

numWater.Social

A numeric model for water supply systems including considerations about social interactions and drinking behavior

Robert Egel

01.14.2020

1 Developing the Social System

As a basis for a more realistic approach to modeling a water system for a rural school in Nyamasheke, Rwanda, a region with very irregular rain precipitation, the existing complex dynamic model was complemented with social behavior. Additionally, it was simplified intensively: the processes *microscopic water inflow/outflow* and *macroscopic water inflow/outflow* were combined to a single *water inflow/outflow (I/O)* process. Thereby the interaction between the processes is eased since now everything is calculated on the same time scale. Furthermore the time scales of the simulation is altered, because the calculations of the social system are very intensive for high numbers of users and small time steps are necessary. Since bottles fill up very quickly, the length of time steps was varied systematically between 1 second and 10 seconds. 5 seconds proved to be a good compromise between accuracy and computation effort. The whole simulation time was set to cover one school day. The schedule was altered slightly: Many students will stay longer in school to do homework collectively and help each other than necessary. This is why users are assumed to arrive at 7:00 am, classes start an hour later. They also stay in school until 4:00 pm although classes end at 4:15 pm. Besides, the model was changed from a differential version utilizing Euler integration to a difference equation model using iterations because the previous model gained non-interpretable results for the social aspects.

1.1 Behavioral aspects and social preferences of users

The social model works by generating a data frame with character preferences for each person of the 2000 students and teachers. Each person is assigned to a bottle of random size and random initial fill. Throughout the day they drink a random amount from that bottle. The fill level is monitored and refreshed each time step. If the fill level falls below a certain margin, the person gets queued in a waiting line, which goes into effect when the next school break starts. At that point in time, the water taps are opened and the water runs into the bottles of the first few people in line. The amount of people who get their bottles filled at the same time depends on the number of water taps. As soon as the filling level of the first people in line reaches their respective bottle volumes, they are removed from the waiting line and other people can start filling their bottles. In case the fill level of a persons bottle reaches zero, they become unsatisfied with the water system. Thereby a satisfaction rate can be calculated and used as a system metric easily and effectively.

From a collection of usual bottle sizes, each person is randomly assigned an individual bottle. Thereby the chances of getting a 0.5 or 1.0 liter bottle is quite high with probabilities of 0.3 and 0.35, where the probability of getting a 0.75 liter bottle is lower because they are not as popular. The initial fill of the bottle follows a random normal distribution with mean being half full, the chance of having a negative initial fill level is zero because that error is caught using a minimum function mapped over the array. How much water is actually used by an individual in each time step determined by a random normal distribution as well. It is assumed that people drink around two liters of water each day. This value is assumed to be constant throughout the day. The daily water use is converted to water use per second for easier computations in the model. This randomness provides individual character preferences for the whole group of people using the schools water system. To make comparisons easier and results reproducible, the generators for random events are influenced by a static seed using the `set.seed()`-function. Therefore each model run with the same parameters yields the same results. Initial satisfaction is set to TRUE for all people, since everyone's bottles are about half full.

1.2 Social norms and group behavior

A problem that arises with this system is the order in which people get lined up. To solve this and show effects of different societal norms on the system, at the start of each break, the waiting line is ordered in a way according to a predefined social preference. The default manner keeps the initial order: who's fill level reaches the threshold first, gets in line as the first person.

Another option is shuffling the line at the start of each break by using re-sampling without replacement, simulating the behavior of many people from a variety of locations approaching the water taps. Instead of offering people with very low fill heights the first spaces in line, no one will switch places. Therefore this social norm is called *egoistic* behavior.

As defined by Shaver,¹ egoism “claims that each person has but one ultimate aim: her own welfare”. It is assumed to generate lower general satisfaction levels than the default behavior.

The last implemented option tried to simulate an *altruistic* society as an opposite to egoism: the group let’s people with empty or almost empty bottles get to the taps first. Altruism is defined as “behavior undertaken deliberately to help someone other than the agent for that other individual’s sake.”² Thereby the weakest get lots of helpful support, especially considering that in short breaks not everyone will be able to get any water. It is assumed to boost satisfaction rate in comparison to default behavior.

Nonetheless it seems highly unlikely to me that such a group could be able to show such behavior. I guess the most realistic scenario is egoistic behavior unless people are somehow enforced to help each other and cooperate on this problem.

1.3 Definition of High Performance Criteria

The High Performance Criteria (HPC) introduced before were reused but a key indicator was changed. The objective *provide enough water for each individual within school time* is measured by the aforementioned *satisfaction rate*: everyone is happy unless their bottle goes empty. Anyhow, some people might drink significantly more than others but have smaller bottles, making them unsatisfied quite quickly. At the end of the day, the satisfaction rate will rarely reach 1.0 in optimal conditions only, which is why the goal was set to 0.9. In real life, people might also share their bottles in case a friend gets empty, but this behavior is not covered by the model.

Using deterministic calculations, each high performance criterion except *provide healthy water* could be satisfied. This is due to the fact that according to Blaustein et al.³ E.coli-bacterial growth is not temperature-dependent enough to be controlled entirely by the temperature in the water tank: no matter the temperature, according to the model, the bacteria will keep growing exponentially. Thereby the limit of bacterial concentration will be exceeded after a few days. Options to improve this situation could be filtering or exposure to UV light, which would eventually inactivate all E.coli bacteria. The temperature criterion has been proved to be fulfilled by many system configurations in deterministic calculations, is not affected deeply by the social model and will therefore not be considered in the latter analyses. In Table 2, all High Performance Criteria can be observed.

Table 2: High Performance Criteria

OBJECTIVE	KEY INDICATOR	UNIT (SI)	EXTENT
provide enough water for each individual within school time	satisfaction rate	1	≥ 0.9
provide healthy water	bacteria concentration c	num/m^3	$< 25 \cdot 10^{-4}$
provide a pleasant drinking temperature	temperature	°C	≤ 17

1.4 Alternative approaches to computing and visualization

Parallel computing was utilized to speed up simulations. Used packages are `doSNOW`⁴ (based on `snow`⁵) and `foreach`⁶. This method allows to use all processor cores in parallel, which speeds up running the parameter variation scripts significantly. Furthermore the package `plotly`⁷ used in the tutorial was replaced by `ggplot2`⁸ and `cowplot`⁹ because it is not able to handle big amounts of data effectively.

2 Analysis of the models behavior

Figure 1 depicts the models behavior with default parameter settings and default social norm. Looking at the water height (h) inside the tank and number of bottles ($n_{Bottles}$) drawn from the system, the relation to both the stochastic and the deterministic models gets clear. Just as before, h declines over time, as $n_{Bottles}$ increases, both happens stepwise because of the underlying schedule, which represents the unique character of a school’s water supply system.

The new outputs are *rate of satisfaction* and *people standing in waiting line*. These social metrics help to understand the new aspects of the model. The High Performance Criterion for providing enough water for everyone is now measured by a satisfaction rate. As stated before, this value is initially set to 1.0 (100%). Considering this metric ought not fall below the threshold of 90%, the system with default parameters does not perform not well enough to be classified as high

¹“Egoism.”

²Kraut, “Altruism.”

³“Escherichia Coli Survival in Waters.”

⁴Ooi, Corporation, and Weston, *doSNOW*.

⁵Tierney et al., *Snow*.

⁶Ooi, Microsoft, and Weston, *foreach*.

⁷Sievert et al., *Plotly*.

⁸Wickham et al., *Ggplot2*.

⁹Wilke, *Cowplot*.

performing, but that should be adjustable easily by altering a few parameters. Interesting is that this output also shows stepwise evolution with increasing step size. The existence of the steps can be explained by the behavior of the waiting line at starts of breaks: the decline of the satisfaction rate slows down when people fill their bottles. Especially because in the default social norm, the persons who's bottle fill level falls below the threshold first, get in line first and the order is not changed.

Another interesting aspect is the strong correlation between number of opened water taps (n_{taps}) and outflow (Q_{out}). These outputs are almost proportional to each other because Q_{out} the only dynamic parameters it depends on are h and n_{taps} . The jitter around time 7h and 17h in n_{taps} and Q_{out} can be explained by the length of the waiting line at those times: the actual n_{taps} used in each time step is set to the minimum of the preset n_{taps} (default value is 10) and the length of the waiting line. Because the "breaks" before and after classes are quite long, the waiting line gets very short or

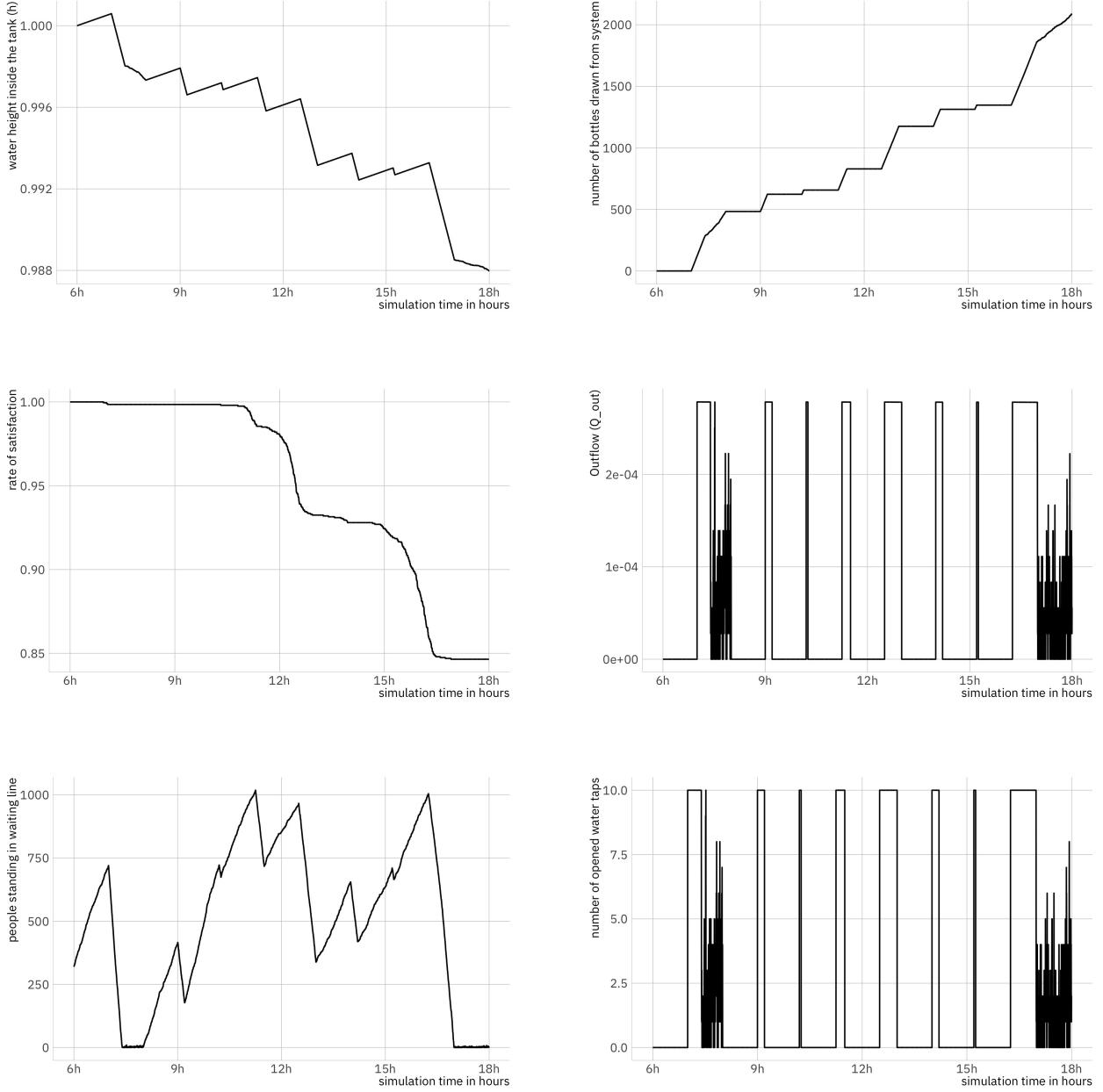


Figure 1: Model behavior with default settings

2.1 Parameter analysis and suggestions for optimization

Since in the two previous projects, z , n_{taps} and A_{roof} were identified as very influential parameters,¹⁰ these are chosen for a further parameter analysis. Additionally the social preferences are taken into account by adding it to the analyzed parameters and inspecting socially relevant parameters in detail for each norm.

¹⁰Egel, “NumWater.Stochastic.”

Table 3: Overview of parameters taken into account in the parameter analysis

parameter	social norms	extent of variations
A_{roof}	default	100 - 2000 m ²
n_{taps}	default	5 - 25
z	default	0 - 10 m
social norm	-	default, altruistic, egoistic
n_{taps} (detailed analysis)	default	10 - 15
n_{taps} (detailed analysis)	altruistic	9 - 12
n_{taps} (detailed analysis)	egoistic	10 - 15
z (detailed analysis)	default	3.0 - 4.0 m
z (detailed analysis)	altruistic	2.5 - 4.0 m
z (detailed analysis)	egoistic	4.5 - 6.0 m

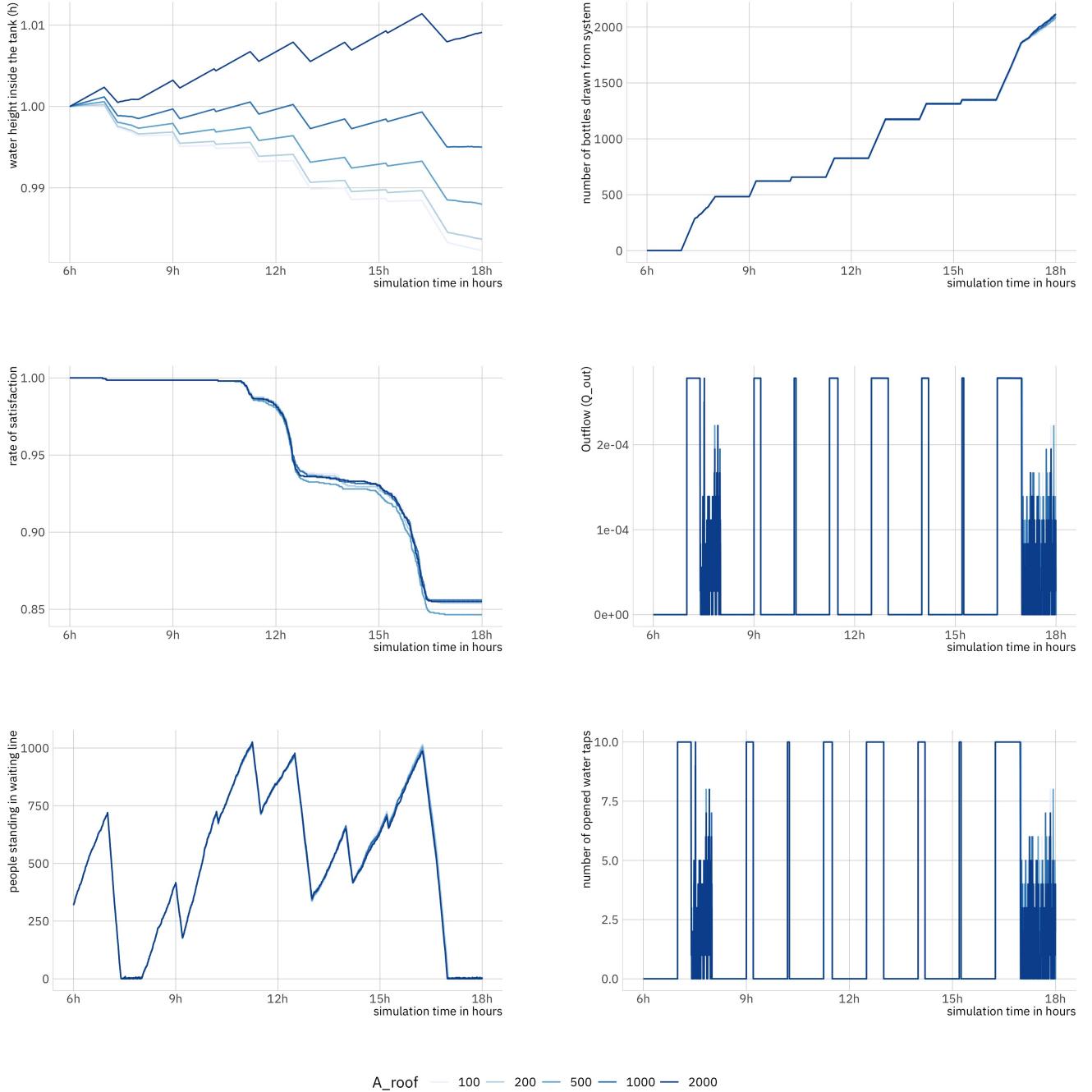


Figure 2: Variation of $A_{rooftop}$

Just as in the stochastic model, a huge influence of the roof area on the water height is shown in Figure 2. Contrarily to that, it has almost no effect on the number of bottles drawn from the system. This due to the social system: $A_{rooftop}$ has

practically no impact on the peoples behavior, as can be seen in the satisfaction rate and length of the waiting line. People don't consume more water just because there is more available in the tank they drink from. This is plausible because the number of bottles filled depends on whether people are actually getting water, this metric does not change depending on the water height, as it did before.

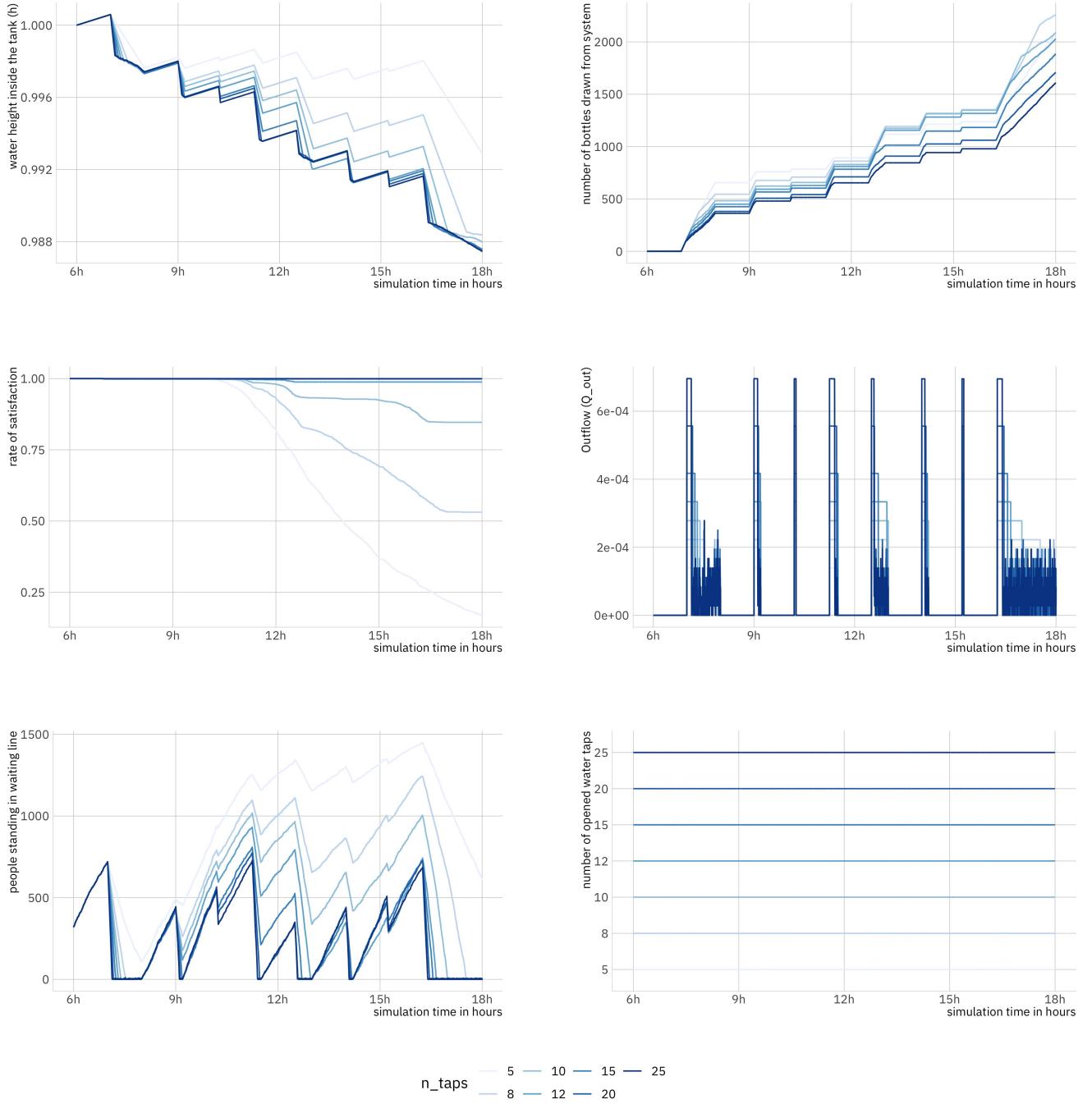


Figure 3: Variation of n_{taps}

In contrast to A_{roof} , n_{taps} shows a huge impact on the social aspects of the system, which is obvious, since it limits the number of water taps the people can use interact with the system. The dependency is quite clear: there are more taps in the system \rightarrow more people can fill their bottles at the same time \rightarrow shorter waiting times and less people waiting in line \rightarrow higher satisfaction rate. All of this can be seen in Figure 3.

In comparison to the previous models, the number of water taps does not have a significant influence on water height and number of bottles per day . With one exception that will be analyzed in detail later, both do not show a trend with growing n_{taps} . Again, that's because of the social aspects of the model: the water stops running if no one is standing in line. The aforementioned exception is the divergence of the water height for $n_{taps} = 5$. All other variations converge to the same water height in the end of the day. That one line diverges because at 18h there is still a waiting line if there are just five taps. Therefore the demand for water is not covered and the water needed to cover it is obviously still in the tank, hence the difference in water height.

Because of the strong dependence of Q_{out} on n_{taps} , the scale of the outflow changes with the number of taps, as shown in the plot. The smaller the number of taps, the smaller the scale of the outflow and the longer the duration of Q_{out} being at

it's respective maximum, as can be seen around times 7h, 13h and 17h. Accordingly, the length of the waiting line reaches zero sooner and more often for higher number of taps.

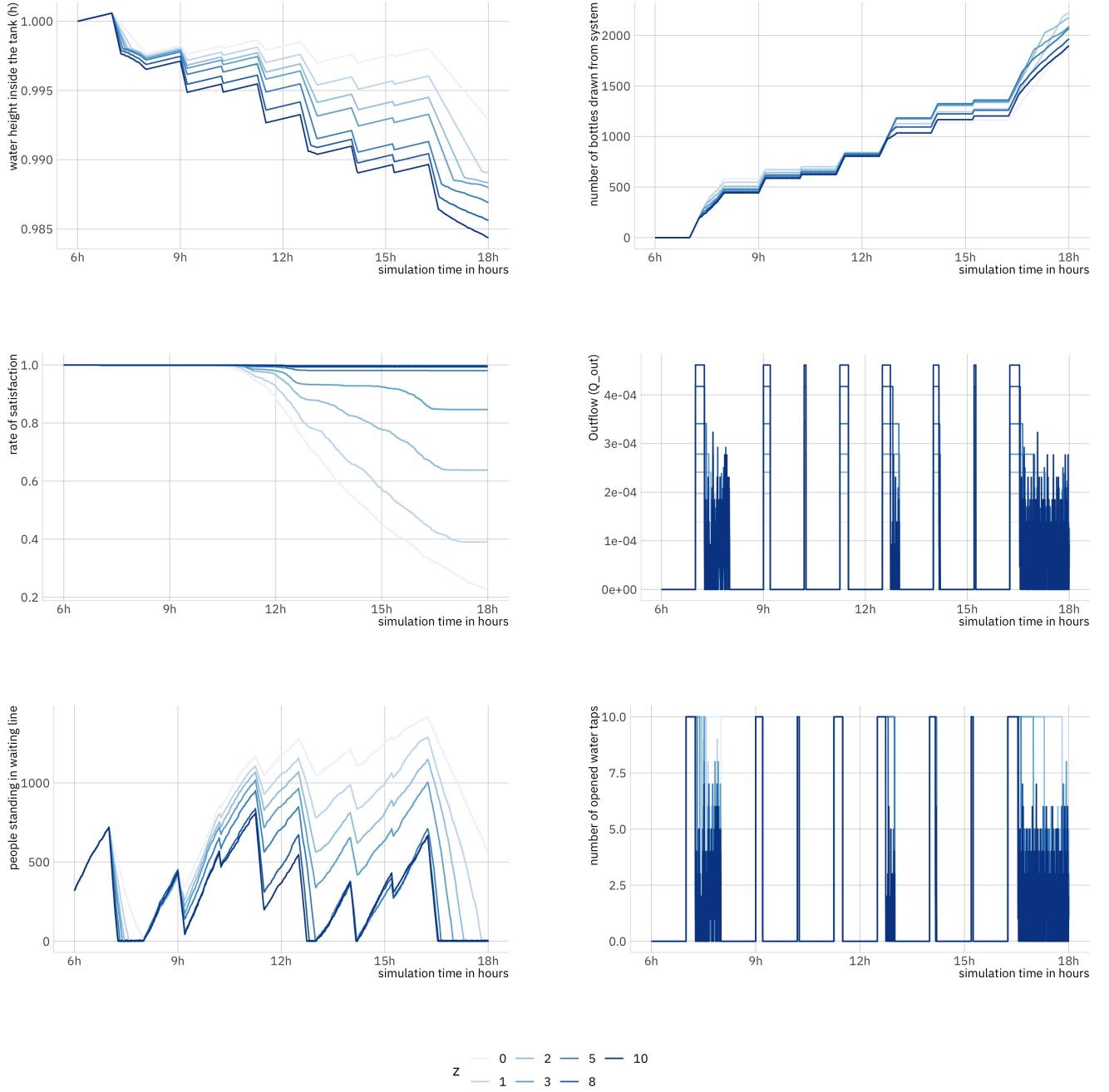


Figure 4: Variation of z

Contrarily to the stochastic model, where variations of the elevation difference between bottom of the tank and taps showed a behavior almost identic to n_{taps} , in this project there are some differences, as illustrated in Figure 4. The influence on the social system is very similar, because z , just like n_{taps} , strongly determines the maximum of Q_{out} , because it increases pressure on the taps. Therefore, outflow, waiting line and satisfaction behavior are very close to the variations of n_{taps} .

The difference lies in the water height at the end of the day. While almost all lines converge to the same value for variations of the number of taps, this does not happen with z . Instead, as z increases, the final water height declines. Again this can be explained by the increased pressure \rightarrow increased outflow \rightarrow bottles fill up more quickly \rightarrow more bottles get too much water, because the time steps do not change. The water is practically spilled, which is a good simile, since with increased pressure, water tends to spray out of the taps and gets lost.

Because of this unpleasant behavior, the number of water taps is assumed to be the better parameter to optimize the system according to the HPC of fulfilling the peoples need for water.

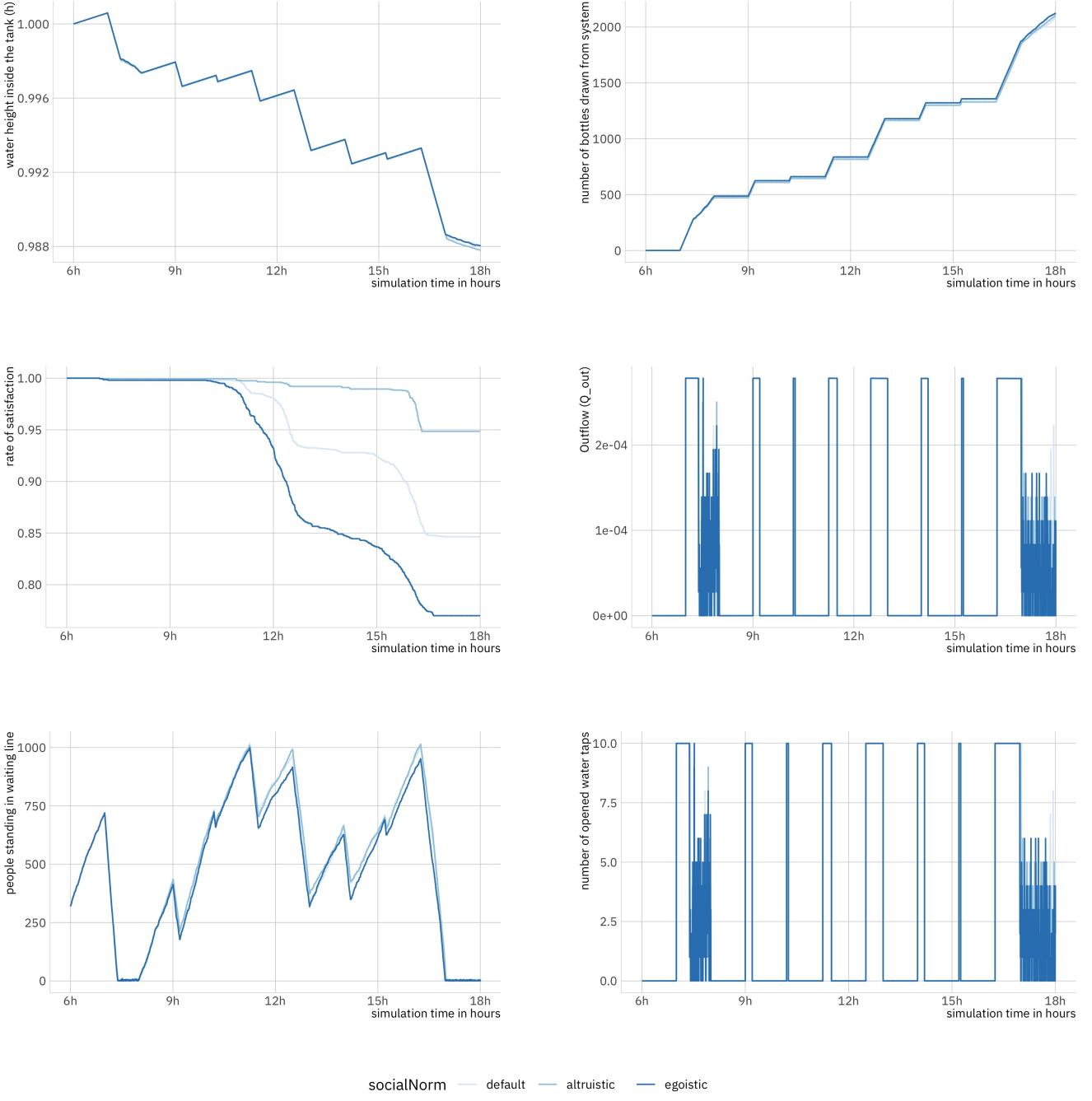


Figure 5: Variation of social norm

The last parameter analysis in Figure 5 is for the newly implemented social norms. Just as expected, these don't have any significant impact on the inflow and outflow aspects of the model. Furthermore, the length of the waiting line isn't altered either, since only the order of it is changed. The complete demand for water stays the same.

The only output that shows significant effects to the change of the social norm is the satisfaction rate. The *altruistic* model shows, the more each individual cares about others, the higher the whole population's satisfaction. This shows that a society grows stronger and improves the well-being of everyone by helping their weakest individuals by giving them small privileges. Therefore it is advisable to motivate people to act mindfully. Nonetheless, reaching such behavior is not very realistic because sorting the line would take too much time and would not be easy with such a large group of people. Another option to reach a similar behavior would be structural or legislative changes like an "express waiting line" for people with low fill levels. Anyhow, considering the other extreme social norm, *egoism*, those structural changes would not help either because people would intentionally spill their water to get in line first.

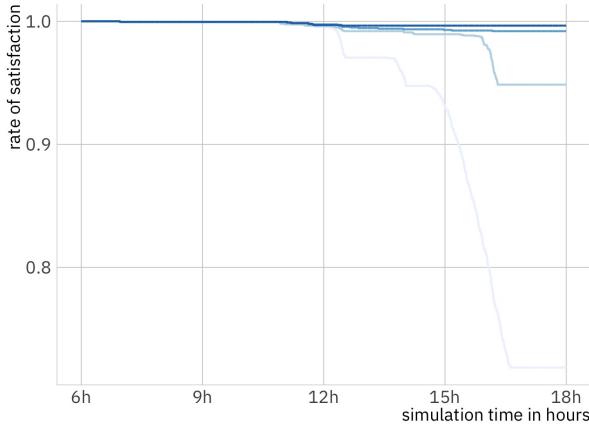
Anyhow, egoism would model the society closer to reality than the other social models. Especially because the taps would not be placed in one location but a variety and people could not sort like this anyway. Additionally, people approach the taps from different locations when breaks start and arrive at diverging points of time, which also supports the method of shuffling the waiting line.

Conclusively the suggestions for values for n_{taps} and z are shown by more detailed variation analyses following the *one factor at a time*-approach (OFAT): all parameters except one are fixed. Parameters are optimized to meet the High Performance Criterion *provide enough water for each individual within school time*. The suggested values are the minimum to reach that goal, of course choosing higher values results in higher satisfaction rates as shown by plots.

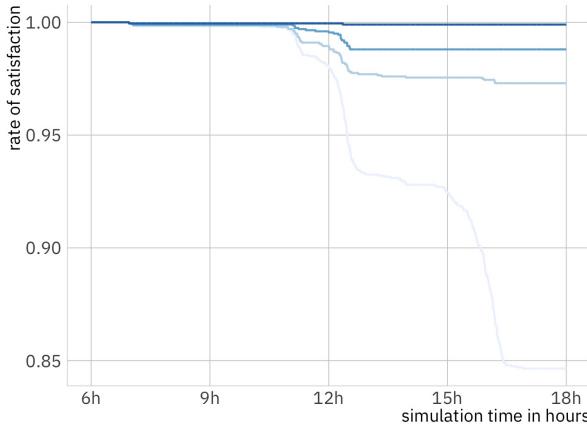
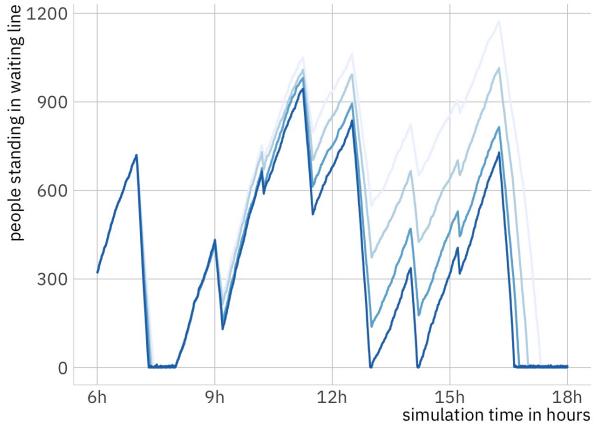
Results for the number of taps are depicted in Figure 6 for each social habitus implemented. The differences induced by the societies behavior are large enough to make the minimum number of taps differ even though they are integers. While ten taps are enough for an altruistic group, the default behavior demands at least eleven and an egoistic group even twelve. Interestingly, the stepwise decrease of gets even clearer for altruistic societies than default ones. That might be because the first in line are not the ones who decided to get there first, but the one's who are most likely to get unsatisfied in the nearest future.

Similarly, the minimum elevation difference is also strongly dependent on the social norm, as shown in Figure 7. In this case, the difference between altruism and egoism amounts even two meters or an increase by 73 percent. The minimum suggested value for z for an altruistic society lies at 2.75 meters, for the default one at 3.5 meters and for egoism at 4.75 meters.

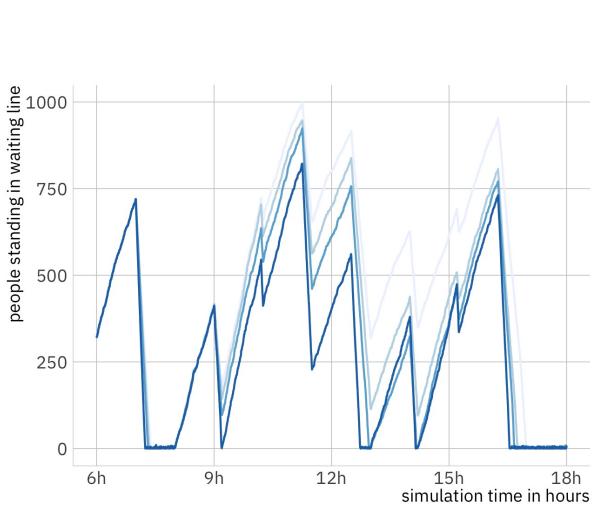
As stated formerly, the egoistic society preference seems to be the one closest to reality while also showing *worst case*-behavior. Therefore decisions and optimizations should be based on results yielded by this option of the model.



n_taps 9 10 11 12
(a) social norm: altruism

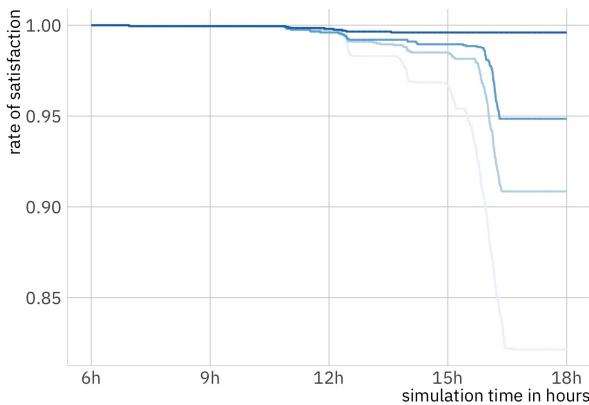


n_taps 10 11 12 15
(b) social norm: default



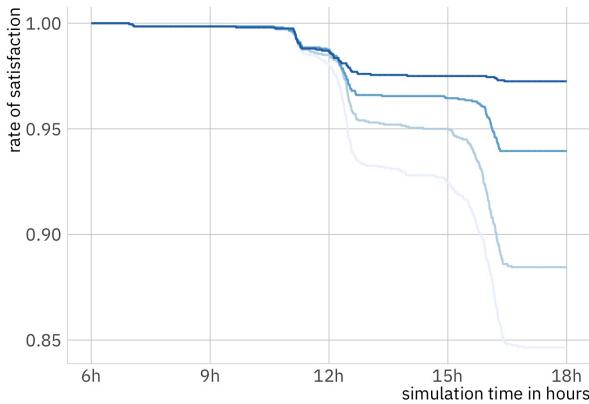
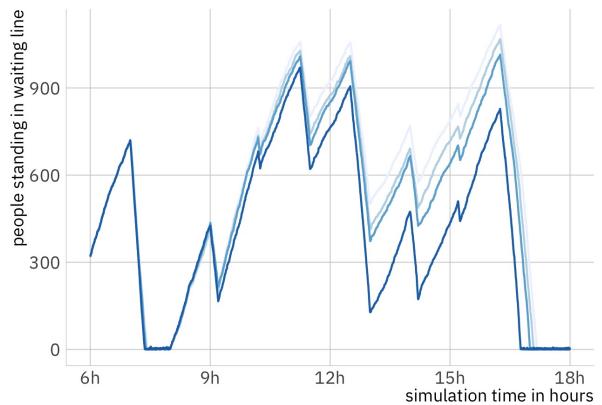
n_taps 10 11 12 15
(c) social norm: egoistic

Figure 6: Detailed variation of n_{taps} with influence of different social norms



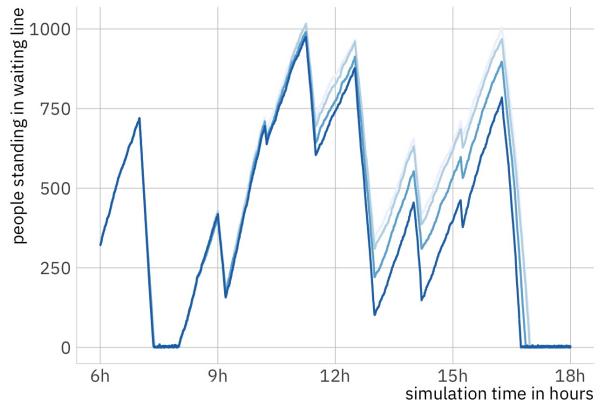
z — 2.5 — 2.75 — 3 — 4

(a) social norm: altruism



z — 3 — 3.25 — 3.5 — 4

(b) social norm: default



z — 4.5 — 4.75 — 5 — 6

(c) social norm: egoistic

Figure 7: Detailed variation of z with influence of different social norms

3 List of References

- Blaustein, R. A., Y. Pachepsky, R. L. Hill, D. R. Shelton, and G. Whelan. “Escherichia Coli Survival in Waters: Temperature Dependence.” *Water Research* 47, no. 2 (February 2013): 569–78. <https://doi.org/10.1016/j.watres.2012.10.027>.
- Egel, Robert. “NumWater.Stochastic: A Numeric Model for Water Supply Systems Including Considerations About Random Precipitation,” January 14, 2020.
- Géron, Aurélien. *Hands-on Machine Learning with Scikit-Learn and TensorFlow Concepts, Tools, and Techniques to Build Intelligent Systems*, 2019.
- Kraut, Richard. “Altruism.” In *The Stanford Encyclopedia of Philosophy*, edited by Edward N. Zalta, Spring 2020. Metaphysics Research Lab, Stanford University, 2020. <https://plato.stanford.edu/archives/spr2020/entries/altruism/>.
- Ooi, Hong, Microsoft Corporation, and Stephen Weston. *doSNOW: Foreach Parallel Adaptor for the 'Snow' Package* (version 1.0.18), 2019. <https://CRAN.R-project.org/package=doSNOW>.
- Ooi, Hong, Microsoft, and Steve Weston. *Foreach: Provides Foreach Looping Construct* (version 1.4.7), 2019. <https://CRAN.R-project.org/package=foreach>.
- Peterson, Brian G., Peter Carl, Kris Boudt, Ross Bennett, Joshua Ulrich, Eric Zivot, Dries Cornilly, et al. *PerformanceAnalytics: Econometric Tools for Performance and Risk Analysis* (version 1.5.3), 2019. <https://CRAN.R-project.org/package=PerformanceAnalytics>.
- Renard, Philippe, Andres Alcolea, and David Ginsbourger. “Stochastic Versus Deterministic Approaches.” In *Environmental Modelling*, edited by John Wainwright and Mark Mulligan, 133–49. Chichester, UK: John Wiley & Sons, Ltd, 2013. <https://doi.org/10.1002/9781118351475.ch8>.
- Shaver, Robert. “Egoism.” In *The Stanford Encyclopedia of Philosophy*, edited by Edward N. Zalta, Spring 2019. Metaphysics Research Lab, Stanford University, 2019. <https://plato.stanford.edu/archives/spr2019/entries/egoism/>.
- Sievert, Carson, Chris Parmer, Toby Hocking, Scott Chamberlain, Karthik Ram, Marianne Corvellec, Pedro Despouy, and Plotly Technologies Inc. *Plotly: Create Interactive Web Graphics via 'Plotly.js'* (version 4.9.1), 2019. <https://CRAN.R-project.org/package=plotly>.
- Taddy, Robert B. Gramacy and Matt A. *Tgp: Bayesian Treed Gaussian Process Models* (version 2.4-14), 2016. <https://CRAN.R-project.org/package=tgp>.
- Tierney, Luke, A. J. Rossini, Na Li, and H. Sevcikova. *Snow: Simple Network of Workstations* (version 0.4-3), 2018. <https://CRAN.R-project.org/package=snow>.
- Wickham, Hadley, Winston Chang, Lionel Henry, Thomas Lin Pedersen, Kohske Takahashi, Claus Wilke, Kara Woo, Hiroaki Yutani, and RStudio. *Ggplot2: Create Elegant Data Visualisations Using the Grammar of Graphics* (version 3.2.1), 2019. <https://CRAN.R-project.org/package=ggplot2>.
- Wilke, Claus O. *Cowplot: Streamlined Plot Theme and Plot Annotations for 'Ggplot2'* (version 1.0.0), 2019. <https://CRAN.R-project.org/package=cowplot>.