Estimating regional potential for micro-hydropower energy recovery in irrigation networks on a large geographical scale

Miguel Crespo Chacón [a, *], Juan Antonio Rodríguez Díaz [b], Jorge García Morillo [b], Aonghus McNabola [a]

[a] Department of Civil, Structural and Environmental Engineering, Trinity College Dublin, Dublin, D02 PN40, Ireland
[b] Department of Agronomy, University of Córdoba, International Campus of Excellence ceiA3, 14071, Córdoba, Spain

A B S T R A C T

Micro-hydropower has been highlighted as a potential technology suitable for installation in irrigation networks to reduce system overpressures and to reduce the net energy consumption of the irrigation process. However, the full impact of this technology on a large regional scale is unknown. Artificial Neural Networks and regression models were used in this research to predict the energy recovery potential for micro-hydropower in on-demand pressurised irrigation networks across a large spatial scale. Predictors of energy recovery potential across spatial unit areas included: Irrigated land surface area, irrigation crop water requirements, rainfall, evapotranspiration, and mean topographical slope. The model was used to predict the energy recovery potential across the 164,000 ha of the Spanish provinces of Seville and Cordoba in the absence of hydraulic models. A total of 21.05 GWh was identified as the energy potential which could have been recovered using micro-hydropower during the 2018 irrigation season. This amount of energy would have potentially reduced the energy consumption of the irrigation process in this region by approximately 12.8%. A reduction in energy consumption in the agriculture sector of this magnitude could have significant impacts on food production and climate change. The main novelty of this paper lies in the assessment of micro hydropower resources in operating irrigation networks on a large geographical scale, in areas where no information is available. It provides an approximation of the existing potential using computational methods.

© 2020 Elsevier Ltd. All rights reserved.

1. Introduction

The water industry became one of the most energy intensive sectors during the last decades, consuming 4% of global electricity in 2014 for the extraction, distribution and treatment processes. This consumption is projected to more than the double, with more than 100 million tonnes of oil equivalent of thermal energy in 2040 [1]. Focusing on irrigation, some research carried out in Southern Spain studied how the water costs have changed due to energy consumption, when the irrigation districts started using pressurised pipelines systems. The analysis of numerous pressurised irrigation districts concluded that the energy cost comprised around

40% of the total water costs on average, reaching peaks of 65% in some cases [2,3]. To counteract these high percentages represented by energy, previous investigations have focused on the analysis of different measures to reduce energy costs and energy dependency. Different investigations reported that important energy savings could be achieved applying irrigation scheduling and improvements in the configuration of the hydraulic network with savings varying from 3.5% to 36.4% [4—8]. Renewable energy was also applied in other research to improve the energy efficiency in irrigation systems, developing smart solar irrigation management systems able to fulfil the water requirements during the entire irrigation season, thus avoiding 100% of the CO_2 emissions [9].

During the last years, several investigations have also studied micro hydropower (MHP) as a solution for offsetting this growing trend of energy dependency in the water industry. This technology provides a potential solution in water networks by taking the advantage of existing system overpressures. Within drinking water supply networks, Corcoran et al. [10] assessed the energy recovery

---

* Corresponding author. Engineering School Office, First Floor, Museum Building, Trinity College, Dublin 2 (D02), Dublin, Ireland.
E-mail addresses: crespocm@tcd.ie (M. Crespo Chacón), jarodriguez@uco.es (J.A. Rodríguez Díaz), jgmorillo@uco.es (J. García Morillo), amcnabol@tcd.ie (A. McNabola).

Nomenclature

ADAM	Adaptive Moment Estimation
ANN	Artificial Neural Network
GHG	Greenhouse Gases
MAE	Mean Absolute Error
MHP	Micro Hydropower
MSE	Mean Squared Error
PAT	Pump-as-Turbine
R^2	Coefficient of Determination
ReLU	Rectified Linear Unit
VOS	Variable Operating Strategy

potential of applying MHP at 95 sites in Ireland and the UK to reduce overpressures and produce electricity. A yearly energy of 6.75 GWh could be recovered, just considering the 12 highest power output sites assessed in Dublin city, during 2011. Power et al. [11] evaluated the existing potential in the wastewater sector, analysing data from 100 wastewater treatment plants, reporting a total hydropower potential of 1.75 GWh in just 14 of them, which was said to be equivalent to the energy demand of 350 households in Ireland, avoiding over 900 t eCO$_2$. Gallagher et al. [12] estimated an annual energy recovery potential of 20.1 GWh in 238 sites in water and wastewater networks in Ireland and the UK, which would be capable of supplying energy to 4702 households in Ireland and Wales. However, these studies were limited to drinking water and assessing the potential from measured flow and pressure data in networks with existing hydraulics models in some cases. These investigations also covered just a small part of the existing infrastructure in those locations. Different studies also assessed the potential of MHP as measure to improve the energy efficiency in pressurised irrigation networks. These reported energy potential which varied between 2.12 MWh year^{-1} and 281 MWh year^{-1} in different individual case studies with irrigated surface areas from 68 to 2691 ha [13—18]. Nonetheless, each of them used different approaches to quantify the potential, and different input flow and head data. These encompassed predicted or recorded annual mean values, or predicted or annual recorded values. This fact makes joining all these results in one study to assess the large-scale potential, significantly complex.

As previously mentioned, none of the previous investigations examined this impact beyond a single case study or on a large regional scale. The difficulty in obtaining the detailed water network information required for such investigations is a major barrier to conducting large-scale assessments. Network information on pipe size, layout, water demands and pressure, are often absent, not recorded, or not publicly available. As such, an alternative approach to MHP potential prediction is required using proxy measures of key variables. Mitrovic et al. [19] analysed the linear correlation between MHP power potential and different proxy variables, such as population, population density and land topography, as predictors of water demand and system overpressure, key variables in MHP potential for drinking water networks. The 238 sites studied by Gallagher et al. [12], were analysed by Mitrovic et al. [16] with a view to predicting large scale energy potential in the absence of network data in Ireland. The results showed a coefficient of determination (R^2) of 0.26 for the population as a proxy measure of MHP potential, and in general failed to offer a reliable prediction of MHP potential for the drinking water sector using this method. Nevertheless, these works were focussed on drinking water and there are not previous investigations for the irrigation sector.

This paper aims to develop a model to predict the energy recovery potential in pressurised irrigation networks on a large-scale using proxy indicators of irrigation demand and network pressure. Its main novelty falls on being the first paper exploring the large geographical scale energy recovery prediction in pressurised irrigation networks using hydropower, providing an approximation of the existing resources for such technology. Furthermore, two approaches were assessed: Single linear regression models, as well as non-linear analysis through artificial neural networks (ANNs), were evaluated. These models were evaluated using a database with 177 observations obtained from the detailed hydraulic model of 18 irrigation networks employing data from the 2018 irrigation season. Three different variables were utilised to evaluate their relationship with the energy recovery potential, measuring distinct statistical metrics in each model. ANNs provided the best results and were finally applied to prediction of large-scale the energy recovery potential. The prediction was conducted for every municipality forming the provinces of Seville and Cordoba, in Southern Spain. The potential in more than 160,000 ha of irrigated surface was evaluated, assessing the economic and environmental benefits.

2. Materials and methods

2.1. Study area

The observations required to develop and test a prediction model were obtained from 18 pressurised on-demand irrigation networks, most of them located within the provinces of Cordoba and Seville, in Southern Spain. Two of the networks were out of this region, one located in Southern Portugal and the other in South Western Spain (see Fig. 1). The annual energy recovery potential was calculated for these networks for the 2018 irrigation season, using the methodology developed and validated by Crespo Chacon et al. [17,18]. The aforementioned methodology aimed to predict the flow distribution along the irrigation season, assessing every possible flow value predicted as a best efficiency flow for different theoretical hydropower turbines. The methodology in particular relies on the use of pump-as-turbines (PATs), conventional pumps operated in reverse as turbines, which have been shown to be suited to the micro scale applications present in irrigation networks [16—21], and also be to be economically viable in this setting due to their low-cost nature [20]. The Crespo Chacon et al. [17,18] methodology enables the selection of the PAT that returned the minimum payback period from all possible best efficiency flows within the analysed network. It used a simplified variable operating strategy (VOS) [21,22], which considered the whole flow and head distribution, simulating the theoretical behaviour of the machine for these values.

The 18 networks irrigated a total surface of 36,536 ha, where a wide distribution of crops were cultivated. The infrastructure was either gravity fed or supplied through direct pumping, depending on the network. The service pressure required at hydrant level in every case was 35 m. The irrigation networks worked as 18 independent hydraulic infrastructures, corresponding to nine different irrigation districts. Eight of the networks belonged to nine different districts, while the other ten were different sectors within the same district. The different districts analysed were: Genil Margen Izquierda (GMI), Bembézar Margen Izquierda (BMI), Bembézar Margen Derecha (BMD), El Villar (EV), Genil-Cabra (GC), Guadalmellato (GU), Fuente Palmera (FP), Aboro (AB) and Zújar (ZJ). A summary of each irrigation district, their crops and characteristics, can be seen in Table 1. In addition, all the networks analysed were fed by surface water coming from different infrastructures (rivers or irrigation channels), from which the water was pumped either to a reservoir, if the network was gravity fed, or straight pumped into
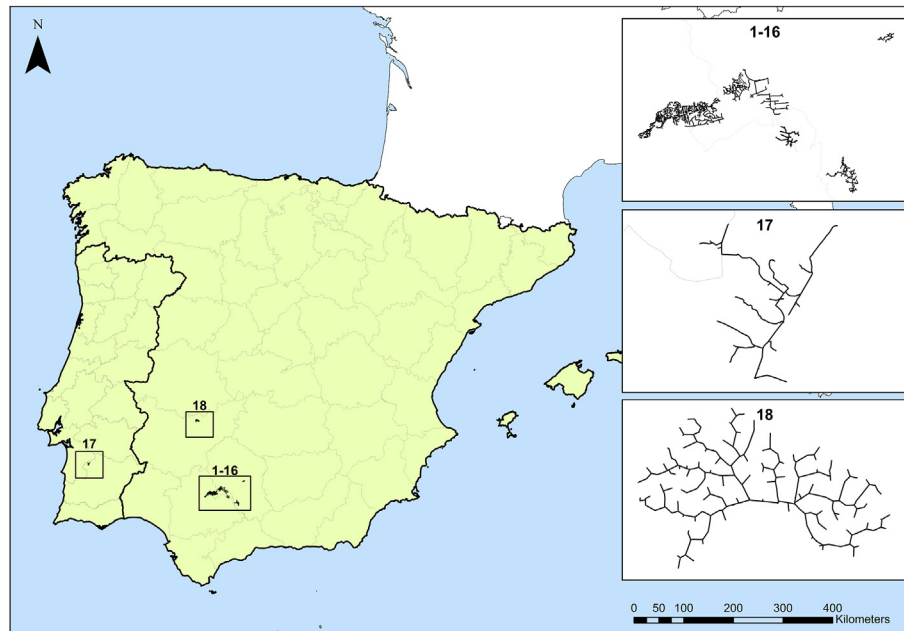
Fig. 1. Location and summary of the networks analysed to obtain the observations.

Table 1
Summary of the main properties of the irrigation districts assessed.

District	Networks Analysed	Irrigated Surface (ha)	Dominant Crops	Feeding system	Country
Genil Margen Izquierda	1	4450	Citrus, Almond, Olive, Walnuts	Gravity	Spain
Bembézar Margen Izquierda	1	3900	Citrus, Maize, Olive, Sunflower	Pumping	Spain
Bembézar Margen Derecha	10	11,163	Citrus, Maize, Cotton, Sunflower	Pumping	Spain
El Villar	1	2726	Cereals, Cotton	Pumping	Spain
Genil-Cabra	1	4320	Cotton, Sunflower, Wheat	Pumping	Spain
Guadalmellato	1	475	Maize, Cotton, Sunflower, Wheat	Pumping	Spain
Fuente Palmera	1	5611	Cotton, Sunflower, Wheat	Pumping	Spain
Aboro	1	1200	Olive, Maize, Almond	Gravity	Portugal
Zujar	1	2691	Tomatoes, Maize, Vine, Fruit, Rice	Pumping	Spain

the network. The networks were designed for a high demand, 1–1.2 l s^{-1} ha^{-1} on demand (24 h per day) and simultaneity of 100%, which means that all the hydrants could be open at the same time. Lastly, the dominant irrigation system in all the networks was drip irrigation.

2.2. Potential MHP locations

Applying the Crespo Chacon et al. [17,18] methodology to the 18 irrigation networks resulted in the identification of 177 specific locations where the installation of a PAT was economically viable. The irrigated surface encompassed by the 177 potential points for micro-hydropower energy recovery was 27,417 ha, accounting for an energy recovery potential of 6.11 GWh. Table 2 shows a summary of the results obtained for each independent irrigation network, showing among others the number of viable points for MHP application, total energy potential or percentage of surface where potential was found. Relating the irrigated surface, where energy recovery potential was found, with the total irrigated surface analysed in the 18 networks, resulted in a ratio of 0.75. However, looking individually to each network analysed, the average value of this factor decreased up to 0.70. This factor showed the portion of irrigated surface where MHP potential was found with a payback period less than 10 years. The mean power found per

observation was 12.6 kW, with minimum and maximum power of 1.6 kW and maximum of 62.3 kW respectively. With respect to energy, the average amount found per location was 29.1 MWh, with minimum and maximum values of 3.8 MWh and 214.7 MWh respectively.

These 177 potential MHP installations were used as the basis for the assessment of the large-scale prediction methodology described in the following sections. Using linear and non-linear techniques, proxy variables were used to attempt to predicted this potential energy production in the absence of specific measured irrigation network data on flow, pressure, pipe layout, pipe diameter, etc.

2.3. Proxy variables definition

Like in any prediction model, the definition of one or several explanatory variables, determined as inputs, was required to predict the output or response variable. To allow the application of this model to regions with pressurised irrigation networks, the explanatory variables were chosen considering the possibility of easily gathering these independently of the area studied. The output variable used was the MHP energy recovery potential. The explanatory variables selected in this model were selected to characterise different aspects of the area where the potential was

Table 2
Results summary obtained for the 18 irrigation networks.

Network	Points	Energy (MWh)	Surface (ha)	Surface with MHP potential	Average Power (kW)	Average Energy (MWh)
GMI	17	662	4450	62.0%	15.4	39.0
BMI	15	744	3900	88.7%	24.5	49.6
BMD-S3	4	46	631	56.8%	2.9	11.4
BMD-S4	8	98	1679	48.6%	5.5	12.3
BMD-S5	3	59	1186	47.8%	8.4	19.5
BMD-S6.1	15	452	726	92.9%	17.1	30.1
BMD-S6.2	3	107	924	92.5%	11.3	35.5
BMD-S7	5	94	922	66.3%	5.0	18.8
BMD-S8.1	4	123	1141	70.1%	9.9	30.8
BMD-S8.2	8	127	1686	53.7%	15.8	15.8
BMD-S9	5	132	1275	83.0%	7.1	26.5
BMD-S10	3	80	993	70.2%	8.9	26.7
EV	13	917	2726	94.3%	28.8	70.5
GC	34	1165	4320	88.4%	24.9	34.3
GU	1	16	475	21.1%	6.5	16.3
FP	26	934	5611	91.0%	20.0	35.9
AB	4	79	1200	74.6%	6.4	19.7
ZJ	9	281	2691	58.2%	8.1	31.2
Total/Average	**177**	**6114**	**36,536**	**70%**	**16.9**	**34.6**

analysed. Therefore, these variables were: the irrigated surface area; theoretical irrigation requirements; and mean slope. The irrigated surface area, in hectares, was the first variable considered, since the larger surface, the higher the probability to find MHP potential. Irrigated surface area was a proxy measure of pipe flow rate as large surface areas will require greater flows and therefore could have a higher potential for hydropower production. The irrigation requirements, in m^3 $year^{-1}$, depended on the crops cultivated and on the agro-climatic parameters (rainfall and evapotranspiration) of the area studied. Crops with higher irrigation requirements would lead to a greater irrigation time, which would again affect flow rates and potential energy production in a turbine. On the other hand, in areas with high rainfall and low evapotranspiration, the irrigation requirements would be lower and so the irrigation time and vice versa. Therefore, this variable also considered climatic conditions as a proxy. Finally, the mean slope in percentage was introduced in the model to represent how the terrains topography affected the energy recovery potential. It was assumed that the mean ground slope would be related to potential overpressure within the network and areas with higher slopes were more likely to contain overpressures. In this way, the model was applied in each area analysed, once the proxy variables were gathered, identifying more or less potential depending on the distribution of these variables in each specific site.

Downstream of the 177 potential MHP locations found in each network, the irrigated surface was considered as a unique plot, independent of the number of hydrants found. Therefore, the irrigated surface, defined as one of the explanatory variables, was obtained. To calculate theoretical irrigation requirements for every location, the crop distribution as well as the agro-climatic parameters (rainfall and evapotranspiration) were required. The crop distribution was known for every location from the development of the 18 hydraulic models. Regarding the agro-climatic parameters, the information was gathered from the closest weather stations in each case. The theoretical irrigation requirements were then calculated applying the method proposed by Allen et al. [23] using the CROPWAT software [24]. The mean slope in percentage was calculated for each point considering the distance and the height difference between the water source and the most critical hydrant for each location. This was the hydrant with the lowest head available of the branch assessed, when 100% of the hydrants were open simultaneously.

2.4. Linear analysis

The first stage of the study was conducted assuming the relationship among variables was linear, which reduced the complexity of the problem. Thus, single linear analysis was first carried. The input variables were related one on one to the response variable. The correlation coefficient (r), coefficient of determination (R^2), the mean squared error and the mean absolute error (MAE) were used as statistical metrics to evaluate the existing relationship. High values of the coefficient of determination would indicate linearity, against low values, which would indicate no linearity in their relationship. The entire sample of 177 locations was used to do this analysis. The different linear models evaluated followed Equation (1), where x_i referred to the different input variables used.

$$Y(x_i) = a + bx_i \tag{1}$$

2.5. Non-linear analysis

In order to consider non-linearity among the variables selected, Artificial Neural Networks (ANNs) were used for this analysis. ANNs are structures or models used for the learning process carried out in machine and deep learning approaches, structured in layers stacked on top of each other [25]. The general structure of these models is composed by an input layer, which corresponds to the explanatory variables, hidden layers, used to transform the inputs into outputs, and an output layer, which is the expected value. Each of these layers has a number of neurons, which should be defined specifically for each problem. ANNs are able to predict an output using different input variables, capable of adapting to non-linear relationship between output and input. ANNs have been applied in several engineering fields, such as rainfall forecasting, time variables prediction or water demand forecasting in irrigation networks [26–37].

2.5.1. Data transformation

As the different inputs variables had different units, this fact could lead to some difficulties during the ANN learning process, since the range of values for each input variable could be widely different. There are several methods to avoid such problems while

improving the accuracy of the model, such as normalisation or transformation. In this case, logarithmic transformation was applied just for the irrigation requirements, since the range of values found for this variable was normally much higher than the other two variables. Using this transformation, the values were brought into a more similar value range to the other explanatory variables, following Equation (2).

$$X' = \log(X) \tag{2}$$

Where X' is the value transformed, X is the actual input data.

2.5.2. ANN network structure

During the model definition, we have to set, among others, the number of neurons on each layer, the number of hidden layers and the number of epochs. In addition, inputs and outputs were previously defined. The number of hidden layers tested varied between one and two, in which different numbers of neurons were tried, varying between 2 and 64. The sample was divided into a training and validation set. Different size distributions of these sets were tested, analysing the minimum squared error (MSE) for each distribution in each fold and selecting the one returning the minimum mean value. Four different folds were randomly selected, optimising the objective function for each number of hidden layer, neurons and sample distribution in every fold. Since the size of the sample was small, the scores obtained for each fold could vary from one fold to another. Thus, average results for the four folds were considered in order to obtain more accurate results. The minimum average of the four folds was calculated, whose structure was fixed as the optimal. The number of epochs tried oscillated between 1 and 300 for each possible configuration.

The gradient descent optimisation algorithm ADAM (Adaptive Moment Estimation) [38] was implemented in Python for the learning process, aiming to obtain an optimal prediction of the energy recovery existing potential. The objective function was used to measure the performance on the training and validation data. The MSE was used as the objective function (Equation (3)), which measured the average of the square of the errors. The optimal configuration was provided by the structure whose MSE was the minimum. The relationship between the input and output in a neuron was analysed using an activation function. To consider non-linearity, the Rectified Linear Unit function (ReLU) was used, since it was more computationally efficient than other non-linear functions. Mathematically, this function is expressed as per Equation (4).

$$MSE = \frac{1}{N} \sum (y_i - \widehat{y}_i)^2 \tag{3}$$

$$f(x) = \max(0, x) \tag{4}$$

2.5.3. K-fold cross validation

The validation of the ANN model was carried out running k-fold cross validation, consisting of the partition of the whole sample into k equal sets randomly selected. For this purpose, the observations were split into training and validation sets, corresponding to the distribution, which returned the best results during the network architecture definition. For each fold, the input data corresponded to the training data of the explanatory variables and the output to the energy recovery potential of the same set. Using the validation data, the output variable was predicted inputting the data corresponding to the explanatory variables of the same set. Comparing the predicted values with the observed ones, two

statistical metrics were calculated: the mean absolute error, expressed in the same units to the output variable; and the coefficient of determination, whose value varies within the range [0−1].

2.6. Application to large geographical scale predictions

Finally, the models were compared, selecting the one that provided the best metrics. This was used to predict the energy recovery potential in every municipality in the whole province of Seville and the province of Cordoba. The potential of 180 municipalities was predicted, 105 of them corresponded to Seville and 75 to Cordoba. The input variables were gathered for these municipalities from the SIGPAC platform [39], where different information, such as crop cultivated, mean slope or irrigation coefficient were found for all the plots of each municipality. A database with around 20 million data points was compiled and analysed for the whole region. Thus, the surfaces with crop cultivations were extracted, calculating the theoretical irrigation requirements for all of them. The agro-climatic parameters were obtained from 29 weather stations distributed around Seville and Cordoba. For the mean slope of each municipality, a weighted measure was calculated, considering the mean slope for each plot containing crops, as per Equation (5). The irrigated surface found in every municipality was corrected with the average relation factor aforementioned, which showed the ratio of surface area with viable energy recovery potential found for the networks analysed individually (0.70). This correction was necessary, since otherwise the whole irrigated surface found for each municipality would have potential, which is unrealistic. The dominant crops found were olive trees and citrus, which occupied 66.7% and 23.4% of the total irrigated surface respectively.

$$S_m = \frac{A_1}{A_T} S_1 + \frac{A_2}{A_T} S_2 + \dots + \frac{A_n}{A_T} S_n \tag{5}$$

Where S_m is the mean slope for each municipality; A_1, A_2, A_n correspond to the area of the plots 1,2 and n respectively, where crops were found; A_T is the total irrigated area with crops; and S_1, S_2, S_n were the mean slope for the plots 1, 2 and n.

2.7. Energetic and economic analysis

To assess the potential benefits associated with the adoption of MHP technology, two analyses were carried out for the outputs predicted for Seville and Cordoba. For the energy analysis, a general comparison between the energy consumption and the potential recovery was conducted. Rodriguez et al. [2] estimated an average energy consumption per unit of irrigated area of 1003 kW ha^{-1} for ten pressurised irrigation districts located in Southern Spain, which contained 11 of the networks used in this research, all of them located in Andalusia. Regarding the economic analysis, the potential savings for the energy cost per irrigated surface unit was calculated. Fernandez Garcia et al. [3] studied how the energy cost, related to the total water cost, per unit of irrigated area, changed after the modernisation process. Its value was reported for five pressurised irrigation districts, including some of the networks analysed in this research. This value for pressurised infrastructure varied between €48.9 ha^{-1} and €147.6 ha^{-1}. A mean weighted value of €127.5 ha^{-1} was used.

3. Results

3.1. Linear analysis

The results of simple linear regression approach varied widely depending on the variable considered. The highest r and R^2 were

Table 3
Linear analysis results for single models.

Single	a	b	r	R^2	MSE	MAE
Irrigated Surface	9.691	0.157	0.754	0.569	479.67	15.27
Irrigation Requirement	−249.464	48.95	0.609	0.37	700.36	18.85
Slope	36.754	−2.749	0.071	0.005	1106.92	22.69

obtained for the irrigated surface area (0.754 and 0.569 respectively) whilst the lowest was obtained for the slope (0.0071 and 0.005). Nevertheless, it could be seen that the relationship between the irrigated surface area and the energy potential was not strongly linear. Concerning the MSE and MAE, the model using the irrigated surface area as an input returned the best results for both metrics, with 479.67 MWh and 15.27 MWh respectively. The outcomes of this analysis were shown in Table 3.

3.2. Non-linear analysis

The different results from the non-linear analysis carried out to define the ANN network structure can be seen in Fig. 2, where each line represented a different number of neuron configurations. The average MSE for the four folds is represented for each configuration for every number of epochs run for two hidden layers. The observation distributions among training and validation sets that provided the best results yielded 74%–26% respectively. In order to evaluate how this distribution could affect the model performance, the mean and standard deviation was calculated for each fold, obtaining slight differences between the training and validation sets. More specifically, the maximum differences were obtained in fold 2 for the mean value (17.6%) and in fold 4 for the standard deviation (17%). Therefore, the training data set included 130 observations and the validation set included 47. Two hidden layers composed the structure with the minimum MSE, with 26 and 18 neurons respectively. The minimum MSE obtained by the aforementioned structure can be seen and compared with the rest of the structures tested in the zoom plot of Fig. 2. Concerning the overfitting, all of the lines trended to increase their slope after reaching a certain number of epochs, so a greater number of epochs for those would be translated in an overfitting on the model. The minimum MSE was achieved by the previously defined structure after running 36 epochs, with an average value of 383.3 MWh for the four folds.
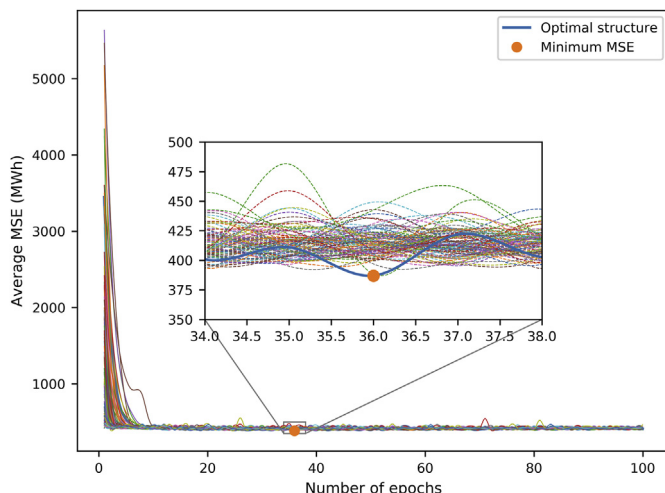

Fig. 2. Average MSE of the four folds for each configuration of neurons in each epoch.

Once the structure of network was defined, the validation was carried out. The results of MSE, which was the objective function, was obtained for each epoch and fold run. The metrics R^2 and MAE were also calculated for every fold. The predicted values compared to the validation ones can be graphically seen in Fig. 3. The results showed how the R^2 changed as per the training and validation data. The best approach was obtained using the sets configured in the third fold, with an R^2 of 0.736. The fourth fold prediction was the weakest with R^2 equal to 0.462. The average R^2 obtained from the four folds was 0.631. On the other hand, the MAE of the different folds varied between 13.25 MWh and 15.59 MWh, with an average value of 14.52 MWh. This average obtained from the four folds MAE was used to correct the predicted values of each the municipality in Seville and Cordoba, since it represented a more reliable metric than using the MAE obtained from a unique fold. The correction was carried out considering both, positive and negative MAE, thus adding or subtracting it to the value predicted.

Comparing the test and predicted values for the folds considered and analysing the errors aforementioned, it was found that the sum of energy of the test set was lower than the predicted for the first fold and greater for the three remaining. The difference found between the total amount of energy for both, test and predicted sets, varied between 5% and 18%. Although the MAE showed high relative errors for single observations (44% in the worst case), when the whole set was compared, these errors significantly decreased. Furthermore, when the total average energy value, for test and predicted sets, of the four folds were compared, the difference found was just 6.5%, which could be considered acceptable for prediction purposed. This overall error was also used to correct the potential forecasted in the municipalities.

The objective function (MSE) got its minimum in the third fold (320.90 MWh), with an average value of 383.3 MWh. The difference between the results obtained for first three folds and the fourth one could come from different distribution of MHP locations on the training set used. The distribution of MHP turbine energy around the mean value is uniform in the first three folds (see Fig. 4), with a percentage of observations below and over the mean value of 62.3%–37.7%, 61.5%–38.5% and 63.8%–36.2% respectively. Although for the fourth fold this percentage is similar to the other folds (64.6%–35.4%), analysing the standard deviation (in MWh), it could be deduced where this difference in the coefficient of determination is coming from. The values obtained for the first three folds were 32.4, 32.8 and 32.4, while for the last one it was 34.8. This fact showed a wider distribution of the training set points of the fourth fold from its mean value.

On the other hand, when the validation sets were analysed the results showed opposite trends than in the training sets. Thus, the standard deviation of the first three folds was 35.6, 34.4 and 35.8 respectively, while for the fourth it was 28.9. A summary for each fold for the different results is shown in Table 4, as well as the mean and standard deviation values for the train and test sets.

Comparing the results obtained using ANNs with the linear analysis, it can be seen how the MSE values are significantly lower. The MSE achieved in the third fold was around 36% lower than the MSE accomplished in the multivariate linear analysis using all the variables. This difference kept increasing when the number of variables decreased. The four folds average minimum MSE achieved in this last model was 20.5% lower than the MSE obtained in the three variables linear multivariate analysis. When other metrics were compared, ANN also showed better results. The maximum R^2 attained in the third fold was almost 30% greater than the best value achieved in the linear analysis. When the average result was equated, the results improved by 10%. However, a 10% improvement could be considered large in this case, as the sample was split into two sets for the ANN model but used in its entirety for the
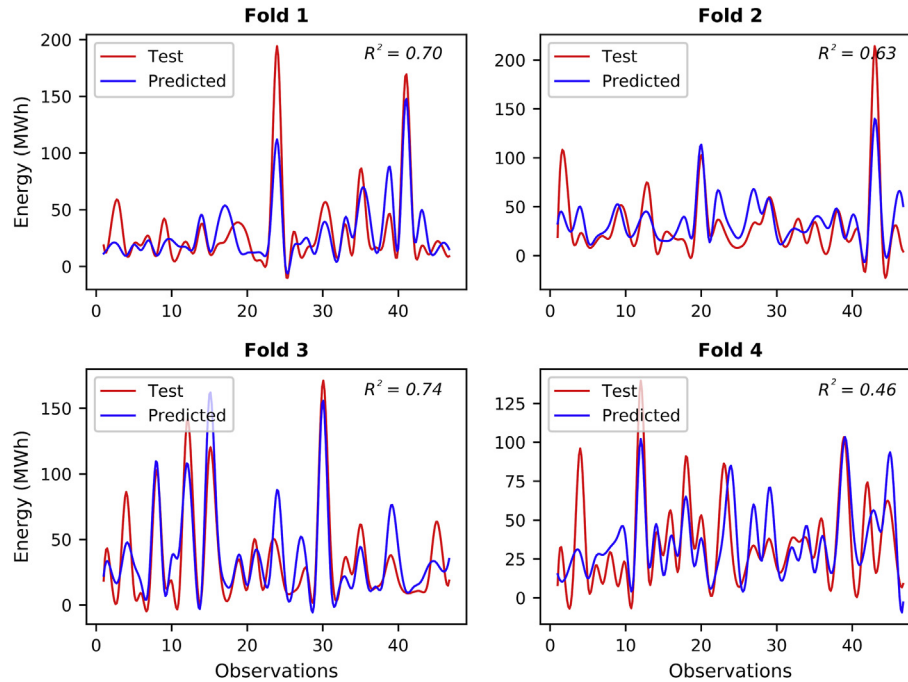
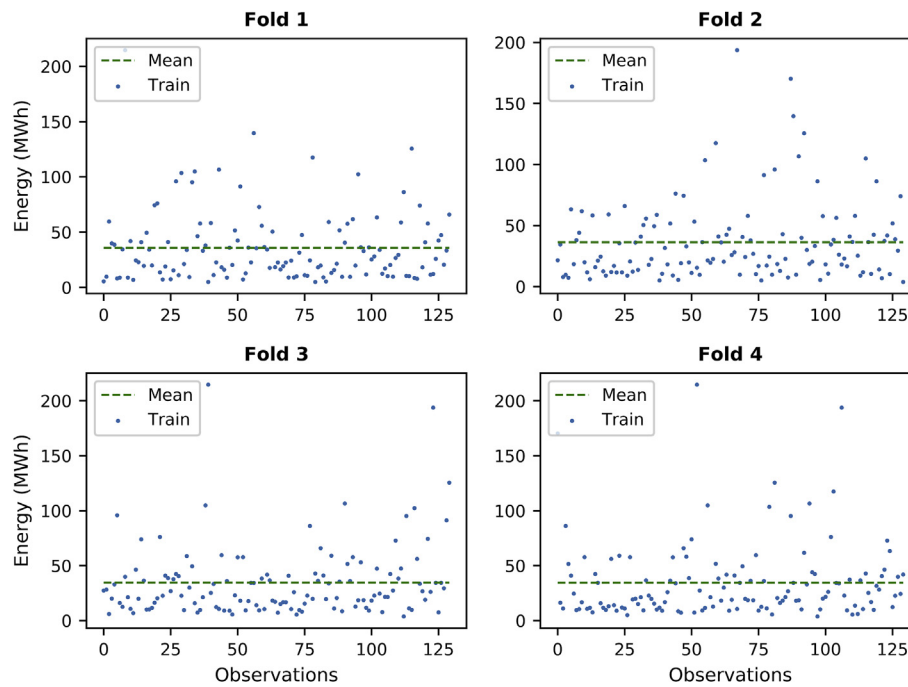**Fig. 3.** Comparison between test observations and predicted values for each fold.

Fig. 4. Training set distributions from the mean value for each fold.

linear analysis. Finally, the MAE obtained in the first fold was 13.5% smaller than the best result obtained in the linear analysis, decreasing this difference down to 5% when the average MAE was compared.

3.3. Prediction in municipalities

With the network already defined and validated, the energy recovery potential for every municipality of the provinces of Seville

and Cordoba was predicted. The pressurised irrigated surface raised up to 163,472 ha, composing around 6% of the total surfaced encompassed in both provinces, from which 114,430 ha were analysed, obtained after applying the correction factor outlined in the Materials and Methods Section (0.7).

The total energy potential predicted for the whole region varied between 19.64 and 22.38 GWh, depending on the fold used. The average potential predicted was 21.05 GWh for 2018. Applying corrections for the previous errors observed between test and

Table 4
Mean values and standard deviations obtained for each fold for the train and test sets.

K	R^2	MAE	MSE	Mean		Standard Deviation	
				Train	Test	Train	Test
1st Fold	0.698	13.25	369.8	35.9	30.8	32.4	35.6
2nd Fold	0.627	15.43	396.8	36.3	29.9	32.8	34.4
3rd Fold	0.7361	13.82	320.9	34.5	34.7	32.4	35.8
4th Fold	0.462	15.59	445.8	34.5	34.8	34.8	28.9
Average	0.631	14.52	383.3	—	—	—	—

predicted sets for the four folds, 6.5%, the energy varied between 19.7 and 22.4 MWh. Applying the average MAE from the folds analysis, 14.52 MWh to the average value obtained for each municipality as correction measure, a minimum and maximum energy of 18.95 GWh and 23.70 GWh were obtained. This value was distributed among the 180 municipalities, among which, six were found not to have irrigated surface or energy recovery potential (see Fig. 5). It can be observed how the different values obtained after applying different correction measures were similar, with differences oscillating between 6.5% and 12.5%. An average value of 114.4 MWh per municipality was found.

The greatest potential was predicted to be found in Ecija (Seville), with an annual value of 1.63 GWh. If the variables were analysed, it seems rational that the maximum potential was found there, since more than 9000 ha were found to be irrigated, whilst the average irrigated surface per municipality was 636 ha. Its average slope was 7.8%, whilst the average slope of the municipalities was 5.5% However, most of the municipalities showed a potential lower than the average, accounting for around 71.7%, with a potential lower than 100 MWh per year. More than 50% of the potential predicted was concentrated in 20 of the sites analysed, which irrigated around 60% of the surface considered. Evaluating the case of Ecija and inputting a slope of zero, the energy potential estimated increases in every fold, giving an average value of 1.82 GWh per year. Therefore, it could be extracted that high slope

values return lower values of predicted energy recovery. The map showing the predicted potential is presented in Fig. 5.

3.4. Energetic and economic analysis

Considering the energy consumed in the entire irrigated surface found in Seville and Cordoba, the percentage of energy that could have been saved in 2018 in the irrigation sector was 12.8%. Introducing this percentage of energy savings into the energy cost index per unit of irrigated surface area, it was estimated that this index could be reduced from €127.5 ha^{-1} to €111.1 ha^{-1} using the average value given by Fernandez Garcia et al. [3].

However, it would be optimistic to suggest that all of this energy potential could be directly introduced to the grid or used at the point of production as self-consumption. In the irrigation sector located in agricultural and often rural areas, the likelihood lack of grid connections or local energy demands close to points where MHP turbines could be installed is only moderate.

4. Discussion

Sustainability requires, among others factors, enhancement of the use of non-fossil fuel energy sources. The existing water networks present, in many cases, an overpressure that is being dissipated in different ways. Micro hydropower (MHP) appears as a potential solution for renewable energy to be implemented in different fields within the water industry, transforming part of the potential energy, represented in the form of overpressure in pipelines, into electricity. Previous research for assessing MHP potential were focused in the analysis of those locations where detailed network information was gathered. However, the lack of larger scale assessment in the different sectors encompassed within this industry makes having a clear idea of the existing potential and its benefits more difficult.

An analysis for Andalusia's historical electricity consumption, GHG emissions and renewable energy generation was made from 2013 to 2018 (see Fig. 6). The Spanish emission factors for the different years were used to quantify the emissions of Andalusia
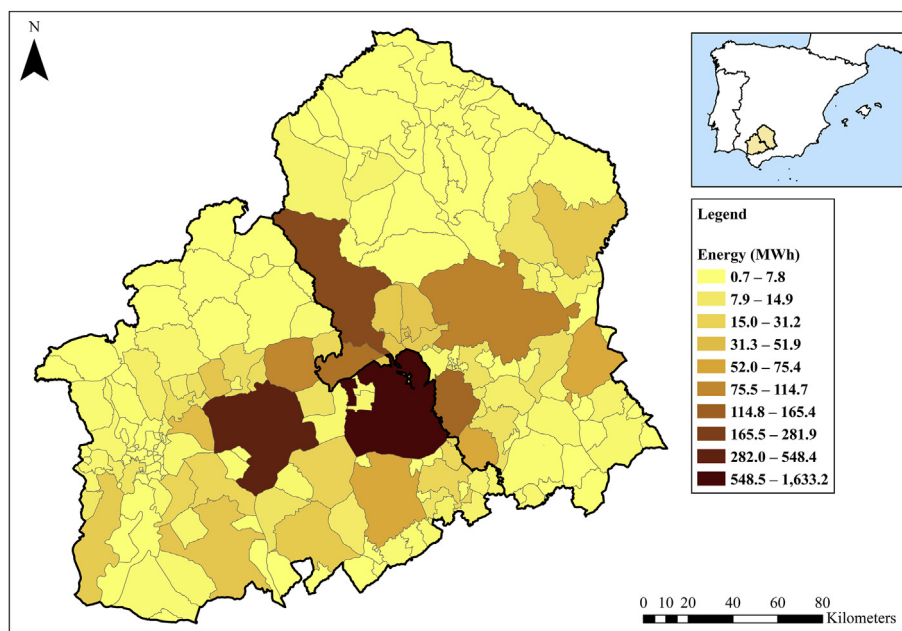

Fig. 5. Energy recovery potential found for every municipality of Seville and Cordoba during the year 2018.
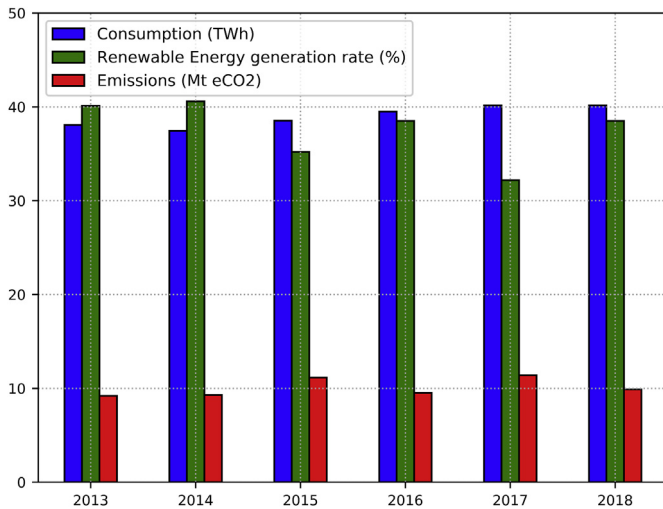
Fig. 6. Historical levels of electricity consumption, renewable energy generation rate and emissions for Andalusia.

[40]. The energy consumption tended to decrease related to 2013 levels, however from 2014 onwards, it kept increasing by around 1 TWh per year until 2017. The 2018 levels kept constant when compared to 2017. Nevertheless, the GHG emissions, measured in Mt eCO_2 did not follow the same trend as the consumption. They were more related to the ratio of renewable energy generation. When this ratio decreased, it could be seen how the emissions for that year increased and vice versa. Thus, both parameters had a variable trend along those years. Considering that irrigation is the main economic and energy consuming activity in Andalusia and that more activities are included within the water industry, MHP could make an important contribution for the renewable energy sector and the sustainability of irrigation activities.

Although the methodology used in this research provided theoretical results, which might not be fully exploited or may not completely match with the actual existing resources in reality, it helped to have an approximated idea of these resources. Large-scale assessments estimating the existing resources could help to foster hydropower as potential solution to improve the energy efficiency in water industry. In this research, the estimated energy savings would affect the operational cost of irrigation networks, which suffered an increase of 500% after replacing traditional open channels with pressurised networks in Andalusia [41]. Going further, Andalusia accounts for more than a million ha of irrigated surface, of which 84% is pressurised [42]. However, the prediction here was carried out just for 15% of this pressurised irrigated surface. The method proposed in this research was employed in regions with similar irrigation infrastructures and design parameters for networks. All of them were design for 100% simultaneity and a high design demand, which ranged between 1 and 1.2 l s^{-1} ha^{-1}, as previously stated. Thus, the application of the trained ANN obtained would be limited to regions with similar networks' properties, where the water source was superficial and the same simultaneity index and design demand. The method could also be applied to others areas with different networks' characteristics, but the ANN would have to be defined again using networks typical from the region to be analysed. Future research could be focused on the analysis of networks from different regions and the quantification of MHP's benefits in the whole Andalusia.

The results obtained in this research showed a mean R^2 of 0.63, which reached its highest value (0.74) in the third fold of the cross validation. The potential predicted for provinces accounted

21.05 GWh found in an irrigated surface of 163,472 ha during 2018. The potential found per irrigated unit area was 0.129 MWh ha^{-1}. This compares well with the measured potential of 0.167 MWh ha^{-1} in the 18 detailed hydraulic networks. The difference between the observations and the predicted values was of 0.038 MW ha^{-1}. Comparing the MAE corrected potential, the energy recovery per unit irrigated area varied from 0.116 MWh ha^{-1} to 0.145 MWh ha^{-1}. The energy recovery potential per unit irrigated surface indices found in previous research, in MWh ha^{-1}, varied as 0.65, 0.08, 0.10 and 0.11 in Ref. [14,16—18] respectively. The results obtained in this research remained within acceptable values when compared to these previous ratios. If the index extracted from the 18 networks would have been used to predict the existing potential in the irrigated surface for both provinces, 7.5 GWh more would be estimated. Thus, the method proposed in this research, is able to go beyond the simple assumption of a linear trend between the irrigated surface and the existing potential, through a deep learning process and the introduction of other proxy variables. In addition, the correction factor of 0.7 was used to limit the predicted output to show just energy recovery potential which was economically viable. Therefore, the surface taken into account in Seville and Cordoba could be increased if the economical parameters used in the methodology developed by Crespo Chacon et al. [17] changed (i.e. economical savings per energy unit, installation costs, grants, etc).

On the other hand, it would be very complex to find points with existing potential within the irrigation sector where the energy recovered could be directly sold to the grid, stored or used for direct purposes, such as pumping. For either use, the installation costs could be importantly increased, making too many of the plants not economically attractive for investors. MHP solutions together with storage systems could be a potential way to apply the energy recovered at points where energy is required. Nevertheless, costs and logistic make this solution unviable currently for those points where no energy is needed. An attractive use for recovering energy could be at farm levels, where farmers with no access to electric grid tend to use diesel generators if some energy consumption is required. Adopting this alternative would reduce the amount of energy to be recovered using MHP as turbines located at farm level will inevitably have less flow and pressure available than those located higher up in the pipe network. However, this could still be considered as a potential measure to reduce the energy dependency of irrigation networks.

Comparing the energy savings obtained by MHP with other measures for improving the energy dependency in irrigation, the results obtained here showed that MHP could be an important solution. Furthermore, it could also be applied in tandem with the different energy saving measures previously highlighted in the introduction. For example, irrigation scheduling together with MHP could be a potential solution to improve energy dependency. Concerning the photovoltaic solution, this is limited for power production just during the sunlight hours. In big irrigation infrastructures, pumping using photovoltaic energy is considered as a potential solution to reduce the energy dependency. The addition of MHP to this solution in the networks could lead to an important reduction of the energy consumption in this sector. Coupling both technologies could be of special interest for future research.

5. Conclusion

Sustainable development requires clean energy sources for GHG emissions to be reduced in the short term, and completely avoided in the long term. Hydropower accounted for almost 40% of the total renewable generation in the EU [43]. Nonetheless, micro hydro resources are not very well exploited yet. It is first necessary to

conduct a large geographical scale assessment of these available resources in different sectors, quantifying the existing potential and its intrinsic environmental and economic benefits, in order to allow targeted investment in micro-hydropower.

This research explored energy recovery in pressurised irrigation networks on a large geographical scale using hydropower without flow or head information. This is the main novelty of this paper, as previous studies were focused on specific locations with available information, providing an approximation of the existing resources and potential impacts on a small geographical scale. Linear models, single and multivariate, and ANNs were studied in this research for predicting the MHP energy recovery in pressurised irrigation networks. Three variables that could be easily obtained for the different irrigated areas in many regions around the world were used as input data, through which the energy recovery potential was predicted. Irrigated surface area of pressurised systems, irrigation requirements (directly related to the crops and the agroclimatic parameters), and mean slope of the area were the input variables. Inputs and outputs, were first obtained for 18 irrigation networks, using detailed hydraulic models, where 177 potential MHP installation (as observations), composed the database of economically viable sites. ANNs showed the best results and was used for large-scale prediction in two provinces in Andalusia. Using the ADAM optimisation algorithm and minimising the mean squared error as objective function, the network was able to predict the energy recovery potential with a R^2 varying from 0.46 to 0.74, with an average value of 0.63. The minimum MSE varied between 320.9 and 445.8 MWh, with a mean value of 383.3 MWh. Two hidden layers with 26 neurons in the first and 18 in the second composed the network's structure. A total potential of 21.05 GWh during 2018 for the regions of Seville and Cordoba, in Southern Spain, was predicted. Important environmental and economic benefits would be linked to this energy recovery, with more than 5000 t eCO_2 per year and more than 12% of reduction in energy costs in irrigation.

Declaraion of competing interest

No conflict of interest exists.

CRediT authorship contribution statement

Miguel Crespo Chacón: Conceptualization, Methodology, Writing - original draft. **Juan Antonio Rodríguez Díaz:** Formal analysis. **Jorge García Morillo:** Resources. **Aonghus McNabola:** Supervision, Writing - review & editing.

Acknowledgements

References

[1] International Energy Agency, World Energy Outlook: water-energy Nexus [Online]. Available, https://www.iea.org/weo/water/. (Accessed 19 June 2019).

[2] J.A. Rodríguez Díaz, E. Camacho Poyato, M. Blanco Pérez, Evaluation of water and energy use in pressurized irrigation networks in Southern Spain, J. Irrigat. Drain. Eng. 137 (10) (2011) 644—650.

[3] I. Fernández García, J.A. Rodríguez Díaz, E. Camacho Poyato, P. Montesinos, J. Berbel, Effects of modernization and medium term perspectives on water and energy use in irrigation districts, Agric. Syst. 131 (2014) 56—63.

[4] J.A.R. Díaz, R.L. Luque, M.T.C. Cobo, P. Montesinos, E.C. Poyato, Exploring energy saving scenarios for on-demand pressurised irrigation networks, Biosyst. Eng. 104 (4) (2009) 552—561.

[5] M.A. Jiménez-Bello, F. Martínez Alzamora, V. Bou Soler, H.J.B. Ayala, Methodology for grouping intakes of pressurised irrigation networks into sectors to minimise energy consumption, Biosyst. Eng. 105 (4) (2010) 429—438, https://doi.org/10.1016/j.biosystemseng.2009.12.014.

[6] M.A. Moreno, J.I. Córcoles, J.M. Tarjuelo, J.F. Ortega, Energy efficiency of pressurised irrigation networks managed on-demand and under a rotation schedule, Biosyst. Eng. 107 (4) (2010) 349—363, https://doi.org/10.1016/j.biosystemseng.2010.09.009.

[7] J.M. Navarro Navajas, P. Montesinos, E.C. Poyato, J.A. Rodríguez Díaz, Impacts of irrigation network sectoring as an energy saving measure on olive grove production, J. Environ. Manag. 111 (2012) 1—9, https://doi.org/10.1016/j.jenvman.2012.06.034.

[8] I. Fernández García, M.A. Moreno, J.A. Rodríguez Díaz, Optimum pumping station management for irrigation networks sectoring: case of Bembezar MI (Spain), Agric. Water Manag. 144 (2014) 150—158, https://doi.org/10.1016/j.agwat.2014.06.006.

[9] A. Mérida García, I. Fernández García, E. Camacho Poyato, P. Montesinos Barrios, J.A. Rodríguez Díaz, Coupling irrigation scheduling with solar energy production in a smart irrigation management system, J. Clean. Prod. 175 (2018) 670—682.

[10] L. Corcoran, P. Coughlan, A. McNabola, Energy recovery potential using micro hydropower in water supply networks in the UK and Ireland, Water Sci. Technol. Water Supply 13 (2) (2013) 552—560.

[11] C. Power, A. McNabola, P. Coughlan, Development of an evaluation method for hydropower energy recovery in wastewater treatment plants: case studies in Ireland and the UK, Sustain. Energy Technol. Assess. 7 (2014) 166—177.

[12] J. Gallagher, I.M. Harris, A.J. Packwood, A. McNabola, A.P. Williams, A strategic assessment of micro-hydropower in the UK and Irish water industry: identifying technical and economic constraints, Renew. Energy 81 (2015) 808—815.

[13] F. Estrada Tarragó, H.M. Ramos, Micro-hydro solutions in Alqueva Multipurpose Project (AMP) towards water-energy-environmental efficiency improvements., MSc Thesis, Escola Técnica Superior d'Enginyeria Industrial de Barcelona (2015).

[14] M. Pérez-Sánchez, F.J. Sánchez-Romero, H.M. Ramos, P.A. López-Jiménez, Modeling irrigation networks for the quantification of potential energy recovering: a case study, Water (Switzerland) 8 (6) (2016).

[15] M. Pérez-Sánchez, F.J. Sánchez-Romero, H.M. Ramos, P.A. López-Jiménez, Optimization strategy for improving the energy efficiency of irrigation systems by micro hydropower: practical application, Water (Switzerland) 9 (10) (2017) 1—22.

[16] J. García Morillo, A. McNabola, E. Camacho, P. Montesinos, J.A. Rodríguez Díaz, Hydro-power energy recovery in pressurized irrigation networks: a case study of an Irrigation District in the South of Spain, Agric. Water Manag. 204 (December 2017) (2018) 17—27.

[17] M. Crespo Chacón, J.A. Rodríguez Díaz, J. García Morillo, A. McNabola, Pump-as-turbine selection methodology for energy recovery in irrigation networks: minimising the payback period, Water (Switzerland) 11 (1) (Jan. 2019) 1—20, https://doi.org/10.3390/w11010149.

[18] M. Crespo Chacón, J.A. Rodríguez Díaz, J. Garcia-Morillo, A. Mcnabola, Hydropower energy recovery in irrigation networks: validation of a methodology for flow prediction and pump as turbine selection, Renew. Energy 147 (2020) 1728—1738, https://doi.org/10.1016/j.renene.2019.09.119.

[19] D. Mitrovic, et al., Hydropower energy recovery in water pipe networks : spatial regression analysis using GIS, assessing the correlation between energy recovery potential and geographical data, in: Water Efficiency Conference, 2018.

[20] D. Novara, A. Carravetta, A. McNabola, H.M. Ramos, Cost model for pumps as turbines in run-of-river and in-pipe microhydropower applications, J. Water Resour. Plann. Manag. 145 (5) (2019) 4019012.

[21] A. Carravetta, G. Del Giudice, O. Fecarotta, H.M. Ramos, Pump as turbine (PAT) design in water distribution network by system effectiveness, Water (Switzerland) 5 (3) (2013) 1211—1225.

[22] O. Fecarotta, A. Carravetta, H.M. Ramos, R. Martino, An improved affinity model to enhance variable operating strategy for pumps used as turbines, J. Hydraul. Res. 54 (3) (2016) 332—341, https://doi.org/10.1080/00221686.2016.1141804.

[23] R.G. Allen, L.S. Pereira, D. Raes, M. Smith, Crop Evapotranspiration: Guidelines for Computing Crop Requirements, 1998.

[24] M. Smith, CROPWAT: A Computer Program for Irrigation Planning and Management, 1992.

[25] F. Chollet, Deep Learning with Python, First, Manning Publications Co., New York, 2018.

[26] M.N. French, W.F. Krajewski, R.R. Cuykendall, Rainfall forecasting in space and time using a neural network, J. Hydrol. 137 (1—4) (1992) 1—31.

[27] M. Zhang, J. Fulcher, R.A. Scofield, Rainfall estimation using artificial neural network group, Neurocomputing 16 (2) (1997) 97—115.

[28] J.C. Gutiérrez Estrada, E.J. de Pedro Sanz, R. López Luque, I. Pulido Calvo, Estimación a corto plazo de la temperatura del agua. Aplicación en sistemas de producción en medio acuático, Ing. del agua 12 (1) (2015) 77.

[29] R.G. Perea, E.C. Poyato, P. Montesinos, J.A.R. Díaz, Irrigation demand forecasting using artificial neuro-genetic networks, Water Resour. Manag. 29 (15) (2015) 5551—5567.

[30] H.H. Park, Analysis and prediction of walleye pollock (Theragra chalcogramma) landings in Korea by time series analysis, Fish. Res. 38 (1) (1998) 1—7,

https://doi.org/10.1016/S0165-7836(98)00118-0.

[31] R.J. Abrahart, L. See, Comparing neural network and autoregressive moving average techniques for the provision of continuous river flow forecasts in two contrasting catchments, Hydrol. Process. 14 (11—12) (2000) 2157—2172.

[32] R.J. Kuligowski, A.P. Barros, Experiments in short-term precipitation forecasting using artificial neural networks, Mon. Weather Rev. 126 (2) (2002) 470—482.

[33] I. Pulido, J. Roldán, R. López, J. Gutiérrez, Técnicas de predicción a corto plazo de la demanda de agua. Aplicación al uso agrícola, Ing. del agua 9 (3) (2002) 319—331.

[34] I. Pulido-Calvo, J. Roldan, R. Lopez-Luque, J.C. Gutierrez-Estrada, Demand forecasting for irrigation water distribution systems, J. Irrigat. Drain. Eng. 129 (December) (2003) 422—431.

[35] I. Pulido-Calvo, P. Montesinos, J. Roldán, F. Ruiz-Navarro, Linear regressions and neural approaches to water demand forecasting in irrigation districts with telemetry systems, Biosyst. Eng. 97 (2) (2007) 283—293, https://doi.org/10.1016/j.biosystemseng.2007.03.003.

[36] I. Pulido-Calvo, M.M. Portela, Application of neural approaches to one-step daily flow forecasting in Portuguese watersheds, J. Hydrol. 332 (1—2) (2007) 1—15, https://doi.org/10.1016/j.jhydrol.2006.06.015.

[37] I. Pulido-Calvo, J.C. Gutiérrez-Estrada, Improved irrigation water demand forecasting using a soft-computing hybrid model, Biosyst. Eng. 102 (2) (2009) 202—218.

[38] D.P. Kingma, J. Ba, Adam: A Method for Stochastic Optimization, 2014, pp. 1—15.

[39] Junta de Andalucia, Consejeria de Agricultura, Ganadería, Pesca y Desarrollo Sostenible, 2019 [Online]. Available, https://www.juntadeandalucia.es/organismos/agriculturaganaderiapescaydesarrollosostenible/areas/politica-agraria-comun/paginas/sigpac-descarga-informacion-geografica-shapes-provincias.html. (Accessed 10 June 2019).

[40] Red Eléctrica de España, Series estadísticas nacionales, 2019 [Online]. Available, https://www.ree.es/es/estadisticas-del-sistema-electrico-espanol/series-estadisticas/series-estadisticas-nacionales. (Accessed 15 May 2019).

[41] J.A. Rodríguez-Díaz, E. Camacho-Poyato, R. López-Luque, L. Pérez-Urrestarazu, Benchmarking and multivariate data analysis techniques for improving the efficiency of irrigation districts: an application in Spain, Agric. Syst. 96 (1—3) (2008) 250—259, https://doi.org/10.1016/j.agsy.2007.07.010.

[42] A. y M. A. Ministerio de Agricultura y Pesca, Encuesta sobre superficie y rendimientos de cultivos. Informe sobre regadíos en España., 2018.

[43] European Commission, EU in Figures. Statistical Pocketbook, 2018.