

Machine learning para modelizar el efecto
de medidas de gestión

Planificación y gestión de recursos hídricos a partir de análisis de datos de IoT (WATERoT)

Dr. D. José M. Cecilia.
Investigador Ramón y Cajal.
Grupo de Redes de Computadores (GRC)
Universitat Politècnica de València

WaterOT: Ministerio de Ciencia, convocatoria Retos-Colaboración (RTC-2017-6389-5)

Empresas



Universidades y centros tecnológicos



Equipo multidisciplinar de investigación



Computación, IA, PLN, IoT ...



Planificación y gestión hidrológica

Motivación

Mejorar la *planificación y gestión de recursos hídricos* gracias al análisis de datos. Pero...

1. ¿Con qué datos? → *Temperatura, Precipitación,..., Opinion mining*
2. ¿Cómo accedemos a esos datos? → *IoT, Redes Sociales, OpenData*
3. ¿Qué calidad tienen esos datos? → *IA, Combinación de fuentes (Fusion)*
4. ¿Qué técnicas aplicamos? → *Machine learning, Modelos físicos, PLN ...*
5. ¿Para mejorar qué? → *Aplicaciones útiles para la sociedad*
(*Predicción de heladas, sedimentos, evaporización en balsas...*)

Ejemplos de aplicaciones.

Modelado de la eutrofización del Mar Menor



International Journal of
*Environmental Research
and Public Health*

Submit to this Journal

Review for this Journal

Edit a Special Issue

Article Menu

Article Overview

- Abstract
- Open Access and Permissions
- Share and Cite
- Article Metrics
- Related Articles
- Order Article Reprints

Open Access Article

Using Machine-Learning Algorithms for Eutrophication Modeling: Case Study of Mar Menor Lagoon (Spain)

by  Patricia Jimeno-Sáez ^{1,*}  ,  Javier Senent-Aparicio ¹ ,  José M. Cecilia ²  and  Julio Pérez-Sánchez ¹ 

¹ Department of Civil Engineering, Universidad Católica San Antonio de Murcia, Campus de los Jerónimos s/n, 30107 Guadalupe, Murcia, Spain

² Department of Computer Engineering, Universitat Politècnica de València, Camí de Vera, s/n, 46022 Valencia, Spain

* Author to whom correspondence should be addressed.

Int. J. Environ. Res. Public Health **2020**, *17*(4), 1189; <https://doi.org/10.3390/ijerph17041189>

Received: 16 January 2020 / Revised: 7 February 2020 / Accepted: 9 February 2020 / Published: 13 February 2020

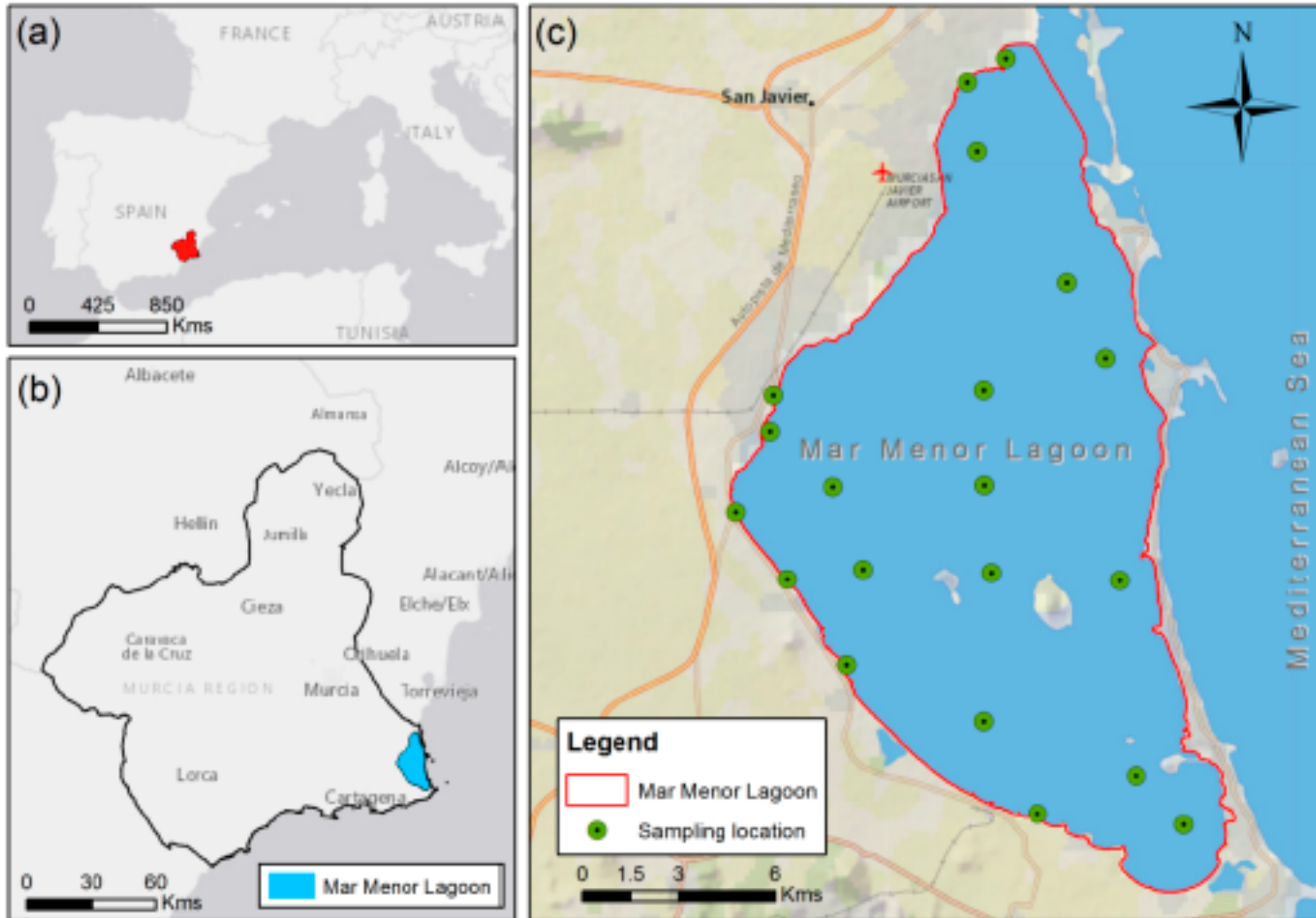
View Full-Text

Download PDF

Browse Figures

Cite This Paper

Área de estudio y recogida de datos



<http://www.canalmarmenor.es/web/canalmarmenor/parametros>

Enfoques de modelado y selección de características



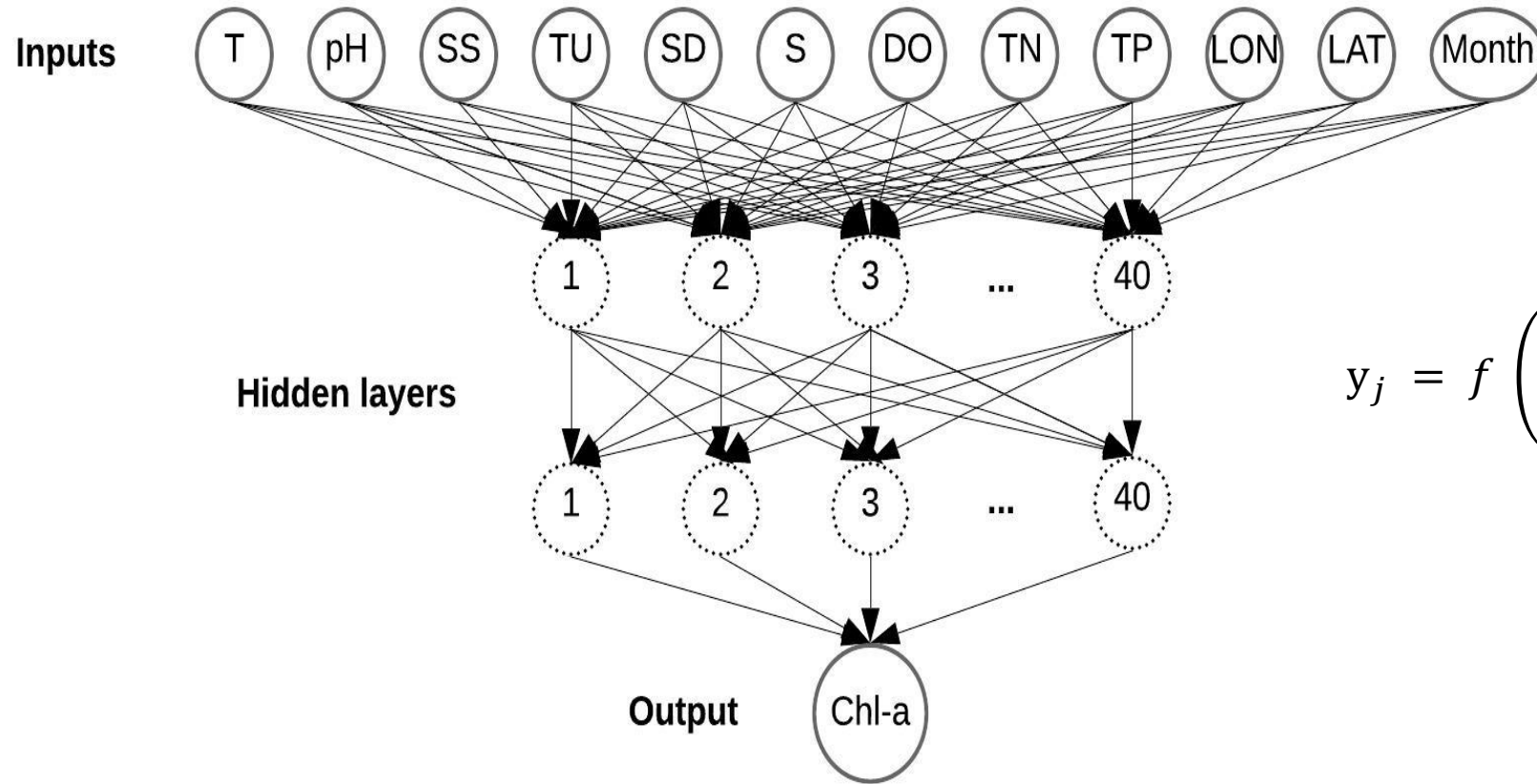
Análisis descriptivo

Parámetros	Xmax	Xmin	Xmean	St.Dev	C.V	CC
Chl-a(mg/m ³)	7.50	0.13	2.02	1.43	0.71	1.00
T(°C)	27.70	11.00	20.09	6.41	0.32	-0.08
pH	8.46	7.83	8.17	0.12	0.01	-0.02
SS (mg/l)	35.35	5.00	8.64	4.94	0.57	0.22
TU(NTU)	24.00	0.50	2.82	3.37	1.19	0.002
SD (m)	6.50	0.30	2.39	1.63	0.68	-0.55
S (PSU)	46.38	41.86	44.21	0.95	0.02	-0.35
DO(mg/l)	8.12	4.25	6.55	0.80	0.12	-0.17
TN (mg N/l)	8.84	0.16	0.59	0.80	1.35	0.09
TP (mg P/l) ^o	0.07	0.01	0.01	0.01	0.70	0.25

Selección de características

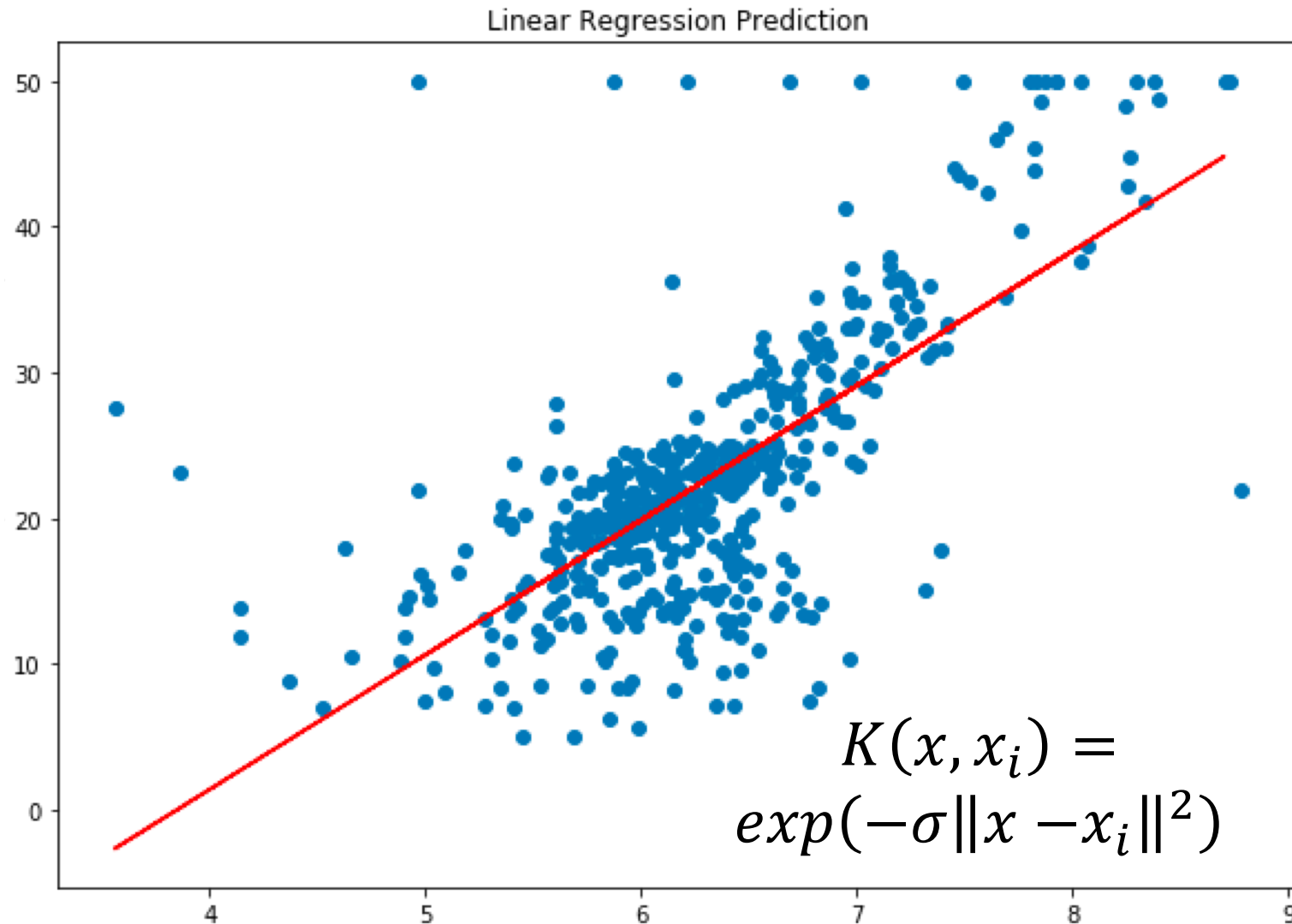
Algorithm	Nº SF	Selected features	Input scenario
Without feature selection	12	T, pH, SS, SD, S, DO, TU, TN, TP, LON, LAT, month	M1
The most highly correlated features	4	SD,S,TP,SS	M2
RFE_RF	9	SD, S, SS,TN, T, pH, DO, TP, LON	M3
RFE_SVM	9	SD, T, SS, S, TP, pH, TN, LON, TU	M4
GA_RF	9	LON, T, pH, SS, TU, SD, S, DO, TP	M5
GA_SVM	10	Month, LAT, LON, T, pH, TU, SD, DO, TN, TP	M6
SA_RF	8	LAT, LON, pH, TU, SD, S, DO, TN	M7
SA_SVM	7	LON, T, SS, TU, S, TN, TP	M8

Modelo: Artificial Neural Networks



$$y_j = f \left(\sum_{i=1}^n x_i \cdot w_i - b_j \right)$$

Modelo: Support Vector Machine



Evaluación del modelo

✓ 5-fold cross-validation

Performance metric	Equation	Range
R^2_{CV}	$\frac{[\sum_{i=1}^n (O_i - \bar{O}) \cdot (E_i - \bar{E})]^2}{\left[[\sum_{i=1}^n (O_i - \bar{O})^2]^{0.5} \cdot [\sum_{i=1}^n (E_i - \bar{E})^2]^{0.5} \right]^2}$	$[0, 1]$
$RMSE_{CV}$	$\sqrt{\frac{\sum_{i=1}^n (O_i - E_i)^2}{n}}$	$[0, \infty]$
MAE_{CV}	$\frac{\sum_{i=1}^n O_i - E_i }{n}$	$[0, \infty]$

Evaluación ANNs

Model-Input scenario	Architecture (I-H1-H2-O) ¹	Training Phase			Testing Phase		
		R^2_{cv}	RMSE _{cv} (mg/m ³)	MAE _{cv} (mg/m ³)	R^2_{cv}	RMSE _{cv} (mg/m ³)	MAE _{cv} (mg/m ³)
MLNN-M1	[12-16-27-1]	0.63 ± 0.14	0.85 ± 0.21	0.59 ± 0.21	0.53 ± 0.16	0.98 ± 0.31	0.70 ± 0.22
MLNN-M2	[9-31-23-1]	0.67 ± 0.17	0.80 ± 0.25	0.54 ± 0.15	0.60 ± 0.23	0.89 ± 0.41	0.66 ± 0.24
MLNN-M3	[9-32-39-1]	0.72 ± 0.07	0.76 ± 0.14	0.50 ± 0.06	0.61 ± 0.16	0.89 ± 0.35	0.63 ± 0.18
MLNN-M4	[9-40-39-1]	0.74 ± 0.09	0.73 ± 0.17	0.49 ± 0.13	0.62 ± 0.13	0.89 ± 0.26	0.66 ± 0.14
MLNN-M5	[9-40-33-1]	0.72 ± 0.13	0.74 ± 0.20	0.51 ± 0.20	0.55 ± 0.21	0.99 ± 0.34	0.76 ± 0.19
MLNN-M6	[8-39-16-1]	0.72 ± 0.08	0.78 ± 0.15	0.57 ± 0.13	0.54 ± 0.16	0.96 ± 0.32	0.71 ± 0.18
MLNN-M7	[7-25-30-1]	0.56 ± 0.16	0.93 ± 0.19	0.66 ± 0.17	0.49 ± 0.11	1.07 ± 0.18	0.74 ± 0.10

Evaluación SVR

Model-Input scenario	Model parameters	Training Phase			Testing Phase		
		R^2_{cv}	RMSE _{cv} (mg/m ³)	MAE _{cv} (mg/m ³)	R^2_{cv}	RMSE _{cv} (mg/m ³)	MAE _{cv} (mg/m ³)
SVR-M1	$\sigma = 0.07$	0.53 ±	0.99 ±	0.70 ±	0.56 ±	0.85 ±	0.62 ±
	C = 3.31	0.05	0.07	0.03	0.09	0.26	0.14
SVR-M2	$\sigma = 0.13$	0.56 ±	0.96 ±	0.66 ±	0.65 ±	0.82 ±	0.58 ±
	C = 2.54	0.05	0.12	0.05	0.11	0.30	0.16
SVR-M3	$\sigma = 0.10$	0.58 ±	0.94 ±	0.66 ±	0.68 ±	0.81 ±	0.56 ±
	C = 2.59	0.06	0.10	0.04	0.10	0.32	0.17
SVR-M4	$\sigma = 0.10$	0.58 ±	0.92 ±	0.64 ±	0.67 ±	0.82 ±	0.57 ±
	C=3.01	0.05	0.09	0.04	0.09	0.30	0.18
SVR-M5	$\sigma = 0.10$	0.52 ±	0.98 ±	0.69 ±	0.61 ±	0.86 ±	0.61 ±
	C=4.20	0.09	0.11	0.06	0.16	0.31	0.15
SVR-M6	$\sigma = 0.13$	0.51 ±	1.03 ±	0.72 ±	0.57 ±	0.90 ±	0.64 ±
	C=3.38	0.10	0.15	0.08	0.20	0.37	0.21
SVR-M7	$\sigma = 0.30$	0.47 ±	1.06 ±	0.74 ±	0.54 ±	0.94 ±	0.66 ±
	C = 2.63	0.05	0.07	0.04	0.13	0.21	0.13

Conclusiones

- ✓ Necesitamos entender la interacción entre los parámetros de calidad del agua para adoptar estrategias de gestión sostenibles.
- ✓ Los métodos tradicionales de regresión no dan buenos resultados en variables con tan baja correlación y no-lineales.
- ✓ Con R^2 de 0,7 como punto de partida, demuestra que los métodos de ML pueden ser una buena alternativa para predecir la Chl-a.
- ✓ Con métodos de medición más precisos y continuos, se podrían hacer mejores predicciones a distintas escalas temporales.

Agradecimientos



Agencia de Ciencia y Tecnología
Región de Murcia

