



**ARGOS RESEARCH NOTE: NUMBER 20, NOVEMBER 2006**

## Management of Data in ARGOS

The data management in ARGOS is centered around a Microsoft Access database and an ArcGIS geodatabase. This research note focuses on the Access database, and briefly touches on how this database is linked with a Geographic Information System (GIS).

The initiative for an integrated approach to data management in the project were sparked by a need to safeguard data integrity/security, have a central storage for all data, and to facilitate cross linkages between datasets, within disciplines, as well as between the economic, ecological and social domain of the research.

### Database development

The development focus for the database has been on designing appropriate entity structures that safeguards data integrity and flexibility in relationship building between separate datasets, rather than a user-friendly front end. There are a couple of reasons for this focus;

1. The database is developed alongside the development of the research programme, i.e. the data and the structure of the data is not determined in advance and any elaborate front end would risk becoming obsolete very quickly.
2. The data will very seldom be entered directly in the database by

researchers or field personnel. In most cases the database administrator will import data from an Excel spreadsheet.

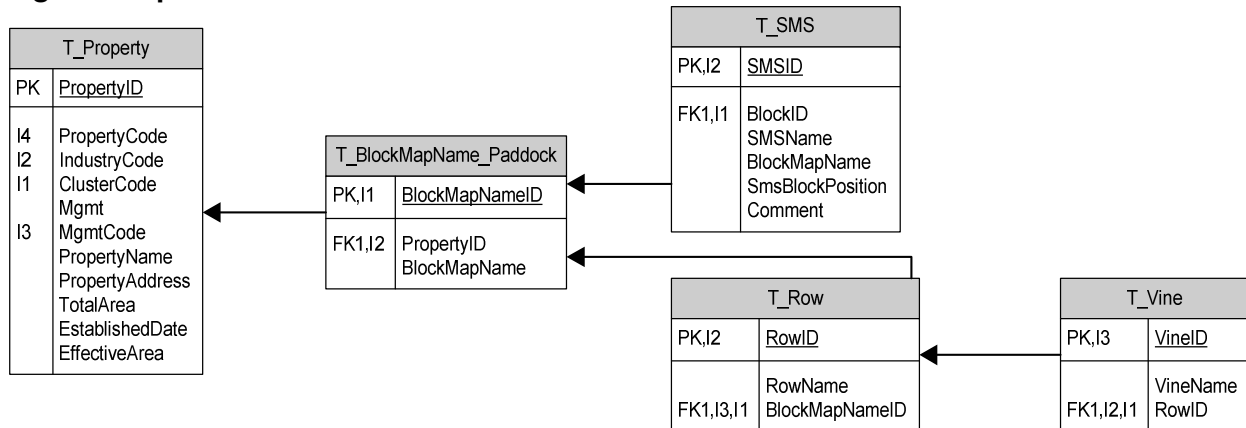
3. The users of the data, the researchers, will require customised datasets rather than repeated use of standard 'views'.
4. ArcGIS, the GIS system ARGOS is using, provides a spatial front end where most data in the Microsoft Access database can be visualised/displayed.

### Database structure

The central entity structure in the database relates to the spatial structure of the research design (see Figure 1 on next page). The main subject of interrogation is properties, may they be Sheep and Beef farms or Kiwifruit orchards. Each property is divided into management units (blocks or paddocks). Furthermore, a management unit may contain monitoring sites, such as Soil Monitoring Sites (SMS). For the Kiwifruit orchards there is additional spatial structure in the form of rows within blocks, and vines within a row.

The spatial structure is supplemented with a subject structure of farm owners and their families, as well as ARGOS researchers. Each dataset that ARGOS collects will be related to this basic structure of properties and people.

**Figure 1. Spatial structure in the database**



### GIS integration

The primary keys (unique identifiers) of properties, paddocks/blocks, soil monitoring sites and vegetation elements, that are created in the Microsoft Access database is used in the GIS system as unique identifiers that allow us to relate all ARGOS datasets, to all the elements that are stored in the geodatabase. On the other hand, the GIS can provide spatial information such as management units' areas or fence lengths that can be imported into the Access Database.

### Data management process

The database sets no restrictions on survey design. When a survey has been designed and data collection is about to begin, the database administrator is involved in the setup of the data templates. Templates are designed in either Microsoft Access or Microsoft Excel.

When a survey has been completed the dataset is sent to the database administrator in the agreed template, from which it is imported in the database.

The database administrator designs queries and exports data from the database for any analysis that involves one or more datasets. An example of such a query could be a dataset that links data on soil status, fertiliser management, canopy management, yield, and dry matter for the kiwifruit orchards in the project. This involves linking 5 different datasets across 3

different spatial units (block, soil monitoring sites, property). Most queries can be generated within a days work and can be saved for future use. This has the distinct advantage of providing the project with a 'push of the button' capability for generating cross linked datasets for when individual surveys are repeated and the longitudinal aspects of the ARGOS project becomes the focus.

### Summary

The key to management of ARGOS data has been to create a database that is flexible enough to handle the data storage needs of a research project that has an exploratory component to its surveys and potentially will go on for 20 years.

The use of Microsoft Access and a dedicated database administrator is a cost effective way of achieving professional data management. With linkages to a modern GIS system the power of the database setup is significantly increased.

This Research Note was prepared by **Martin Emanuelsson** and **Tania Maegli**

For further information, please contact;

Martin Emanuelsson, Programme Manager  
 The AgriBusiness Group, PO Box 4354,  
 Christchurch.  
 Phone: 03 365 6808  
 Email: martin@agribusinessgroup.com