

Obtención de similitudes entre eventos como apoyo a la toma de decisiones

Maray Garrido Monagas

E-MAIL: maray@civil.cujae.edu.cu

Universidad Tecnológica de La Habana José Antonio Echeverría (Cujae), Cuba.

Carlos Javier Fernández de Córdoba Webster

E-MAIL: jfernand@etapa.net.ec

Empresa Pública Municipal de Telecomunicaciones, Agua Potable, Alcantarillado y Saneamiento de Cuenca. Universidad del Azuay. Cuenca. Ecuador.

Modesto R. Gómez Crespo

E-MAIL: modesto@tesla.cujae.edu.cu

Universidad Tecnológica de La Habana José Antonio Echeverría (Cujae), Cuba.

RESUMEN

Este artículo tiene como objetivo fundamental ofrecer una visión general del procedimiento para obtener similitudes entre eventos extremos para el apoyo a la toma de decisiones ante la ocurrencia de inundaciones, con la Plataforma modular integrada (PMI). También se muestran los resultados conseguidos en un estudio realizado en la cuenca del río Yanuncay en Ecuador.

PALABRAS CLAVES: Eventos hidrometeorológicos, inundaciones, razonamiento basado en casos, sistemas expertos, toma de decisiones.

Obtaining similarities between events as support for decision making

ABSTRACT

The main objective of this article is to offer an overview of the procedure to obtain similarities between extreme events to support decision-making in the event of floods, with the Integrated Modular Platform (PMI). The results obtained in a study carried out in the Yanuncay river basin in Ecuador are also shown.

KEYWORDS: Hydrometeorological events, floods, case based reasoning, expert systems, decision making.

01 INTRODUCCIÓN

Las inundaciones son fenómenos destructivos. Su estudio, control y prevención, es una prioridad para nuestro país, pues se ve afectado por ellas con frecuencia. El conocimiento temprano de los efectos que puede provocar una inundación, se estudia aplicando diversas herramientas. La adquisición, transmisión y procesamiento de la información de la precipitación, de los niveles y caudales que se alcanzan en los ríos de una cuenca en particular, constituyen la base inicial sobre la cual se construyen los modelos matemáticos que permiten simular tanto el proceso hidrológico lluvia-escurrimiento, como el proceso hidráulico del tránsito de este escurrimiento a través de la red hidrográfica. Este conjunto de herramientas, previamente adaptadas a las características específicas del entorno donde se estudia la inundación, permite una predicción temprana y oportuna que puede llevar a la toma de medidas que protejan las vidas humanas y los recursos materiales del ambiente geográfico en cuestión.

Desde el año 2012, comienza en el Centro de Investigaciones Hidráulicas (CIH) el desarrollo de un sistema (Garrido, 2012), que a partir de la integración de herramientas informáticas, permite la alerta temprana ante el peligro de inundaciones. Este sistema se basa en el principio de trabajo de la Plataforma modular integrada (PMI), desarrollada por (Gómez, 2008) para la gestión de la operación de sistemas fuentes de recursos hidráulicos. El sistema desarrollado en (Garrido, 2012) presenta una solución más completa al estudio de las inundaciones en una cuenca, pues no solo se centra en la modelación, sino que a partir de la integración del sistema de adquisición de datos, los modelos de simulación, el Sistema de información geográfica (SIG) y una base de datos histórica actualizada es capaz de brindar resultados cuando el evento aún se encuentra lejano del territorio.

En este artículo se presenta la aplicación del sistema desarrollado por (Garrido, 2016) en la cuenca del río Yanuncay en Ecuador (Fernández, 2020), el cual a partir de un análisis de similitudes permite brindar apoyo al proceso de toma de decisiones.

02 METODOLOGÍA

EL PROCESO DE TOMA DE DECISIONES.

Se define la toma de decisiones como “el proceso para identificar y solucionar un curso de acción para resolver un problema específico”.

El proceso de toma de decisiones cuenta con diferentes pasos o etapas tal como muestra la figura 1.

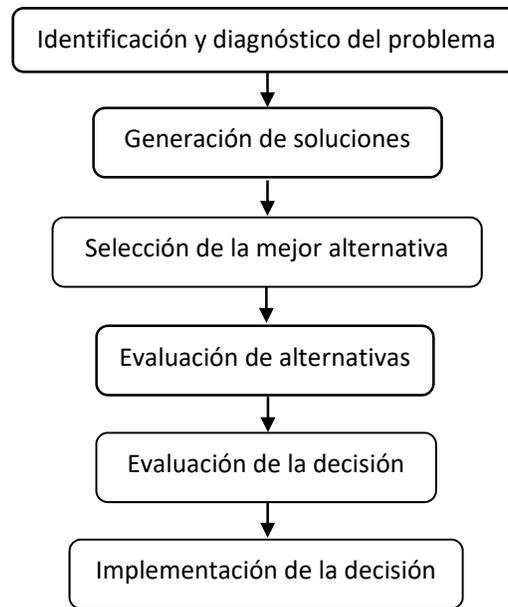


Figura 1. Etapas del proceso de toma de decisiones.

En el proceso de toma de decisiones no siempre se dispone de toda la información requerida y mientras más difícil es la decisión, más complejo resulta el conocimiento de las alternativas; de ahí la necesidad de disponer de mayor información con la calidad necesaria, que aporte elementos de juicio sobre el problema a resolver. De esta manera, se incrementa la probabilidad de que la decisión sea más racional para el logro del objetivo deseado.

De acuerdo con (Gutiérrez et al., 2002) la base del proceso de toma de decisiones es la información que se tiene del dominio de aplicación. A más y mejor información, mayor calidad en la definición del problema, en las propuestas de solución, en el análisis de variantes y en la selección de la acción más conveniente.

LOS SISTEMAS DE RAZONAMIENTO BASADO EN CASOS EN EL APOYO A LA DECISIÓN.

A los sistemas de información interactivos que apoyan al decisor al utilizar datos y modelos para resolver problemas de decisión, se les denomina sistemas de ayuda a la toma de decisiones (Power et al., 2015).

Los sistemas de ayuda a la decisión son una disciplina que utiliza las tecnologías de la información y las comunicaciones para sustentar o apoyar la toma de decisiones. Michael Scott-Morton los definió en la década del 70, como “sistemas que unifican los recursos intelectuales de expertos humanos, con todas las ventajas que brindan los sistemas informáticos para proporcionar decisiones de calidad” (Padrón, 2006).

Los sistemas de ayuda a la decisión tuvieron sus inicios en el área de los negocios, no obstante, se utilizan actualmente en otras ramas, en especial la ingeniería. Pueden clasificarse en dos grandes grupos (Franz, 2012), tal como muestra la figura 2.

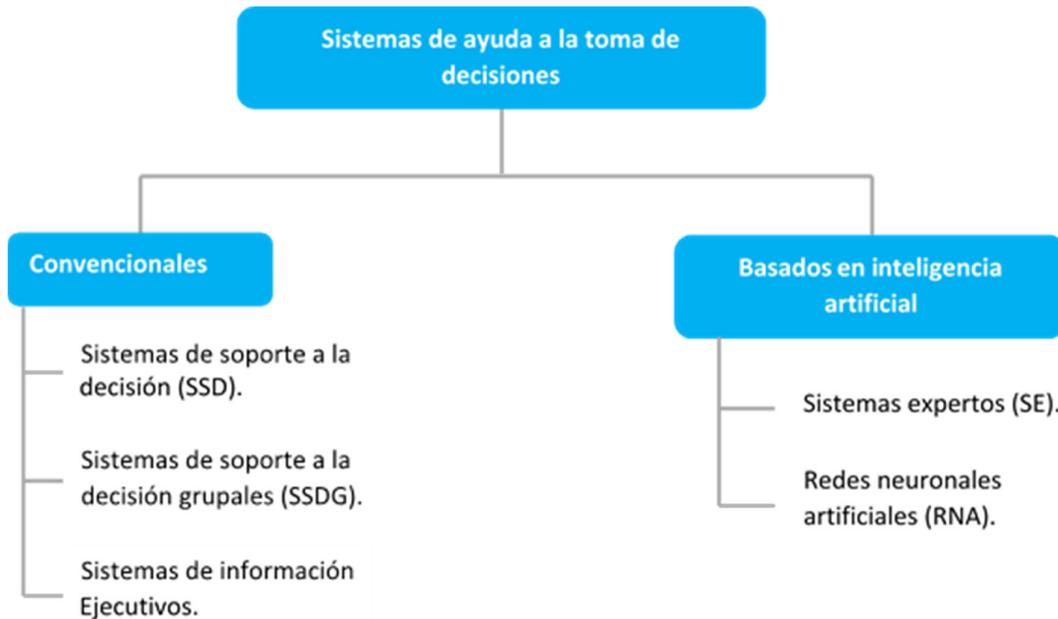


Figura 2. Clasificación de los sistemas de ayuda a la decisión.

Teniendo en cuenta las características de la toma de decisiones ante inundaciones, que es compleja por la cantidad de variables que involucra y con una estructura que no permite una evaluación clara de alternativas por medio de modelos, se decide emplear los sistemas expertos (Castillo et al., 2012) y dentro de ellos los sistemas de razonamiento basado en casos (RBC) para crear una herramienta que permita, a partir del análisis de similitudes apoyar esta importante tarea.

El razonamiento basado en casos se define en (Pal and Shiu, 2004) como un modelo de razonamiento que integra resolución de problemas, entendimiento y aprendizaje con procesos de memoria; tareas que se realizan en base a situaciones típicas, llamadas casos.

En un nivel de abstracción alto los sistemas RBC están compuestos por un mecanismo de razonamiento y tres componentes externos: caso problema, base de casos y solución derivada. El mecanismo de razonamiento se divide internamente en dos partes: el recuperador y el razonador de casos tal como muestra la figura 3.

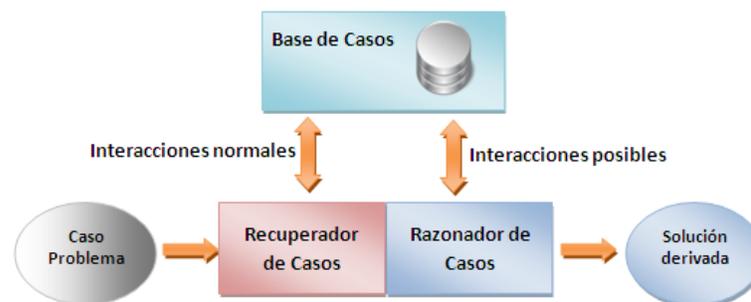


Figura 3. Componentes internos del RBC tomado de (Pal and Shiu, 2004)

La tarea del recuperador de casos es buscar el caso apropiado en la base de casos, mientras que el razonador utiliza los casos recuperados para encontrar una solución a un problema determinado. Este proceso de razonamiento en general, implica tanto la determinación de las diferencias entre los casos recuperados y el caso actual, como la modificación de la solución. El proceso de razonamiento puede, o no, implicar la recuperación de casos adicionales o partes de los casos de la base de casos (Riesbeck and Schank, 2013).

El ciclo de vida para la solución de problemas usando un sistema RBC consta de cuatro estados:

- Recuperación de casos similares partiendo de una base de casos previos.
- Reutilización de casos mediante copia o integración de soluciones a partir de los casos recuperados.
- Revisión o adaptación de las soluciones recuperadas para resolver el nuevo problema.
- Retención de una nueva solución, una vez haya sido confirmada o validada.

Los casos se definen como una pieza de conocimiento contextualizado que representa una experiencia significativa (Pal and Shiu, 2004).

Cuando se define un caso se deben tener en cuenta tres elementos: la representación del problema, el contenido de las soluciones y el resultado de los casos (Riesbeck and Schank, 2013).

La similitud es el concepto fundamental en RBC; esta se puede definir como una relación donde el numerador es el número de atributos que dos objetos tienen en común y donde el denominador es el número total de atributos, tal como se ve en la ecuación 1.

$$similitud_{ep,eq} = \frac{\alpha(A)}{\alpha(A)+\beta(B)} \quad (1)$$

Donde:

A: atributos comunes.

B: atributos diferentes.

α y β : pesos determinados por un algoritmo de aprendizaje, un experto o la fuerza de la relación.

ep, *eq*: representan los casos.

El cálculo de la similitud en un RBC es un aspecto de suma importancia en el proceso de recuperación de los casos; la eficacia de una medida de similitud está determinada por la utilidad de un caso recuperado en la resolución de un nuevo problema.

El sistema que se propone está compuesto por tres módulos tal como muestra la figura 4.

El módulo interfaz de usuario es el encargado de establecer la comunicación entre el usuario y el sistema, en ella se introducen los datos del evento a comparar y luego se muestran los resultados de la búsqueda realizada.

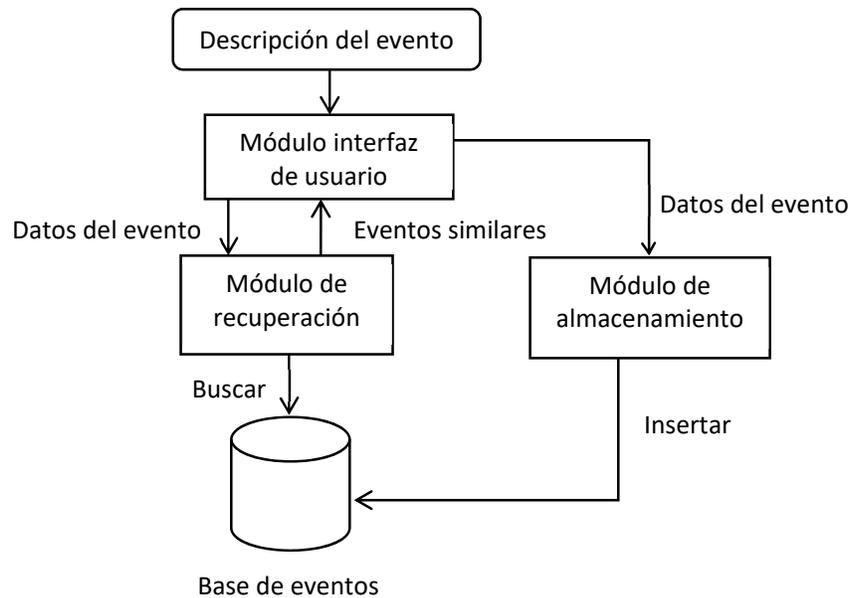


Figura 4. Arquitectura del sistema para la obtención de eventos similares.

En el módulo de recuperación se implementan los mecanismos necesarios para realizar la búsqueda por similitud, una vez obtenidos los resultados se envían al módulo interfaz para mostrarlos al usuario.

El módulo de almacenamiento es el responsable de insertar en la base de eventos los nuevos eventos.

03 EL SAD-PMI EN EL APOYO A LA TOMA DE DECISIONES ANTE LA OCURRENCIA DE INUNDACIONES.

Para lograr el apoyo en las fases iniciales del proceso de toma de decisiones se desarrolló e incorporó a la Plataforma modular integrada desarrollada en el CIH (Garrido, 2016) un sistema de razonamiento basado en casos denominado SAD-PMI.

Sad-PMI es un módulo de la PMI desarrollado con el lenguaje de programación C#, multiplataforma y portable. Brinda un conjunto de funcionalidades que se enumeran a continuación:

- La caracterización de la cuenca.
- El ingreso, modificación y visualización de los eventos.
- La clasificación de los eventos a partir de un proceso de agrupamiento.
- La obtención de eventos similares.

La caracterización de la cuenca incluye la definición del nombre, área, país y las secciones que el usuario desee crear.

Con respecto a los eventos, el sistema permite insertar, modificar y ver los eventos existentes en la base de eventos. La base de eventos es la clave principal de este proceso, pues cuando se implanta la PMI en la cuenca es necesario contar con una base de datos inicial. En esta primera etapa la base de eventos, que está contenida en la base de datos inicial, se comienza a llenar con los eventos obtenidos a partir de los datos de lluvia históricos de la cuenca y luego se va actualizando con los

eventos que modelan los especialistas. Estos últimos pueden ser eventos reales que afecten la región o hipotéticos, con el fin de comprobar el comportamiento de la cuenca ante una situación determinada. En la figura 5 se muestran las vías de actualización de la base de eventos del sistema.



Figura 5. Vías de actualización de la base de eventos.

La clasificación de los eventos, se realiza a través de un proceso de agrupamiento, utilizando el algoritmo k-means (Benítez, 2005). Para ello el usuario selecciona las propiedades del evento por las cuales se realizará el agrupamiento y define el número de grupos que desea crear. Como resultado del proceso, la base de eventos se organiza, ubicando cada evento en un grupo.

El proceso de obtención de similares inicia con la introducción de las propiedades del evento que se desea clasificar.

Se pueden presentar cuatro situaciones fundamentales al analizar la inclusión de un evento en un grupo:

- No cumple con los valores de los atributos del grupo.
- Cumple con todos los atributos del grupo.
- Cumple con al menos un atributo.
- Cumple con más de un atributo.

Los criterios a seguir fueron los siguientes:

- Si no cumple con los valores de los atributos del grupo, no pertenece a él.
- Si cumple con todos los valores de los atributos del grupo, se incluye en este.
- Si hay más de un grupo donde se cumpla con todos los atributos, se aplica un criterio de selección.
- Si no hay grupo donde se cumpla con todos los atributos, seleccionar el grupo donde hay mayor cantidad de coincidencias.
- Si hay más de un grupo donde se presente la situación anterior, se aplica un criterio de selección.

El criterio de selección asumido consiste en el cálculo de la distancia euclidiana que existe entre el evento y los centroides de cada uno de los grupos. El evento se ubica en el grupo con la menor distancia.

En la figura 6 se resume el proceso para la obtención de eventos similares antes descrito.

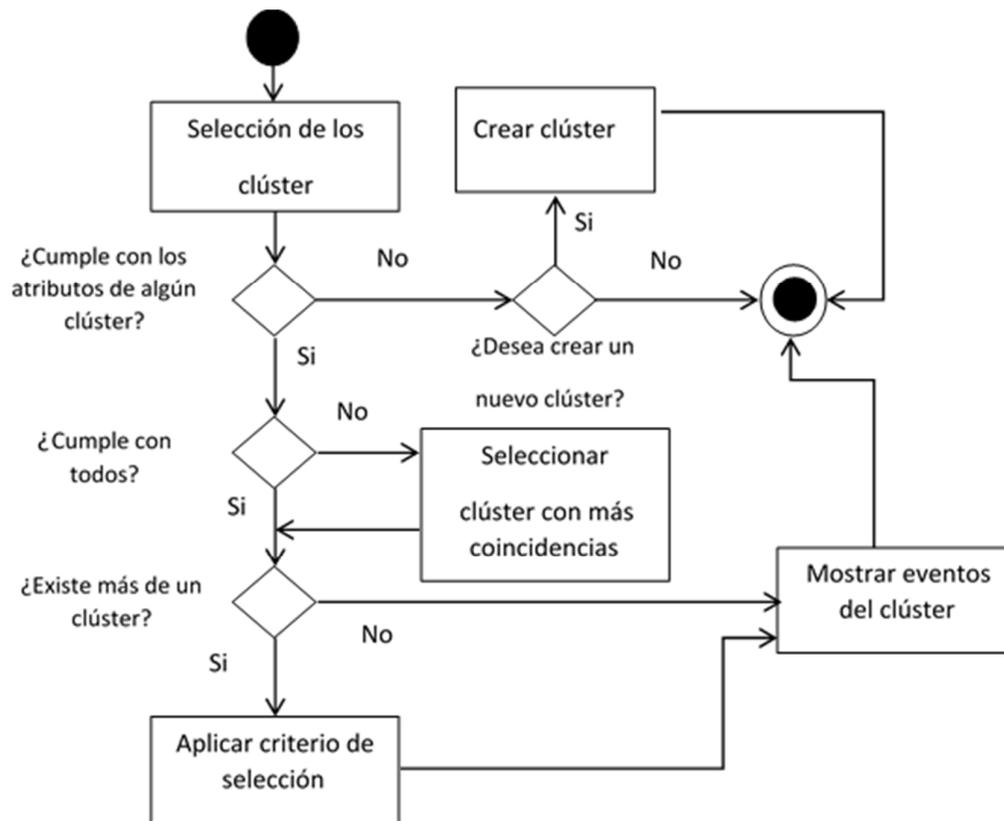


Figura 6. Proceso de obtención de eventos similares

Una vez finalizado el proceso de recuperación se muestra la información de los eventos contenidos en el grupo seleccionado, dando la posibilidad de insertar el evento que se analiza en la base de eventos como parte de este.

APLICACIÓN DEL SAD-PMI EN LA CUENCA DEL RÍO YANUNCAY EN ECUADOR.

El río Yanuncay es uno de los más importantes que atraviesa la ciudad de Cuenca y se utiliza para el suministro de agua potable, el riego, la pesca y la recreación; previéndose para un futuro cercano su empleo para la generación de energía hidroeléctrica. El cauce principal del río tiene una longitud aproximada de 58 km y en su trayecto por la ciudad atraviesa varias comunidades y poblados como son: Soldados, Chictarrumi, Bayán, Sayán, Barabón y San José (Fernández, 2020).

La Empresa Pública Municipal de Telecomunicaciones, Agua Potable, Alcantarillado y Saneamiento de Cuenca – ETAPA EP mantiene y opera una red de estaciones de monitoreo hidrometeorológico en gran parte de la cuenca del río Paute a la que pertenece el río Yanuncay. En la cuenca de río Yanuncay existen cuatro estaciones meteorológicas, un pluviómetro, una estación mixta que registra lluvia, niveles y caudales, y una estación hidrológica que mide niveles y caudales (Fernández, 2020).

En la figura 7, se detalla el proceso seguido para la generación de los eventos que conformarán la base de datos inicial de la PMI implantada en la cuenca del río Yanuncay (Garrido,2020).

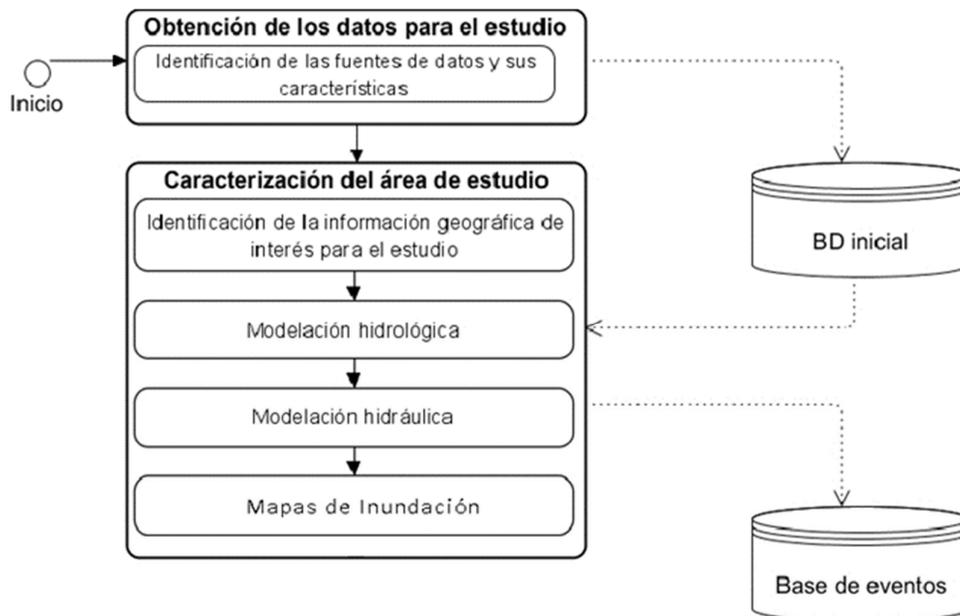


Figura 7. Diagrama general del proceso de obtención de la Base de eventos.

Siguiendo el procedimiento se generaron 170 eventos resumidos en la tabla 1, caracterizados por los atributos siguientes:

- Velocidad de movimiento de la tormenta.
- Dirección de la tormenta.
- Intensidad máxima.
- Lluvia acumulada.
- Duración de la precipitación.
- Condiciones de Precipitación Antecedente.
- Diámetro aproximado de la tormenta.
- Caudal máximo de la estación de Pucán.
- Caudal máximo de la estación de Bomberos.
- Tiempo hasta la máxima inundación desde el centro del hietograma de precipitación de los tres sectores susceptibles a inundaciones.
- Tiempo de regreso del agua al cauce principal desde la máxima inundación de los tres sectores susceptibles a inundaciones.
- Mapas de inundación de los tres sectores susceptibles a inundaciones.

Tabla 1. Eventos generados para la base de datos de la PMI.

Lluvia acumulada (mm)	Número de días de la precipitación antecedente	Precipitación antecedente (mm)	Velocidad de la tormenta (m/s)	Dirección de la tormenta	Intensidad máxima (mm/h)	N° de eventos
20	1	10 y 20	2 y 5	O-E, E-O, NO-SE, NE-SO	27, 28, 34 y 35	28
20	3	20, 40 y 60	2 y 5	O-E, E-O, NO-SE, NE-SO	27, 28, 34 y 35	60
20	5	20, 40 y 60	2 y 5	O-E, E-O, NO-SE, NE-SO	27, 28, 34 y 35	60
30	1	10 y 20	2 y 5	O-E, E-O, NO-SE, NE-SO	41, 42, 50 y 52	22

La cuenca quedó caracterizada tal como muestra la figura 8.

Cuenca

Información general de la cuenca

Nombre: Área (Km2): País:

Secciones de medición de caudal

Nombre	Latitud	Longitud
Pucán	-2.95	-79.17
Barabón	-2.91	-79.08
Bomberos	-2.92	-79.00

Secciones Críticas

Nombre	Latitud	longitud
Barabón	-2.91	-79.08
Av. de las Américas	-2.91	-79.04
Puente Felipe II	-2.91	-79.01

Modificar Cuenca

Figura 8. Caracterización de la cuenca del Río Yanuncay.

Una vez definida la cuenca y conformada la base de eventos, se procede a realizar el proceso de agrupamiento. Para ello se selecciona la opción Agrupamiento del menú principal del SAD-PMI. El sistema muestra una ventana donde el especialista debe indicar la cantidad de grupos que desea crear y escoger los criterios que se tendrán en cuenta en el procedimiento. Como puede observarse en la figura 9 estos criterios son tantos como características tiene cada uno de los eventos y podrán seleccionarse todos o un conjunto de ellos.

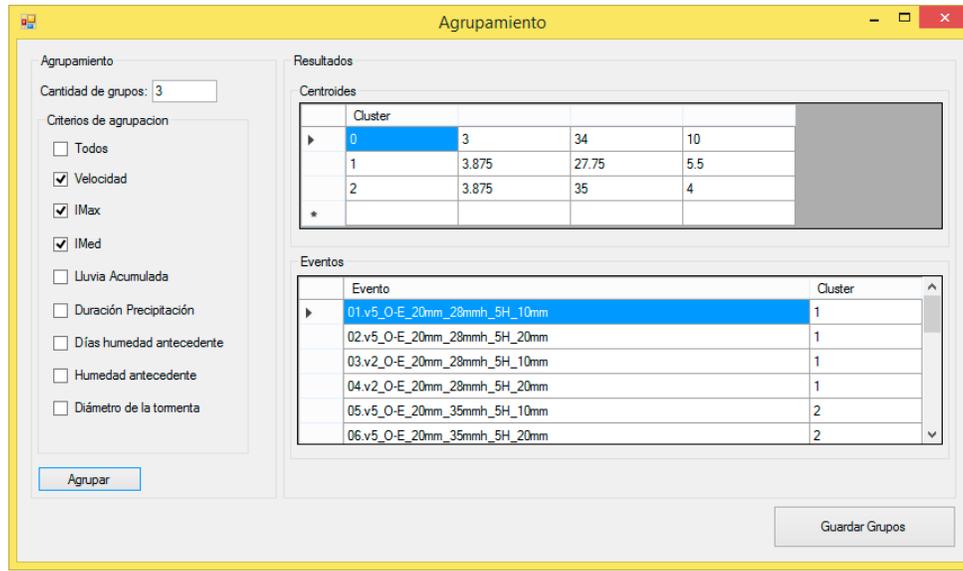


Figura 9. Proceso de agrupamiento realizado en la cuenca del Río Yanuncay.

Como puede observarse, en este caso se crearon 3 grupos y se tuvieron en cuenta la velocidad, intensidad máxima (IMax) e intensidad media (IMed) como criterios de agrupamiento.

Si se ha realizado el agrupamiento al menos una vez, se puede proceder a obtener los eventos similares. Es oportuno señalar que el agrupamiento no tiene que realizarse siempre que se vaya a buscar similitudes entre eventos, pues como se mencionó anteriormente en la base de datos de la PMI queda almacenada la información de los grupos y los eventos que componen cada uno de ellos, por tanto, este procedimiento se realiza inicialmente al conformar la base de eventos y luego a criterio de los especialistas.

Para realizar la búsqueda de similares se selecciona la opción Eventos similares del menú principal del SAD-PMI. El sistema muestra una ventana donde el especialista ingresa los datos del evento que desea buscar, tal como muestra la figura 10.

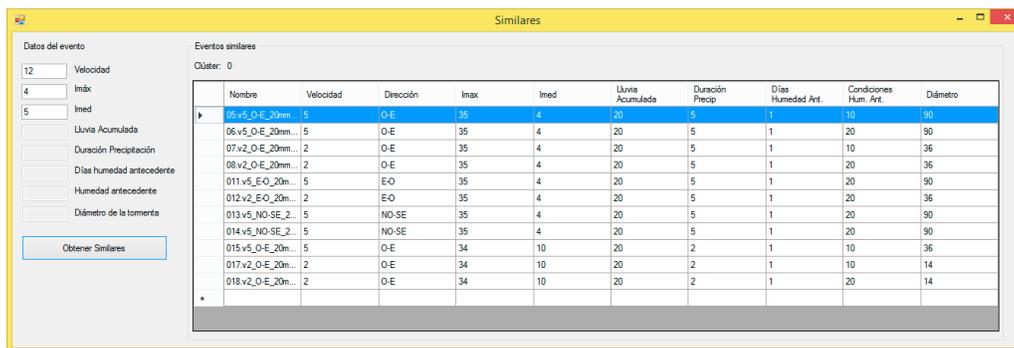


Figura 10. Resultado de la búsqueda de eventos similares.

Como puede apreciarse, en la sección de los datos del evento solo están habilitadas las propiedades que sirvieron como criterio para realizar el agrupamiento. En este caso, teniendo en cuenta los valores de cada una de ellas, el resultado de la búsqueda fue que es similar a los eventos agrupados en el clúster 0, los cuales se listan todos en una tabla.

Al hacer clic sobre cualquiera de ellos, el sistema muestra una ventana donde se detalla la información relacionada con el mismo, tal como muestra la figura 11.

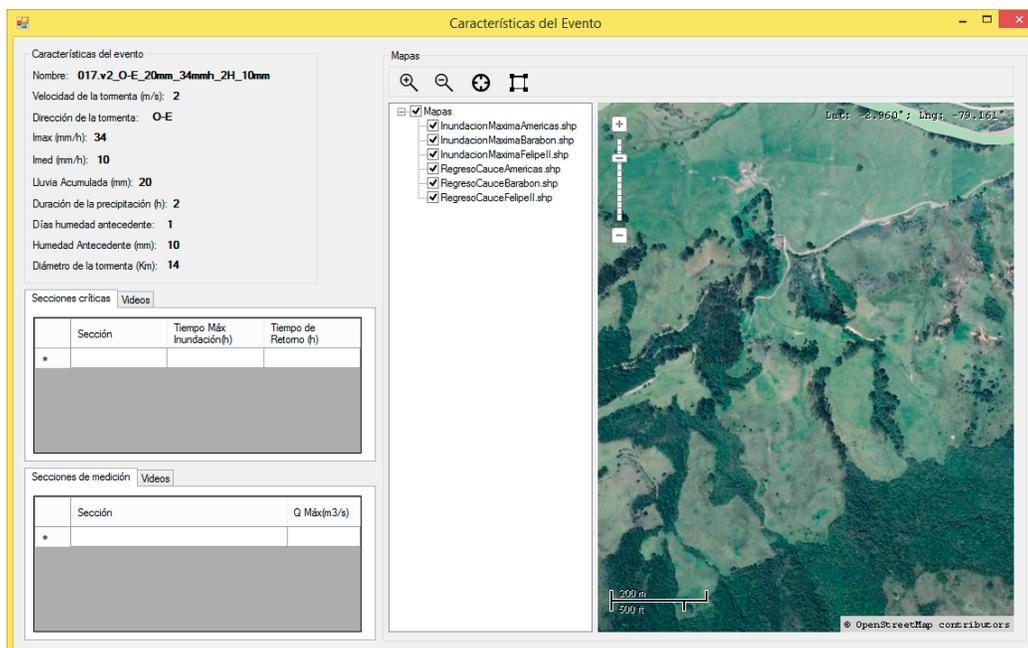


Figura 11. Detalles del evento.

Esta información es de gran importancia para los decisores, pues no solo se muestran las características generales del evento, sino que también puede almacenarse información con respecto a la inundación en las secciones que se definieron como críticas y de medición, incluyendo videos. Además de los mapas generados en los procesos de modelación.

La aplicación del SAD-PMI en la cuenca del Río Yanuncay demostró que la PMI puede ser aplicada para generar una alerta temprana ante inundaciones provocadas por lluvias intensas con relativa confiabilidad, pronosticando inundaciones con un cierto margen de seguridad (Fernández, 2020).

04 CONCLUSIONES

Se reconocen las ventajas de los RBC en cuanto a la facilidad de adquisición de los conocimientos, la flexibilidad en el modelado, la eficiencia en el proceso de razonamiento, y facilidad en el aprendizaje y mantenimiento de la base de conocimientos, por lo que se decide emplearlos en el proceso que propone.

Se propone el SAD-PMI como una herramienta para el apoyo a la toma de decisiones, dado que posibilita contar con la información de cómo se comportó la cuenca ante una situación similar y las medidas que se adoptaron, lo que representa un punto de partida para la toma de decisiones ante una posible inundación.

Se demostró la aplicabilidad del SAD-PMI para las fases iniciales de la toma de decisiones ante el peligro de inundaciones fluviales, al aplicarse en la cuenca del Río Yanuncay en Ecuador, dado que es posible conocer el comportamiento de la cuenca en momentos en que el evento aún está distante del territorio, lo que aporta elementos que perfecciona la identificación de posibles problemas y la generación de alternativas.

05 REFERENCIAS

- Benítez I. S.** (2005). "Técnicas de Agrupamiento para el Análisis de Datos Cuantitativos y Cualitativos". Trabajo de investigación. Universidad Politécnica de Valencia. España. Extraído de: <https://www.researchgate.net>: junio 2016.
- Castillo E., Gutierrez J. and Hadi A.** (2012). "Expert systems and probabilistic network models". Springer Science & Business Media. ISBN: 978-1-4612-2270- 5. New York.
- Fernández de Córdoba C.** (2020). "Propuesta de un sistema de alerta temprana ante inundaciones por lluvias intensas. Caso de estudio: Río Yanuncay en Cuenca - Ecuador". Tesis de doctorado. Facultad de Ingeniería Civil. Universidad Tecnológica de La Habana "José A. Echeverría" (Cujae). La Habana.
- Franz U. R.** (2012). "Decision-making support systems: Theory & practice". Bookboon. Dinamarca. Extraído de: <https://bookboon.com/es/decision-making-support-systems-ebook> en: abril 2015.
- Garrido M.** (2012). "Integración de herramientas informáticas para la alerta temprana ante el peligro de inundaciones". Tesis de maestría. Facultad de Informática. Instituto Superior Politécnico "José A. Echeverría" (Cujae). La Habana.
- Garrido M.** (2016). "Análisis de la información hidrológica para el apoyo a las decisiones ante inundaciones de origen fluvial". Tesis de doctorado. Facultad de Ingeniería Civil. Universidad Tecnológica de La Habana "José A. Echeverría" (Cujae). La Habana.
- Garrido M., Fernández de Córdoba J. y León A.** (2020). Generación de una base de eventos hidrometeorológicos en la cuenca del río Yanuncay. Ingeniería Hidráulica y Ambiental, 41(3), 97–110. ISSN: 2788-6050. La Habana. Cuba. Recuperado a partir de <https://riha.cujae.edu.cu/index.php/riha/article/view/528>
- Gómez M.** (2008). "Plataforma modular integrada para la gestión de campos de pozos". Tesis Doctoral. Facultad de Ingeniería Civil. Instituto Superior Politécnico "José Antonio Echeverría" (Cujae). La Habana.
- Gutiérrez I., Bello R. y Tellería R.** (2002). "Un Sistema basado en casos para la toma de decisiones en condiciones de incertidumbre." Revista Investigación Operacional, 2002. Vol 23. No.2. 19pp. ISSN: 2224- 5405. Departamento de Matemática Aplicada. Facultad de Matemática y Computación. Universidad de la Habana. Cuba.
- Padrón S.** (2006). "Arquitectura de un Sistema Multi-Agente para el soporte a la decisión". Tesis de Maestría. Facultad de Ingeniería Eléctrica. Universidad Tecnológica de La Habana "José Antonio Echeverría" (Cujae). La Habana. Cuba.
- Pal S. K. and Shiu S. C. K.** (2004). "Foundations of soft case-based reasoning". Wiley series on intelligent systems. ISBN 0-471-08635-5. John Wiley & Sons Inc. Hoboken. New Jersey. Estados Unidos.
- Power D. J., Sharda R. and Burstein F.** (2015). "Decision support systems". Wiley Encyclopedia of Management. Vol 7:1–4. ISBN-online: 9781118785317. Estados Unidos.

Riesbeck C. K. and Schank R. C. (2013). " Inside case-based reasoning". The Institute for the Learning Sciences. Northwestern University. Evanston Laurence Earlbaum associates publishers. Illinois. Estados Unidos. Consultado desde: <https://books.google.com.cu> en: mayo 2015.

CONFLICTO DE INTERESES

Los autores declaran que no existen conflictos de intereses.

CONTRIBUCIÓN DE LOS AUTORES

Maray Garrido Monagas <https://orcid.org/0000-0003-0349-7676>

Participó en el diseño de la investigación y en el estudio y selección de los aspectos a tratar. Realizó contribuciones en la interpretación y procesamiento de los datos. Participó en la redacción y revisión del artículo.

Carlos Javier Fernández de Córdova Webster <http://orcid.org/0000-0003-3288-6822>

Realizó contribuciones en la interpretación y procesamiento de los datos.

Modesto R. Gómez Crespo <https://orcid.org/0000-0002-9220-3942>

Participó en el diseño de la investigación.