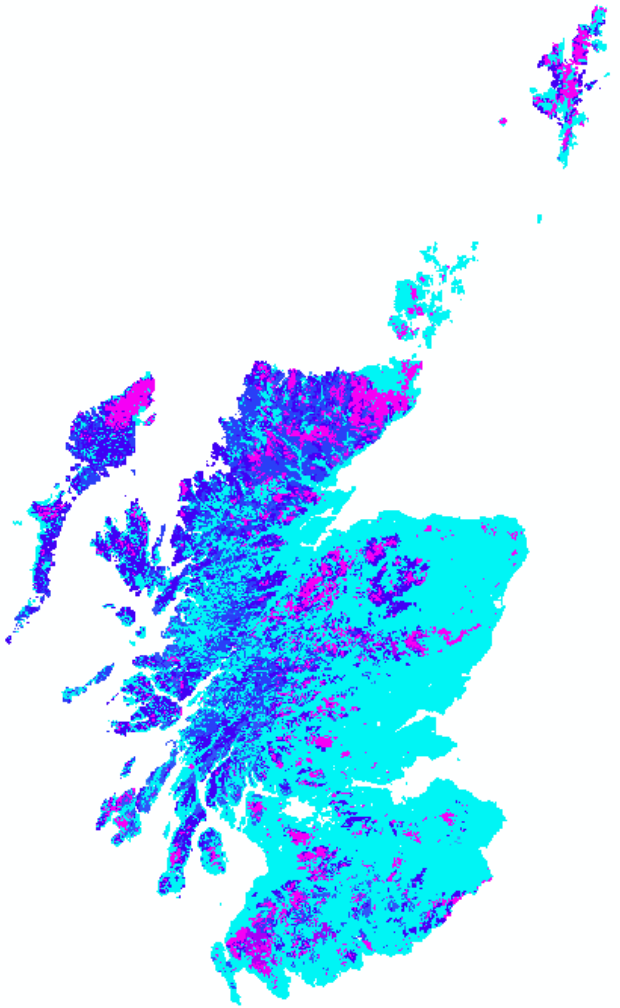
A Digital Soil Mapping neural network approach for peat depth mapping in Scotland

Matt Aitkenhead
Matt Saunders
Jagadeesh Yeluripati
John Bell

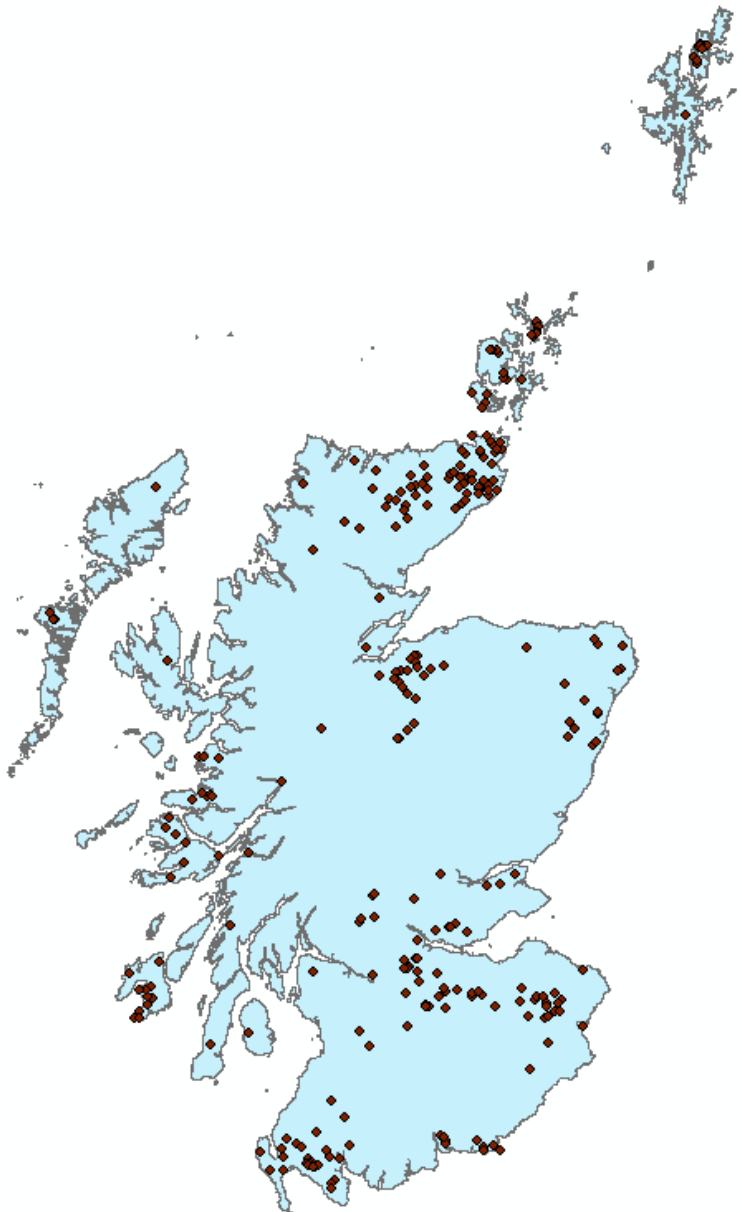
The James Hutton Institute

Scotland's peat

- Peat bogs cover ~22% of Scotland, contain 60% of UK peat and ~3000 million tons of C
- Different characteristics (pH, water content, bulk density)
- Different vegetation/land use
- Different depths
- Condition varies greatly
- Accurate depth mapping not achieved
- Difficult to assess overall C sequestration/storage potential



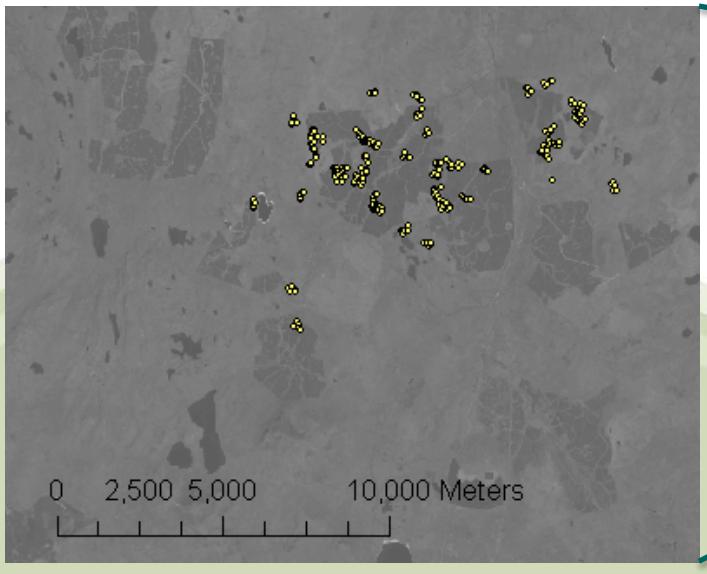
Peat survey data



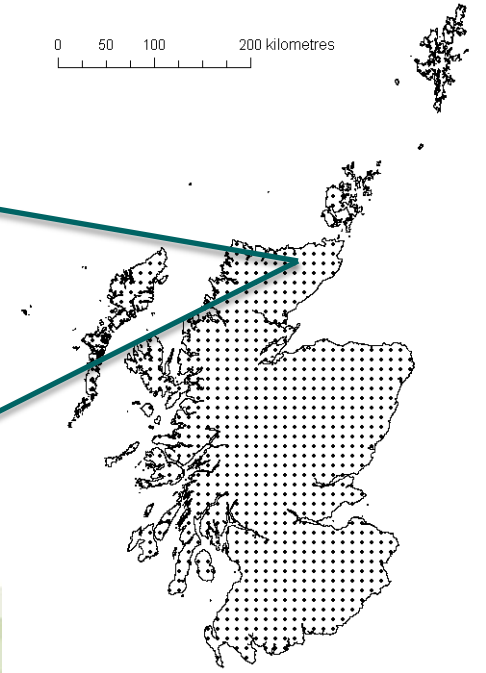
- Multiple peat surveys (1964-1968)
- More recent work focussed on specific bogs or types of bog
- Data from 277 peat bogs used
- Mean depths along transects
- Number of sample points varied with size of bog
- Good distribution between lowland, raised and upland peat bogs (blanket & basin peats)

The big assumption

- That using sufficient mean depth values from peat bogs across Scotland is:
  - Going to provide sufficient depth variation for modelling
  - Qualitatively equivalent to using individual depth values
- The assumption is tested against peat depth sampling data from two surveys (Forsinard and NSIS)



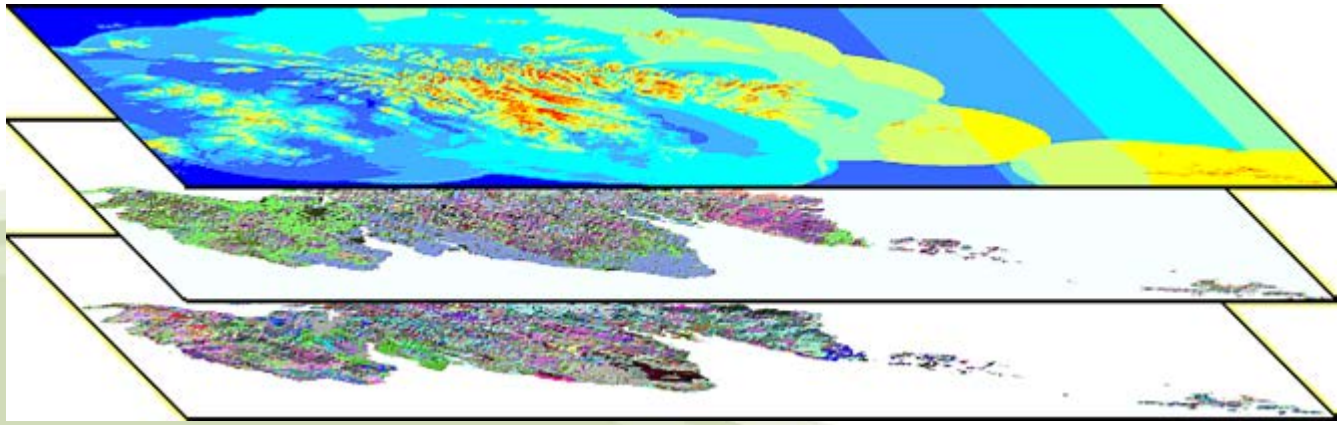
Forsinard sample data



National Soil Inventory of Scotland (NSIS 1)

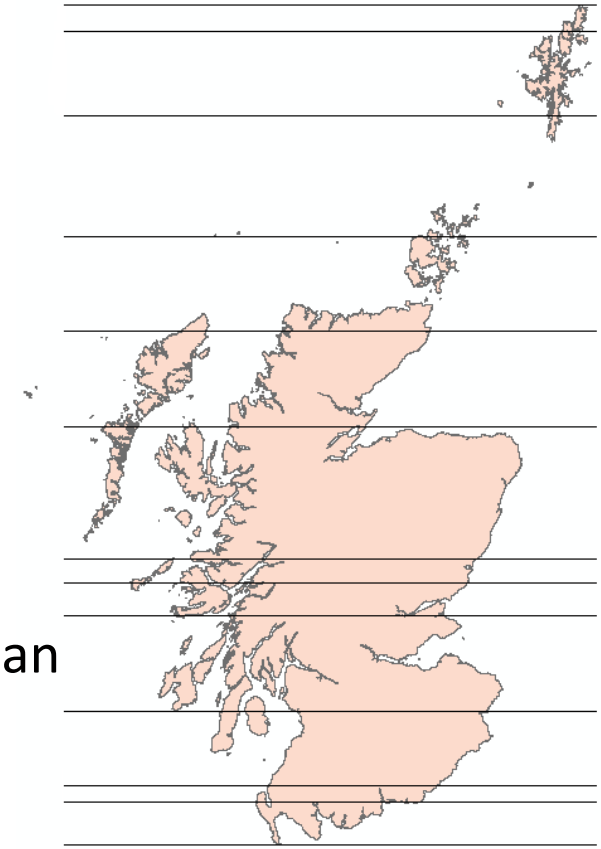
Ancillary datasets

- Topography (elevation, slope, aspect, curvature, flow integration)
- Soil map data (percentage of nine broad classes within map units)
- Vegetation (CORINE, 10 broad classes)
- Climate (monthly mean temperature & rainfall)
- Geology (19 broad classes derived from soil map of Scotland)



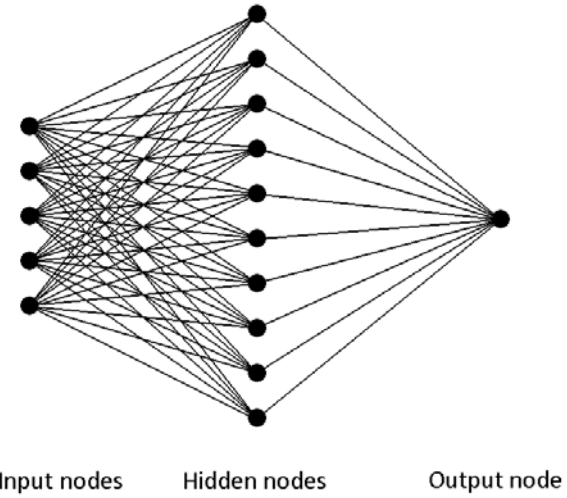
Data ‘strip’ generation

- Ancillary datasets used to produce strips of data
- Each strip is 425km x 100m
- Contains 4250 lines with 80 parameters
- 6927 strips produced
- Speeds up data access for model training/testing
- Many small (1.5 MB) files easier to handle than one big (10.4 GB) file
- Can be used to rapidly generate maps using trained models



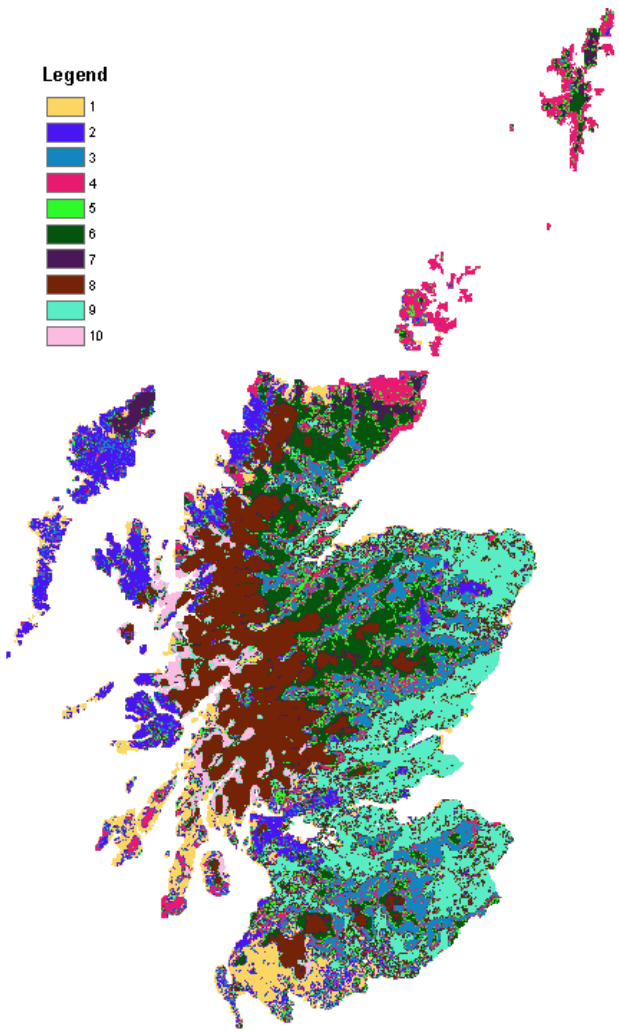
Neural network training

- Backpropagation NN used
- 10-fold cross validation
- Each ‘fold’ has 10 networks for consensus
- NN architecture 80:10:10:1
- Trained for 100k steps, tested every 1k and best taken
- Data normalisation within range [0.1, 0.9]
- Depth values distribution-normalised by using square root



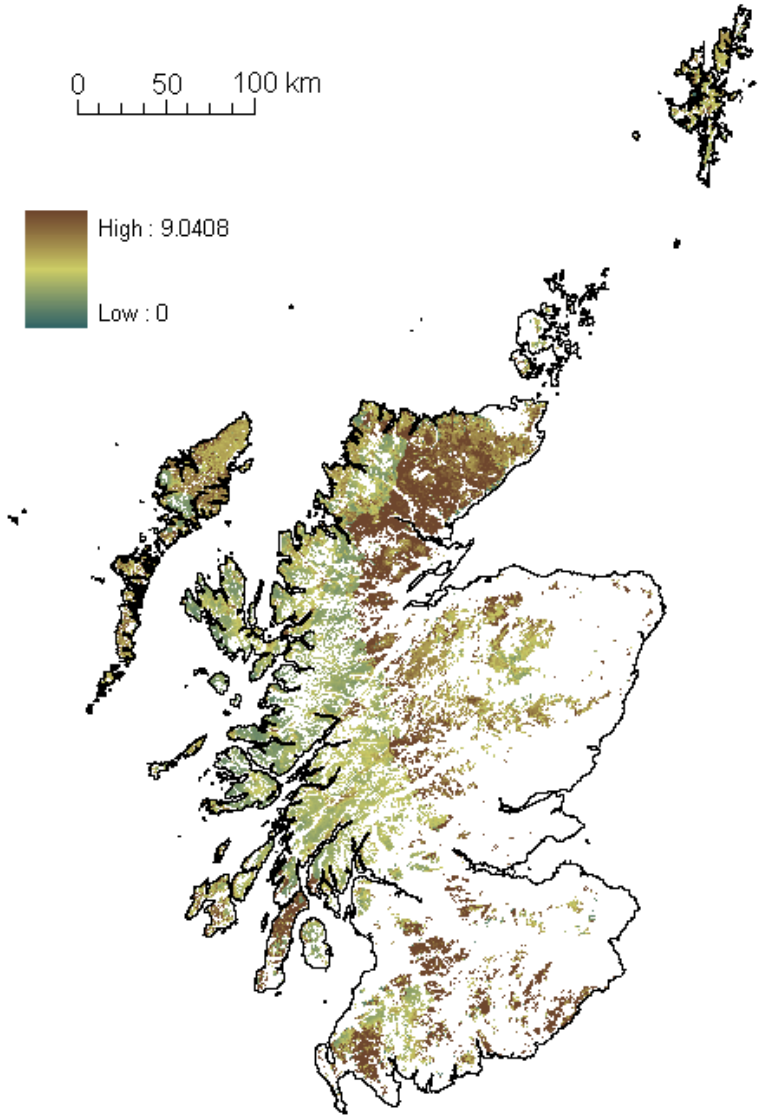
Stratification of data

- Fuzzy k-mean clustering of training data
- 10 clusters, based on input parameters
- Several factors influence stratification
  - Vegetation
  - Temperature
  - Rainfall
  - Slope
- Assessed with & without stratification



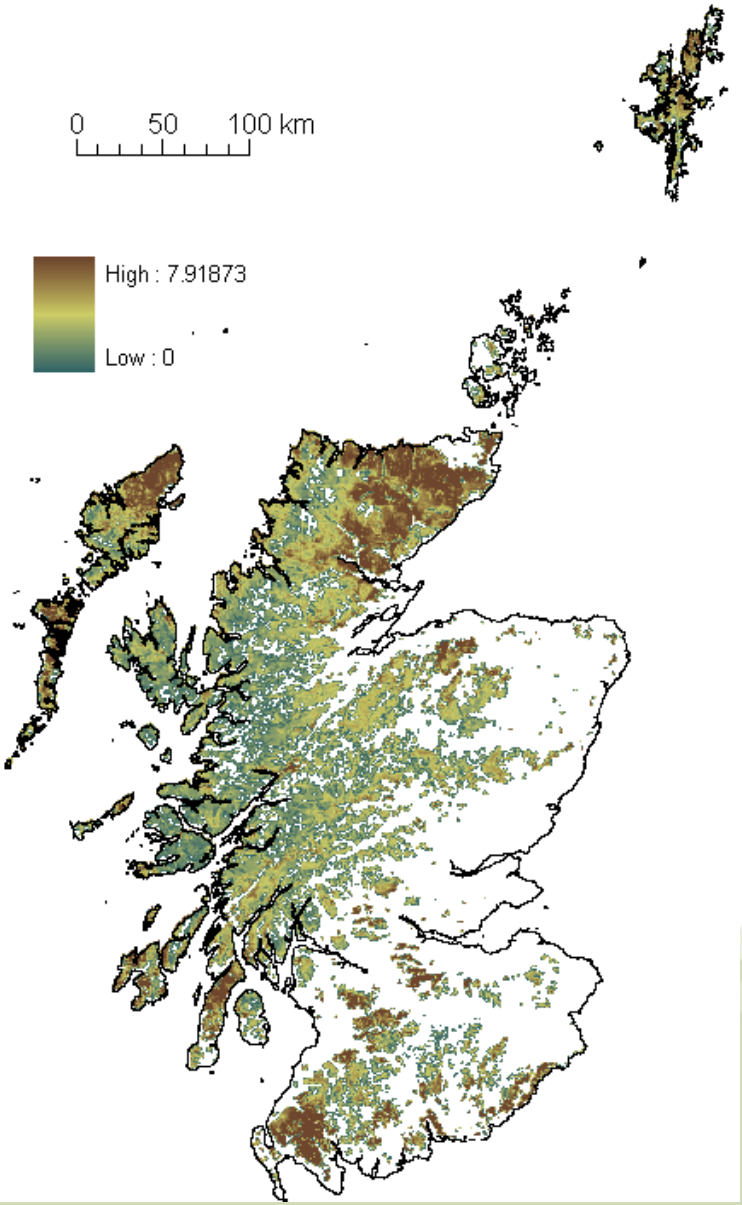
The James Hutton Institute

Mapping – with stratification



- Predicted depth for all map units where peat > 0%
- Range 0 - 9 metres predicted
- Issues with stratification boundaries

Mapping – without stratification



- Predicted depth for all map units where peat > 0%
- Range 0 - 8 metres predicted
- No issues with stratification boundaries
- Overall agreement with stratification map

Results (I)

Normalised (SQRT)	Stratification	No stratification
RSQ	0.754	0.720
RMSE	1.37	1.46 m
MAE	0.98	1.07 m
RPD*	2.01	1.93
RPIQ**	2.76	2.57

Linear scale	Stratification	No stratification
RSQ	0.730	0.685
RMSE	1.39	1.52 m
MAE	1.04	1.18 m
RPD*	1.90	1.78
RPIQ**	2.53	2.31

*RPD = std. dev. of observations / std. err. in predictions (redundant with RSQ)
**RPIQ = (Q75% - Q25%) / std. err. in predictions (Bellon-Maurel et al. 2010, TrAC Trends in Analytical Chemistry)