# Groundwater withdrawals prediction in semi-arid basins using machine learning algorithms and integrated water management models

**Kiara Tesén** [1,2] **; Francisco Suárez** [1]

[1] Pontificia Universidad Católica de Chile, Departamento de Ingeniería Hidráulica y Ambiental, Santiago, Chile
[2] Universidad de Piura, Instituto de Hidráulica, Hidrología e Ingeniería Sanitaria, Piura, Perú

# Integrated Management of Water Resources **(IMWR)**

problems related to

- Resources availability
- Institutional capacities
- Represent systems through **models**

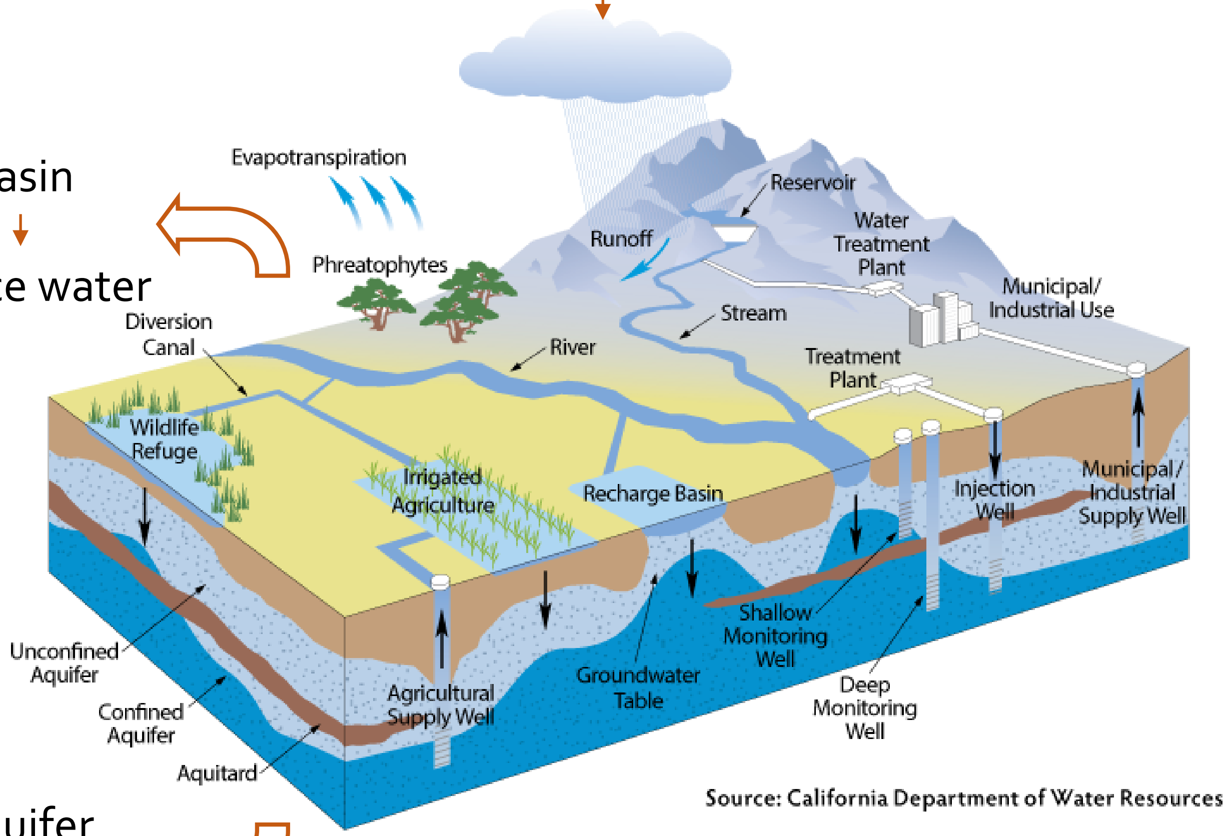
Problems?

Lack of data

**Groundwater models that fail to capture the aquifer dynamics**

Basin

Surface water




Source: California Department of Water Resources

Aquifer

Groundwater

## Machine Learning Algorithms

Supervised Learning

Regression

Analysis of continuous variables

- Lineal Regression Models
- Regression Trees
- Support Vector Machines Regression (SVM)
- Gaussian Process Regression Model (GPR)
- Kernel Approximation Regression Models
- Ensembles of Trees
- Artificial Neural Networks (ANN)

- Streamflow prediction
- Rain forecast
- Short-term groundwater level prediction

(Wang, et al., 2006;
Lange y Sippel, 2020;
Band et al., 2021)

Classification

# Objective

Estimate groundwater withdrawals using machine learning algorithms and integrated water management models.

# Part I - Integrated model of water resources

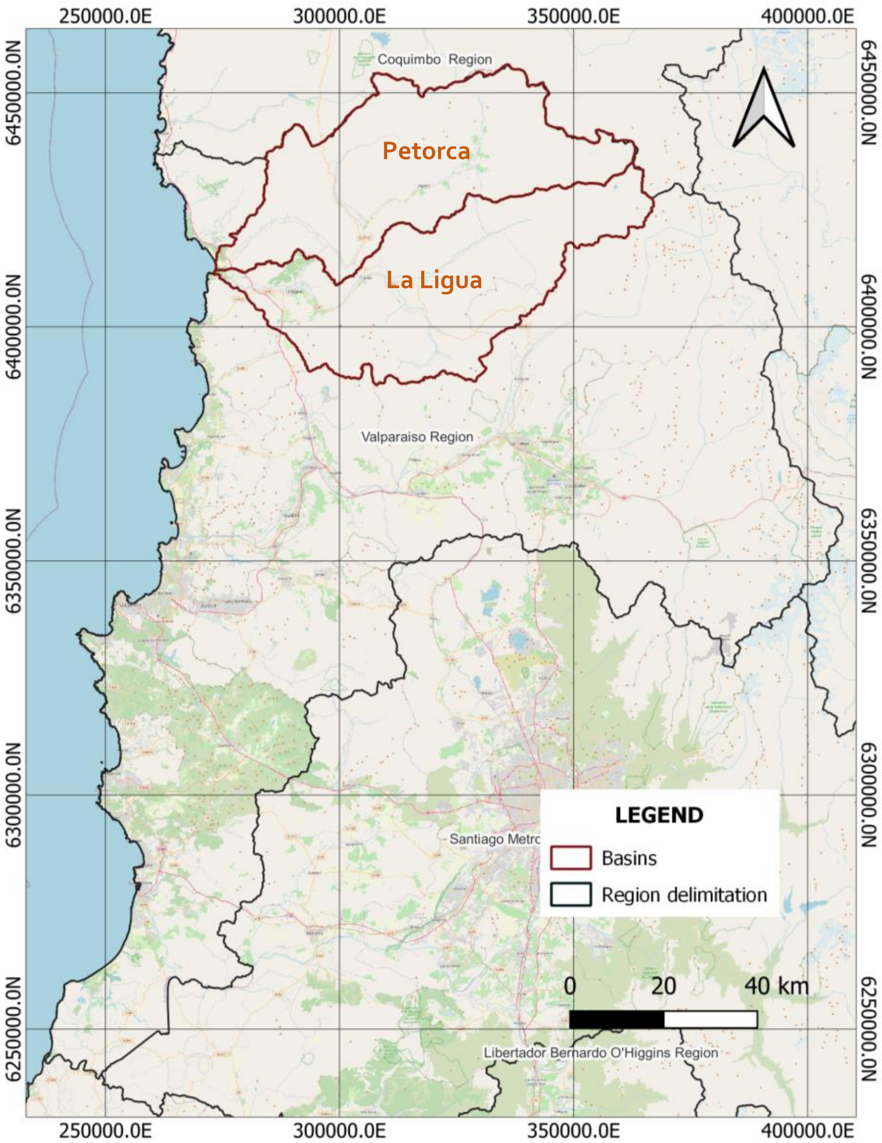
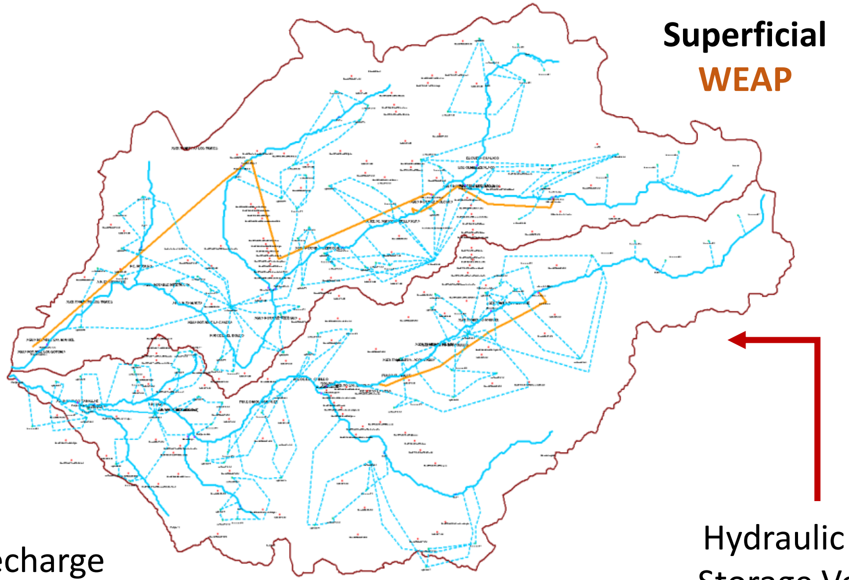



Figure: Pilot basins.





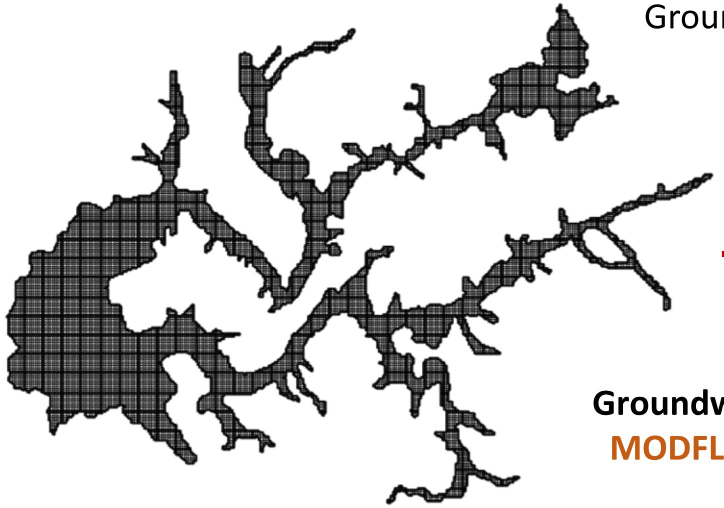




Figure: Schematic configuration of the integrated model of WEAP-MODFLOW.

# Part II – Machine Learning algorithms - Training

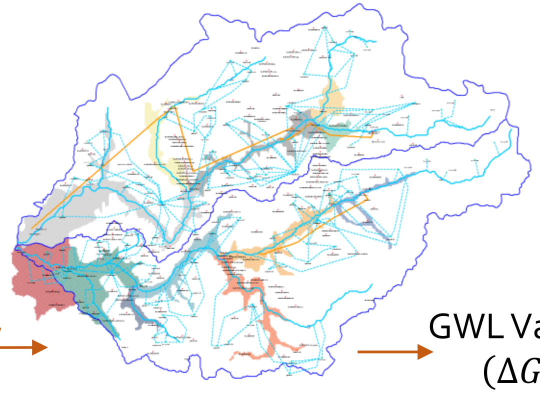
**Obtaining data within WEAP – MODFLOW model**

*Demand sites*

Water Use Rights

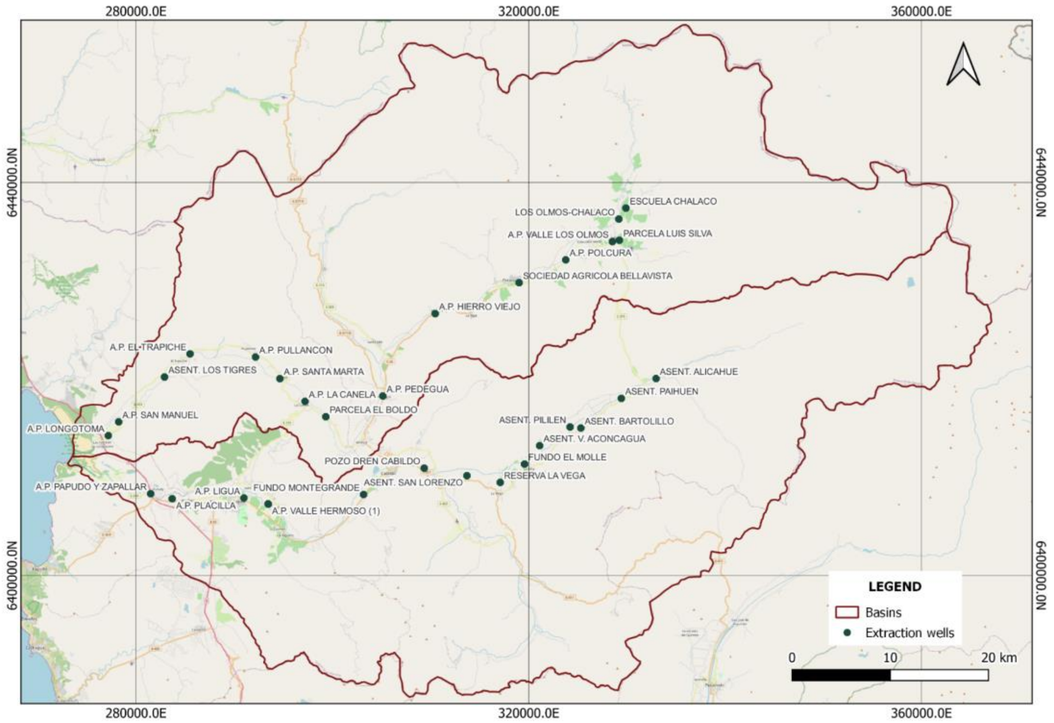
Demand Variations ($\Delta$WUR)

WEAP – MODFLOW Execution



GWL Variations ($\Delta GWL$)

**Output** ← **Input**




**Training process**

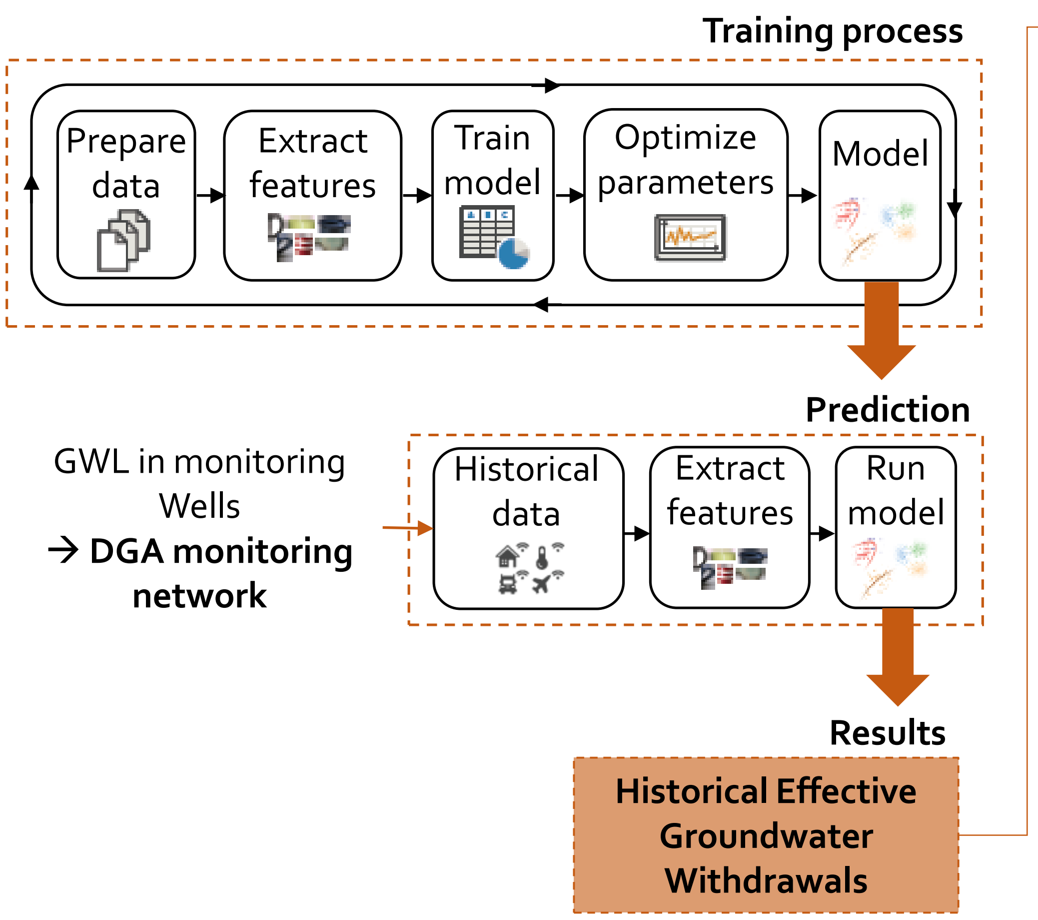
Prepare data → Extract features → Train model → Optimize parameters → Model

**Metrics:**

- $R^2$
- RMSE
- MSE
- MAE

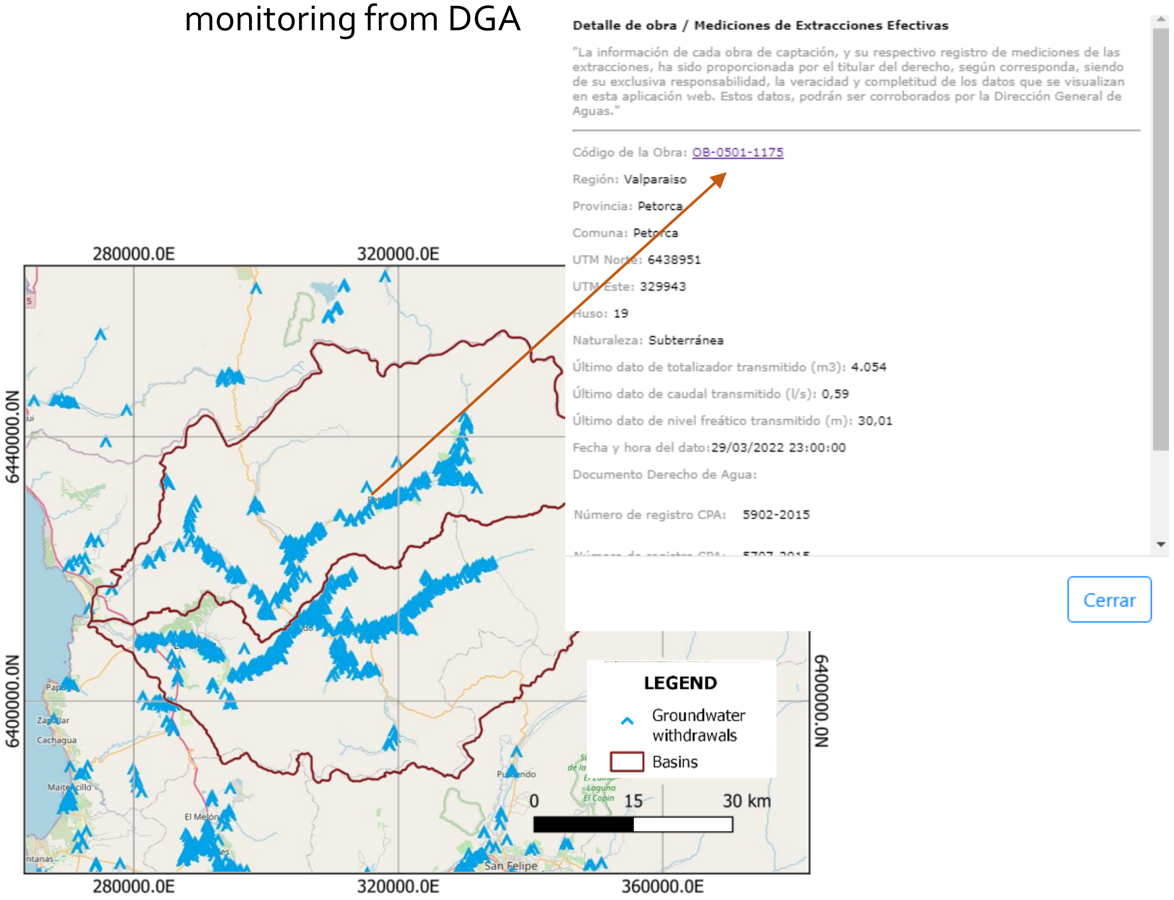
- *Gaussian Process Regression Model* (*GPR*)
- *Artificial Neural Networks* (*ANN*)

# Part II – Machine Learning algorithms - Prediction

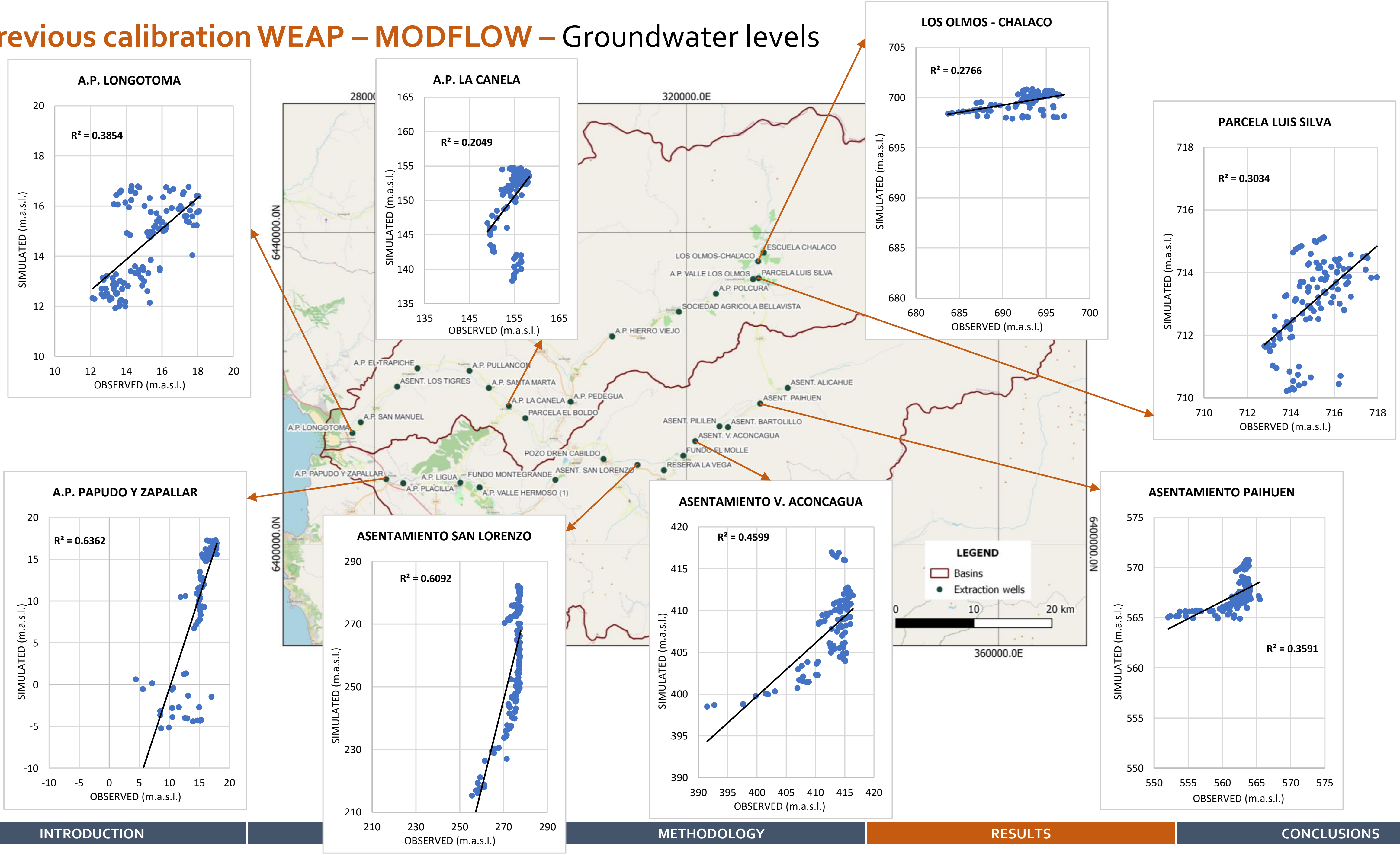



**Calibration →** WEAP – MODFLOW model
Historic period: 1980 – 2021

**Validation →** Record of effective extraction monitoring from DGA



# Previous calibration WEAP – MODFLOW – Groundwater levels



# Metrics – Training algorithms

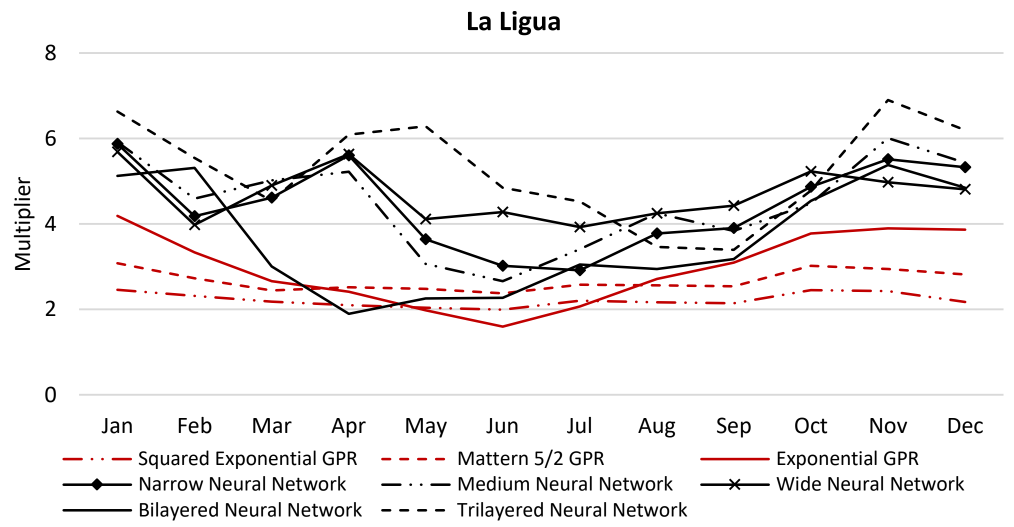
Table. Metrics obtained from the training algorithms process in **La Ligua** river basin.

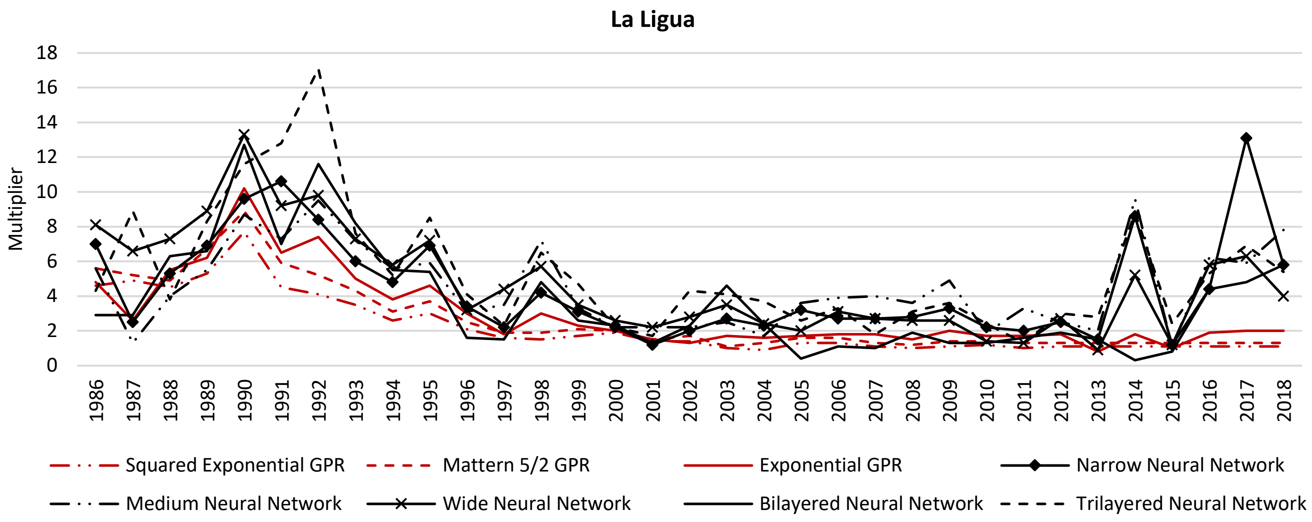
| Algorithm | Configuration | Training results | | | |
|---|---|---|---|---|---|
| | | RMSE | $R^2$ | MSE | MAE |
| GPR model | Squared Exponential GPR | 0.3768 | 0.95 | 0.1420 | 0.2318 |
| | Matern 5/2 GPR | 0.3426 | 0.96 | 0.1174 | 0.1977 |
| | Exponential GPR | 0.2895 | 0.97 | 0.0838 | 0.1342 |
| Artificial Neural Network | Narrow Neural Network | 0.4357 | 0.93 | 0.1898 | 0.2841 |
| | Medium Neural Network | 0.3999 | 0.94 | 0.1599 | 0.2514 |
| | Wide Neural Network | 0.3627 | 0.95 | 0.1316 | 0.2235 |
| | Bilayered Neural Network | 0.3961 | 0.94 | 0.1569 | 0.2518 |
| | Trilayered Neural Network | 0.3840 | 0.95 | 0.1474 | 0.2387 |

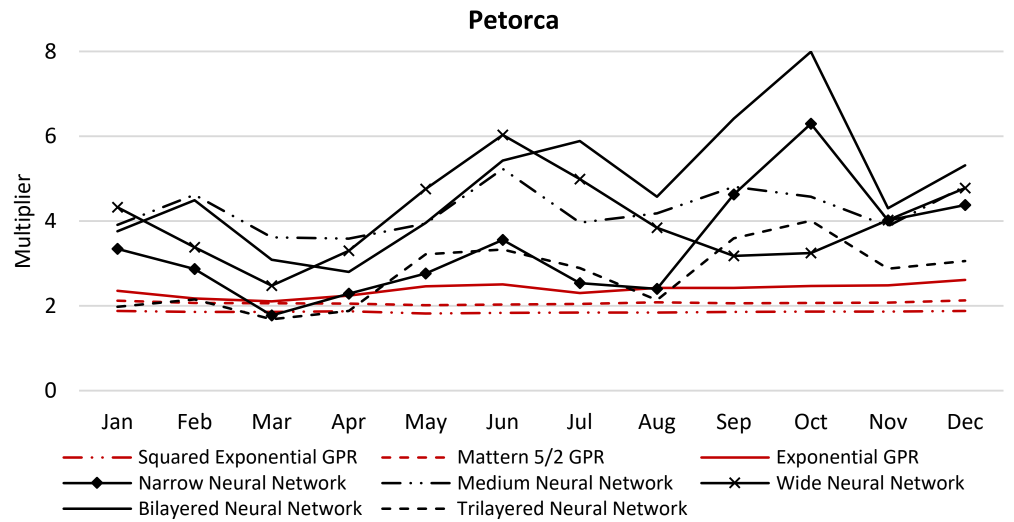
Table. Metrics obtained from the training algorithms process in **Petorca** river basin.

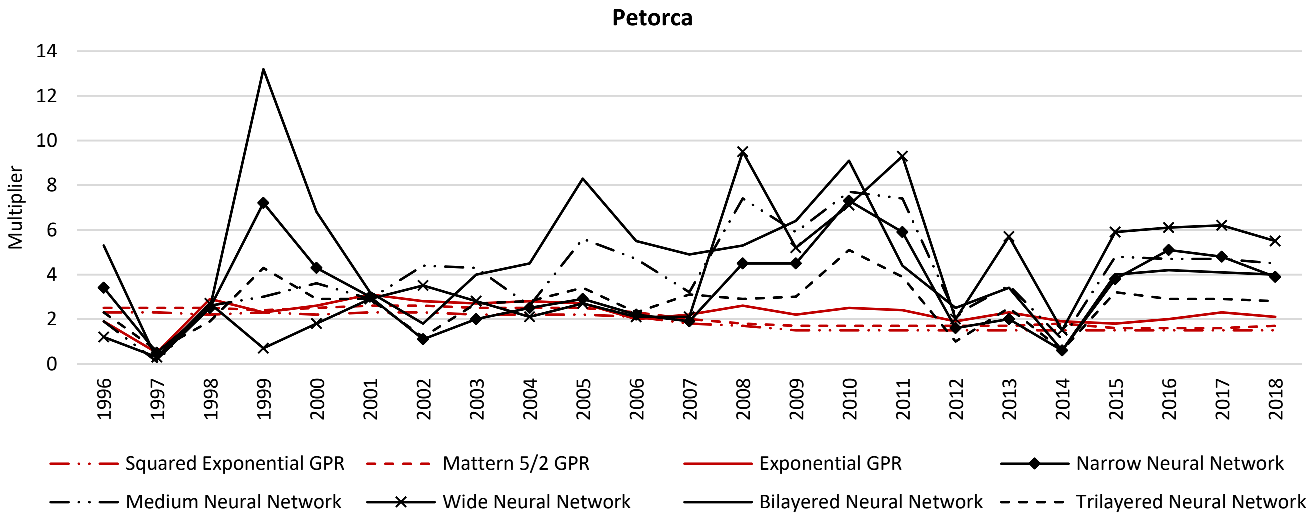
| Algorithm | Configuration | Training results | | | |
|---|---|---|---|---|---|
| | | RMSE | $R^2$ | MSE | MAE |
| GPR model | Squared Exponential GPR | 0.1603 | 0.98 | 0.0257 | 0.0749 |
| | Matern 5/2 GPR | 0.1341 | 0.99 | 0.0180 | 0.0599 |
| | Exponential GPR | 0.1311 | 0.99 | 0.0172 | 0.0618 |
| Artificial Neural Network | Narrow Neural Network | 0.2678 | 0.94 | 0.0717 | 0.1732 |
| | Medium Neural Network | 0.2035 | 0.97 | 0.0414 | 0.1306 |
| | Wide Neural Network | 0.1639 | 0.98 | 0.0268 | 0.1011 |
| | Bilayered Neural Network | 0.2103 | 0.96 | 0.0442 | 0.1329 |
| | Trilayered Neural Network | 0.2105 | 0.96 | 0.0443 | 0.1302 |

# **Groundwater withdrawals prediction -** Effective groundwater withdrawals predictions show an increase compared to assigned water rights.

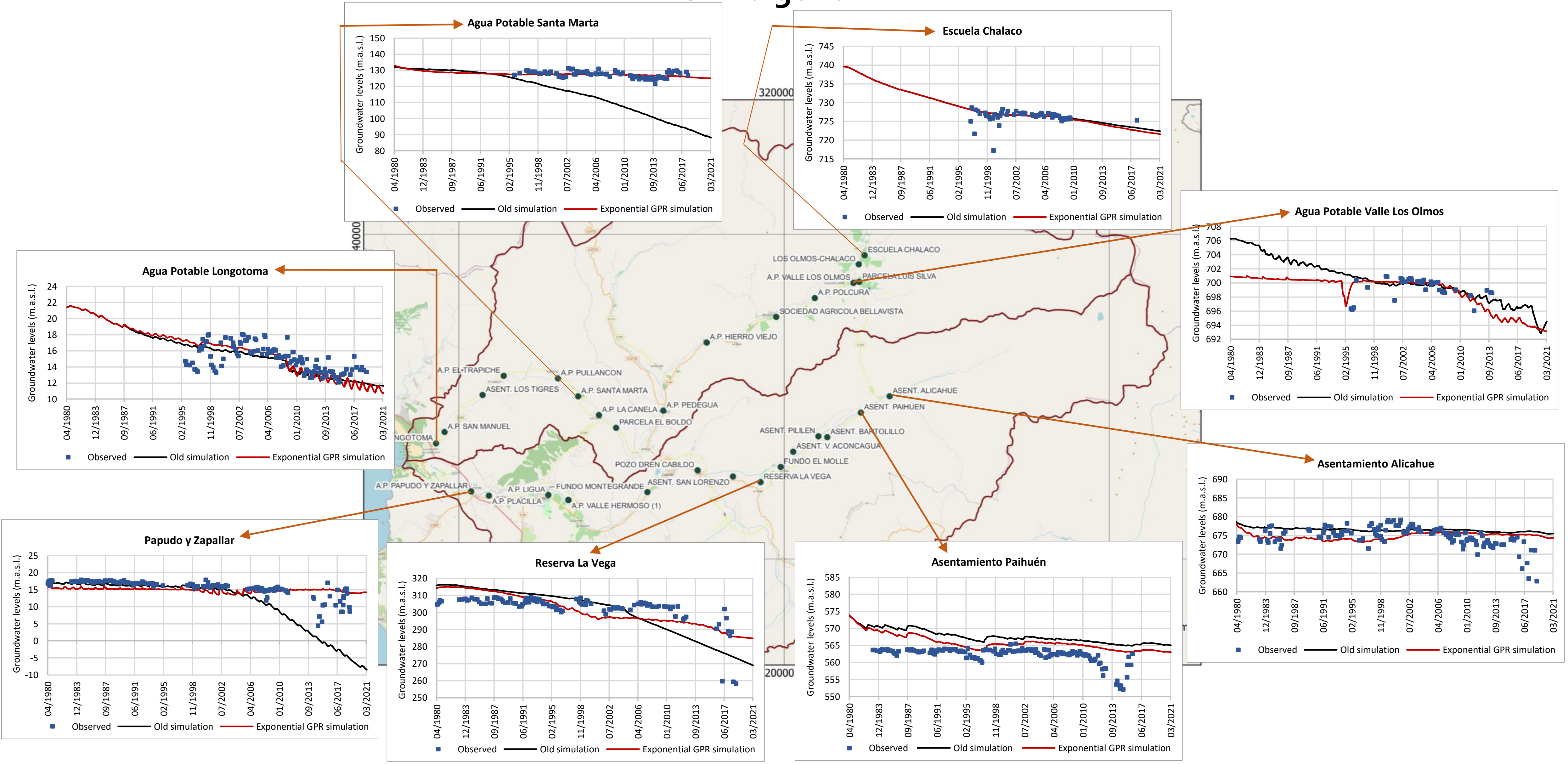


# **Re-calibration WEAP – MODFLOW model –** Effective groundwater withdrawals obtained from **Exponential GPR algorithm**



# Conclusions

- Our results show that the **GPR model**, compared to ANNs, adequately estimates the spatiotemporal distribution of groundwater withdrawals in the pilot basins.
- Thus, the use of machine learning algorithms improves the performance of integrated water resources management models.

- The initial years show a high increase in the effective groundwater withdrawals compared with the water rights in La Ligua river basin. This could be because, with time, the water rights registration has been improved.
- We can assume or deduce that the increase calculated in the effective groundwater withdrawals for the algorithms is due to the illegal extractions in those areas or to temporal water rights awarded from the DGA.

# Next steps

- Disaggregated analysis by aquifer sectors.

- Application of other machine learning algorithms, e.g., Support Vector Machines Regression, Linear Regression Models, among others.

- Improve the based model with more observed data and with the support of the stakeholders (water users).

# Groundwater withdrawals prediction in semi-arid basins using machine learning algorithms and integrated water management models

PONTIFICIA UNIVERSIDAD CATÓLICA DE CHILE

UNIVERSIDAD DE PIURA

**Kiara Tesén Arámbulo**

katesen@uc.cl
kiara.tesen@udep.edu.pe

# **Annex –** Algorithms configuration

Table. GPR model

| Algorithm | Configuration | Model Hyperparameters | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| | | Basis function | Kernel function | Use isotropic kernel | Kernel scale | Signal standard deviation | Sigma | Standardize | Optimize numeric parameters |
| GPR model | Rational Quadratic GPR | Constant | Rational Quadratic | true | Automatic | Automatic | Automatic | true | true |
| | Squared Exponential GPR | Constant | Squared Exponential | true | Automatic | Automatic | Automatic | true | true |
| | Matern 5/2 GPR | Constant | Matern 5/2 | true | Automatic | Automatic | Automatic | true | true |
| | Exponential GPR | Constant | Exponential | true | Automatic | Automatic | Automatic | true | true |

Table. Artifitial Neural Network

| Algorithm | Configuration | Model Hyperparameters | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| | | Number of fully connected layers | First layer size | Second layer size | Third layer size | Activation | Iteration limit | Regularization strength (Lambda) | Standardize data |
| Artificial Neural Network | Narrow Neural Network | 1 | 10 | | | ReLU | 1000 | 0 | Yes |
| | Medium Neural Network | 1 | 25 | | | ReLU | 1000 | 0 | Yes |
| | Wide Neural Network | 1 | 100 | | | ReLU | 1000 | 0 | Yes |
| | Bilayered Neural Network | 2 | 10 | 10 | | ReLU | 1000 | 0 | Yes |
| | Trilayered Neural Network | 3 | 10 | 10 | 10 | ReLU | 1000 | 0 | Yes |