Julie Wilson

What's your name and can you tell us a little about yourself and your field of study?

My name is Julie Wilson, I am a professor in applied statistics in the Department of Mathematics at the University of York. My research interests lie in the application of mathematical modelling and statistical methods, often applied to large biological and chemical data. I develop methods for data pre-processing and analysis and have experience in statistical pattern recognition, classification and machine learning techniques.

Can you give us an overview of the project?

This is an interdisciplinary project in collaboration with my PhD student, Laura Wiggins, who is co-supervised by Dr William Brackenbury in Biology and Dr Peter O'Toole of the Technology Facility. We are characterising cells based on their morphology and movement monitored by time-lapse imaging. The software imports tracking information from multiple segmentation and tracking algorithms and includes automated recognition and removal of erroneous cell boundaries. We provide an extensive list of features extracted from individual cell time series, with custom feature selection to identify variables that provide greatest discrimination for a particular problem. For example, we have been classifying different cancer cell lines that have been treated with different chemotherapy drugs.

Once you started working with the data did you encounter any specific problems or challenges? How might an RSE have been a useful addition to the project?

The data analysis was not a problem, but the code had evolved throughout the project. We solved problems as we came to them using the programming language that seemed most appropriate for the task: Matlab as we could use existing code, C for fast image processing and R to exploit existing packages for classification. The need for different platforms was a major limitation to making the software available to others and working with an RSE earlier in the project might have led to more coherent planning.

What was the benefit of working with an RSE? Were there specific tools, software or outcomes you found that an RSE could provide?

We worked with both Killian Murphy and Stuart Lacy and their knowledge has been invaluable. After originally trying to hang on to long and complicated C code, they eventually convinced me to rewrite the software in R. Having not done much programming in R before, my code was not very efficient and Stuart was able to massively reduce the computation time. In fact, there were many specific tools, software or outcomes that Killian and Stuart either led us through or even did for us, from R packages and writing documentation to Github and Shiny App development.

What tools and software did you use for your analysis? Is the software, code and data that you used available for others to reproduce your work?

Yes, this was exactly how working with Killian and Stuart helped. Thanks to them we now have an R package, CellPhe, that we can make available to others, not just to reproduce our work, but to use for their own research projects.

Having worked with an RSE, will it change your approach in the future?

Yes, I would consider how any programmes I might develop as part of my research might eventually become software packages.

Can you tell us more about your current or any future research projects?

I work on various interdisciplinary projects in collaboration with colleagues across the University, but particularly Chemistry, Biology and Computer Science. I also work with external partners and current industrial collaborations involve Fera Science, AstraZeneca, Croda and Eluceda.