Supplementary material for: Generative AI for the optimal design of seawater desalination processes

Valentin ZARLENGA[a,b,c*], Antonio ROCHA AZEVEDO[a,b], Alvaro MARTINEZ-TRIANA[b] and Thibaut NEVEUX[a,b]

[a] Université de Lorraine, CNRS, LRGP, F-54000, Nancy, France
[b] EDF R&D Chatou, 6 quai Watier, 78400 Chatou, France
[c] Université de Strasbourg, ECPM, 25 rue Becquerel, 67200 Strasbourg, France
* Corresponding Author: valentin.zarlenga@univ-lorraine.fr.

1. PHYSICAL VALIDITY PROCESS RULES

Processes generated may contain some physical inconsistencies. To ensure their validity, several tests are performed to determined whether:

- The conversion between SFILES notation and its graph representation is performed correctly
- The number of inlets is equal to 1 and correspond to the "Raw" token
- The number of outlets is equal to 2 and correspond to the "Perm" and "Ret" token
- {P} and {R} tokens are placed at the exits of desalination units
- The specific connectivity to the equipment is respected (see Table 1)
- The two outputs of a splitter aren't linked to the two inputs of a mixer
- The mixer output flow do not become one of its input flows

2. THERMAL SAMPLING STRATEGY

The thermal sampling strategy has been described in the paper of Hinton and al [1]. In fact, it is possible to influence the probability distribution of token prediction by using the "Temperature" hyperparameter (control randomness in predictions) thank to this equation:

$$P(\text{token}) = \frac{e^{\frac{P(token)}{T}}}{\sum e^{\frac{P(token)}{T}}} \quad (1)$$

If the "Temperature" is <1, the distribution will be more predictable for the token selection (pointed distribution). If the "Temperature" is >1, the distribution will be more random so more innovate for the token selection (flatter distribution).

An experiment was conducted to see if increasing the temperature would result in new SFILES sequences being found, using the method describes in section 2.5 of the article. A temperature of 1.5 generates new design areas (see Figure 1) compared to the normal way. A more in-depth study could increase this gain but that is out of this work's scope. Indeed, high temperature flattens the probability distribution, allowing more randomness and novelty, but can also lead more mistakes.

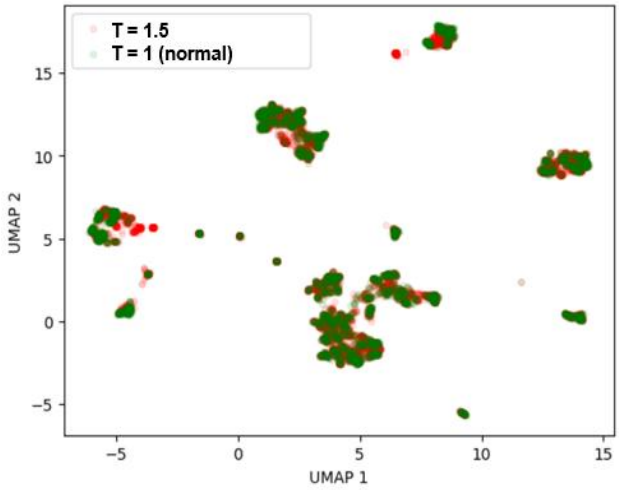

Figure 1. Evaluation of search space exploration for a temperature of 1 (green points) and 1.5 (red points).

3. OBJECTIVE FUNCTION WEIGHTS

The weights selected in the study are:

- 10 for the recovery rate (noted k_{Tr})
- 100 for the electrical power consumption (noted k_w)

	Token associated	Number of inlets	Number of outlets	Exit tags
Seawater solution	(raw)	0	1	
Pump	(pump)	1	1	
Reverse osmosis	(RO)	1	2	{R} and {P}
Mixer	(mix)	2	1	
Splitter	(splt)	1	2	
Permeate	(Perm)	1	0	
Retentate	(Ret)	1	0	
Membrane distillation	(MD)	1	2	{R} and {P}

Table 1. Specific connectivity to the equipment used in this work

- 4 for the total surface area used (noted k_S)
- 1 for the heat duty required (noted k_h)
- 3 for the NaCl concentration (noted k_c/m)

There are two terms for the NaCl concentration because a condition was imposed. If a permeate concentration less than 1000 is obtained, the penalty will be slowly initiated (slow-growing exponential function). Otherwise, the penalty is stronger with a linear function (see Figure 2).

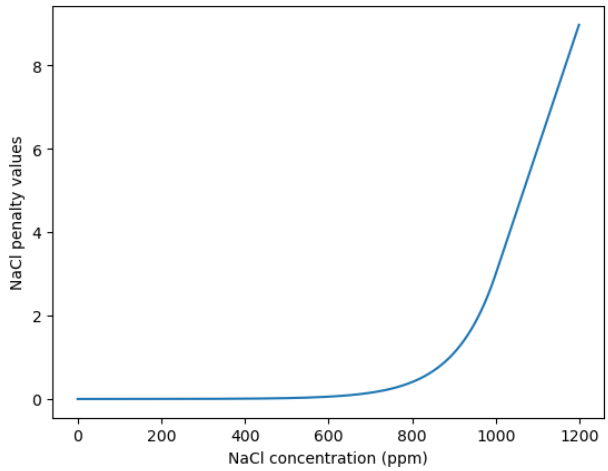

Figure 2. NaCl penalty in function of NaCl concentration

4. PARAMETERS OF THE SELECTED PROCESSES

The parameters of the presented processes in the article are given in the Figure 3-7. The parameters represented are:

- The discharge pressure (P, in bar) and below the electric work (E, in kW) for the pump
- The stage cut (RR, in %) and the calculated membrane surface (S, in m^2) for the reverse osmosis
- The permeate flowrate (q_{perm}, in kg/h) and the NaCl concentration for the permeate exit
- The initial flowrate (q_{init}, in kg/h) and the NaCl concentration for the entry

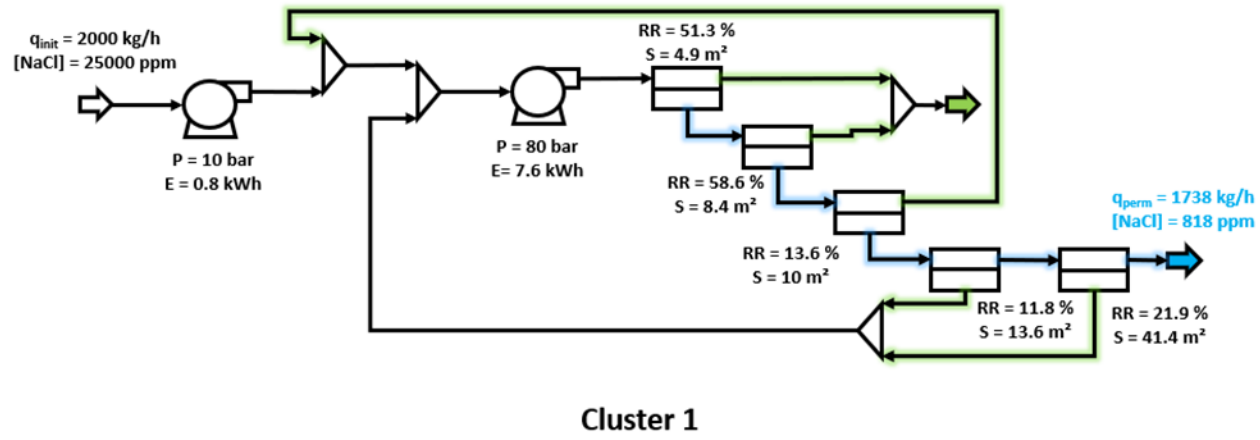

Figure 3. KPI of the cluster 1 process

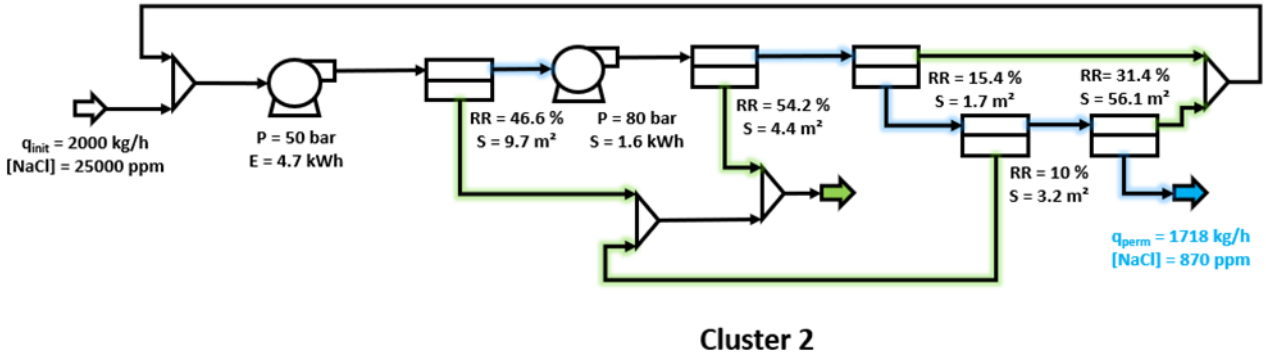

Figure 4. KPI of the cluster 2 process

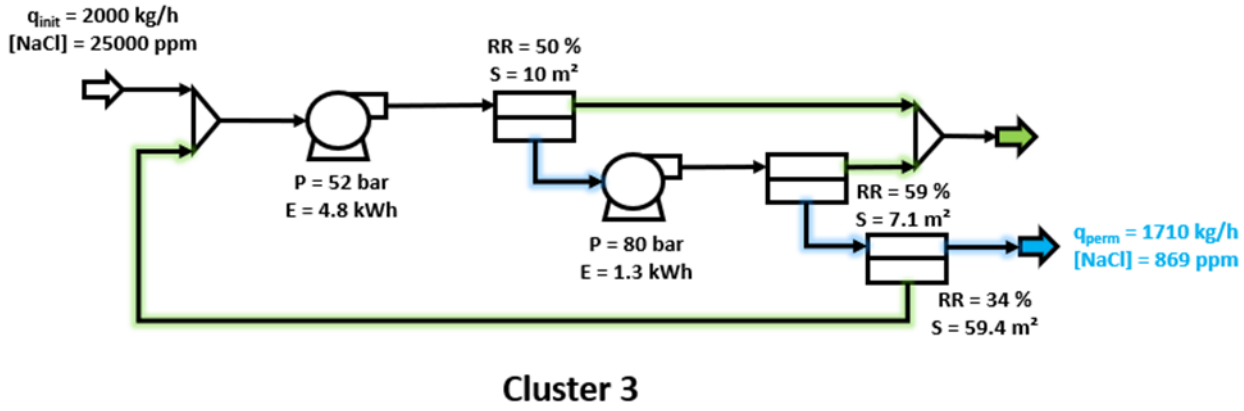

Figure 5. KPI of the cluster 3 process

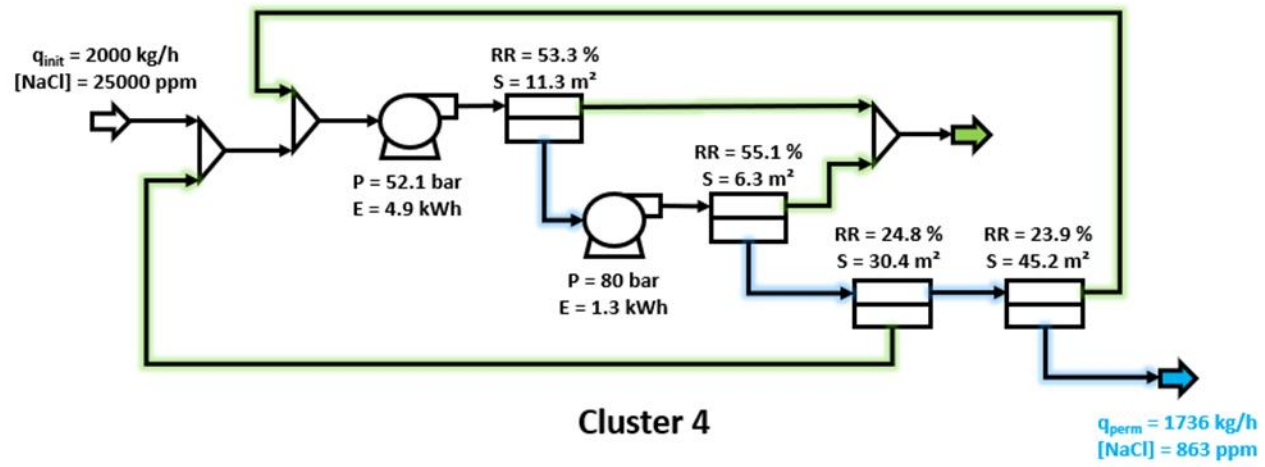

Figure 6. KPI of the cluster 4 process

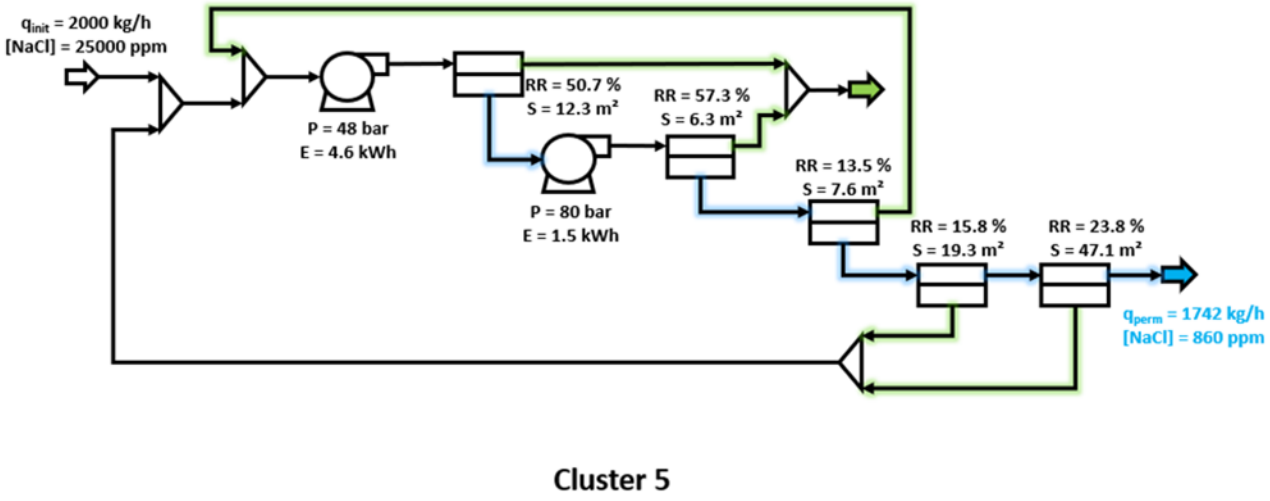

Figure 7. KPI of the cluster 5 process

5. REFERENCES

1. Hinton, Geoffrey & Dean, Jeff & Vinyals, Oriol & Rachmad, Yoesoep. (2014). Distilling the Knowledge in a Neural Network. 1-9. url: https://arxiv.org/pdf/1503.02531.