

# Analysis of Influences on Groundwater Recharge: A Case Study in Contextual Reusability of Big Data Systems

Walter Garrido<sup>1</sup>, Agustina Buccella<sup>1</sup> Alejandra Cechich<sup>1</sup>, and Ayelen Montenegro<sup>2</sup>

<sup>1</sup> GIISCO Research Group, Departamento de Ingeniería de Sistemas, Facultad de Informática, Universidad Nacional del Comahue, Neuquen, Argentina,  
agustina.buccella@fi.uncoma.edu.ar,  
<https://giisco.fi.uncoma.edu.ar/>

<sup>2</sup> Instituto Nacional de Tecnología Agropecuaria (INTA), Alto Valle de Río Negro y Neuquén, Argentina,  
montenegro.ayelen@inta.gob.ar,

**Abstract.** Precision agriculture is a field characterized by the use of technologies to improve various practices associated with crop development. These technologies generate large volumes of data that, when properly managed, can be converted into valuable knowledge for organizations. In this context, this article applies a Big Data Systems (BDS) development methodology, defined in previous work, which allows for the development of application cases in the domain. The objective is to analyze the influence of climatic factors and river flow on the behavior of the groundwater table in the Upper Valley region of Río Negro and Neuquén provinces.

**Keywords:** Precision agriculture, Big Data Systems, Reusability, Context

## Análisis de Influencias en la Recarga de las Napas Freáticas: Un caso de estudio en reusabilidad contextual de sistemas big data

**Resumen.** La agricultura de precisión es un dominio que se caracteriza por la utilización de tecnologías para mejorar las diferentes prácticas asociadas al desarrollo de los cultivos. Estas tecnologías generan grandes volúmenes de datos que, cuando se gestionan de forma adecuada, pueden convertirse en conocimiento valioso para las organizaciones. En este contexto, el presente artículo aplica una metodología de desarrollo de Sistemas de Big Data (SBD), definida en trabajos previos, que permite desarrollar casos de aplicación en el dominio. El objetivo es analizar la influencia de los factores climáticos y del caudal del río en el comportamiento de la napa freática en la región del Alto Valle de Río Negro y

Neuquén.

**Palabras clave:** Agricultura de Precisión, Sistemas Big Data, Reusabilidad, Contexto

## 1 Introducción

Desde la perspectiva de la ingeniería de software, la agricultura de precisión se puede ver como un dominio de trabajo que requiere del soporte de procesos software y tecnologías para evaluar y mejorar las formas y circunstancias de aplicación de las técnicas y prácticas agrícolas. Los grandes volúmenes de datos generados por los distintos dispositivos electrónicos como sensores, artefactos automáticos y manuales, y los propios producidos por el hombre, deben ser soportados por metodologías que permitan su almacenamiento, integración y explotación.

El análisis de esos datos mediante Sistemas de Big Data (SBD) es un enfoque hoy difundido en la construcción de sistemas para la toma de decisiones. Sin embargo, esa construcción puede demandar tiempo y esfuerzo que algunas veces impiden la transferencia y uso de los beneficios que supone la analítica de datos. Como un abordaje a esa problemática, se propone una metodología para reusar SBD basados en datos contextuales [1]. Así, un sistema construido para una región podría adaptarse para ser usado en otra región, con menor esfuerzo y tiempo comparado con un desarrollo completo.

En este trabajo se describe la aplicación de una metodología de desarrollo, presentada en trabajos previos [1, 6], para dar solución a requerimientos solicitados por expertos del INTA en su estación experimental del Alto Valle. En particular, desde el INTA querían *analizar la influencia de los factores climáticos y del caudal del Río Negro en el comportamiento de la napa freática en la región*. Este requerimiento fue abordado contextualmente considerando dos áreas: (1) el período sin riego, comprendido entre mayo y agosto, donde las plantaciones se encuentran en reposo vegetativo, y (2) el período donde hay riego artificial (entre septiembre y abril) debido a la escasez de lluvias.

La sección siguiente presenta trabajos relacionados y antecedentes de nuestra propuesta de SBDs reusables. Luego, la sección 3 desarrolla los casos de estudio. Finalmente, se abordan conclusiones.

## 2 Trabajos Relacionados y Antecedentes

Considerando primero los trabajos relacionados acorde al requerimiento planteado por el INTA, se pueden citar varias propuestas de la literatura. Por ejemplo, en [4] se estableció como objetivo el uso de datos temporales de freatómetros y de dos estaciones meteorológicas para calcular el nivel del agua subterránea en la cuenca del norte de París-Francia y predecir el comportamiento del nivel de las aguas subterráneas. Para esto utilizaron ANN (redes neuronales artificiales - en inglés Artificial Neural Network) con datos provenientes de las mediciones de

freatímetros, lluvia, temperatura y evaporación. Luego, en el trabajo presentado en [2] propusieron el uso de datos temporales obtenidos de freáticos para predecir el nivel de las napas. Para esto utilizaron datos del nivel de agua en pozos cercanos, nivel del río, nivel del canal, precipitación, evapotranspiración, y temperatura del aire, y aplicaron dos modelos de aprendizaje automático: ANN y ANFIS (sistemas de inferencia neuro-difuso adaptativo - en inglés Adaptive Neuro-Fuzzy Inference System). Con los modelos lograron predecir el nivel de las napas con una precisión similar a los modelos físicos y concluir que no hay una diferencia significativa entre ambas técnicas en este estudio. También podemos citar el artículo presentado en [7] donde los autores utilizaron datos temporales obtenidos de freáticos y estaciones meteorológicas con el objetivo de demostrar la aplicabilidad de la filtración fluvial y su potencial en la modelización de las fluctuaciones del nivel de agua. Para esto utilizaron un modelo ANN y datos tales como lluvia, nivel del río, flujo superficial y la temperatura. Los resultados respaldaron la relevancia primordial de la lluvia y el flujo de agua superficial como factores clave que influyen en los niveles de agua subterránea en el área de estudio.

Continuando, en [3] el objetivo fue evaluar la eficacia de los modelos de aprendizaje automático en la predicción del nivel del agua subterránea. Para ello, compararon, clasificaron y evaluaron la incertidumbre asociada a los mismos. Se utilizaron datos sobre el nivel del agua subterránea, precipitación, caudal del río, tasa de evaporación y tasa de extracción de agua subterránea obtenidos de estaciones meteorológicas y freáticos. El análisis de incertidumbre enfatizó la eficacia de los modelos de aprendizaje automático en la predicción del nivel del agua subterránea. Por último, podemos nombrar al artículo presentado en [5], donde se estableció el objetivo de comparar la capacidad predictiva de diferentes modelos de aprendizaje automático y aprendizaje profundo para identificar factores, simular el nivel del agua subterránea y predecir niveles futuros del mismo.

Luego, considerando los trabajos previos [1, 6], en [1] se presentó la metodología desarrollada para la construcción de SBDs en la forma de casos de aplicación nuevos o reusables. Dicha metodología posee dos enfoques de desarrollo, *top-down* y *bottom-up*. Cada uno de ellas depende de la información de la cual se cuenta al iniciar el proceso y de lo que se quiere obtener. Por ejemplo, en el enfoque *bottom-up*, surge de la necesidad de los expertos de conocer qué puedo extraer de los datos sin conocer fehacientemente lo que deseamos encontrar. En cambio, un enfoque *top-down*, parte de hipótesis bien definidas por los expertos para ver si los datos verifican las mismas. En la Figura 1 se muestra el enfoque top-down en donde se puede observar, además de los cinco pasos o actividades, una serie de componentes y recursos que hacen falta describir.

Primero, es necesario remarcar que nuestra metodología toma como base a las líneas de producto de software, donde el desarrollo de los sistemas posee una fase de modelado e implementación de activos del dominio (*ingeniería de dominio*) y una fase de instanciación de dichos activos para la construcción de un caso de aplicación en particular (*ingeniería de aplicación*). Entonces, como

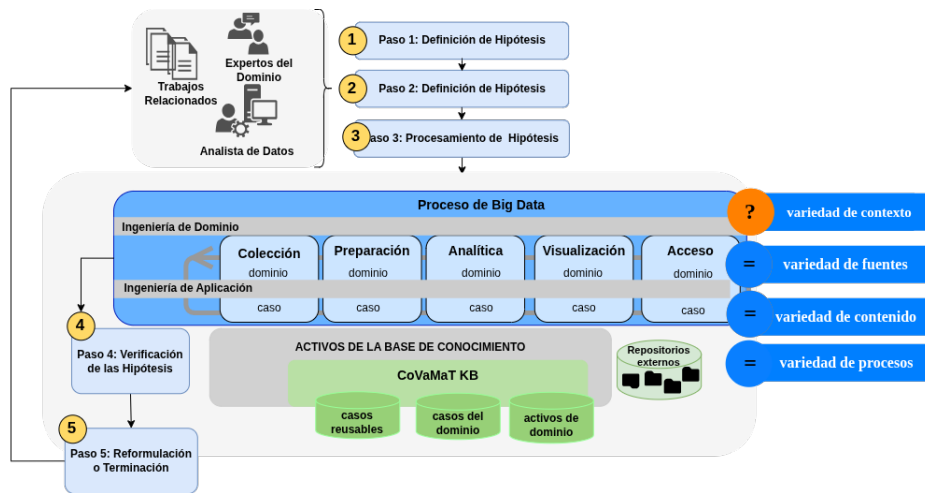


Fig. 1. Enfoque top-down para el desarrollo de SBD reusables

podemos observar en el centro azul de la figura, el *proceso de big data* tiene sus cinco actividades divididas en dos fases: (1) la *ingeniería de dominios* donde se identifican y modelan los activos para todos los casos (SBDs) a ser desarrollados en el dominio, y (2) la *ingeniería de aplicación* donde los activos se instancian para crear casos de aplicación particulares (un SBD). Estas dos fases, junto con las cuatro *variedades* (a la derecha de la figura) definidas, son la base de la reusabilidad.

El concepto de *variedad* en sí mismo, se define como la diversidad o multiplicidad de alternativas dentro de dominios específicos que pueden ser identificados para ser reusados [1]. En particular, los trabajos previos se han centrado en el modelado y gestión de la *variedad* considerando cuatro tipos posibles: *variedad de las fuentes* son los diversos conjuntos de datos que se utilizan para llevar adelante la construcción de un caso de aplicación; *variedad de contenido*, se centra en las variables o factores utilizados, las variables objetivo, los umbrales y otras características relevantes; *variedad de procesos*, se enfoca en las técnicas de análisis y preparación utilizadas para procesar y extraer conocimiento a partir de los datos; y *variedad de contexto* incorpora el contexto semántico al definir las observaciones en función de su entorno, permitiendo diferenciar y agrupar casos similares.

Por último, podemos observar en la figura, en verde, la base de conocimiento junto con su herramienta de soporte denominada *CoVaMaT* (Context-based Variety Management Tool) [6]. La misma permite: (1) definir y documentar la variedad en cada actividad del proceso de big data durante la ingeniería de dominio, y (2) consultar e instanciar la misma durante las actividades de la ingeniería de la aplicación. *CoVaMat* documenta tres componentes principales: (1) los *activos del dominio*, que corresponden a las variantes encontradas dentro

de un dominio en particular, (2) los *casos del dominio* que se refieren a los casos de aplicación instanciados con las variantes identificadas previamente y (3) los *casos reusables* que se refieren a la reutilización de activos o casos del dominio para el desarrollo de un nuevo caso. Estos tres componentes son el eje principal de nuestra metodología.

En la siguiente sección se describe la aplicación del enfoque top-down (Figura 1) mediante un caso de estudio que incluye dos casos de dominio.

### 3 Reusabilidad top-down en la aplicación de los casos

**Paso 1: Definición del Problema del Dominio** En este caso, los requerimientos de los expertos de dominio (INTA) estaban orientados a conocer las causas de las fluctuaciones en las napas freáticas. Para esto, se contaba con datos obtenidos de estaciones meteorológicas, caudales del Río Negro, y registros obtenidos de mediciones de freatómetros instalados.

Al mismo tiempo, se relevó la información provista en los trabajos relacionados (Sección 2), para conocer aplicaciones similares con sus objetivos y logros. Del análisis de los mismos, se destacó la importancia de variables como la temperatura, precipitación, nivel del río, flujo superficial y evaporación. Esas variables fueron fundamentales para comprender los procesos hidrológicos y las fluctuaciones en los niveles freáticos, ya que influyeron directamente en el comportamiento del agua subterránea en diferentes entornos.

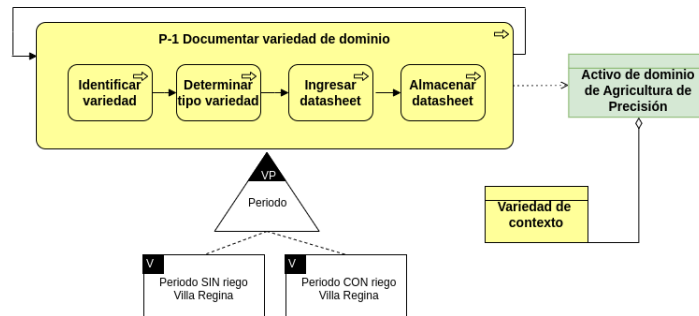
De esta forma, considerando los datos y requerimientos de los usuarios expertos, y del relevamiento de la literatura, se definió como principal objetivo del estudio: *analizar la influencia de los factores climáticos y del caudal del Río Negro en el comportamiento de la napa freática.*

Este objetivo fue enmarcado dentro de un problema de *variedad de contexto*, donde los procesos y activos creados en un primer caso se documentaron y reusaron en el segundo. La variación se definió basada en dos períodos principales: (1) *Caso de Dominio 1 - Período SIN Riego* comprendido entre mayo y agosto. En estos meses, en la zona del Alto Valle, las condiciones productivas no demandan riego ya que los cultivos no se encuentran en su período vegetativo; y (2) *Caso de Dominio 2 - Período CON Riego* comprendido entre septiembre y abril, coincidiendo con la máxima demanda de agua por parte de los cultivos.

Luego, se utilizó el proceso *P-1 (Documentar variedad de dominio)* de CoVaMaT para crear la variedad de contexto con dos variantes: *Período sin riego Villa Regina* y *Período con riego Villa Regina* (Figura 2). Como se puede observar en los diagramas, se utilizaron *datasheets* [1] con la notación usada en los modelos de variabilidad ortogonal.

#### 3.1 Caso de Dominio 1 - Período SIN Riego

**Paso 2: Definición de las Hipótesis.** Para este caso se definieron dos hipótesis enmarcadas en el objetivo principal del caso de estudio. La primera de ellas se refiere al análisis del Río Negro con respecto a la napa freática de



**Fig. 2.** Proceso P-1 para documentar la variedad de contexto como un nuevo *Activo de dominio de Agricultura de Precisión* en CoVaMaT

forma tal de conocer su influencia acorde a los niveles obtenidos de los datos y las distancias en donde fueron obtenidos los mismos. La segunda hipótesis está dirigida a conocer cuáles variables climáticas tienen mayor influencia en la napa junto con la relación con el nivel del río y las distancias. De esta forma, se plantearon:

- **Hipótesis 1 (H1):** ¿El nivel del Río Negro tiene influencia en la recarga de la napa freática? ¿Qué impacto tiene la distancia entre el río y los freáticos en la variación del nivel de la napa?
- **Hipótesis 2 (H2):** ¿Cuáles son las variables climáticas que tienen mayor-menor influencia que el nivel del Río Negro en la fluctuación del nivel freático?

**Pasos 3 y 4: Procesamiento de las Hipótesis.** Una vez definidas las hipótesis se realizaron las cuatro actividades del proceso de big data (en azul en la Figura 1).

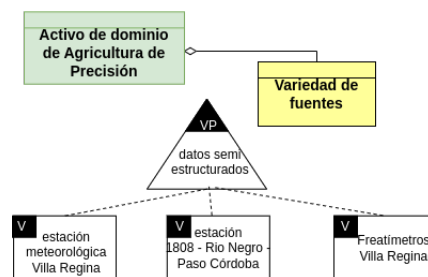
**Proceso de Big Data: Actividad Colección.** El conjunto de datos utilizado se obtuvo de tres fuentes diferentes:

1. *Freatímetros Villa Regina:* La zona de Villa Regina posee instalados 77 freáticos que son revisados y registrados mensualmente. Estos permiten medir el nivel de la napa freática en la zona. Para este caso se tomaron los valores registrados en el período 2010 al 2023, lo que equivale a un total de 168 registros de datos por freático. Los datos poseen 4 variables: *identificador, cota, mes y altura* de la napa en metros.
2. *La estación hidrométrica 1808 - Río Negro - Paso Córdoba:* La estación toma registros del nivel del Río Negro en forma diaria. Para este caso se seleccionaron los valores desde el 01 de enero de 2010 al 31 de Diciembre del

2023; lo que equivale a un total de 5112 registros. Este conjunto de datos posee 4 variables: *fecha*, *altura*, *cota IGN*<sup>3</sup> y *cota IGN corregida*<sup>4</sup>.

3. *La estación meteorológica de Villa Regina*: La estación registra 33 variables climáticas cada 10 minutos. En total, el periodo de datos abarca desde 2010 hasta 2023, acumulando 520715 registros.

Finalmente y para concluir con esta actividad, se utilizó nuevamente el proceso P-1 de CoVaMaT para documentar la variedad de las fuentes dentro del *Activo de dominio de Agricultura de Precisión* creado previamente. En este caso, como podemos observar en la Figura 3, se creó un nuevo punto de variación (*datos semi-estructurados*) que posee una variante por cada una de las fuentes de información.



**Fig. 3.** Proceso P-1 donde se documenta la variedad de las fuentes dentro del *Activo de dominio de Agricultura de Precisión* en CoVaMaT

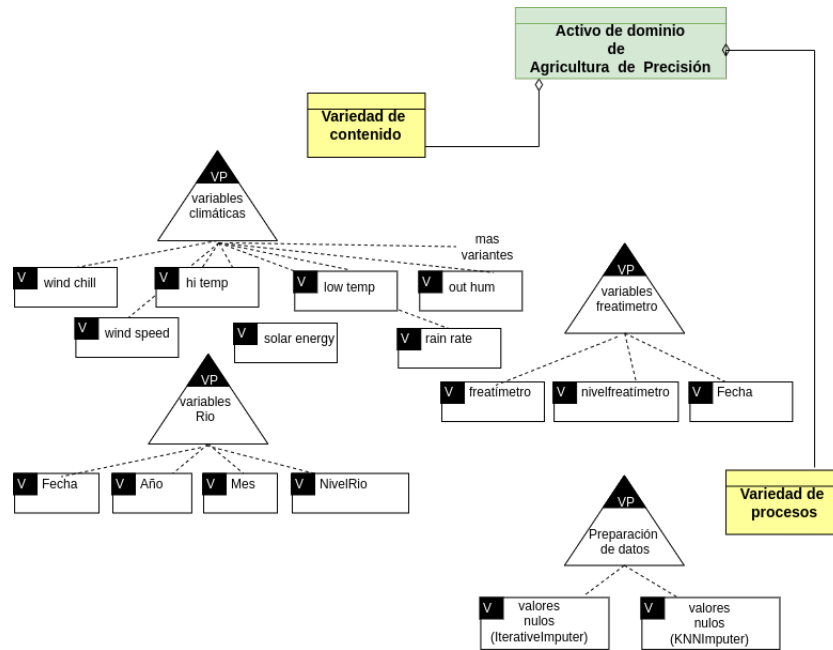
**Proceso de Big Data: Actividad Preparación.** En esta actividad se prepararon los datos obtenidos de las fuentes para su limpieza, normalización y estructuración de forma que sean comparables entre ellas. En primer lugar, se seleccionó la primera fuente con los freaímetros que se encuentran en la zona comprendida entre el Río Negro y un límite de 2 km de distancia desde su cauce<sup>5</sup>. De esta manera, se seleccionaron 19 freaímetros de los cuales se eliminaron valores nulos y algunos freaímetros que no habían tenido registros en más de un año. Como resultado, quedaron 15 freaímetros.

Continuando, para la segunda fuente de datos (nivel del río), se decidió cambiar la escala de días (el caudal se mide diariamente) a meses con el objetivo de poder unificar este conjunto de datos con el creado previamente. Para la tercera fuente de datos se partió del conjunto inicial de 33 variables climáticas registradas por la estación meteorológica ubicada en Villa Regina y lo redujimos a 14 variables. Esta simplificación fue realizada en base a criterios metodológicos,

<sup>3</sup> *msnm*: metros sobre el nivel del mar

<sup>4</sup> *msnm*: cota y la cota corregida miden la altura sobre el nivel medio del mar. La cota corregida se refiere a la cota ajustada teniendo en cuenta factores como errores de medición o cambios en el plano de referencia

<sup>5</sup> criterio que se adoptó por sugerencia de los expertos en el dominio



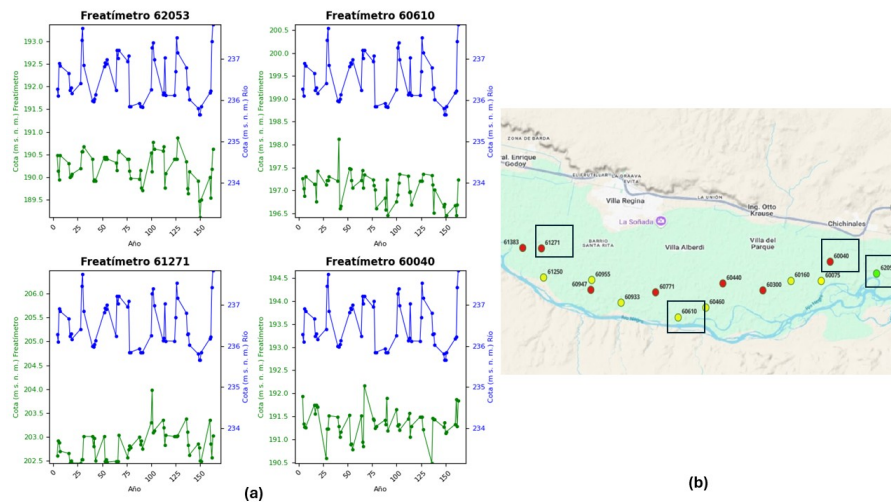
**Fig. 4.** Proceso P-1 donde se documentó la variedad de contenido y de proceso como activo de dominio de Agricultura de Precisión en CoVaMaT

la revisión de trabajos relacionados y el asesoramiento de los expertos en el dominio.

Por último, se volvió a utilizar CoVaMaT para registrar la *variedad de contenido* y la *variedad de proceso* (Figura 4).

**Proceso de Big Data: Actividad Analítica.** En esta actividad, se exploraron los conjuntos de datos con el objetivo de validar las dos hipótesis. Para analizar si el nivel del Río Negro influye en la recarga de la napa freática, se comenzó con un análisis visual para comparar el comportamiento de la misma con el nivel del río, utilizando gráficos que muestran sus fluctuaciones en el período sin riego. A modo de ejemplo, en la Figura 5(a) vemos la fluctuación de 4 de los 15 freaímetros con respecto a la napa en dicho periodo. Cada serie de cada gráfico representa, en azul, la variación del nivel del río, y en verde el nivel de la napa. El primer y segundo gráfico representan los freaímetros 62053 y 60610 con un comportamiento similar entre ambos niveles, lo que indicaría que las fluctuaciones del río tienen un impacto en la dinámica de la napa.

Luego, se analizó el impacto de la correlación con respecto a la distancia de los freaímetros con el río. El mapa de la Figura 5(b) presenta una muestra de la distribución espacial de los mismos, categorizados en tres grupos de acuerdo con su nivel de correlación, siendo: verde (correlación fuerte) mayor a  $|0.5|$ , amarillo (correlación moderada) entre  $|0.5|$  y  $|0.2|$  y rojo (correlación baja) entre  $|0.2|$



**Fig. 5.** Nivel del Río Negro en 4 freaímetros (a) - Distribución geográfica (b)

y 0. De esta forma, se puede identificar una tendencia donde los puntos con correlación fuerte y moderada son los más cercanos al río entre (0-250 m) y (250-900 m), mientras que los más alejados (900-2000 m) suelen tener correlaciones más bajas. Esto sugiere que la influencia del río sobre el nivel freático disminuye a medida que aumenta la distancia.

Sin embargo, el freaímetro 61383, ubicado a 600 m, presenta una correlación negativa (-0.0300), lo que sugiere que en este punto no hay una relación directa entre el nivel del río y el nivel freático. Una posible explicación es que se deba a la composición del suelo o la presencia de otra fuente de recarga, como un canal de riego o infiltración de precipitaciones.

En conclusión, los datos sugieren que la proximidad al río es un factor importante en la intensidad de la correlación entre los niveles del río y la napa freática pero no el único. De esta forma, se abordó la segunda hipótesis para analizar si las variables climáticas podrían también influir. De este estudio, los mayores niveles de asociación encontrados fueron con las variables Humedad (*Out Hum*), Punto de rocío (*Dew Pt.*), y Tasa de precipitación (*Rain Rate*). En particular, la variable climática *Punto de rocío (Dew Pt.)* mostró una relación con el nivel freático, con diferencias notables entre los distintos freaímetros.

En la Tabla 1 se muestra la correlación de dicha variable por cada freaímetro en comparación también con el nivel del río obtenido en la hipótesis anterior. En algunos puntos específicos, como 61271, 60300 y 60771, se observa una correlación moderada positiva, lo que sugiere que un mayor punto de rocío (*Dew Pt.*) podría estar relacionado con un aumento en el nivel freático, posiblemente por reducción de la evaporación o mayor condensación. Sin embargo, en otros freaímetros, como 60610, 61250 y 61383, la correlación es cercana a cero o negativa, lo que sugiere que este efecto no es uniforme en toda la región. También podemos

observar que en 5 de los 15 freaímetros, la correlación entre el nivel freático y el punto de rocío es mayor que la correlación con el nivel del río (comparación de tercera y cuarta columna de la tabla). Este patrón se observa principalmente en freaímetros más alejados del río, como los casos de 60300, 60955, 61271, 60771 y 60440. Por ejemplo, en el freaímetro 60771, que se encuentra a 1900 metros del río, la correlación con el punto de rocío alcanza un valor de 0.41, mientras que con el nivel del río es de apenas 0.19. Del mismo modo, en el freaímetro 61271, a 2000 metros de distancia, el punto de rocío presenta una correlación de 0.33, muy superior a la del río, que es apenas 0.048. Estos casos sugieren que, en zonas donde la correlación del nivel del río es baja o nula, el punto de rocío adquiere un rol más importante como factor asociado a la variación del nivel freático.

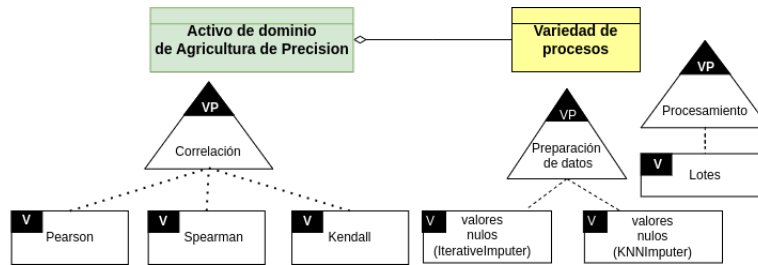
Una vez terminadas de analizar ambas hipótesis, se documentó la nueva *variedad de proceso* implementada en CoVaMaT. Como podemos observar en la Figura 6, estas variedades, que son parte del *Activo de dominio de Agricultura de Precisión*, registran las técnicas de correlación utilizadas.

Freaímetro	Distancia (m)	Correlación Río	Correlación Dew Pt.
62053	250	0.6100	0.06
60075	350	0.2300	-0.09
60610	400	0.4100	-0.04
61250	600	0.3300	0.00
61383	600	-0.0300	-0.03
60460	700	0.3700	0.03
60933	900	0.2000	0.19
60947	1200	0.1500	0.12
60040	1350	-0.0940	0.10
60440	1550	0.1400	0.28
60300	1600	-0.0068	0.44
60955	1900	0.2700	0.28
60771	1900	0.1900	0.41
60160	2000	0.3200	0.03
61271	2000	0.0480	0.33

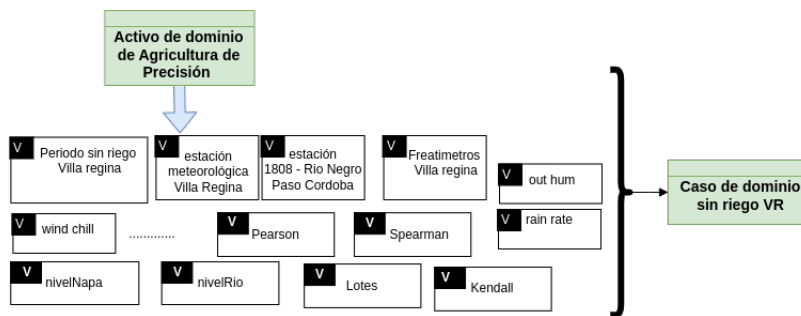
**Table 1.** Comparación de correlaciones del nivel del río y Punto de rocío (Dew Pt.), según la distancia al río.

**Proceso de Big Data: Actividades de Visualización y Acceso.** Esta actividad fue realizada sólo con los gráficos presentados en las actividades anteriores, quedando pendiente la implementación de un sistema de visualización para los expertos.

**Paso 4 y 5: Verificación de las Hipótesis y Terminación.** En estos pasos se realizó el análisis de las hipótesis planteadas. Resumiendo, se destaca una clara tendencia donde: *la influencia del río es más evidente en los puntos*



**Fig. 6.** Proceso P-1 donde se documentó la variedad de proceso como activo de dominio de Agricultura de Precisión en CoVaMaT



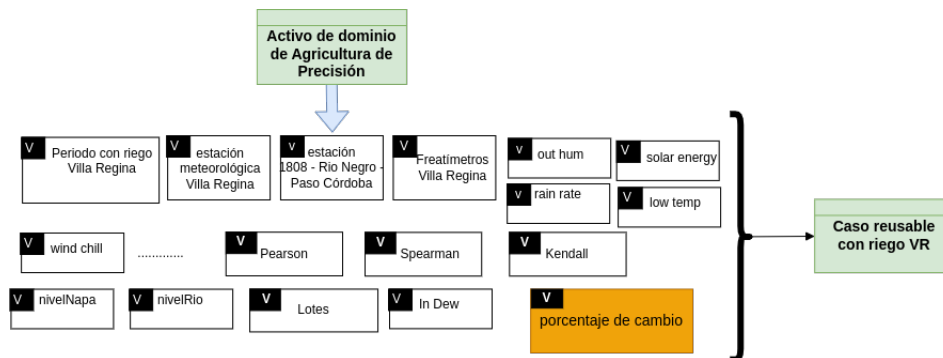
**Fig. 7.** Caso de dominio sin riego VR creado a partir de las instancias del Activo de dominio de Agricultura de Precisión

más cercanos al cauce, disminuyendo a medida que aumenta la distancia. Esto apoya la idea de que el Río Negro ayuda a recargar la napa. Al mismo tiempo, se detecta que existen otros factores influyentes como la humedad y el punto de rocío, que se vuelven evidentes sobre la napa en los lugares más alejados del río.

Por último, se decidió dar por concluido el análisis para este caso de dominio, y se generó en en CoVaMaT. Es decir, hasta el momento se generaron los activos del caso como parte de la *ingeniería de dominio*, realizando cada una de las actividades del proceso y sumándolos al dominio de la Agricultura de Precisión (Figura 1). Luego, y para documentar el *Caso de dominio sin riego VR* y como parte de la *ingeniería de aplicación*, se instanciaron las variedades para seleccionar solo aquellas utilizadas en este caso. Para esto, se utilizó el proceso *P-2 (Definir caso)* de CoVaMaT, el cual recuperó los activos del dominio, y permitió seleccionar las variedades utilizadas y documentarlas (Figura 7).

### 3.2 Caso de Dominio 2 - Período CON Riego

En este caso de dominio, se aplicó la variedad de contexto de forma tal de crear un nuevo caso, a partir del ya creado en el periodo sin riego. Para ello, se tomaron en cuenta los meses desde septiembre a abril, ya que en ese periodo la zona se



**Fig. 8.** Caso reusable con riego VR creado a partir de las instancias del Activo de dominio de Agricultura de Precisión y del primer caso

riega en forma continua mediante una red de canales que se abastecen del canal principal de riego. De esta forma, se aplicaron los cuatro pasos (del 2 al 5) de la metodología top-down, pero instanciando ahora el contexto *Periodo CON riego Villa Regina*.

**Paso 2: Definición de las Hipótesis.** En este caso, se aplicaron las mismas hipótesis ( $H1$  y  $H2$ ) para analizar el comportamiento de la napa freática, pero considerando que la zona de estudio posee riego constante. A su vez, se agregó una nueva hipótesis que involucró la comparación de los resultados de este caso con el caso previo en el periodo sin riego. La nueva hipótesis postulaba:

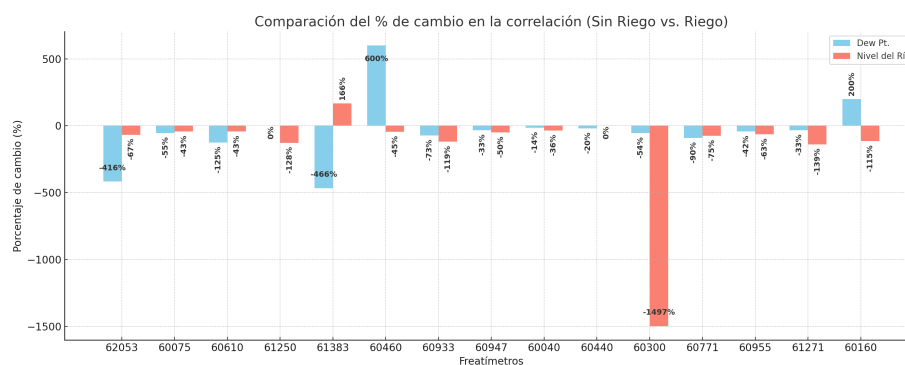
- **Hipótesis 3 ( $H3$ ):** *¿Los factores analizados en el periodo sin riego, poseen la misma influencia en el periodo con riego en la zona de estudio?*

**Pasos 3 y 4: Procesamiento de las Hipótesis.** Para el desarrollo de estos pasos, y por cuestiones de espacio, solo mostraremos los resultados de la  $H3$ , que es quien muestra las diferencias en los dos contextos definidos. Sin embargo, cabe aclarar que para la realización de este caso, se tiene la particularidad de que existe en el repositorio de CoVaMaT el Activo en el dominio de Agricultura de Precisión, con varios activos para cada una de las variedades definidas en cada actividad del proceso de big data. Para esto, se utilizó la funcionalidad  $F-3$  (Funcionalidad para reutilización de casos) y los procesos  $P-3$  (Definir caso de reuso) y  $P-4$  (Consultar activos de dominio en casos similares) para consultar y recuperar los activos del dominio y/o casos de dominio existentes y similares.

La realización de las actividades de colección, preparación y analítica del proceso de big data generó la instanciación de las mismas variantes que se pueden observar en la Figura 8.

La realización de la variante *porcentaje de cambio* corresponde a la respuesta de la  $H3$  y busca responder si las variables analizadas durante el período sin

riego (mayo-agosto) conservan la misma influencia en el período con riego (septiembre-abril). Para ello, el foco de análisis se centró en el nivel del río y el punto de rocío (Dew Pt.). En la Figura 9 podemos observar la comparación de cada freaímetro con respecto al nivel del río-nivel freático (en rojo), punto de rocío-nivel freático (en azul), considerando los períodos sin y con riego. Un valor negativo indica que la correlación disminuye durante el período de riego, mientras que un valor positivo refleja que la relación aumenta. Podemos observar por ejemplo, que el freaímetro 62053, el cual es el más cercano al río, posee un porcentaje de cambio de -67.21 (en color rojo), es decir que el nivel del río tiene un 67% menos de influencia cuando se riega la zona que cuando no se riega. A su vez, los valores de correlación con el punto de rocío (Dew Pt.) (en azul en la figura) muestran un comportamiento más variado. Por ejemplo, el freaímetro 60460 (a 700 metros del río) presenta un incremento de 600%, lo que indica un aumento importante en la influencia del punto de rocío durante el periodo con riego.



**Fig. 9.** Comparación del porcentaje de cambio entre los periodos sin riego y con riego para cada freaímetro

## 4 Conclusiones y Trabajo Futuro

En este artículo, se han desarrollado dos casos de estudio dentro de la agricultura de precisión siguiendo una metodología orientada a maximizar la reusabilidad en la construcción de sistemas de big data. En base a la variedad contextual definida en periodos con y sin riego en la zona de estudio, se ha podido reusar gran parte del trabajo realizado en el primer caso (sin riego) y almacenado como activos del dominio, y reusado para el desarrollo del segundo caso (con riego) como caso reusable. Para lograr el reuso, se ha seguido una metodología propuesta en trabajos previos junto con un repositorio de activos de dominio (CoVaMAT). Dicho repositorio es la clave para el reuso, permitiendo que los

activos de dominio creados en casos ya desarrollados puedan volver a utilizarse en la construcción de nuevos casos. De esta forma se genera un ahorro en tiempo y recursos, permitiendo obtener comparaciones más rápidas y útiles. En particular, el primer caso desarrollado analizó la influencia del Río Negro y de variables climáticas en la napa freática durante el período que no se riega en la zona, para luego crear un nuevo caso donde se analizaron variables similares, pero considerando que la zona se irriga.

Los hallazgos nos permitieron concluir que la dinámica del sistema freático es sensible al contexto temporal, y que tanto el nivel del río como las variables climáticas pueden modificar su grado de influencia según la presencia o ausencia de riego. Por lo tanto, es esencial considerar esta variabilidad al momento de construir modelos predictivos o al diseñar estrategias de gestión del recurso hídrico.

## References

1. Buccella, A., Cechich, A., Saurin, F., Montenegro, A., Rodríguez, A., noz, A.M.: A context-based perspective on frost analysis in reuse-oriented big data-system developments. *Information* **15**(11) (2024). DOI 10.3390/info15110661. URL <https://www.mdpi.com/2078-2489/15/11/661>
2. Djurovic, N., Domazet, M., Stricevic, R., Pocuca, V., Spalevic, V., Pivic, R., Gregoric, E., Domazet, U.: Comparison of groundwater level models based on artificial neural networks and anfis. *The Scientific World Journal* **2015**(742138), 13 (2015). URL <https://doi.org/10.1155/2015/742138>
3. El Bilali, A., Taleb, A., Brouziyne, Y.: Comparing four machine learning model performances in forecasting the alluvial aquifer level in a semi-arid region. *Journal of African Earth Sciences* **181**, 104,244 (2021). DOI <https://doi.org/10.1016/j.jafrearsci.2021.104244>. URL <https://www.sciencedirect.com/science/article/pii/S1464343X2100145X>
4. Lallahem, S., Mania, J., Hani, A., Najjar, Y.: On the use of neural networks to evaluate groundwater levels in fractured media. *Journal of Hydrology* **307**(1), 92–111 (2005). DOI <https://doi.org/10.1016/j.jhydrol.2004.10.005>. URL <https://www.sciencedirect.com/science/article/pii/S0022169404004925>
5. Liu, Q., Gui, D., Zhang, L., Niu, J., Dai, H., Wei, G., Hu, B.X.: Simulation of regional groundwater levels in arid regions using interpretable machine learning models. *Science of The Total Environment* **831**, 154,902 (2022). DOI <https://doi.org/10.1016/j.scitotenv.2022.154902>. URL <https://www.sciencedirect.com/science/article/pii/S0048969722019957>
6. Osycka, L., Cechich, A., Buccella, A., Montenegro, A., Muñoz, A.: Covamat: Functionality for variety reuse through a supporting tool. In: XI Conference on Cloud Computing, Big Data & Emerging Topics (JCC-BD&ET) (2023)
7. Shamsuddin, M.k.n., Mohd Kusin, F., Sulaiman, W., Ramli, M., Tajul Baharuddin, M.F., Adnan, M.S.: Forecasting of groundwater level using artificial neural network by incorporating river recharge and river bank infiltration. *MATEC Web of Conferences* **103**, 04,007 (2017). DOI 10.1051/mateconf/201710304007