Biomathematics and Statistics Scotland (BioSS)

D.A. Elston & J.W. McNicol

Biomathematics and Statistics Scotland (BioSS) contributes statistical and mathematical expertise to the SEERAD-sponsored research programme. BioSS staff are based in Edinburgh, Aberdeen, Ayr and Dundee to facilitate close collaboration with scientists at the five SABRIs and all three SAC sites. BioSS research is carried out under three themes, namely 'Systems Modelling and Risk', 'Spatial and Temporal Models' and 'Statistical Genomics and Bioinformatics'. The last of these is led from SCRI, and was the subject of a workshop with our French counterparts, INRA's Département de Biométrie et Intelligence Artificielle.

Detection of putatitive recombinant sequences in multiple alignment data has been the target of three statistical methods developed at SCRI during recent years. With the appointment of a Bioinformatics programmer, we started to develop a Java-based graphical interface, called TOPALi, to these methods that would run on all computer platforms. The initial version allows plotting of the graphs in real time, and the estimation and visualization of phylogenetic trees. Our future plans include the automation of many aspects of the analysis. BioSS also increased its efforts in Microbial Genomics via a joint BioSS/SCRI PhD studentship to analyse the newly sequenced Erwinia carotovora genome. Initial work has been on the detection of regulatory elements using existing statistical methods, including Weight Matrices and Hidden Markov Models.

In linkage disequilibrium mapping, a model has been developed to simulate the flow of genetic markers through a plant breeding population. Our aim is to create a simulated barley population to see if the reduced recombination rate in barley makes the method feasible for detecting traits of economic interest. The model estimates the proportion of heterozygotes, allele frequencies and Hardy-Weinberg equilibrium in each generation. The results of the model are consistent with genetic theory, with linkage disequilibrium decaying at a rate that is proportional to the selfing rate and homozygosity increasing under selfing.

The link between vegetation and microbial community structures in upland grassland received an increased BioSS involvement with the start of phase III of the SEERAD funded Micronet project. As well as designing sampling strategies, geostatistical and multivariate methods were used to identify the extent and scale of coupling between soil microbial communities and plant species in upland grasslands. This work has been extended to investigate the effect of arbuscular mycorrhizal fungi on rhizoplane bacterial community structure.

Image processing and analysis were used to identify void and solid space in soil samples and geostatistical tools were used to describe the spatial distribution of voids within highly heterogeneous images of soil aggregates obtained using X-ray microtomography at the Advanced Photon Source, Argonne National Laboratory, Chicago

Under the heading of ecosystem management, a model was constructed to investigate the processes

responsible for the persistence of genetically modified (GM) oilseed rape volunteers and to predict the impurity that would result in subsequent crops. A stagestructured, life-history approach was taken which also includes the effect of various management practices on persistence and contamination. The simulated oilseed rape populations are subject therefore to a set of life-history processes and events associated with the management of an arable rotation that in combination dictate the growth rate of the volunteer population. Results from the numerical analysis of the model indicate that stringent control measures would be necessary to suppress GM volunteer populations sufficiently to achieve the EU thresholds for the adventitious presence of GM material

Analysis of metabolic profiles has become an increasingly important aspect of the BioSS contribution to the Quality, Health and Nutrition programme at SCRI. We have adopted the approach of analysing each compound individually by analysis of variance and summarising the results by simple significance plots for main effects and interactions. We also use principal components to examine how the compounds react in combination with each other, and plot the component scores against the individual compound values to identify the relationships among compounds, components and experimental treatments.