

## Supplementary Material

# Power to the programmer: Modeller's perspective on automating the setup of hydrodynamic models for Dutch water authorities

## A: Interview guide

The interview guide consisted of three parts: an introductory part covering the background of the interviewee, a specific part covering the full modelling process and a final part concerning general question about modelling. The questions that were used for this research are highlighted in bold text. The questions are subdivided in main questions and sub-questions to probe the interviewee if necessary.

### Questions regarding the background:

1. How would you describe your current position?
2. What is your background?
  - (a) Have you worked at other companies/institutes before?
  - (b) Where and what did you study?
3. Can you describe your experience as a modeller?
  - (a) Which type of models?
  - (b) Which models models exactly?
  - (c) How many years?
4. For which type of questions do you use a model?

For interviewees from water authorities, we also asked how much and which parts of the modelling were done in-house or externally.

### Questions about the full modelling process:

#### Model Software

1. Which software do you generally use?
2. How did you make this decision?
  - (a) Why this software?
  - (b) Are there other options?
  - (c) Why for this purpose?
3. Did you have certain settings in your model?
  - (a) If no:
    - i. Why not?
    - ii. When would you have certain settings?
    - iii. **If this step were to be automated would you use it / execute this step?**
      - A. Would you find this useful?**
      - B. Why (not)?**
  - (b) If yes:
    - i. Why these settings?
    - ii. In which situation(s) might you chose different settings?

- iii. **Do you think this step can be automated in the future?**
  - A. Why (not)?**
  - B. To what extent?**
  - C. What are possibilities for automation?**

4. In which situation(s) might you chose a different software?

Forcing Data

1. Which forcing data do you generally use?

- (a) Type?
- (b) Source?
- (c) Resolution of data, spatial and temporal?
- (d) Availability?

2. How did you make this decision?

3. In which situation(s) might you chose different data?

4. **Do you think this step can be automated in the future?**

- (a) Why (not)?**
- (b) To what extent?**
- (c) What are possibilities for automation?**

5. Do you perform any pre-processing on the forcing data?

- (a) If no:
  - i. Why not?
  - ii. When would you pre-process the data?
  - iii. If this step were to be automated would you use it / execute this step?
    - A. Would you find this useful?
    - B. Why (not)?

(b) If yes:

- i. How do you generally execute any pre-processing?
- ii. How did you make this decision?
- iii. In which situation(s) might you chose a different method?
- iv. **Do you think this step can be automated in the future?**
  - A. Why (not)?**
  - B. To what extent?**
  - C. What are possibilities for automation?**

Static Data

1. Which static data do you generally use?

- (a) Type?
- (b) Source?
- (c) Resolution of data, spatial and temporal?
- (d) Availability?

2. How did you make this decision?

3. In which situation(s) might you chose different data?

4. **Do you think this step can be automated in the future?**

- (a) Why (not)?**
- (b) To what extent?**
- (c) What are possibilities for automation?**

5. Do you perform any pre-processing on the static data?
- (a) If no:
- i. Why not?
  - ii. When would you pre-process the data?
  - iii. **If this step were to be automated would you use it / execute this step?**
    - A. Would you find this useful?**
    - B. Why (not)?**
- (b) If yes:
- i. How do you generally execute any pre-processing?
  - ii. How did you make this decision?
  - iii. In which situation(s) might you chose a different method?
  - iv. **Do you think this step can be automated in the future?**
    - A. Why (not)?**
    - B. To what extent?**
    - C. What are possibilities for automation?**

#### Observation Data

1. Which observation data do you generally use?
  - (a) Type?
  - (b) Source?
  - (c) Resolution of data, spatial and temporal?
  - (d) Availability?
2. How did you make this decision?
3. In which situation(s) might you chose different data?
4. **Do you think this step can be automated in the future?**
  - (a) Why (not)?**
  - (b) To what extent?**
  - (c) What are possibilities for automation?**
5. Do you perform any pre-processing on the observation data?
  - (a) If no:
    - i. Why not?
    - ii. When would you pre-process the data?
    - iii. **If this step were to be automated would you use it / execute this step?**
      - A. Would you find this useful?**
      - B. Why (not)?**
  - (b) If yes:
    - i. How do you generally execute any pre-processing?
    - ii. How did you make this decision?
    - iii. In which situation(s) might you chose a different method?
    - iv. **Do you think this step can be automated in the future?**
      - A. Why (not)?**
      - B. To what extent?**
      - C. What are possibilities for automation?**

#### Simulation Period

1. What simulation period do you generally use?
2. How did you make this decision?

3. In which situation(s) might you chose a longer/shorter period?

4. **Do you think this step can be automated in the future?**

(a) **Why (not)?**

(b) **To what extent?**

(c) **What are possibilities for automation?**

#### ***Temporal Resolution***

1. What temporal resolution do you generally use?

2. How did you make this decision?

3. In which situation(s) might you chose a different resolution?

4. **Do you think this step can be automated in the future?**

(a) **Why (not)?**

(b) **To what extent?**

(c) **What are possibilities for automation?**

#### ***Spatial Resolution***

1. What spatial resolution do you generally use?

2. How did you make this decision?

3. In which situation(s) might you chose a different resolution?

4. **Do you think this step can be automated in the future?**

(a) **Why (not)?**

(b) **To what extent?**

(c) **What are possibilities for automation?**

#### ***Sensitivity Analysis***

1. Is this a step you normally execute?

(a) If no:

i. Why not?

ii. When would you pre-process the data?

iii. **If this step were to be automated would you use it / execute this step?**

**A. Would you find this useful?**

**B. Why (not)?**

(b) If yes:

i. How do you generally execute any sensitivity analysis?

ii. How did you make this decision?

iii. In which situation(s) might you chose a different method?

iv. **Do you think this step can be automated in the future?**

**A. Why (not)?**

**B. To what extent?**

**C. What are possibilities for automation?**

#### ***Calibration***

1. Do you normally calibrate the model?

(a) If no:

- i. Why not?
- ii. When would you pre-process the data?
- iii. **If this step were to be automated would you use it / execute this step?**
  - A. Would you find this useful?**
  - B. Why (not)?**

(b) If yes:

- i. How do you generally execute any calibration?
- ii. How many parameters do you calibrate on?
- iii. How did you make this decision?
- iv. In which situation(s) might you chose a different method?
- v. **Do you think this step can be automated in the future?**
  - A. Why (not)?**
  - B. To what extent?**
  - C. What are possibilities for automation?**

### Uncertainty Analysis

1. Is this a step you normally execute?

(a) If no:

- i. Why not?
- ii. When would you pre-process the data?
- iii. **If this step were to be automated would you use it / execute this step?**
  - A. Would you find this useful?**
  - B. Why (not)?**

(b) If yes:

- i. How do you generally execute any uncertainty analysis?
- ii. How did you make this decision?
- iii. In which situation(s) might you chose a different method?
- iv. **Do you think this step can be automated in the future?**
  - A. Why (not)?**
  - B. To what extent?**
  - C. What are possibilities for automation?**

### Validation

1. Is this a step you normally execute?

(a) If no:

- i. Why not?
- ii. When would you pre-process the data?
- iii. **If this step were to be automated would you use it / execute this step?**
  - A. Would you find this useful?**
  - B. Why (not)?**

(b) If yes:

- i. How do you generally validate the model?
- ii. How did you make this decision?
- iii. In which situation(s) might you chose a different method?
- iv. **Do you think this step can be automated in the future?**
  - A. Why (not)?**
  - B. To what extent?**
  - C. What are possibilities for automation?**

### Results/Conclusion

1. How do you reach your final conclusions?

- (a) Purely the model?
- (b) Multiple model runs?

(c) Expert judgement in combination with the model?

**Questions regarding modelling in general:**

1. Do you have confidence in a model and its simulations?
2. How do you estimate your influence as a modeller on the final outcomes?
  - (a) If another modeller would have executed the same study, would the results be different?
  - (b) And what about the conclusions?
3. **How do you estimate the influence of a modeller on the model in comparison to the programmer?**
4. Does your organisation use a certain modelling workflow?
  - (a) If yes, what does it look like?
  - (b) What is it based on?
5. Are you familiar with the Dutch handbook ‘Good Modelling Practices’? This was published by STOWA.
  - (a) If no:
    - i. What would a handbook be useful for?
    - ii. In what format would you use a handbook?
  - (b) If yes:
    - i. Have you used it in the past? Why (not)?
    - ii. How could the handbook become more relevant for practical use?
6. Do you have any further remarks or additions?

If it became apparent that automation scripts were already being used, these additional follow-up questions would be asked:

1. **What has been automated?**
2. **What is it based on?**
3. **How long do you already use it?**
4. **How is the automation script maintained?**

**B: ATLAS.ti Codes**

**B.1 Deductive interview codes**

In Table 1, the interview codes based on Pagano et al. (2016) can be found. These were used for the deductive content analysis in this study.

Table 1: List of the deductive interview codes used in our content analysis. The deductive interview codes are based on Pagano et al. (2016). In the third column the aspect(s) to which the interview codes are linked are given. I1, 2 and 3 stand for respectively the different issues described by Pagano et al. (2016): the role of the modeller, the change in modellers’ behaviour and the perception of trustworthiness. BP1, 2, 3, 4 and 5 correspond to respectively these best practices from Pagano et al. (2016): have transparent systems, no peeking at the answer, evaluate your results, never stop learning the science and redefine your role.

Code Name	Description	Aspects of Pagano et al. (2016)
Need to execute different tasks	Due to the automation, the modeller needs to execute different tasks than they previously did.	I1

<b>New tasks outside own expertise</b>	The modeller's current tasks after automation are outside their own expertise	I1
<b>Different skills obtained by modeller (in future)</b>	Due to automation, the modeller obtains different or more/less skills than they previously did without automation	I1
<b>Different knowledge obtained by modeller (in future)</b>	The modeller obtains different, more or less hydrological knowledge than they previously did without automation	I1
<b>Capacity to take over automation</b>	A modeller would be less capable to take over if the automation fails	I1
<b>Capacity to interpret results</b>	Any change in if the modeller's ability to interpret results due to automation	BP4, BP5
<b>Test of own understanding (give hypothesis)</b>	Before the automation is started, the modeller should give their own hypothesis to test this.	BP2
<b>Bias due to output</b>	The modeller's judgement is influenced by the output they see from the automation, especially if they didn't have an initial idea of the potential outcome	BP2
<b>Redefinition of the modeller's role</b>	After the automation is incorporated, how is the modeller's role redefined?	BP5
<b>Transfer of who makes the modelling choices</b>	The choices within the modelling process are taken by someone else (often the programmer)	I2
<b>Different set of choices that have to be made by modeller</b>	The modellers is faced with other (modelling) decisions than previously	I1
<b>Change in communication of results</b>	Change in what and how the results are communicated	I2, BP5
<b>(False) 1st impression</b>	The first impression a modeller has of the automation is incorrect	I3
<b>Transference of 1st impression</b>	The first impression of a modeller is copied by other modeller(s)	I3
<b>Change of 1st impression</b>	A modeller changes their view from their initial impression	I3
<b>Transparency of automation process</b>	Clarity of what happens within the automation, i.e. what choices were made? How does the automation process work?	BP1
<b>Obtain intermediate results</b>	While the automation runs, results should be given after different steps to give more insight into the automation.	BP1
<b>Evaluation</b>	Evaluation of the automation, does this happen? How?	BP3, BP4

## B.2 Inductive interview codes

In Table 2, the interview codes created during the inductive content analysis of this study are shown. The interview codes were subdivided in seven different groups: Extent of automation, Implementation of automation, Interaction between water authorities and consulting company, Levels in automation development, Motivations for (not) developing automation, Role of modeller and programmer, and Usage of automation. Some interview codes were not subdivided in the groups. These are in the miscellaneous group.

Table 2: List of the inductive interview codes used in our content analysis.

Group	Code Name	Description
Extent of automation	Current extent of automation	Indication that the modelling process is automated (to some extent).
	Can be automated	The modelling process can be automated according to the modeller.
	Can't be automated	The modelling process cannot be automated according to the modeller.
	Too extensive	Indication to which extent the modelling process has been automated according to the modeller.
	Completely	Indication to which extent the modelling process has been or can be automated according to the modeller.
	Almost completely	Indication to which extent the modelling process has been or can be automated according to the modeller.
	Mostly	Indication to which extent the modelling process has been or can be automated according to the modeller.
	In development	Indication to which extent the modelling process has been or can be automated according to the modeller.
	Partly	Indication to which extent the modelling process has been or can be automated according to the modeller.
	Not that far	Indication to which extent the modelling process has been or can be automated according to the modeller.
	Not at all	Indication to which extent the modelling process has been or can be automated according to the modeller.
	Executed manually	A certain aspect of the modelling process is carried out manually, without automation.
	Execution of modelling decisions in automation	The automation only covers the steps to executed certain choice made by the modeller.
	Less complex components	According to the modeller only less complex components are automated.
No transfer of modelling decision	In automating the modelling process no transfer of modelling decision has occurred for the modeller.	



Group	Code Name	Description
Implementation of automation	Context dependent	How the automation script development is done depends on context (resources, funding, capabilities)
	Dependent on model software	How an automation script is developed depends on with which model software it should be compatible.
	External organisation makes automation script	The automation script development is done by an external organisation.
	First gain insight, then automate	Before an automation script is developed, insight into the process to be automated should exist.
	Iteratively	An automation script is developed over time, improving it whenever it is used.
	Quick and dirty	The automation is developed quickly and for its then current purposes reliable.
	What is the purpose?	An automation script is developed according to a certain purpose.
Interaction between water authorities and consulting companies	Executed by Consulting company (originally)	Regardless of automation or not, this step is generally executed by a consulting company.
	Executed by Research Institute	Regardless of automation or not, this step is generally executed by a research institute.
	Executed by Water Authority (originally)	Regardless of automation or not, this step is generally executed by a water authority.
	Ownership of automation	Who is the owner and therefore responsible for the automation.
Levels in automation development	Organised on personal level	The automation is developed and used at the personal level.
	Organised on team level	The automation is developed and used at the team level.
	Organised on organizational level	The automation is developed and used at the organisational level.
	Organised on inter-organisational level	The automation is developed and used at the inter-organisational level.
	Capacity to automate	What resources (e.g.) funding and computer capacity) are available to automate within an organisation.
	Create documentation	Along with any automation script, documentation should be written to communicate the purpose of automation script and how it works.
	Differences between organisations	There are differences in modelling practices and perspectives between different organisations.
	Differences within department	There are differences in modelling practices and perspectives within a department.
	Differences within organisation	There are differences in modelling practices and perspectives within an organisation.
	One-time use	An automation script is developed (initially) to be used only one time or for one project
	Open-Source	The automation is available open-source.

Group	Code Name	Description
Levels in automation development (continuation)	Overlap in tools	Within different departments or organisations similar automation scripts or tools exist.
	Uniformise automation scripts	When scripts from an Individual or Team level are rewritten in such a way that they can be used at a(n) (Inter-)organisational level.
Motivations for (not) developing automation	Automated with previous method	It used to be automated with a previous method, so it is expected that is, will be automated in the new method as well.
	Do not execute this ourselves	Modellers do not carry out a certain aspect of the modelling process themselves and therefore, they care less about if and how this aspect is automated.
	Consistency	Automating a certain aspect of the modelling process will create more consistency in the execution of that aspect.
	Easier to do manually	It is considered easier to execute a modelling step manually than to automate the modelling step.
	If needed	Modelling process is automated once it seems necessary.
	Maintain control	The modeller wants to maintain control over the modelling process.
	Not time efficient	Automating a certain part of the modelling process would not be time efficient.
	Not useful	Automating a certain part of the modelling process is deemed not useful.
	Gain insight	Developing and using automation can give the modeller insight into the modelling process, uncertainties and results.
	Ease of use	Automation is easy to use according to the modeller.
	More accurate	Automation will give a more accurate result.
	Objectivity	Automating the modelling process introduces more objectivity into the whole process.
	Reduce human errors	Automation limits the potential for human errors.
	Reproducibility	Automation increases the reproducibility of the modelling results.
	Time efficient	Automation will save time for the modellers.
	Transparency	Automation makes the modelling process more transparent.
	Useful to have	The automation of (a part of) the modelling process is deemed useful.
	Budget	The financial resources to develop an automation.
	Does it matter?	The method or the outcome of the method are not that important.
Data are limiting factor	Data quality and quantity limit the potential to develop and use automation.	
Level of difficulty	How difficult the automation is will influence to which extent if at all a certain part of the modelling process is automated.	

Group	Code Name	Description
Motivations for (not) developing automation (continuation)	Should we want this?	Consideration if the automation of certain parts of the modelling process would be desirably.
	Too complex	Automating certain parts of the modelling process would have to many options or be too specific or require the modeller's expertise to execute that part.
Role of modeller and programmer	Programmer most influential	Modeller's perception that programmer is more influential.
	Modelling steps afterwards	The modelling steps taken after the programmer has developed an automation script are more influential than the decisions made in the automation script.
	Programmer follows hydrological laws	The modeller's trust that the programmer follows hydrological laws.
	Modeller's responsibility	What is considered to be the modeller's responsibility in using and developing an automation script.
	Programmer's responsibility	What is considered to be the programmer's responsibility in using and developing an automation script.
Usage of automation	Was not used	A previous automation script was not used.
	Would not use it	A modeller would not use a certain automation script when automated.
	Would execute this modelling step when automated	If a certain modelling step were to be automated, a modeller would use it in the future.
	Not sure if they'd use it	If a certain aspect is automated, a modeller is not sure they would use it.
	Couple of times	An automation script would be used a couple of times immediately after development.
	Irregular	The use of the automation script would be / is irregular.
	Repeated use	The automation script would be / is used frequently.
	Use default value	Within an automation script the default value was used even if there was a choice.
	Check by modeller	The modeller would check the output after the automation has given it.
	Modeller gives some input	When using an automation script, a modeler will give some inputs to the automation script before it runs.
	Make modelling decision yourself	The modeller wants to make the modelling decision themselves.
	Expert knowledge necessary	The experience and expertise of a modeller is necessary in a particular modelling step.
	Remain critical	When using an automation script, a modeler should keep checking the automation itself.

Group	Code Name	Description
Usage of automation (continuation)	Understand the automation	When using an automation script, the modeler should understand the automation.
	Use automation to advise you	Use the results of an automated aspect to inform you as a modeller to make a next decision and maybe adapt certain aspects of the step just executed with the automation script.
Miscellaneous	Data doesn't receive enough attention	A modeller perceives that the data are payed enough attention.
	No choice	A modeller does not have a choice in a particular modelling step, which makes automation illogical or difficult.
	Not automation, but rather standardisation	A modeller perceives that standardisation is essential, not automation necessarily.
	Not known by interviewee	The interviewee does not know (exactly).
	Surprised not organised on inter-organisational level	A modeller is surprised that a certain part of the modelling process is not developed or used at an inter-organisational level.
	Trust in scientific literature high	A modellers does not trust certain aspects of science with regards to the modelling process or automation.
	Trust in scientific literature low	A modellers trusts certain aspects of the modelling process or automation because they are underpinned by science.