

REVIEWS

Simple...but complex

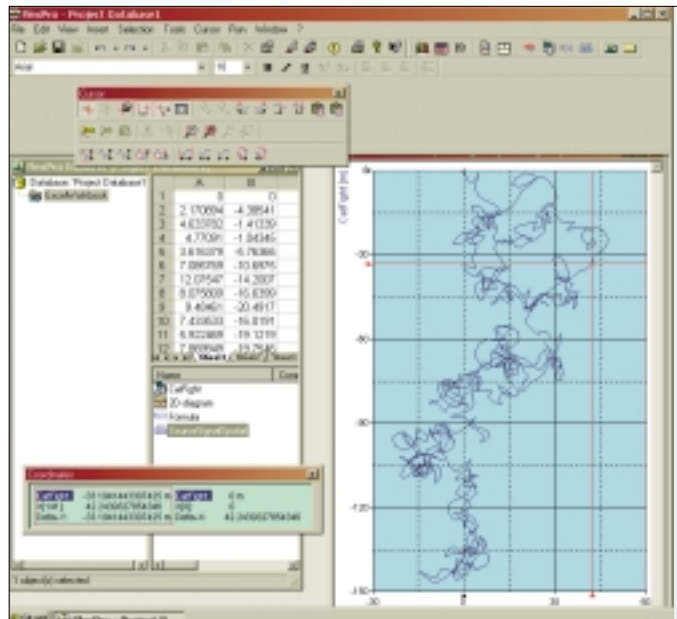
Felix Grant finds the simplicity of FlexPro 5.0 belies its hidden depths

FlexPro 5.0, from Weisang and Co, is one of those products which aim to serve an often ignored range of data users: those who, in FlexPro's words, are interested in 'documenting, analysing and archiving data in the simplest way possible'. The online help system is clearly designed to promote the product in this market segment, with a very clear introduction from first principles and a hands-on tutorial, and the live project to which it was applied was selected with this in mind.

The performance of dry cell batteries in handheld field equipment is of crucial importance to my work; hundreds of thousands of datapoints, generated by colleagues and myself, accumulating in sundry data stores over a period of more than five years, were overdue for organisation. Yet actual battery planning and usage were still based largely on instinctive and anecdotal bases. If it was ever to get properly done justice to, this vital but neglected chore certainly cried out to be done 'in the simplest way possible'. The arrival of FlexPro seemed like an omen: the time for documenting, analysing and archiving that morass of operational detail had arrived.

Statistics packages tend to present themselves, at least on first glance, as spreadsheets. This is practical marketing, since the spreadsheet is familiar to most users. On the other hand, it is also a conceptual sleight of hand since the data structure has more in common with a database, which those users find far more daunting. There are products that bravely make this fact explicit, but they are rare. This is the first crucial difference between FlexPro and the general run of data analysis products: it presents not as a specialised spreadsheet, but as a specialised database. Data, analyses, formulae and visualisations appear as packaged objects, which can then be manipulated without direct reference to their contents.

There is tight integration with Excel, so users do not have to abandon their existing data



Mapping an invisible event: I set up a sound-triggered system in a concrete yard, the site of nightly mating-season contests among local cats. The WAV file, imported as a raw signal, was processed to generate moving mean points whose amplitudes were converted to postulated (x,y) co-ordinates and fitted with a spline curve, providing an approximate 'map' of the fight.

container, but incoming data are split up into several strictly typed data sets in FlexPro. These data sets are then saved together in a FlexPro database, along with other structures and objects created thereafter. An Excel workbook opened in the FlexPro-Explorer triggers background initiation of Excel (97 or later) itself, running within FlexPro. The workbook appears first in its native form, including sheet tabs, but the data are assigned to data sets as they are selected for FlexPro operations.

Excel is not mandatory. Data import can be direct from text files, ODBC sources and a list of file types not usually found in generic analysis products. Gould oscilloscope files are on offer, for example. Forsaking the batteries for a while, I have spent some time on exciting and highly encouraging experiments with microphone and other transducer data recorded to high quality (128 44) MP3 files on a cheap and lightweight 20 gigabyte consumer jukebox machine and then imported to FlexPro as a WAV file. Data can also be dragged, or copied and pasted, from most grid-oriented programs via the Windows clipboard. Given the claims to data archiving, I tested the mention of data sets one million cases in size with a 30-variable example; then doubled it; FlexPro didn't seem

to mind. I suspect that the way variables are handled (as database items inert until acted upon, rather than constantly live across a computational grid) is responsible for the marked lack of degradation in performance at this size.

The choice of supported file import formats is one of several features that reveal a central concern with handling automated metrological logging. There are data structures not familiar to spreadsheet users as well, aimed at applied data usage. Chief amongst these is the 'signal', which pulls in two or three data series comprising an (x,y) pair or (x,y,z) triple and packages them. A signal can, therefore, be displayed and handled as a rational entity without explicitly specifying the series involved. The name 'signal' is obviously derived from relevance to direct data logging devices but the structure itself is useful in many ways, and for many purposes, well beyond those literally implied. In the task on hand, for example, the representation of data series groups such as (days, volts) and (Days, %life, %voltage) as single signal structures as 'DaysVolts' and 'Day%Life%Volt' simplified handling and discussion of many issues. The option to aggregate data series into signals is offered as

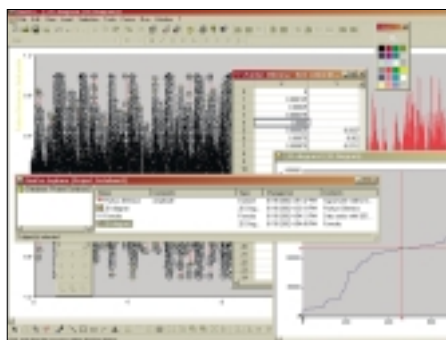
part of data import, in many cases; and assignment of one data set to another is an alternative route.

Data structures can be operated upon by formulae, the distinction between the two blurred by the fact that a formula often produces, or functions as, a new data set. The spreadsheet variable '%voltage' (actual voltage as a percentage of initial value) in my data, for instance, could be generated with FlexPro using the formula $100 * \text{Maximum}(\text{Volts}) / \text{Volts}$ – this formula could then be treated as if it were a data series. The separate series within a signal structure are not lost; they can be accessed individually from within the structure when required, using formulae such as DaysVolts.x or DaysVolts.y . Conversely, the DaysVolts signal structure itself can be emulated from the separate series using the formula $\text{Signal}(\text{Days}, \text{Volts})$ which then becomes a similarly manipulable object.

With the data objects in place, visualisation and analysis objects can be added. I use that form of words deliberately, to reflect the philosophy at work; you will find 'statistics' not in its familiar position on the top menu or the 'tools' drop down, but under 'insert'. Analysis objects can also, it will come as no surprise to hear, become data sets.

In analysis, once again, there is the clear orientation towards automated metrology. Data series can, metaphorically, be seen as wavicles; statistics software tends to view data as particulate, with continuity taking a second order place in their scheme of things, but that way of looking at things is reversed here – the wave leading, the particle serving as reference.

The 'Event Isolation' object, to take a more or less random example, pulls out occurrences within the data set that meet criteria of specified interest. Instances where values meet or cross class boundaries or other specified threshold levels; local maxima or minima; specified rates of change in variable value; and so on. The results comprise a new data set, or just mark up the original in various ways. The 'extrema' function searches for local minima, maxima, or both, in a data set, generating a new result



Partly as an familiarisation exercise, partly for illustration purposes, and partly for relief from endless battery data on a lovely sunny day atop a cliff, I analysed a section from Simon and Garfunkel's /Punky's Dilemma/ (the section which starts 'if I were a first lieutenant...', for any fans out there). In the background, a detailed plot of the sampling data (seen tabulated in the x/y block just right of centre) is overlaying the raw plot of the signal wave. In front of that, a cumulative plot of crossings at the -0.75 level; in the foreground, the database listing of the objects involved.

subset. Arguments to the function include hysteresis (allowing control of the degree to which values may rise or fall alongside a peak or trough before it is identified as a local extremum) and control over whether the result set includes (or excludes) only the extrema themselves, or identifies them through (for example) indices.

Graphic objects are broadly what you'd expect, with few surprises: line, spline, column, bar, area, colour field matrix, 3D and contour surfaces, 3D curve, bubble, vector, error limit, each with up to eight subtypes. I was surprised to find no true histogram, although spreadsheet users will be used to working around this one and accept the lack as a price of the other advantages on offer.

Documentation is as important to FlexPro as analysis, and is supported by a document editor which talks directly to both internal data objects and external sources. Drawings, text, diagrams and tables can be stored as objects in

the database or embedded from other originating programs. The document itself is one more database object. Material generated in FlexPro can also be embedded in external programs (word processors, for instance, or web page editors). References back to values from within text can be interactively generated using a trace cursor on a graph or chart, without necessity for opening the data sets concerned; there is a worksheet view which allows, amongst other things, simultaneous access of this kind to multiple objects.

Because it offers and delivers simplicity I made the serious mistake, at first approach, of thinking this a simple and straightforward package. Nothing could be further from the truth; the more I became acquainted with it, the more intriguing its hidden depths and byways became. It is unique in several ways and distinctive in several more. The object structure, in conjunction with a facility to 'activate' and 'deactivate' data folders, makes manual reuse of analyses natural and painless; but ActiveX and DDE interface are provided through which automation and external control are possible.

And the batteries? Yes, we learned a great deal, and sustained a number of surprises, about effective procurement, storage, utilisation and replacement regimes. Changes are on the way, in anticipation of significant cost reductions and operational efficiency improvements.

• **It takes a while to put a piece of software properly through its paces. Just as this issue of *Scientific Computing World* was being passed for press, Weisang informed us that FlexPro 6 was being made available.**

Its main new features are HTML export and the automation object model with Visual Basic integration. The HTML export wizard will make it much easier for users to share their results by using a standard internet browser. FlexPro 6 now has a built-in macro-recording and playback feature, so that any user interaction with FlexPro can be recorded as a macro. FlexPro 6 primarily addresses industrial customers where automation is of high importance. More details to follow.

The competition

The most obvious comparison to FlexPro is StatView (see feature on selecting and using software, *SCW February 2001*). Like FlexPro, it takes an object oriented, variable-centred view of data and analyses and the mode of use is very similar. Its analysis and documentation repertoires are very different, however.

Another product which approaches the operational analysis of data objects in a comparable manner, and further extends the 'signal' idea, is DaDiSP (see 'Looking

Beyond Excel', *SCW January 2002*). This is an even more radical departure than FlexPro, and not for the faint-hearted. The truest equivalent is an analytic package (or a graphics product), its data stores located in a dedicated database manager, linked to a word processor or DTP package. Properly done, this gives the same facilities and potentially greater reach – but will be less convenience and cost far more than Flexpro.

The 'lite' alternative: A parallel product from Weisang, FlexPro-View, offers all the

documentation options of FlexPro but a very restricted range of analysis functions, and without the interface customisation and user-admin options. FPV appears a useful supplement to the full version; sitting on satellite workstations, it would be an economic expansion of data access to consumers. Unfortunately, this is not feasible since databases created in the full version cannot be opened in FlexPro-View. The 'lite' users would have to comprise an initial archiving and documenting operation, or a separate one with its own data requiring only nonanalytic handling.