Evaluation of Danube River in the Gabcikovo Region

<u>GROUP 2</u>

Dennis Zanutto

Emmanuel Nyero

Francesco Capone

Hosea Naimoi

Politecnico di Milano

Table of Contents

The objectives of the MATLAB Project are listed below.

I. Build a one day ahead forecast model of the river streamflow

II. Discuss the impacts of damming up the river

1. MODEL SELECTION FOR RIVER DANUBE STREAMFLOW FORECAST AT GABCIKOVO

1.1. INTRODUCTION

Figure 1: Gabcikovo-Nagymaros Dams

Source: Google Earth

Flowing over 2.850 kilometers from the Black Forest in Germany to the Black Sea in Romania, the Danube is Europe's second-largest river after the Volga. It ranks 21st in the world. The Danube basin drains an area of 817.000 km2 and transfers water from the non-riparian countries of the Czech Republic, Slovenia, Albania, Macedonia, Italy, Switzerland, and Poland. Over 300 tributaries flow into the Danube, and 80 million people live in the river basin. Throughout its length, the Danube River provides a valuable resource for many competing uses. Downstream from Slovakia, the river is the primary source of drinking water in all the countries except Bulgaria, and it is a crucial source in Austria and Slovakia. The river is also used extensively for

irrigation, especially in the Hungarian plain. Fisheries are a valuable source of food and income at its lower reaches, and the Danube Delta at the Black Sea is a significant tourist area.
The Danube is also vital for industry, including hydroelectric generation, industrial cooling, and waste disposal. The mountainous character of the Danube in its upper reaches and a large number of tributaries further downstream combine to make the energy potential of the river significant. There are over 40 hydropower stations on the upper Danube, which are matched in energy output by the two enormous Iron Gate stations between Serbia-Montenegro and Romania. There are also a large number of dikes, navigation locks, and other hydraulic structures to aid navigation.

Figure 2: Bird's Eye View of the Danube River at the Gabcikovo-Nagymaros Dams

1.2. MODELING

Natural processes are often described as a stochastic process, where a function of time is associated with every outcome of a random experiment. The evolution of a random system in time can be described using mathematical models to understand the system and use forecasting and simulation to contribute in the final decision-making. The available data are 9856 values of the historical trajectories of average daily streamflow (m3/s), average daily precipitation (mm/day), and average daily temperature (◦C) in the period between January 1st, 1990 and December 31st, 2016. Ignoring leap years, an equal number of days in all years (365 days per year), is considered, making the entire data trajectory 27 years.

1.3. METHODOLOGY

To build a model that captures the main dynamics of the system, provided that it ideally reproduces the measured input into measured outputs, the model parameters are determined by comparing measures and model output, then the model's ability to reproduce the system behaviour is tested using different data. The historical time series of net inflow, precipitation, and the temperature were modelled and can be seen in the following figures.

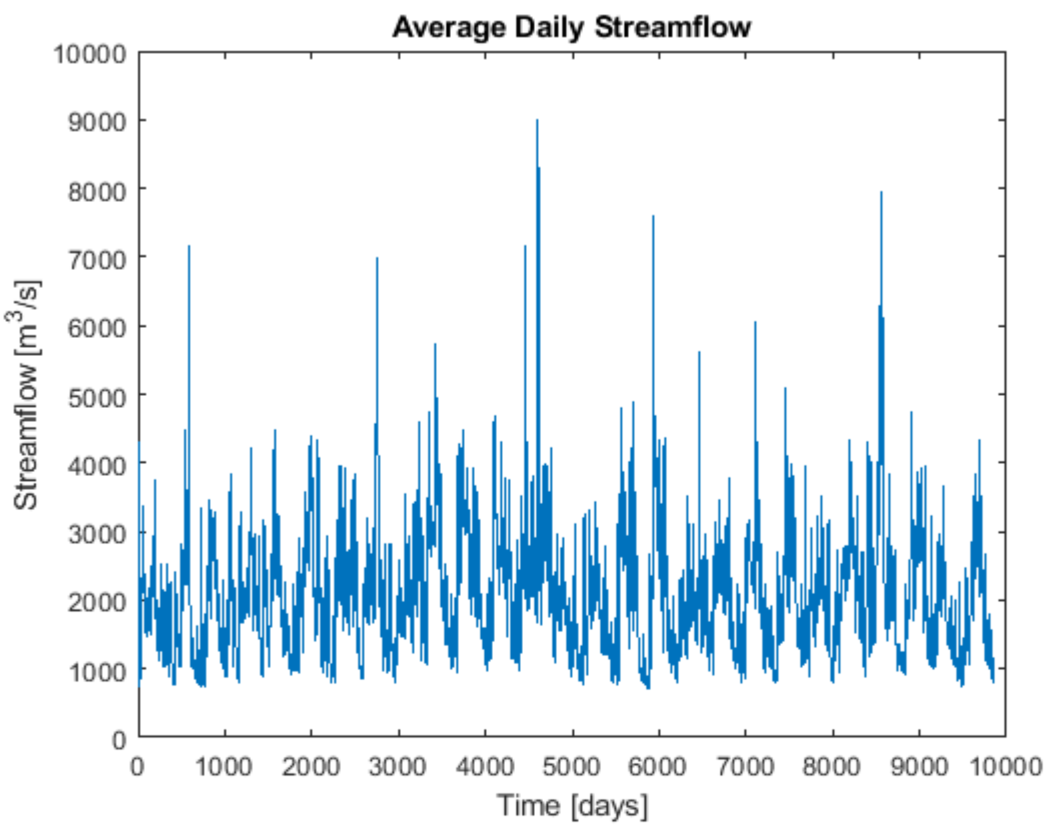


Figure 3: Average Daily Streamflow

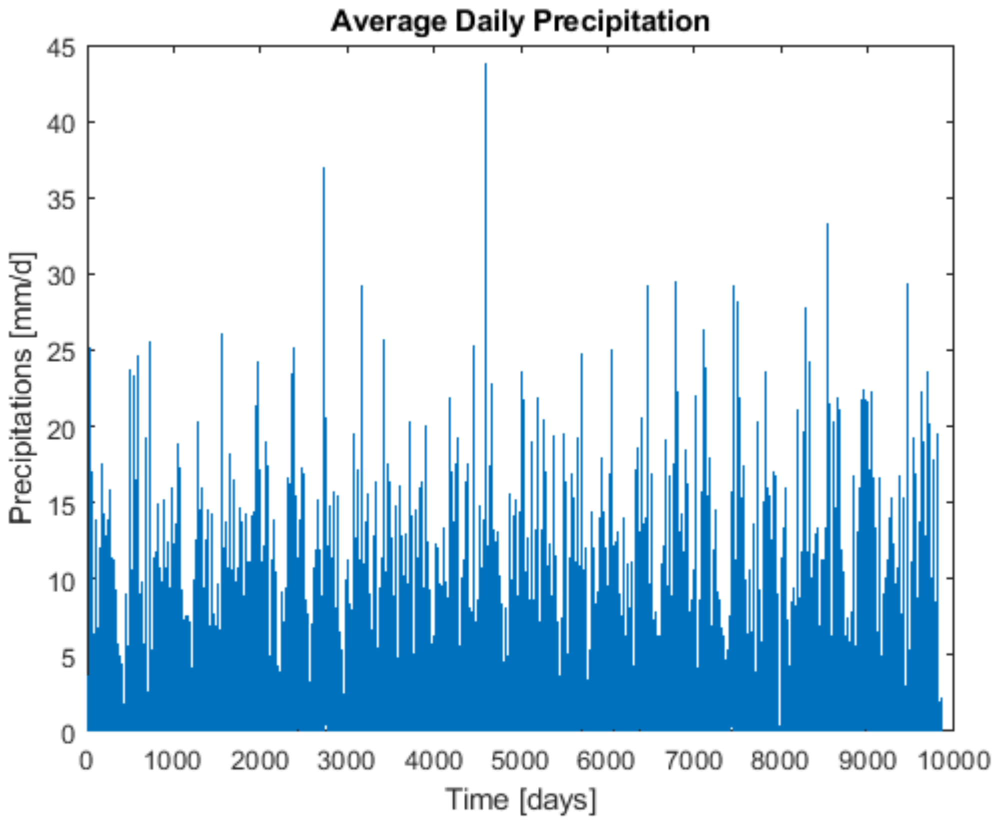


Figure 4: Average Daily Precipitation

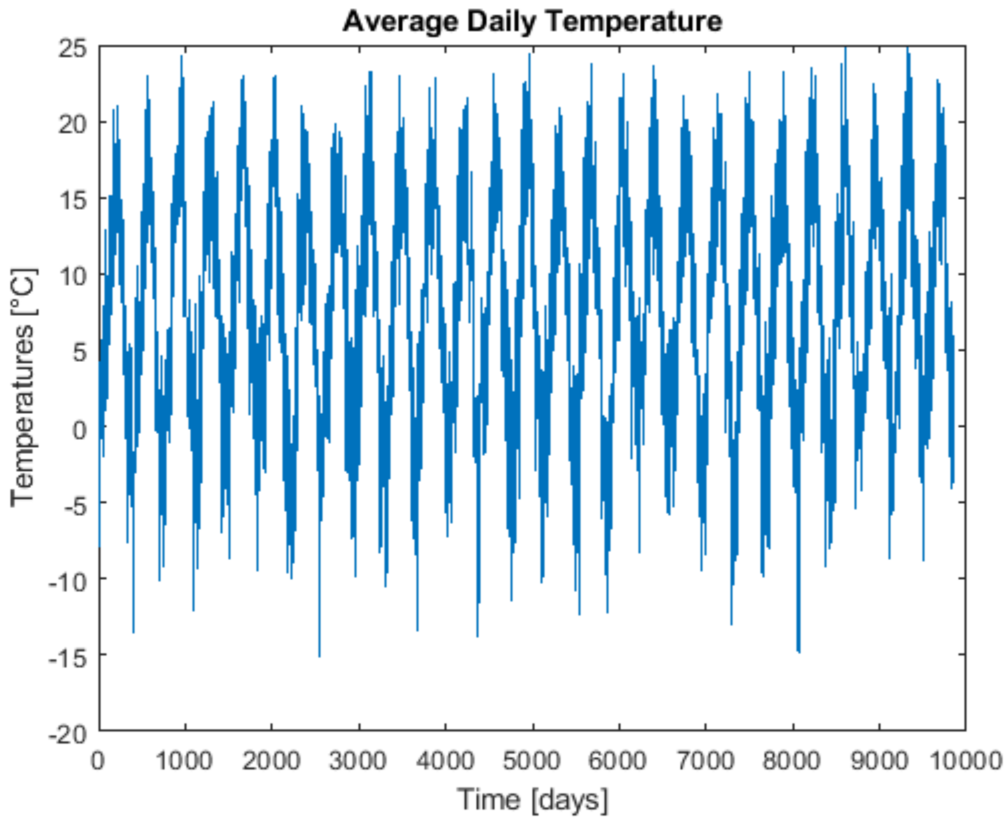


Figure 5: Average Daily Temperature

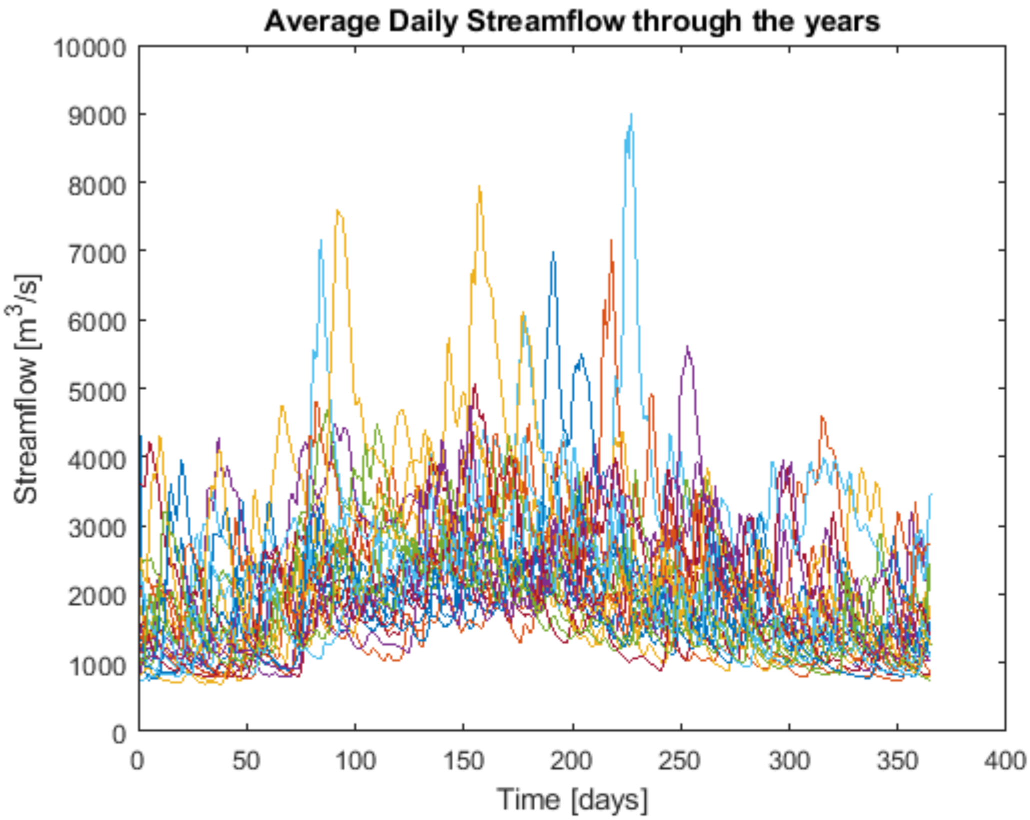


Figure 6: Average Daily Streamflow for 27 Years

To deseasonalize the data, the general deterministic trend in data must be removed using the information of the system itself. Under the hypothesis of cyclo-stationarity, graphs are plotted with a 2, 4, 10, and 20—days window from the first day to the last of each of the 27 years. Hence, a moving-window average approach was used to estimate a replicate.

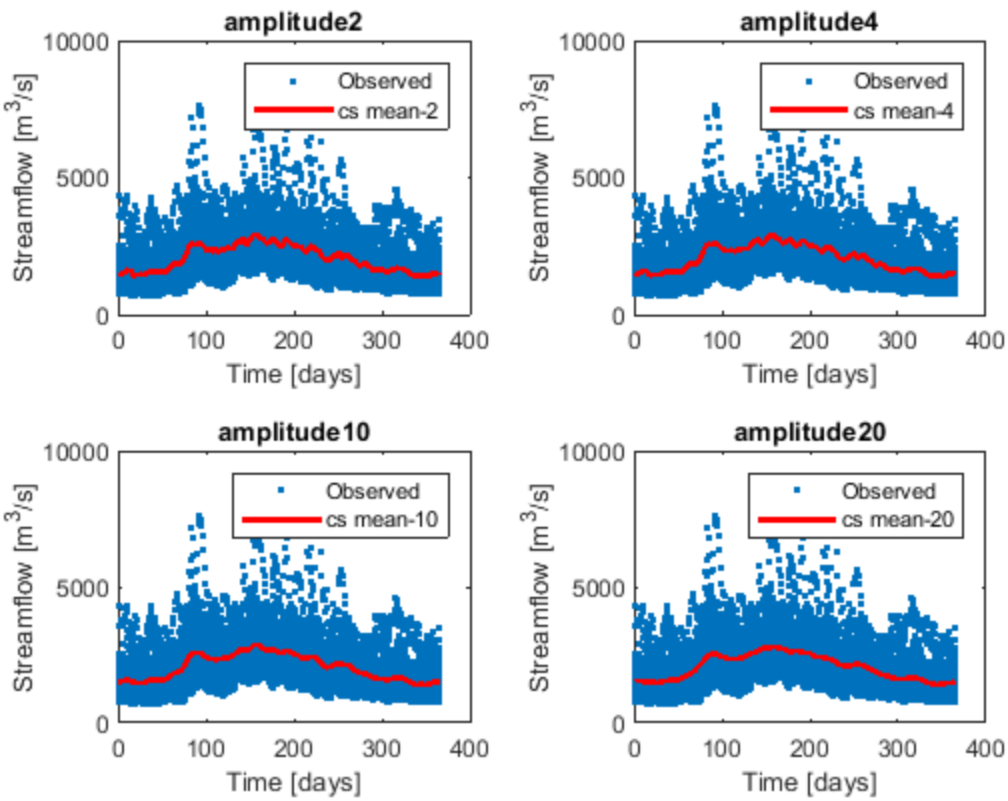


Figure 7: Cyclostationary Mean with Different Values of Moving Average Windows

After removing trends from data, temporal behaviour is evaluated by investigating correlation within the data.

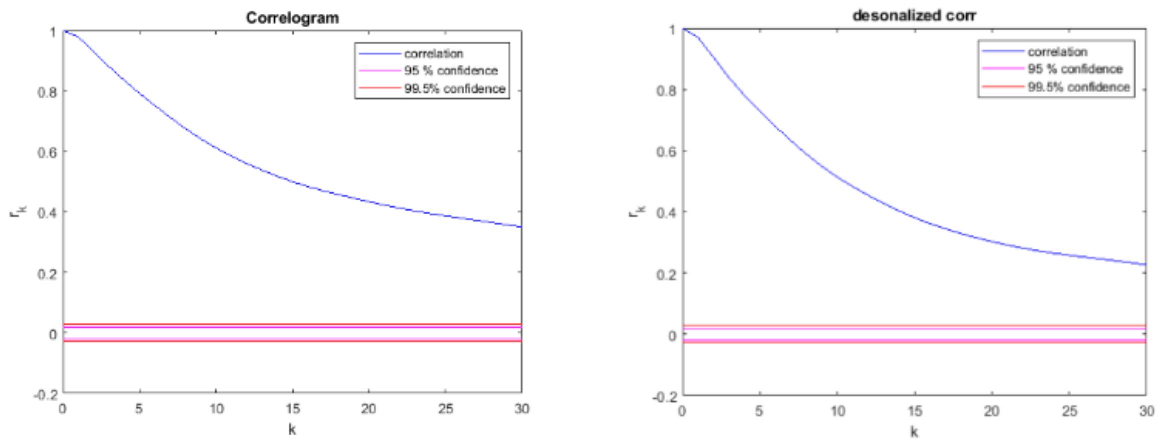


Figure 8: Seasonalized vs. Deseasonalized Correlation of Streamflow

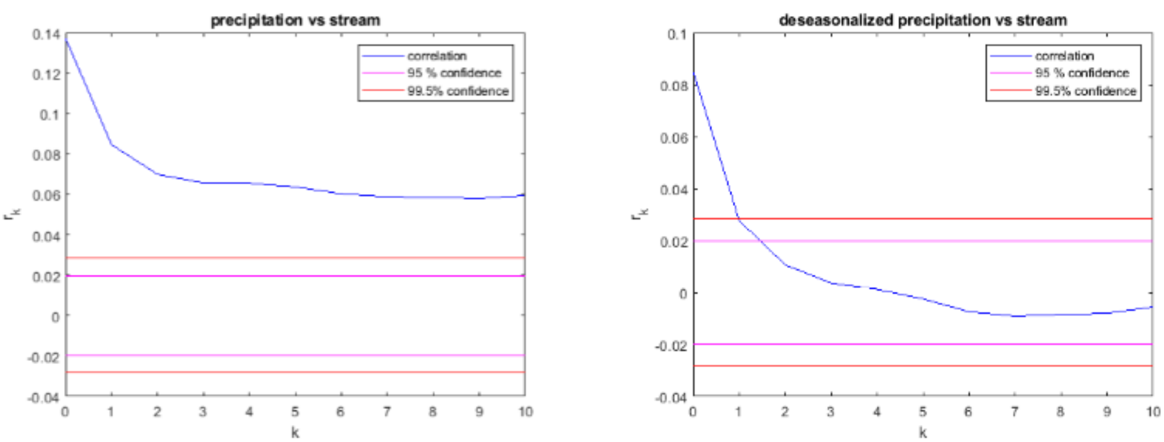


Figure 9: Seasonalized vs. Deseasonalized Streamflow-Precipitation Correlation

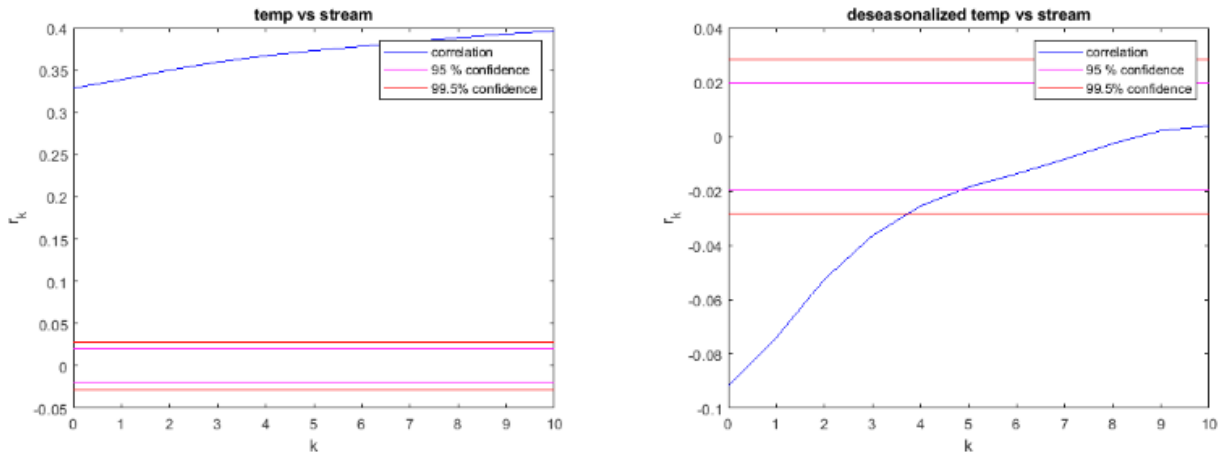


Figure 10: Seasonalized vs. Deseasonalized Streamflow-Temperature Correlation

The correlation seems to gradually decrease until time lag *k* = 15 days while a sharper declination can be observed in the normalized data before it asymptotically flattens at k=20 days (herein, streamflow strongly influences itself over 20 days). Besides, correlation does not reach the zero in the lag of 20 days considered in the figure; correlation extends beyond 20 days which comes to no surprise since the system is subjected to seasonality, where wet seasons (fall and spring) have a high flow for approximately 90 days and adversely for dry seasons (summer and winter).

1.4. LINEAR AUTOREGRESSIVE (AR) MODEL

Starting from the deseasonalized streamflow, under the hypothesis of autocorrelation and taking into account the analysis of streamflow correlogram (correlation is not zero), AR models could be suitable to model the temporal evolution. To measure the performance of a model performance indices of Mean Squared Error and coefficient of determination (R2) are calculated to compare the model's results. The program automatically computes optimal parameter value p using K-cross validation with k = 9. After that, it plots the value of the objective function MSE and R2 with the order of the model (for both calibration and validation).

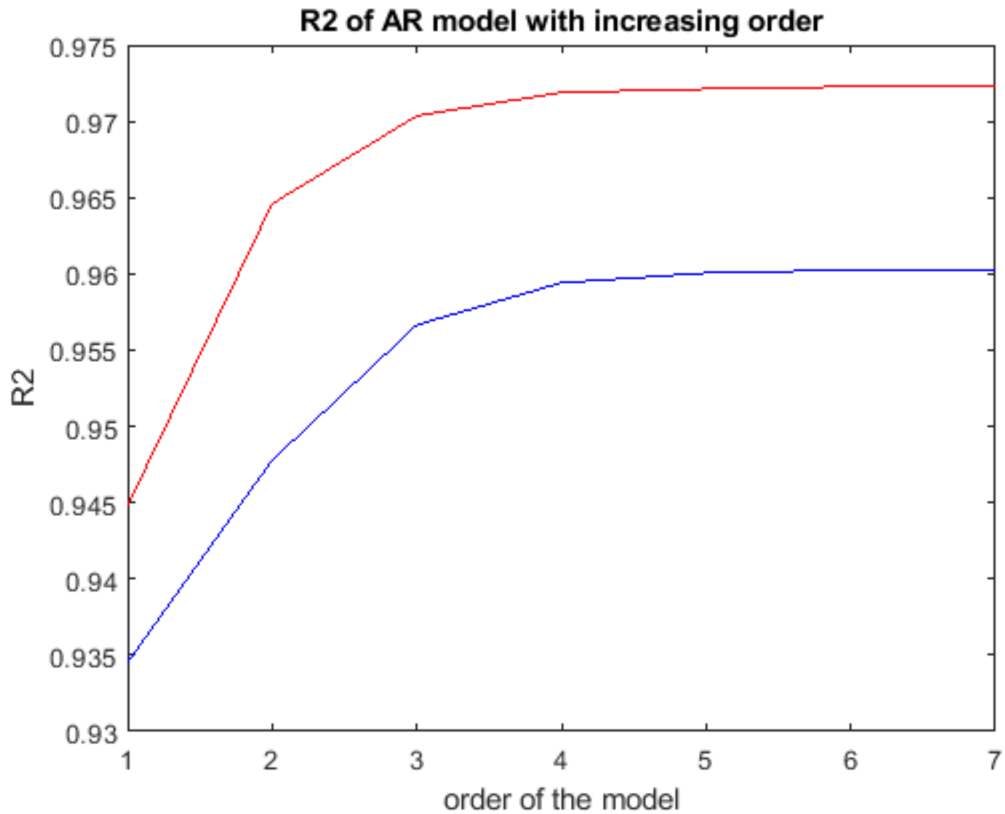


Figure 11: R^2 of AR Model with Increasing Order

Starting from the AR (1) and increasing the order, the value of R2 increases until it stabilizes at 0.972 in calibration and 0.96 in validation, both in AR (5) and maintaining their values until AR (7). Having the maximum R2 value also reached in AR (4) could indicate that correlation has a significant influence on the second time step as well. Increasing the model order, correlation is not high enough to improve the results.

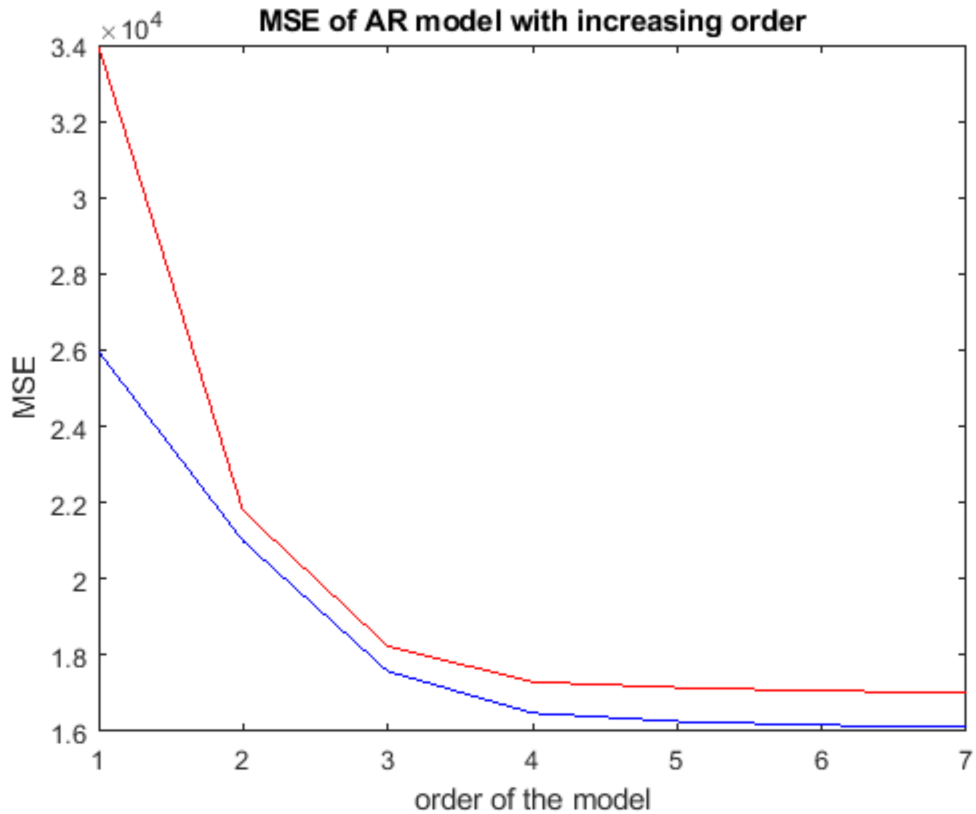


Figure 12: Mean Squared Error of AR Model with Increasing Order

The same is behaviour is obviously noticed in the MSE plot.

1.5. LINEAR AUTOREGRESSIVE EXOGENOUS MODEL

To enhance the performances of the identified model structures in AR models discussed in the previous section, exogenous information, in our case precipitation and temperature data, are employed. First, we compute ARX with precipitation as the exogenous part, then with both precipitation and temperature. Since adding temperature does result in any significant variation, as seen in Figure 13, we ran ARX with only temperature and compared it with only precipitation, as shown in Figure 14. It was identified that using precipitation leads to better results; thus, only precipitation would be used from this point henceforth.

The following table gives the corresponding parameter values of p and q which are used to plot R2 for figure 13.

index for the plot	**p**	**q**

1	1	1
2	1	2
3	1	3
4	1	4
5	2	1
6	4	1
7	6	1
8	6	2
9	6	3
10	9	1
11	9	2
12	16	1
13	16	2

Table 1: Corresponding Parameter Values for p and q as represented in Figure 13

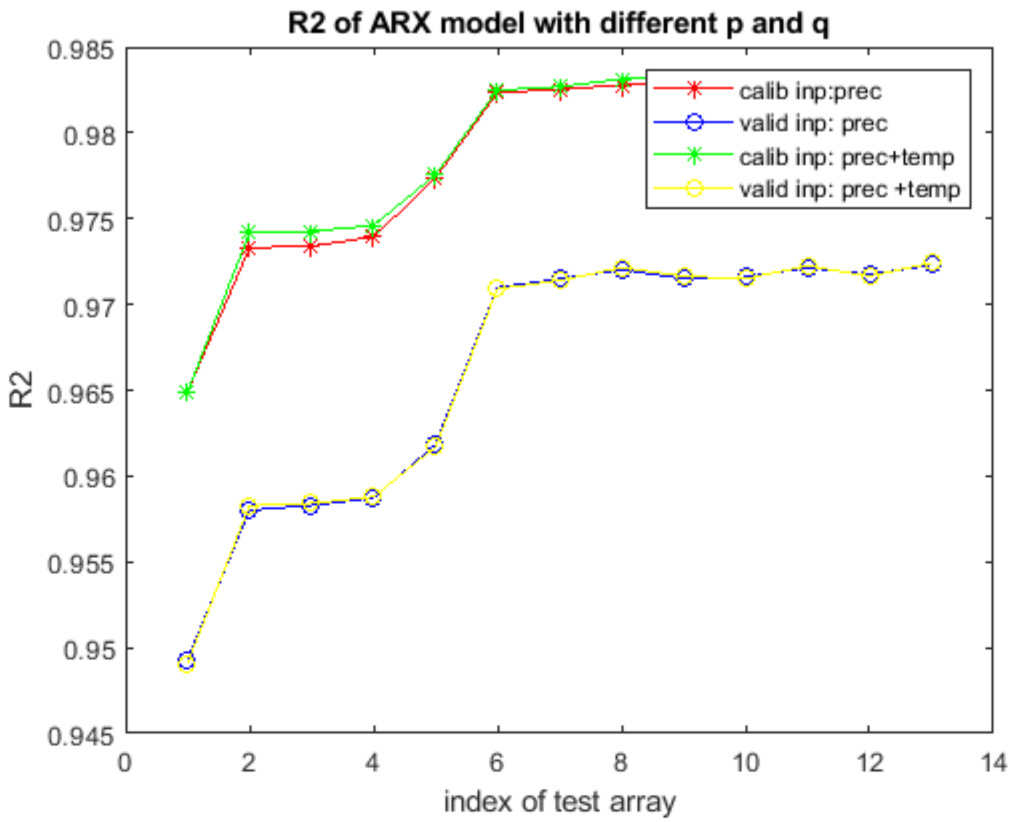


Figure 13: R2 of ARX Model with Different Values of p and q

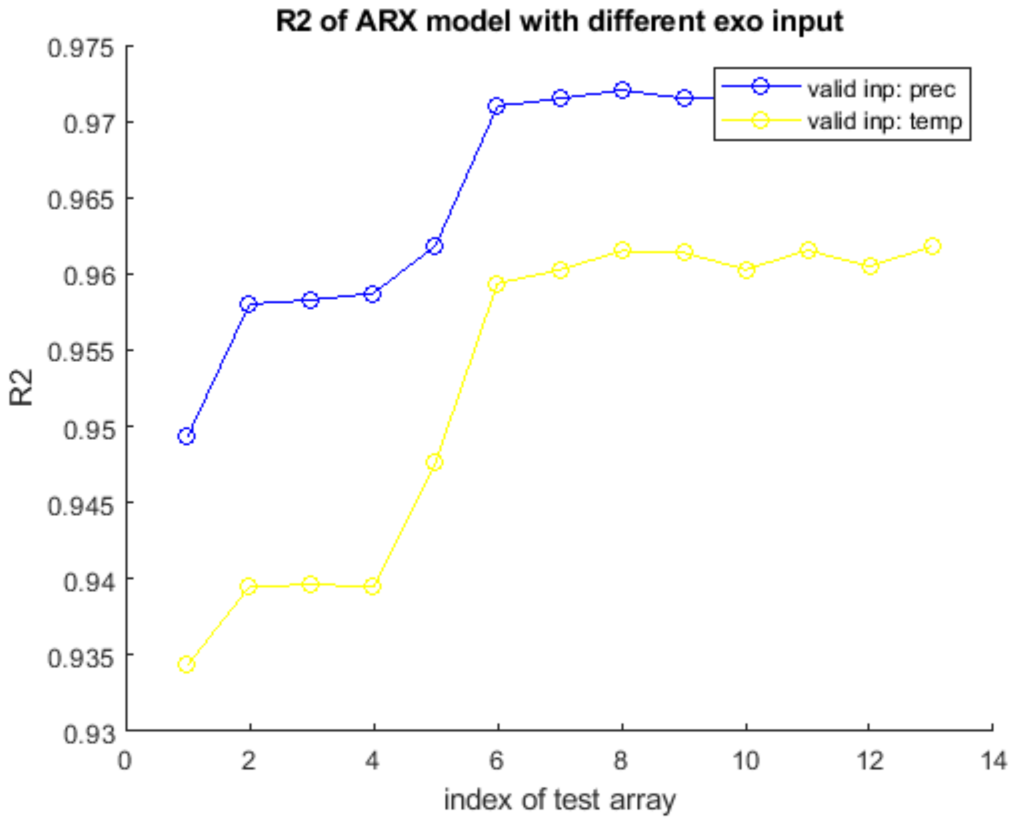
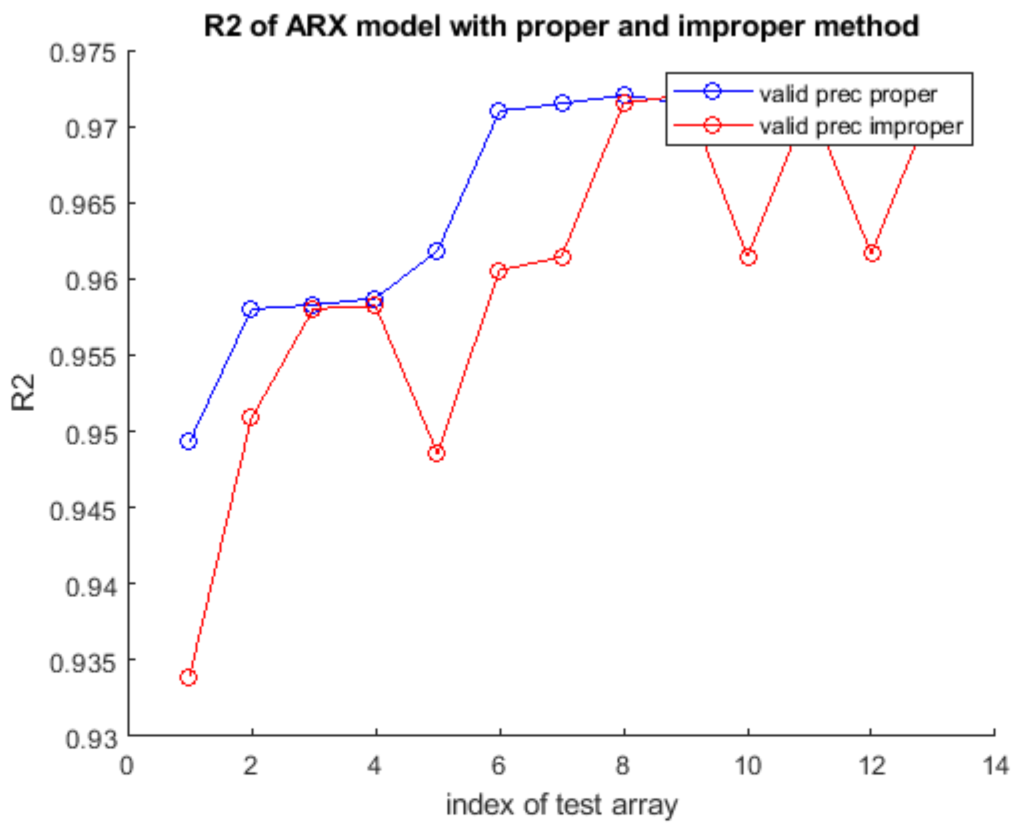


Figure 14: R2 of ARX Model with Different Values of Exogenous Input

Finally, we ran the ARX model with precipitation but as an improper model and compared it with a proper model (Figure 14). The main finding is that the proper model is slightly better for low values of q.



1.6. ARTIFICIAL NEURAL NETWORKS (ANN) MODEL

ANN is an empirical non-linear mathematical model inspired by biological neural networks. Input and Output have a linear relation, processed through a non-linear basis function called Neurons, which can be displayed in layers. ANN calibration is automatically performed in MATLAB, using the Mean Squared Error (MSE) as an objective function and solved using Levenberg-Marquardt (LMA) technique. Furthermore, for this model the MATLAB training algorithm for the ANN already divides the sample in calibration, validation and train sets, so we don't need to divide our time series before feeding it to the training algorithm. Our program tries wit han increasing number of neurons until the R2 flats or decreases. In the end, the best number of neuron is the minimum 3. Since ANN uses a stochastic algorithm, we ran it multiple times. We observed that the result remains the same despite using 10 or more neurons.

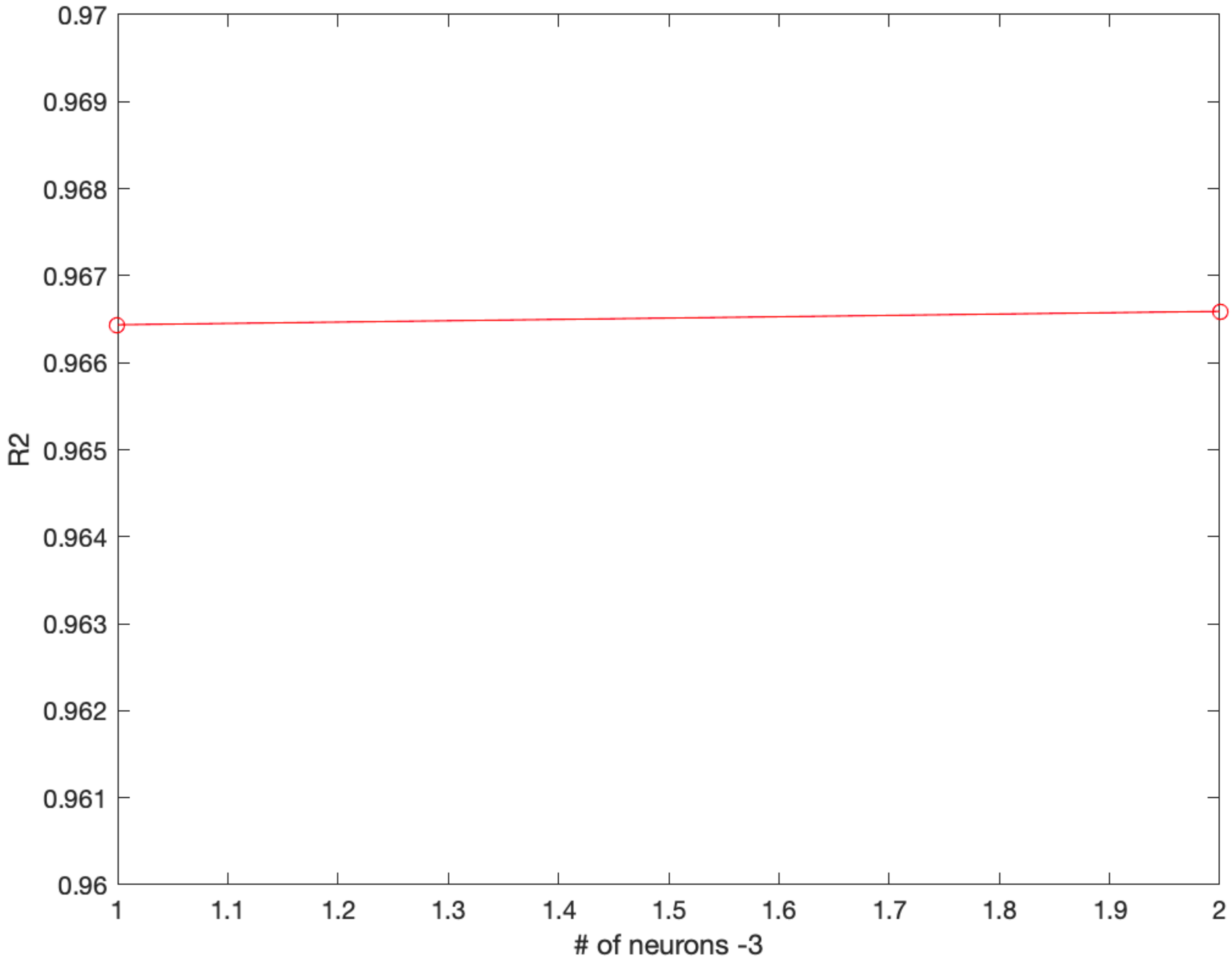


Figure 15: Artificial Neural Network R2 for 4 and 5 Neurons

1.7 RESULTS AND COMPARISONS

A plot was made for comparisons to identify the best model among the different models used.

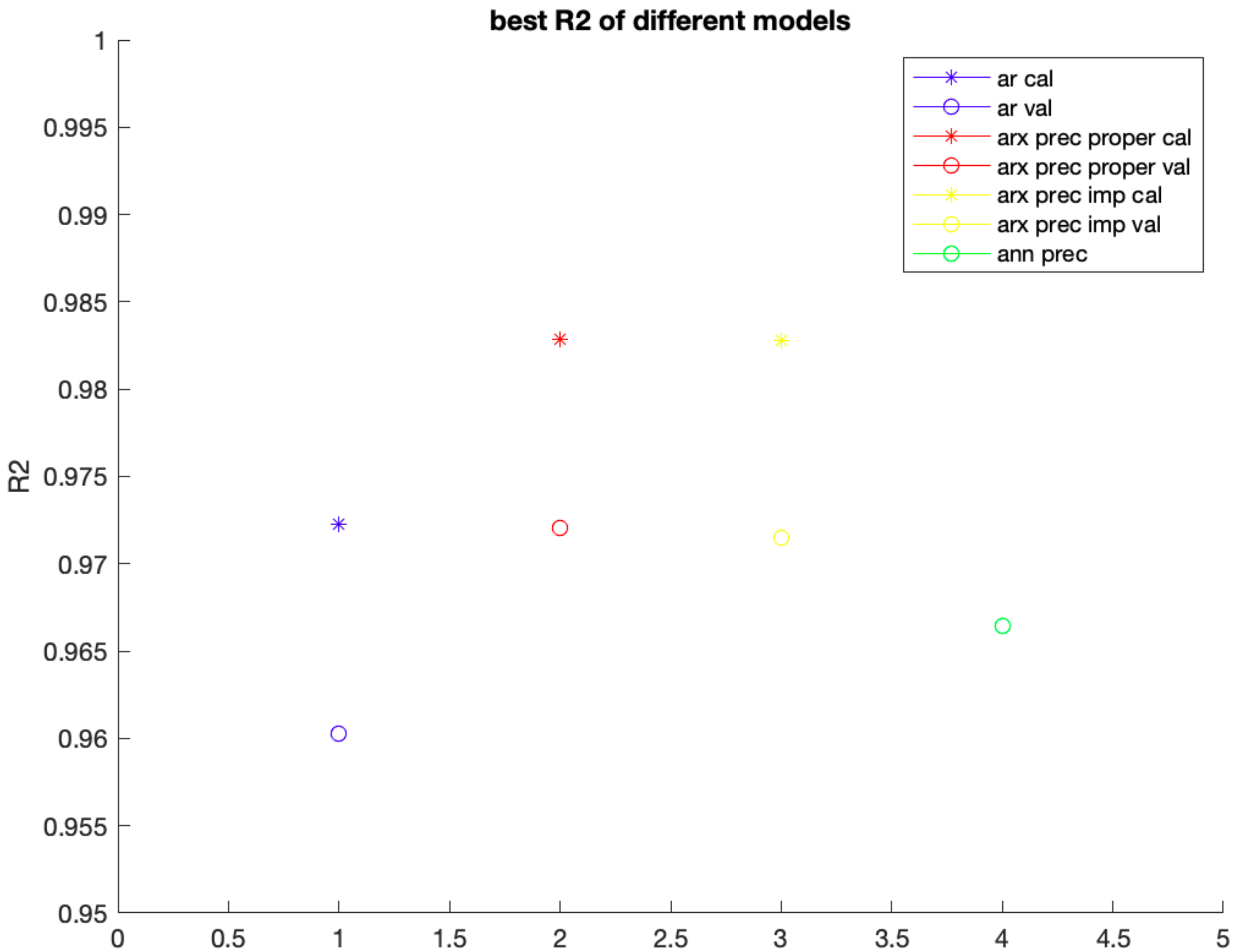


Figure 16: Best R2 (Calibration and Validation) for Different Models

Regarding the coefficient of determination R2, increasing the model order leads to higher R2, which means better identification. However, increasing the model complexity by using a non-linear model does not lead to an improvement in the determination coefficient compared to ARX (since it performs better than ANN). Therefore, a proper ARX has the best performance among the different models.

2. IMPACTS OF DAM CONSTRUCTION ON RIVER DANUBE AT GABCIKOVO

The second objective of the project is to select an appropriate design and adjusting operating policies of the Gabcikovo Dam on the Danube River in the Gabcikovo region. Once the dam is dimensioned the different performance will be evaluated by the release regulating policy, taking to account different stakeholders. For this purpose, first the release regulating policy are obtain by changing the required parameters manually then for considering all of the possible alternatives in objective space and not choosing the parameters manually, Evolutionary Direct Policy Search Algorithm (EMODPS) has been implemented in MATLAB to find the Pareto Frontier of optimal decisions, minimizing irrigation water supply and flood risk. Three of these optimal decisions have been analyzed in detail, leaving to the decision maker the possibility to choose the one which best matches stakeholders' interests.

2.1. PROJECT INDICATORS

To assess the satisfaction of a particular stakeholder in each alternative and to compare between them some indicators are introduced, they are the quantitative representation of their interests.

2.1.1. WATER SUPPLY INDICATORS

- **Reliability ($\boldsymbol{Irel}$):** number of days in which irrigation water demand is satisfied divided by total number of days.
- **Vulnerability1 ($\boldsymbol{Ivul}$):** probable failure of the demand supply.
- **Vulnerability2 ($\boldsymbol{Idef}$):** daily water deficit average squared to quantify water scarcity risk.
- **Resilience ($\boldsymbol{Ires}$):** probability that if the system is in an unsatisfactory state (water supply not provided) the next state will be satisfied.

2.1.2. FLOODING RISK INDICATORS

- **Flooding Risk 1 ($\boldsymbol{IF1}$):** The average number of days per year when flood is recorded.
- **Flooding Risk 1 ($\boldsymbol{IF2}$):** mean flooded area.

For Alternative Zero (A0), none of the flooding risk indicators was obviously calculated, and for the other alternatives only $IF1$ was calculated as the function of flooded surface of the river is not available.

2.1.3. ENVIRONMENTAL IMPACT INDICATORS

- **Environmental Indicator 1 ($IE1$):** Water Scarcity Impact (Less Than 25° Percentile), Low Pulses (LP) on River's Biologic Processes.
- **Environmental indicator 2 ($IE2$):** Impact of water overflow (higher than 75° percentile), High Pulses (HP) in the given habitat by the change of solid transport and nutrients.

It should be mentioned only the values of reliability ($Irel$) and resilience ($Ires$) must be maximized in order to satisfy the stakeholder. The rest of the indicators must be minimized.

2.2. ALTERNATIVE ZERO (A0) – RIVER NATURAL FLOW

The A0 scenario assumes that there is no action on the system – only actual conditions of Danube River. Table 2 shows the values of the mentioned indicators to quantify the current state (A0), subsequently used as a base point for comparing the obtained alternatives of this study.

ALTERNATIVE/ INDICATORS	**WATER SUPPLY**				**FLOODING RISK**	**ENVIRONMENTAL IMPACTS**	
	Irel	*Ivul*	*Idef*	*Ires*	*IF1*	*IE1*	*IE2*
A0	0.55	497	1.47*10^5	0.055		91	91

Table 2: Performance of A0

The values of these indicators provide a picture of the situation that it is beneficial to construct a dam. For example, it is observed that in average, the water demand is not satisfied in 45% time of the year ($\boldsymbol{Irel}$). With proper sizing and adjustment of a dam, it is known a priori that it is possible to improve it significantly, which relates to the improvement of stakeholder's satisfaction.

2.3. DAM SIZING METHOD

The aim of constructing a dam is reducing the water deficit during the "dry period" which is the interest of the farmers although the level of flood risk should be kept in a reasonable level to protect the safety of residents in downstream areas of the Danube River. First, the daily data were transformed into monthly data and plotted over the 27 years to evidence mean flow, then the water demand (w) of people in the downstream areas should be assumed as the 45th percentile of the inflow dataset. The following figure shows the monthly flow rate of the given data for this river along with the level of water demand.

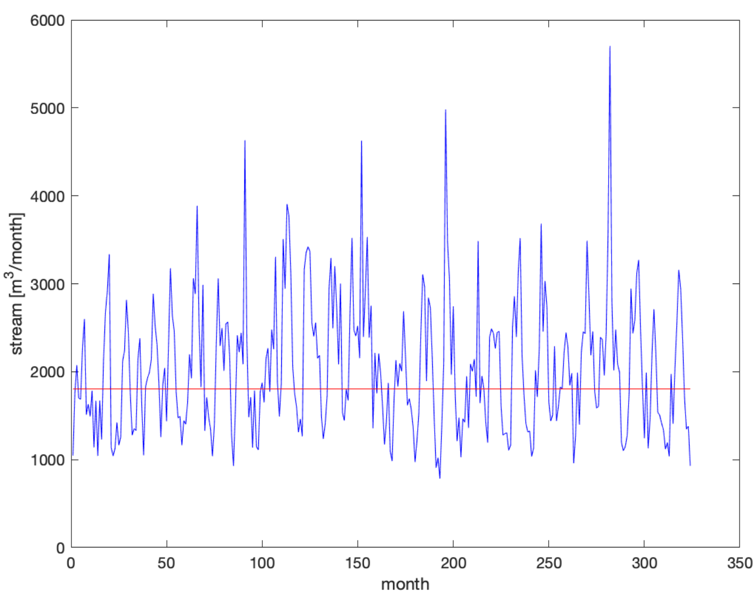


Figure 17: monthly flowrate with considered water demand

By using a sequent peak analysis, the maximum capacity of the dam was found, and suitable dam height was proposed. For the sake of simplicity in calculations, the cylindric surface shape for the river assumed. The main data regarding the dam are reported in the Table 3.

Demand (m^3/s)	**Capacity (m^3)**	**Maximum Height (m)**	**Surface (km^2)**
1800	9.87*10^9	0.0741	133260

Table 3: Characteristics of the Dam

In order to draw the maximum dam release curve (assumption of a linear shape), the height of the dam was considered as *hmax* and 99th percentile of the inflow dataset was chosen for the *rmax*, consequently, the value of the angle, m, is calculated by dividing maximum release by maximum height. The maximum release curve is shown below which is followed by values for maximum release showing in Table 4.

Doing this analysis led to the same result as using the ripple method. The height of the desired dam is really low, 74 mm. This data seems to be strange, but compared to the huge surface available it's more clear, the Danube is a really long river and we are using the cylindric assumption. Anyway this "dam" should be seen as an increase of the height of the existing one, more than a completely new one. Another considerations is about the resolution of the policy. Since the height is really low, changes in the order of millimetres in the policy parameters don't make a lot of sense. Despite this concerns, a mathematical analysis can be run anyway, so we have tried to see how the behaviour of the river changes.

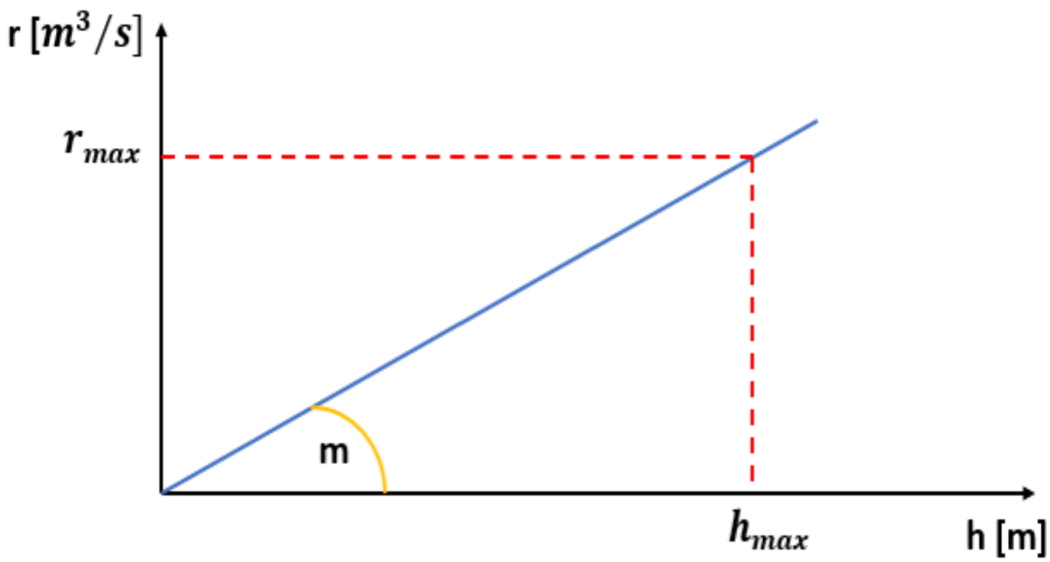


Figure 18: Maximum Release Curve

Maximum Height (h_{max}) (m)	Minimum Height (h_{min}) (m)	Maximum Release (r_{max}) (m^3/s)	m

0.0741	0.02	4857	65806

Table 4: Dam Dimensions

2.4 STANDARD OPERATING POLICY (MANUAL SELECTION OF PARAMETERS)

Regulation Operating Policy is defined by some parameters which are shown in Figure 19.

- $\boldsymbol{h_{max}}$: level above which spillways are activated (set equal to 0.074 m in each alternative)
- $\boldsymbol{h_{min}}$: minimum reservoir level for which the flow can be delivered (set equal to 0.02 m in each alternative)
- $\boldsymbol{h_1}$, $\boldsymbol{h_2}$: (normative constrains) minimum and maximum level for which exact water demand (w) is delivered
- $\boldsymbol{m_1}$, $\boldsymbol{m_2}$:: slopes of the lines of release for lower and higher levels to h_1, h_2

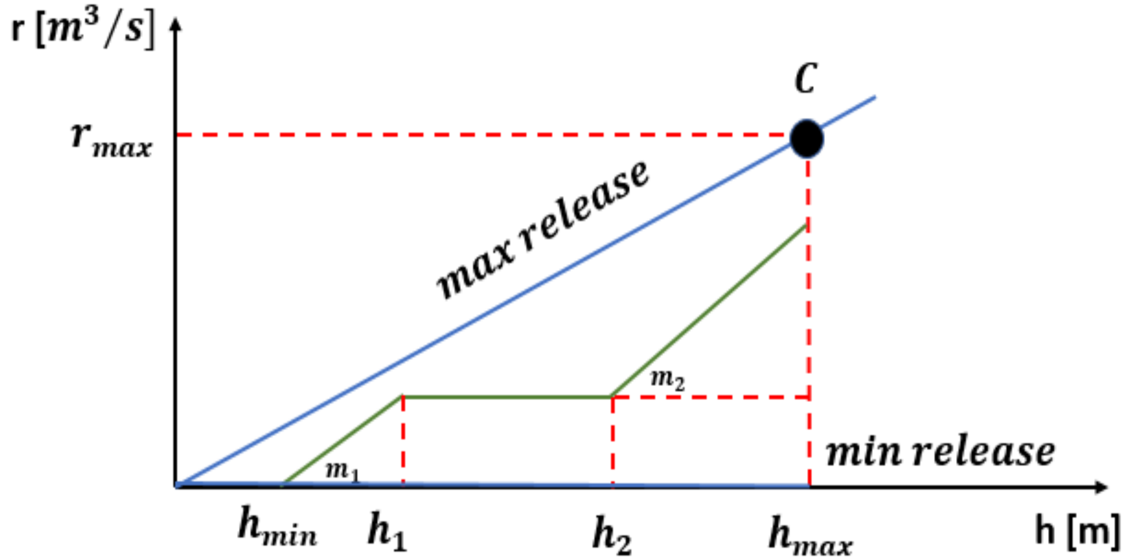


Figure 19: Dam's Operating Policy Curve

The alternatives presented below were obtained by identifying 3 different policies by changing the above regulated parameters. Two of them are focused on the satisfaction of a specific stakeholder (A1, A2) and the last one a compromise between the parties (A3).

The value of the demand was fixed for all alternatives w=1800 $m3s⁄$ and the initial level of the water reservoir as nearly the half of the dam's height $hinit$ = 0.037 m.

2.4.1. ALTERNATIVE 1 – DECREASING THE WATER SUPPLY DEFICIT IMPACT

The A1 has been set in such a way to mitigate the impact of the water deficit. The control law should be set to maximize the period that demand is satisfied. The horizontal line set by demand's value (w) should be as long as possible, similarly $m1$ *and* $m2$ as large as possible.

Table 5 shows the parameters of the control law for A1, and the standard operation policy graph for this alternative shows in subsequent figure.

h_1 (m)	h_2 (m)	m_1	m_2
0.0274	0.0667	131610	400000

Table 5: Operating Policy Parameters for Alternative 1

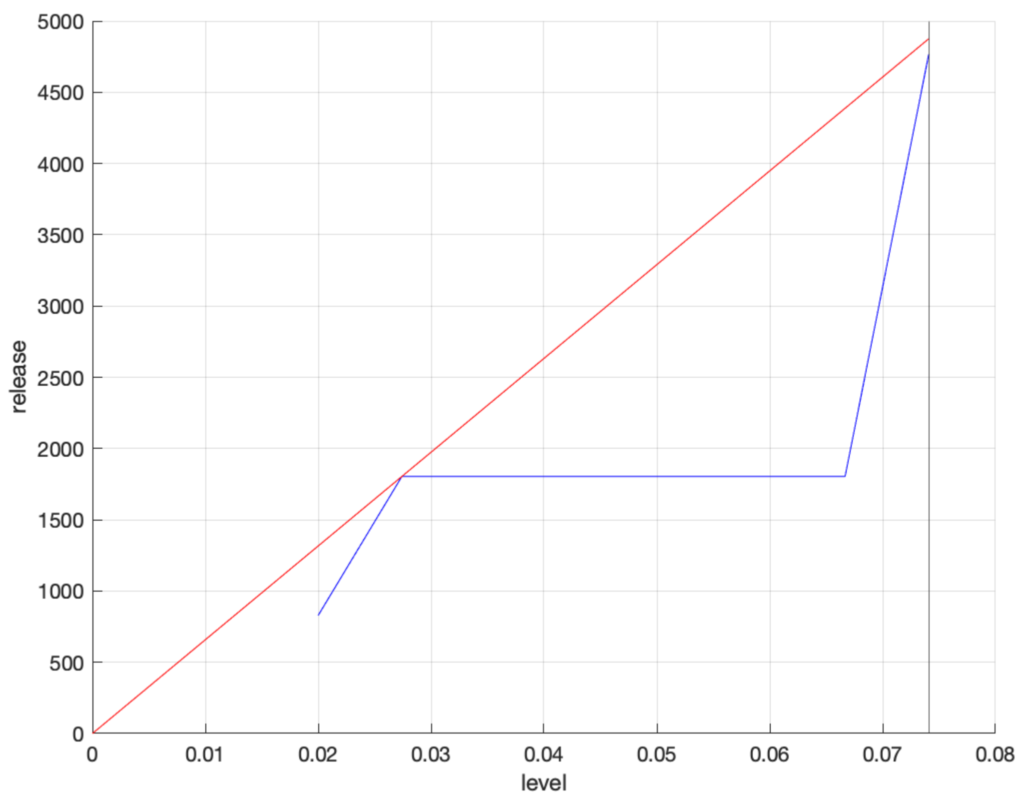


Figure 20: Standard operating policy graph for A1 i.e Release (m^3) vs Reservoir Level (m).

The following table shows the calculated indicators for A1 and compares them with A0.

ALTERNATIVE/ INDICATORS	**WATER SUPPLY**				**FLOODING RISK**	**ENVIRONMENTAL IMPACTS**	
	Irel	*Ivul*	*Idef*	*Ires*	*IF1*	*IE1*	*IE2*
A0	0.55	497	1.47*10^5	0.055		91	91
A1	0.89	411	2.44*10^4	0.024	4	16	58

Table 6: Alternatives' Comparison

It is well-observed that water supply indicators are better than A0 values. The deficit days increase considerably ($Irel$), similarly the annual average of daily deficit vulnerability ($Idef$) decreases with a order of magnitude compared to A0. Moreover, the Impact of water scarcity ($IE1$) and water overflow ($IE2$) decrease which is environmentally friendly.

2.4.2. ALTERNATIVE 2 – FLOOD RISK MITIGATION

A2 has been set in such a way to minimize the risk of flooding on the downstream of a river which is an important matter for citizens who live there, this has been done by reducing a lot the h2 value. Table 7 shows the parameters of the control law for A2, and the standard operation policy graph for the alternative are shown in the subsequent figure.

h_1 (m)	h_2 (m)	m_1	m_2
0.035	0.037	32903	250000

Table 7: Operating Policy Parameters for Alternative

Figure 21: Standard operating policy graph for A2 i.e Release (m^3) vs Reservoir Level (m).

Table 8 shows the calculated indicators for A2 and compares them with A0 and A1.

ALTERNATIVE/ INDICATORS	**WATER SUPPLY**	**FLOODING RISK**	**ENVIRONMENTAL IMPACTS**

	Irel	*Ivul*	*Idef*	*Ires*	*IF1*	*IE1*	*IE2*
A0	0.55	497	1.47*10^5	0.055		91	91
A1	0.89	411	2.44*10^4	0.024	4	16	58
A2	0.52	291	7.97*10^4	0.009	0.18	31	93

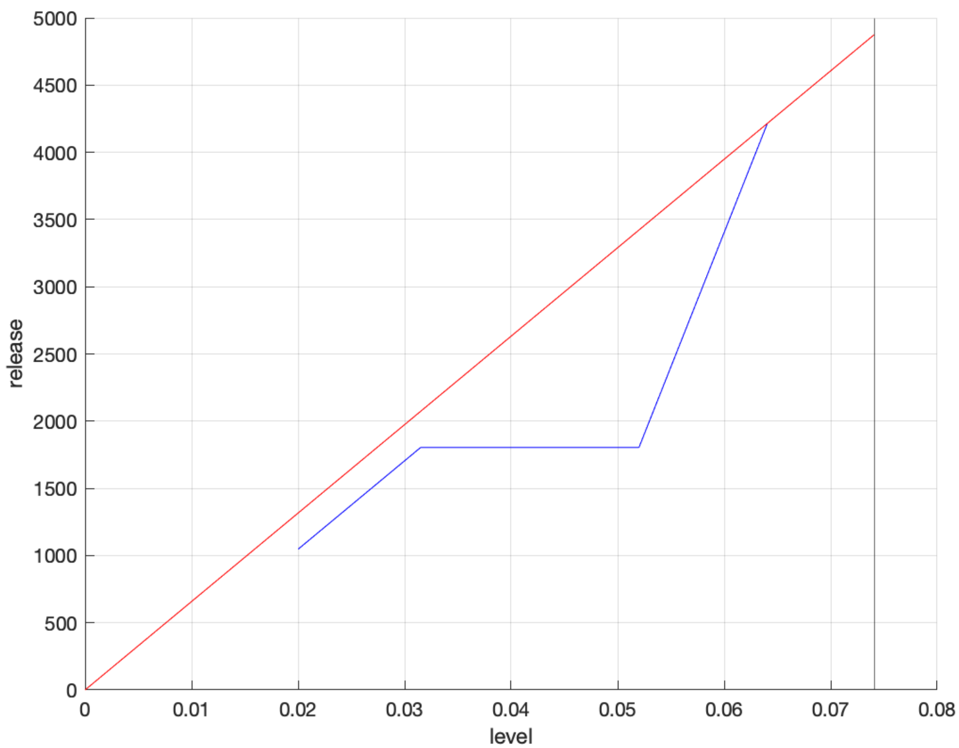


Table 8: A0, A1 & A2 Alternatives' Comparison

The idea was to release water sooner to avoid filling up the storage too much. The normative constrain $\boldsymbol{h}_2$ was set very low. It shorted the section where exactly the demand flow value is released and by setting steep gradients outside of it, the policy tends to empty the reservoir before it is almost full, but in the same time not optimally using the water volume that could be used for irrigators in the dry season. Thus, we expect the water supply indicators worsen whereas the flood risk indicator improves. The Table 8 proves this interpretation.

2.4.3. ALTERNATIVE 3 – COMPROMISE CONDITION

The last alternative analysed in this section is a mix of the previous alternatives. It has been set up to obtain good projected indicators values. The goal is obtaining the optimal solution that minimized the trade between the stakeholders. Table 9 shows the parameters of the control law for A3, and the standard operation policy graph for this alternative is shown in the subsequent figure.

h_1 (m)	h_2 (m)	m_1	m_2
0.031	0.052	65806	200000

Table 9: Operating Policy Parameters for Alternative 3

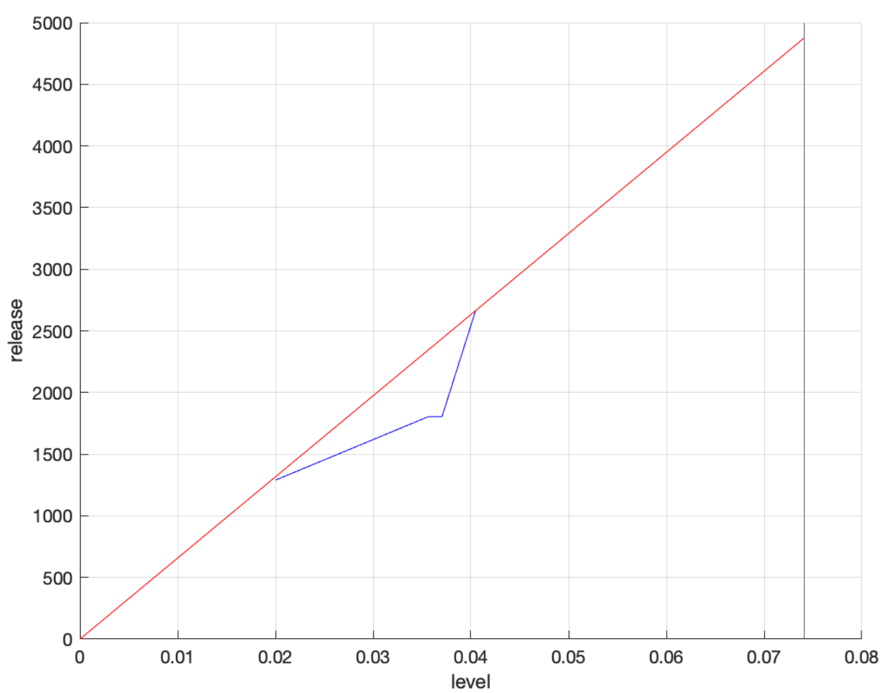


Figure 22: Standard operating policy graph for A3 i.e Release (m^3) vs Reservoir Level (m).

The following table shows the calculated indicators for A3 and compares them with the other alternatives.

ALTERNATIVE/ INDICATORS	**WATER SUPPLY**				**FLOODING RISK**	**ENVIRONMENTAL IMPACTS**	
	Irel	*Ivul*	*Idef*	*Ires*	*IF1*	*IE1*	*IE2*
A0	0.55	497	1.47*10^5	0.055		91	91
A1	0.89	411	2.44*10^4	0.024	4	16	58
A2	0.52	291	7.97*10^4	0.009	0.18	31	93
A3	0.77	362	4.39*10^4	0.014	1.4	30	67

Table 10: A0, A1, A2 & A3 Alternatives' Comparison

2.5. EMODPS (Evolutionary Multi-Objective Direct Policy Search)

In the previous section, the required parameters were chosen manually. In this part, instead of randomly choosing the parameters to define the control law, it has been decided to generate a set of optimal policies using an optimization algorithm.

The optimization was conducted minimizing two objective functions:

- Daily water deficit average squared to quantify water scarcity risk, $\boldsymbol{Idef}$
- The average number of days per year when flood is recorded, $\boldsymbol{IF1}$

EMODPS (NSGA2) has been implemented in MATLAB. After considering the limit for policy parameters ($\boldsymbol{h}_1$, $\boldsymbol{h}_2$, $\boldsymbol{m}_1$, $\boldsymbol{m}_2$), setting the number of populations on 30 and performing 25 iterations (generations) leading to obtaining something really close to the Pareto front. It includes many optimal solutions in order to give free choice to decision makers. The following figure shows the distribution of initial populations and the final approximated Pareto front in the objectives space.

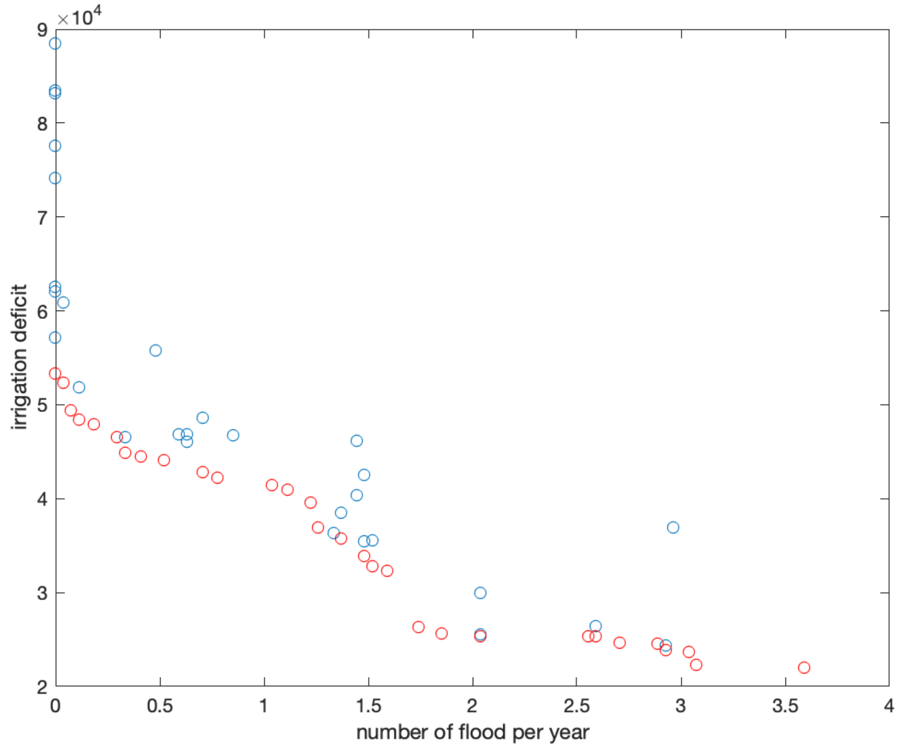


Figure 23: Initial Population and Final Pareto Front

As the same procedure of the previous section, it has been decided to evaluate the indicators for three alternatives with below characteristics, which will be compared to the Alternative zero (no dam construction).

1. Alternative one (A1): minimizing the irrigation water deficit

2. Alternative two (A2): minimizing the number of the flooding days

3. Alternative three (A3): the compromise between A1 and A2

The parameters found for each alternative are reported in the following table.

ALTERNATIVES	h_1 (m)	h_2 (m)	m_1	m_2
A1_EMODPS	0.033	0.067	58021	281640
A2_EMODPS	0.031	0.047	70176	307270
A3_EMODPS	0.034	0.064	50000	381440

Table 11: Alternatives' Parameter Values

2.5.1. COMPARISON OF THE INDICATORS BETWEEN MANUALLY-SELECTED AND EMODPS PARAMETERS

ALTERNATIVE/ INDICATORS	**WATER SUPPLY**				**FLOODING RISK**	**ENVIRONMENTAL IMPACTS**	
	Irel	*Ivul*	*Idef*	*Ires*	*IF1*	*IE1*	*IE2*
A0	0.55	497	1.47*10^5	0.055		91	91
A1	0.89	411	2.44*10^4	0.024	4	16	58
A2	0.52	291	7.97*10^4	0.009	0.18	31	93
A3	0.77	362	4.39*10^4	0.014	1.4	30	67
A1_EMODPS	0.85	317	2.19*10^4	0.017	5.44	14	60
A2_EMODPS	0.72	374	5.33*10^4	0.013	0.44	36	74
A3_EMODPS	0.82	308	2.63*10^4	0.012	2.77	16	61

Table 12: A0, A1, A2 & A3 Alternatives' Comparison

Surprisingly the manual selected policy has more extreme results, in fact they minimise some parameters but increase a lot other, this is because of the nature of the selection was very extreme. The algorithm instead finds some policy more "smooth".

The changes in the parameters between every test is resorted to third - fourth decimal value and so it doesn't make a lot of sense switching from one policy to another, they are basically the same, even if they change a lot the indicators. Here you can see the different policies solution of the approximated Pareto front. Basically the algorithm only tries to change the parameter h2. This can be due to a too much narrow variability range for the other parameters. In fact, we allowed h1 from 0.027 to only 0.036 meters, instead h2 has a much wider range from 0.037 to 0.067.

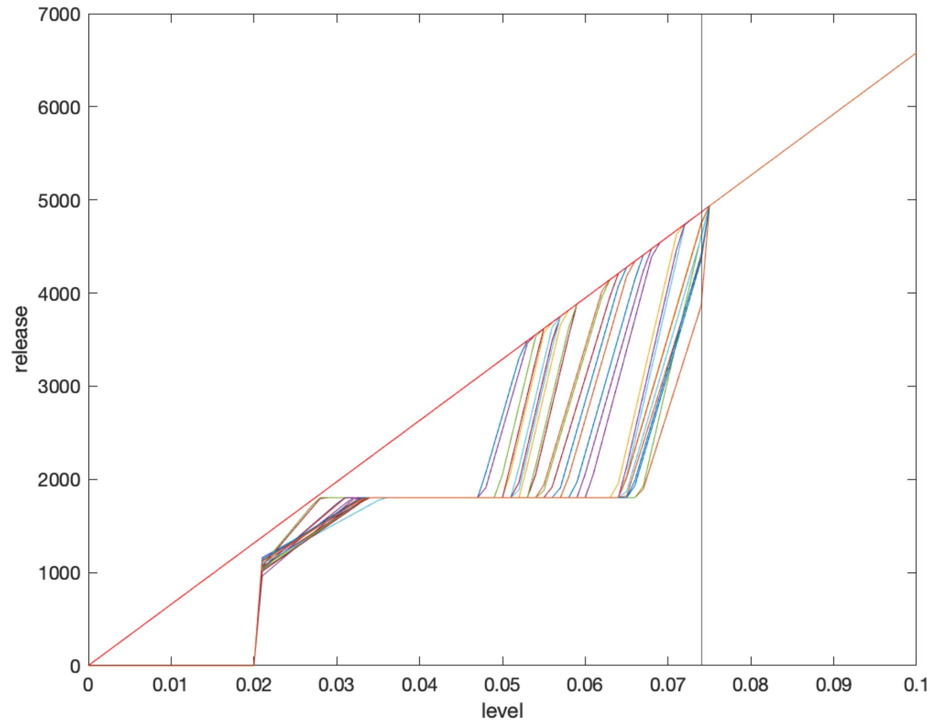


Figure 22: different policies at the last iteration of the EMODPS algorithm

2.6. DISCUSSION AND CONCLUSION

Following the analysis, building a dam at Gabckikovo on the Danube River would bring improvements to the system. For considering the interest of different stakeholders, the compromise alternatives are proposed in this study which considering the trade-off between various objectives. It is also important to note that the resulting compromise alternative will be different if one adds other interests in the analysis, for example also the population water supply or the industrial ones. To include them, they should be represented by appropriate indicators and their demand should be considered in the estimate of the needed water supply. First, the policy created by manual selection of the policy parameters then use EMODPS method to generate a set of optimal policies using an optimization algorithm. Based on the comparison between two methods, the latter optimize the policy better and it seems more efficient.

Despite the data and the mathematical behaviour of this model some considerations has to be done about the Gabcikovo dam anyway. The Danube is a really long river and feed a lot of countries along its path so any change and construction has a huge impact on all the nations and each decision has to be taken accordingly to the other countries. In this case the change of the operating policy, letting the dam filling up a little bit more has a great impact on the river, especially for upstream countries. A lot of concerns and discussions on the existing Gabcikovo dam has been done during the years, more precisely about its environmental impact.

References

Bertoni, F., Castelletti, A., Giuliani, M., & Reed, P. M. (2019, January). Do we understand performance dependencies, trade-offs, and robustness in dam design and operation?. In *Geophysical Research Abstracts* (Vol. 21).

Fitzmaurice, J. (2018). *Damming the Danube: Gabcikovo/Nagymaros and post-communist politics in Europe*. Routledge.

Giuliani, M., Castelletti, A., Pianosi, F., Mason, E., & Reed, P. M. (2016). Curses, tradeoffs, and scalable management: Advancing evolutionary multiobjective direct policy search to improve water reservoir operations. *Journal of Water Resources Planning and Management*, *142*(2), 04015050.