USING INTERACTIVE JUPYTER NOTEBOOKS TO FACILITATE KNOWLEDGE TRANSFER AND TRAINING OF RAINFALL-RUNOFF MODELLING

A. Peñuela[1]* y F. Pianosi[2]

[1]Dpto de Agronomía, Universidad de Córdoba, Edificio da Vinci. Ctra. Madrid km 396, 14071 Córdoba, España. e-mail: apenuela@uco.es
[2]Department of Civil Engineering, University of Bristol, Bristol, BS8 1TR, United Kingdom. e-mail: francesca.pianosi@bristol.ac.uk

RESUMEN. El campo de la computación interactiva se ha centrado recientemente en Jupyter Notebooks, un entorno de programación literaria que combina en un solo documento código ejecutable, contenido multimedia, resultados, gráficas y texto explicativo. Otra mejora que proporcionan respecto a la documentación estática de modelos es la implementación de visualización interactiva. Esto abre la posibilidad de su uso para la transferencia de conocimientos y la enseñanza. En este estudio utilizamos un Jupyter Notebook interactivo incluido en iRONS (Reservoir Operations Notebooks and Software) toolbox y una encuesta online para evaluar la eficacia de los Notebooks interactivos para facilitar la comunicación de conceptos complejos en la modelización de lluvia-escorrentía. Los resultados muestran que la mayoría de los participantes consideraron los Notebooks interactivos más efectivos que otros métodos. Los participantes aumentaron tanto su familiaridad como sus conocimientos. Estos resultados sugieren que el aprendizaje con Notebooks interactivos facilita la comunicación de conceptos complejos de modelización.

ABSTRACT. The field of interactive computing has recently coalesced around the Jupyter Notebooks, a literate programming environment that combines executable code, rich media, model outputs, figures and explanatory text in a single document. Another enhancement provided by Notebooks over model static documentation, is the possibility to include interactive visualization. This opens the potential use of interactive Jupyter Notebooks for knowledge transfer and teaching purposes. In this study we used the interactive knowledge transfer Notebooks included in the iRONS (Reservoir Operations Notebooks and Software) toolbox and an online survey to evaluate how interactive Jupyter Notebooks contribute to facilitate the communication of complex concepts in rainfall-runoff modelling. The results of the survey show that most participants considered the interactive Notebooks more effective for learning than other methods. Participants increased both their familiarity and knowledge. These results suggest that learning from interactive Jupyter Notebooks facilitate the communication of complex modelling concepts.

1.- Introduction

Advances in the development of interactive computing environments are facilitating the interaction between users and models. By "interactive computing," we mean the sort of exploratory analysis that involves a "human in the loop. The field of interactive computing has recently coalesced around the Jupyter Notebooks (https://jupyter.org/). Jupyter Notebooks combines executable code, rich media, computational output and explanatory text in a single document. With Jupyter Notebooks users execute the code, see right away what happens, modify and repeat in a kind of iterative conversation between user and data (Perkel, 2018). The result is a computational narrative that builds stronger links between model, data and results (Perez and Granger, 2007, Kluyver et al., 2016). Moreover, Jupyter Notebooks can be run on the cloud by using online platforms, such as Binder (https://mybinder.org/), so that they are accessible by a web browser without requiring the installation of Python. But Jupyter Notebooks are not the best communication tool for all audiences, in particular for non-technical users, who may be put off by the presence of code.

However, this can be solved by the use of interactive visualization libraries such as Plotly (https://plotly.com/) and Bqplot (https://bqplot.readthedocs.io/) that allow the implementation of more intuitive elements such as sliders, buttons and menus to interact with the model and results. This provides a mechanism for sharing Jupyter Notebooks as standalone applications and may provide a more effective mechanism to communicate complex modelling concepts to a non-technical audience, such as students and decision-makers.

Previous studies in education have shown mixed conclusions in terms of the efficacy of interactivity for learning (Chou, 2003). For example, Chien and Chang (2012) investigated students learning to use a topographic instrument by watching visualizations with different degrees of interactivity. Compared to animations, full interactions of dragging and controlling the topographic instrument, was the most effective. Similarly, in a study with engineers given a Lego truck to disassembly, Akinlofa et al. (2013) observed that compared to animations and videos, an interactive virtual space, where participants could drag the computer replicas of the Lego pieces was the most effective. However, Pedra et al. (2015) showed that incorporating sophisticated interactivity features into lessons on hand-held devices increased the interest of students, but this was not translated into better learning outcomes.

In this study, we aim to demonstrate of the potential of interactive Jupyter Notebooks for training and knowledge transfer. We will use the knowledge transfer (KT) Notebooks included in the Reservoir Operation Notebooks and Software (iRONS) toolbox (https://github.com/AndresPenuela/iRONS) and an online survey to evaluate how effective interactive Jupyter Notebooks in communicating complex concepts of rainfall-runoff modelling to students.

2.- Methods

2.1. iRONS toolbox

iRONS is a Python toolbox that implements several functions for rainfall-runoff modelling and the simulation and optimisation of reservoir operations, and is based on the principles of modularity, minimalism, openness and accessibility. It is organised into two parts: a suite of functions (the 'Software') implementing several tasks related to simulation and optimisation; and a set of Jupyter Notebooks (the 'Notebooks') that demonstrate key functionalities of the software through practical examples, and that can be run either locally or remotely via a web browser.

The Jupyter Notebooks are divided in two sections:

1) Knowledge Transfer Notebooks: a set of simple examples to demonstrate the value of simulation and optimisation tools for reservoir operations by application to 'proof-of-concept' systems. The Notebooks cover a range of concepts relevant to reservoir operation, such as: manual vs automatic calibration of rainfall-runoff models used to generate reservoir inflows; what-if analysis vs optimisation of reservoir releases; optimisation under conflicting objectives and under uncertainty; optimisation of release releases scheduling vs optimisation of an operating policy; different shapes of operating policies for different reservoir purposes such as domestic or irrigation supply, flood control, or hydropower production.

2) Implementation Notebooks: a set of workflow examples showing how to apply the iRONS functions to real-world data and problems, including: generating inflow forecasts through a rainfall-runoff model, including bias correcting weather forecasts; optimising release scheduling against an inflow scenario or a forecast ensemble; optimising an operating policy against time series of historical or synthetic inflows. These Notebooks are meant to serve as a 'learn-by-doing' alternative to a user manual and a starting point for the user's own application workflows.

In this study we used one of the iRONS Knowledge Transfer Notebooks, "Calibration and evaluation of a rainfall-runoff model", which covers the concepts of interdependency between model parameters, goodness-of-fit between the observed and the simulated hydrograph, manual and automatic model calibration, model evaluation and trade-off between conflicting objectives of the model calibration. Through a simple example the Notebooks asks the user to simulate the natural inflows into a water reservoir, knowing the amount of rainfall that has fallen in the reservoir's catchment area, and to tailor the rainfall-runoff model to the catchment through model calibration, first manually and then using automatic optimisation. Optimisation is performed by linking iRONS to the Platypus Multi-Objective Evolutionary Algorithm (MOEA) (https://github.com/Project-Platypus/Platypus) and the rainfall-runoff applied is the HBV model (Bergström and Singh, 1995). Interactive visualization tools and the implementation intuitive elements, such as sliders to manually change the value of the model parameters (Fig. 1a) or interactive scatter plots to facilitate the exploration of the automatic model calibration results (Fig. 1b), are used to facilitate the communication of the concepts covered in the Notebook.

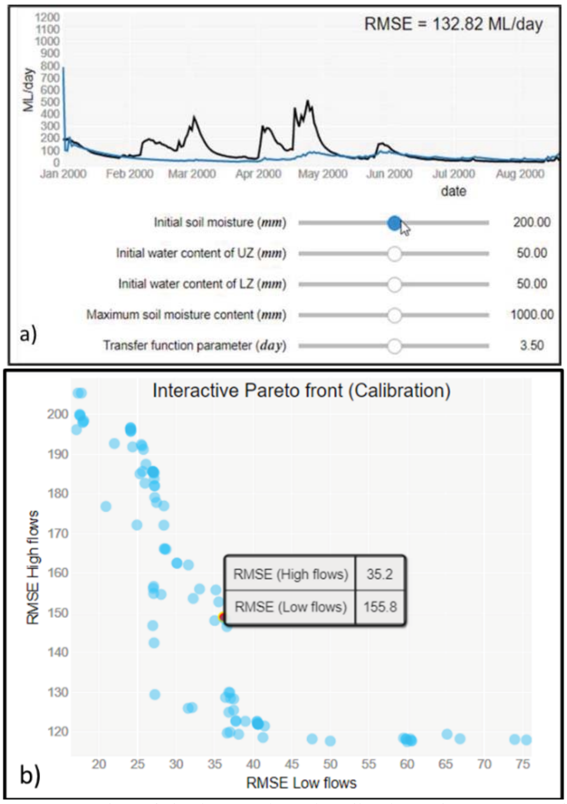

Fig. 1. Examples of the interactive visualisations that were used to convey some of the key concepts: a) manual calibration, goodness-of-fit and parameter interdependency and b) trade-off between conflicting calibration objectives.

2.2. Questionnaire

Through both email invitations and in-person workshops we let participants go through the Notebook at their pace. Before and after going through the Notebook, we asked participants to fill in a questionnaire to measure their familiarity and understanding (Fig. 2) of the concepts covered. Then, by comparing responses before and after, we can assess whether their familiarity and knowledge were increased by using the Notebook. Overall, we obtained responses from 25 participants - consisting of 19 PhD students (11 in Water Resource Management, 3 in Biochemical flows, 3 in Agronomy and 2 in Fine Chemistry) and 6 master's degree students (Ms in Environmental Hydraulics).

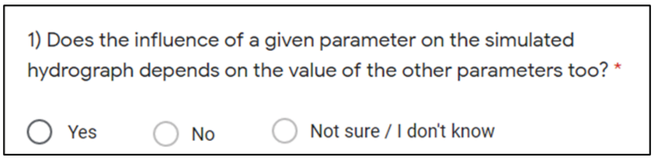

Fig. 2. Example of a question to measure the increase of understanding of participants after using the interactive Jupyter Notebook.

Familiarity to a concept is self-assessed by the participants using a scale from 1 ("not confident") to 5 ("very confident"). Understanding of a concept is instead measured by the ability to give the correct answer to a close-ended question about that concept. The survey also includes direct questions to the participants about their opinion of our interactive Notebooks compared to other learning methods such as lecture slides, textbooks or online videos.

3.- Results

The results of the survey show that the majority of participants (82.4%) considered the interactive Notebooks more effective for learning than other methods, such as lecture slides, videos or books, thanks mainly to the code explanations, interactive visualization and step-by-step structure of Jupyter Notebooks. In general, the participants increased both their familiarity and knowledge about the rainfall-runoff modelling concepts covered by the Notebook (Fig. 3) however, there is not necessarily a clear relationship between the two.

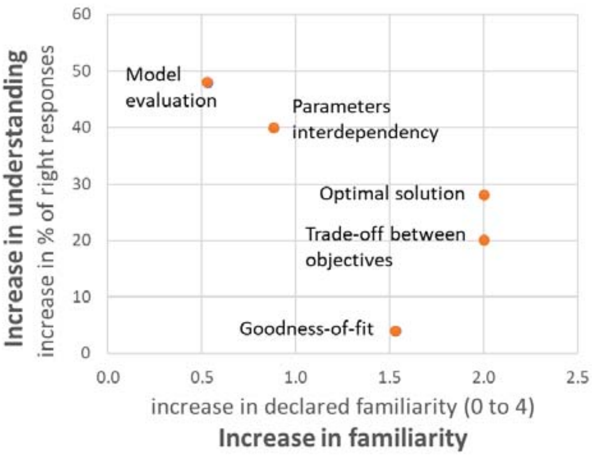

Fig. 3. Results of the questionnaire after using the "Calibration and evaluation of a rainfall-runoff model" Knowledge transfer Notebook - Increase in familiarity to some key concepts versus increase in knowledge of some key concepts (next to each dot the concept covered is indicated).

4.- Conclusions

These results suggest that the characteristics of the Notebooks (literate programming and step-by-step structure) combined with visual interactivity do enhance learning and are appreciated by users. These results also demonstrate the potential of interactive Jupyter Notebooks to facilitate knowledge transfer and training of complex hydrological modelling concepts.
Future work will further evaluate the efficacy of the interactive Jupyter Notebooks particularly as a knowledge transfer tool for hydrology researchers and practitioners.

*Acknowledgements***.** This work was funded by the Engineering and Physical Sciences Research Council (EPSRC), grant EP/R007330/1. The authors are also grateful to the workshop participants who contributed to the evaluation of interactive Jupyter Notebooks.

5.- References

Akinlofa, O.R., Holt, P.O.B. & Elyan, E. 2013. Domain expertise and the effectiveness of dynamic simulator interfaces in the acquisition of procedural motor skills. British Journal of Educational Technology, 44, 810-820.
Bergström, S. & Singh, V. 1995. The HBV model. Computer models of watershed hydrology., 443-476.
Chien, Y.-T. & Chang, C.-Y. 2012. Comparison of Different Instructional Multimedia Designs for Improving Student Science-Process Skill Learning. Journal of Science Education and Technology, 21, 106-113.
Chou, C. 2003. Interactivity and interactive functions in web-based learning systems: a technical framework for designers. British Journal of Educational Technology, 34, 265-279.
Kluyver, T., Ragan-Kelley, B., Pérez, F., Granger, B.E., Bussonnier, M., Frederic, J., Kelley, K., Hamrick, J.B., Grout, J. & Corlay, S. Jupyter Notebooks-a publishing format for reproducible computational workflows. ELPUB, 2016. 87-90.
Pedra, A., Mayer, R.E. & Albertin, A.L. 2015. Role of Interactivity in Learning from Engineering Animations. Applied Cognitive Psychology, 29, 614-620.
Perez, F. & Granger, B.E. 2007. IPython: A System for Interactive Scientific Computing. Computing in Science & Engineering, 9, 21-29.
Perkel, J.M. 2018. Why Jupyter is data scientists' computational notebook of choice. Nature, 563, 145-147.