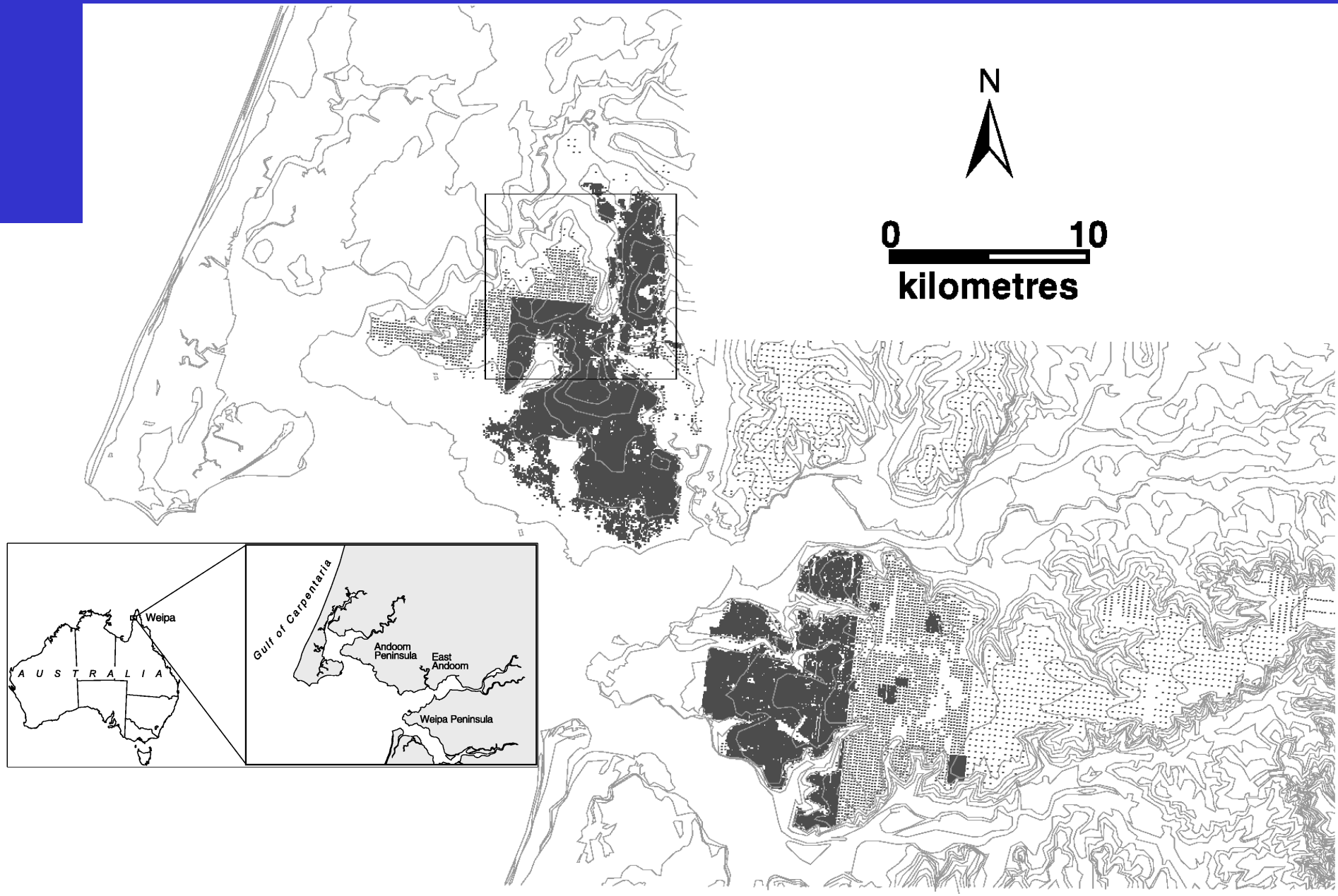
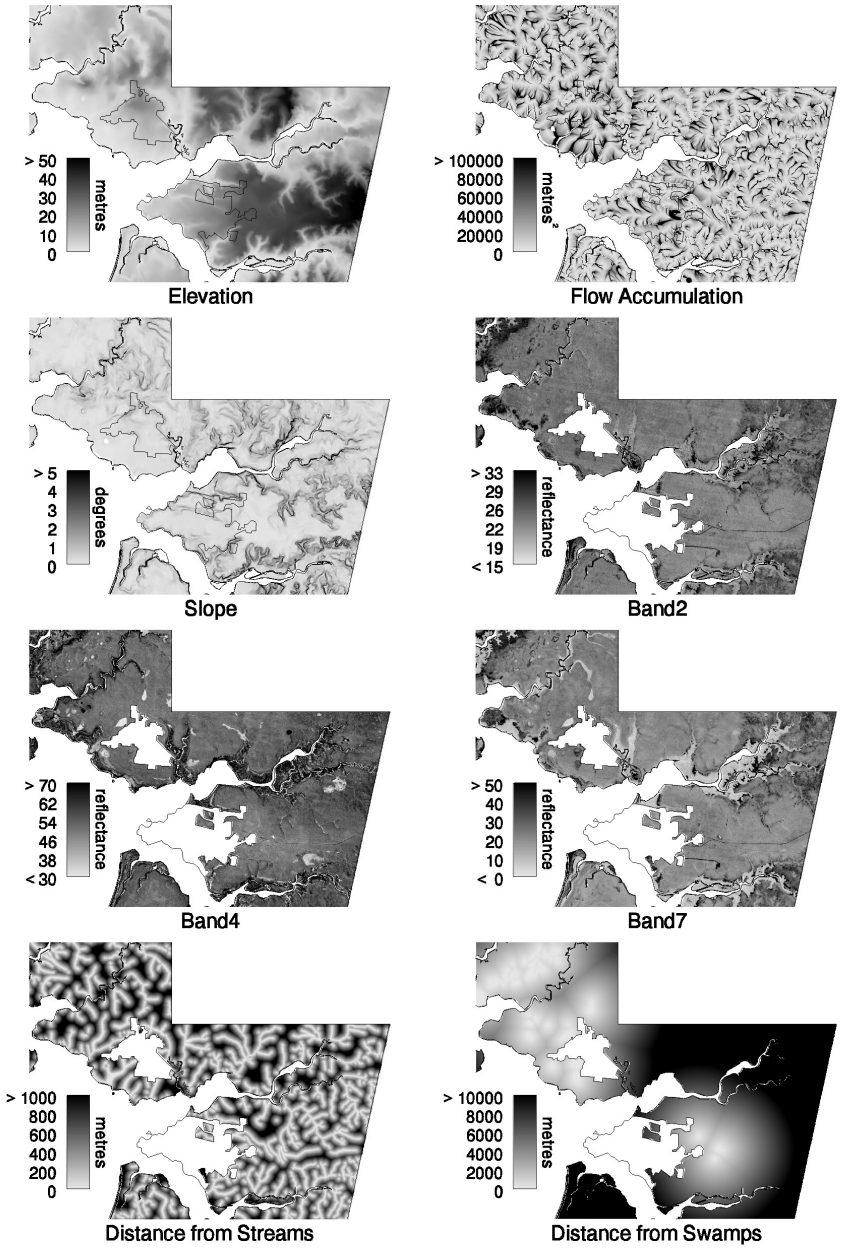
Exploiting the data explosion using geographically local analyses

Shawn Laffan

THE UNIVERSITY OF NEW SOUTH WALES
SYDNEY · AUSTRALIA

ANU
THE AUSTRALIAN NATIONAL UNIVERSITY


N
0
10
kilometres
Weipa
A U S T R A L I A
Gulf of Carpentaria
Andoom Peninsula
East Andoom
Weipa Peninsula


> 50
40
30
20
10
0
metres
Elevation
> 100000
80000
60000
40000
20000
0
metres²
Flow Accumulation
> 5
4
3
2
1
0
degrees
Slope
> 33
29
26
22
19
< 15
reflectance
Band2
> 70
62
54
46
38
< 30
reflectance
Band4
> 50
40
30
20
10
< 0
reflectance
Band7
> 1000
800
600
400
200
0
metres
Distance from Streams
> 10000
8000
6000
4000
2000
0
metres
Distance from Swamps
N
0
10
kilometres
Coastline,
Mining Areas
and Study Boundary

Local v global

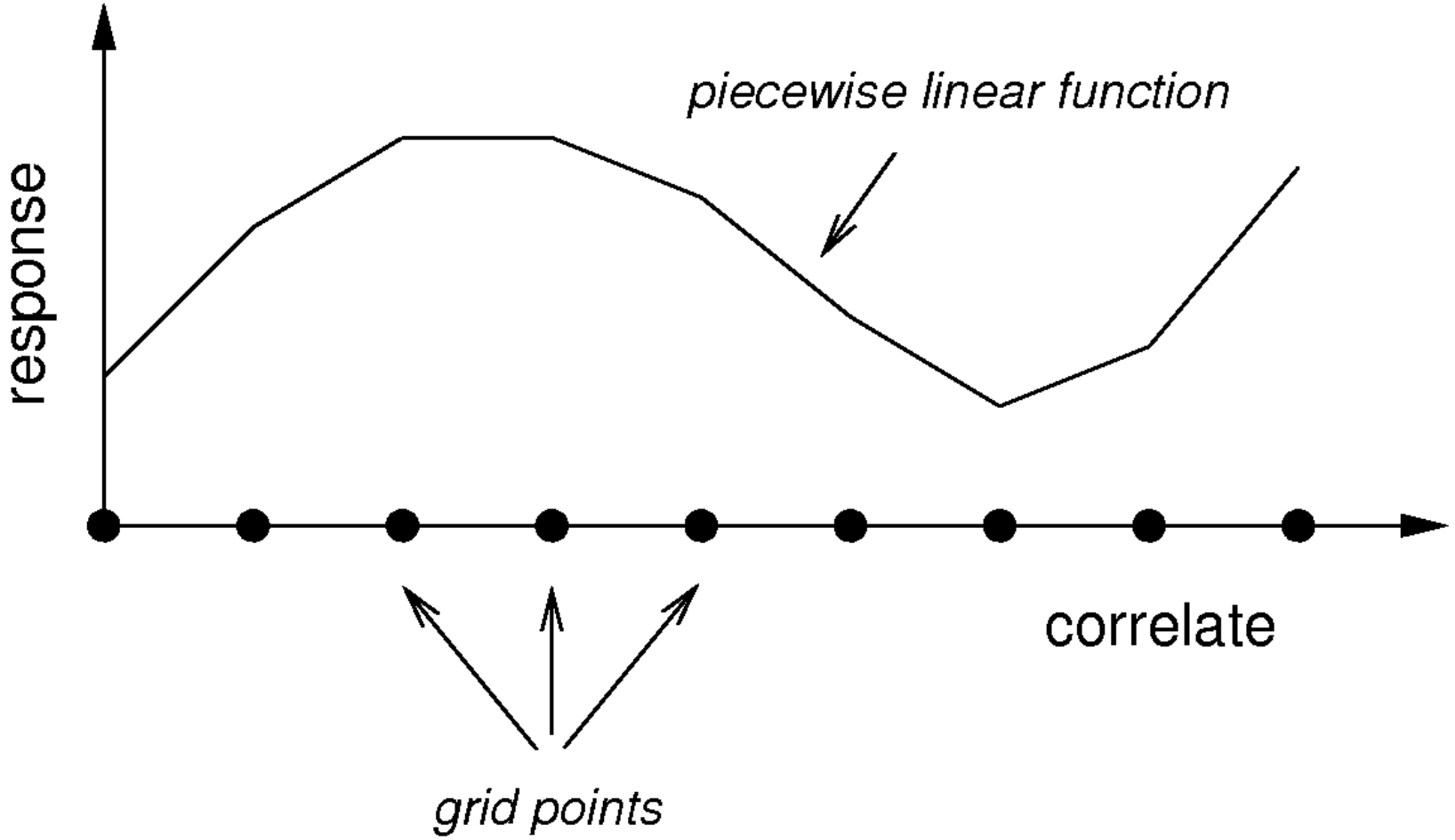
- Moving window vs lumped
  - spatial non-stationarity
- Surface of models vs single model

- Need more data
- Still subject to autocorrelation effects

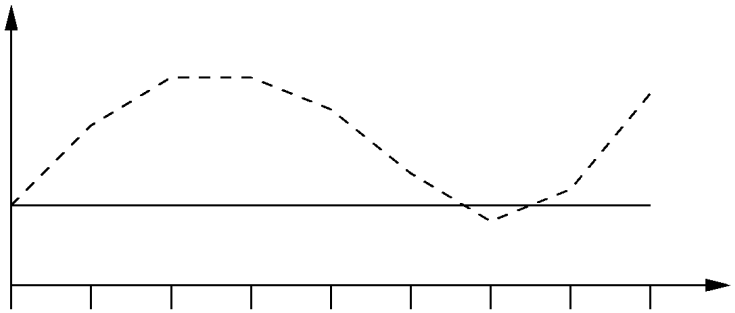
Sparse Grids

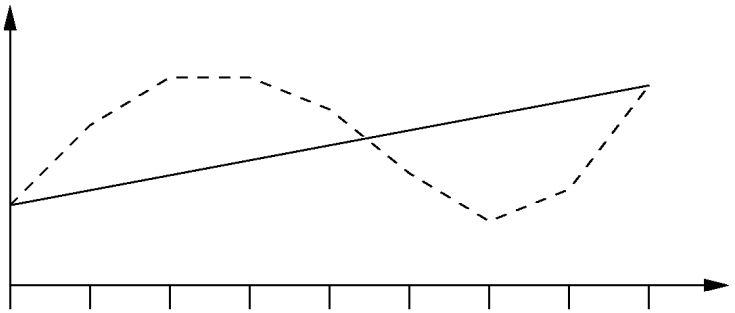
- Machine learning
  - do not assume uncorrelated errors
- Approximate high dimensional relationships using functions on grids
  - in attribute space
- Additive
  - similar to many other models
- Use fewer parameters than regular grid functions
  - are collections of regular grid functions

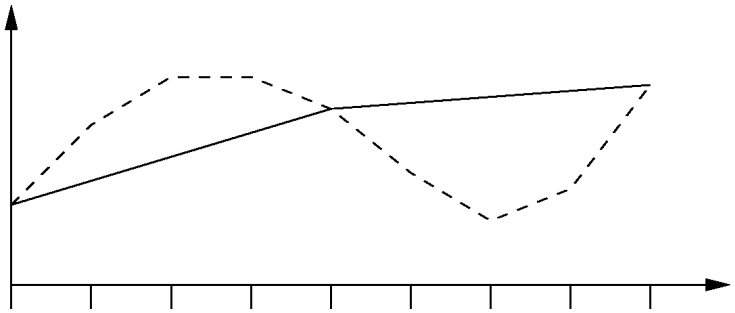
Piecewise linear functions

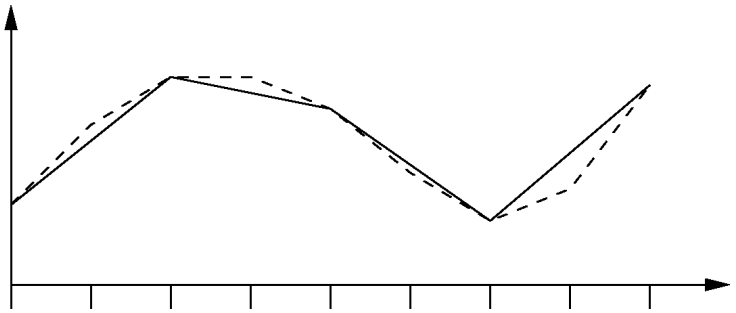


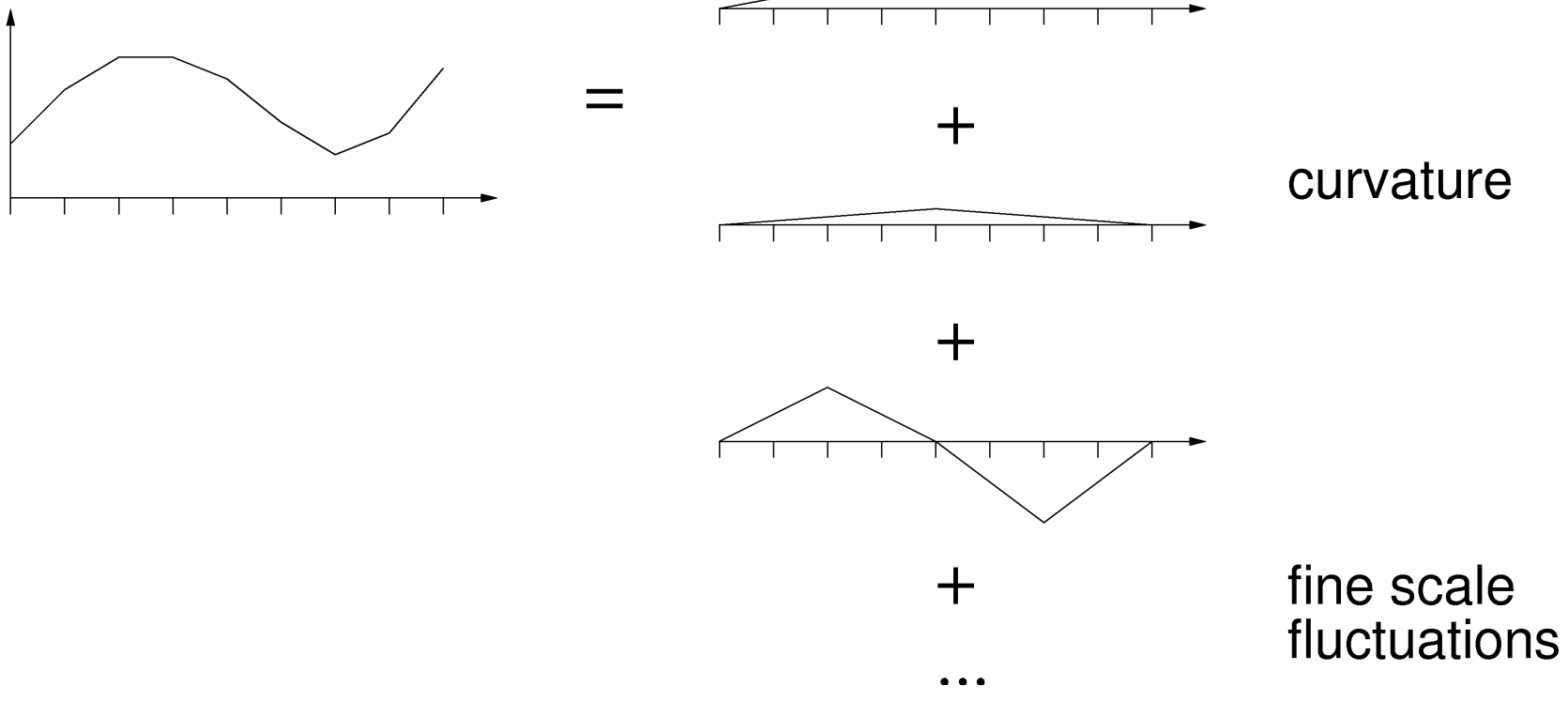
Decomposition of piecewise linear functions





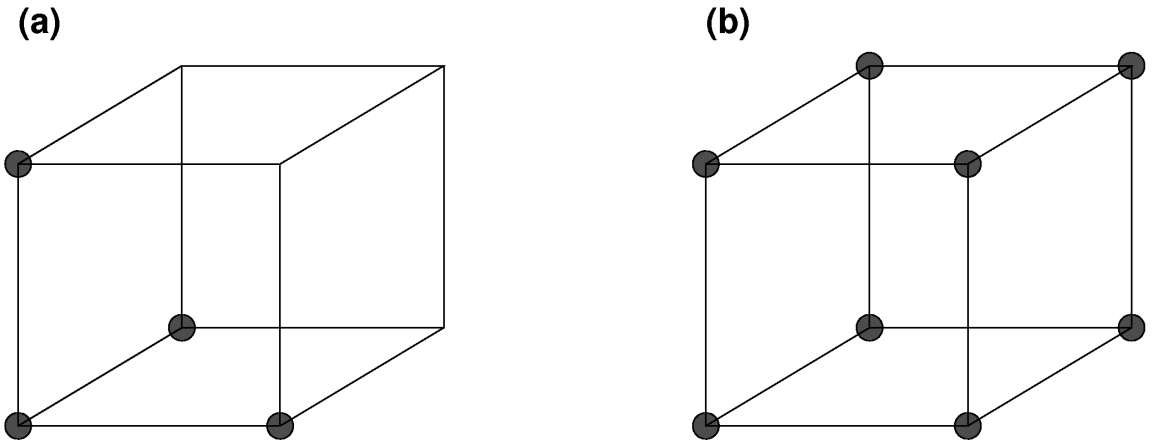


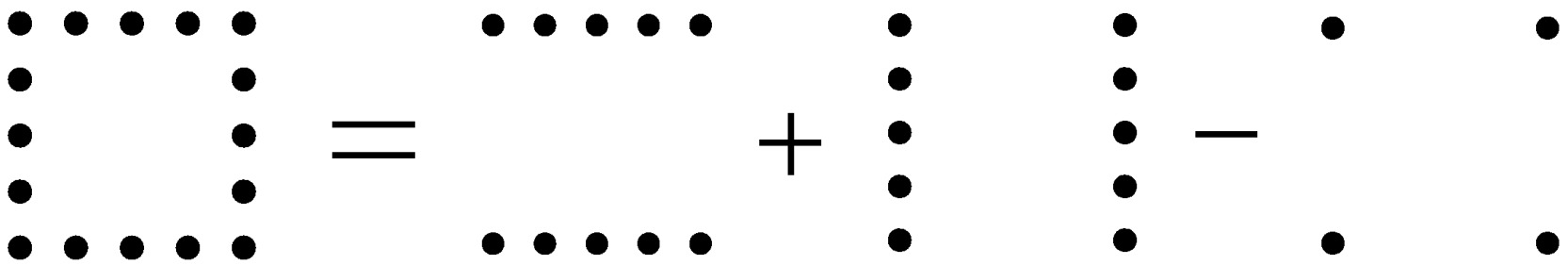


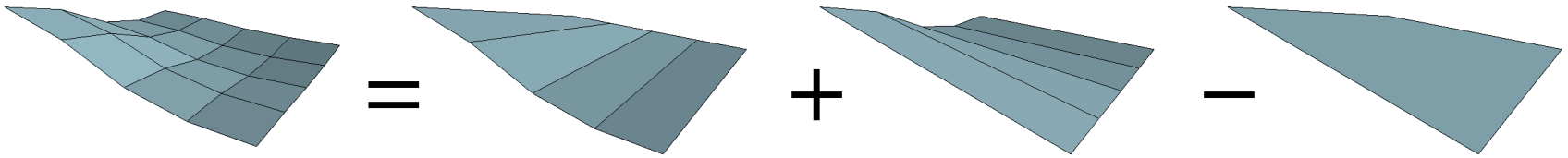

=
+
+
+
+
...
height
slope
curvature
fine scale
fluctuations

Sparse vs regular grids

- To define a function in three dimensions
  - Sparse grid uses 4 parameters
  - Regular grid uses 8 parameters

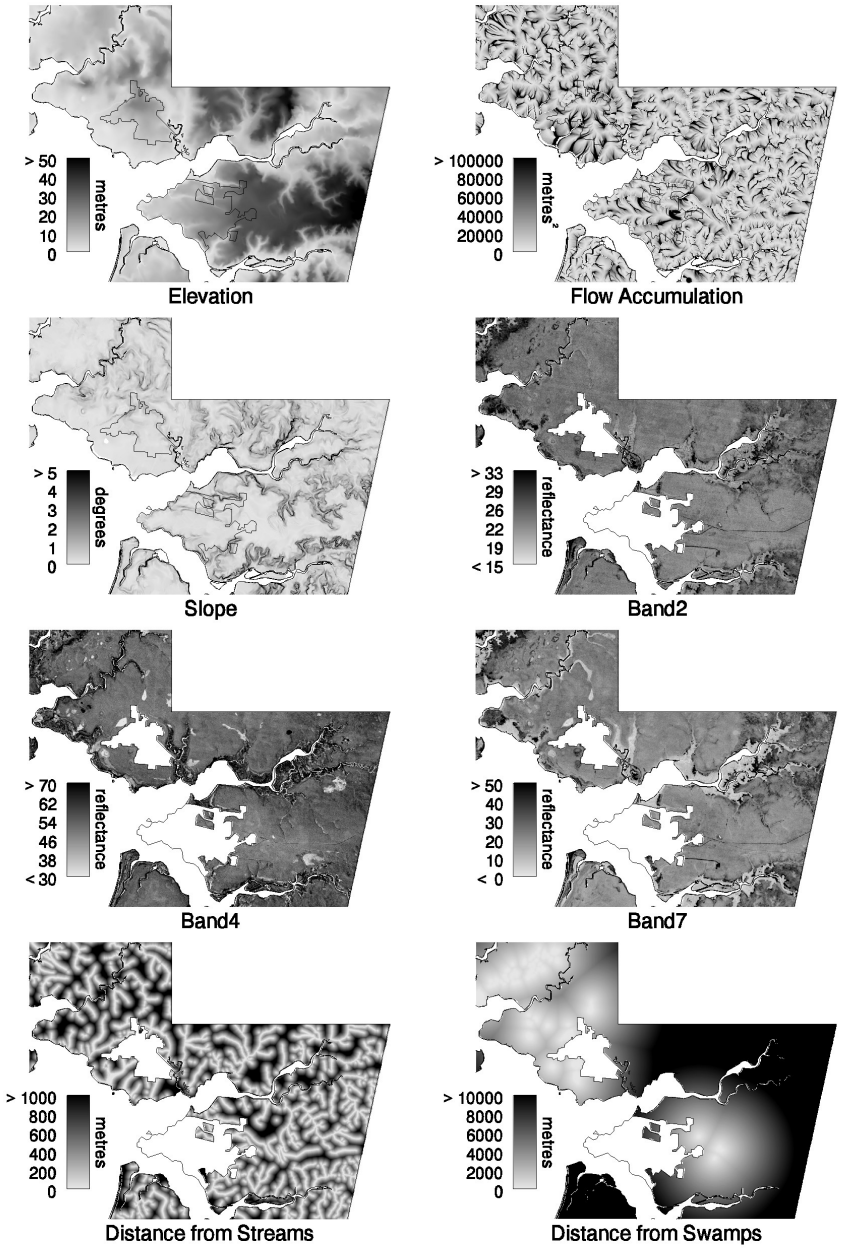






Sparse grid functions

- Can have an arbitrary number of grid points in any dimension
- So, we can have functions that are
  - order 0 (constant) in 6 dimensions
    - 1 grid point
  - order 3 and 5 in other dimensions
    - 5 & 17 grid points
  - V(0,0,0,3,0,0,5,0)

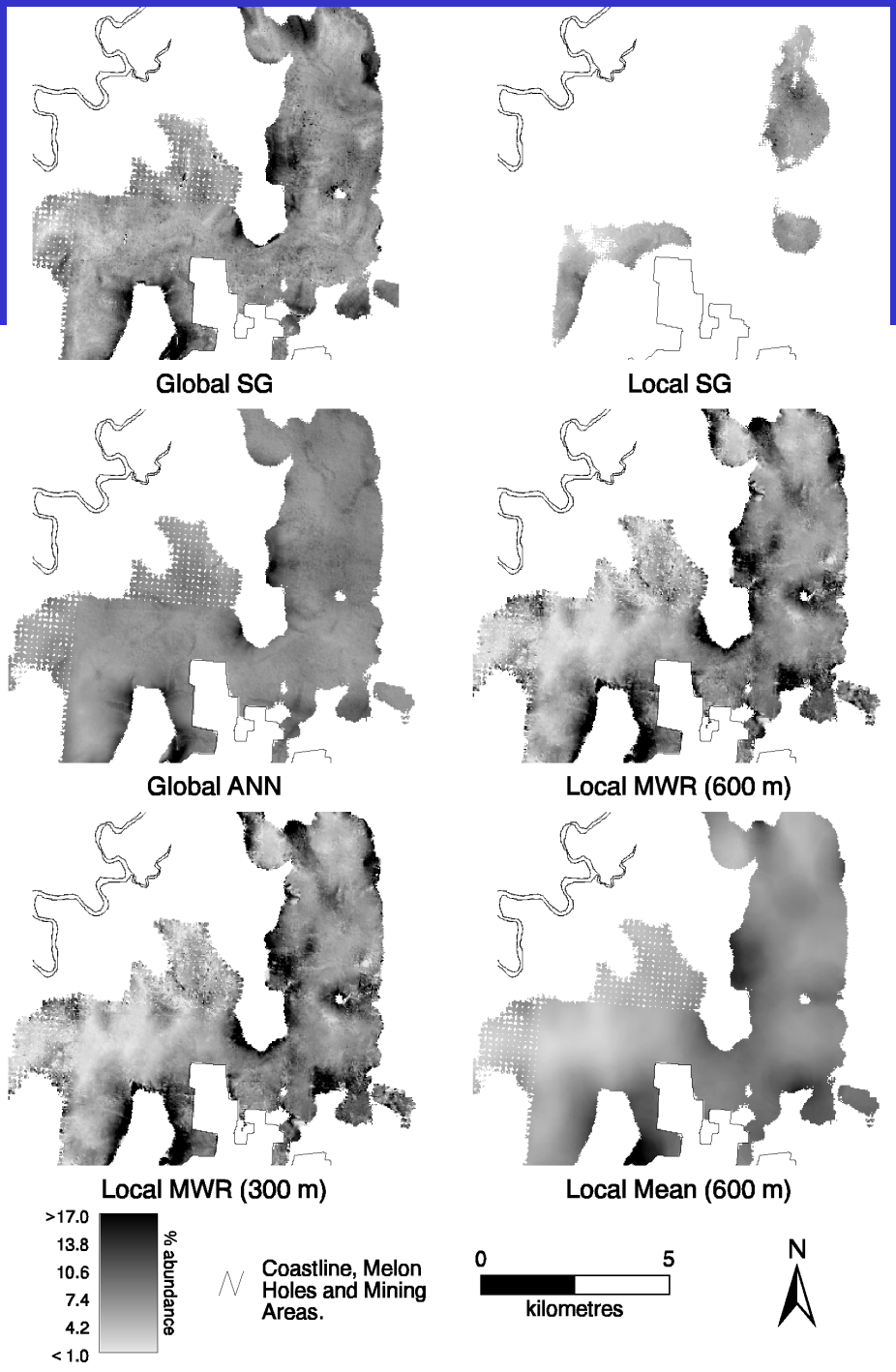

> 50
40
30
20
10
0
metres
Elevation
> 100000
80000
60000
40000
20000
0
metres²
Flow Accumulation
> 5
4
3
2
1
0
degrees
Slope
> 33
29
26
22
19
< 15
reflectance
Band2
> 70
62
54
46
38
< 30
reflectance
Band4
> 50
40
30
20
10
< 0
reflectance
Band7
> 1000
800
600
400
200
0
metres
Distance from Streams
> 10000
8000
6000
4000
2000
0
metres
Distance from Swamps
N
0
10
kilometres
Coastline,
Mining Areas
and Study Boundary

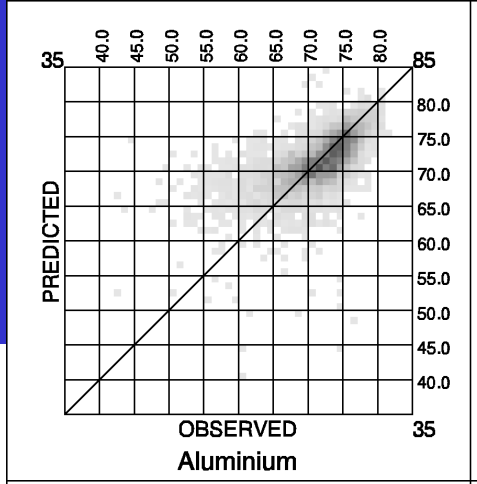
Application - Global

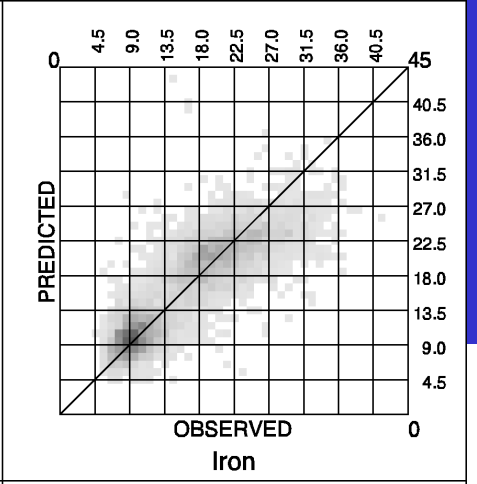
- Used 37 sparse grids
- All single variable grids, eg. V(5,0,0,0,0,0,0,0)
  - Order V(5) = 17 grid points
- All two variable interaction grids,
  - eg. V(5,5,0,0,0,0,0,0) , V(5,0,0,0,0,5,0,0)
- One constant grid
- Trained with 9,889 points
- Tested with 4,944 points
- 7297 parameters

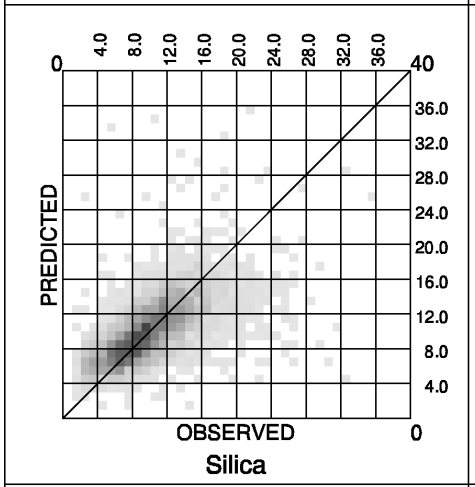
Application - Local

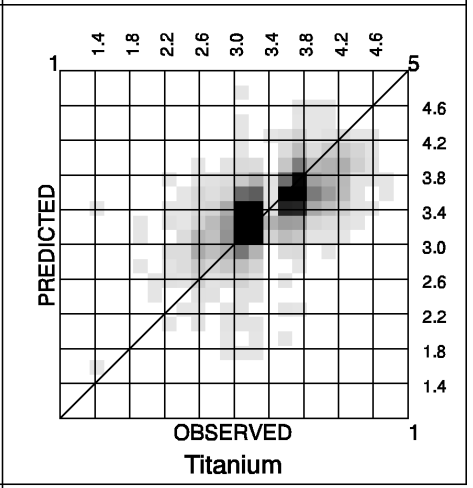
- All single variable grids, eg. V(2,0,0,0,0,0,0,0)
  - Order V(2) = 3 grid points
- All two variable interaction grids,
  - eg. V(2,2,0,0,0,0,0,0) , V(2,0,0,0,0,2,0,0)
- One constant grid
- Sample window of 600m radius
- Trained with 1232 points
- 20,994 models
- 129 parameters per model

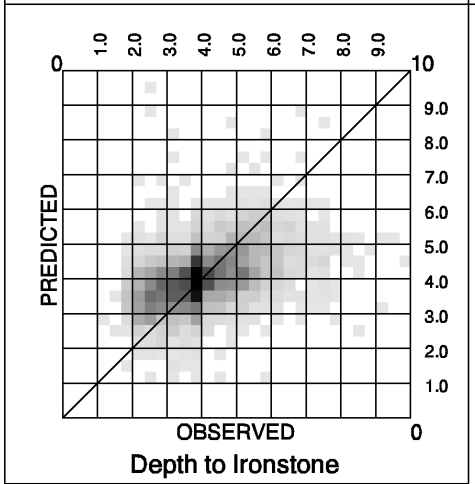

Global SG
Local SG
Global ANN
Local MWR (600 m)
Local MWR (300 m)
Local Mean (600 m)
>17.0
13.8
10.6
7.4
4.2
< 1.0
% abundance
Coastline, Melon Holes and Mining Areas.
0
5
kilometres
N

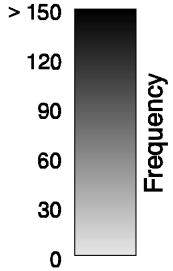




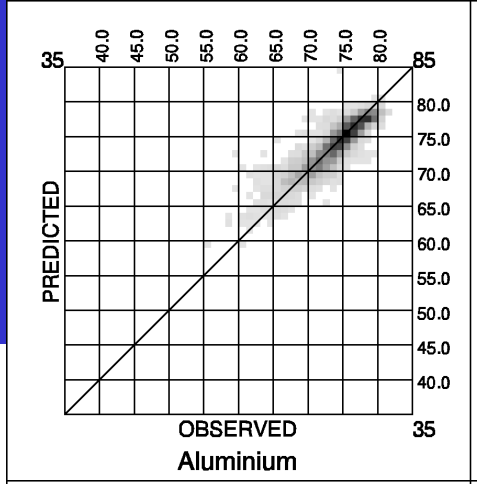


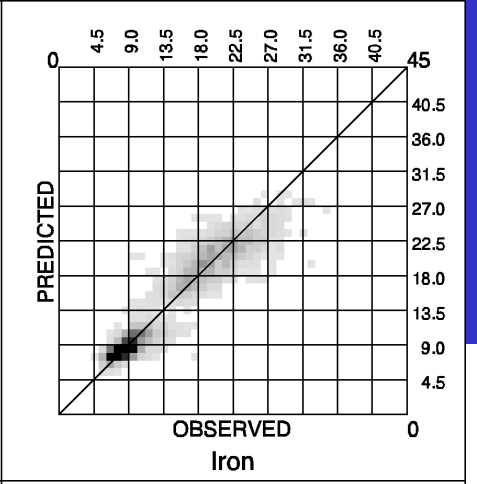


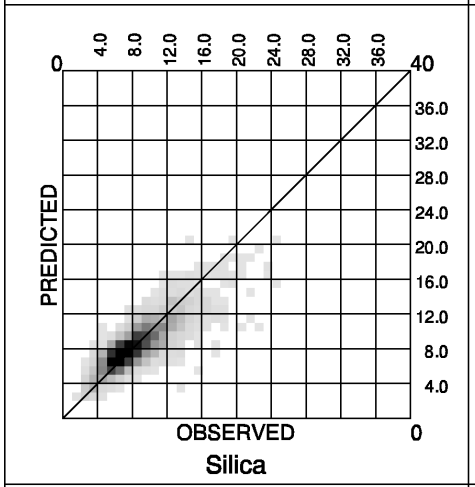


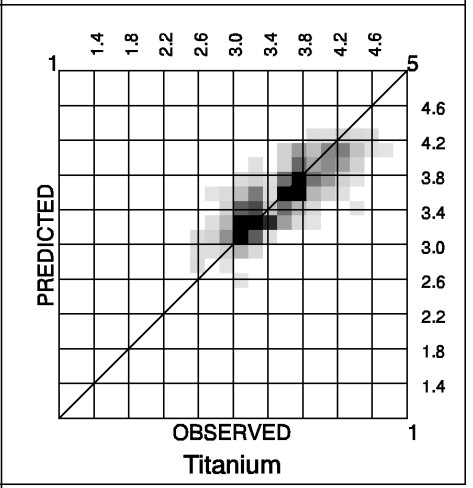


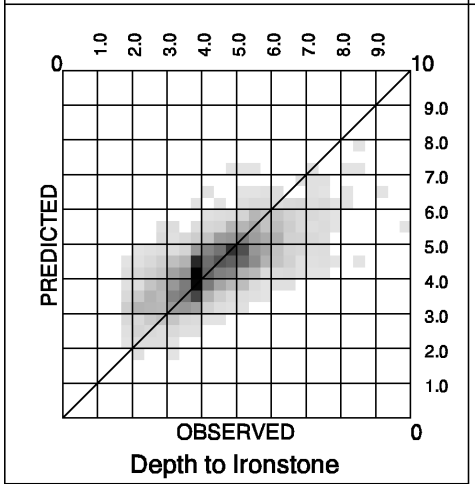
Global

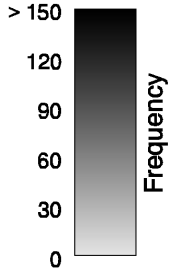




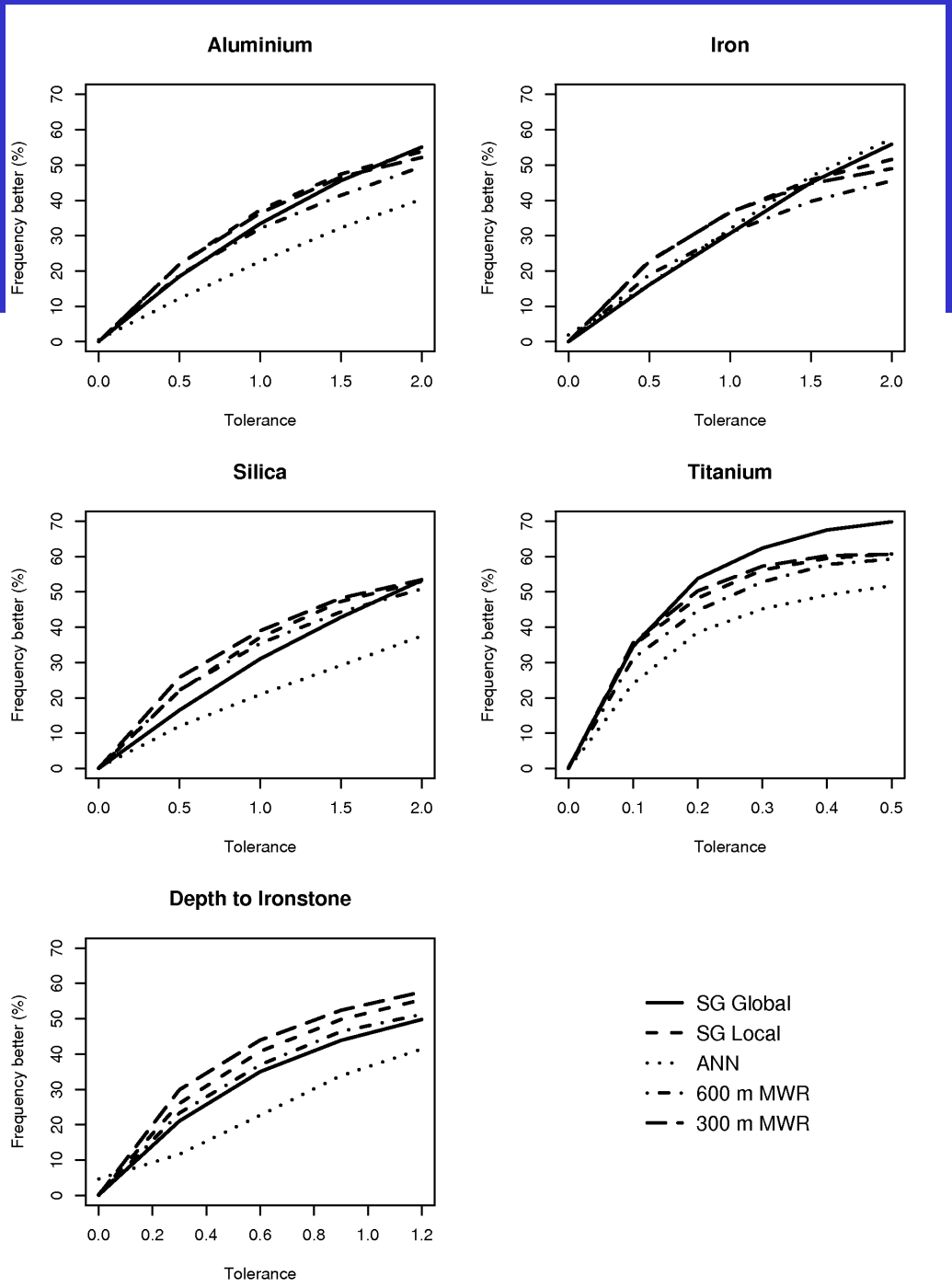


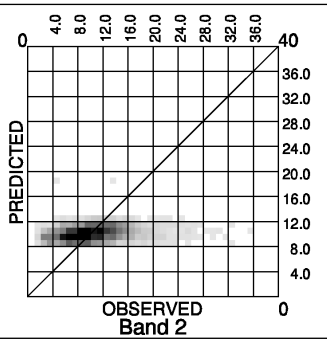


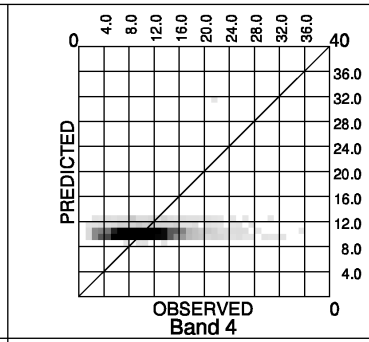


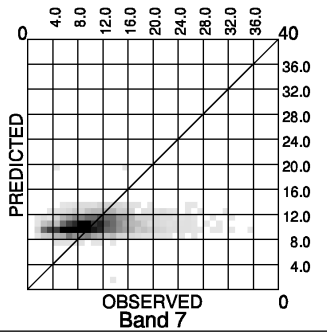


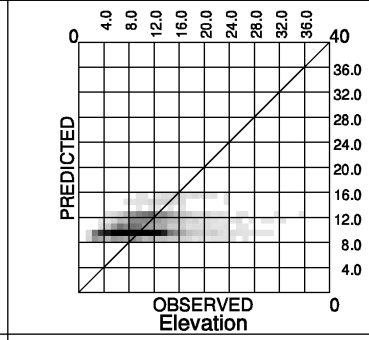
Local

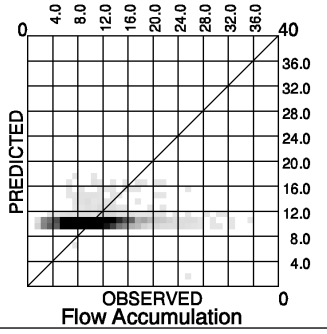

Aluminium
Iron
Silica
Titanium
Depth to Ironstone
Frequency better (%)
Tolerance
SG Global
SG Local
ANN
600 m MWR
300 m MWR

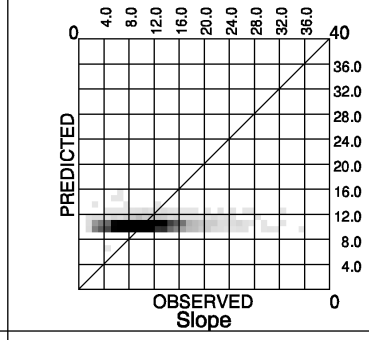

PREDICTED
OBSERVED
Band 2

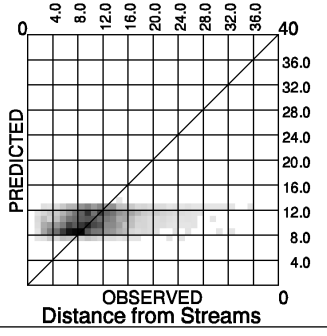

PREDICTED
OBSERVED
Band 4

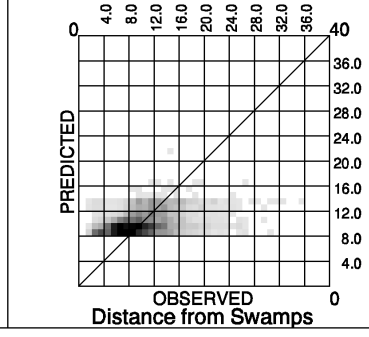

PREDICTED
OBSERVED
Band 7

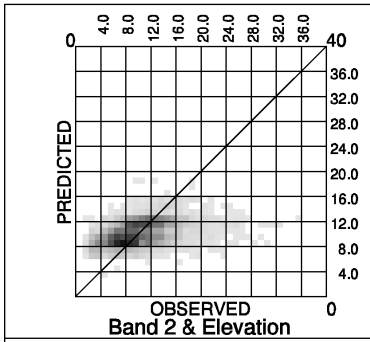

PREDICTED
OBSERVED
Elevation

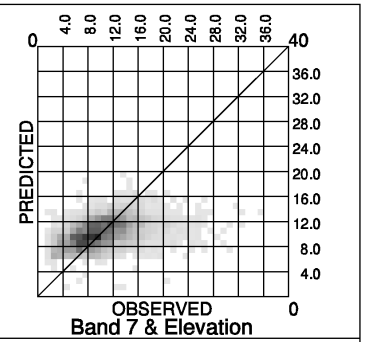

PREDICTED
OBSERVED
Flow Accumulation

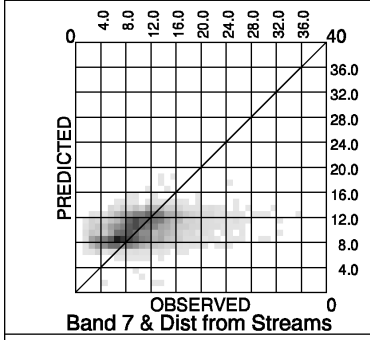

PREDICTED
OBSERVED
Slope

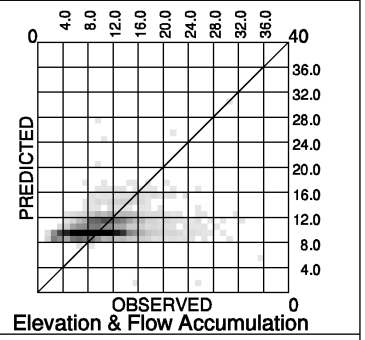

PREDICTED
OBSERVED
Distance from Streams

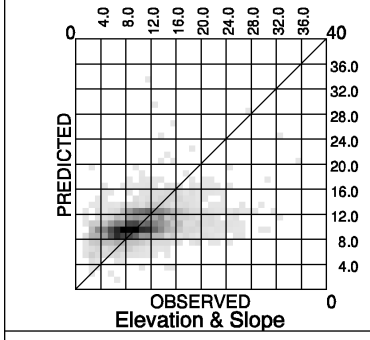

PREDICTED
OBSERVED
Distance from Swamps

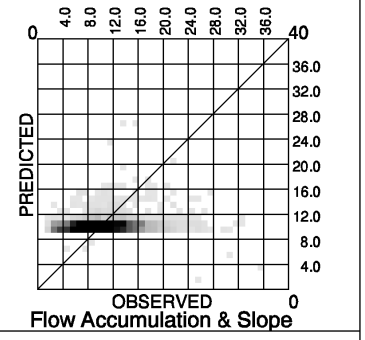

PREDICTED
OBSERVED
Band 2 & Elevation

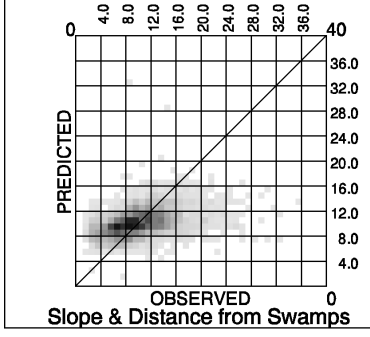

PREDICTED
OBSERVED
Band 7 & Elevation

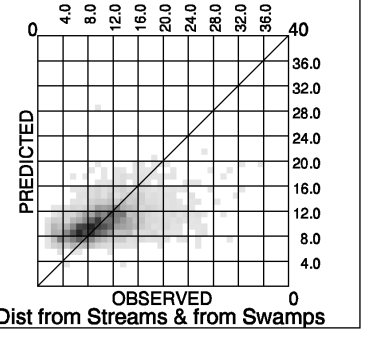

PREDICTED
OBSERVED
Band 7 & Dist from Streams

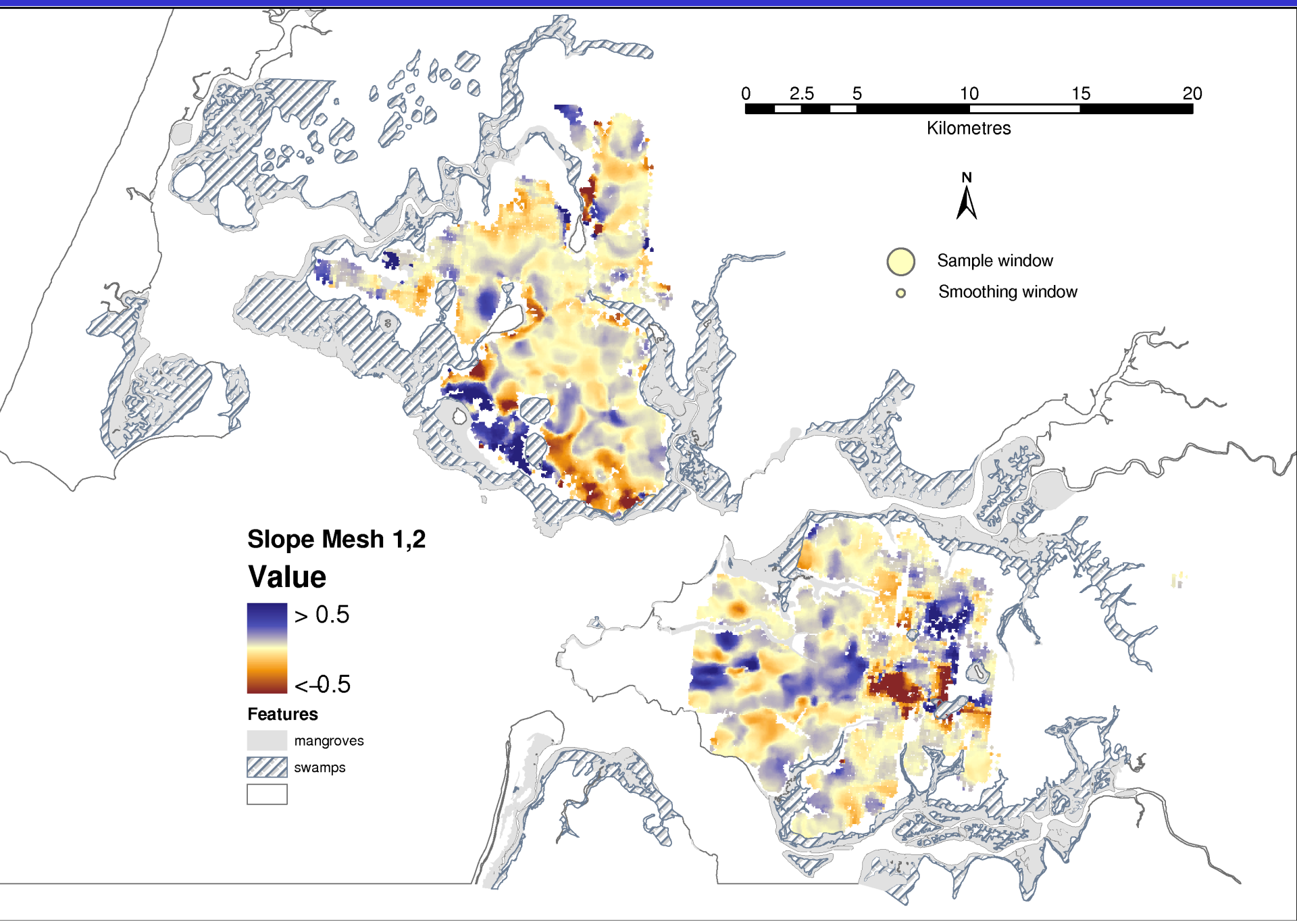

PREDICTED
OBSERVED
Elevation & Flow Accumulation

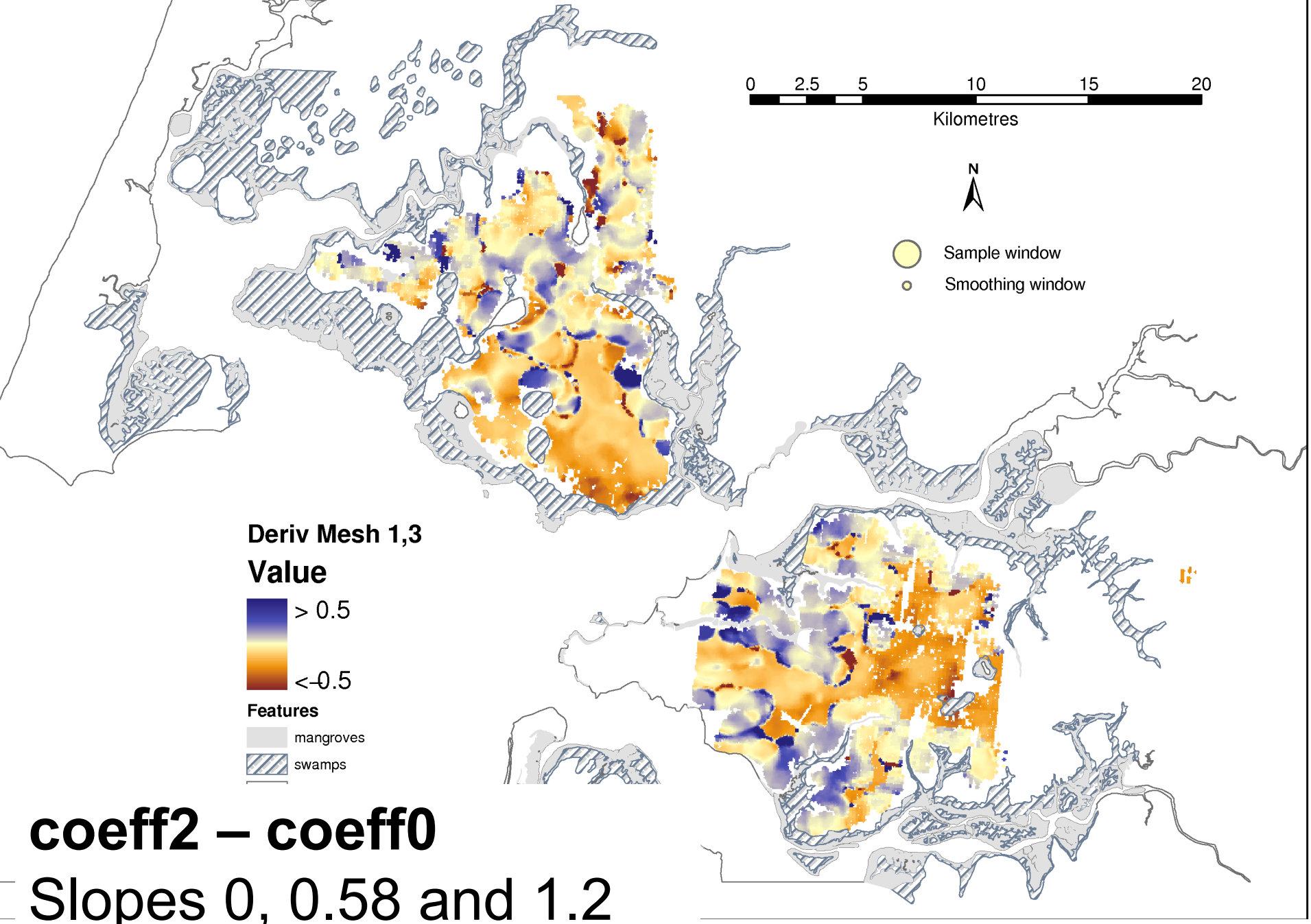

PREDICTED
OBSERVED
Elevation & Slope


PREDICTED
OBSERVED
Flow Accumulation & Slope


PREDICTED
OBSERVED
Slope & Distance from Swamps

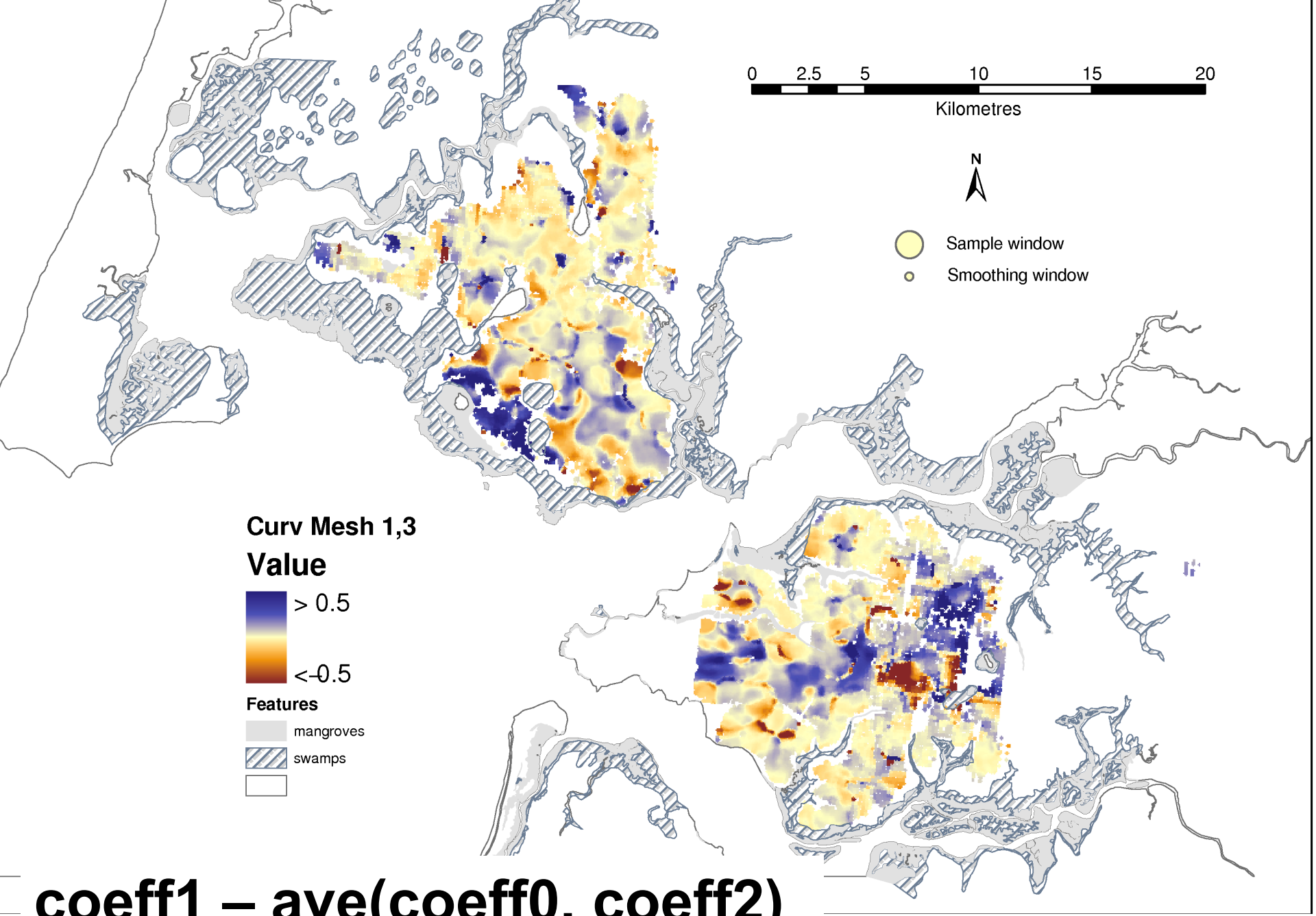

PREDICTED
OBSERVED
Dist from Streams & from Swamps


0 2.5 5 10 15 20
Kilometres
N
Sample window
Smoothing window
Slope Mesh 1,2
Value
> 0.5
<-0.5
Features
mangroves
swamps

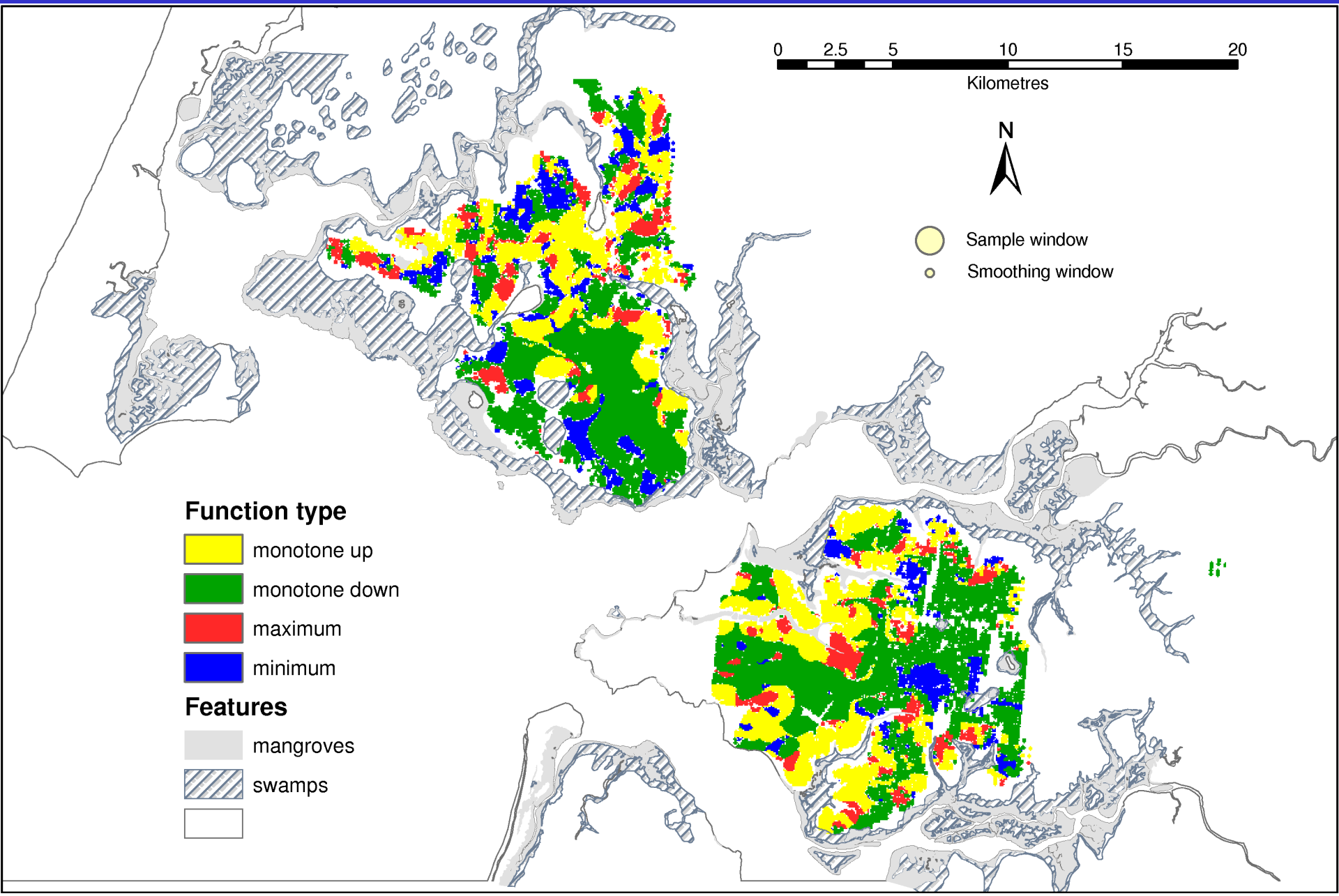


coeff2 – coeff0

Slopes 0, 0.58 and 1.2



coeff1 – ave(coeff0, coeff2)


0 2.5 5 10 15 20
Kilometres
N
Sample window
Smoothing window
Function type
monotone up
monotone down
maximum
minimum
Features
mangroves
swamps

Summary

- Sparse grids are a promising tool for the analysis of geographic data
- Potential to understand “scale” of relationships in attribute space
  - use number of grid points required in each dimension
- Parallel implementations possible
  - for the impatient