

FOCUS Version Control Group Testing

Peter Rainbird (Syngenta) & Wim Beltman (Wageningen Environmental Research)
Piacenza, 2 September 2019



Introduction

- Two parts to this presentation
 1. Overview of what testing aims to achieve and how the FOCUS VCG is aiming to increase the rigour to improve quality
 2. Outline of developer release candidate testing as an example of good practice
- Open the floor for discussion

What is Software Testing and Why is it Needed?

- Software testing is a process to evaluate the functionality of a software application with the intent to determine whether the software meets the specified requirements or not and to identify any defects in order to produce a quality product
- It is not there to show that the requirements are correct or not, just that whatever has been specified has been implemented
- It is dependent on a proper definition of clear requirements which should, ideally, be unambiguous
- The result of any test must, preferably, be a simple pass/fail criteria. Any test that requires a subjective interpretation is not a reliable test.
- Testing might appear to be a quick and easy process, but it requires considerable thought in order to be effective and reduce the risk of encountering problems after a piece of software goes live

FOCUS VCG Proposed Two Stage Strategy

- Current situation is that a release candidate is sent round the VCG for assessment, but this tends to be undertaken using whatever a VCG member happens to be working on. It then goes for a quality check
- This lacks rigour and has a potential to miss issues (*e.g.* R2 climate file)
- Also, there can be significant iteration if the QC turns up issues that need to be addressed
- New proposal is for a two stage process
- Stage 1 is similar to the quality check in that the behaviour of the release candidate is looked at in terms of its operation and robustness. The main question here is whether the model is robust enough for users to work with.
- Stage 2 (undertaken after Stage 1 is passed) is a formalised set of defined runs covering a range of crops and scenarios which will be repeated for all release candidates going forward to assess the impact of changes to the model outputs

Two Stage Strategy - Details

- Stage1 is quite open and potentially iterative so needs a flexible time frame
- Stage 2 is relatively formalised and therefore could operate on a fixed time line, but there would have to be an allowance if repeats are needed (*e.g.* if unexpected differences between versions are identified)
- In order for this to work well, it is imperative that model developers system test thoroughly (and in a documented manner) before submitting to the VCG
- The VCG purpose is not to undertake software testing on behalf of the model developers
- However, the VCG recognises the need to support model developers in this process rather than making life difficult
- Developer testing is focused on making sure that specific updates to the model have been implemented correctly and therefore the test framework is targeted to demonstrating that. Step2 testing is looking at the portfolio impact of the holistic update to the model which is why there has to be consistency between version tests otherwise we don't see any drift in the output or if there are significant changes, they need to be explained/justified.

Testing of Changes by Model Developers

- Before implementing, changes are described, including
 - a test -> e.g. for implementing TWA tables in the summary file
 - or a test plan -> e.g. for implementing all changes in the application tables in SWASH
- Thereafter changes are implemented, tested and reported.

Aim: test that changes have been implemented correctly

Testing of Changes by Model Developers

- Run Test Suite TOXSWA
 - A. Basic test; comparison with analytical solutions for water and for sediment
 - B. For each of the model applications of TOXSWA one or a few typical runs
 - -> for FOCUS e.g.; pond, ditch, stream
 - [in total 16 model runs, report is prepared automatically]
- ***Aim: test that all model applications are still running and give the same results***

This is done frequently
in periods that changes
are implemented



Testing of TOXSWA before submitting to Version Control 1

- Run series of runs with kernel covering:
 - broadness of model applications of TOXSWA (for NL scenario, GEM, etc.)
 - outer values of substance properties combined with outer values of scenario; e.g. substance with DegT50 in water of 0.1 d and stream with highest flow velocity
 - runs that were reported to fail in the past

Aim: test the robustness of the kernel

Database and GUI are not tested

Testing of TOXSWA before submitting to Version Control 2

- Run Test Suite TOXSWA

Aim: test that all model applications are still running and give the same results

- Runs to fill in Format change form
 - Preset runs, using MACRO and PRZM output files downloaded from the VC website

Aim: test that calculation results of TOXSWA have not changed

Backup Slides

Proposed Stage 2 Testing Substances (SW example)

Reference	FOCUS VC 1	FOCUS VC 2	FOCUS VC 3	FOCUS VC 4	FOCUS VC 5	FOCUS VC 6	FOCUS VC 7	FOCUS VC 8 ^a
Based Upon:	FOCUS Test Compound B	FOCUS Test Compound H	FOCUS Test Compound 3	FOCUS Test Compound 4	VFSMod Test Compound 3	VFSMod Test Compound 3 with K _{foc} adjusted to 10,000 L/kg	FOCUS Test Compound 6 and 6met	FOCUS Test Compound F and Fmet
Soil persistence	Impersistent	Very persistent	Impersistent	Moderate persistence	Slightly persistent	Slightly persistent	Moderately persistent parent and metabolite	Moderately persistent parent and metabolite
Water phase persistence				Impersistent	Very persistent	Very persistent		Variable ^a
Sediment phase persistence				Very persistent				
Sorption	Moderate	Moderate	Very mobile	Extremely sorptive	Slightly mobile	Non-mobile	Mobile parent, slightly mobile metabolite	Slightly mobile parent and metabolite
Rationale	Rapid degradation	Persistence in soil	Extremely low sorption - impersistence in soil	Extreme sorption – impersistence in soil	Persistence in wat/sed	Persistence in wat/sed	Moderate persistence and mobility incl. metabolite	Compartmental dynamics testing in TOXSWA ^a

In order to better understand the implications of constraints with reliability of two-compartment kinetics in water sediment studies and how TOXSWA responds to this it is proposed that further testing is carried out as follows:

- 8a; 1 simulation set using separate water DT50 (10 d) and sediment DT50 (30 d) – metabolite parameter unchanged
- 8b; 1 simulation set using total system DT50 in water (22 d) and 1000 d in sediment – metabolite parameters unchanged
- 8c; 1 simulation set using total system DT50 in sediment (22 d) and 1000 d in water – metabolite parameters unchanged

Stage 2 Testing Crops

Reference	FOCUS VC 1	FOCUS VC 2	FOCUS VC 3	FOCUS VC 4	FOCUS VC 5	FOCUS VC 6	FOCUS VC 7	FOCUS VC 8 ^a
Based Upon:	FOCUS Test Compound B	FOCUS Test Compound H	FOCUS Test Compound 3	FOCUS Test Compound 4	VFSMod Test Compound 3	VFSMod Test Compound 3 with K_{foc} adjusted to 10,000 L/kg	FOCUS Test Compound 6 and 6met	FOCUS Test Compound F and Fmet
GAP 1 (potatoes)	Y	Y	Y	Y	Y	Y	Y	Y
GAP 2 (maize)	Y	Y	Y	Y	Y	Y	Y	Y
GAP 3 (winter wheat)	Y	Y					Y	Y
GAP 4 (apples)	Y	Y	Y		Y			
GAP 5 (vines)	Y	Y		Y		Y		
GAP 6 (spring OSR)	Y	Y			Y		Y	
GAP 7 (leafy veg.)	Y	Y	Y	Y		Y		Y