

Impact case study (REF3b)

<p>Institution: Edge Hill University</p>
<p>Unit of Assessment: 11 - Computer Science and Informatics</p>
<p>Title of case study: Improving quality assurance of a large software model through relative debugging.</p>
<p>1. Summary of the impact As a result of collaborative commissioned research, the lead developers of a major atmospheric research and operational weather forecasting model have changed their approach to quality assuring model source code. Drawing directly on the research findings, the lead developer has taken the decision to adopt a new approach to the correction of inconsistencies and inefficiencies in source code and to alter the software build procedure to be followed by a large model development community. An additional impact, in the form of improved business competitiveness, is felt by a British software and consultancy company, which has been able to enhance a key tool used in their quality assurance and platform migration work with a global client base.</p>
<p>2. Underpinning research The findings underpinning the impacts described in Section 4 are the product of research undertaken at Edge Hill University (EHU) by Dr Mark Anderson on the behaviour and structure of the source code of the Weather Research and Forecasting (WRF) model. Work commenced in June 2011, ending in June 2013. Anderson joined EHU in June 2008 as a Senior Lecturer, was employed by EHU continuously throughout the assessment period and remains employed in the Department of Computing as a Reader.</p> <p>The WRF model (a next-generation mesoscale numerical weather prediction system) serves the atmospheric research and operational weather forecasting needs of its users. Development has been led by the National Center for Atmospheric Research (NCAR, US), the National Oceanic and Atmospheric Administration (US) and the US Air Force Weather Agency, supported by a wider community of developers. The model has been designed to support implementation on a broad set of hardware (single processor laptops to multi-processor super computers). Its modelling system is complex and exploits many Fortran capabilities that permit compile-time checking, including modules and interfaces. Code started development in 1998 and is largely Fortran compliant, evidenced by the number of successful compiler ports. However, capabilities and code have been added by many different contributors in the development community. The complicated nature and the origin of the source code means that the software has accumulated inconsistencies and inefficiencies with respect to the strict Fortran standards.</p> <p>Anderson's research was undertaken as part of a project led by the University Corporation for Atmospheric Research (UCAR, a consortium of 100+ universities and academic affiliates) and funded by the US Air Force. Anderson used a FPT toolkit developed by SimCon Ltd to explore the extent and nature of inconsistencies within the WRF model. This required the manual instrumentation of the code and execution of a limited test case (covering 18% of the model's code). The outcome of this investigation included the finding that inconsistencies in output generated by the model could be caused by the absence of a single and uniformly applied set of Fortran code standards, compiler inconsistencies or by building the code on different systems/hardware. Some of these inconsistencies were found to be highly significant, creating exposure to model problems due to inconsistent treatments within the source code e.g. branches of code being executed in different environments using the same input data. These findings are reported in research output 1 in Section 3 below.</p> <p>Further work was then commissioned, funded by the US Air Force and collaborating with SimCon Ltd who would develop an enhanced FTP toolkit and visualisation tool. A Research Assistant (under the direction of Anderson) was appointed in April 2012 to configure the software environment, perform test runs, determine differences between the outputs and develop a software tool to visualise the differences between outputs in the model. The EHU testing environment was reconfigured to integrate an extensive test suite developed within UCAR to increase the coverage of the tests. Anderson worked with the development leader at UCAR to effect these changes to the Edge Hill testing environment, facilitating extended coverage of the</p>

Impact case study (REF3b)

tests executed from 18% to 80% of code and extending the analysis of the code base to include chemistry and physics modules used by the user community. The findings from this research, reported in research output 2 in Section 3, were the classification of the issues that existed across the suite of modules within the WRF model. This led us to the development of the automated test procedure which would support testing by the UCAR developers in future releases of the model to the user community.

3. References to the research

The quality of the research is evidenced in the project's inclusion in UCAR's successful funding application to the US Air Force, and additional funding for the follow-up work. The work had influence on policy and practice for UCAR and the wider development community as the WRF development lead, leading to changes in the development processes for the WRF model, enhancing the testing procedure to align with practices adopted in the development of engineering software models for incremental code releases. Work was scrutinised by target audiences in both the engineering and academic communities. Outputs accepted following peer-review.

[1] Conference Contribution, Journal Article: Collins, J., Farrimond, B., Anderson, M., Owens, D., Bayliss, D. and Gill, D. "Automated Quality Assurance Analysis: WRF – a case study". In Proceedings of 5th International Conference on Computer Engineering and Technology (ICCET), Vancouver, Canada, 13-14 April 2013 (also *Journal of Software*, JSW, ISSN 1796-217X). doi:10.4304/jsw.8.9.2177-2184

[2] Journal Article: Collins, J., Farrimond, B., Flower, D., Anderson, M. and Gill, D. "Removal of Numerical Drift from Scientific Models: A case study using WRF". *International Journal of Software Engineering and Applications* (IJSEA), 4(2), pp 51-61, March 2013. DOI 10.5121/ijsea.2013.4204

4. Details of the impact

The WRF model was developed as a collaborative partnership between federally funded US agencies, principally NCAR, the National Oceanic and Atmospheric Administration, US Air Force Weather Agency, Naval Research Laboratory, and the Federal Aviation Authority. It has a worldwide community of 20,000 users in 130 countries, is used extensively for research and real-time forecasting and is the most widely used Numeric Weather Prediction model in the world. Users include organisations engaged in climate and weather research, as well as civil (government, industry) and military organisations who use the model to inform operational decisions. Prominent users include US National Centers for Environmental Protection, other forecasting centres (including the Korean Meteorological Administration), the Japan Agency for Marine-Earth Science and Technology, NASA, US Air Force Weather Agency, and US Army Research Laboratory.

Forecast accuracy and reliability is vital to users (and to the communities and public they serve). EHU's collaboration with SimCon has shaped the way in which the model is developed and built in order to improve the quality of the WRF model source code. Impact will increase in reach and significance as the enhancements and new procedures are implemented with the development community and reflected in improvements to output generated by and for the model's users.

The following impacts occurred as a result of findings in the underpinning research between June 2011 (when research commenced) and 31 July 2013:

Impact 1: An enhanced FPT toolkit is now being integrated into the development of new versions of the WRF model .

EHU and SimCon collaborated on the evaluation and analysis of the testing output, identifying additional features which needed to be implemented within the FPT toolkit to cater for updates to the Fortran coding standards (between FORTRAN 77 and Fortran 90). The EHU team also demonstrated the value of integrating automated testing utilising relative debugging techniques to assist in the identification of issues in incremental releases of the WRF model. This has resulted in a more consistent output being generated by the WRF model irrespective of the underlying hardware architecture or compilers used by the developers to implement the model. As a result

Impact case study (REF3b)

the toolkit was updated by SimCon and adopted by UCAR in July 2013. UCAR are currently planning the roll-out of an amended build procedure with the development community.

The initial beneficiaries of this impact are the software development team at UCAR which implements and maintains the code base. Outside the assessment period impact will widen to benefit the wider WRF development community, WRF users and those who rely on WRF model output.

Corroboration is provided by Factual Statement 1.

Impact 2: The WRF model's development lead has commissioned changes to the software build procedure.

The underpinning research demonstrated that issues existed within the WRF model code, that an automated approach to relative debugging would aid the removal of such issues, and that the existing build procedure would not support the automated instrumentation of the WRF model code.

Drawing on these research findings the development lead (at UCAR) for the WRF model took the decision to commission changes to the build procedure for the code base in the following ways:

- The compilation of modules within the WRF model was initially implemented in a hierarchical process – the EHU team demonstrated that this resulted in an unwieldy build procedure. As a result SimCon developed a more linear build procedure that facilitates the integration of automated testing
- The implementation of the build procedure made use of pre-processing to modify the source code to support the broad range of environments in which the model has been migrated.

This decision was taken in Summer 2012 and the work to implement this change is ongoing. The decision was taken in order to provide the WRF developer and user communities with new, more effective and more efficient ways to identify and reduce/eliminate non-standard Fortran, reducing exposure to model problems due to inconsistent treatments within the source code.

Corroboration is provided by Factual Statement 1 and Factual Statement 2.

Impact 3: SimCon have improved their WinFPT toolkit and can deploy it in other projects

Simcon Ltd is a British software products and consultancy company, specialising in real-time simulation and scientific computing. Its key service areas are Fortran Analysis, Engineering and Migration; key customers are Airbus, Boeing, DENEL OTR, QinetiQ and the US Navy. Simcon collaborated with Edge Hill on the WRF software validation project described above. The company deploys its WinFPT toolkit in their work on the WRF model and in projects for other clients (particularly migration and quality assurance). In the course of the software validation project, the research undertaken by Edge Hill enabled SimCon to make a number of improvements to this toolkit.

The most important development was the addition of cpp pre-processor support to the toolkit. Programs such as the WRF use the cpp pre-processor to adapt the code to different parallel computing environments. Before this project SimCon's WinFPT toolkit could analyse the WRF program after pre-processing but could not re-engineer WRF because any changes made by WinFPT would be overwritten by the pre-processor every time WRF was rebuilt. Errors could be found but could not be fixed, and because correction could not be automated the correction rate was lower than is now possible.

Collaborating with UCAR and SimCon, Anderson designed and implemented the extended test suite and then undertook test runs of the WRF model, identifying the relative difference in output generated by the model between test runs and determining the significance of these differences. As testing proceeded it was found that certain conditions were not being correctly identified by the WinFPT toolkit as the WRF model utilised later coding standards (Fortran 90). Working with SimCon the relevant updates to the standard were identified and the toolkit was then updated by SimCon to support new language features.

Impact case study (REF3b)

The improvements made to WinFPT now allow permanent changes to be made in the WRF code base. Following incremental testing and analysis at EHU of these updates, the addition of cpp pre-processor support was completed in July 2013, making significant contributions to the validation of different compilers for WRF. As a result, the way in which errors in the WRF code base are fixed, code is re-written, and the extent and implications of changes made are identified has been changed. The direct beneficiaries of this impact are SimCon Ltd and WRF user and developer communities. SimCon are now able to use the enhanced toolkit in their migration and quality assurance projects for other clients in the scientific and engineering community where later versions of Fortran coding standard are in use. For example, it is now used in work with GE Oil & Gas. The improvements to the WinFPT toolkit and their application/benefits are confirmed in Factual Statement 2.

5. Sources to corroborate the impact**Factual Statements**

- 1) Software Engineer, NCAR Earth Systems Laboratory – for Impacts 1 and 2.
- 2) Technical Director, SimCon Ltd – for Impacts 2 and 3.