

# Tessellations

News And Technical Updates From Tessella

Edition 63, Spring 2008

## Groundwater modelling for the 21<sup>st</sup> Century

'Groundwater modelling' involves modelling what water actually does in the ground, taking into account the geology and physics of each location, together with the flows of water from:

- Precipitation (e.g. past and predicted rainfall) and evaporation
- Rivers, streams, lakes and reservoirs (including their location and how much water they contain)
- Abstractions (including removing water from the ground via boreholes)

Groundwater models produce details of the flows in and out of 3D sections of ground, and estimate the amount of water contained within a defined volume of ground. The Environment Agency for England and Wales (EA) uses groundwater modelling to:

- Develop mitigation scenarios for possible droughts (an issue that seems to alternate in importance with flooding!)
- Assess the impact of abstracting large quantities of water (e.g. when the EA are considering an application to sink a new borehole, or when a farmer applies to increase the volume of water removed from an existing borehole)
- Enable it to comply with the European Union Habitats Directive (which aims to protect the wild plants, animals and habitats that make up our diverse natural environment)

### Working in isolation

Groundwater models have been in use by the EA for many years. However, until recently, models were typically developed individually (depending on the knowledge of specific staff for their continued use) and ran on isolated dedicated environments on ageing IT systems.

Numerous modellers were working around the regions of the EA, and within the UK water companies, however it was difficult for good practice, models, systems or results to be shared and exploited for common good. The EA could see the potential benefits of bringing the groundwater modelling communities together.

### Exploiting the NFFS

The EA's National Flood Forecasting System (NFFS<sup>\*</sup>) is now in use throughout England and Wales, and proved valuable during the July 2007 floods of the Humber, Severn and Thames rivers. A key aspect of the NFFS is that it allows models to run under a managed architecture, without the individual modeller needing to worry about *how* or *where* their model runs.

\* The NFFS is based on the Delft Hydraulics FEWS system.

It was realized that the architecture underpinning the NFFS might be able to be adapted to support groundwater modelling. Building on the partnership delivering the NFFS, Tessella and Delft Hydraulics have supported the EA's investigations throughout, working with them on the feasibility study, requirements collection, design and pilot, and now the implementation stage of the National Groundwater Modelling System (NGMS) for England and Wales.



The feasibility study proved that the NGMS could indeed be built on the existing architecture. The requirements phase established what groundwater modellers (and their models) needed, which was different

from flood forecasting. By the end of the design stage, the NGMS architecture was in place, with a 'central server' located in Peterborough and with model runs farmed out to a number of 'model servers'. As part of the system's roll out, more models were migrated from April 2007 onwards to run on the new system, allowing the pilot system to evolve into a full-blown implementation (scheduled to continue until 2009).

In parallel with the IT development, the EA's groundwater modellers have been refining their working practices to enable them to take full advantage of the new system, including developing standard nomenclatures for terms, and standard GUI interfaces for their models.

### The technology

The NGMS provides a centralized managed run-time environment for the execution, storage and analysis, of models and results, based on standard IT hardware and infrastructure (which is therefore easier for the EA's IT staff to support). Modellers can submit their jobs to run either immediately or at a pre-defined time, and can easily view input and results datasets, and extract subsets of data. There is central storage of model datasets, results, reports and configurations. The system includes good visualization facilities for input and results datasets. Due to the large input/output datasets, efficient data transfer and synchronization mechanisms have been expanded from their original implementation in the earlier Flood Forecasting System.

### Reaping benefits

The introduction of the National Groundwater Modelling System for England and Wales will ensure a greater standardization of future groundwater models within the UK, which promises to make them easier to keep up-to-date and ready for operational use. The wider availability of tried and tested models will enable more staff to use them to make key regulatory decisions. To find out more please email [info@tessella.com](mailto:info@tessella.com)

**ARCHIVING NEWS:** Building on our experience of working with the UK, US, Netherlands and Malaysian national archives, on their digital archiving challenges, Tessella is delighted to be supplying our **Safety Deposit Box** technology to the British Library, working with the Swiss Federal Archives on the design for their digital archive, and consulting to the UK Parliament regarding their preservation needs.

## Optimizing drug discovery

The process of drug discovery requires a combination of scientific knowledge and inspiration to guide the investigation, together with a great deal of perspiration, time and money to carry out experiments. High Throughput Experimentation systems are increasingly being used to reduce the time and effort to identify potential new drugs (known as 'drug candidates').

GlaxoSmithKline (GSK) is a world-leading research-based pharmaceutical company. At the heart of their business is the need to discover compounds that are biologically active so that they can then be developed into drug candidates.

GSK's Technology Development Department was investigating new platforms for the synthesis and bioassay of novel compounds. These platforms were based on microfluidic technology, and hence consumed much smaller quantities of materials than conventional systems, and results could be obtained much more rapidly. To create a fully-integrated closed-loop synthesis-assay system, an optimization module was required. It would contain a selection algorithm that would decide which compound – of a potentially very large pool of 'virtual' compounds – to make in the next synthesis/assay iteration. The algorithm would also be able to base its selections on data from all previous iterations, and thus discover compounds with high biological activity as quickly as possible.

### Applying academic research

Based on some existing literature on optimizing searches in combinatorial chemistry\*, GSK chose to use a genetic algorithm as the basis for one of their optimization modules.

Genetic algorithms use strings of letters to represent individuals in a population. The letters represent a code that defines the characteristics of the individual, in the same way that the 'letters' of a DNA molecule represent genetic information. Genetic algorithms mimic biological processes by crossing the strings of individuals (exchanging sequences of letters between strings) and applying other genetic processes like mutation. The probability of an individual surviving to the next 'generation' and being allowed to cross depends upon its fitness, in other words how well it performs in the problem that you are trying to solve. The result is that each generation adapts and becomes fitter than the previous, by a process similar to natural selection.

The automated discovery system represents each product as an entity – with 'DNA' and a related fitness. The reagents used to make a particular product are equivalent to its genes. Having tested a number of products, their fitness can be assessed by the experimental results. The genetic algorithm is used to produce the next generation of products to be tested. As the process is repeated the level of measured biological activity will increase.

### Developing a software solution

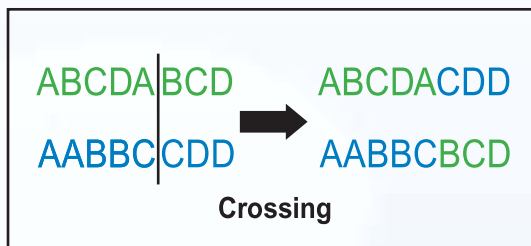
The scientific backgrounds of the Tessella team members significantly aided the process of translating the academic paper into a software solution and allowed the system to be developed interactively with GSK's staff to ensure that their business needs were met. Tessella's previous experience in developing optimization systems allowed the team to build on the basic algorithms documented in the academic paper. The resulting Genetic Algorithm Optimizer not only meets the original use case, but also provides a flexible optimizer which can be applied to other scientific problems.

**Microfluidics** is a multidisciplinary field which intersects physics, chemistry, microtechnology and biotechnology. It has practical applications for the design of systems in which microlitre and nanolitre volumes of fluids will be used.

In **combinatorial chemistry** known chemical entities are combined in as many different combinations as possible, potentially resulting in thousands, even millions, of new chemical compounds. These compounds can then be tested for efficacy as drugs; a process called **high throughput screening**.

The new system can optimize a wide range of search domains including discrete, continuous and grouped variables, allowing searches to be optimized for both combinatorial and parallel synthesis. The system makes use of a number of innovative features developed by Tessella and GSK team members working in close collaboration, including:

- The ability to identify significant groups of reagents
- As a search proceeds, the system automatically modifies the way in which it searches to achieve an efficient searching strategy
- The ability for chemists to specify the level of similarity between reagents or between products
- The ability to manage experimental errors



*Genetic Algorithms – the fittest members of the population, represented by strings, are crossed to produce new, possibly better, individuals.*

No optimization system can replace the expertise and creativity of an experienced chemist. However, the Genetic Algorithm Optimizer has been designed to allow chemists to enhance their searches by providing chemically relevant information about the reagents and products. The system also provides feedback that assists the chemists to understand better the underlying chemistry and why certain products are biologically active.

### Achieving even higher efficiencies

It is important to note that in many applications, genetic algorithms are run to convergence. To be effective in lead optimization, useful results must be obtained in a relatively small number of iterations (generations), even with new technology platforms that may reduce the typical synthesis/assay cycle time. After validation runs using historic or *in silico* generated data, the new software module was used in a real optimization experiment to iteratively select ten generations of compounds for synthesis and assay. This demonstrated that, in combination with suitable hardware, statistical analysis and visualization tools, the Genetic Algorithm Optimizer has the potential to be an effective tool for lead optimization.

Further developments are possible using novel optimization techniques to achieve even higher efficiencies than those achievable with conventional genetic algorithms. This should allow scientists to assess the relationship between different reagents, products and their properties, in determining the levels of biological activity, further enhancing the chemists' understanding of their drug discovery problems. To find out how Tessella can help with your potential requirements, or to request a technical supplement on HTE, please return the enclosed order form or email [info@tessella.com](mailto:info@tessella.com)



David  
Project Manager  
Tessella

(\*Illgen, K.; Enderle, T.; Broger, C., and Weber, L. Chemistry & Biology 2000, 7, 433-441)

*If you would like further information, or back issues of Tessellations, please email [info@tessella.com](mailto:info@tessella.com)*

## Improving decision making – via real-time insight

In many organizations – both research and commercial – there is a need to deal with ever-increasing volumes of data. Interpreting and analyzing such data is a challenge in itself, but for many event-driven applications the data must be interpreted and responded to in real time.

Complex Event Processing (CEP) is a technique which enables organizations to extract key information from streams of real-time data, as soon as that data becomes available.

Each new data item can be thought of as an *event*. Whilst an organization is usually not interested in any one of these 'simple events', they can potentially benefit from watching for patterns of simple events; e.g. two or more specific events arriving within a certain time, or in a given order, or having certain characteristics. A set of simple events that match a pre-defined pattern can be treated as one 'complex event', which can highlight both opportunities and threats to the organization.

To give a concrete example, consider credit card fraud detection, where the simple events are purchases being automatically reported back to individuals' banks. A CEP system can watch for the same credit card being used within an hour at locations more than 100 miles apart, such that this complex event can be flagged up as highly suspicious. More subtle patterns – such as purchases which are much larger than average for a particular cardholder or at locations they don't usually frequent – can also be watched out for. The response taken by the bank will depend on their interpretation of the severity of the complex event.

A CEP system gives its users the power to specify rules for recognizing specific patterns in the data stream, watches for and detects those patterns in real-time from the incoming flow of data, and responds appropriately depending on the type of complex event it captures.

In the main, CEP was initially used for highly-specialized mission-critical applications, however the potential of the technique is now being appreciated in an ever-growing number of industry sectors and for a widening range of applications. As the volume and complexity of the data that we generate grows, and as both business and consumer increasingly expect 24/7 operations, being able to react quickly gives organizations a key advantage.

In the **FMCG** (Fast Moving Consumer Goods) sector, there is much interest in the possibilities of using CEP to track the progress of consumer products through the supply chain, monitoring data from RFID tags to find and react to problems or inefficiencies immediately.



Within the **transport sector**, Southern and HSBC Rail are turning raw operational data into the intelligence they require to streamline and maximize the benefits of their rolling stock maintenance regime. The award-winning 'Train Automatic Performance Analysis System', featured in Tessellations 62, is being used to monitor remotely and report the condition of the fleet of trains, each fitted with an 'On-Train Monitoring and Recording' data logger.

Within the **chemical sector**, Tessella has employed CEP techniques to build a proof-of-concept application for the automatic monitoring of safety systems in chemical process plants. The prototype allows signals coming in from various instruments around a plant to be interpreted into higher-level complex events which are used to verify that systems are responding properly. The gathered data can be used to optimize testing schedules, leading to significant savings from reduced plant downtime.

Within the **space sector**, CEP techniques are frequently employed in Failure Detection, Isolation and Recovery systems for spacecraft, where potentially mission critical failures must be detected and acted on appropriately to safeguard the system. Tessella has developed such techniques on a number of ESA missions, analyzing the data streams from attitude sensors and actuators to detect and identify failures and initiate the appropriate corrective or remedial action.

In any industry where there is a wealth of data to sift through, and responding to it rapidly is a priority, CEP can provide a way to define what you are interested in and give a centralized overview of the relevant facts as they happen. By gaining real-time insight, improved decision making and an agile business can become a reality.

Tessella has 27 years experience of developing bespoke/custom software and integrating commercial-off-the-shelf components to produce unique solutions.

We are vendor independent and always offer the solution that is best for our clients.

To discuss your potential requirements please email [info@tessella.com](mailto:info@tessella.com)



Neil  
Associate Analyst  
Programmer, Tessella



### Innovative Chemical Registration

In 2004, UCB, the Brussels-based chemicals and pharmaceuticals group, transitioned itself into a pure biopharmaceutical company, with a focus on both antibody-based and small chemically-derived drug discovery. In that year UCB divested its surface specialities and films businesses and acquired Celltech, the UK's leading biotechnology company.

As part of the process to integrate Celltech, UCB reviewed the suite of software systems in use at each company. It was decided that neither of the two companies' existing chemical registration systems were suitable for adoption by the new company. The systems were closely aligned with legacy business practices and built with ageing technologies.

No commercial off-the-shelf system met the expanded company's business model sufficiently. Hence, in 2005, Tessella was asked to help design and implement a custom-built chemical registration system, which was:

- Designed to be 'material-oriented' rather than 'structure-oriented' – in order that data are associated with the physical material rather than the chemical structure
- Built on modern IT architecture principles, such as service-orientation, technology independence, and messaging, to aid integration with other internal systems and simplify system deployment

The key innovative step taken by UCB was to implement a dynamic association of chemical structure and material, and to associate all data with material assayed. This facilitates the capture of increased knowledge of the chemical make-up of the material over time, without compromising data integrity.

Using a phased rollout, the first of UCB's legacy systems was replaced within a year. Old and new data was also migrated from the second legacy system. Following further enhancements, the final rollout of the new chemical registration system was successfully delivered in summer 2007, on time and on budget.

UCB is dedicated to the R&D and commercialization of innovative products in three key therapeutic areas, and has made use of Portfolio Management tools developed with Tessella.

For further details please see the client story in Tessellations issue 62.

### Groundbreaking Clinical Trials

Adaptive Clinical Trials (ACT) are a groundbreaking progression, attempting to increase the efficiency of clinical trials by breaking out of the confines of conventional experimental design. For almost a decade, Tessella has been working with pharmaceutical clients on adaptive trial designs, including supporting four of the top-ten international pharmaceutical companies.



Tessella continues to win new clients in this area, and now has projects underway with two further top pharma, helping them to run adaptive dose finding studies.

For a good introduction to ACT please see Tessellations 57. Readers with a more advanced knowledge of ACT should find [www.tessella.com/Services/Discipline/AdaptiveTrials.htm](http://www.tessella.com/Services/Discipline/AdaptiveTrials.htm) of interest.

### Explosives Risk Management

In meeting home and overseas military demands, the UK Ministry of Defence (MoD) has a duty of care to ensure that the highest safety standards are applied in Explosive Risk Management. This presents significant challenges, particularly when potential threats come in many forms.

The MoD's Defence Equipment & Support, Defence Ordnance Survey Group (DOSG) is charged with providing assurance on the safety of munitions sites, including ensuring that the tools for performing Quantitative Risk Assessment (QRA) are fit for purpose.

The kernel of a site QRA is the calculation of the risk of death or injury as a result of an unplanned explosion. Tessella has worked with DOSG for a number of years, and more recently we have been helping to identify the requirements for a new software application to support the QRA process. The outcome will be a more modular and flexible tool.

### Tessella – Providing innovative solutions to scientific, technical and engineering problems

Tessella uses its unique blend of scientific, engineering and IT skills to solve the most complex of technical and business problems in a highly cost-effective way. We have a proven 27-year history of excellence, adding value to demanding public sector and commercial R&D based customers.

Tessella comprises Tessella Support Services plc and Tessella Inc. Our space and defence business, Analyticon, is now fully integrated into Tessella.

The group's services include software design and development, mathematical modelling and simulation, algorithm development, infrastructure support, project management and consultancy.

Our enviable reputation for providing high-quality, low-risk, value for money services is backed up by many successful, high-profile projects, plus a high level of repeat business.

For each client problem we develop a fundamental understanding within the 'big picture' context – so our solutions fit. We focus on the details (however intricate) so our solutions work. Our ultimate aim is that the systems we deliver are used by our clients with great enthusiasm.

\*\*\*\*\*

[www.tessella.com](http://www.tessella.com)

[info@tessella.com](mailto:info@tessella.com)

\*\*\*\*\*

#### Abingdon, UK (Head Office)

Tel: +44 (0)1235 555511

#### Burton upon Trent, UK

Tel: +44 (0)1283 553300

#### Cambridge, UK

Tel: +44 (0)1223 897770

#### Stevenage, UK

Tel: +44 (0)1438 749886

#### Warrington, UK

Tel: +44 (0)1925 286800

#### Winchester, UK

Tel: +44 (0)1962 850055

#### Den Haag, the Netherlands

Tel: +31 70 354 2296

#### Boston, USA

Tel: +1 617 454 1220

#### Washington DC, USA

Tel: +1 240 235 6052

**Tessellations** is published by Tessella Support Services plc. Our aim is to provide you with interesting information on topical technology issues and to outline key projects which we hope you will find of use. We depend on the feedback from our readers to help us develop Tessellations and to maximize its usefulness. Your input is always appreciated; please send to The Editor (Alison Smith) at [info@tessella.com](mailto:info@tessella.com)