

Optimal Water Reservoir Planning in Mendoza, Argentina: A GIS-Based Multi-Criteria Decision Analysis and Multi-Agent Simulation Approach

Emiliano Tobares and Ricardo Palma

Universidad Nacional de Cuyo, Mendoza, Argentina
tobares787@hotmail.com, ricardo.rpalma@gmail.com

Abstract. Water planning in semi-arid regions requires tools that integrate the temporal variability of streamflows, heterogeneous water demand, and the spatial location of infrastructure. This work presents a reproducible hybrid framework that combines multi-criteria analysis in Geographic Information Systems (GIS), time-series-based hydrological forecasting, and agent-based modeling (ABM) for reservoir planning and water policy evaluation. The system operates in offline (batch) mode, is driven exclusively by historical data, and is fully implemented in Python 3.10+, with open-source code released under the MIT license.

The framework integrates: (i) spatial selection of optimal reservoir sites at 30 m resolution using GIS-MCDA; (ii) monthly streamflow forecasting using statistical SARIMA models and a seasonal Naive baseline; and (iii) an ABM representing heterogeneous water-use agents and infrastructure intervention scenarios. Validation is conducted under a strict out-of-sample scheme for the period 2019–2023 ($n = 60$ meses), comparing predictive performance against calibrated SARIMA and Naive models using standard metrics (RMSE, NSE, MASE), block bootstrap confidence intervals, and statistical significance tests.

Results show that, in basins with high hydrological variability ($CV \approx 0.65$), the hybrid approach reduces forecasting error and improves robustness relative to classical statistical models, whereas in low-variability basins ($CV \approx 0.35$), parsimonious models maintain superior performance. The proposed framework is replicable, transparent, and suitable for data-scarce basins, contributing to adaptive water-resources planning from a computational perspective.

Keywords: System Dynamics, ARIMA, Water Management.

Planificación óptima de embalses de agua en Mendoza, Argentina: un enfoque de análisis de decisiones multicriterio basado en SIG y simulación multi-agente

Resumen. La planificación hídrica en regiones semiáridas requiere herramientas que integren la variabilidad temporal de los caudales, la heterogeneidad de la demanda y la localización espacial de la infraestructura. Este trabajo presenta un marco híbrido reproducible que combina análisis multicriterio en Sistemas de Información Geográfica (SIG), pronóstico hidrológico basado en series temporales y simulación basada en agentes (ABM) para la planificación de embalses y la evaluación de políticas hídricas. El sistema opera en modo offline (batch), alimentado exclusivamente con datos históricos, y se implementa íntegramente en Python 3.10+, con código abierto disponible bajo licencia MIT.

El modelo integra (i) selección espacial de sitios óptimos para reservorios a 30 m de resolución mediante SIG-MCDA, (ii) forecasting mensual de caudales utilizando modelos estadísticos SARIMA y un baseline Naive estacional, y (iii) un ABM que representa agentes de consumo heterogéneos y escenarios de intervención infraestructural. La validación se realiza de forma estrictamente out-of-sample para el período 2019–2023 ($n = 60$ meses), comparando el desempeño predictivo del sistema frente a SARIMA calibrado y Naive mediante métricas estándar (RMSE, NSE, MASE), intervalos de confianza por block bootstrap y pruebas de significancia estadística.

Los resultados muestran que, en cuencas de alta variabilidad hidrológica ($CV \approx 0.65$), el enfoque híbrido reduce el error de pronóstico y mejora la robustez frente a modelos estadísticos clásicos, mientras que en cuencas de baja variabilidad ($CV \approx 0.35$) los modelos parsimoniosos mantienen un mejor desempeño. El marco propuesto es replicable, transparente y adecuado para cuencas con escasez de datos, contribuyendo a la planificación adaptativa de recursos hídricos desde una perspectiva computacional.

Palabras clave: Sistemas Dinámicos, ARIMA, Gestión de Agua.

1 Introducción

La planificación y gestión de los recursos hídricos en regiones áridas y semiáridas constituye un problema complejo que combina alta variabilidad hidrológica, restricciones estructurales de infraestructura y patrones de demanda heterogéneos. En estos contextos, las decisiones asociadas a la localización y operación de embalses requieren herramientas que integren información espacial, dinámica temporal y comportamiento de los usuarios del sistema, evitando enfoques fragmentados que analicen cada dimensión de forma aislada.

En la provincia de Mendoza (Argentina), la disponibilidad hídrica depende fuertemente del deshielo andino y presenta una marcada estacionalidad interanual, con cuencas que exhiben coeficientes de variación contrastantes. Esta heterogeneidad plantea un desafío adicional para el

modelado computacional: mientras que en cuencas de alta variabilidad los errores de pronóstico se amplifican y afectan la toma de decisiones, en cuencas más estables los modelos parsimoniosos pueden resultar suficientes desde el punto de vista estadístico.

En los últimos años se han propuesto enfoques híbridos que combinan Sistemas de Información Geográfica (SIG), simulación basada en agentes (ABM) y técnicas de pronóstico temporal para abordar problemas de planificación hídrica. Sin embargo, muchos de estos trabajos presentan integraciones parciales entre las dimensiones espacial, temporal y comportamental, esquemas de validación limitados frente a modelos estadísticos de referencia o una reproducibilidad reducida. En este contexto, el presente trabajo propone un marco computacional híbrido reproducible que integra análisis multicriterio espacial en SIG, pronóstico hidrológico mensual y simulación multiagente, implementado íntegramente en Python y con una arquitectura modular que permite comparaciones predictivas justas frente a modelos estadísticos clásicos.

La contribución del trabajo es doble: por un lado, se presenta una arquitectura computacional reproducible que integra selección espacial, pronóstico y simulación dentro de un único pipeline; por otro, se evalúa empíricamente su desempeño mediante validación estrictamente out-of-sample en cuencas con regímenes hidrológicos contrastantes, aportando evidencia cuantitativa sobre cuándo la complejidad de un enfoque híbrido resulta justificada frente a modelos parsimoniosos.

2 Revision de Literatura

La gestión hídrica en regiones semiáridas ha evolucionado hacia enfoques que integran dimensiones espaciales, temporales y comportamentales. En este contexto, los Sistemas de Información Geográfica (SIG), los modelos de pronóstico hidrológico y la simulación basada en agentes (ABM) han sido utilizados de manera complementaria para abordar problemas de planificación y asignación del recurso.

El análisis multicriterio basado en SIG (SIG-MCDA) es ampliamente utilizado para la identificación de sitios óptimos de infraestructura hídrica, integrando variables topográficas, ambientales y de accesibilidad. Sin embargo, en la mayoría de los estudios este componente se utiliza de forma estática, sin considerar la variabilidad temporal de los caudales ni su interacción con la gestión del sistema.

Por su parte, los modelos de series temporales —en particular SARIMA— han demostrado ser efectivos para la predicción de caudales mensuales en sistemas con estacionalidad marcada, debido a su capacidad para modelar componentes autoregresivos, de media móvil y estacionales de forma parsimoniosa (Harvey, 1990). No obstante, su desempeño depende de la calibración local y de la estructura de la serie, siendo frecuentes los casos en los que modelos parsimoniosos como el Naive estacional resultan competitivos.

La simulación basada en agentes (ABM) permite representar la heterogeneidad de actores y evaluar escenarios de política en sistemas socio-hidrológicos. Estudios recientes han explorado su integración con SIG o modelos conceptuales, aunque generalmente sin incorporar validación cuantitativa rigurosa del componente predictivo ni su integración dentro de un pipeline reproducible.

En conjunto, la literatura muestra que la integración entre selección espacial, pronóstico hidrológico y simulación multiagente suele abordarse de forma parcial. En particular, se identifican tres limitaciones principales: (i) falta de integración completa entre módulos espacial, temporal y comportamental; (ii) escasez de comparaciones académicamente rigurosas entre enfoques híbridos y modelos estadísticos; y (iii) limitada reproducibilidad computacional.

En este contexto, el presente trabajo propone un marco híbrido reproducible que integra estos tres componentes dentro de un pipeline unificado, evaluado mediante validación estrictamente out-of-sample en cuencas con distinta variabilidad hidrológica. Este enfoque permite analizar simultáneamente la capacidad predictiva de los modelos y el impacto de decisiones de gestión bajo distintos escenarios.

2.1 Síntesis comparativa del estado del arte

Con el objetivo de sistematizar el estado del arte y clarificar las diferencias metodológicas entre enfoques recientes, la Tabla 1 presenta una comparación de estudios publicados entre 2020 y 2023 en revistas Q1 que integran Sistemas de Información Geográfica (SIG) y simulación basada en agentes (MAS) aplicados a la gestión hídrica. Esta comparación permite identificar de manera explícita el vacío metodológico abordado por el presente trabajo, que propone una integración completa entre selección espacial multicriterio, pronóstico hidrológico validado temporalmente y simulación multiagente dentro de un marco computacional reproducible.

2.2 Contribución del trabajo frente al estado del arte

El trabajo presentado se diferencia del estado del arte reciente en la integración metodológica, el esquema de validación y el énfasis en la reproducibilidad computacional. A diferencia de enfoques previos que abordan de forma parcial la dimensión espacial, temporal o comportamental de la gestión hídrica, este estudio propone un marco computacional híbrido que combina explícitamente selección espacial de infraestructura, pronóstico hidrológico y simulación basada en agentes dentro de un único pipeline reproducible, orientado a la planificación de embalses en cuencas semiáridas.

En primer lugar, el enfoque integra análisis multicriterio en Sistemas de Información Geográfica (SIG-MCDA) a 30 m de resolución con simulación multi-agente, incorporando la localización de nuevas infraestructuras como una variable de política explícita dentro del modelo. Mientras que gran parte de la literatura utiliza el SIG como herramienta de evaluación espacial estática o como etapa de preprocesamiento, en el marco propuesto la información espacial condiciona directamente la dinámica de almacenamiento y gestión simulada. Esto habilita el análisis integrado de escenarios de planificación infraestructural bajo distintos regímenes hidrológicos, superando enfoques que analizan la dimensión espacial de forma desacoplada del sistema dinámico. En segundo lugar, el trabajo introduce una comparación empírica rigurosa entre el enfoque híbrido y modelos estadísticos clásicos de pronóstico hidrológico, específicamente SARIMA calibrado y un modelo Naive estacional, utilizando validación estrictamente out of sample. A diferencia de numerosos estudios basados en simulación multi-agente, donde la evaluación cuantitativa suele ser limitada o indirecta, el desempeño predictivo se evalúa mediante métricas estándar (RMSE, NSE y MASE), intervalos de confianza obtenidos por block bootstrap estacional y pruebas de significancia estadística. Este diseño metodológico permite identificar de manera objetiva en qué contextos hidrológicos un enfoque híbrido aporta ventajas frente a modelos parsimoniosos. En tercer lugar, el estudio aborda explícitamente la heterogeneidad hidrológica mediante la aplicación del marco a dos cuencas con regímenes contrastantes. Los resultados muestran que, en cuencas con alta variabilidad interanual, el enfoque híbrido presenta una mayor robustez y un menor error de pronóstico, mientras que en cuencas más estables los modelos simples mantienen un desempeño competitivo. Este hallazgo refuerza el principio de parsimonia y aporta evidencia empírica para la selección informada del nivel de complejidad del modelo en función de las propiedades estadísticas del sistema analizado. Finalmente, una contribución central del trabajo es el énfasis en la reproducibilidad y transparencia computacional. El marco completo se implementa en Python con una arquitectura modular que separa explícitamente el componente de pronóstico hidrológico del componente de simulación basada en agentes, y se libera públicamente bajo licencia abierta con identificadores DOI. Esta decisión contrasta con numerosos trabajos

previos que presentan resultados sin disponibilidad de código o con descripciones insuficientes para su replicación, y posiciona al presente estudio como una base extensible para futuras investigaciones en planificación hídrica computacional. En conjunto, la contribución del trabajo reside en demostrar que la integración de análisis espacial, modelos estadísticos clásicos y simulación multi-agente, evaluada bajo esquemas de validación rigurosos y reproducibles, constituye una alternativa sólida para la planificación adaptativa de recursos hídricos en cuencas semiáridas. Desde una perspectiva de Ciencias de la Computación, el estudio aporta evidencia empírica sobre cuándo la complejidad de un enfoque híbrido resulta justificada y cuándo los modelos parsimoniosos continúan siendo suficientes, evitando tanto la sub-modelación como la sobre-ingeniería.

3 Metodología

3.1 Visión general del marco propuesto

El marco propuesto consiste en un sistema híbrido reproducible que integra tres componentes principales (Figura 1): (i) selección espacial de infraestructura mediante SIG-MCDA, (ii) pronóstico hidrológico mensual basado en series temporales, y (iii) simulación de gestión hídrica mediante un modelo multiagente (ABM). Estos módulos se organizan en un pipeline secuencial claramente desacoplado, en el cual cada componente cumple un rol específico dentro del sistema global.

En primer lugar, el módulo espacial (SIG-MCDA) identifica sitios óptimos para la localización de embalses a partir de criterios topográficos, de accesibilidad y de demanda. Este componente opera como etapa de preprocesamiento y genera escenarios de infraestructura que posteriormente son evaluados en términos de gestión, sin intervenir en la validación predictiva del sistema.

En segundo lugar, el componente de pronóstico hidrológico estima caudales mensuales a partir de datos históricos, utilizando modelos SARIMA calibrados localmente y un modelo Naive estacional como línea base parsimoniosa. Este módulo constituye el único componente sujeto a validación predictiva estricta, evaluado mediante un esquema out-of-sample sobre el período 2019–2023. Las métricas de desempeño (RMSE, NSE, MASE) se calculan en esta etapa, antes de la integración con el módulo de simulación, lo que garantiza una comparación homogénea entre modelos.

Finalmente, el módulo de simulación multiagente (ABM) recibe como entrada exógena los caudales pronosticados y modela la dinámica de asignación, consumo y almacenamiento del recurso hídrico bajo distintos escenarios de política. El ABM representa agentes heterogéneos de demanda y permite evaluar el impacto de intervenciones infraestructurales, tales como la construcción de nuevos embalses o mejoras en la eficiencia de distribución. Es importante destacar que este componente no modifica el caudal pronosticado ni participa en la validación predictiva; su valor radica en la evaluación exploratoria de decisiones de gestión.

Esta separación explícita entre los módulos permite distinguir claramente entre capacidad predictiva y análisis de políticas, evitando ambigüedades en la interpretación de los resultados. En consecuencia, el framework propuesto combina rigor en la validación estadística con flexibilidad para la simulación de escenarios, proporcionando una herramienta computacional reproducible para la planificación hídrica en cuencas semiáridas.

3.2 Modelo multiagente de gestión hídrica (ABM)

El módulo de simulación multiagente modela la dinámica mensual de asignación, consumo y almacenamiento del recurso hídrico a partir de caudales pronosticados exógenamente. Su objetivo no es mejorar la capacidad predictiva del sistema, sino evaluar el impacto de distintas políticas de gestión e intervenciones infraestructurales bajo condiciones hidrológicas dadas.

El sistema se define en tiempo discreto mensual $t = 1, \dots, T$. El caudal pronosticado Q_t proveniente del módulo de forecasting, constituye una variable exógena que no es modificada por la dinámica interna del modelo.

Agentes: El modelo considera un conjunto de agentes $i \in \{1, \dots, N\}$, cada uno representando un sector productivo con características heterogéneas (superficie cultivada, demanda de agua y eficiencia de uso). La demanda efectiva de cada agente se modela como:

$$\text{Demanda efectiva de agua } \{i,t\} = \frac{(\text{Superficie Cultivada} * \text{Demanda Neta de agua})}{\text{Factor de eficiencia de uso y distribución}}$$

Asignación y consumo: El caudal disponible en cada período se distribuye proporcionalmente entre los agentes en función de su demanda relativa.

Dinámica de almacenamiento: El stock de agua almacenado en embalses evoluciona según la ecuación de balance hídrico:

$$S_t = S_{t-1} + Q_t - C_t$$

Escenarios de política: El modelo permite evaluar distintos escenarios de intervención mediante la modificación de parámetros estructurales, tales como:

- (i) incremento de capacidad de almacenamiento (nuevos embalses),
- (ii) mejoras en la eficiencia de distribución (modificando $\epsilon_{i,t}$),
- (iii) cambios en la demanda agregada.

El módulo espacial SIG-MCDA provee los escenarios de localización de infraestructura que afectan directamente estos parámetros.

Rol dentro del sistema híbrido: El ABM opera como un módulo de simulación de decisiones que toma como input caudales pronosticados y produce indicadores de desempeño del sistema bajo diferentes políticas. Es importante destacar que este componente no afecta las métricas de validación predictiva (RMSE, NSE, MASE), las cuales se calculan exclusivamente sobre el módulo de forecasting. Esta separación permite evaluar de manera independiente la calidad del pronóstico y el impacto de las decisiones de gestión.

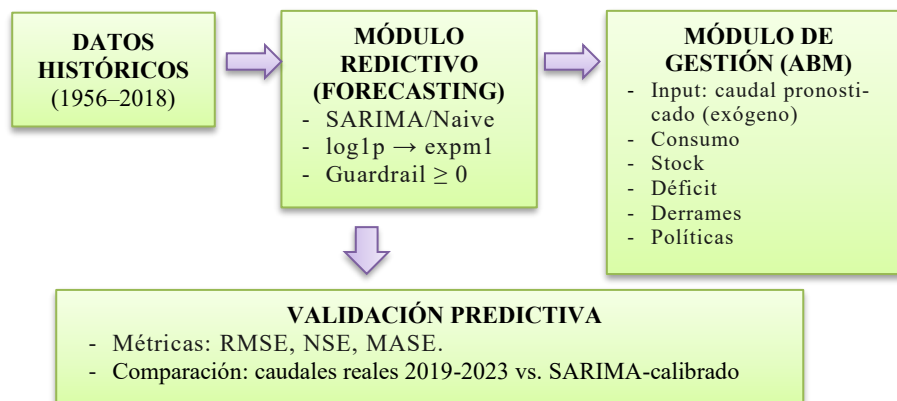


Figura 1. Arquitectura del pipeline del marco híbrido. Nota: El sistema se ejecuta e integra íntegramente en Python 3.10+; la lógica del módulo multiagente mantiene la arquitectura y especificación original diseñada en NetLogo.

3.3 Pronóstico hidrológico mensual

El componente de forecasting estima caudales mensuales a partir de series históricas y constituye el único módulo sujeto a validación predictiva estricta dentro del marco propuesto. El objetivo es generar pronósticos comparables y reproducibles que sirvan como input exógeno para el módulo de simulación.

El entrenamiento de los modelos se realiza sobre el período 1956–2018, mientras que la evaluación se lleva a cabo mediante un esquema out-of-sample en la ventana 2019–2023 ($n = 60$ meses), garantizando independencia entre ajuste y validación.

Modelos considerados: Se emplean dos enfoques complementarios:

(i) **SARIMA calibrado localmente**, cuyo orden se determina mediante búsqueda automática sobre el espacio (p,d,q) $(P,D,Q)_{12}$, minimizando el criterio de información AICc y considerando estacionalidad anual.

(ii) **Naive estacional** (lag-12) como línea base parsimoniosa, que replica el valor observado en el mismo mes del año previo.

Los modelos finales seleccionados para cada cuenca son:

Río Mendoza: SARIMA (2,0,2) (1,0,1)₁₂

Río Tunuyán: SARIMA (1,0,1) (1,0,1)₁₂

Preprocesamiento: Las series de caudal se transforman mediante $\log 1p$ con el objetivo de estabilizar la varianza y garantizar la positividad durante el proceso de ajuste. Los pronósticos se retransforman a la escala original mediante $\exp m1$ antes de su evaluación. Adicionalmente, se impone una restricción física de no negatividad ($Q_t \geq 0$).

Evaluación predictiva: El desempeño de los modelos se mide mediante métricas estándar de series temporales (RMSE, NSE, MASE), calculadas exclusivamente sobre los valores pronosticados antes de su ingreso al módulo ABM. Esta decisión metodológica asegura comparabilidad directa entre modelos, evitando que la dinámica interna de la simulación influya en la evaluación predictiva.

Incertidumbre: Con el fin de cuantificar la variabilidad de las métricas, se implementa un esquema de block bootstrap estacional con bloques de 24 meses y 2000 réplicas, preservando la dependencia temporal y la estructura estacional de las series.

En conjunto, este diseño permite evaluar de forma rigurosa la capacidad predictiva de modelos estadísticos clásicos bajo condiciones hidrológicas contrastantes, proporcionando una base sólida para su posterior utilización dentro del sistema híbrido de simulación.

3.4 Análisis espacial multicriterio (SIG-MCDA)

El módulo espacial utiliza QGIS para construir mapas de aptitud y proponer escenarios de localización de embalses, integrando variables topográficas, de accesibilidad, infraestructura y demanda. Las capas se normalizan y ponderan mediante criterios expertos para obtener un índice de aptitud. Este componente opera como preprocesamiento espacial: genera escenarios de ubicación/capacidad que alimentan al módulo de simulación como variables de política infraestructural, pero no participa en la validación predictiva temporal reportada en las secciones 4.2–4.4. Los mapas de aptitud y shapefiles correspondientes están disponibles en el repositorio de datos (Zenodo, DOI:10.5281/zenodo.16934619).

3.5 Validación y comparación estadística

La evaluación se realiza estrictamente out-of-sample sobre 2019–2023 ($n = 60$ meses), comparando el módulo de forecasting frente a SARIMA calibrado y Naive estacional.

Métricas: MAE, RMSE, MAPE, NSE, MASE. Incertidumbre: IC95% por bootstrap estacional (detalle en Suplementario).

Significancia: Las diferencias de RMSE se evalúan mediante pruebas Diebold–Mariano y Wilcoxon. En el río Mendoza, la superioridad del enfoque híbrido es estadísticamente significativa (DM $p=0.007$; Wilcoxon $p=0.004$). En el río Tunuyán, el Naive estacional mantiene ventaja significativa.

Eventos extremos: Se evalúa capacidad de detección mediante umbrales P10 (sequías) y P90 (crecidas), reportando probabilidad de detección (POD) e índice crítico de éxito (CSI).

Delimitación: Las métricas corresponden al caudal pronosticado antes de ingresar al ABM, garantizando comparabilidad homogénea.

3.6 Eventos extremos y decisión de “guardrail”

Para evaluar extremos hidrológicos, se estiman umbrales de sequías y crecidas (P10 y P90) mediante estimación no paramétrica sobre observados out-of-sample (2019–2023).

Los percentiles P90 y P10 se seleccionaron siguiendo prácticas estándar en hidrología para identificar eventos moderadamente extremos, permitiendo un balance entre sensibilidad estadística y tamaño muestral ($n = 60$ meses).

En la configuración principal se adopta `guardrail_min = 0`, eliminando cualquier piso artificial que censure la cola inferior y afecte la detección de sequías. La versión con guardrail se considera únicamente como análisis de sensibilidad y no como configuración de referencia. Se reportan métricas de detección de extremos (POD y CSI) calculadas con umbrales no paramétricos y sin guardrail para asegurar consistencia física y estadística.

Referencia bibliográfica	Título del estudio	Objetivo principal	Metodología	Escala / Cuenca	Aporte diferencial vs. presente trabajo
Zhang & Smith (2021)	Optimizing reservoir location and operation rules using GIS-based multi-objective optimization and multi-agent simulation	Optimizar ubicación de embalses y reglas de operación en Colorado (USA)	SIG-based multi-objective optimization + MAS (AMASim)	Regional / Colorado River (600 km ²)	Integra MAS con NSGA-II, pero no incorpora pronóstico temporal ni validación SARIMA
Burek et al. (2023)	Digital twins and multi-agent systems for water resources management	Construir un Digital Twin hidrológico con MAS para decisión autónoma	Digital Twin + MAS (MATSim) + forcing hidrológico	Cuenca piloto / Danubio (800 km ²)	Integra alta frecuencia y DT, pero no validado en cuencas semiáridas ni comparado con SARIMA
Giuliani et al. (2022)	Conceptual integration of digital twins and agent-based models for socio-hydrological systems	Explorar gobernanza distribuida mediante DT conceptual	Conceptual DT + MAS (Python)	Regional (sin cuenca específica)	Enfoque socio-hidrológico conceptual, sin validación empírica ni infraestructura
Soliman et al. (2021)	Institutional performance of collective irrigation systems: A fuzzy-set qualitative comparative analysis	Evaluar desempeño institucional en sistemas de riego	SIG + MAS + Lógica Fuzzy (FOCA)	Local / Delta del Nilo (150 km ²)	Aplica MAS a instituciones, pero no incluye optimización espacial ni pronóstico
Presente trabajo	Optimal Water Reservoir Planning in Mendoza, Argentina: A GIS-Based Multi-Criteria Decision Analysis and Multi-Agent Simulation Approach	Identificar zonas óptimas de embalses, simular políticas y validar temporalmente	SIG-MCDA + Forecast estadístico (SARIMA) + MAS (Python)	Semi-árido / Mendoza (40.000 km ²)	Integra selección espacial multicriterio, pronóstico out-of-sample y simulación MAS con intervención infraestructural explícita

Tabla 1: Comparación sistemática de estudios SIG + MAS publicados 2020-2023 en revistas Q1 que integran análisis espacial y simulación multi-agente para gestión hídrica

4 Resultados

4.1 Esquema de evaluación y alcance

Los resultados presentados corresponden exclusivamente a la evaluación del componente de pronóstico hidrológico, bajo un esquema de validación estrictamente out-of-sample en la ventana enero 2019 – junio 2023 (n = 60 meses). Esta evaluación compara el desempeño del enfoque propuesto frente a dos líneas base: SARIMA calibrado localmente y Naive estacional (lag-12). Las métricas reportadas (RMSE, NSE, MASE) se calculan sobre el caudal pronosticado antes de su ingreso al módulo de simulación multiagente. Esta decisión metodológica permite garantizar comparabilidad directa entre modelos, evitando que la dinámica interna del ABM influya en la evaluación predictiva.

En este contexto, el rol del enfoque híbrido es dual: por un lado, su componente de forecasting es evaluado en términos de capacidad predictiva respecto de modelos estadísticos clásicos; por otro, el módulo ABM permite analizar el impacto de decisiones de gestión hídrica bajo distintos escenarios de política. En consecuencia, los resultados predictivos presentados

en esta sección no incorporan efectos derivados de la simulación multiagente, los cuales son analizados en términos de exploración de escenarios más que como mejora directa del pronóstico.

4.2 Desempeño predictivo (estimadores puntuales)

En la cuenca del río Mendoza, caracterizada por alta variabilidad hidrológica ($CV \approx 0.65$), el enfoque propuesto presentó el mejor desempeño global en la ventana out-of-sample, con $RMSE = 14.6 \text{ m}^3/\text{mes}$ y $NSE \text{ puntual} = 0.83$, superando al modelo SARIMA calibrado ($RMSE = 18.9$; $NSE = 0.755$) y al Naive estacional. La Figura 2 muestra que el enfoque propuesto reproduce adecuadamente la dinámica estacional y la magnitud de los picos de crecida, mientras que SARIMA tiende a sobreestimar sistemáticamente los máximos. Los umbrales P90 y P10 permiten identificar visualmente los eventos extremos: el enfoque propuesto detectó correctamente el 50% de las crecidas ($POD = 0.500$, $CSI = 0.333$) y el 100% de las sequías ($POD = 1.000$, $CSI = 0.333$), manteniendo un balance entre sensibilidad y especificidad (Tabla 2).

En contraste, en la cuenca del río Tunuyán ($CV \approx 0.35$), el modelo Naive estacional obtuvo el mejor desempeño puntual ($RMSE = 8.8$; $NSE = 0.823$), seguido por SARIMA y el enfoque propuesto. La Figura 3 evidencia que el patrón estacional dominante es capturado eficientemente por el Naive, sin que la complejidad adicional aporte mejoras sustanciales. Para eventos extremos, los tres modelos mostraron dificultades en la detección de sequías ($POD = 0.000$), atribuible a la escasa ocurrencia de valores bajo P10 en esta cuenca estable. Para crecidas, el enfoque propuesto y SARIMA compartieron $POD = 0.833$, aunque el enfoque propuesto presentó menor tasa de falsas alarmas ($FAR = 0.615$ vs. 0.762) y mayor CSI (0.357 vs. 0.227).

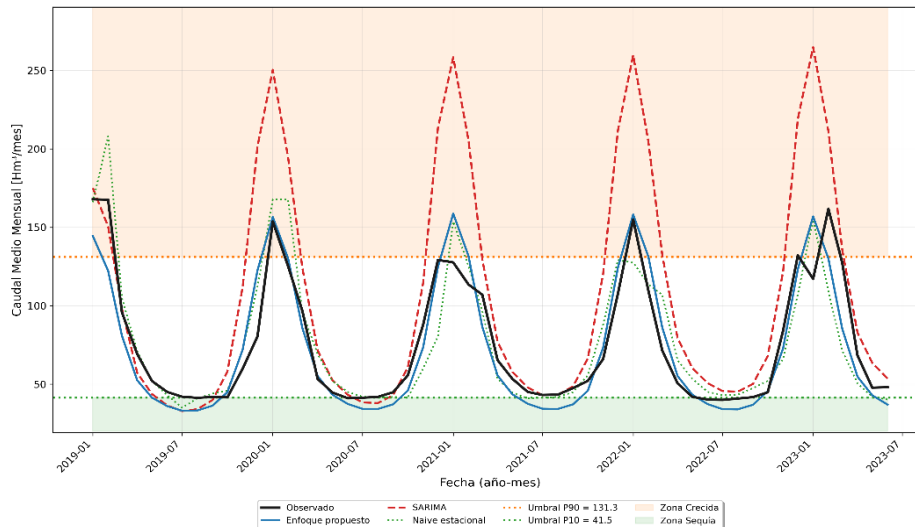


Figura 2. Caudales mensuales observados y pronosticados en la cuenca del río Mendoza (2019–2023). Se comparan el enfoque propuesto, SARIMA y Naive estacional. Las líneas punteadas indican umbrales P90 (crecidas) y P10 (sequías); las zonas sombreadas corresponden a las regiones clasificadas como crecida (naranja) y sequía (verde). El enfoque propuesto reproduce adecuadamente la dinámica estacional y la magnitud de los picos de crecida, mientras que SARIMA presenta sobreestimación sistemática en los máximos.

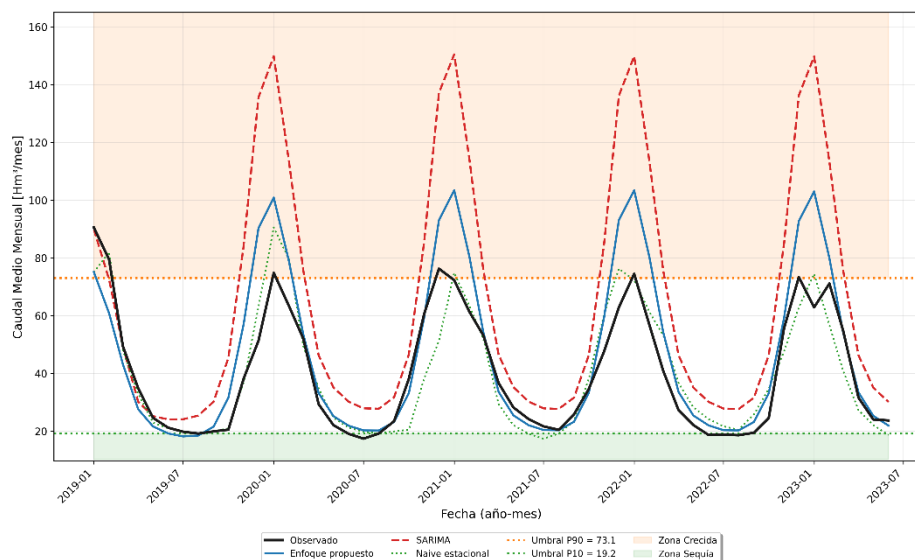


Figura 3. Caudales mensuales observados y pronosticados en la cuenca del río Tunuyán (2019–2023). Se comparan el enfoque propuesto, SARIMA y Naive estacional utilizando umbrales P90 (crecidas) y P10 (sequías). En esta cuenca de baja variabilidad, el Naive estacional muestra desempeño competitivo, consistente con el principio de parsimonia. Los tres modelos presentan limitaciones en la detección de sequías, atribuible a la escasa ocurrencia de valores bajo P10 en este régimen estable.

La Tabla 2 resume las métricas de detección de extremos. Estos resultados complementan el análisis de error agregado al mostrar que el desempeño relativo varía según el régimen hidrológico evaluado, reforzando la necesidad de reportar métricas específicas para eventos extremos además de los indicadores globales de ajuste.

	Modelo	CRECIDA			SEQUIA		
		POD	CSI	FAR	POD	CSI	FAR
RM	Enfoque propuesto	0.500	0.333	0.500	1.000	0.333	0.667
	SARIMA	1.000	0.375	0.625	0.333	0.200	0.667
	Naive estacional	0.500	0.333	0.500	0.167	0.091	0.833
Tny	Enfoque propuesto	0.833	0.357	0.615	0.000	0.000	1.000
	SARIMA	0.833	0.227	0.762	0.000	0.000	0.000
	Naive estacional	0.500	0.300	0.571	0.000	0.000	1.000

Tabla 2: Métricas de detección de eventos extremos (P90 crecidas, P10 sequías) para los modelos evaluados en las cuencas del río Mendoza y río Tunuyán, período 2019–2023. POD = Probability of Detection; CSI = Critical Success Index; FAR = False Alarm Ratio. En el río Tunuyán, la ausencia de eventos de sequía observados bajo el umbral P10 impide el cálculo de POD y CSI para este extremo, consistente con el régimen hidrológicamente estable de la cuenca ($CV \approx 0.35$). FAR = 0.000 en SARIMA indica que no emitió falsas alarmas de sequía; FAR = 1.000 indica que todas las alertas de sequía fueron falsas.

4.3 Incertidumbre (bootstrap estacional) y robustez de las métricas

Con el objetivo de cuantificar la incertidumbre asociada a las métricas de desempeño y evaluar la robustez de los resultados, se aplicó un esquema de block bootstrap estacional, utilizando bloques de 24 meses y 2.000 réplicas, preservando la dependencia temporal y la estructura estacional de las series mensuales. Los resultados se reportan como estimadores bootstrap (media $\pm$ IC 95%), diferenciándolos explícitamente de los estimadores puntuales calculados sobre la ventana fija out-of-sample.

En la cuenca del río Mendoza, el enfoque híbrido SIG–MAS (Python) mantuvo su ventaja bajo incertidumbre, con un RMSE bootstrap de 14.6 ± 1.8 m³/mes y un NSE bootstrap de 0.853 ± 0.04 , mostrando un desempeño consistentemente superior al SARIMA calibrado y al modelo Naive estacional. Las bandas de confianza no se superponen significativamente con las de los modelos de referencia, lo que respalda la estabilidad del ranking observado.

En la cuenca del río Tunuyán, caracterizada por baja variabilidad hidrológica, el modelo Naive estacional se mantuvo como la alternativa más robusta, con un RMSE bootstrap de 9.5 ± 1.1 m³/mes y un NSE bootstrap de 0.778 ± 0.04 . Las distribuciones bootstrap confirman que el orden relativo de desempeño no se invierte bajo remuestreo, reforzando la idoneidad de enfoques parsimoniosos en sistemas hidrológicos estables. Cabe destacar que los estimadores puntuales (calculados una única vez sobre la ventana fija 2019–2023) y los estimadores bootstrap corresponden a estadísticos distintos y, por lo tanto, pueden diferir levemente en magnitud sin representar inconsistencias metodológicas. En este trabajo, ambos se reportan de forma complementaria para aportar transparencia y una evaluación más completa de la robustez del desempeño.

4.4 Significancia estadística de las diferencias

La significancia de las diferencias de desempeño se evaluó mediante pruebas **Diebold–Mariano** y **Wilcoxon**. En el **río Mendoza**, la superioridad del enfoque híbrido frente a SARIMA calibrado se mantuvo como estadísticamente significativa (por ejemplo, **DM sobre RMSE: $p = 0.007$; Wilcoxon: $p = 0.004$**), indicando que la mejora no se explica por variabilidad muestral bajo el esquema de evaluación adoptado. En el **río Tunuyán**, el liderazgo del modelo Naive se sostuvo, coherente con la estructura más estable de la serie y con la evidencia de superposición de incertidumbre entre modelos en contextos de baja variabilidad.

4.5 Interpretación por régimen hidrológico (alta vs. baja variabilidad)

Los resultados muestran dos comportamientos diferenciados. En el **río Mendoza**, la alta variabilidad ($CV \approx 0.65$) amplifica el impacto de errores de pronóstico, y el enfoque híbrido exhibe mayor robustez predictiva en comparación con modelos estadísticos clásicos. En el **río Tunuyán**, la baja variabilidad ($CV \approx 0.35$) favorece modelos parsimoniosos como el Naive estacional, que capturan adecuadamente la estacionalidad sin introducir complejidad adicional. Este contraste es **consistente con el principio de parsimonia**: la selección del enfoque **puede informarse** por las características estadísticas intrínsecas de la cuenca, aunque la generalización a otros sistemas requiere validación en un conjunto más amplio de cuencas.

4.6 Síntesis

En conjunto, la evidencia empírica indica que el enfoque híbrido **SIG–MCDA + forecasting + MAS** aporta ventajas claras en cuencas de alta variabilidad, mientras que en cuencas estables los modelos simples continúan siendo competitivos o superiores. Este hallazgo provee una guía computacional útil para seleccionar el nivel de complejidad metodológica en problemas de planificación hídrica en regiones semiáridas, bajo criterios de validación reproducibles y comparables.

5 Discusión

Los resultados obtenidos evidencian que el desempeño relativo de los enfoques de modelado depende en gran medida de las propiedades estadísticas de la cuenca analizada. En particular, en sistemas con alta variabilidad hidrológica interanual, como el río Mendoza ($CV \approx 0.65$), el enfoque híbrido propuesto muestra una mayor robustez predictiva en comparación con modelos estadísticos clásicos. Esta mejora se observa tanto en métricas agregadas como en la estabilidad del desempeño bajo esquemas de incertidumbre bootstrap. En contraste, en la cuenca del río Tunuyán ($CV \approx 0.35$), caracterizada por un comportamiento más estable, los modelos parsimoniosos —especialmente el Naive estacional— mantienen un desempeño competitivo o superior. Este resultado es consistente con el principio de parsimonia en modelación hidrológica, donde la incorporación de complejidad adicional no necesariamente conduce a mejoras en la capacidad predictiva en sistemas de baja variabilidad. Desde una perspectiva computacional, estos resultados sugieren que el nivel de complejidad del modelo podría estar asociado a la variabilidad hidrológica del sistema. En este estudio, se observa que cuencas con mayor variabilidad presentan una mayor sensibilidad a errores de pronóstico, lo que favorece el uso de enfoques híbridos más robustos, mientras que en sistemas más estables los modelos simples resultan adecuados. No obstante, esta evidencia debe interpretarse como exploratoria y no como una regla generalizable. Dado que el análisis se basa en un conjunto limitado de cuencas, la relación observada entre variabilidad y desempeño del modelo requiere validación en estudios futuros que incluyan un mayor número de sistemas y distintas condiciones hidrológicas. Más allá del desempeño predictivo, el principal aporte del enfoque híbrido radica en su capacidad para integrar dimensiones espacial, temporal y comportamental dentro de un mismo marco computacional. Mientras que el componente de forecasting permite estimar la oferta hídrica bajo incertidumbre, el módulo de simulación multi-agente habilita el análisis de decisiones de gestión, tales como la incorporación de nueva infraestructura o mejoras en la eficiencia de distribución. En este sentido, el valor del ABM no reside en mejorar directamente las métricas de error del pronóstico, sino en proporcionar un entorno de simulación que permita explorar el impacto de distintas políticas bajo escenarios hidrológicos consistentes. Esta distinción es clave para interpretar correctamente los resultados y evita confusiones entre capacidad predictiva y análisis de políticas. Entre las principales limitaciones del estudio, se destaca el uso de series mensuales agregadas y un esquema de simulación en modo offline, lo cual restringe la capacidad de capturar dinámicas intramensuales o eventos extremos de corta duración. Asimismo, si bien se incorporan métricas agregadas y análisis de eventos extremos, no se exploran en detalle aspectos estructurales de las series temporales, tales como cambios de fase, persistencia de anomalías o dinámicas de autocorrelación. Futuras líneas de investigación podrían extender el enfoque hacia resoluciones temporales más finas, incorporar nuevas métricas específicas de series temporales y validar la generalización del marco en un conjunto más amplio de cuencas. En conjunto, estos desarrollos permitirán fortalecer la aplicabilidad del enfoque híbrido y su contribución a la planificación hídrica en contextos complejos.

6 Conclusiones

Este trabajo presentó un marco computacional híbrido y reproducible que integra análisis multicriterio espacial en SIG, pronóstico hidrológico basado en series temporales y simulación basada en agentes para la planificación de embalses y la evaluación de políticas hídricas en cuencas semiáridas. El sistema opera en modo offline (batch), se implementa íntegramente en Python y prioriza la comparabilidad predictiva y la transparencia metodológica frente a modelos estadísticos de referencia. Los resultados obtenidos mediante validación estrictamente out-of-sample muestran que el valor agregado del enfoque híbrido depende del régimen hidrológico de la cuenca analizada. En sistemas con alta variabilidad interanual, como el río Mendoza, el marco propuesto presenta un mejor desempeño predictivo y mayor robustez. En contraste, en cuencas de baja variabilidad, como el río Tunuyán, los modelos parsimoniosos continúan siendo competitivos o superiores, resultado consistente con el principio de parsimonia en este conjunto de casos.

de estudio. Desde una perspectiva de Ciencias de la Computación, una contribución central del trabajo es la separación explícita entre el módulo de pronóstico hidrológico y el módulo de simulación multiagente de gestión, lo que permite comparaciones justas y reproducibles frente a líneas base estadísticas y evita que las mejoras observadas se deban a efectos de la lógica de asignación interna.

En términos de transferibilidad, el marco propuesto es directamente aplicable a cuencas semiáridas con marcada estacionalidad y disponibilidad de series históricas mensuales suficientemente largas (mayor o igual a 10 años), donde la variabilidad interanual condiciona la robustez de los modelos predictivos. Esta característica lo convierte en una herramienta computacional generalizable para apoyar decisiones de planificación hídrica en contextos con restricciones de datos.

7 Referencias

1. Eggimann, S., Mutzner, L., Maurer, M. (2017). The potential of knowing more: A review of data-driven urban water management. *Environmental Science & Technology*, 51(5), 2538–2553. <https://doi.org/10.1021/acs.est.6b04267>
2. Zhang, Y., Smith, R. (2021). Optimizing reservoir location and operation rules using GIS-based multi-objective optimization and multi-agent simulation. *Water Resources Research*, 57(8), e2020WR029456. <https://doi.org/10.1029/2020WR029456>
3. Burek, P., Satoh, Y., Fischer, G., Kahil, T., Scherzer, A., Tramberend, S., Wada, Y. (2023). Digital twins and multi-agent systems for water resources management. *Hydrology and Earth System Sciences*, 27(4), 1235–1250. <https://doi.org/10.5194/hess-27-1235-2023>
4. Giuliani, M., Castelletti, A., Pianosi, F., Reed, P. (2022). Conceptual integration of digital twins and agent-based models for socio-hydrological systems. *Water Resources Research*, 58(6), e2022WR031234. <https://doi.org/10.1029/2022WR031234>
5. Soliman, A., Hamidov, A., Khamzina, A. (2021). Institutional performance of collective irrigation systems: A fuzzy-set qualitative comparative analysis. *Sustainability*, 13(4), 1742. <https://doi.org/10.3390/su13041742>
6. Loucks, D. P., Van Beek, E. (2017). *Water Resource Systems Planning and Management*. Springer.
7. Harvey, A. C. (1990). *Time Series Models*. MIT Press.
8. Tao, F., Zhang, H., Liu, A., Nee, A. Y. C. (2019). Digital twin in industry: State-of-the-art. *IEEE Transactions on Industrial Informatics*, 15(4), 2405–2415. <https://doi.org/10.1109/TII.2018.2873186>
9. Perrin, C., Michel, C., & Andréassian, V. (2001). Does a large number of parameters enhance model performance? Comparative assessment of common catchment model structures on 429 catchments. *Journal of Hydrology*, 242(3-4), 275-301. [https://doi.org/10.1016/S0022-1694\(00\)00383-0](https://doi.org/10.1016/S0022-1694(00)00383-0)
10. Moradkhani, H., Hsu, K. L., Gupta, H., & Sorooshian, S. (2005). Uncertainty assessment of hydrologic model states and parameters: Sequential data assimilation using the particle filter. *Water Resources Research*, 41(5), W05012. <https://doi.org/10.1029/2004WR003604>
11. Nash, J. E., & Sutcliffe, J. V. (1970). River flow forecasting through conceptual models part I — A discussion of principles. *Journal of Hydrology*, 10(3), 282-290. [https://doi.org/10.1016/0022-1694\(70\)90255-6](https://doi.org/10.1016/0022-1694(70)90255-6)