

MODELOS PREDICTIVOS AVANZADOS PARA LA GESTIÓN INTELIGENTE DEL CONSUMO DE AGUA POTABLE

ADVANCED PREDICTIVE MODELS FOR THE SMART MANAGEMENT OF DRINKING WATER CONSUMPTION

José Hernández^a, Silvia Simón^a, Carlos Carossio^a, Gabriela Minetti^b, Carolina Salto^b y Mercedes Carnero^a

^a*Universidad Nacional de Río Cuarto, Facultad de Ingeniería. Río Cuarto, Argentina,
mcarnero@ing.unrc.edu.ar*

^b*CONICET-Universidad Nacional de La Pampa, Facultad de Ingeniería. General Pico, Argentina,
minettig@ing.unlpam.edu.ar*

Palabras clave: Aprendizaje automático, Series temporales, Sistemas de distribución de agua, Bosque aleatorio, Prophet

Resumen. La predicción precisa del consumo de agua potable constituye un desafío clave para las distribuidoras, especialmente en contextos donde la medición manual continúa siendo el método predominante. En muchas regiones de Argentina, las cooperativas aún dependen de estos registros, lo que introduce errores en la facturación, dificulta la planificación operativa y genera conflictos con los usuarios. Este trabajo aborda dicha problemática mediante la evaluación y comparación de tres metodologías de pronóstico: un método estadístico clásico (Medias Móviles), un modelo de aprendizaje automático de conjunto (Bosque Aleatorio o Random Forest) y un procedimiento de series temporales especializado (Prophet). Los resultados demuestran que los modelos basados en aprendizaje automático, como Bosque Aleatorio, reducen de manera significativa los errores de predicción respecto a los métodos tradicionales. Se concluye que la adopción de estas herramientas aporta información valiosa para optimizar la toma de decisiones estratégicas en la gestión cooperativa del agua.

Keywords: Machine Learning, Time series, Water distribution systems, Random Forest, Prophet

Abstract. Accurate prediction of drinking water consumption is a key challenge for utility providers, especially in contexts where manual metering remains the predominant method. In many regions of Argentina, cooperatives still rely on these records, which leads to billing errors, hinders operational planning, and creates conflicts with users. This work addresses this issue by evaluating and comparing three forecasting methodologies: a classic statistical method (Moving Averages), an ensemble machine learning model (Random Forest), and a specialized time-series procedure (Prophet). The results show that machine learning-based models, such as Random Forest, significantly reduce prediction errors compared to traditional methods. It is concluded that the adoption of these tools provides valuable insights to optimize strategic decision-making in cooperative water management.

INTRODUCCIÓN

La facturación precisa del consumo de agua potable es un pilar fundamental en la relación entre las cooperativas distribuidoras y sus usuarios. Sin embargo, en gran parte de Argentina, la dependencia de lecturas manuales de medidores introduce una fuente de error recurrente. Estas imprecisiones no solo derivan en errores de facturación y los consecuentes reclamos por parte de los usuarios, sino que también limitan la capacidad de la organización para una planificación operativa eficiente.

Para mitigar estos problemas, la adopción de modelos predictivos que estimen el consumo esperado de cada usuario se presenta como una solución estratégica. Al comparar una predicción fiable con la lectura manual registrada, es posible identificar anomalías, detectar posibles errores de lectura antes de emitir la factura y, en definitiva, minimizar los conflictos.

Los enfoques estadísticos clásicos, como las medias móviles, son útiles para identificar tendencias simples, pero demuestran ser insuficientes frente a dinámicas de consumo no lineales o al no poder incorporar variables externas influyentes (ej. temperatura, estacionalidad). En este contexto, las técnicas de Aprendizaje Automático (Machine Learning, ML) y los modelos especializados en series temporales surgen como alternativas con un potencial significativamente mayor.

El presente trabajo tiene como objetivo principal evaluar y comparar el desempeño de tres enfoques distintos para el pronóstico del consumo mensual a nivel de usuario domiciliario: (i) un método estadístico tradicional de Medias Móviles, (MA); (ii) un modelo de aprendizaje de conjunto, Bosque Aleatorio (Random Forest, RF); y (iii) un modelo especializado en series temporales, Prophet, (PH). El análisis busca determinar la viabilidad de estos modelos como herramientas para la validación de lecturas y la reducción proactiva de errores de facturación, utilizando para ello un conjunto de datos reales de una cooperativa de servicios públicos del centro de Argentina.

1 MARCO TEÓRICO Y ANTECEDENTES

La predicción de la demanda de agua es un área de investigación consolidada y de vital importancia para la gestión eficiente de los recursos hídricos. La literatura académica ha abordado este problema desde múltiples perspectivas, evolucionando desde modelos estadísticos clásicos hacia técnicas avanzadas de aprendizaje automático. A pesar de la extensa investigación en la predicción de la demanda de agua a nivel de sistema o ciudad, se identifican en este trabajo dos aportes principales. El primero, relacionado con la escala de análisis: la mayoría de los estudios se centran en la demanda agregada de una red o un sector, y no en la predicción a escala de usuario domiciliario individual, que es fundamental para la gestión de la facturación y minimización de reclamos. El segundo aspecto es el contexto de aplicación operativo específico de las cooperativas de servicios públicos, donde la lectura manual de medidores sigue siendo una práctica extendida y una fuente importante de errores y conflictos.

Los enfoques Box-Jenkins, del tipo Modelo Autorregresivo Integrado de Media Móvil (o modelos ARIMA por sus siglas en inglés) y sus extensiones han sido la base histórica para demanda de agua por su capacidad de capturar autocorrelación y componentes estocásticos bajo supuestos de linealidad y cuasi estacionariedad (Box et al, 2015). Cuando la estacionalidad es fuerte y el patrón es relativamente estable, los modelos SARIMA (Seasonal ARIMA o ARIMA estacional), suelen rendir bien, aunque su desempeño se degrada frente a no linealidades, cambios de régimen y choques exógenos no modelados explícitamente. La construcción de modelos y el diagnóstico (identificación, estimación y validación) siguen las guías consolidadas en la literatura clásica de series temporales. Dentro de esta categoría se encuentran también métodos más sencillos como el suavizado exponencial y las MA. Si bien estos modelos son fáciles de implementar e interpretar, su principal limitación es que asumen que los patrones del

pasado continuarán en el futuro de manera lineal, y tienen dificultades para incorporar variables externas (exógenas) que influyen en el consumo, como las condiciones climáticas. Son considerados, en la literatura reciente, como un punto de referencia (baseline) para evaluar el rendimiento de modelos más complejos.

Con el aumento de la capacidad computacional y la disponibilidad de datos, los modelos de ML han ganado un enorme protagonismo. A diferencia de los modelos clásicos, estas técnicas son capaces de capturar relaciones no lineales complejas entre múltiples variables. Las Redes Neuronales Artificiales (ANN) y Máquinas de Vectores de Soporte (SVM) han demostrado ser eficaces para predecir la demanda de agua a corto y mediano plazo, especialmente al incluir variables como la temperatura, la precipitación y el día de la semana (Ghiassi et al., 2008).

Los modelos basados en árboles de decisión, como RF y Gradient Boosting (GB), son particularmente valorados por su habilidad para modelar relaciones no lineales e integrar covariables (p. ej., temperatura, precipitaciones, feriados, indicadores socioeconómicos). La literatura destaca que RF es robusto frente al sobreajuste y puede manejar un gran número de variables de entrada, identificando automáticamente las más relevantes. Su aplicación en la predicción de consumo ha mostrado resultados superiores a los métodos estadísticos en contextos con alta variabilidad (Brentan et al., 2017). El consenso general en la investigación es que los modelos de ML superan a los tradicionales en precisión, siempre que se disponga de datos de calidad y de variables explicativas relevantes. Revisiones recientes del área reportan ventajas de ML frente a ARIMA cuando hay no linealidades marcadas y múltiples variables exógenas, especialmente en horizontes cortos y medios (Ghannam & Hussain, 2024).

En los últimos años han surgido modelos específicamente diseñados para facilitar el pronóstico de series temporales, los cuales buscan combinar la robustez estadística con la flexibilidad de ML. Uno de los ejemplos paradigmáticos es PH (Taylor y Letham, 2018), desarrollado por *Facebook*, que se ha consolidado como una herramienta ampliamente utilizada en distintos campos de aplicación. Este modelo está concebido para trabajar con series que presentan fuertes componentes estacionales (por ejemplo, anuales, mensuales o semanales), mostrando además una notable robustez frente a datos faltantes y frente a cambios abruptos de tendencia. La literatura lo describe como un recurso potente y accesible, ya que automatiza gran parte de los pasos complejos del modelado de series temporales sin exigir una formación estadística avanzada por parte del usuario.

PH se basa en un marco aditivo y descomponible, que divide la serie en tres componentes principales: tendencia, estacionalidad y efectos de calendario (como feriados o eventos locales). Este enfoque permite que los resultados sean interpretables y que el modelo pueda ajustarse con relativa facilidad, incluso en contextos donde existen puntos de cambio (changepoints) en la dinámica de la serie. PH ha sido ampliamente adoptado en diferentes contextos de predicción de demanda. En el sector energético, su combinación con redes neuronales de memoria a corto y largo Plazo, ha permitido mejorar significativamente las métricas predictivas, (Albahli, 2025). En contextos industriales y de salud, PH fue parte esencial en modelos de planificación avanzada y demanda farmacéutica sensible al clima (Nadal et al, 2025; Schisa & Farnè, 2025)

Aunque no fue diseñado de manera específica para el dominio hídrico, su arquitectura modular lo hace particularmente útil en este campo, ya que admite la incorporación de variables exógenas relevantes. Asimismo, su desempeño predictivo competitivo, junto con la transparencia en la interpretación de sus resultados, ha llevado a que se emplee con frecuencia como punto de referencia metodológico (benchmark) frente a otros modelos más complejos.

La literatura revisada revela que, si bien la predicción de la demanda agregada es un tema ampliamente cubierto, la aplicación de estos modelos a escala de usuario individual y para el contexto operativo de las cooperativas representa un área de especial interés y menor exploración. El presente estudio se concentra en este dominio específico, aplicando y

comparando un modelo clásico, uno de aprendizaje automático y uno especializado como herramienta para la gestión de la facturación.

2 MATERIALES Y MÉTODOS

Esta sección detalla el conjunto de datos utilizado, los modelos de pronóstico implementados y la estrategia de evaluación empleada para comparar su desempeño en la predicción del consumo de agua domiciliario.

2.1 Descripción del conjunto de datos

Para la realización de este trabajo se cuenta con registros históricos de consumo mensual de los usuarios de una cooperativa en la región sur de la provincia de Córdoba, Argentina abarcando el período que va desde febrero de 2019 hasta agosto de 2025. Se organiza la información en un conjunto de datos almacenado en un archivo de valores separados por comas (*comma separated values, csv*) en cinco columnas cuya descripción se muestra en la tabla 1.

#	Nombre	Descripción	Tipo
1	Usuario	Identif del usuario	Entero
2	Año	Año de la lectura	Entero
3	Mes	Mes de la lectura	Entero
4	Consumo	Consumo [m ³]	Entero
5	Temperatura	Temperatura [°C]	Real

Tabla 1: Descripción del conjunto de datos.

Usuario se refiere a la identificación de cada usuario dentro de la cooperativa. La fecha del período considerado queda determinada por el mes y el año de la lectura. El consumo es calculado como simple diferencia entre lecturas y expresado en metros cúbicos. La última columna corresponde a la temperatura promedio para cada período medido y es una variable exógena para el modelo que se expresa en °C.

2.2 Preprocesamiento de los datos

Previo al análisis, es necesario llevar a cabo un proceso de transformación de los datos, cuyo objetivo es depurar y estandarizar los registros disponibles en la base. Esta etapa permite convertir la información en bruto en un conjunto de datos limpio, coherente y estructurado, garantizando su idoneidad para el modelado y el análisis posterior. Se ordenan cronológicamente para asegurar la correcta estructura de la serie temporal, se realiza una verificación del tipo de datos y se calcula algunas estadísticas básicas para cada columna numérica que permita detectar rápidamente valores atípicos o absurdos tales como valores negativos de consumo, valor de mes fuera de rango, temperaturas fuera de los valores aceptados en la región. Posteriormente se debe proceder al tratamiento de datos faltantes: en este trabajo no se permite faltantes de temperatura, ya que se han obtenido a partir de una fuente meteorológica externa, y en el caso de consumos se imputa el faltante por interpolación entre los valores anterior y posterior al dato faltante. Finalmente se quitan algunos datos atípicos identificados como claros errores de lecturas o registro. Para detectarlos se utilizan dos métodos: en el caso de procesamiento de un usuario individual se utiliza un método visual, graficando el consumo a lo largo del tiempo y observando los picos anómalos. Para procesar un grupo de usuarios se utiliza el rango intercuartílico (IQR) que considera un valor atípico si está por encima de Q3+1.5 IQR o por debajo de Q1-1.5 IQR. Los valores atípicos se reemplazan

por los valores máximos o mínimos “creíbles”. Se considera válido el consumo mínimo en 0 y se reemplaza un valor atípico por exceso por el percentil 99 del usuario.

2.3 Modelos implementados

Medias Móviles. Se implementó MA como línea de base para representar un enfoque estadístico simple y que es normalmente utilizado como método para estimar consumos en las cooperativas que deriva de una práctica común en el caso de lectura de medidores de energía eléctrica de usuarios rurales en los cuales es común reemplazar el consumo para un período dado, tomando el promedio de los últimos 3 o 6 meses. En este trabajo se utiliza varias configuraciones de ventana (3, 6 y 12 meses), donde la predicción para el mes siguiente corresponde al promedio del consumo en la configuración de ventana adoptada.

Bosque Aleatorio. Se utiliza un modelo de RF de la librería scikit-learn, representativo de los enfoques de aprendizaje automático. El modelo es entrenado utilizando como variables predictoras (features) el año, el mes y la temperatura, con el consumo como variable objetivo (target). Se emplea una partición cronológica, reservando los últimos 12 meses de datos como conjunto de prueba (test set). Se configura el modelo con 200 estimadores.

Prophet. Finalmente, se implementa PH, un modelo especializado en series temporales desarrollado por Facebook. El modelo se ajustó para un usuario individual, configurado para capturar la estacionalidad anual (`yearly_seasonality=True`) y la tendencia subyacente de la serie de consumo. Se genera un pronóstico a un mes de horizonte.

2.4 Estrategia de Evaluación y Métricas de Error

Para evaluar y comparar cuantitativamente la precisión de los modelos, se calculan dos métricas de error estándar sobre el conjunto de prueba (los últimos 12 meses). En primer lugar, se usa el Error Absoluto Medio (MAE) que mide el promedio de los errores absolutos entre los valores reales y los predichos puesto que es fácil de interpretar, ya que se expresa en las mismas unidades que la variable (en este caso, metros cúbicos). También se utiliza la Raíz del Error Cuadrático Medio (RMSE) que permite penalizar en mayor medida los errores grandes al elevarlos al cuadrado antes de promediálos. También se expresa en las mismas unidades que la variable.

3 EXPERIMENTOS Y RESULTADOS

En esta sección se presentan los resultados obtenidos al aplicar la metodología descrita sobre el conjunto de datos de prueba. El objetivo es evaluar y comparar el desempeño de los tres modelos de pronóstico, de manera objetiva.

Se ha tomado, como ejemplo los datos en el período descrito para un usuario y se exponen los resultados cuantitativos a través de las métricas de error MAE y RMSE para establecer una comparación directa de la precisión predictiva. A continuación, se realiza un análisis visual mediante gráficos que ilustran el comportamiento de cada modelo frente a los datos reales de consumo.

3.1 Desempeño Predictivo y Métricas de Error

Para evaluar y comparar de manera objetiva la precisión de los modelos, se calculan las métricas MAE y RMSE sobre el conjunto de datos de prueba. Estos valores cuantifican el error promedio de predicción, donde valores más bajos indican un mejor desempeño.

La Tabla 2 resume los resultados de este análisis para cinco configuraciones de modelos: tres variantes del método de MA (con ventanas de 3, 6 y 12 meses, respectivamente), el modelo

RF y PH.

Modelo	MAE	RMSE
MA (ventana=3)	1.34	1.93
MA (ventana=6)	1.50	2.02
MA (ventana=12)	1.29	1.76
RF	0.60	0.83
PH	1.08	1.43

Tabla 2: Comparación de las métricas para las 3 metodologías.

De acuerdo con los resultados obtenidos, se observa que el modelo RF se destaca como el de mayor precisión, obteniendo los valores más bajos en ambas métricas de error. Registró un MAE de 0.60 y un RMSE de 0.83, indicando que sus predicciones tuvieron, en promedio, la menor desviación respecto a los valores reales y fue el más efectivo en minimizar errores de gran magnitud.

El siguiente modelo con mejor desempeño fue PH, con un MAE de 1.08 y un RMSE de 1.43. Si bien demostró ser una alternativa robusta, sus errores fueron notablemente superiores a los de RF.

Dentro de las variantes de MA, el modelo con una ventana de 12 meses ($MAE = 1.29$, $RMSE = 1.76$) resultó ser el más preciso de los tres, superando a las configuraciones con ventanas de 3 y 6 meses. No obstante, todos los modelos de MA presentaron errores significativamente más altos que los obtenidos por los dos modelos más avanzados.

3.2 Análisis visual de las predicciones

Para complementar la evaluación cuantitativa, esta sección presenta un análisis visual del comportamiento de los modelos. Los gráficos permiten observar de manera intuitiva cómo cada método de pronóstico se ajusta a la serie temporal real, facilitando la identificación de patrones como el seguimiento de la tendencia, la captura de la estacionalidad y la presencia de retardos en la predicción. A continuación, se muestran los resultados para MA, RF y PH.

La Figura 1 presenta la serie temporal del consumo real de agua superpuesta con las predicciones generadas por los tres modelos de MA con ventanas de 3, 6 y 12 meses. Se puede observar que las predicciones de los modelos de MA son versiones suavizadas de la serie original. Ninguno de los modelos logra capturar completamente los picos y valles (la estacionalidad) del consumo real, sino que tienden a aplandar estas fluctuaciones.

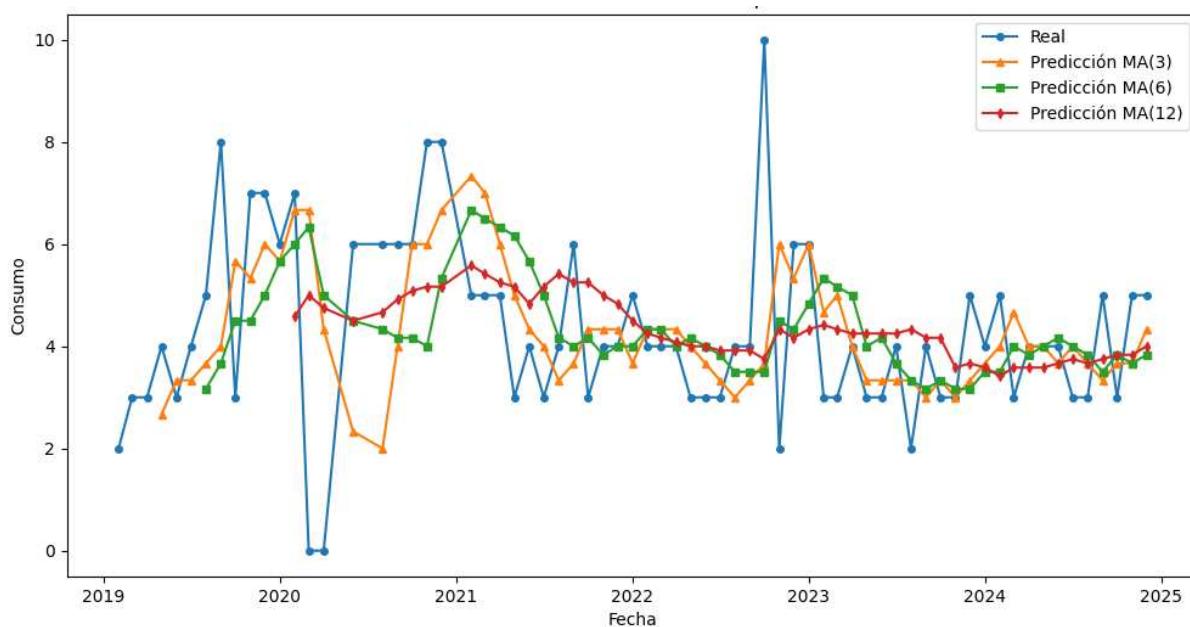


Figura 1: Comparación de Predicciones de MA vs. Consumo Real.

Se aprecian las siguientes diferencias clave entre las tres ventanas:

MA(3): La línea de predicción con ventana de 3 meses es la que sigue más de cerca las variaciones del consumo real, aunque presenta un ligero retraso. Es la más reactiva a los cambios a corto plazo.

MA(6): Aumentar la ventana a 6 meses produce una línea de predicción más suave que la anterior, mostrando un mayor aplanamiento de los picos y un retraso más pronunciado.

MA(12): La predicción con una ventana de 12 meses es la más plana y estable de las tres. Si bien es la que peor captura la volatilidad mensual, es la que mejor representa la tendencia a largo plazo de la serie, confirmando los resultados de la Tabla 2, donde este modelo obtuvo el menor error.

En conjunto, el gráfico ilustra la característica fundamental de las medias móviles: su incapacidad para predecir picos estacionales y su tendencia a reaccionar con retraso a los cambios en los datos.

La Figura 2 muestra la comparación entre los valores de consumo real y las predicciones generadas por el modelo de RF en el conjunto de prueba. A diferencia del comportamiento suavizado de las MA, se observa que la predicción del RF se ajusta de manera mucho más precisa a la serie de datos reales. El modelo es capaz de capturar las fluctuaciones mensuales, incluyendo los picos y valles estacionales, con una fidelidad considerablemente mayor.

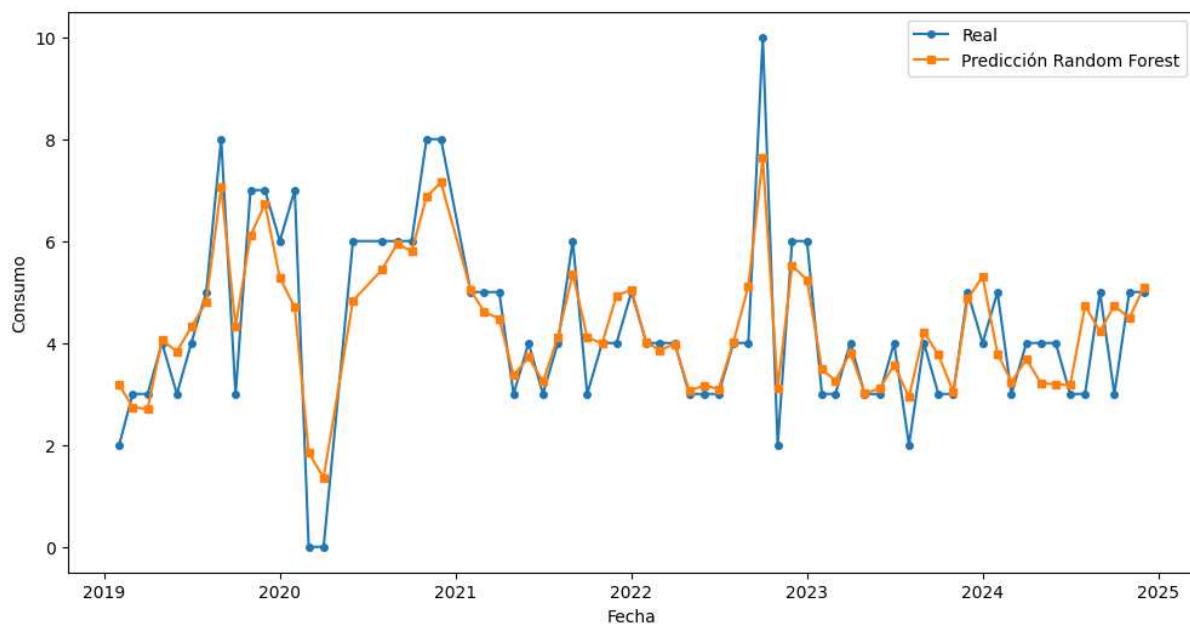


Figura 2: Comparación de la Predicción de RF vs. Consumo Real.

Los puntos clave que revela el gráfico son, por una parte, una alta capacidad de ajuste: la línea de predicción sigue muy de cerca a la línea de consumo real, lo que demuestra la habilidad del modelo para aprender las relaciones complejas y no lineales entre las variables de entrada (año, mes, temperatura) y el consumo. Por otra parte, no se aprecia el retardo característico de las medias móviles. El modelo predice los puntos altos y bajos del consumo en el momento en que ocurren, lo que es crucial para una validación precisa de las lecturas.

En resumen, el análisis visual confirma los resultados cuantitativos de la Tabla 2, mostrando que RF ofrece un modelo predictivo robusto y reactivo para la serie de consumo de agua analizada.

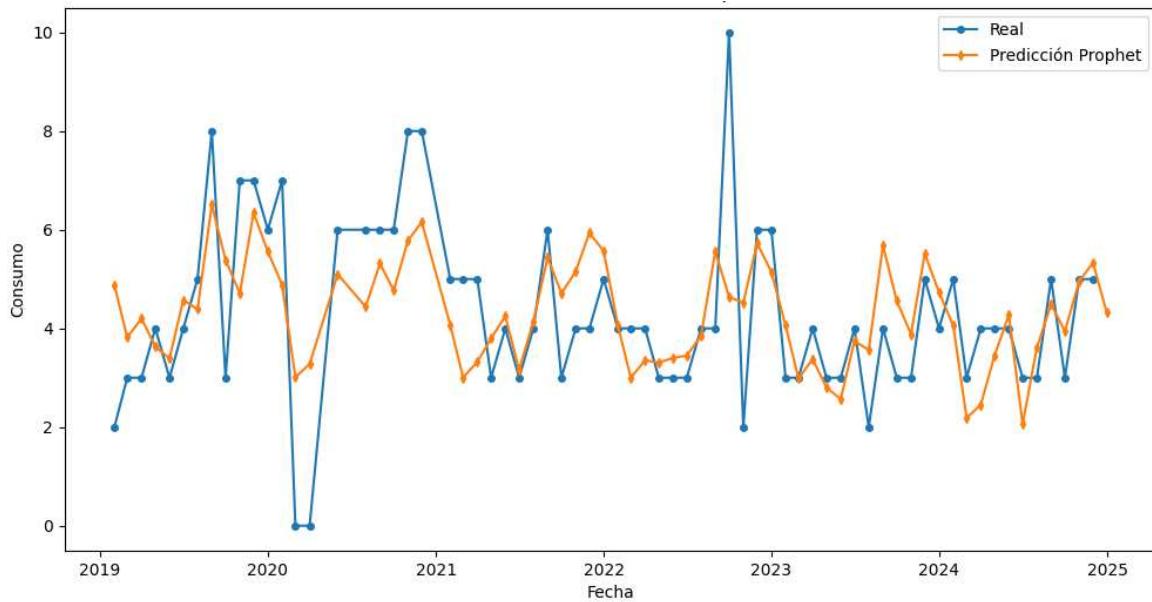


Figura 3: Comparación de la Predicción de PH vs. Consumo Real.

La Figura 3 ilustra los resultados del modelo PH, mostrando la serie de consumo real junto con la predicción del modelo. Visualmente, la predicción de PH se caracteriza por una curva

suave y cíclica que modela los patrones subyacentes de la serie temporal, en lugar de seguir las fluctuaciones de alta frecuencia.

El primer aspecto destacable en la utilización de esta estrategia es la capacidad de capturar la estacionalidad. El modelo identifica y reproduce con claridad el patrón estacional anual del consumo. La curva de predicción sube y baja en sintonía con los picos y valles estacionales de los datos reales, demostrando su capacidad para modelar ciclos a largo plazo. También puede observarse que la predicción sigue la tendencia general de la serie, ya sea esta creciente, decreciente o estable a lo largo del tiempo.

En conclusión, el gráfico muestra que PH es efectivo para modelar la estructura fundamental de la serie temporal (tendencia + estacionalidad), aunque no se ajusta a cada punto individual con la misma precisión que el RF, proporciona un pronóstico robusto y estructuralmente coherente.

4 CONCLUSIONES

Se evaluaron y compararon tres metodologías de pronóstico (MA, RF y PH) con el fin práctico de desarrollar una herramienta para validar lecturas manuales de medidores de agua y reducir los errores de facturación en el contexto de cooperativas de servicios públicos argentinas. Los resultados demuestran que los modelos de aprendizaje automático ofrecen una ventaja significativa sobre los métodos estadísticos tradicionales para esta tarea. El modelo de RF se destacó como la alternativa más precisa, obteniendo los valores más bajos en el error absoluto medio y en el error cuadrático medio, mostrando en el análisis visual una capacidad superior para capturar las fluctuaciones del consumo real sin el retardo inherente a otros métodos. Por su parte, PH también se mostró como un modelo robusto, especialmente eficaz para modelar la estructura estacional del consumo, aunque con una precisión puntual ligeramente inferior a la de RF. En contraste, los modelos de Medias Móviles, si bien son simples de implementar, resultaron insuficientes para la tarea de validación precisa debido a su incapacidad para modelar la estacionalidad y su retardo en la respuesta.

La principal implicancia de estos resultados es de naturaleza práctica y operativa. Un modelo como RF puede ser implementado como una herramienta de validación automática. Al comparar la lectura manual de un medidor con la predicción del modelo para ese usuario y período, el sistema puede generar alertas cuando la discrepancia supere un umbral predefinido. Esto permitiría a la cooperativa detectar errores de lectura o posibles fugas de agua de manera proactiva, mejorando la precisión del sistema de facturación y reduciendo significativamente los conflictos con los usuarios.

Es importante reconocer las limitaciones de este estudio para contextualizar los resultados. El análisis se ha centrado en datos de una única cooperativa y en un usuario como caso testigo para la visualización, por lo que la generalización de los resultados requiere estudios adicionales en contextos más amplios. Asimismo, no se ha realizado una optimización exhaustiva de los hiperparámetros de los modelos.

Como líneas de trabajo futuro, se propone la validación de esta metodología en un conjunto de datos que abarque un mayor número de usuarios y diversas cooperativas. Además, se podría explorar la incorporación de variables exógenas adicionales (como precipitaciones o eventos locales) y el desarrollo de un prototipo de software que integre el modelo predictivo directamente en el sistema de gestión comercial de la cooperativa, pasando de un estudio de viabilidad a una herramienta operativa funcional.

AGRADECIMIENTOS

Los autores agradecen el apoyo de la Secretaría de Ciencia y Técnica de la Universidad

Nacional de Río Cuarto, la Universidad Nacional de La Pampa y el Consejo Nacional de Investigaciones Científicas y Técnicas (CONICET), cuyo respaldo ha sido fundamental para el desarrollo de esta investigación, especialmente en el desafiante contexto actual que atraviesa el sistema académico-científico-tecnológico.

REFERENCIAS

- Box, G., Jenkins, G., Reinsel, G. and Ljung, G. Time series analysis: forecasting and control. John Wiley & Sons. 2015
- Ghannam, S., Hussain, F.. Short-term water demand forecasting: a review. *Australasian Journal of Water Resources*, 1–19, 2024. <https://doi.org/10.1080/13241583.2024.2350102>
- Brentan, B., Luvizotto, E., Herrera, M., Izquierdo, J., Pérez-García, R. Hybrid regression model for near real-time urban water demand forecasting. *Journal of Computational and Applied Mathematics*. Volume 309, Pages 532-541, ISSN 0377-0427. 2017.
- Albahli, Saleh. (2025). LSTM vs. Prophet: Achieving Superior Accuracy in Dynamic Electricity Demand Forecasting. *Energies*. 18. 278. 10.3390/en18020278.
- Nadar, Ajesh & Chandane, Soham & Raj, Gabriel & Pasi, Nihar & Patil, Yash. (2025). Automated Energy Billing with Blockchain and the Prophet Forecasting Model: A Holistic Approach. 10.48550/arXiv.2506.16649.
- Schisa, V. & Farnè, M., 2025. *The Impact of Climatic Factors on Respiratory Pharmaceutical Demand: A Comparison of Forecasting Models for Greece*. arXiv preprint arXiv:2505.10642.
- Taylor, S.J. & Letham, B., 2018. *Forecasting at scale*. *The American Statistician*, 72(1), pp.37–45. <https://doi.org/10.1080/00031305.2017.1380080>