Raster data and analysis

This is lecture 7

How do we talk about raster data analysis?
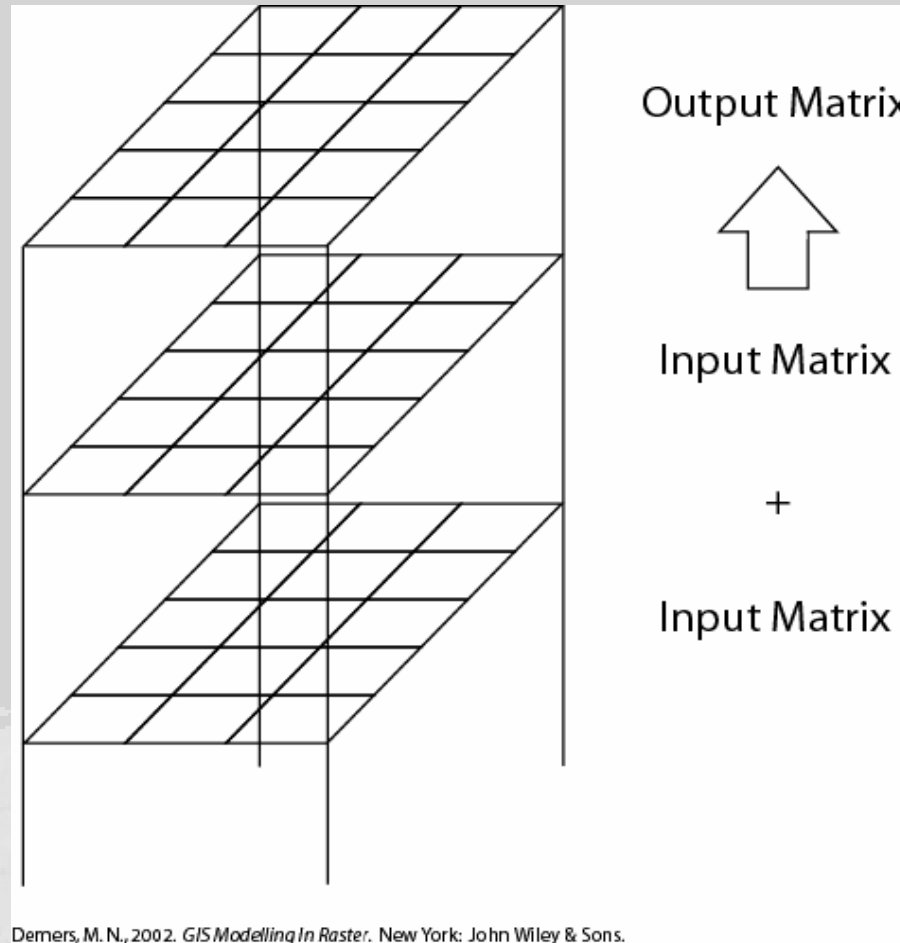
- Several methods of classifying raster analysis.
- First is based on type of operation
  (i) Boolean operations (AND, OR, NOT)
  (ii) Arithmetic operations (+, -, /, *)
  (iii) Overlay (using either arithmetic or Boolean)
  (iv) Geometric properties (such as shape, size or topology)
  (v) Geometric transformations (changes to the spatial properties such as scale changes, warping, rotation etc.)
  (vi) Geometric derivations (creating new features by using generalization, filtering or interpolation)

More common classification – based on number of cells used

- Based on local, neighbourhood, and local neighbourhood operations.

Local operations (a review)

- Based on cell by cell analysis
- Two main types of local operation:
- (i) reclassification and (ii) overlay

Output Matrix

↑

Input Matrix

+

Input Matrix

Demers, M. N., 2002. *GIS Modelling In Raster*. New York: John Wiley & Sons.

Local functions work on a cell by cell basis. The
most common is simple map algebra.

4	5	0
1	3	1
6	2	1

+

5	3	4
1	7	1
3	2	6

=

9	8	4
2	10	2
9	4	7

Demers, M. N., 2002. *GIS Modelling In Raster*. New York: John Wiley & Sons.

Uses of local functions

- Local functions for raster GIS seem simple, but are the basis of much raster analysis.

- Potential operators fall into six categories:
  (i) trigonometric;
  (ii) exponential and logarithmic (ex. exp (x));
  (iii) reclassification;
  (iv) selection (based on a condition);
  (v) statistical (ex. mean, median, mode);
  (vi) arithmetic (ex. absolute value of x)

For example…

1	0	1	1
2	4		1
1	2	4	2
2	1	4	2

Sin

=

0.8	0	0.8	0.8
0.9	-0.8		0.8
0.9	-0.8	-0.8	0.9
0.9	0.8	-0.8	0.9

☐ = No Data

Demers, M. N., 2002. *GIS Modelling In Raster*. New York: John Wiley & Sons.

How map algebra lets you use local functions

- Map algebra or performing operations on layers lets us use this wide variety of mathematical and classification tools in a local context.

- Reclassification is the most widely used local operator.

Soil Types

1	0	1	1
2	4		1
1	3	4	2
2	3	4	2

Zoning Suitability

1	0	1	1
2	4		1
1	3	4	2
2	3	4	2

Select >=2

Input Matrices

1	0	1	1
2	4		1
1	3	4	2
2	3	4	2

Output Matrices

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

Reclass:
(Selection = 1,
not = 0)

☐ = No Data

Demers, M. N., 2002. *GIS Modelling In Raster*. New York: John Wiley & Sons.

1) Soil types are compared with zoning suitability to find those

Reclassification (2)

- The whole layer does not need to be evaluated.

- A subset of a layer can be selected for reclassification.

- You can choose a subset of cell values to reclassify all the cells that represent high level of urbanization or all cells that represent agricultural land use.

Input Matrix

3	5	7	2
6	3	4	9
6	7	9	2
5	6	2	7

Select:
Values = 3:6

Output Matrix

3	5	7	2
6	3	4	9
6	7	9	2
5	6	2	7

☐ = No Data

Demers, M. N., 2002. *GIS Modelling in Raster*. New York: John Wiley & Sons.

Reclassification (3)

- Another approach is to use position as a basis for subset selection of a layer for reclassification.

- For instance, a 2x2 or 3x3 filter can be delineated, and cells within the filter reclassified.

Input Matrix

3	5	7	2
6	3	4	9
6	7	9	2
5	6	2	7

Select:
Center (2x2)

Output Matrix

3	5	7	2
6	3	4	9
6	7	9	2
5	6	2	7

Demers, M. N., 2002. *GIS Modelling in Raster*. New York: John Wiley & Sons.

Reclassification (4)

- Statistical local functions are useful for comparing two or more input grids. You can return minimum, maximum, mean, median etc. values for each grid cell.

- This allows you to compare layers to other constants or an indicator value.

Input Matrix

5	6	2	1
5	4		1
6	7	2	3
9	5	3	7

Output Matrix

4.5	6	1	3
6	4		1
6	4.5	2	2.5
7.5	5	5.5	4

Mean =

Input Matrix

4	6	0	5
7	4		1
6	2	4	2
6	5	8	1

Demers, M.N., 2002. *GIS Modelling In Raster*. New York: John Wiley & Sons.

Local neighbourhood functions

- These are also called focal functions.
- Local neighbourhood functions examine the target cell and immediate neighbours.
- A value is returned based on their combined evaluation.
- Neighbourhoods can be defined by rectangles, circles, wedges, doughnut shapes (annulus) etc.

Neighbourhood (Rectangle, 3, 3)

= Focal cell

Demers, M. N., 2002. *GIS Modelling In Raster*. New York: John Wiley & Sons.

Other types of neighbourhoods

- A doughnut shape can be drawn around a focal cell, and the values of the surrounding shape (referred to an annulus neighbourhood) are used to reclassify the focal cell.

Input Matrix

4	6	2	5
7	4	7	1
2	5	4	2
6	5	8	1

FOCALSUM
(grid, annulus)

Output Matrix

4	6	2	5
7	37	7	1
2	5	4	2
6	5	8	1

Demers, M.N., 2002. *GIS Modelling In Raster*. New York: John Wiley & Sons.

Local Neighbourhood Functions

- Local neighbourhood operations can be based on operators such as Min, Max, Mean, Average etc.

FOCALMAJORITY (Grid, Neighbourhood, Rectangle, 3, 3)

Input Matrix

4	7	2	1	9
7	2	3	2	7
3	2	5	3	5
4	1	2	2	4
9	5	4	6	2

Output Matrix

4	7	2	1	9
7	2	3	2	7
3	2	2	3	5
4	1	2	2	4
9	5	4	6	2

FOCALMIN (Grid, Neighbourhood, Rectangle, 3, 3)

Input Matrix

4	7	2	1	9
7	2	3	2	7
3	2	5	3	5
4	1	2	2	4
9	5	4	6	2

Output Matrix

4	7	2	1	9
7	2	3	2	7
3	2	1	3	5
4	1	2	2	4
9	5	4	6	2

FOCALMEAN (Grid, Neighbourhood, Rectangle, 3, 3)

Input Matrix

4	7	2	1	9
7	2	3	2	7
3	2	5	3	5
4	1	2	2	4

Output Matrix

4	7	2	1	9
7	2	3	2	7
3	2	2.4	3	5
4	1	2	2	4

Neighbourhood operations

- Just like local neighbourhood operators, neighbourhood operations (or zonal operations) use surrounding cells to reassign values. Often they deal with regions or zones.

- These are defined in the raster GIS context as groups of cells that share the same values (remember 354 definition of zones or classes).

- All the same operators/operations apply; they just take into account a wider range of cells.

Input Matrix 1
Zonal Grid

4	4	2	2
7	7	7	2
7	7	2	2
2	2	2	2

Input Matrix 2
Value Grid

4	6	2	5
3	4	9	1
2	5	4	2
6	5	8	1

ZONALMAX
(Zonal grid,
Value grid)   =

Output Matrix

6	6	8	8
9	9	9	8
9	9	8	8
8	8	8	8

Demers, M. N., 2002. *GIS Modelling in Raster*. New York: John Wiley & Sons.

Input Matrix

4	4	2	2
7	7	7	2
7	7	2	2
2	2	2	2

Zonal Grid

ZONALAREA =

Output Matrix

200	200	900	900
500	500	500	900
500	500	900	900
900	900	900	900

Zone	Cell Size	Cell Count	Area
2	10 m²	9	900 m
4	10 m²	2	200 m
7	10 m²	5	500 m

Demers, M. N., 2002. *GIS Modelling In Raster*. New York: John Wiley & Sons.

Distance operations

- Figuring out distance is an extended neighbourhood operation.

- For straight line distances, cells are merely counted. For diagonal distances, the Pythagorean theorem is used.

True
Euclidian
Distance

Y

X

= No Data

Demers, M. N., 2002. *GIS Modelling In Raster*. New York: John Wiley & Sons.
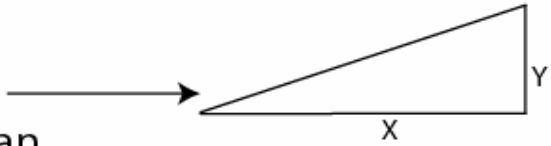
Weighted Distance Functions

- Based on idea that there is an accumulated "travel cost" between cells.

- Concerned with functional rather than Euclidean distance.

- Travel costs can be thought of as *frictions*.

- To estimate a travel cost, require a source grid and corresponding friction grid.

- Raster GIS calculates cost distance by adding Euclidean distance to friction values in the friction grid.

Calculating Weighted Distance

$a1 = (cost1 + cost2)/2$

- That is, if you are travelling between two cells in a horizontal or vertical direction, you add each grid cell (distance and friction) and divide by 2.

- Accumulated cost is found by adding interval costs $(a1 + a2 + a3)$.

Diagonal Weighted Distance

- The Pythagorean theorem is used to calculate the length of the hypotenuse. This creates a multiplier of approx. 1.414216 (or the $\sqrt{2}$).

- In this case, the diagonal cost between the first two grids cells is
  $$a1 = 1.414216\,(cost1 + cost2)/2$$

- This is the same formula as for simple weighted distance, but is multiplied by $\sqrt{2}$).

Iterative nature of cost distance analysis

- Remember that cost distance analysis always looks for the lowest cost between two points.

- It is an iterative process that begins with a source cell, selects the lowest cost cell in the grid and then moves to that cell, and begins the process again.
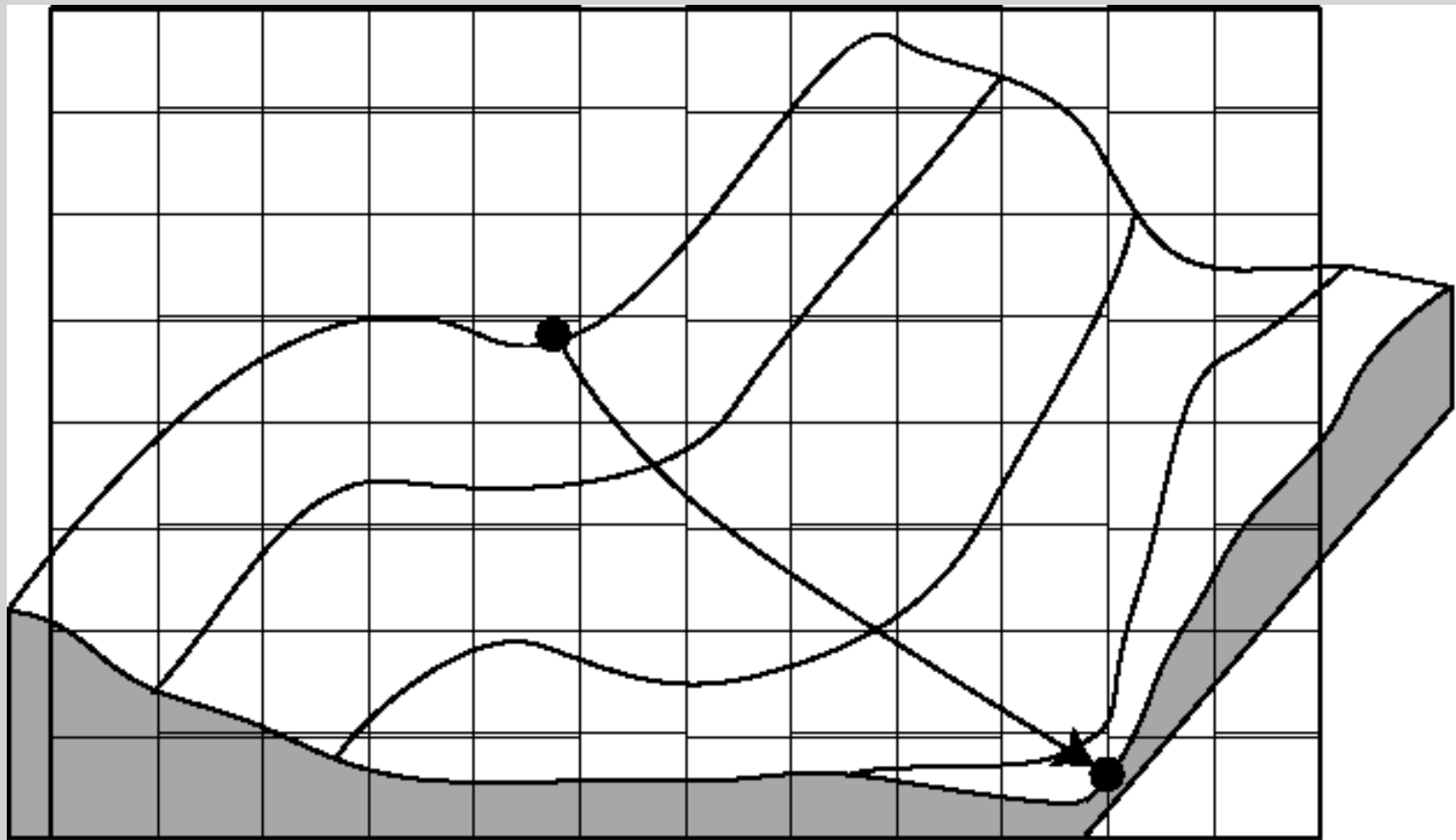
Least cost path over a raster surface. The first step is to compare elevation values around the starting point. The lowest value is chosen and the process is iteratively repeated until the end point is reached.

Example

- A car is travelling between points A and B.
- We want to determine the effect (cost) of distance on the amount of fuel consumed.
- A simplified solution is to take elevation into consideration.
- So, assuming that the distance *AB* is 25 km, and that it cost 1 litre of fuel, the Pythagorean theorem can be used to estimate the increased travel distance when there is an elevation effect.
- The added fuel cost can then be calculated proportionately.

Elevation Matrices

200	200	150	150
150	150	100	100
100	100	50	50
50			50

Direction of Travel

Horiozntal Distance
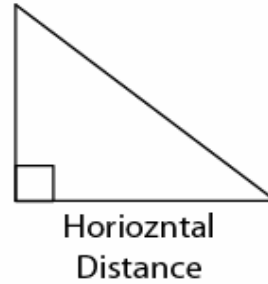
Total Surface Distance =
Horizontal Distance
(No Vertical Distance)

200	200	150	150
150	100	100
100	100		50
50	50	50	50

Direction of Travel
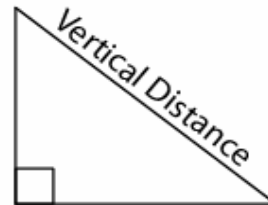
Vertical Distance

Total Surface Distance =
Horizontal Distance +
Vertical Distance

Demers, M. N., 2002. *GIS Modelling In Raster*. New York: John Wiley & Sons.

Problems with simplification

- This example is overly simplified. Real cost distance must account for surface roughness, wind, etc.
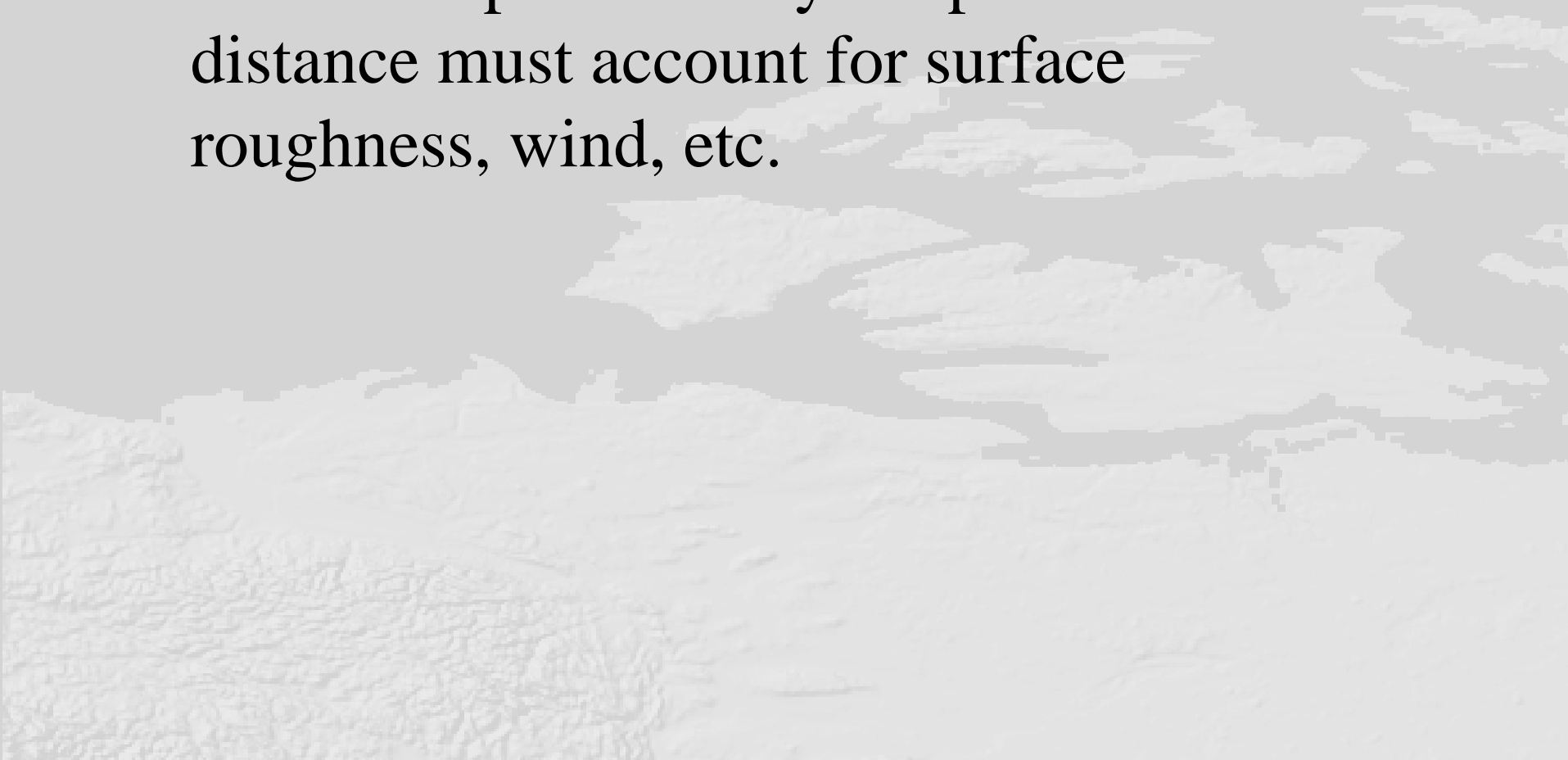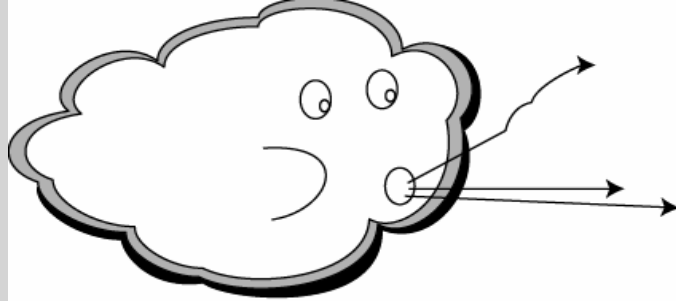
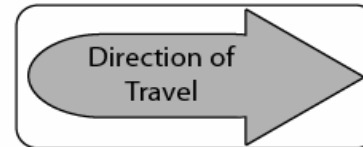Tailwinds are a POSITIVE Force → → REDUCED Fuel Consumption

Direction of Travel
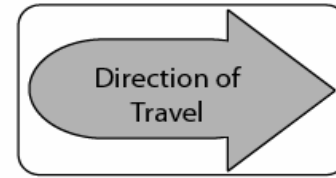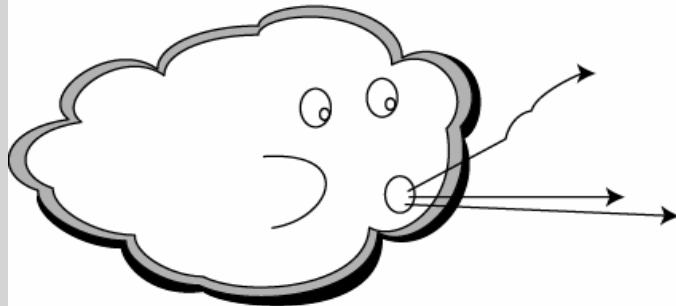
Headwinds are a NEGATIVE Force → ← INCREASED Fuel Consumption

Direction of Travel
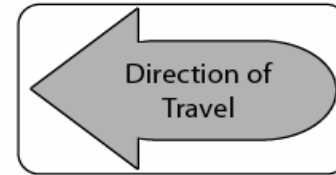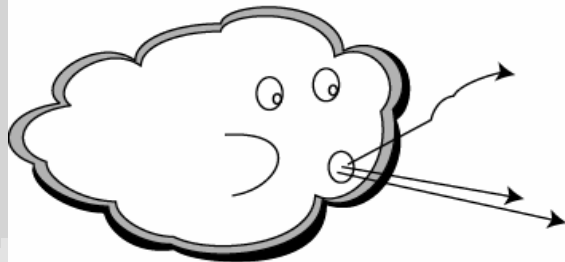
Crosswind with a NEGATIVE Force → ← INCREASED Fuel Consumption

Direction of Travel
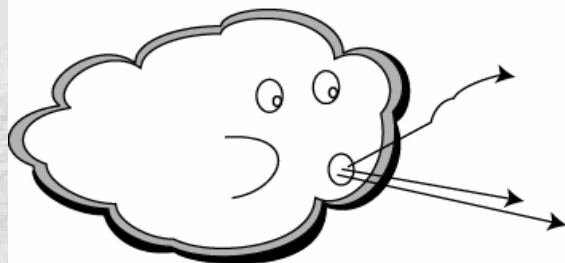
Crosswind with a POSITIVE Force → → REDUCED Fuel Consumption

Direction of Travel
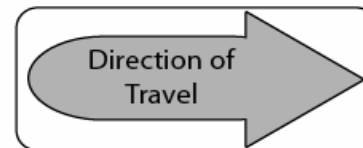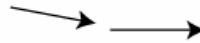
Remember

- All friction distances are subjective, and simplified versions of impedance.