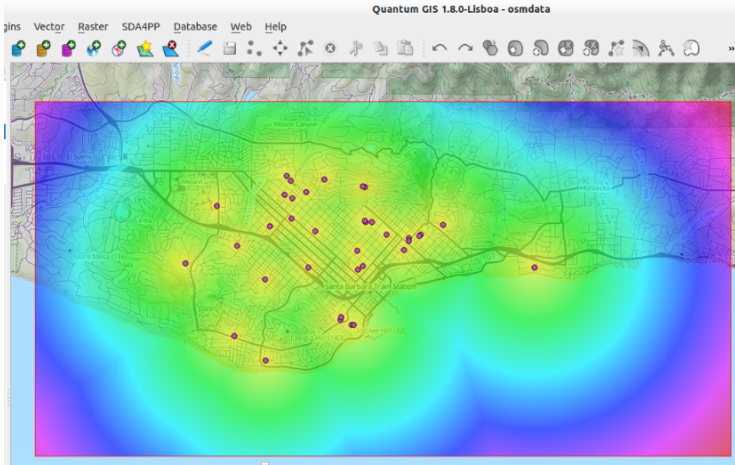
Note how this buffer is about one specific school.

Point-Based Vector Buffer Example



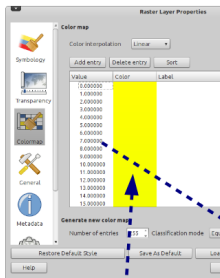
These buffers are **dissolved** using **common feature type**.

Proximity Map



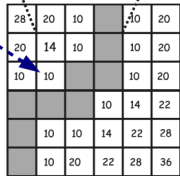
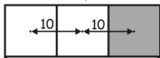
What do the **colors** mean; what are the **values** of these cells?

Proximity Map

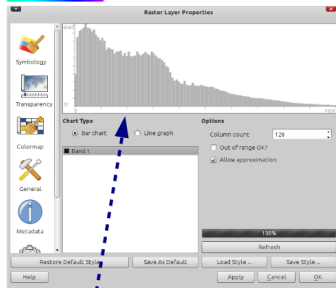


Classes

$$\text{distance} = \sqrt{x^2 + y^2}$$



distance from nearest
target cell

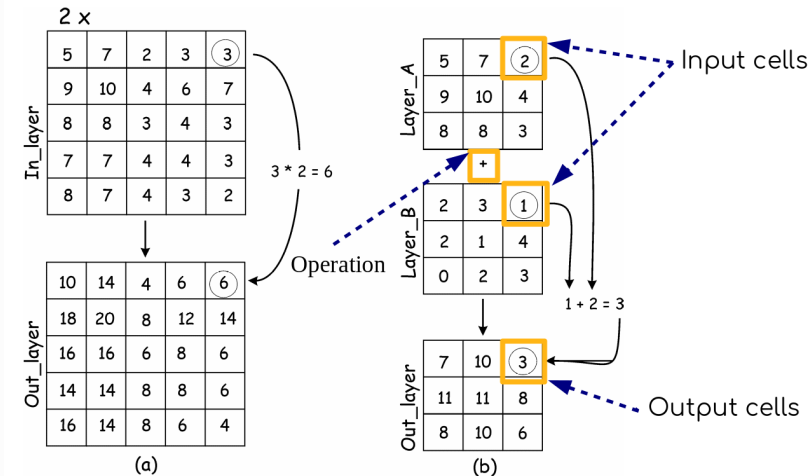


Distribution

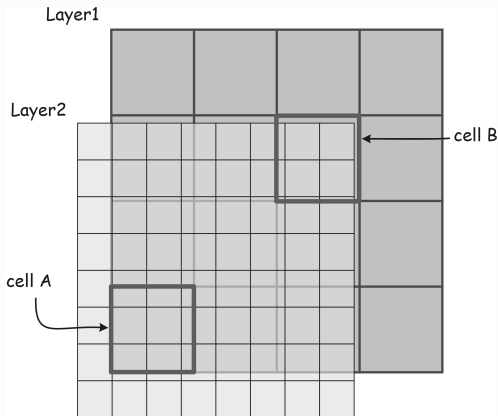
Map Algebra

- **Map algebra** is a set-based algebra introduced by Tomlin in the 1980s. It describes how to perform (and combine) primitive operations over **1..*** geographic (raster) data **layers**
- These operations often include:
 - **Arithmetic operations** such as addition or multiplication
 - **Statistical operations** such as means, maxima or minima
 - **Relational operations** such as greater than or (not) equal
 - ...
- can be used to combine expressions or perform different operations based on some conditional statement (**IF...THEN...**)
- Use cases range from everything starting from data cleaning to visualization.

Map Algebra – Combining Cell Values



Map Algebra – Requirements

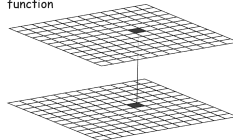


- Before combining cell values, both layers must be **geo-registered** and of the same spatial **resolution**.
- You may have to **resample** the layer.

Scopes of Raster Operations

- Similarly, to **vector** operations **raster** operations have a local, neighborhood, or global **scope**.
- **Local** and **neighborhood** scopes are most often used in raster analysis.

Local function



e.g.

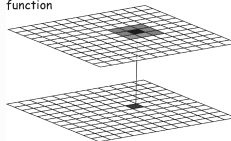
10	12	42
30	9	4
-12	8	15

plus 4

↓

14	16	46
34	13	8
-8	12	19

Neighborhood function



e.g.

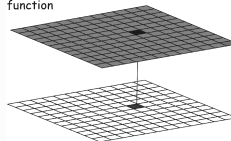
10	12	42
30	9	4
-12	8	15

neighborhood maximum

↓

33	42	42
30	42	42
30	30	17

Global function



e.g.

10	12	42
30	9	4
-12	8	15

global maximum

↓

42	42	42
42	42	42
42	42	42

Local Functions – Boolean Operators

a) **Input**

1	3	1	1
0	N	2	-1
1	2	5	0
0	1	N	N

AND

0	1	0	9
0	5	2	5
0	2	N	2
0	-3	4	8

= ?

b)

1	3	1	1
0	N	2	-1
1	2	5	0
0	1	N	N

OR

0	1	0	9
0	5	2	5
0	2	N	2
0	-3	4	8

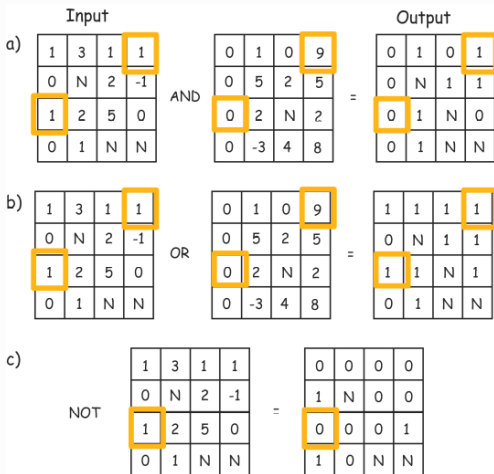
= ?

c) **NOT**

1	3	1	1
0	N	2	-1
1	2	5	0
0	1	N	N

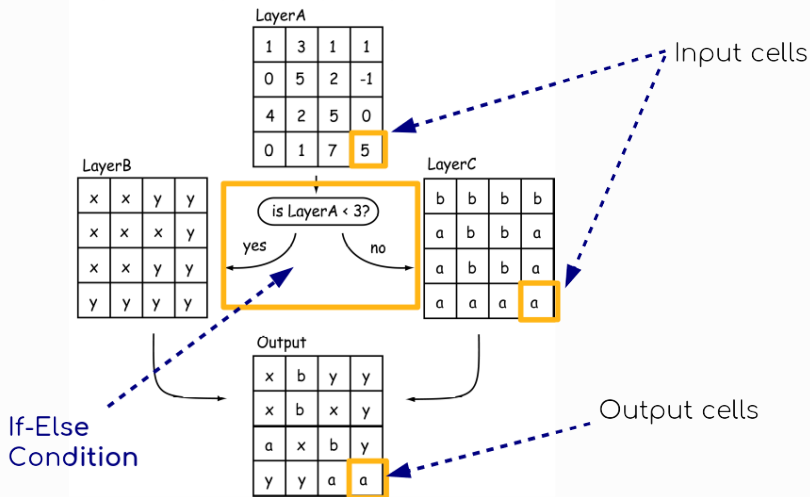
= ?

Local Functions – Boolean Operators

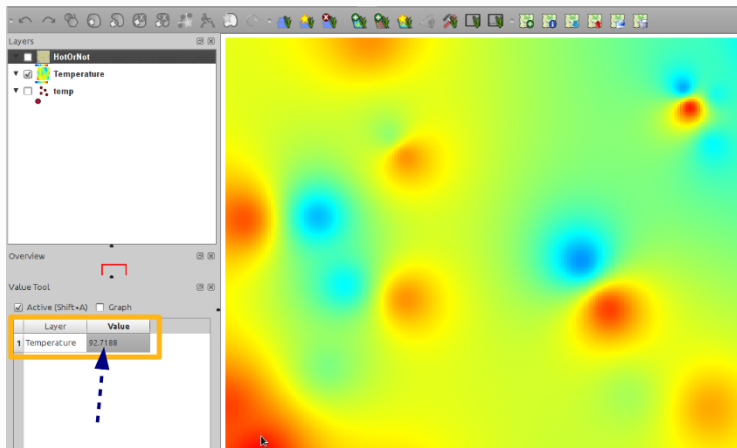


Introducing Conditions/ Control Statements

Output = CON (LayerA < 3, LayerB, LayerC)

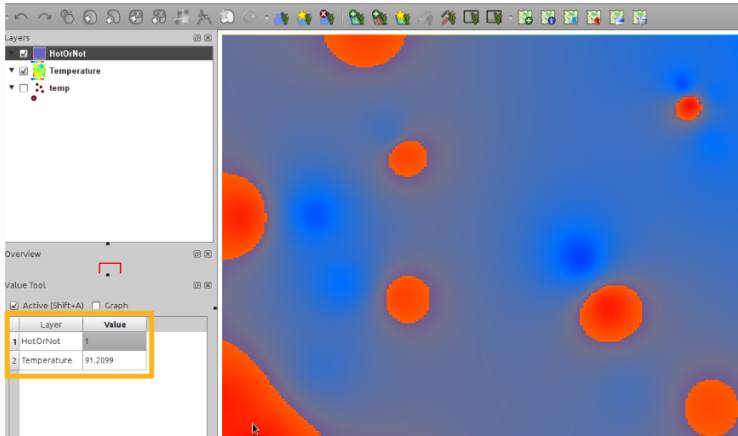


Classify Temperature Example



Based on your personal ranking, is this a very warm/hot temperature spot?

Classify Temperature Example



How to create a raster map showing hot areas based on your ranking?

Classify Temperature Example

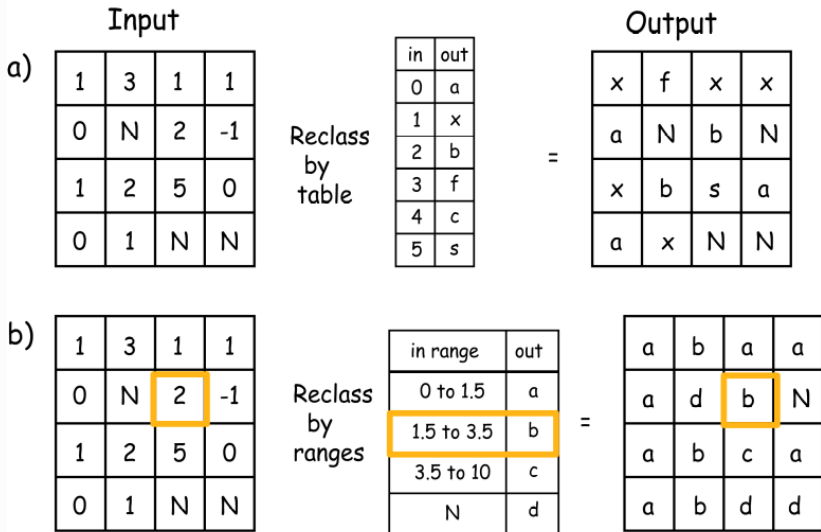
The screenshot shows the QGIS interface. On the left, the Layers panel lists 'HotOrNot' (blue), 'Temperature' (yellow), and 'temp' (red). The Overview map shows a red and blue area. The Value Tool panel shows a table with the following data:

Layer	Value
1 HotOrNot	0
2 Temperature	75.8282

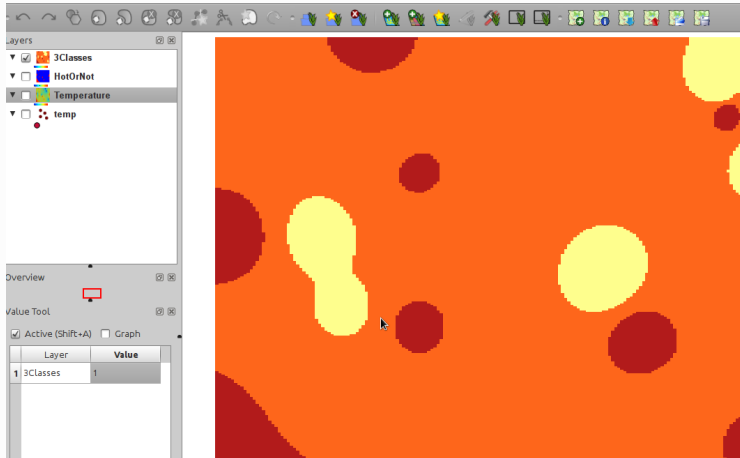
On the right, the rmapcalculator dialog box is open. The Formula field contains the expression `if(F>85,1,0)`. The Name for output raster map field contains `TBS`. Dashed blue arrows point from the text below to the formula and output name fields.

If a cell temperature value is **greater than 85**, set the cell value to **1**, else **0**.

Reclassification Operations

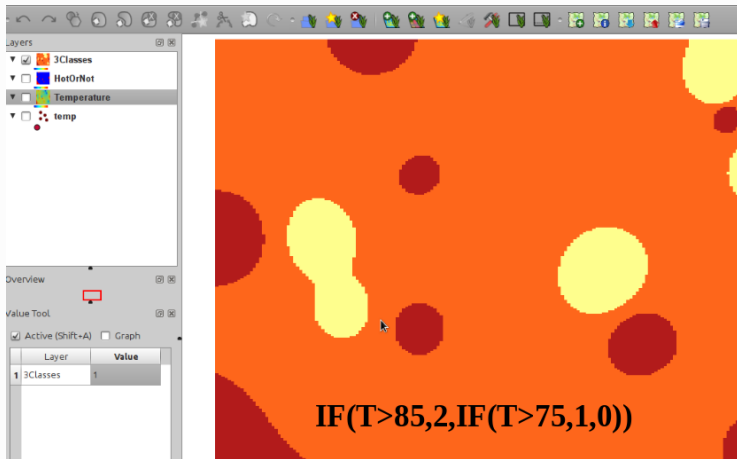


Classify Temperature Example with 3 Classes



How to create a raster map showing **hot**, **warm**, and **cold** areas ?

Classify Temperature Example with 3 Classes



How to create a raster map showing **hot**, **warm**, and **cold** areas ?

The Map Algebra Realization of a Clip Overlay

Input raster

2	2	2	8	8	2	2	2
2	2	2	8	8	8	2	2
2	3	3	3	8	8	8	7
2	3	3	3	8	8	8	7
3	3	3	6	6	6	7	7
3	3	3	3	6	6	6	7
3	6	3	6	6	6	6	6
3	6	6	6	6	6	6	6

Clip raster

0	0	0	0	1	1	1	1
0	0	0	0	1	1	1	1
0	0	1	1	1	1	1	0
0	0	1	1	1	1	0	0
0	0	1	1	1	0	0	0
0	0	1	1	0	0	0	0
0	1	0	0	0	0	0	0
0	0	0	0	0	0	0	0

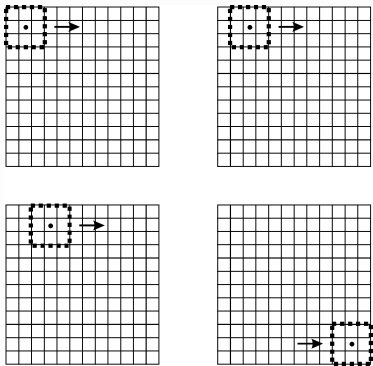
x

Output raster

0	0	0	0	8	2	2	2
0	0	0	0	8	8	2	2
0	0	3	3	8	8	8	0
0	0	3	3	3	8	0	0
0	0	3	6	6	0	0	0
0	0	3	3	0	0	0	0
0	6	0	0	0	0	0	0
0	0	0	0	0	0	0	0

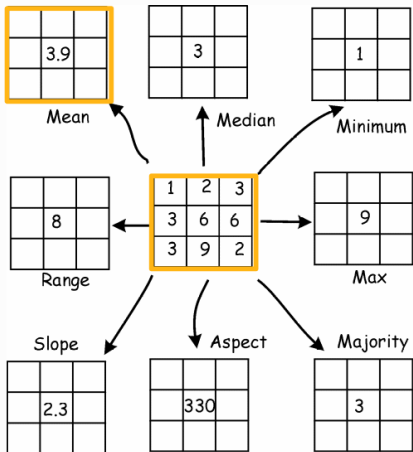
How to use just one operator to realize an **erase overlay**?

Neighborhood Functions – Moving Window



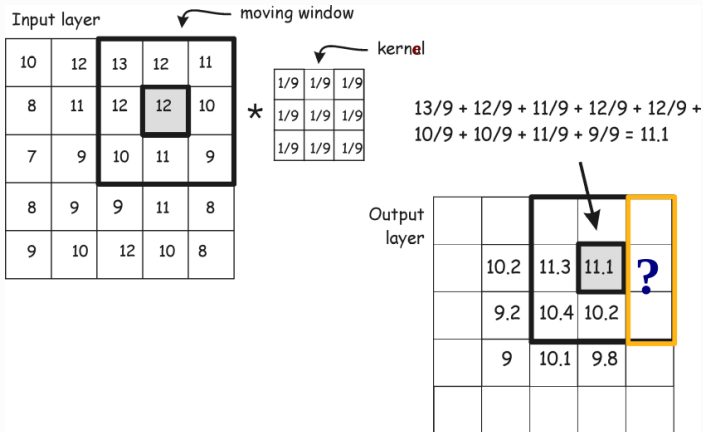
- The **value** for the cell in the **center** of each window is **computed** by all the 3x3 **neighbor cells**.
- The **window moves** cell by cell for each row and column
- **Why 3x3** or 5x5 cells but not 4x4 cells?

Moving Windows and Neighborhood Functions



How would you realize the **mean** using map algebra?

Kernels



A **kernel** is the set of **constants** for the cell values in a given **window**.

Different Kernels for Corners and Margins

Mean function kernels

corner

$\frac{1}{4}$	$\frac{1}{4}$
$\frac{1}{4}$	$\frac{1}{4}$

margin

1/6	1/6	1/6
1/6	1/6	1/6

corner

$\frac{1}{4}$	$\frac{1}{4}$
$\frac{1}{4}$	$\frac{1}{4}$

example application,
lower right corner

10	12	13	12	11
8	11	12	12	10
7	9	10	11	9
8	9	9	11	8
9	10	12	10	8

margin

1/6	1/6
1/6	1/6
1/6	1/6

main

1/9	1/9	1/9
1/9	1/9	1/9
1/9	1/9	1/9

margin

1/6	1/6
1/6	1/6
1/6	1/6

corner

$\frac{1}{4}$	$\frac{1}{4}$
$\frac{1}{4}$	$\frac{1}{4}$

margin

1/6	1/6	1/6
1/6	1/6	1/6

corner

$\frac{1}{4}$	$\frac{1}{4}$
$\frac{1}{4}$	$\frac{1}{4}$

			$9\frac{1}{4}$

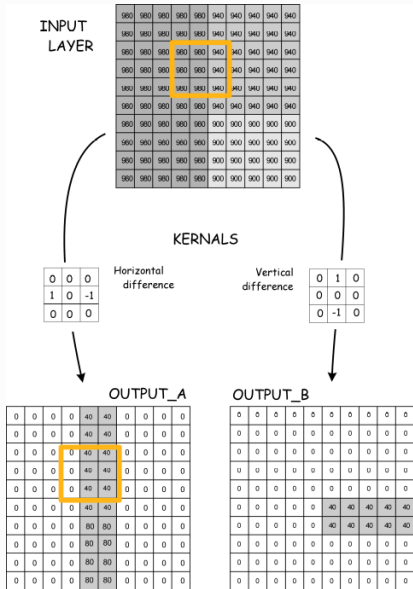
$$\frac{1}{4} \cdot 11 + \frac{1}{4} \cdot 8 + \frac{1}{4} \cdot 10 + \frac{1}{4} \cdot 8 = 9\frac{1}{4}$$

Select a different kernel window for corners or margins or a **larger study area.**

Kernel-Based Edge Detection

- Detect **abrupt changes** in cell values using a kernel.

- **Example:**
 $(980*1) + (940*-1)$
 $= 980 - 940$
 $= 40$



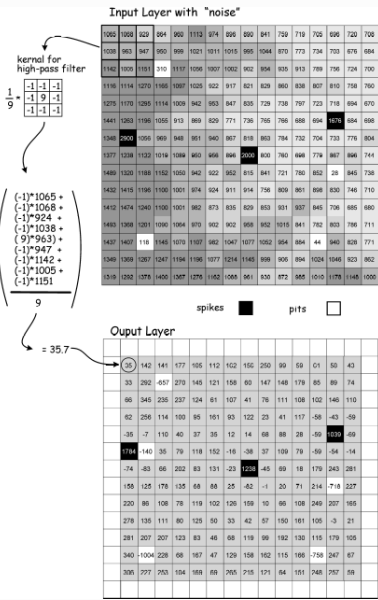
Kernel-based High-Pass Filter

100	100	100
100	100	100
100	100	100

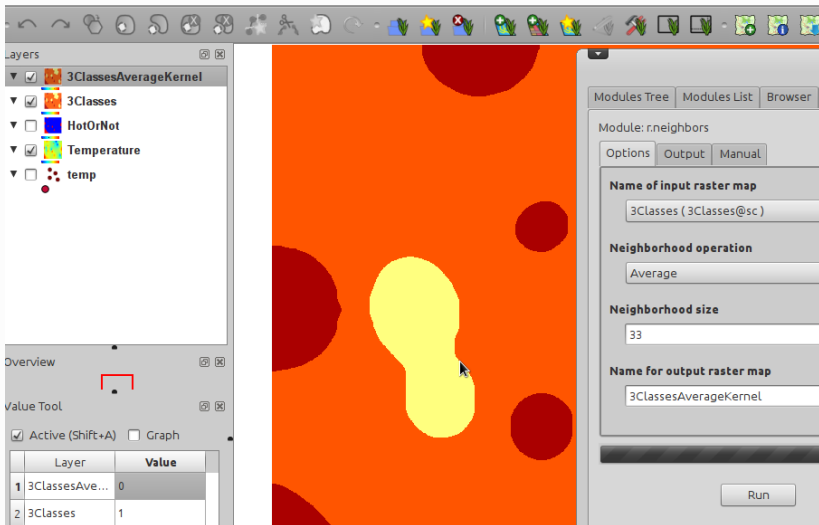
$(-800+900) / 9 = 11$

100	100	100
100	1000	100
100	100	100

$(-800+9000) / 9 = 911$



Using an Averaging Kernel for Smoothing



The screenshot displays the QGIS interface with the 'r.neighbors' module configuration window open. The main map area shows a 3-class raster map with a central yellow region and several red circular regions on an orange background. The 'r.neighbors' module is configured with the following settings:

- Module: r.neighbors
- Options tab selected
- Name of input raster map: 3Classes (3Classes@sc)
- Neighborhood operation: Average
- Neighborhood size: 33
- Name for output raster map: 3ClassesAverageKernel
- Run button

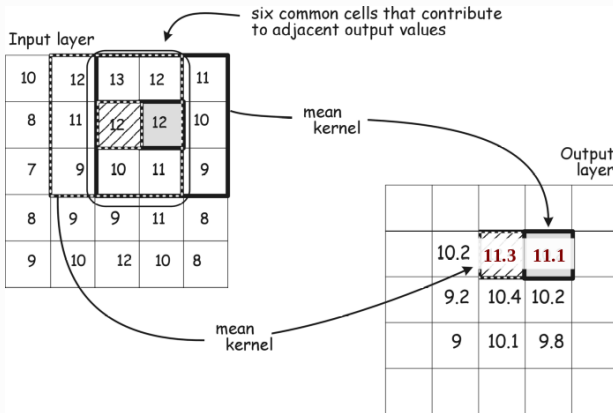
The Layers panel on the left shows the following layers:

- 3ClassesAverageKernel
- 3Classes
- HotOrNot
- Temperature
- temp

The Value Tool panel at the bottom left shows the following table:

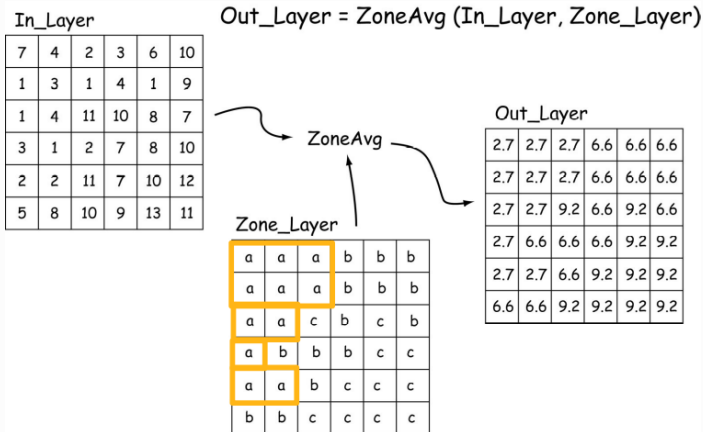
Layer	Value
1 3ClassesAve...	0
2 3Classes	1

Kernel Functions and Spatial Covariance



Keep in mind that kernel functions will **increase** the **spatial covariance** of cell values. For adjacent cells, **6 out of 9 cells** used to compute the new value will be the same.

Zonal Functions

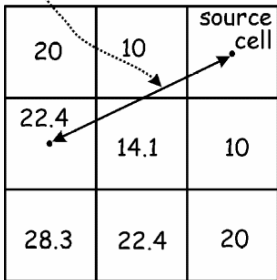


The term **neighborhood** may also be **generalized** for **regions** or zones.

Cost Surfaces As Analogy to Network Analysis

$$\text{distance} = \sqrt{(x^2 + y^2)}$$

$$\text{e.g., } D = \sqrt{(20^2 + 10^2)} \\ = 22.4$$

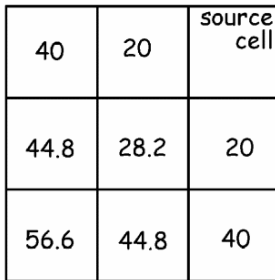


10
units

$$\text{cost} = \text{distance} * \text{fixed cost factor}$$

e.g.,

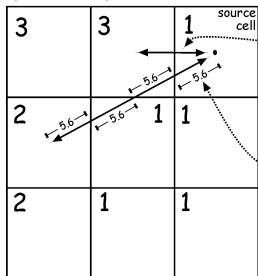
$$\text{cost} = \text{distance} * 2$$



Cost Surfaces As Analogy to Network Analysis

cost = cell distance * friction

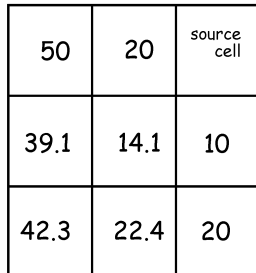
friction surface



$$\begin{aligned} \text{Cost} &= (5 * 1) \\ &+ (5 * 3) \\ &\hline &20 \end{aligned}$$

$$\begin{aligned} \text{Cost} &= (5.6 * 1) \\ &+ (5.6 * 3) \\ &+ (5.6 * 1) \\ &+ (5.6 * 2) \\ &\hline &39.1 \end{aligned}$$

output cost surface



Instead of a cost constant, **costs** can be modeled as **cell values**.

Row-Column Distance

$$\text{cost} = \text{row/column distance} * \text{friction}$$

friction surface

3	3	1 <small>source cell</small>
2		1 1
2	1	1

← 10 units →

Cost =

$$\begin{aligned}
 &(5 * 1) \\
 &+ (5 * 3) \\
 &+ (5 * 3) \\
 &+ (5 * 1) \\
 &+ (5 * 1) \\
 &+ (5 * 2) \\
 \hline
 &55
 \end{aligned}$$

output cost surface

50	20	<small>source cell</small>
55	40	10
45	30	20

← 10 units →

All edges have the same length – either **5** or **10** units.