

Predicción de inundaciones fluviales en un núcleo costero mediante un modelo de regresión de Máquinas de Vectores Soporte de Mínimos Cuadrados (LS-SVM)

María Bermúdez^{a1}, Luis Cea^{a2} y Jerónimo Puertas^{a3}

^aGrupo de Ingeniería del Agua y del Medio Ambiente, ETS de Ingenieros de Caminos, Canales y Puertos, Universidade da Coruña, Campus de Elviña s/n. E-mail: ^{a1}maria.bermudez@udc.es, ^{a2}luis.cea@udc.es, ^{a3}jeronimo.puertas@udc.es

Línea temática | Tema Monográfico

RESUMEN

En este trabajo se propone y aplica un modelo de regresión no paramétrico basado en máquinas de vectores soporte de mínimos cuadrados (modelo LS-SVM) como metamodelo de un modelo de aguas someras bidimensional (modelo SWE) para el estudio de inundaciones. Como caso de estudio se considera el núcleo urbano costero de Vilagarcía de Arousa (Pontevedra). El modelo utiliza los hidrogramas y el nivel de marea prescritos en los contornos del modelo SWE como variables predictoras para estimar los calados y velocidades máximas. Las predicciones obtenidas muestran un buen ajuste con los resultados del modelo SWE, a pesar de la complejidad y la diversidad de las relaciones entre las variables predictoras y la variable respuesta en los distintos puntos. Los tiempos de cálculo son del orden de unos segundos, lo que supone una reducción del orden de 1000 veces con respecto al modelo SWE.

Palabras clave | modelo de inundación; modelo Iber; ecuaciones de aguas someras; máquinas de vectores soporte.

INTRODUCCIÓN

Los modelos bidimensionales de aguas someras han demostrado en los últimos años unas buenas capacidades para la predicción de inundaciones en zonas urbanas. Sin embargo, a pesar de que las técnicas de computación de alto rendimiento han reducido muy significativamente los tiempos de cálculo, este tipo de modelos continúan siendo poco prácticos para aplicaciones que requieren simulaciones de larga duración, resultados en un tiempo muy reducido (previsión en tiempo real), o un gran número de ejecuciones (técnicas de calibración tipo Monte Carlo). Para este tipo de aplicaciones una posible forma de reducir significativamente el tiempo de computación es el desarrollo de modelos sustitutos ('surrogate models') computacionalmente más eficientes de estos modelos.

Un primer enfoque es el desarrollo de modelos físicos de menor fidelidad, que son modelos simplificados del sistema original (Razavi et al., 2012). En el contexto de la modelización de inundaciones urbanas, esto suele implicar el uso de mallas gruesas y el desarrollo de modelos submalla que incorporan la variabilidad topográfica que no es posible resolver con la malla (Cea y Vazquez-Cendón, 2009; Schubert y Sanders, 2012). En estudios que requieren un dominio computacional extenso también es posible emplear modelos anidados, de forma que se consiga una resolución espacial alta en las áreas de interés sin incurrir en el coste computacional de una malla fina en todo el dominio (Bermúdez et al., 2017; Comer et al., 2017). Otros trabajos han adoptado un modelo de autómatas celulares para minimizar la necesidad de resolver ecuaciones complejas (Ghimire et al., 2013; Guidolin et al., 2016). En lugar de las ecuaciones de aguas someras, se emplean reglas de transición (por ejemplo, la fórmula de Manning) para determinar el movimiento del flujo desde una célula de malla a otra.

Para muchas aplicaciones, la mejora computacional conseguida mediante un modelo físico de menor fidelidad puede ser insuficiente. Además, en muchas aplicaciones el principal objetivo es obtener predicciones precisas de inundación, con un interés limitado en los procesos físicos representados en el modelo original. En este contexto surge la posibilidad de desarrollar metamodelos guiados por datos, computacionalmente más eficientes, que relacionan directamente datos de entrada y salida sin analizar los procesos físicos involucrados. Las técnicas de inteligencia artificial como las máquinas de vectores soporte (SVM) o las redes de neuronas artificiales ya han sido aplicadas con éxito para la modelización de la transformación lluvia-escorrentía y la predicción de caudales (véase la revisión de Yaseen et al., 2015), pero apenas se encuentran en la literatura aplicaciones a la modelización de inundaciones. Liu y Pender (2015) investigaron el uso de la regresión de máquinas de vectores soporte para predecir la evolución del calado y velocidad obtenidos a partir de un modelo de aguas someras de malla fina en distintos puntos. El método propuesto obtuvo buenas predicciones y proporcionó considerables ahorros en tiempo de cálculo. En los trabajos de Lin et al. (2013) y Jhong et al. (2017) las SVM se utilizan para predecir la profundidad de inundación en diversos puntos de control, que posteriormente se traducen en un mapa de inundación en base a información geográfica. Se demostró que el método era capaz de generar predicciones de inundación precisas con suficiente tiempo de preaviso. Estos estudios han demostrado que este tipo de modelos, una vez calibrados convenientemente, pueden resultar en predicciones comparables a los modelos de base física. En comparación con otros métodos de inteligencia artificial, las SVM han demostrado una alta precisión y eficiencia, lo que significa que pueden ser más rápidamente entrenados con datos en tiempo real y son más adecuados para su integración en sistemas de alerta de desastres (Jhong et al., 2017; Lin et al., 2013). En Raghavendra y Deka (2014) puede encontrarse una revisión de aplicaciones de las SVM en el amplio campo de la hidrología.

En este trabajo se profundiza en la idoneidad de las técnicas de inteligencia artificial para la simulación de inundaciones. Se propone y aplica un modelo de regresión no paramétrico basado en máquinas de vectores soporte de mínimos cuadrados (modelo LS-SVM) como metamodelo de un modelo de aguas someras (modelo SWE) para el estudio de inundaciones. Como caso de estudio se considera el núcleo urbano costero de Vilagarcía de Arousa (Pontevedra).

METODOLOGÍA

Caso de estudio

El núcleo urbano costero de Vilagarcía de Arousa (Pontevedra) está frecuentemente afectado por inundaciones. Está, de hecho, clasificado como Área con Riesgo Potencial de Inundación dentro del Plan Hidrológico de Galicia Costa. Se encuentra en la desembocadura del río Con en la Ría de Arousa. La cuenca del río Con tiene un área de 24 km² y su topografía es relativamente abrupta, siendo la cota máxima del terreno de aproximadamente 640 m.s.n.m. Se divide en tres subcuentas, siendo la principal la del río Con que abarca toda la parte central y Norte de la cuenca, con una extensión de aproximadamente 16.3 km² (salida Q1 en la Figura 1). Las características hidrológicas de la cuenca, con un bajo tiempo de concentración y pendientes pronunciadas, así como el nivel de marea, juegan un papel en la inundación del área urbana.

Modelo de aguas someras bidimensional

El modelo de aguas someras del área urbana de Vilagarcía Arousa (Pontevedra) está basado en el software Iber (Bladé et al., 2014). Dicho modelo resuelve las ecuaciones de aguas someras bidimensionales para calcular el calado y las dos componentes horizontales de la velocidad promediada en profundidad. En Bladé et al. (2014) puede encontrarse una descripción detallada del modelo, junto con ejemplos de aplicación y validación.

El modelo incorpora la complejidad topográfica del área urbana y representa los flujos a escala de cada edificio. Para ello el dominio se discretiza mediante una malla no estructurada formada por ~112000 elementos triangulares, con tamaños de entre 1 y 10 m. La topografía se define a partir de un modelo digital de elevaciones (MDE) obtenido mediante LIDAR con una resolución de 1 m. En el MDE se integra la batimetría de los principales cauces y la geometría de las estructuras y edificios,

representados estos últimos como huecos en la malla. La fricción de fondo se calcula con la fórmula de Manning, utilizando valores típicos para el coeficiente de rugosidad: $0.025 \text{ s}\cdot\text{m}^{-1/3}$ para los cauces, $0.032 \text{ s}\cdot\text{m}^{-1/3}$ para las zonas con vegetación y $0.020 \text{ s}\cdot\text{m}^{-1/3}$ para las zonas pavimentadas.

En los contornos abiertos de aguas arriba, se imponen los hidrogramas de los tres principales cursos de agua de la cuenca (Q1, Q2 y Q3 en la Figura 1). En el contorno de aguas abajo se fija la serie temporal de nivel del mar.

Generación de datos para los modelos de regresión

Para generar los datos necesarios para la calibración y validación del metamodelo, se simulan un total de 100 eventos de inundación en el modelo 2D-SWE. Cada ejecución del modelo viene definida por tres series temporales de caudal (los hidrogramas aguas arriba) y una serie temporal de nivel del mar (la marea). Los hidrogramas se obtienen mediante un modelo hidrológico distribuido de la cuenca del río Con, utilizando series temporales de precipitación y dos parámetros constantes, la abstracción inicial y la infiltración potencial. Este modelo hidrológico proporciona datos de entrada realistas para el modelo hidráulico. Los caudales máximos oscilan entre 1 y $106 \text{ m}^3/\text{s}$ para Q1, de 0.3 a $24 \text{ m}^3/\text{s}$ para Q2 y de 0.5 a $22 \text{ m}^3/\text{s}$ para Q3 (Figura 1). En el contorno del mar, se definen cuatro series temporales de nivel de mar, correspondientes a marea viva o marea muerta (3 m frente a 1.8 m de carrera de marea) y comenzando en marea vaciante o marea creciente (desplazamiento temporal de 6 horas).

El calado y la velocidad máxima durante la inundación obtenidas con el modelo 2D-SWE constituyen la variable dependiente de los modelos de regresión. Para cada ejecución del modelo, se extraen mapas ráster de estas variables con una resolución espacial de 1 m. Las variables predictoras (o independientes) se obtienen de los datos de caudal y marea prescritos en los límites del dominio 2D-SWE. Se consideraron diferentes vectores de entrada para desarrollar los modelos de regresión y se evaluó su desempeño en Cea et al. (2016). En base a este análisis previo, se selecciona una combinación de cuatro parámetros. Para cada uno de los tres hidrogramas, se toma el caudal máximo que se supera durante una hora como variable predictora ($Q1_{\text{max}}$, $Q2_{\text{max}}$ y $Q3_{\text{max}}$). En el caso de las series temporales de nivel del mar, se considera el valor máximo (Z_{max}). Estas variables se normalizan entre [0,1] con una transformación lineal simple.

Con el fin de dividir los datos de entrada entre validación y calibración, se utiliza un algoritmo de máxima disimilitud (MDA) (Kennard y Stone, 1969). El objetivo del MDA es seleccionar un subconjunto de una base de datos que represente la diversidad de los datos. Esto se hace maximizando la disimilitud entre los vectores en el subconjunto. En este trabajo se elige la versión MaxMin del algoritmo (Willet, 1996) y se utiliza el algoritmo desarrollado por Polinsky et al. (1996) para calcular la disimilitud entre un vector y un subconjunto.

Modelo de regresión basado en máquinas de vectores soporte de mínimos cuadrados

El objetivo de los modelos de regresión es evaluar si una variable dependiente puede predecirse a partir de un conjunto de variables independientes (o predictoras). En este caso, la variable dependiente Y es el calado máximo (h) o la velocidad máxima (v) calculada por el modelo 2D-SWE en cada punto del área urbana. El vector predictor es $X=(Q1_{\text{max}}, Q2_{\text{max}}, Q3_{\text{max}}, Z_{\text{max}})$. Se seleccionaron aproximadamente 25000 puntos uniformemente distribuidos en el área de estudio y se desarrolló un modelo de regresión no paramétrico basado en máquinas de vectores soporte de mínimos cuadrados (Suykens et al., 2012) para cada punto. Este tipo de modelos de regresión provienen del modelo de máquinas de vectores soporte original (Vapnik, 1998), una técnica de aprendizaje basada en transformaciones con núcleos ('kernel-based learning machines') para la solución de problemas de clasificación lineal y no-lineal. Para una descripción detallada de estos modelos el lector puede consultar Suykens et al. (2012) En este trabajo los modelos fueron desarrollados utilizando en software Matlab (versión 2011a) y la toolbox StatLSSVM (Brabanter et al., 2013).

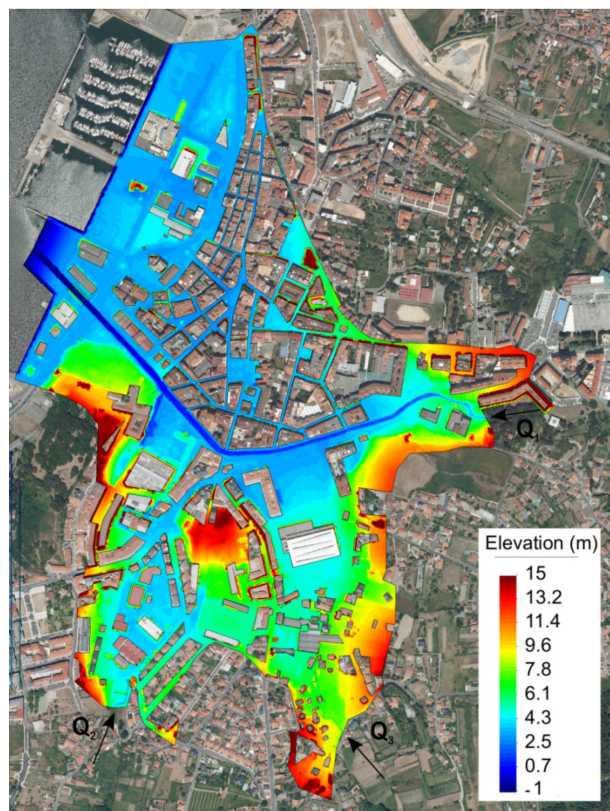


Figura 1 | Batimetría y topografía del modelo de aguas someras e imagen aérea del área de estudio. Las flechas indican las condiciones de contorno de aguas arriba del modelo, que se corresponden con las salidas de los tres cursos de agua principales de la cuenca del río Con.

RESULTADOS Y DISCUSIÓN

Número de ejecuciones de calibración

El algoritmo de máxima disimilitud se utilizó para dividir los datos de entrada en conjuntos de calibración y validación. Con el fin de determinar el número óptimo de ejecuciones de calibración a seleccionar, se realizaron varias pruebas cambiando el número de ejecuciones de calibración (k) de 5 a 50. As restantes ejecuciones se utilizaron para validar el modelo generado. El desempeño global del modelo se cuantifica mediante el error absoluto medio global, considerando todos los puntos, en calibración ($MAE_{global-cal}$) y en validación ($MAE_{global-val}$), definidos como:

$$\begin{aligned}
 MAE_{global-cal} &= \frac{1}{k} \sum_{j=1}^k \left(\frac{1}{n} \sum_{i=1}^n |\hat{Y}_{i,j} - Y_{i,j}| \right) \\
 MAE_{global-val} &= \frac{1}{100-k} \sum_{j=1}^{100-k} \left(\frac{1}{n} \sum_{i=1}^n |\hat{Y}_{i,j} - Y_{i,j}| \right)
 \end{aligned} \tag{1}$$

Donde $Y_{i,j}$ es el calado o velocidad predicha por el modelo LS-SVM en el punto i para la ejecución j , $\hat{Y}_{i,j}$ es el calado o velocidad calculado por el modelo 2D-SWE en el punto i para la ejecución j , n es el número de puntos y k es el número de ejecuciones de calibración.

En la fase de calibración, el error absoluto medio global de las predicciones de calado y velocidad muestra muy poca variación con el número de ejecuciones de calibración, siendo del orden de 0.015 m y 0.02 m/s, respectivamente (Figura 2). En fase de validación, el error aumenta a medida que disminuye el número de ejecuciones de calibración, variando entre 0.016 m ($k=50$) y 0.053 m ($k=5$) y desde 0.019 m/s ($k=50$) hasta 0.050 m/s ($k=5$) para las predicciones de calado y velocidad, V Jornadas de Ingeniería del Agua. 24-26 de Octubre. A Coruña

respectivamente. Sin embargo, los errores permanecen prácticamente constantes en el intervalo $k=40-50$, por lo que no se logra una mejora significativa si el número de ejecuciones de calibración es superior a 40 (Figura 2). En base a este análisis, los modelos LS-SVM se calibran con los datos de 40 eventos de inundación y se validan usando los 60 eventos de inundación restantes.

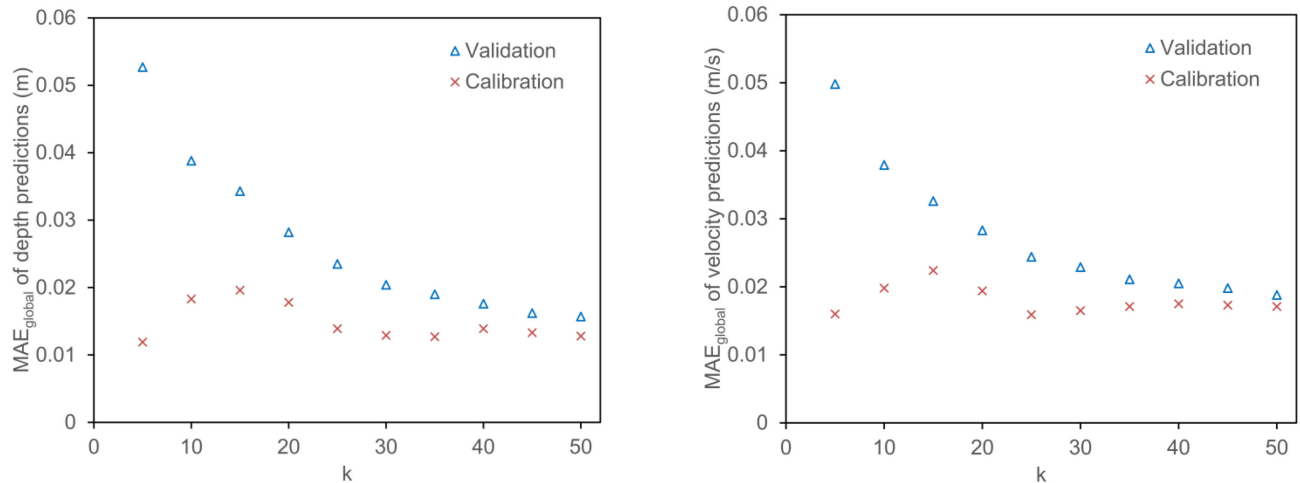


Figura 2 | Error absoluto medio global de las predicciones de calado (izquierda) y velocidad (derecha) en calibración y validación en función del número de ejecuciones de calibración.

Desempeño del modelo

Las predicciones en todos los puntos para las ejecuciones de calibración y validación se muestran en la Figura 3. En estas gráficas se comparan los calados y velocidades calculados por el modelo LS-SVM y el modelo 2D-SWE. Muestran una concordancia razonablemente buena entre los dos modelos. En el caso de las predicciones de calado, los puntos se distribuyen uniformemente alrededor de la línea 1:1 y la dispersión es bastante uniforme para todo el rango de calados. El modelo LS-SVM sobrestima ligeramente los calados en el rango superior ($h > 3$ m) en validación, pero la dispersión es por lo demás bastante simétrica, con aproximadamente el mismo número de puntos por encima y por debajo de la línea 1:1. De hecho, si se ajusta una regresión lineal pasando por el origen se obtiene un coeficiente de regresión (pendiente) de 0.99 tanto en validación como en calibración. En el caso de las predicciones de velocidad, los datos también siguen la línea 1:1. La dispersión es mayor en el rango inferior de velocidades ($v < 2$ m/s), pero no hay una subestimación o sobreestimación obvia en todo el rango de velocidades. La recta de regresión está muy cerca de la línea 1:1, con un coeficiente de regresión de 0.99 y 0.97 en calibración y validación, respectivamente.

El desempeño del modelo LS-SVM en cada punto en las fases de calibración y validación se cuantificó utilizando el error absoluto medio, definido de la siguiente forma:

$$\begin{aligned} \text{MAE}_{i-\text{cal}} &= \frac{1}{k} \sum_{j=1}^k \left| \hat{Y}_{i,j} - Y_{i,j} \right| \\ \text{MAE}_{i-\text{val}} &= \frac{1}{100-k} \sum_{j=1}^{100-k} \left| \hat{Y}_{i,j} - Y_{i,j} \right| \end{aligned} \quad (2)$$

Donde $Y_{i,j}$ es el calado o velocidad predicha por el modelo LS-SVM en el punto i para la ejecución j , $\hat{Y}_{i,j}$ es el calado o velocidad calculado por el modelo 2D-SWE en el punto i para la ejecución j y k es el número de ejecuciones de calibración.

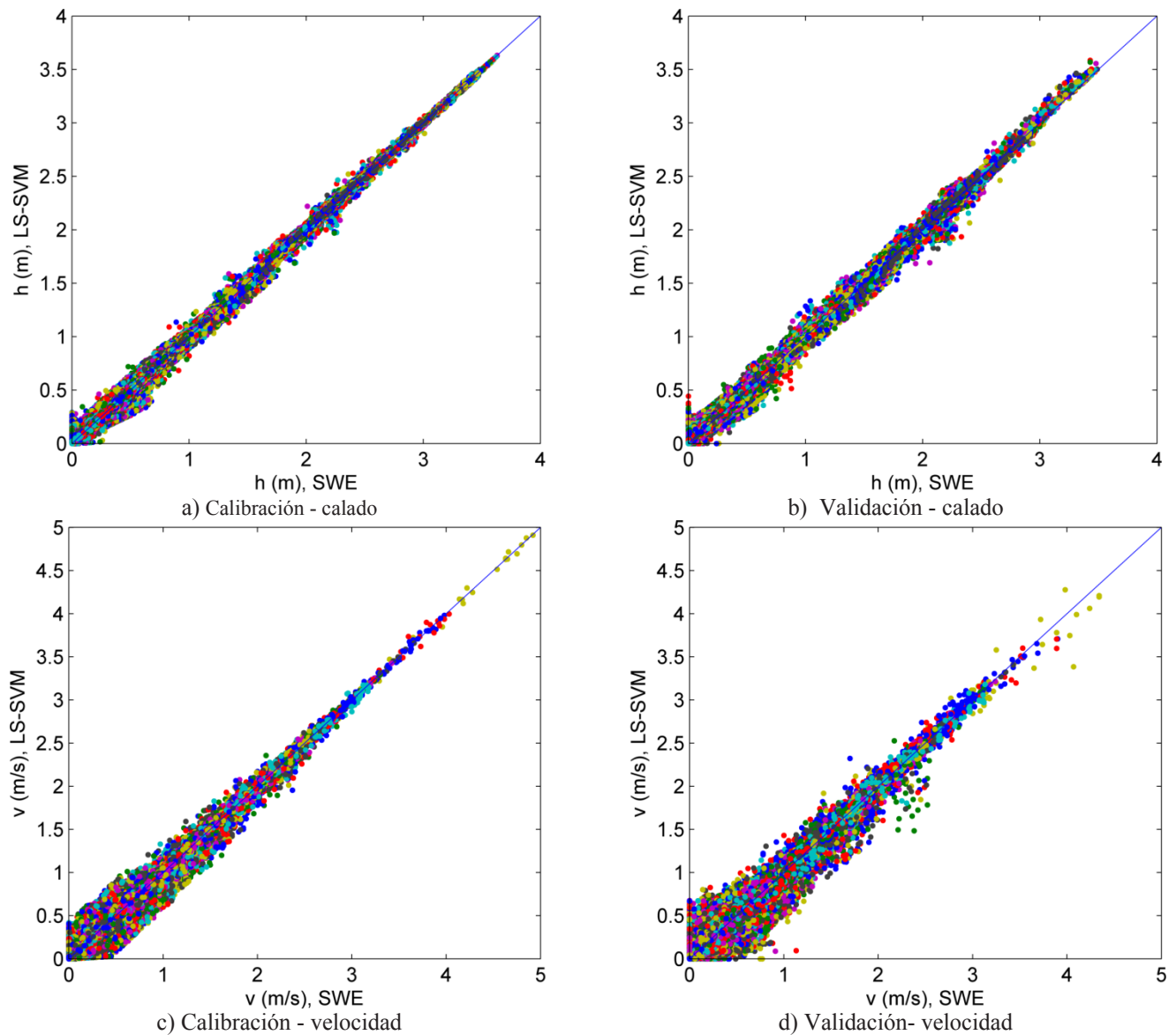


Figura 3 | Diagrama de dispersión de las predicciones LS-SVM vs. 2D-SWE en calibración (izquierda) y validación (derecha) para el calado (arriba) y la velocidad (abajo). Cada color representa un punto diferente.

La Figura 4 muestra los histogramas de error en calado y velocidad (MAE_i) para los eventos de calibración y validación. Aunque el MAE puede alcanzar hasta 0.115 m en calado y 0.235 m/s en velocidad, su valor es significativamente menor en la mayoría de puntos. El error absoluto medio global de las predicciones de calado y velocidad es de sólo 0.018 m y 0.020 m/s, respectivamente, mientras que el percentil 90 es 0.034 m y 0.05 m/s. Esto indica que la mayoría de las predicciones del modelo LS-SVM guardan un buen acuerdo con los resultados del modelo 2D-SWE.

La distribución espacial de estos errores se muestra en la Figura 5. Esta representación espacial proporciona información sobre el desempeño local del modelo y permite diagnosticar dónde se producen los peores ajustes. En la Figura 5 se puede ver cómo el modelo LS-SVM proporciona una buena predicción global. Los mayores errores no se localizan en una región en particular, sino que se distribuyen por todo el dominio del modelo. Existe una correspondencia entre el mapa de MAE de calado y el mapa de MAE de velocidad (entre las Figuras 5a y 5b), es decir, entre los puntos de alto MAE en calado y en velocidad. Sin embargo, esto no siempre es así, y en algunos puntos se logra una buena predicción de sólo una de las dos variables.

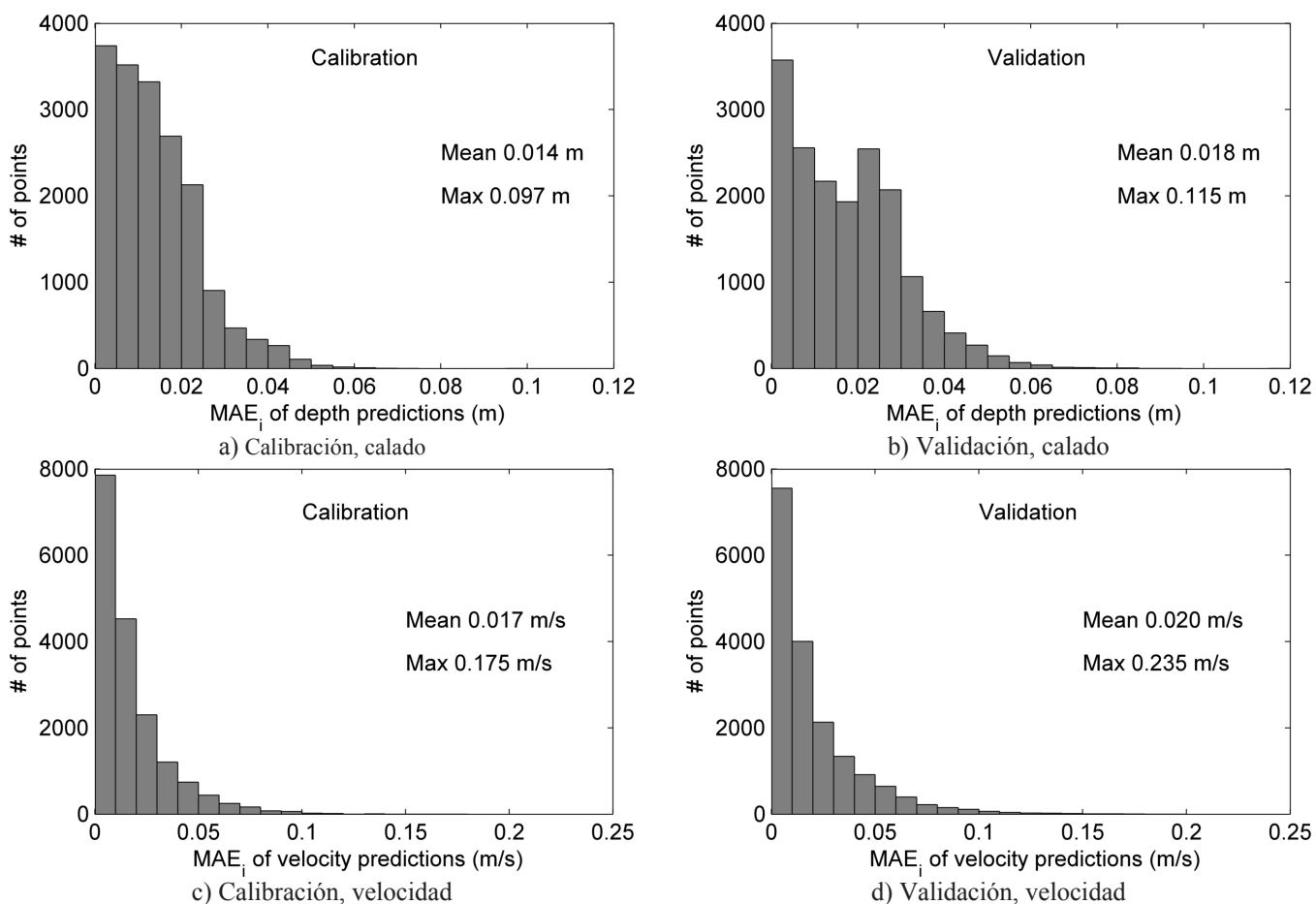


Figura 4 | Histogramas de error absoluto medio en calado y velocidad en los diferentes puntos.

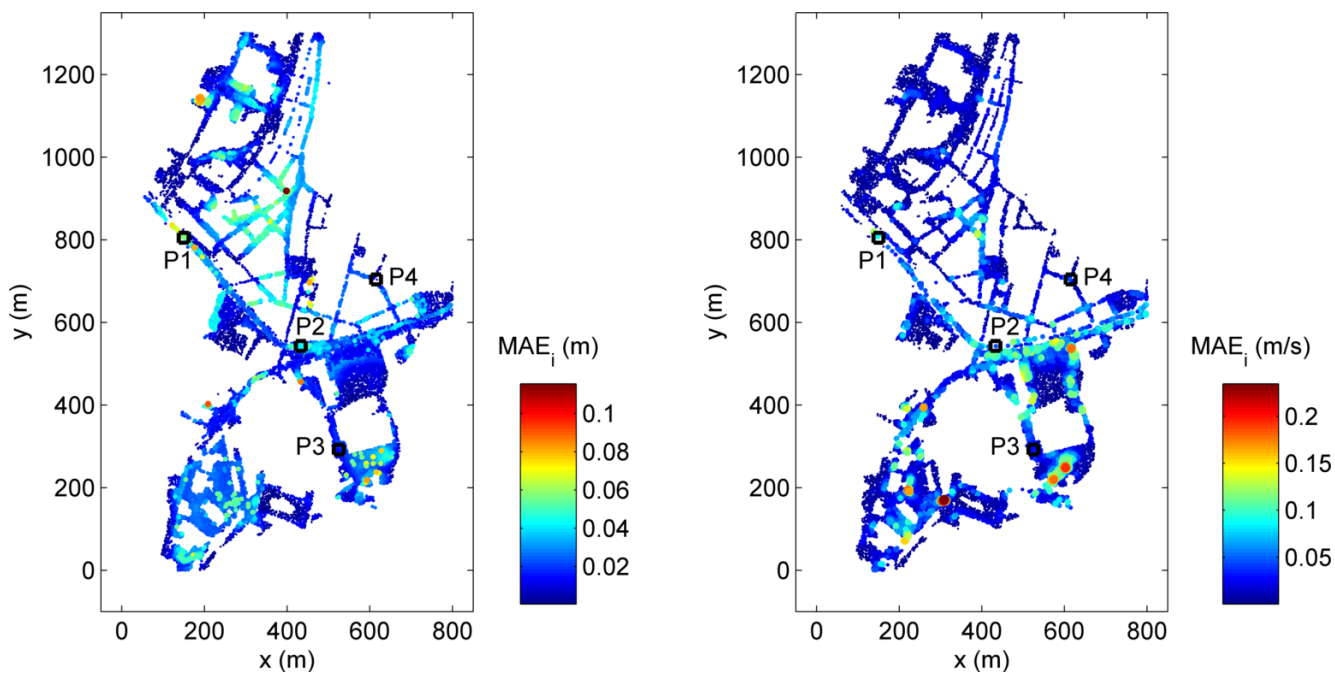


Figura 5 | Mapa de error absoluto medio en calado (izquierda) y velocidad (derecha) en validación. El tamaño del marcador es proporcional al error en cada punto. Se indica la localización de los puntos analizados en la figura 6.

El modelo LS-SVM es capaz de emular diversas y complejas relaciones entre las variables predictoras y dependientes, como se ilustra en las Figuras 6. Estas relaciones no pueden reproducirse adecuadamente usando técnicas de regresión estándar (por ejemplo, una regresión lineal múltiple), como se mostró en el análisis previo de Cea et al. (2016). En los puntos cercanos al mar se observa claramente la influencia de la marea (Figura 6a). Para el rango inferior de los caudales (Q_1), las predicciones de calado máximo (h) dependen en gran medida del nivel máximo de marea, y la relación entre h y Q_1 se bifurca. La rama inferior está asociada a la marea muerta, con un nivel máximo de marea de 1.26 m, mientras que la rama superior corresponde a las simulaciones con marea viva, con un nivel máximo de marea de 1.81 m. El modelo LS-SVM es capaz de capturar esta compleja relación, como se muestra en la Figura 6a.

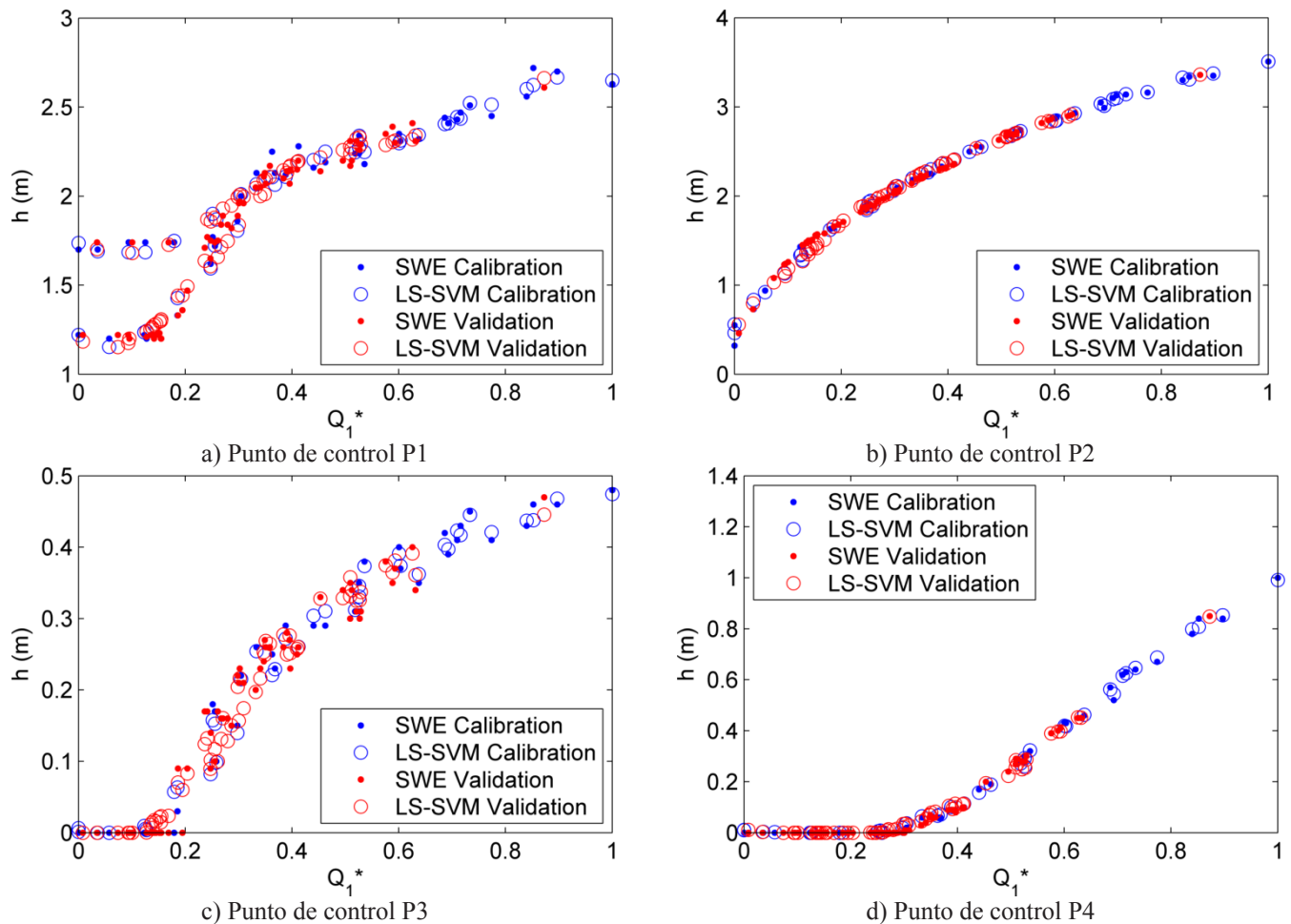


Figura 6 | Predicciones de calado (h) en 4 puntos del área de estudio (puntos de control P1-P4 en la Figura 5) con el modelo 2D-SWE y el modelo LS-SVM en función del caudal máximo normalizado de uno de los hidrogramas de entrada (Q_1^*).

La influencia de la marea en las predicciones de calado disminuye a medida que nos movemos río arriba a lo largo del cauce principal. Su influencia se percibe a lo largo de 600 m del cauce. En los puntos situados más aguas arriba, dentro del cauce principal, la relación entre en flujo Q_1 y el calado h sigue una curva única con muy poca dispersión (Figura 6b). Las predicciones de calado en esta zona dependen por tanto principalmente del caudal máximo del cauce principal. En las regiones de aguas arriba de la llanura derecha, el calado máximo h también depende principalmente de Q_1 , aunque los calados alcanzados son significativamente más bajos (Figura 6d). La influencia de los caudales en los otros dos afluentes (Q_2 y Q_3)

está limitada espacialmente a las regiones más aguas arriba de la llanura de inundación izquierda. Su efecto se refleja en una dispersión mayor en la relación Q1-h (Figura 6c).

Globalmente, el modelo LS-SVM muestra un buen desempeño en la predicción de calados y velocidades máximas en todo el área estudiada. El modelo reproduce las distintas relaciones con las variables predictoras, reflejadas en las formas de las curvas mostradas en las Figura 6. Este buen desempeño es independiente del rango de profundidades y velocidades alcanzados en los puntos estudiados.

CONCLUSIONES

En este trabajo se presenta y evalúa un modelo de regresión no paramétrico como metamodelo de un modelo de aguas someras bidimensional (modelo SWE) para el estudio de inundaciones. El modelo está basado en máquinas de vectores soporte de mínimos cuadrados (modelo LS-SVM) y utiliza los hidrogramas y el nivel de marea prescritos en los contornos del modelo SWE como variables predictoras para estimar los calados y velocidades máximas. El modelo es capaz de emular relaciones diversas y complejas entre las variables predictoras y dependientes, consiguiendo errores absolutos medios de sólo 0.018 m en calado y 0.020 m/s en velocidad. Debido a su simplicidad, tiempos de cálculo reducidos y buen desempeño, este modelo presenta un gran potencial para aplicaciones tales como la previsión de inundaciones en tiempo real o la propagación de incertidumbre en un contexto de tipo Monte Carlo.

AGRADECIMIENTOS

Este trabajo ha sido financiado por el Ministerio de Economía y Competitividad a través del proyecto “CAPRI: Cálculo probabilístico de inundaciones con modelos hidrológicos de alta resolución espacial a partir de estimaciones de precipitación de radar” (referencia CGL2013-46245-R). María Bermúdez agradece a la Xunta de Galicia la concesión de su contrato postdoctoral (referencia ED481B 2014/156-0).

REFERENCIAS

- Bermúdez, M., Neal, J.C., Bates, P.D., Coxon, G., Freer, J., Cea, L., Puertas, J. 2017. Quantifying local rainfall dynamics and uncertain boundary conditions into a nested regional-local flood modeling system. *Water Resources Research*, 53(4), 2770–2785.
- Bladé, E., Cea, L., Corestein, G., Escolano, E., Puertas, J., Vázquez-Cendón, E., Dolz, J., Coll, A. 2014. Iber: herramienta de simulación numérica del flujo en ríos. *Revista Internacional de Métodos Numéricos para Cálculo y Diseño en Ingeniería* 30(1), 1-10, doi: 10.1016/j.rimni.2012.07.004
- Brabanter, K. De, Suykens, J.A.K., Moor, B. De. 2013. Nonparametric Regression via StatLSSVM. *Journal of Statistical Software*, 55(2), 1–21.
- Cea, L., Bermúdez, M., Puertas, J., Fraga, I., Coquerez, S. 2016. Rapid flood inundation modelling in a coastal urban area using a surrogate model of the 2D shallow water equations. In *Proceedings of the 4th European Congress of the International Association of Hydroenvironment engineering and Research*, IAHR 2016. pp. 850–855.
- Cea, L., Vázquez-Cendón, M.E. 2009. Unstructured finite volume discretization of two-dimensional depth-averaged shallow water equations with porosity. *International Journal for Numerical Methods in Fluids*, 63(8), 903-930.
- Comer, J., Olbert, A.I., Nash, S., Hartnett, M. 2017. Development of high-resolution multi-scale modelling system for simulation of coastal-fluvial urban flooding. *Natural Hazards and Earth System Sciences*, 17(2), 205–224.

- Ghimire, B., Chen, A.S., Guidolin, M., Keedwell, E.C., Djordjevic, S., Savic, D.A. 2013. Formulation of a fast 2D urban pluvial flood model using a cellular automata approach. *Journal of Hydroinformatics*, 15(3), 676-686.
- Guidolin, M., Chen, A.S., Ghimire, B., Keedwell, E.C., Djordjevic, S., Savic, D.A. 2016. A weighted cellular automata 2D inundation model for rapid flood analysis. *Environmental Modelling & Software*, 84, 378-394.
- Jhong, B.-C., Wang, J.-H., Lin, G.-F. 2017. An integrated two-stage support vector machine approach to forecast inundation maps during typhoons. *Journal of Hydrology* 547, 236-252, doi: 10.1016/j.jhydrol.2017.01.057.
- Kennard, R.W., Stone, L.A. 1969. Computer Aided Design of Experiments. *Technometrics*, 11(1).
- Lin, G.-F., Lin, H.-Y., Chou, Y.-C. 2013. Development of a real-time regional-inundation forecasting model for the inundation warning system. *Journal of Hydroinformatics*, 15(4), 1391-1407.
- Liu, Y., Pender, G. 2015. A flood inundation modelling using v-support vector machine regression model. *Engineering Applications of Artificial Intelligence* 46, 223-231, doi: 10.1016/j.engappai.2015.09.014.
- Polinsky, A. et al., 1996. Librain: software for automated design of exploratory and targeted combinatorial libraries. In I. Chaiken & K. D. Janda, eds. *Molecular Diversity and Combinatorial Chemistry: Libraries and Drug Discovery*, American Chemical Society, Washington, D.C. pp. 219-232.
- Raghavendra, N, S., Deka, P.C., 2014. Support vector machine applications in the field of hydrology: A review. *Applied Soft Computing*, 19, 372-386.
- Razavi, S., Tolson, B.A. & Burn, D.H. 2012. Review of surrogate modeling in water resources. *Water Resources Research*, 48(7), W07401, doi: 10.1029/2011WR011527.
- Schubert, J.E., Sanders, B.F. 2012. Building treatments for urban flood inundation models and implications for predictive skill and modeling efficiency. *Advances in Water Resources*, 41, 49-64.
- Suykens, J.A.K. et al. 2002. *Least Squares Support Vector Machines*. World Scientific.
- Vapnik, V.N. 1998. *Statistical learning theory*. Wiley.
- Willett, P. 1996. Molecular diversity techniques for chemical databases. *Information Research*, 2(3). Available at: <http://informationr.net/ir/2-3/paper19.html>
- Yaseen, Z. M., El-shafie, A., Jaafar, O., Afan, H. A., Sayl, K. N. 2015. Artificial intelligence based models for stream-flow forecasting: 2000-2015. *Journal of Hydrology* 530, 829-844, doi: 10.1016/j.jhydrol.2015.10.038.