Supplementary material

Several input variables were considered and tested to replicate past values of S.Giustina water stored and outflows. Beyond the three variables reported in Table 1 (inflow to, outflows from and volume in S.Giustina), outflows from an upstream dam reservoir (i.e. Careser), temperature, precipitation and the national single energy price (PUN) were initially considered and tested as predictors. However, their low predictive importance and limited temporal overlap with the response variables affected a robust simulation and validation of historical values of both S. Giustina volume and outflows for hydropower use.

Table S1 – Summary table with all tested variables, unit, temporal coverage and their data source. * Data for temperature and precipitation considered the closest weather station (#T0236) to the S.Giustina reservoir.

#	Variable	Variable name	Unit	Temporal coverage	Source and link
1	Simulated inflows to S.Giustina [m^3/s]	Inflow	[$m^3 \cdot s^{-1}$]	1981-2010	GeoTransf hydrological model (Bellin et al., 2016)
2	S.Giustina outflows for hydropower	Outflow	[$m^3 \cdot s^{-1}$]	1981-2017	Province of Trento – Agency for water resource and energy
3	S.Giustina volume	Volume	[Mm^3]	1999-2004 2009-2017	
4	Upstream dam reservoir (Careser) outflows for hydropower	Outflow Careser	[$m^3 \cdot s^{-1}$]	1990-2013	
5	Temperature*	Temp	[°C]	1986-2017	Province of Trento – Weather service https://www.meteotrentino.it/index.html#!/content?menuItemDesktop=111
6	Precipitation*	Prep	[mm]	1986-2017	

7	National Single energy Price (PUN)	PUN	[€·MWh^{-1}]	2004-2017	Energy market operator http://www.mercatoelettrico.org/En/download/DatiStorici.aspx

The combination of all the variables was not considered due to the limited temporal length of the final dataset (i.e. 2004, 2009, 2010) and hence its limited representativeness for the response variables.

For the prediction of the variable Outflow, Inflow, Temp and Volume were the most correlated variables. However, given the correlation between Inflow and Temp and the principle of selecting the most parsimonious models, the variable Inflow was selected and Temp not further considered for predicting Outflow values. A similar process was followed for the prediction of Volume values, given the correlation between Outflow Careser and Month and the higher influence of the variable Month on the prediction of Volume.

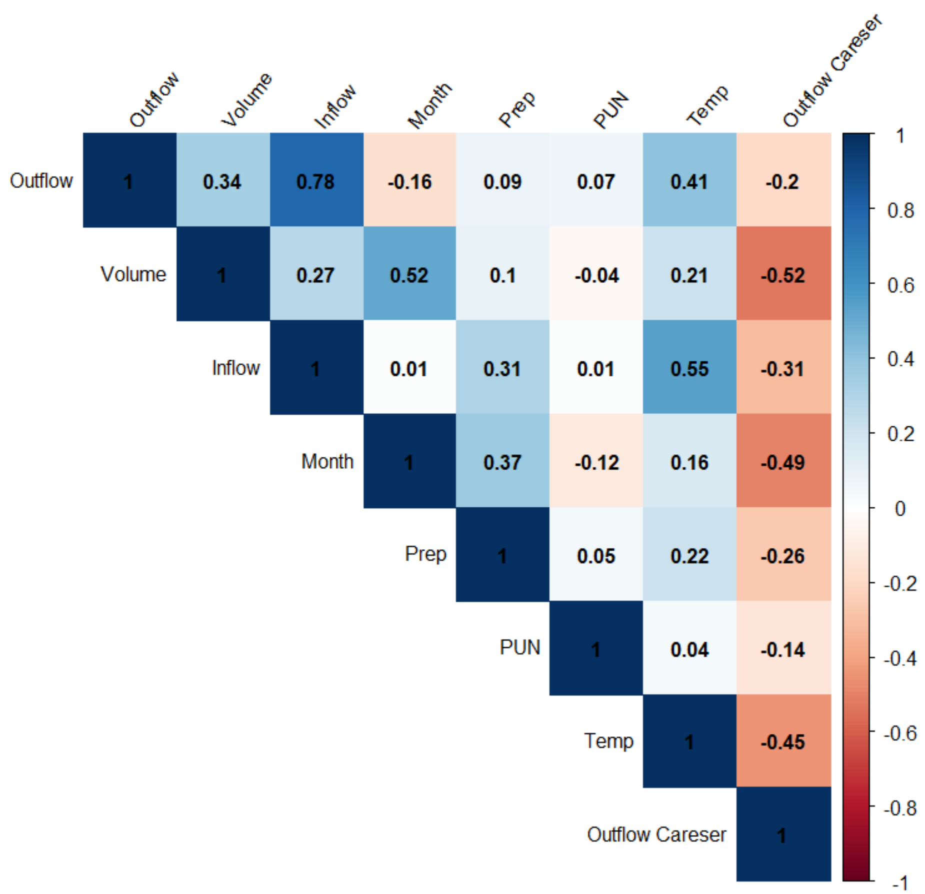


Figure S1 – Correlation matrices displaying positive correlations in blue and negative correlations in red color. Color intensity is proportional to the correlation coefficients reported within each square.

Here a summary of the best models implemented and tested is reported. Starting from the variables shown in Table 1, a stepwise procedure for model testing was implemented for each model type. Moreover, a moving window approach for the assessment of model performance indicators (i.e. R-squared and RMSE) allowed to evaluate the model performances considering an increasing set of training and testing datasets.

Table S2 - Summary table with best developed and tested models for each model type for predicting water volume stored in the S.Giustina reservoir. Final models chosen for the assessment are reported in Table2.

Models type	#	Formulas	Adjusted-R^2	RMSE ($\cdot 10^6$)
Multi-linear model	1.	lm (Volume ~ Geotransf_inflows)	0.13	21.03
	2.	lm (Volume ~ lag(Geotransf_inflows))	0.35	18.37
	3.	lm (Volume ~ Geotransf_inflows + Outflow)	0.17	20.78
	4.	lm (Volume ~ lag(Geotransf_inflows) + Outflow)	0.35	18.30
	5.	lm (Volume ~ lag(Geotransf_inflows) + lag(Outflow))	0.36	18.20
Linear mixed effect model	6.	lmer (Volume ~ Geotransf_inflows + (1|month))	0.68	12.12
	7.	lmer (Volume ~ lag(Geotransf_inflows)+ (1|month))	0.71	13.59
	8.	lmer (Volume ~ Geotransf_inflows + Outflow + (1|month))	0.67	14.22
Generalized additive model	9.	gam(Volume ~ s(Geotransf_inflows))	0.12	28.15
	10.	gam(Volume ~ s(lag(Geotransf_inflows))	0.37	23.84
	11.	gam(Volume ~ s(Geotransf_inflows) + s(Outflow))	0.15	27.40
	12.	gam(Volume ~ s(lag(Geotransf_inflows))+ s(Outflow))	0.36	23.60
	13.	gam(Volume ~ s(lag(Geotransf_inflows)) + s(lag(Outflow))	0.44	22.13
Generalized additive mixed model	14.	gam(Volume ~ s(Geotransf_inflows) + s(mo, bs="re"))	0.31	24.83
	15.	gam(Volume ~ s(lag(Geotransf_inflows)) +s(mo, bs="re"))	0.45	22.22
	16.	gam(Volume ~ s(Geotransf_inflows) + s(Outflow) + s(mo, bs="re"))	0.45	24.89

	18.	gam(Volume ~ s(lag(Geotransf_inflows)) + s(lag(Outflow)) + s(mo, bs="re"))	0.50	20.88

Table S3 - Summary table with best developed and tested models for each model type for predicting turbined water outflow from the S.Giustina reservoir. Final models chosen for the assessment are reported in Table2.

Models type	**#**	**Formulas**	**Adjusted-R^2**	**RMSE ($\cdot 10^6$)**
Multi-linear model	1.	lm(Outflow ~ Geotransf_inflows)	0.61	15.55
	2.	lm (Outflow ~ lag(Geotransf_inflows))	0.56	17.90
	3.	lm (Outflow ~ Geotransf_inflows + Volume)	0.64	15.30
	4.	lm (Outflow ~ lag(Geotransf_inflows) + Volume)	0.56	17.85
	5.	lm (Outflow ~ Geotransf_inflows + lag(Volume))	0.72	13.09
	6.	lm (Outflow ~ lag(Geotransf_inflows) + lag(Volume))	0.55	18.05
Linear mixed effect model	7.	lmer (Outflow ~ Geotransf_inflows + (1|month))	0.72	13.24
	8.	lmer (Outflow ~ lag(Geotransf_inflows) + (1|month))	0.53	17.80
	9.	lmer (Outflow ~ Geotransf_inflows + Volume + (1|month))	0.73	13.20
	10.	lmer (Outflow ~ Geotransf_inflows + lag(Volume) + (1|month))	0.74	12.35
Generalized additive model	11.	gam(Outflow ~ s(Geotransf_inflows) + s(mo))	0.55	21.47
	12.	gam(Outflow ~ s(lag(Geotransf_inflows)) + s(mo))	0.51	22.51
	13.	gam(Outflow ~ s(Geotransf_inflows) +s(Volume) + s(mo))	0.62	19.31
	14.	gam(Outflow ~ s(lag(Geotransf_inflows)) + s(Volume) + s(mo))	0.61	19.83
	15.	gam(Outflow ~ s(Geotransf_inflows) + s(lag(Volume) + s(mo))	0.72	16.86
	16.	gam(Outflow ~ s(Geotransf_inflows) +s(mo, bs="re"))	0.55	21.47
	17.	gam(Outflow ~ s(lag(Geotransf_inflows)) +s(mo, bs="re"))	0.51	22.51

Generalized additive mixed model	18. gam(Outflow ~ s(Geotransf_inflows) + s(Volume) +s(mo, bs="re"))	0.63	19.31
	19. gam(Outflow ~ s(Geotransf_inflows) + s(lag(Volume)) +s(mo, bs="re"))	0.72	16.86

Moreover, here reported in Figure S2 and S3 the cumulative values for modelled and real values for both volume and outflows during the considered time period 1999-2004 and 2009-2016.

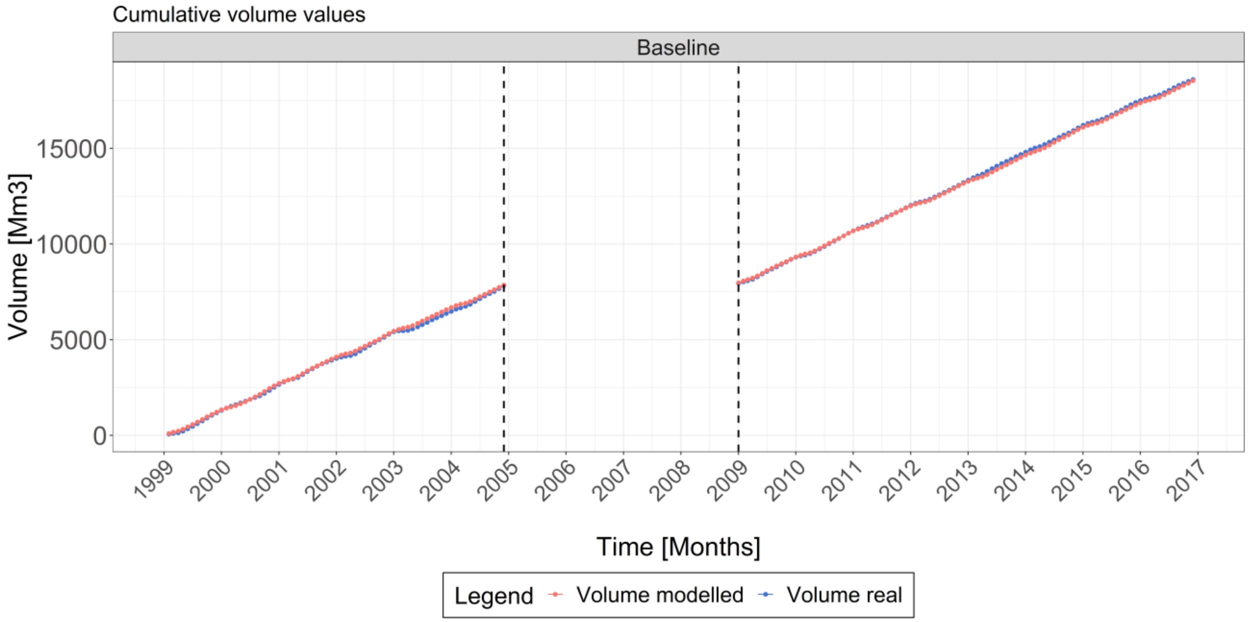
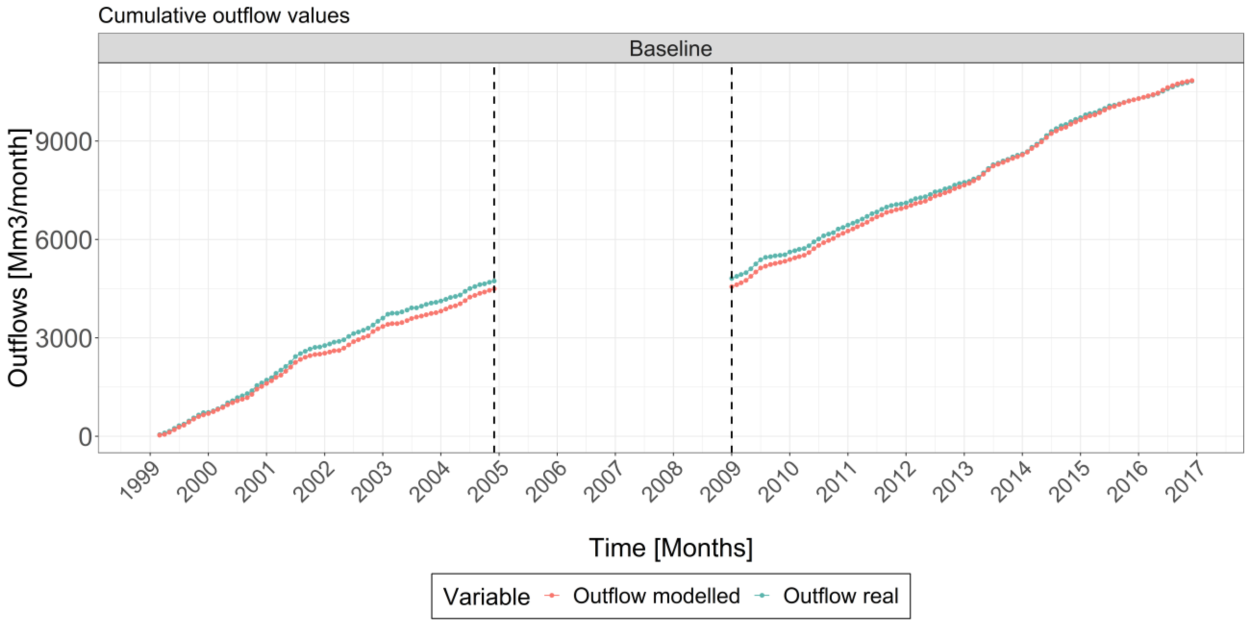


Figure S2 – Cumulative plot for modelled and real volume values, with a final difference between modelled and real volume of 61.47 Mm3. The dotted lines define the start and end of the data gap.



35 **Figure S3 – Cumulative plot for modelled and real volume values, with a final difference between modelled and real outflow of 16.28 Mm3/month. The dotted lines define the start and end of the data gap.**