Improving the accuracy of floating-point computations in *FullSWOF_1D* - no. 2

Frédéric Darboux and Carine Lucas
frederic.darboux@inrae.fr carine.lucas@univ-orleans.fr

2023-03-17

Issue

When running the 12 benchmarks of *FullSWOF_1D*, we obtain, with the >2020 Apple chip (ARM):

```
*** Comparison between reference and standard results ***
Emerged bump at rest: Maximum differences are: 1.350000e-14 (absolute)
                                           and 1.895210e-03 (relative).
Fluvial bump: Results are identical.
Transcritical bump with shock: Results are identical.
MacDonald: Torrential Darcy-Weisbach: Results are identical.
MacDonald: Fluvial/torrential Manning: Results are identical.
MacDonald: Torrential/fluvial Darcy-Weisbach: Results are identical.
MacDonald: Smooth transition and shock Manning: Results are identical.
MacDonald: Rain torrential Darcy-Weisbach: Results are identical.
Dry dam break: Results are identical.
Thacker: Maximum differences are: 3.808100e+07 (absolute) and 2.326270e+04 (relative).
Lajeunesse bedload transport phipsi: Water: Results are identical.
Lajeunesse bedload transport phipsi: Transfer: Results are identical.
Lajeunesse bedload transport cM: Water: Results are identical.
Lajeunesse bedload transport cM: Transfer: Results are identical.
```

See bug 13505.

Testing & Changes

When looking at the results in details, we see that, for the bump at rest, the differences are very small, of order of $10^{-13} - 10^{-15}$ m!
→ We performed a first change, in order to save the results up to 10^{-5} m, which is enough precise! This approximation is only performed in the output files, not for the computation.

For the Thacker benchmark, the results are different up to 10^{-3} m/s for the values of q at the final time. The point is that the long doubles do not exist when using the new Apple configuration, they are replaced by doubles.
→ We replaced long doubles by doubles in `hll.cpp`, `muscl.cpp`, `scheme.cpp`.

In the previous floating points application note (2012-08-13), we explained that long doubles were useful:

- to have the same results using Mac OSX, Linux and Windows,
- to have a symmetry in the dam break problem when water has an opposite velocity (from right to left).

After performing the above modifications, we are able to check the symmetry of the results for the dam break problem. The removal of the > 5 decimals widely helps the reproducibility of the results. The results were validated with Mac OSX, Linux and Windows.

Conclusions & Recommendations

We would like to draw the attention of users to these changes: they affect the results saved in the output files, as they are rounded to 5 decimals, but we believe that it will not affect the comparison to measurements, which are less precise.