TESE DE DOUTORAMENTO

*Desarrollo y optimización del modelo hidráulico
Iber+: Aplicación a eventos de riesgo*

Orlando García Feal

2021

UniversidadeVigo

Universida_{de}Vigo

Escola Internacional de Doutoramento

Orlando García Feal

TESE DE DOUTORAMENTO

Desarrollo y optimización del modelo hidráulico Iber+:

Aplicación a eventos de riesgo

Dirixida polos doutores:

Dr. José Manuel Domínguez Alonso

Dr. Arno Formella

Ano: 2021

Universida_{de}Vigo

Escola Internacional de Doutoramento

José Manuel Domínguez Alonso e Arno Formella:

FAN CONSTAR que o presente traballo, titulado "Desarrollo y optimización del modelo hidráulico Iber+: Aplicación a eventos de riesgo", que presenta Orlando García Feal para a obtención do título de Doutor, foi elaborado baixo a súa dirección no programa de doutoramento "Sistemas de Software Intelixentes e Adaptables" baixo a modalidade de compendio de publicacións.

Ourense 14 de xaneiro de 2021.

Os Directores da tese de doutoramento

Dr. José Manuel Domínguez Alonso   Dr. Arno Formella

Agradecementos

> *Vexo o mundo darredor de min e adoezo por entendelo.*

— Xosé Neira Vilas
Memorias dun neno labrego

Remata con esta tese unha etapa, mais co fin dunha etapa comeza sempre unha nova. O inexorable paso do tempo fai deste mundo un mundo cambiante no que nunca nada volta a ser o que era. Mais o cambio non só suscita que se vertan bágoas na , senon que tamén da lugar a evolución, ó nacemento de vidas, culturas e un senfín de obras diversas. Nese cambio radica a fermosura deste mundo e abonda cunha ollada indagadora para decatarse. E do abraio por este mundo nace, nun afán por coñecelo mellor, a curiosidade e a ciencia. E así, dentro dese atafegante enxame de obras do mundo da investigación científica, achego aquí e de xeito humilde a miña contribución. Apenas unha pinga miúda de orballo dentro do coñecemento da nosa civilización. Pero, acaso non remata esa pinga de orballo formando parte de grandes ríos e océanos? Resulta por tanto gratificante poder contribuir aínda de xeito mínimo ó coñecemento global co obxectivo de mellorar o mundo que nos rodea e comprendelo un chisco mellor. Con todo, este traballo nunca sería posible sen toda a xente que estivo ó meu carón ó longo desta etapa. Nunca debemos esquecer que chegamos a este mundo indefensos e que sen a axuda dos nosos semellantes non estariamos aquí para ser testigos do momento. Compre, por tanto, amosar gratitude e axudar no posible á xente da nosa contorna, pois só así teremos oportunidade de afrontar as dificultades propias dun mundo cambiante. Isto é máis certo ca nunca no contexto de crise que asolaga o mundo nestes intres, mais, cun pouco de axuda, sairemos desta. Resúltame, por tanto, imposible continuar estas verbas sen antes

agradecer todo o seu apoio a Lucía, compañeira de aventuras e desventuras pero sempre loitadora infatigable; a meus pais por darme a vida e abrime a porta ó mundo; a tódolos meus compañeiros e compañeiras do grupo EPhysLab por achegarme ó mundo da investigación; a meus directores de tese por guiarme neste traballo; a tódolos meus familiares, amigos, mestres e coñecidos; a todos aqueles que tristemente xa non están e a todos aqueles que están por vir, pois a todos eles débolles quen son.

Quixera tamén agradecer á Confederación Hidrográfica do Miño-Sil pola súa valiosa colaboración e aportación de datos, os cales foron indispensables para a realización deste traballo.

Índice general

Resumen

La gestión de recursos hídricos es una actividad fundamental en la sociedad humana que reporta múltiples beneficios. Desde la irrigación, mitigación de sequías e inundaciones, generación de energía eléctrica o garantizar la calidad del agua. Son de especial interés las inundaciones al suponer uno de los desastres naturales más peligrosos y cuya frecuencia y gravedad se han ido incrementando en los últimos tiempos. Es por ello que se ha vuelto fundamental el desarrollo de herramientas adecuadas que permitan mejorar la planificación de los recursos hídricos y reducir posibles riesgos.

En este contexto, el modelado numérico es una herramienta fundamental que permite la simulación de múltiples escenarios virtuales con el consiguiente ahorro de recursos en el desarrollo de modelos reales a escala. Además, también permite la medición de multitud de magnitudes físicas que pueden resultar difíciles o imposibles de medir con precisión en el mundo real.

Sin embargo, uno de los mayores inconvenientes del modelado numérico es su elevado coste computacional, lo cual puede limitar su aplicación práctica. No obstante, la utilización de modelos numéricos se ha incrementado considerablemente en las últimas décadas gracias al incremento de la potencia de cálculo de los sistemas informáticos. Desafortunadamente, la utilización efectiva de estos recursos no siempre es trivial, sobre todo en los últimos años en los que las ganancias de rendimiento de los procesadores son debidas principalmente al aumento de la capacidad de cálculo paralelo, ofreciendo un elevado número de núcleos de procesamiento. Por ello es necesario adoptar algoritmos concurrentes y aplicar técnicas de computación de altas prestaciones para sacar partido de estos recursos computacionales.

Iber es un modelo numérico que resuelve las ecuaciones de aguas someras en dos dimensiones utilizando el método de los volúmenes finitos y mallas no estructuradas. Está orientado principalmente a la simulación de flujos de lámina libre en cuencas de ríos y estuarios no estratificados. Además de resolver la hidrodinámica, Iber también incorpora una serie de módulos que amplían su funcionalidad como el módulo de calidad de aguas, el módulo de hidrología o el módulo de transporte de sedimentos. Este modelo se ha empleado con éxito en multitud de aplicaciones y estudios. Sin embargo, una de sus limitaciones principales es su elevado tiempo de ejecución. Al ser diseñado para utilizar un único procesador, Iber no puede usar eficientemente la capacidad de cálculo de las arquitecturas de hardware actuales que presentan un alto nivel de paralelismo. Esta desventaja limita su aplicación efectiva en diversas aplicaciones como pueden ser los sistemas de alerta temprana, calibración y análisis de incertidumbre o la simulación de grandes dominios espacio-temporales.

El objetivo principal de esta tesis es el desarrollo de herramientas de simulación numérica adecuadas para abordar problemas relacionados con la gestión de recursos hídricos y más concretamente el análisis y gestión de eventos de riesgo. Este objetivo se descompone en los siguientes objetivos específicos: aumentar la eficiencia computacional del módulo hidrodinámico de Iber junto con el módulo de calidad de aguas y su aplicación en el desarrollo de un sistema de alerta temprana de inundaciones.

Para alcanzar dichos objetivos se ha desarrollado una nueva implementación del modelo Iber que mediante técnicas de computación de altas prestaciones es capaz de utilizar de forma efectiva la máxima potencia de cálculo tanto de las CPU multi-núcleo como de las GPU o tarjetas gráficas. Esta nueva implementación denominada Iber+ ha demostrado ser capaz de acelerar la velocidad de ejecución de las simulaciones en dos órdenes de magnitud manteniendo la precisión de los resultados, lo que permite abordar simulaciones más grandes en menos tiempo. Además del módulo principal que implementa la hidrodiná-

mica, también se ha realizado una paralelización completa para los módulos de turbulencia, hidrología y calidad de aguas.

El desarrollo de Iber+ ha sido fundamental en la creación de MIDAS, un sistema de alerta temprana de inundaciones para el río Miño. Este sistema es capaz predecir de forma precisa eventos de inundación en zonas urbanas a partir de datos de precipitación (modelos de predicción y/o pluviómetros) y caudal antecedente de estaciones de control. El sistema es capaz de proporcionar a las correspondientes autoridades mapas de riesgo detallados en un margen de tiempo reducido. Además, este sistema puede ser adaptado de forma sencilla a otras regiones. Todos los componentes del sistema de alerta temprana desarrollado en esta tesis están disponibles de forma gratuita y puede ser adaptado fácilmente a los datos disponibles. Por último, la novedosa metodología presentada para la modelización hidrológica, permite simular de forma efectiva eventos de duración arbitraria con un número reducido de parámetros, lo cual es especialmente conveniente en zonas poco estudiadas.

Por último, estos desarrollos se han mostrado eficaces a la hora de estudiar fenómenos extremos de inundación, siendo aplicados con éxito en la reproducción de eventos de inundación históricos. En estos estudios son necesarias un elevado número de simulaciones para la calibración de parámetros debido a la limitada disponibilidad de datos en muchos de estos casos. La utilización de un modelo hidráulico eficiente permite realizar estas simulaciones en un tiempo razonable.

Índice de figuras

Lista de acrónimos

1D	One-dimensional.
2D	Two-dimensional.
3D	Three-dimensional.
API	Application Programming Interface.
ASIC	Application Specific Integrated Circuit.
AVX	Advanced Vector eXtensions.
CEDEX	Centro de Estudios y Experimentación de Obras Públicas.
CIMNE	Centro Internacional de Métodos Numéricos en Ingeniería.
CFD	Computational Fluid Dynamics.
CPU	Central Processing Unit.
CUDA	Compute Unified Device Architecture.
DP	Double Precision.
FPGA	Field-Programmable Gate Array
GEAMA	Grupo de Enxeñaría da Auga e do Medio Ambiente.
GFLOPS	Giga FLoating point OPerations per Second.
GPGPU	General-Purpose computing on Graphics Processing Units.
GPU	Graphics Processing Unit.
HEC-HMS	Hydrologic Engineering Center - Hydrologic Modeling System.
HIP	Heterogeneous-Compute Interface for Portability.
HPC	High Performance Computing.
MPI	Message Passing Interface.
MIDAS	Miño rIver flooD Alert System.
NRMSD	Normalised Root Mean Square Deviation.
NUMA	Non Uniform Memory Access.
OpenACC	Open ACCelerators.

OpenCL   Open Computing Language.

OpenMP   Open MultiProcessing.

PCI   Peripheral Component Interconnect.

ROCm   Radeon Open Compute.

SAICA   Sistema Automático de Información de Calidad de las Aguas.

SAIH   Sistema Automático de Información Hidrológica.

SCS   Soil Conservation Service.

SIMT   Single Instruction Multiple Thread.

SM   Stream Multiprocessor.

SMA   Soil Moisture Accounting.

SMP   Symmetric MultiProcessing.

SP   Simple Precision.

WRF   Weather Research and Forecasting.

Introducción

<div align="right">

1

</div>

El primer capítulo presenta una visión general sobre esta tesis doctoral. En primer lugar, se aborda la motivación que dio lugar a este trabajo siguiendo una breve introducción al modelado numérico en general y a los modelos empleados como base de este trabajo de investigación. Por último, se introduce la estructura general de esta tesis.

1.1. Motivación

La gestión de recursos hídricos se ha convertido en una tarea que reporta múltiples beneficios a la sociedad como pueden ser la mitigación de sequías e inundaciones, irrigación para la agricultura, protección ambiental, agua potable o la generación de energía hidroeléctrica (Bladé i Castellet et al., 2018).

Merecen especial atención las inundaciones, un desastre natural que ha afectado a la humanidad a lo largo de su historia. En perspectiva, este tipo de fenómenos suponen uno de los desastres naturales más peligrosos, causando más de 100.000 muertes en el último siglo (Jonkman, 2005) y miles de millones de euros en pérdidas económicas (Hallegatte, 2012; Wallemacq et al., 2018). Diferentes estudios constatan que su frecuencia y gravedad se han incrementado en los últimos años (Berghuijs et al., 2017; Passerotti et al., 2020). Existen

diversas razones que explican este incremento, aunque las más destacables son: variaciones en los patrones de evaporación y precipitación debidos al cambio climático (Arnell & Gosling, 2016; C. Liu et al., 2018); y por otro lado, cambios significativos en los usos de suelo (debidos a diversos motivos como la deforestación o a la urbanización del terreno) que pueden provocar un aumento del volumen de escorrentía durante eventos de precipitación (Booth & Bledsoe, 2009; Bradshaw et al., 2007; Paix et al., 2013; Rosburg et al., 2017).

Todo ello evidencia la necesidad de planificar la gestión de los recursos hídricos para una mejora de la economía, la conservación ambiental y la protección de la sociedad. Para ello es precisa la obtención de los datos necesarios para caracterizar el área de estudio y la disposición de las herramientas adecuadas que permitan evaluar las posibles estrategias de actuación.

El modelado numérico es una herramienta eficaz en la planificación de recursos hídricos. Este tipo de tecnología permite la simulación de múltiples escenarios virtuales, lo que puede suponer un considerable ahorro de recursos en el desarrollo de modelos reales a escala, especialmente cuando éstos son muy difíciles de construir. Además, también posibilita la medición de multitud de magnitudes físicas que pueden resultar difíciles o imposibles de medir en el mundo real.

Los modelos numéricos resuelven una serie de ecuaciones definidas por un modelo matemático que caracteriza un sistema. Puesto que muchos de los procesos naturales se pueden describir por medio de ecuaciones matemáticas bien conocidas, es posible construir modelos matemáticos que reproduzcan el comportamiento de dichos procesos. Estos modelos matemáticos permiten conocer el estado futuro de un sistema a partir de las condiciones iniciales del mismo y los parámetros pertinentes. Muchas de las ecuaciones de dicho modelo matemático, requieren ser resueltas de forma aproximada por medio de diferentes métodos como por ejemplo el método de volúmenes finitos (Eymard et al., 2000) o el método de diferencias finitas (Smith, 1985) los cuales pueden

ser computacionalmente costosos. Sin embargo, gracias al aumento progresivo de la capacidad de cálculo de los sistemas informáticos, la importancia de los modelos numéricos ha ido aumentando en las últimas décadas, siendo capaces de simular cada vez procesos más complejos y con mayor resolución.

De cara a la gestión de recursos hídricos, es preciso considerar los diferentes procesos típicamente involucrados en el ciclo del agua (ver Fig. 1.1). Estos procesos son numerosos y complejos, entre los que se pueden encontrar fenómenos de precipitación, transporte (de sólidos, sedimentos, sustancias químicas, etc.), escorrentía, infiltración o flujo subterráneo entre otros. Debido a esto, y dependiendo del objeto de estudio es frecuente recurrir a diferentes modelos (p.ej. modelos meteorológicos para calcular la precipitación, modelos hidrológicos para calcular la escorrentía, modelos hidráulicos para calcular calados y velocidades, etc.).
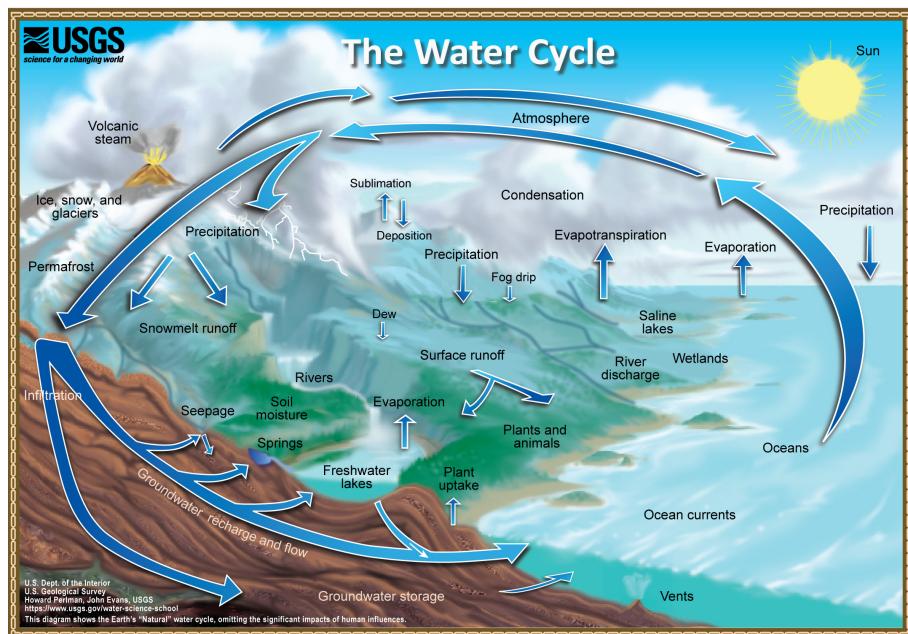
**Figura 1.1:** Ilustración del ciclo hidrológico. Fuente: Perlman, H. y Evans, J. (2019). The water cycle. Obtenido de: https://www.usgs.gov/media/images/water-cycle-natural-water-cycle bajo dominio público.

La dinámica computacional de fluidos, más conocida por sus siglas en inglés CFD (computational fluid dynamics), es una rama de la mecánica de fluidos que

emplea métodos numéricos para el estudio de problemas que estén relacionados con el comportamiento de fluidos. Dentro de los modelos numéricos empleados para el estudio de inundaciones y fenómenos de transporte son de especial interés los modelos CFD distribuidos de dos dimensiones (2D) (Hunter et al., 2008). Estos modelos hidráulicos permiten el estudio del comportamiento del flujo de agua en una topografía concreta teniendo en cuenta de forma explícita las variaciones espaciales que la caracterizan. Estos modelos 2D permiten simular dominios de tamaño medio o grande que no pueden ser abordados por modelos 3D y ofrecen unos resultados con una notable precisión en casos donde las componentes horizontales del flujo predominan sobre las verticales. Con el objetivo de reproducir procesos cada vez más complejos, es frecuente el acoplamiento de este tipo de modelos hidráulicos con otros modelos (p.ej. modelos hidrológicos, de sedimentos, de calidad de aguas, etc.). Las estrategias para realizar este acoplamiento son diversas y dependen de los objetivos del mismo, pudiendo ser un acoplamiento total o parcial y este último a su vez puede ser unidireccional o bidireccional (Amirlatifi, 2013).

Dentro del contexto de la necesidad de disponer de herramientas numéricas para la gestión de los recursos hídricos en el marco de la legislación española, nace el modelo Iber (Bladé et al., 2014). Iber es un trabajo conjunto entre diferentes centros como el CEDEX (Centro de Estudios y Experimentación de Obras Públicas), el grupo GEAMA (Grupo de Enxeñaría da Auga e do Medio Ambiente de la Universidade da Coruña), el grupo Flumen (de la Universitat Politècnica de Catalunya y de la Universitat de Barcelona) y el CIMNE (Centro Internacional de Métodos Numéricos en Ingeniería). El modelo fue desarrolado a partir de los modelos hidráulicos Turbillón (Cea, 2005) y CARPA (Bladé Castellet, 2005) junto con la interfaz para pre y post-procesado GiD (Ribó et al., 1998) (ver Fig. 1.2). Se trata pues de un modelo nacido en el ámbito académico, de código cerrado pero disponible de forma gratuita a través del sitio web oficial (https://iberaula.es).
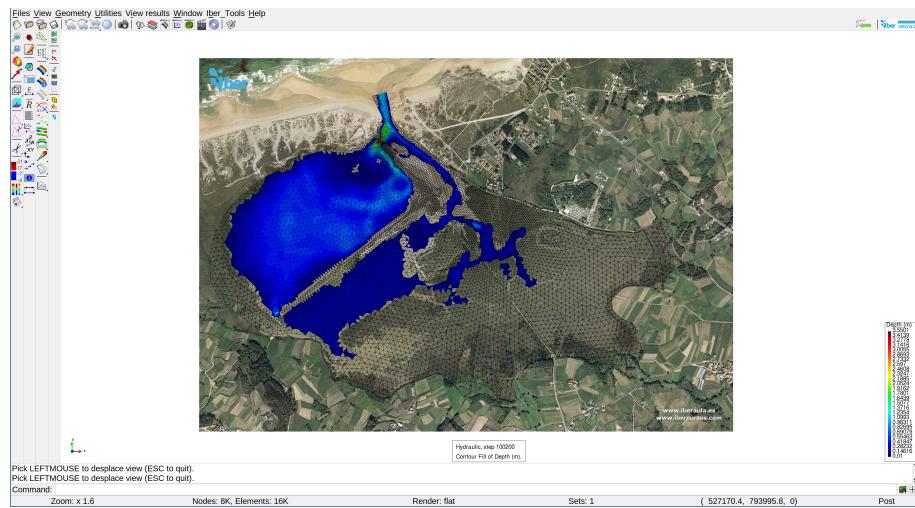
Figura 1.2: Captura de pantalla de la interfaz de Iber.

Iber es fundamentalmente un modelo un modelo hidrodinámico de aguas someras en dos dimensiones. Está orientado principalmente a la simulación de flujos de lámina libre en cuencas de ríos y estuarios no estratificados. Este modelo resuelve las ecuaciones de St. Venant bidimensionales utilizando el método de volúmenes finitos para mallas no estructuradas. Aparte de la simulación hidrodinámica cabe destacar cinco módulos adicionales, totalmente acoplados, que amplían su funcionalidad: turbulencia, hidrología, transporte de sedimentos, calidad de aguas y hábitat fluvial. Para que este tipo de herramientas se puedan emplear en el ámbito científico o en tareas de planificación, es imprescindible comprobar que los modelos sean científicamente sólidos y robustos (N. Moriasi et al., 2007; US EPA, 2002). En este sentido, Iber se ha validado en diversos escenarios y aplicaciones (Bladé Castellet et al., 2014; Bodoque et al., 2016; Cea & Bladé, 2015; Cea & French, 2012; Cea, Legout et al., 2016; Fraga et al., 2016; Sopelana et al., 2018). También se ha comparado en términos de precisión con modelos comerciales utilizando para ello un banco de pruebas diseñado para tal fin (Neelz & Pender, 2010), concluyendo que Iber ofrece resultados comparables al de sus alternativas (Cea et al., 2020). No obstante, aunque similar en términos de precisión, Iber se situaba por detrás en términos de eficiencia respecto a sus alternativas comerciales (García-Feal et al., 2017) como pueden ser TUFLOW

("TUFLOW FV Science Manual", 2013), InfoWorks ("InfoWorks ICM Product Information - Overview", 2018) o JFLOW+ (Crossley et al., 2010). La diferencia de rendimiento se debe a que estos modelos comerciales explotan el paralelismo de las arquitecturas hardware actuales para acelerar los cálculos, mientras que Iber no está preparado para ello.

El modelo Iber fue diseñado originalmente para ser utilizado en un único procesador. Tradicionalmente este tipo de programas se beneficiaban automáticamente las de importantes mejoras en rendimiento conforme mejoraban las arquitecturas de hardware. En 1965, Gordon E. Moore, cofundador de Intel, basándose en observaciones de la propia industria, afirmó que la capacidad de integración de transistores en un chip se duplicaría cada año (Moore, 1965). Más tarde, en 1975 corregiría su propio enunciado incrementando el periodo de duplicación a dos años (Moore, 1975). Esta afirmación se conocería como la ley de Moore e implica que las capacidades de procesamiento de los chips informáticos se incrementan de forma periódica (ver Fig. 1.3). Las mejoras tecnológicas a nivel de microarquitectura y velocidad de reloj de los procesadores constituyeron durante muchos años una mejora inmediata en el rendimiento del software sin precisar cambios en el diseño del mismo. Sin embargo, los incrementos en la velocidad de reloj se redujeron drásticamente en la década de los años 2000, debido principalmente a problemas de temperatura y excesivo consumo de energía (Ross, 2008). Esto impulsó un cambio de paradigma tanto en el diseño del hardware como del software (Sutter, 2005). La capacidad de procesamiento siguió aumentando principalmente debido al diseño de procesadores multinúcleo y en menor medida debido a mejoras en las microarquitecturas. Los procesadores multinúcleo incluyen en un mismo circuito integrado dos o más unidades de procesamiento, denominadas núcleos, gracias a lo cual son capaces de ejecutar múltiples operaciones en paralelo. Sin embargo, para sacar provecho la capacidad de procesamiento que ofrece el paralelismo es necesario realizar cambios sustanciales en el software. Desafortunadamente, la programación concurrente es más difícil que la programación secuencial (Sutter & Larus,

2005) e implica un incremento importante de complejidad. Presenta además grandes desafíos (Belikov et al., 2013; Czarnul et al., 2020) tanto a la hora realizar un programa que funcione correctamente como para que ese programa sea eficiente.


48 Years of Microprocessor Trend Data

Transistors (thousands)

Single-Thread Performance (SpecINT x 10^3)

Frequency (MHz)

Typical Power (Watts)

Number of Logical Cores

Original data up to the year 2010 collected and plotted by M. Horowitz, F. Labonte, O. Shacham, K. Olukotun, L. Hammond, and C. Batten
New plot and data collected for 2010-2019 by K. Rupp
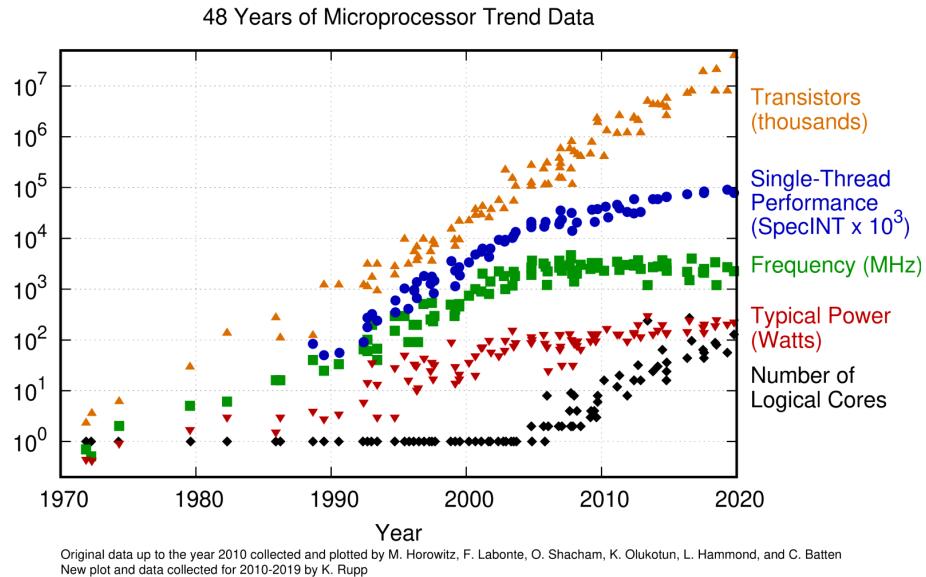
Figura 1.3: Evolución de los microprocesadores en los últimos años. Fuente: Rupp, K. (2020). 48 Years of Microprocessor Trend Data. Obtenido de: https://-github.com/karlrupp/microprocessor-trend-data bajo licencia Creative Commons Attribution 4.0

En este contexto surgen diferentes APIs (Application Programming Interface) para facilitar el aprovechamiento de dicho paralelismo. Algunas de las más comunes son: OpenMP (Open MultiProcessing) (Dagum & Menon, 1998) principalmente para sistemas CPU (Central Processing Unit) de memoria compartida, MPI (Message Passing Interface) (Gropp et al., 1999) para sistemas CPU de memoria distribuida, CUDA (Compute Unified Device Architecture) (NVIDIA, 2020) para tarjetas gráficas o GPUs (Graphics Processing Units) del fabricante Nvidia u OpenCL (Open Computing Language) (Munshi, 2009) standard del grupo Khronos para sistemas heterogéneos entre los que se incluyen CPUs, GPUs y FPGAs (Field-programmable gate array).

El modelo Iber fue parcialmente paralelizado para CPUs multinúcleo, usando para ello el API OpenMP. Sin embargo, al ser concebida como una aplicación

para funcionar en serie, se requerían profundos cambios en el software para lograr una paralelización completa. La paralelización parcial implementada en Iber presenta limitaciones en cuánto a la aceleración máxima que se puede alcanzar. Este límite es bien conocido y definido por la ley de Amdahl (Amdahl, 1967):

$$Aceleración_{paralelo}(f, n) = \frac{1}{(1 - f) + \frac{f}{n}} \tag{1.1}$$

Se calcula la aceleración que se consigue en un código paralelizado donde f es la fracción del programa (relativa al tiempo de ejecución) que está paralelizada y n el número de núcleos de procesamiento. De la ecuación se deduce que cuando la fracción paralelizada f es pequeña, también lo es la aceleración conseguida. Aún con un 75% del programa paralelizado de forma ideal y un número infinito de procesadores sólo sería cuatro veces más rápido (ver Fig. 1.4). Esto evidencia la necesidad de realizar una paralelización del modelo Iber lo más completa posible si lo que se busca es sacar partido a los procesadores multinúcleo actuales, donde las ganancias de rendimiento en actuales y sucesivas generaciones vienen dadas, en gran parte, por el aumento de núcleos de procesamiento.
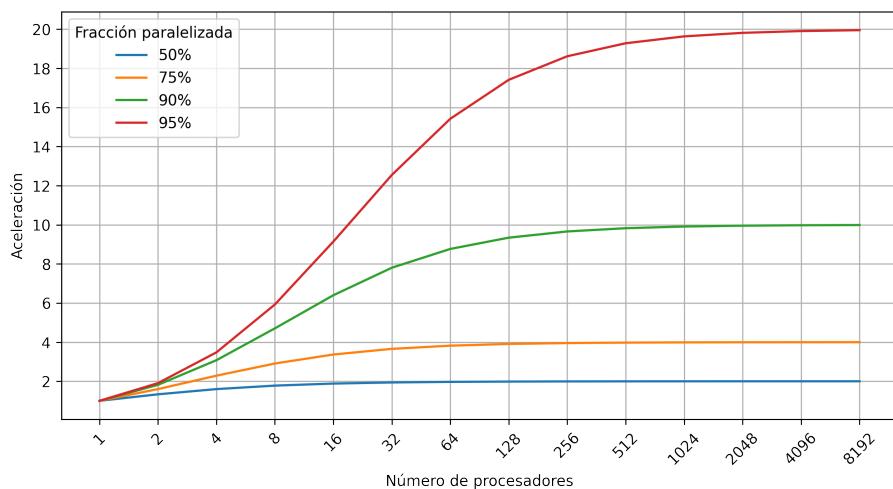

Figura 1.4: Según la ley de Amdahl, la aceleración conseguida con la paralelización está limitada por fracción del programa que se ejecuta de forma secuencial.

De cara al futuro, las actuales tecnologías tienen límites físicos y existen indicios de ralentización en el ritmo de mejora predicho por la ley de Moore (Eeckhout, 2017). Sin embargo, hasta el momento la industria sigue anunciando mejoras en los procesos de integración ("IRDS", 2020; Lapedus, 2019) lo que parece indicar que este progreso todavía continuará en el futuro cercano. Por otro lado, la industria está cada vez más centrada no sólo en la evolución de la ley de Moore sino en el diseño de chips dirigido por la aplicación a la que serán destinados. Por este motivo, en un contexto en el que la cantidad de datos que deben ser procesados es cada vez mayor, es común encontrar unidades de procesamiento especializadas (o aceleradores) para realizar ciertos tipos de cargas de trabajo como pueden ser las GPUs, FPGAs o ASICs (Application Specific Integrated Circuit), con el objetivo de complementar a los procesadores de propósito general (Najafi et al., 2017).

Además de las CPUs multinúcleo de propósito general, otra arquitectura paralela ampliamente extendida son las GPUs (Blake et al., 2009). Su arquitectura está diseñada para ejecutar un gran número de instrucciones en paralelo valiéndose para ello de un gran número de núcleos de procesamiento y un gran ancho de banda de memoria. A diferencia de las CPUs que están diseñadas para ejecutar un *thread* (o hilo de ejecución) con la mayor velocidad posible, las GPUs están diseñadas ejecutar miles de *threads* en paralelo compensando así el menor rendimiento por *thread* (NVIDIA, 2020). Es por ello que este tipo de arquitectura resulta especialmente interesante para aplicaciones de procesamiento de datos. Básicamente, dada una misma tecnología de integración, en una GPU se destina una mayor parte del circuito integrado a la lógica dedica a procesamiento de datos (ver Fig. 1.5) como puede ser el cálculo en punto flotante.

Originalmente las GPUs estaban orientadas exclusivamente al procesamiento de gráficos. No obstante, debido a su naturaleza programable actualmente soportan una gran variedad de aplicaciones, desde el cálculo de aplicaciones científicas a *machine learning*, en cuyo caso suele hablarse de GPGPU (General-Purpose
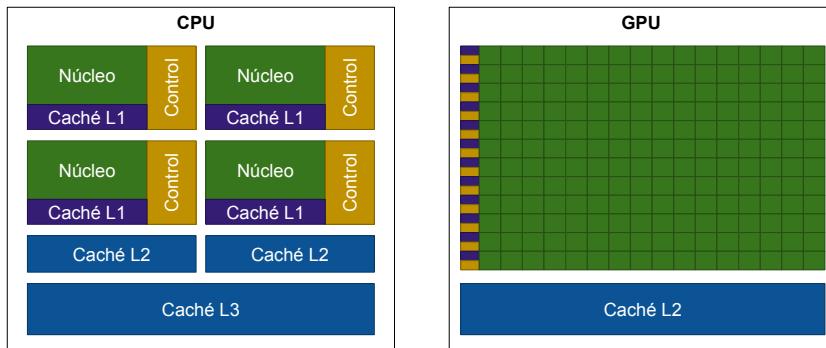
Figura 1.5: Comparativamente, una GPU dedica un mayor porcentaje de transistores a lógica de procesamiento de datos que una CPU.

computing on Graphics Processing Units). En dichas aplicaciones, este tipo de arquitectura goza de popularidad debido a ofrecer una gran capacidad de cálculo en relación con su precio de mercado, especialmente en comparación con las CPUs (Nickolls & Dally, 2010; Sun et al., 2019). Puesto que los chips actuales están limitados por temperatura y consumo eléctrico (Rotem et al., 2013), la eficiencia de una arquitectura en cuanto a operaciones por vatio consumido es clave para estimar cual es capaz de ofrecer una mayor capacidad de cálculo. En este sentido, como muestra la Figura 1.6, y aunque las CPUs han conseguido cerrar ligeramente la brecha en los últimos años gracias a integrar cada vez un mayor número de núcleos de procesamiento y la integración de instrucciones vectoriales de coma flotante específicas como las AVX (Advanced Vector eXtensions), las GPUs son capaces de ofrecer un mayor rendimiento por vatio. Esto las hace especialmente atractivas para ordenadores de escritorio, donde los usuarios podrán disponer de una considerable capacidad de cómputo sin tener que recurrir a centros cálculo. Pero por otro lado, debido a esta mayor eficiencia también son una opción muy utilizada en centros de cálculo. Según el listado TOP500 ("November 2020 | TOP500", 2020) de los supercomputadores más potentes del mundo, a fecha de noviembre del año 2020, seis de los diez más potentes del mundo estaban equipados con GPUs del fabricante NVIDIA. Por otro lado, la clasificación de supercomputadores más eficientes ("November 2020 | TOP500", 2020) está encabezada por un sistema que equipa

GPUs NVIDIA. También en entornos de computación en la nube se ha visto impulsada la incorporación de GPUs y otros aceleradores de cálculo como las FPGAs (Varghese et al., 2015; Varghese & Buyya, 2018).
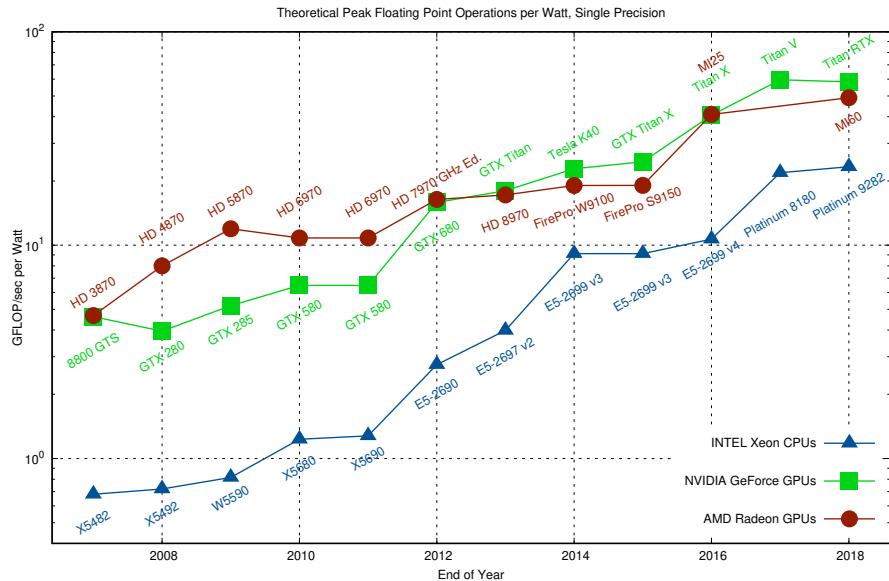

Figura 1.6: Evolución, tanto para CPUs como GPUs, de la capacidad de cálculo en GFLOPS (Giga FLoating point OPerations per Second) con precisión simple (SP) por vatio de energía utilizada. Fuente: Rupp, K. (2019). CPU-GPU-MIC Comparison Charts. Obtenido de: https://github.com/karlrupp/cpu-gpu-mic-comparison bajo licencia Creative Commons Attribution 4.0

Los modelos CFD se benefician de las arquitecturas paralelas y especialmente de las GPUs al contar no sólo con mayor capacidad de cálculo sino que, generalmente también disponen de un mayor ancho de banda de memoria (Lee et al., 2010). Esto ha motivado que se hayan aplicado técnicas de paralelización sobre GPUs en multitud de modelos (Crespo et al., 2015; Crossley et al., 2010; Lacasta et al., 2014; Q. Liu et al., 2018; Vacondio et al., 2014), alcanzando mejoras en la velocidad de ejecución de hasta dos órdenes de magnitud. Dentro del ámbito científico, la solución más comúnmente usada para cálculo con GPUs es la proporcionada por el fabricante NVIDIA a través de su API específica CUDA. Las razones para esta popularidad son: su posición dominante en el mercado de GPUs dedicadas, especialmente en el ámbito de la computación de altas prestaciones o HPC (High Performance Computing) (Harris, 2020); su apuesta

temprana por esta tecnología (la versión inicial data del año 2007); la estabilidad de dicha API y la abundante disponibilidad de herramientas y documentación. Aunque no se trata de un standard abierto como su alternativa OpenCL (Munshi, 2009), a día de hoy incluso AMD (el segundo mayor fabricante de GPUs) propociona soporte para portar código CUDA a su plataforma de cálculo ROCm (Radeon Open Compute) a través de HIP (Heterogeneous-Compute Interface for Portability) (Kuznetsov & Stegailov, 2019).

En cuanto a la programación paralela de arquitecturas basadas en CPUs de memoria compartida, el API OpenMP es el standard de facto en entornos de computación científica. Este API soporta diversos lenguajes de programación y permite paralelizar un código diseñado para ejecución monohilo de forma sencilla mediante la utilización de *pragmas* o directivas. Sin embargo, es frecuente que para lograr explotar al máximo la potencia de cálculo de las CPUs multinúcleo sea necesario estudiar la idoneidad de cada estructura de datos y algoritmos empleados.

En el caso de la programación de GPUs, el proceso de adaptación del código existente es mucho más complejo que en el caso de las CPUs. Aunque existen soluciones como OpenACC (Open ACCelerators) que siguen un concepto similar al utilizado por OpenMP, el rendimiento obtenido es a menudo menor que utilizando el API nativa proporcionada por el fabricante (Hoshino et al., 2013). En el caso de CUDA, los algoritmos deben ser adaptados a esta API de programación teniendo en cuenta las particularidades de las GPUs. Entre estas particularidades destacan: la transferencia de datos entre memoria de CPU y GPU la cual tiene un ancho de banda limitado y puede tener un impacto negativo en la eficiencia del programa; la sincronización de los hilos de ejecución en las GPUs es un proceso computacionalmente mucho más costoso que en las CPUs, por lo que deben considerarse variantes de los algoritmos que necesiten un menor número de estas sincronizaciones, aún a coste de realizar cálculos redundantes; y por último debe estudiarse la organización de los hilos de ejecución a utilizar

y cómo se repartirá el conjunto de datos a procesar, lo cual tiene una gran influencia en el aprovechamiento de los recursos de la GPU.

Este proceso, a pesar de ser laborioso, es un proceso necesario en el actual panorama de la computación. Aquellos programas que no sean capaces de mapear el paralelismo intríseco de un problema al *hardware* actual, no serán capaces de aprovechar la capacidad de cálculo disponible y verán limitado su campo de aplicación.

La utilización de modelos hidráulicos como Iber en determinadas aplicaciones requiere grandes cantidades de recursos computacionales. Para asegurar que el modelo reproduce de forma precisa un escenario dado es necesario realizar una serie de procesos de análisis de sensibilidad, calibración y validación (N. Moriasi et al., 2007). Los análisis de sensibilidad permiten identificar los principales parámetros que afectan a los resultados de la simulación (Saltelli et al., 2008). La calibración del modelo permite obtener una estimación de los parámetros de forma que los resultados obtenidos se ajusten a las mediciones realizadas en el mundo real (Madsen, 2000). La validación consiste en comprobar que la aplicación del modelo específico para un determinado dominio, y utilizando los parámetros obtenidos durante la calibración, es capaz de realizar predicciones suficientemente precisas para períodos de tiempo que no sean los empleados en la calibración (Refsgaard, 1997). Todo esto implica la ejecución de un gran número de simulaciones (generalmente en el orden de decenas) y pone de manifiesto la necesidad de modelos eficientes para reducir los tiempos de este proceso que, de otro modo y dependiendo del detalle necesario, pueden no ser viables.

Requiere especial atención las simulaciones de calidad de aguas, las cuales necesitan a menudo la simulación de diferentes tipos de contaminantes en grandes dominios y durante largos períodos de tiempo. Esto requiere una gran capacidad de cálculo lo que tradicionalmente ha llevado a modelos orientados a este campo a asumir simplificaciones hidrodinámicas, por ejemplo, QUAL2K

(Chapra et al., 2008) emplea un modelo 1D promediando toda la sección del río mientras CE-QUAL-W2 (Cole & Buchak, 1995; Cole & Wells, 2006) emplea un modelo 2D lateralmente promediado. Otros casos de la literatura se centran en modelar un contaminante en concreto o un número reducido de ellos (Abu-Bakar et al., 2017; Gao et al., 2011; Kashefipour et al., 2002; Kay et al., 2005; Wu et al., 2005). Por tanto, para poder simular de forma precisa la hidrodinámica del cauce del río y los parámetros de calidad de las aguas sin que los tiempos de cálculo supongan un factor limitante, es necesario un modelo que implemente técnicas de computación de altas prestaciones.

Modelos distribuidos como Iber también resultan de especial interés para su aplicación en sistemas de alerta temprana de inundaciones o incluso de calidad de aguas ante vertidos detectados por estaciones de control. Estos modelos pueden ofrecer mapas de distribución de calados, velocidad u otras variables de interés. En el caso de los sistemas de alerta temprana, se requiere simular una o múltiples áreas de interés con una resolución adecuada y en un tiempo de cálculo máximo determinado. Sin embargo, el alto coste computacional de los modelos distribuidos, ha motivado que tradicionalmente estos sistemas emplearan modelos hidrológicos agregados. Los modelos agregados no consideran de forma explícita las variaciones espaciales de la topografía a simular y utilizan ciertas características promediadas en su lugar (Brirhet & Benaabidate, 2016), por lo que son computacionalmente muy eficientes. Por contra, estos modelos sólo ofrecen los valores de caudal en puntos determinados del cauce de un río a partir de datos de precipitación (los cuales pueden ser obtenidos de modelos meteorológicos o pluviómetros). Para ofrecer datos de peligrosidad en zonas sensibles (p.ej. zonas urbanas o de recreo) con este tipo de sistemas de alerta, suelen realizarse análisis previos para diversos caudales correspondientes a diferentes períodos de retorno. De este modo es posible ofrecer una indicación aproximada de la peligrosidad en base a los caudales obtenidos del modelo agregado. En literatura pueden encontrarse diversos ejemplos de este tipo de

sistemas aplicados en diferentes escalas (Alfieri et al., 2013; Alfieri et al., 2012; Corral et al., 2019; Thielen et al., 2009).

La aplicación de técnicas de computación de altas prestaciones a modelos hidráulicos distribuidos junto con el incremento de la potencia de cálculo del *hardware* que ha tenido lugar en los últimos años, permite el desarrollo de sistemas de alerta temprana que incorporen este tipo de modelos. De este modo es posible ofrecer detallados mapas de inundación en base a las series de caudal proporcionadas por un modelo hidrológico agregado. Esto permitiría la obtención de datos precisos para informar a las autoridades correspondientes y tomar las medidas necesarias en los lugares oportunos para limitar los daños en caso de inundación.

1.2. Estructura del trabajo

Debido a que la mayor parte del trabajo desarrollado para esta tesis doctoral ya ha sido publicado (o ha sido enviado para su publicación) en revistas de ámbito internacional revisadas por pares, esta tesis es presentada bajo la modalidad de compendio de artículos científicos. Por tanto, este trabajo se estructura de la siguiente forma:

Capítulo 1: ofrece al lector una visión de la motivación de este trabajo y del estado del arte de lo temas tratados.

Capítulo 2: enumera los distintos objetivos que guían el desarrollo de esta tesis.

Capítulo 3: presenta los aspectos claves de la metodología seguida para alcanzar los objetivos establecidos.

Capítulo 4: ofrece una discusión los resultados obtenidos como fruto de esta tesis doctoral.

Capítulo 5: incluye los diferentes artículos científicos que componen la tesis.

Artículo 1: presenta el trabajo de desarrollo del modelo Iber+. Una reimplementación del modelo original Iber optimizada para CPUs multinúcleo y GPUs. Aquí se estudia la nueva implementación en términos de precisión y eficiencia computacional comparando con el modelo original y otros modelos presentes en la literatura.

Artículo 2: introduce el desarrollo inicial de un sistema de alerta temprana de inundaciones diseñado para la cuenca del rio Miño a su paso por la ciudad de Lugo (Galicia, España). Este sistema utiliza un acoplamiento parcial y unidireccional entre un modelo hidrológico agregado (HEC-HMS) e Iber+.

Artículo 3: describe el desarrollo del módulo de calidad de aguas para Iber+ optimizado para CPUs multinúcleo y GPUs. A lo largo de esta publicación se abordan diferentes casos de aplicación en los que se analiza la nueva implementación en términos de eficiencia y precisión.

Artículo 4: presenta el sistema de alerta temprana para el río Miño MIDAS (Miño rIver flooD Alert System), una evolución del sistema introducido en el segundo artículo. Este nuevo sistema incorpora soporte para sistemas fluviales con múltiples cuentas y diferentes zonas de riesgo de inundación. Además también se propone una novedosa metodología hidrológica.

Artículo 5: analiza el evento de inundación histórico acaecido en la ciudad de Badajoz en el año 1876. Mediante los modelos HEC-HMS e Iber+ ha sido posible reproducir este evento, lo que permite una mejor comprensión de este tipo de fenómenos en la zona.

Capítulo 6: expone las principales conclusiones que se pueden extraer del trabajo realizado.

Objetivos

2

> *"There's a right thing to do,"Holden said. "You don't have a right thing, friend,"Miller said. "You've got a whole plateful of maybe a little less wrong."*

— James S.A. Corey
Leviathan Wakes

En este capítulo se enuncian los principales objetivos de esta tesis doctoral. A grandes rasgos, el trabajo presentado busca el desarrollo de herramientas de simulación numérica adecuadas para abordar problemas relacionados con la gestión de recursos hídricos en general y más concretamente el análisis y gestión de eventos de riesgo. Este objetivo general se descompone en una serie de subobjetivos específicos enumerados a continuación:

1. Aumentar la eficiencia computacional del modelo hidráulico Iber. Para ello se emplearán diferentes técnicas de programación de alto rendimiento que permiten explotar el paralelismo de las arquitecturas hardware actuales, más concretamente, sistemas CPU de memoria compartida y sistemas GPU. La consecución de este objetivo permite una reducción significativa de los tiempos de cálculo, abriendo la puerta a una serie de aplicaciones no viables hasta ahora por su elevado tiempo de cálculo. Este objetivo es abordado en el artículo *"An Accelerated Tool for Flood Modelling Based on Iber"* (García-Feal et al., 2018) (ver sección 5.1). Esta publicación detalla el desarrollo de Iber+ y analiza su desempeño en términos de eficiencia y precisión en relación con la implementación de referencia (Iber) y otras alternativas presentes en la literatura.

2. Desarrollo de una implementación eficiente de un acoplamiento entre un modelo hidráulico y un modelo de calidad de aguas. Debido a su elevado coste computacional, este tipo de modelado a menudo requiere asumir ciertas simplificaciones. Con el desarrollo de este objetivo se busca la implementación de una solución eficiente y completa para el modelado de calidad de aguas. Este punto se desarrolla en el artículo *"IberWQ: A GPU Accelerated Tool for 2D Water Quality Modeling in Rivers and Estuaries"* (García-Feal et al., 2020) (ver sección 5.3).

3. Desarrollo de un sistema de alerta temprana de inundaciones mediante el acoplamiento de un modelo hidráulico y un modelo hidrológico agregado. Para ofrecer una respuesta rápida a eventos de inundación, con el objetivo de mitigar los posibles daños, este sistema debe ser capaz de ofrecer simulaciones numéricas detalladas de las zonas sensibles de una forma eficiente. Para alcanzar esto, se utilizará el modelo hidráulico sólo en zonas sensibles a eventos de inundación y el modelo agregado en el resto del dominio. Este objetivo se aborda en las publicaciones *"Towards an automatic early warning system of flood hazards based on precipitation forecast: the case of the Miño River (NW Spain)"* (González-Cao et al., 2019) y *"MIDAS: A New Integrated Flood Early Warning System for the Miño River"* (Fernández-Nóvoa et al., 2020) (consultar secciones 5.2 y 5.4 respectivamente).

Finalmente, la experiencia acumulada y los desarrollos llevados a cabo para la consecución de los objetivos mencionados se aplica en la última publicación *"Numerical reconstruction of historical extreme floods: The Badajoz event of 1876"*. En la que se utiliza el modelado numérico para reconstruir un evento histórico de inundación que permite una mejor comprensión de este tipo de fenómenos en la zona y ofrece información valiosa a la hora de diseñar planes de prevención y contramedidas.

El esquema presentado en la Figura 2.1 relaciona, los desarrollos realizados como parte de este trabajo de tesis doctoral, las tecnologías y modelos empleados y las publicaciones resultantes.
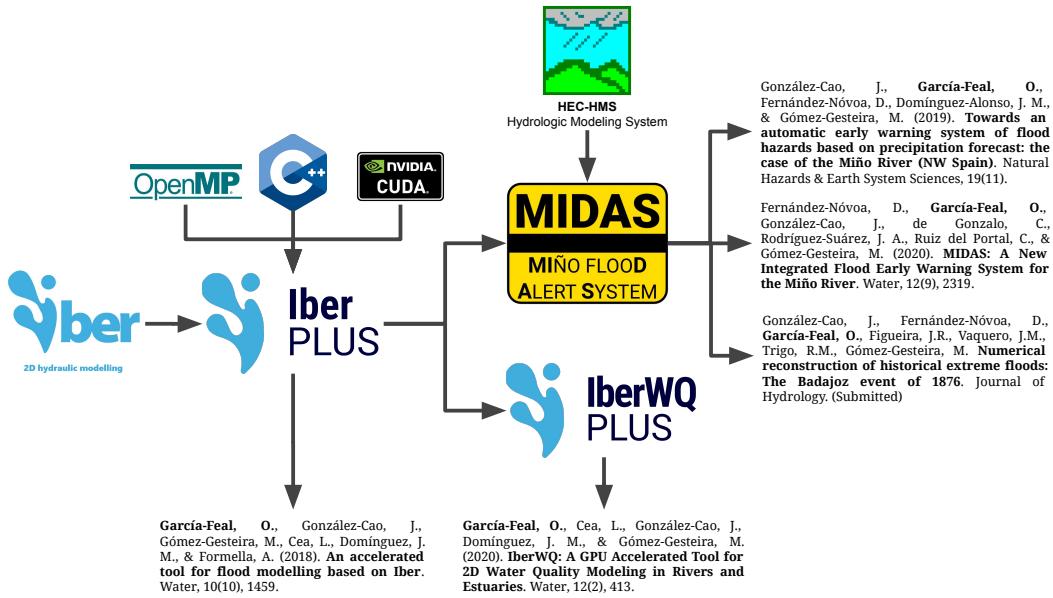

Figura 2.1: Esquema del trabajo desarrollado en la tesis doctoral. A partir del modelo Iber, se desarrolla una versión optimizada del modelo (Iber+) usando C++, CUDA y OpenMP. Esta versión optimizada sirve como base para el desarrollo del modelo de calidad de aguas y, junto con el modelo hidrológico HEC-HMS, el sistema de alerta temprana de inundaciones.

Metodología

<div style="text-align: right;">**3**</div>

> *He had a suspicion of plausible answers; they were*
> *so often wrong.*

<div style="text-align: right;">

— **Arthur C. Clarke**

Rendezvous with Rama

</div>

En este capítulo se recogen los aspectos clave de la metodología empleada para alcanzar los objetivos establecidos en el proyecto de la tesis doctoral.

3.1. El modelo Iber

Iber (Bladé et al., 2014) es un modelo numérico para la simulación de flujos en superficie libre en dos dimensiones en el que se promedia la componente vertical del flujo. Es por tanto adecuado para el estudio de eventos en cuencas fluviales y estuarios no estratificados. Fue desarrollado a partir de los modelos ya existentes Turbillón (Cea, 2005) y CARPA (Bladé Castellet, 2005) además de la interfaz GiD (Ribó et al., 1998) necesaria para realizar el pre y post-procesado de las simulaciones numéricas. Aparte del módulo principal que se encarga de resolver la hidrodinámica, Iber cuenta con diversos módulos que permiten simular múltiples procesos. Entre estos módulos destacan el módulo de turbulencia (Cea et al., 2007), el módulo de transporte de sedimentos (Corestein & Bladé, 2013), el módulo de procesos hidrológicos (Cea & Bladé, 2015), el módulo de calidad de aguas (Cea, Bermúdez et al., 2016) y el módulo para el estudio del hábitats fluviales (Sanz-Ramos et al., 2019).

Iber resuelve las ecuaciones de aguas someras en 2D también conocidas como como *Shallow Water Equations* o ecuaciones de Saint-Venant (De St Venant, 1871) en su forma conservativa. Éstas pueden ser escritas como sigue a continuación:

$$\frac{\partial h}{\partial t} + \frac{\partial hU_x}{\partial x} + \frac{\partial hU_y}{\partial y} = r - i \tag{3.1}$$

$$\frac{\partial hU_x}{\partial t} + \frac{\partial}{\partial x}\left(hU_x^2 + g\frac{h^2}{2}\right) + \frac{\partial}{\partial y}(hU_xU_y) =$$
$$-gh\frac{\partial Z_b}{\partial x} - \frac{\tau_{b,x}}{\rho} + \frac{\partial}{\partial x}\left(v_t h\frac{\partial U_x}{\partial x}\right) + \frac{\partial}{\partial y}\left(v_t h\frac{\partial U_x}{\partial y}\right) \tag{3.2}$$

$$\frac{\partial hU_y}{\partial t} + \frac{\partial}{\partial y}\left(hU_y^2 + g\frac{h^2}{2}\right) + \frac{\partial}{\partial x}(hU_xU_y) =$$
$$-gh\frac{\partial Z_b}{\partial y} - \frac{\tau_{b,y}}{\rho} + \frac{\partial}{\partial x}\left(v_t h\frac{\partial U_y}{\partial x}\right) + \frac{\partial}{\partial y}\left(v_t h\frac{\partial U_y}{\partial y}\right) \tag{3.3}$$

Donde h representa el calado de la lámina de agua, U_x y U_y son las velocidades horizontales promediadas, r es la precipitación, i es la tasa de infiltración del terreno, g es la aceleración de la gravedad, Z_b es la elevación del fondo, τ_b es la fricción originada por el rozamiento con el fondo, ρ es la densidad del agua y v_t es la viscosidad turbulenta. Para el cálculo de la fricción debido al rozamiento con el fondo se emplea la fórmula de Manning (Manning et al., 1890):

$$\tau_{b_x} = \rho g h\frac{n^2 U_x |U|^2}{h^{4/3}}, \qquad \tau_{b_y} = \rho g h\frac{n^2 U_y |U|^2}{h^{4/3}} \tag{3.4}$$

Iber también implementa los términos de fricción por viento, aceleración de Coriolis o presión baroclínica. Sin embargo, éstos no son mostrados en aras de la claridad.

En este trabajo, es de especial interés el modelo de calidad de aguas que se encuentra totalmente acoplado con el modelo hidrodinámico. De este modo, el modelo de calidad de aguas resuelve para cada una de las especies simuladas una ecuación de convección-difusión 2D promediada en profundidad:

$$\frac{\partial hC_i}{\partial t} + \frac{\partial hU_x C_i}{\partial x} + \frac{\partial hU_y C_i}{\partial y} = \frac{\partial}{\partial x}\left(hv_e\frac{\partial C_i}{\partial x}\right) + \frac{\partial}{\partial y}\left(hv_e\frac{\partial C_i}{\partial y}\right) + S_i h \quad (3.5)$$

Donde C_i es la concentración de la especie i de calidad de aguas, S_i es un término de reacción genérico para la especie i mientras que v_e es la viscosidad efectiva calculada utilizando un modelo de turbulencia. Iber implementa los modelos de turbulencia $k - \varepsilon$, longitud de mezcla y parabólico, los cuales se describen de forma pormenorizada en (Cea et al., 2007).

Tanto las ecuaciones que resuelven la hidrodinámica como el transporte de especies de calidad de aguas son resueltas utilizando el método de los volúmenes finitos (LeVeque, 2002). Para lo cual se realiza una discretización espacial del dominio en un conjunto de celdas que forman una malla. Iber trabaja con mallas no estructuras, en las cuales se pueden combinar diferentes celdas o elementos irregulares de tres o cuatro lados. Esto permite modelar de forma sencilla geometrías arbitrariamente complejas lo cual es especialmente ventajoso a la hora de trabajar con entornos fluviales. Para la discretización de los términos de flujo convectivo de la parte hidrodinámica se emplea un esquema descentrado de tipo Godunov, más específicamente el esquema descentrado de Roe (Roe, 1986) (implementados tanto de primer como segundo orden). Los datos de calado h y velocidades (U_x, U_y) necesarios para calcular las ecuaciones de transporte de calidad de aguas son proporcionados por el módulo hidrodinámico siguiendo

el método descrito en (Cea, Bermúdez et al., 2016; Cea & Vázquez-Cendón, 2012).

3.2. El modelo HEC-HMS

El modelo HEC-HMS (Hydrologic Engineering Center - Hydrologic Modeling System) (A. D. Feldman, 2000; Scharffenberg et al., 2018) está diseñado para simular procesos de precipitación-escorrentía en sistemas de cuencas fluviales. Este modelo fue desarrollado por el *U.S. Army Corps of Engineers* como un sucesor más avanzado del antiguo modelo HEC-1 (Goldman & Ely, 1998) cuyos orígenes datan de 1967.

Se considera un modelo hidrológico agregado o semi-distribuido en el que, a diferencia de un modelo distribuido, no considera de forma explícita las variaciones topográficas del terreno, usando en su lugar una serie de características promediadas. HEC-HMS permite realizar una descomposición del dominio en una serie de elementos hidrológicos que forman la red fluvial. Los elementos hidrológicos más destacables son: subcuencas, cauces, intersecciones y embalses.

Los principales procesos hidrológicos modelados por HEC-HMS son:

- Procesos de infiltración con los que se calcula la cantidad de precipitación que el terreno puede absorber y proporcionan la cantidad de dicha lluvia que es objeto de escorrentía. Entre los modelos implementados destacan el método del número de curva del SCS ((SCS), 1985), el método Green Ampt (Green & Ampt, 1911) y el método SMA (Soil Moisture Accounting) (Bennett, 1998) .

- Procesos de transformación en los cuales el exceso de precipitación que no es infiltrada por el terreno se transforma en escorrentía. Aquí se implementan diferentes métodos de hidrograma unitario como son los SCS ((SCS), 1986), Clark (Clark, 1945) o Snyder (Snyder, 1938). De este modo es posible reproducir la curva de respuesta de una subcuenca ante una entrada de precipitación.

- Procesos de flujo base que contribuyen al caudal de salida de una subcuenca que se pueden representar con diferentes técnicas como el método del reservorio lineal (A. D. Feldman, 2000; Zoch, 1934), el método de recesión (Scharffenberg et al., 2018) o el método Boussinesq no lineal (Scharffenberg et al., 2018; Tallaksen, 1995).

- Procesos de enrutado hidrológico que permiten calcular el flujo de agua a través del cauce fluvial. HEC-HMS implementa multitud de métodos, desde métodos sencillos sin atenuación a otros capaces de modelar fenómenos más complejos como el método de Puls modificado (Chow, 1964; A. D. Feldman, 2000; Henderson, 1966) o el método de Muskingum-Cunge (Ponce & Yevjevich, 1978).

Además, HEC-HMS cuenta con una interfaz gráfica de usuario que permite gestionar todos los parámetros y datos de entrada y salida de las simulaciones. Incluye también utilidades de post-procesado para generar informes e representar los resultados obtenidos.

3.3. Computación de altas prestaciones

La computación de altas prestaciones permite emplear sistemas informáticos con una alta capacidad de procesamiento para de resolver problemas complejos. La Unión Europea reconoce la computación de altas prestaciones como una capacidad crítica de cara a la economía y una herramienta esencial en casi todos

los ámbitos de la investigación, proponiendo un marco de colaboración para el desarrollo de esta capacidad estratégica ("The European High Performance Computing Joint Undertaking - EuroHPC", 2018; *Proposal for a Council Regulation on establishing the European High Performance Computing Joint Undertaking*, 2020).

Las técnicas de HPC se aplican a un gran rango de sistemas, desde grandes supercomputadores a ordenadores personales (Severance & Dowd, 2010). Con el objetivo de aprovechar los recursos disponibles, se ha puesto de manifiesto la necesidad de la aplicación de técnicas HPC, tradicionalmente reservadas a entornos de supercomputación, también en el ámbito de los ordenadores personales. Esto es debido a la progresiva tendencia en los últimos años a aumentar las capacidades de cómputo de los sistemas informáticos en forma de recursos de computación en paralelo (p.ej. CPUs multinúcleo o GPUs).

De cara a alcanzar el primer objetivo propuesto de esta tesis, se han evaluado diferentes ténicas HPC para aumentar la eficiencia del modelo Iber. Estas técnicas deben cumplir una serie de requisitos:

- Estar enfocadas a tecnologías maduras. Se trata de un proyecto a largo plazo, que será utilizado en producción y en el que participarán múltiples desarrolladores a lo largo del tiempo. Deben evitarse tecnologías poco extendidas o con futuro incierto ya que podrían requerir un cambio de paradigma a medio plazo. Esto ocurre con cierta frecuencia en un campo que evoluciona rápido, como por ejemplo el caso de Intel Xeon Phi (Cutress & Shilov, 2019; M. Feldman, 2017).

- Usar tecnologías ampliamente extendidas. El objetivo del proyecto es utilizar eficientemente las capacidades de cálculo de los recursos computacionales que ya existen en la mayoría de estaciones de trabajo y centros de cálculo. Esto deja a un lado tecnologías, de momento, menos comunes como las FPGAs.

- Las soluciones implementadas deben poder utilizarse en distintas plataformas. Desde entornos de escritorio bajo el sistema operativo Microsoft Windows a entornos GNU/Linux tanto para escritorio como para servidores.

Dado que el principal ámbito de utilización del modelo Iber es en estaciones de trabajo (*workstations*) se ha descartado inicialmente la utilización computación paralela orientada a sistemas de memoria distribuida como MPI (Snir et al., 1998) la cual se utiliza habitualmente en centros de cálculo y supercomputadores.

Los sistemas de memoria compartida como los equipos con CPUs *multinúcleo* se han extendido a casi todos los sectores de consumo. Estas CPUs cuentan cada vez con un mayor número de núcleos de procesamiento alcanzando cifras de 64 núcleos físicos en la actualidad (Suggs et al., 2020). Todo esto evidencia la necesidad de aplicar técnicas HPC orientadas a este tipo de tecnologías para sacar partido al hardware comercializado en la actualidad.

Otra tecnología ampliamente extendida en el mercado y que goza cada vez de una mayor popularidad son las GPUs. Originalmente se trataban de tarjetas gráficas cuyo único propósito era la visualización de gráficos. Debido a la creciente demanda de potencia de cálculo para mostrar cada vez gráficos 3D más complejos se han transformado en dispositivos con una elevada capacidad de cálculo. Esto ha derivado en que los fabricantes habiliten APIs para la utilización de este tipo de tecnología en tareas de cálculo de propósito general (GPGPU). A día de hoy se han convertido en una opción ampliamente extendida para acelerar cálculos en diversas aplicaciones como la codificación de vídeo, procesamiento de datos, cálculo científico, inteligencia artificial, etc. Es por ello que esta tecnología está disponible en la mayoría de equipos de consumo y centros de cálculo.

De todo esto se concluye que las implementaciones estarán enfocadas a sistemas CPU de memoria compartida y sistemas GPU.

3.3.1. Sistemas de memoria compartida: OpenMP

OpenMP se ha convertido en la opción más extendida para programación de sistemas CPU de memoria compartida en ámbitos científicos (Dagum & Menon, 1998). Es aplicable tanto en sistemas SMP (Symmetric MultiProcessing) donde todos los procesadores acceden a la misma memoria, como en sistemas NUMA (Non Uniform Memory Access) en los cuales cada procesador tiene su memoria local pero todos comparten el mismo espacio de direccionamiento. Está disponible en la mayoría de sistemas operativos y compiladores de C, C++ y Fortran. OpenMP provee una API sencilla y flexible que proporciona una serie de funciones y directivas que permiten la programación paralela en sistemas de memoria compartida. De este modo es posible crear una serie de hilos de ejecución, cada uno de los cuales procesará un subconjunto de datos en paralelo para resolver un problema.

Esta opción permite utilizar eficientemente los recursos disponibles en CPUs multinúcleo sin cambios significativos en el código. Aun así, es necesario revisar todas las estructuras de datos y bucles para evitar patrones de acceso ineficientes a la memoria y problemas derivados de posibles condiciones de carrera. En el primer caso, el rendimiento del programa resultante será deficiente pudiendo ser menor que una versión no paralelizada y en el segundo provocará que el programa sea incorrecto e impredecible.

3.3.2. Sistemas GPU de propósito general: Nvidia CUDA

CUDA (NVIDIA, 2020) es la plataforma para computación de propósito general por medio de GPUs del fabricante Nvidia. Lanzada inicialmente en el año 2007, es a día de hoy la solución GPGPU más usada en las aplicaciones de ámbito científico. Buena parte de este éxito ha sido gracias a la gran compatibilidad entre versiones, lo que permite que programas desarrollados para versiones

antiguas se puedan utilizar en GPUs actuales con poco o ningún esfuerzo. Otro factor a tener en cuenta es la posición dominante de Nvidia en el mercado de las GPUs dedicadas (Harris, 2020), por lo que es la opción más común en centros de cálculo y estaciones de trabajo. No obstante, a pesar de ser una tecnología propietaria y exclusiva de Nvidia, es posible convertir de forma automática código CUDA para usarlo en tarjetas AMD (Kuznetsov & Stegailov, 2019).

Las GPUs de Nvidia cuentan con una arquitectura denominada SIMT (Single Instruction Multiple Thread). Esta arquitectura emplea una gran cantidad de núcleos de procesamiento (en el orden de cientos o miles) organizados en SMs (Stream Multiprocessor). Cada uno de estos SM es capaz de ejecutar diversos hilos de ejecución de forma simultánea y siempre en grupos de 32 denominados *warps*. Sin embargo, los hilos que forman un *warp*, sólo pueden ejecutar una misma instrucción al mismo tiempo, de modo que si se alcanza una instrucción que ramifique el flujo de ejecución (como un *if*), se ejecutarán de forma secuencial los hilos que sigan un camino y a continuación los restantes. Esta es una de las peculiaridades de la programación de GPUs que puede suponer un grave impacto en el rendimiento de un programa si no se tienen en consideración.

Puesto que las tarjetas gráficas dedicadas cuentan con su propia memoria integrada al margen de la memoria del sistema, la transferencia de datos entre la memoria del sistema y la memoria de la GPU debe realizarse a través de un bus dedicado (generalmente PCI-Express) con un ancho de banda limitado que supone un cuello de botella. Por ello, que estas transferencias son costosas y deben limitarse en la medida de lo posible.

CUDA proporciona una serie de extensiones para diferentes lenguajes de programación como C, C++ o Fortran gracias a las cuales es posible programar las GPU. Dichas extensiones proporcionan tres abstracciones fundamentales de la arquitectura de la GPU:

Jerarquía de hilos de ejecución: En CUDA, una rutina que se ejecuta en GPU se denomina *kernel*. A su vez, un mismo *kernel* es ejecutado por un determinado número de hilos de ejecución o *threads*. Estos *threads* se organizan en subconjuntos denominados bloques. A su vez, el conjunto de todos los bloques de *threads* que ejecutan un *kernel* se denomina *grid*. Cada bloque se ejecuta exclusivamente en un sólo SM, sin embargo, cada SM puede ejecutar diversos bloques al mismo tiempo dependiendo del número de recursos de los que disponga.

Modelo de memoria: Existen diversos espacios de memoria a los que un *thread* puede acceder. Básicamente una memoria global a la que pueden acceder todos los *threads*, una memoria compartida a la que pueden acceder los *threads* de un mismo bloque y una memoria local para cada *thread*. Puesto que existen una gran diferencia de rendimiento entre los distintos espacios de memoria es fundamental tener en cuenta esta jerarquía de memorias para obtener un buen rendimiento.

Sincronización: En las GPUs, la sincronización global es un proceso costoso. Debido a ello, implementar ciertos algoritmos, como por ejemplo las reducciones, de forma eficiente no es trivial como en el caso de las CPU. Sin embargo, CUDA provee de mecanismos de sincronización también a nivel de bloque y operaciones atómicas en memoria compartida mucho más eficientes que el algoritmo deberá ser capaz de aprovechar.

Resultados y discusión

<div style="text-align: right">4</div>

> *Paréceme, Sancho, que no hay refrán que no sea*
> *verdadero, porque todos son sentencias sacadas de*
> *la misma experiencia, madre de las ciencias todas.*
>
> — **Miguel de Cervantes**
> El ingenioso hidalgo don Quijote de la
> Mancha

El primer objetivo de esta tesis doctoral se centra en la mejora de la eficiencia computacional del modelo hidráulico Iber. Para ello, partiendo de la implementación original de Iber en Fortran, se realizó una reimplementación del modelo en C++ utilizando el paradigma de programación orientada a objetos. De este modo se buscó crear un marco de trabajo o *framework* modular sobre el cual fuese posible poner a prueba nuevos conceptos y añadir nuevas funcionalidades de forma sencilla.

Esta nueva implementación, denominada Iber+, fue presentada en el artículo incluído en la sección 5.1. Esta primera versión implementó los módulos de hidrodinámica, turbulencia e hidrología del modelo Iber incluyendo una paralelización completa tanto para sistemas CPU de memoria compartida con OpenMP como para sistemas GPU con Nvidia CUDA.

Esta nueva implementación fue analizada en términos de eficiencia y precisión, empleando una batería de pruebas diseñada por la *Environmental Agency* del Reino Unido para evaluar software de simulaciones hidráulicas (Neelz & Pender, 2010). Este paquete de pruebas también se empleó para evaluar la precisión de la implementación de referencia de Iber en García-Feal et al., 2017 y más

tarde en Cea et al., 2020 mostrando resultados muy similares a otros modelos comúnmente usados. Para la evaluación de Iber+ se compararon los diferentes modelos presentados en la tabla 4.1, entre los que se incluyen los modelos comerciales TUFLOW FV ("TUFLOW FV Science Manual", 2013), InfoWorks ICM ("InfoWorks ICM Product Information - Overview", 2018) y JFLOW+ (Crossley et al., 2010). Todos los modelos analizados implementan las ecuaciones de aguas someras utilizando un esquema de volúmenes finitos y todos permiten la utilización de mallas no estructuradas de elementos mixtos salvo JFLOW+.

Modelo	Versión	Aceleración
TUFLOW FV	2012.000b	GPU
InfoWorks ICM	2.5.2	GPU
JFLOW+	2.0	CPU
Iber	2.4.3	CPU
Iber+ CPU	0.7	CPU
Iber+ GPU	0.7	GPU

Tabla 4.1: Modelos comparados para evaluar Iber+.

En términos de precisión, los resultados obtenidos por Iber e Iber+ (tanto en su versión CPU como GPU) son prácticamente idénticos, y muy similares a los resultados ofrecidos por los modelos comerciales.

En cuanto eficiencia, se calcula la aceleración obtenida usando programación paralela respecto al tiempo de ejecución empleado por el modelo Iber usando un sólo hilo de ejecución. Debido a la paralelización parcial de su código, el modelo Iber sólo es capaz de obtener como máximo una aceleración de 1,7 utilizando 16 hilos de ejecución. Sin embargo, Iber+ es capaz de alcanzar una aceleración de hasta 15,4 utilizando 16 hilos y hasta 94,3 utilizando una GPU de gama de consumo. Respecto a los modelos comerciales, Iber+ muestra una eficiencia comparable, obteniendo mejores resultados en el caso más complejo de los analizados.

El hecho de poder alcanzar un aumento de velocidad de cerca de dos órdenes de magnitud utilizando una GPU en un equipo de sobremesa, sin necesidad de

usar un centro de cálculo, abre un amplio abanico de nuevas aplicaciones de cara a modelar eventos de riesgo. En la última parte del artículo se analiza un caso de aplicación en el que se estudia un evento de inundación en el río Aragón (Naverac et al., 2012; Serrano-Muela et al., 2015). En este estudio se realizan un total de 16 simulaciones de 24 horas de tiempo simulado y una malla de medio millón de elementos para calibrar los parámetros necesarios. Iber+ es capaz de realizar cada una de esas simulaciones en menos de 10 minutos mientras que con Iber son necesarias más de 15 horas. Esto implica que Iber+ necesitó menos de tres horas de cálculo para realizar todas las simulaciones, mientras que Iber necesitaría más de diez días.

En la misma línea de desarrollo, partiendo del trabajo previo de Cea, Bermúdez et al., 2016 y utilizando el *framework* creado para Iber+ se implementa el módulo de calidad de aguas para Iber+ (ver sección 5.3). Para ello se realiza un acoplamiento completo del modelo hidrodinámico de Iber+ con un modelo de calidad de aguas que soporta hasta 14 parámetros de calidad de aguas, entre las que se encuentra pH, salinidad, *Escherichia coli*, oxígeno disuelto, nitrógeno orgánico, fósforo, etc.

La nueva implementación con calidad de aguas es evaluada utilizando para ello cuatro casos de prueba. En términos de precisión, la nueva implementación obtiene resultados prácticamente idénticos a la versión no paralelizada, mostrando unos valores promedio de NRMSD (raíz de la desviación cuadrática media normalizada) menores al 0,15 %. Sin embargo, en cuanto a eficiencia, Iber+ es capaz ofrecer importantes mejoras en la velocidad de ejecución, especialmente en casos de mayor tamaño donde al saturar la capacidad de la GPU, Iber+ multiplica la velocidad de Iber por 181.

Este tipo de modelos suponen una herramienta muy útil en la gestión ambiental de los recursos hídricos. Sin embargo, debido a la complejidad de los procesos que se intentan reproducir, deben asumirse una serie de simplificaciones. Por tanto, debe realizarse una correcta parametrización del modelo con el objetivo

de obtener resultados precisos. Este proceso requiere la ejecución de un elevado número de simulaciones de un elevado coste computacional al requerir generalmente la simulación de largos periodos de tiempo sobre dominios relativamente grandes. Gracias a la utilización de técnicas HPC, Iber+ es capaz de reducir la duración de este proceso que puede durar meses a días u horas. Además, debido a la eficiencia mejorada, también abre la puerta a su uso en sistemas de alerta temprana.

El objetivo de aplicar los desarrollos de Iber+ al diseño de un sistema de alerta temprana fue abordado en las publicaciones presentadas en las secciones 5.2 y 5.4. Ambos trabajos están enfocados a la implementación de un sistema de alerta temprana de inundaciones a lo largo del cauce del río Miño. El sistema desarrollado utiliza diferentes fuentes de datos:

- Datos de predicción de lluvia obtenidos a través del modelo WRF (Weather Research and Forecasting) (Skamarock et al., 2005) y ofrecidos por la agencia meteorológica MeteoGalicia ("MeteoGalicia", 2021).

- Datos reales de precipitación obtenidos de la red de pluviómetros de MeteoGalicia.

- Datos de caudales obtenidos en diferentes estaciones de control de la red SAIH (Sistema Automático de Información Hidrológica) de la Confederación Hidrográfica del Miño-Sil ("SAIH del Miño-Sil. Confederación Hidrográfica del Miño-Sil." 2021).

Este sistema es capaz de obtener automáticamente estos datos (predicción, precipitación y caudales) para alimentar el modelo hidrológico HEC-HMS y predecir si el caudal obtenido puede originar una situación de riesgo. En tal caso, automáticamente realiza una simulación de las zonas de interés obteniendo mapas de peligrosidad. Si en las zonas sensibles se detecta una situación de riesgo para los viandantes según el criterio de (Cox et al., 2010), se comunica una alerta a las correspondientes autoridades con información detallada de la

peligrosidad, calados y velocidades. De esta forma, el sistema de alerta desarrollado proporciona la información necesaria para tomar las medidas oportunas que limiten los daños. Gracias al empleo del modelo Iber+, el sistema puede obtener toda esta información en el plazo de tan sólo unos minutos, dando a las autoridades el margen necesario para tomar medidas en el menor tiempo posible.

En el primer estudio se aplica el sistema de alerta temprana al tramo inicial del río Miño hasta la ciudad de Lugo (González-Cao et al., 2019). En el apartado hidrológico se utiliza el método de número de curva para modelar la infiltración, el método del hidrograma unitario del SCS (Soil Conservation Service) para proceso de transformación (Service, 2007) y el método Muskingum-Cunge para el enrutado. Se realizó un proceso de calibración para los parámetros del número de curva y tiempo de retardo para cada subcuenca utilizando eventos extremos ocurridos entre los años 2008 y 2018. Las pruebas con los eventos de validación mostraron unos resultados considerados como *buenos* según el criterio de (N. Moriasi et al., 2007) para la evaluación de modelos hidrológicos. Esta metodología ofreció buenos resultados para cuencas de tamaño relativamente pequeño y eventos de corta duración, teniendo siempre en cuenta la dependencia de la calidad de la predicción meteorológica.

En el segundo estudio se desarrolló un sistema de alerta temprana modular para soportar sistemas fluviales de tamaño arbitrario (Fernández-Nóvoa et al., 2020), el cual fue bautizado como MIDAS (Minho rIver flooD Alert System). En este trabajo se aumentó el dominio de la simulación hasta la ciudad de Ourense. Dadas las limitaciones del modelo de infiltración basado en el número de curva para utilizarlo en eventos largos o de modo continuado, se desarrolló un método novedoso que permite la utilización del método de número de curva en modo continuo. Este método procesa varios eventos consecutivos de forma aislada para luego reconstruir la señal de caudal resultante en la cuenca analizada. Con este método se han obtenido predicciones de caudal a escala

diaria para un período entre los años 2012 y 2019, logrando unos coeficientes de correlación de Spearman mayores a 0,90. Considerando sólo eventos de riesgo en los que el caudal supera el percentil 99 y evaluados a escala horaria se consiguen correlaciones mayores de 0,85. Estos resultado son considerados "muy buenos" según los criterios de (N. Moriasi et al., 2007).

También se realizó un análisis de la capacidad para predecir situaciones de riesgo en zonas inundables. Para esto se llevó a cabo un estudio de los umbrales de caudal que generaban una situación de riesgo estableciendo cuatro niveles de alerta y posteriormente un análisis histórico. Los resultados del análisis indican que el sistema es capaz de predecir correctamente el 76 % de las situaciones de alerta mientras que generó un 28 % de falsos positivos, lo cual indica un buen rendimiento en términos de predictibilidad.

La utilización del modelo Iber+ sobre GPU, permite la obtención de mapas de inundación para cada una de las ciudades analizadas en un plazo de entre 3 y 5 minutos para cada 24 horas simuladas. Esto no sólo permite obtener una amplia información sobre un evento de inundación en un plazo muy breve, sino que también permite realizar un análisis histórico creando mapas de inundabilidad, donde se puede apreciar el número promedio de días que un área permanece inundada.

Por último, la experiencia acumulada en la aplicación de los modelos Iber+ y HEC-HMS a eventos extremos de inundación ha permitido el desarrollo del estudio presentado en la sección 5.5. El estudio de eventos de inundación históricos resulta fundamental a hora de comprender mejor estos fenómenos y qué medidas se pueden adoptar en un futuro para mitigar sus posibles efectos. En este caso se ha realizado un estudio de los eventos de inundación ocurridos en la ciudad de Badajoz en diciembre del año 1876 (Ortega Becerril, 2008; Trigo et al., 2014), correspondiente a uno de los eventos de precipitación más extremos registrados en el oeste de la península ibérica (Fragoso et al., 2010).

Para llevar a cabo el estudio de este evento fue necesaria la reconstrucción espacio-temporal de los datos de precipitación a lo largo de la cuenca del río Guadiana. Para ello se utilizaron los datos de los dos pluviómetros más cercanos disponibles y la base de datos de precipitaciones Spain02 (Herrera et al., 2012). Una vez obtenidos los datos de precipitación, se utilizó el modelo HEC-HMS para estimar el caudal de entrada en Badajoz. A su vez, también se realizó un estudio para estimar el pico de caudal usando el modelo Iber+ y los datos disponibles para la inundación a su paso por Mértola (Portugal), obteniendo valores pico muy similares con ambas metodologías.

Finalmente, con los datos de caudal se realizó una reproducción numérica de la inundación a su paso por Badajoz con el modelo Iber+ teniendo en consideración los posibles cambios en la topografía y los usos de suelo. Los resultados obtenidos están en concordancia con los datos disponibles en la literatura y además arrojan información adicional sobre el evento, demostrando la utilidad del modelo Iber+ a la hora de reproducir este tipo de eventos donde la disponibilidad de datos es limitada.

Conjunto de publicaciones

<div style="text-align: right; font-size: 3em;">5</div>

5.1.   An Accelerated Tool for Flood Modelling Based on Iber

García-Feal, O., González-Cao, J., Gómez-Gesteira, M., Cea, L., Domínguez, J. M. & Formella, A. (2018). An accelerated tool for flood modelling based on Iber [Publisher: Multidisciplinary Digital Publishing Institute]. *Water*, *10*(10), 1459

Información de la revista	
Revista	Water
eISSN	2073-4441
País	Suíza
Editorial	MDPI
Factor de impacto	2,524 (2018)
Cuartil	Q2

An Accelerated Tool for Flood Modelling Based on Iber

Orlando García-Feal [1],*, José González-Cao [1], Moncho Gómez-Gesteira [1], Luis Cea [2] ,
José Manuel Domínguez [1] and Arno Formella [3]

1 Environmental Physics Laboratory (EPHYSLAB), Universidad de Vigo, Campus As Lagoas s/n,
32004 Ourense, Spain; jgcao@uvigo.es (J.G.-C.); mggesteira@uvigo.es (M.G.-G.);
jmdominguez@uvigo.es (J.M.D.)
2 Environmental and Water Engineering Group, Departamento de Ingeniería Civil, Universidade da Coruña,
Campus Elviña s/n, E-15071 A Coruña, Spain; luis.cea@udc.es
3 Laboratorio de Informática Aplicada, Universidade de Vigo, Campus As Lagoas s/n, 32004 Ourense, Spain;
formella@uvigo.es
* Correspondence: orlando@uvigo.es; Tel.: +34-988-372-255

Abstract: This paper presents Iber+, a new parallel code based on the numerical model Iber for two-dimensional (2D) flood inundation modelling. The new implementation, which is coded in C++ and takes advantage of the parallelization functionalities both on CPUs (central processing units) and GPUs (graphics processing units), was validated using different benchmark cases and compared, in terms of numerical output and computational efficiency, with other well-known hydraulic software packages. Depending on the complexity of the specific test case, the new parallel implementation can achieve speedups up to two orders of magnitude when compared with the standard version. The speedup is especially remarkable for the GPU parallelization that uses Nvidia CUDA (compute unified device architecture). The efficiency is as good as the one provided by some of the most popular hydraulic models. We also present the application of Iber+ to model an extreme flash flood that took place in the Spanish Pyrenees in October 2012. The new implementation was used to simulate 24 h of real time in roughly eight minutes of computing time, while the standard version needed more than 15 h. This huge improvement in computational efficiency opens up the possibility of using the code for real-time forecasting of flood events in early-warning systems, in order to help decision making under hazardous events that need a fast intervention to deploy countermeasures.

Keywords: flood; numerical simulation; shallow water equations; Iber+; benchmark; CUDA; OpenMP; finite volume

1. Introduction

Floods are a type of natural disaster that have affected human activity throughout history. In recent years, these phenomena have become more frequent and intense due to climate change [1,2]. The development of numerical tools that are able to simulate these events has become essential. These tools must be accurate, in order to provide useful data, as well as computationally efficient, to be able to obtain results in reasonable computational times. The information provided by the numerical models should help decision makers to design resilient structures, as well as to estimate the intensity of an imminent extreme event in order to implement countermeasures that avoid or mitigate the economic and human losses.

Iber [3] is a numerical model that solves the two-dimensional (2D) depth-averaged shallow water equations with an unstructured explicit finite volume solver. In addition to the hydraulic module,

it has a sediment transport and a water quality module [4] to solve transport processes in free surface shallow flows. It implements the first order and second order extension of the upwind scheme of Roe [5] to model flood inundation, and the DHD (decoupled hydrological discretization) scheme [6] to solve the shallow water equations in rainfall–runoff applications. The algorithms implemented in the model have been extensively validated and applied in previous studies related to river inundation, tidal currents in estuaries and rainfall–runoff modelling [6–10].

One of the main limitations of Iber at the present time is the CPU time needed to perform simulations over complex and large spatial domains (of several km^2). This is because the model lacks the necessary optimizations to take advantage of the parallelism available on current hardware. This hampers the use of the model in interesting applications such as real-time flood forecasting in early warning systems, Monte Carlo-based calibration and uncertainty analysis [10,11], high-resolution rainfall–runoff simulation in medium and large-size watersheds, or continuous simulation methods applied to flood frequency analysis [12].

Hydraulic models are commonly accelerated for shared memory multiprocessor systems with OpenMP (open multi-processing), but the speedup that can be achieved by this procedure is limited. Alternatively, MPI (message passing interface) can be used to take advantage of distributed memory supercomputers, but are expensive and difficult to maintain. Another way to accelerate this kind of code is the use of GPUs (graphical processing units). This technology offers a high amount of parallel processing power in quite inexpensive cards that can be installed in a server or workstation. GPUs have been employed successfully in mesh-free [13] and mesh-based [14–17] models reaching speedups of two orders of magnitude.

In this work, a new parallel implementation of the hydraulics module of Iber, named Iber+, is presented. The new code is written in C++ and it makes use of OpenMP and CUDA [18] to accelerate the simulations. The rest of this paper is organized as follows. In Section 2, the governing equations and the most relevant techniques used in the parallel implementation are detailed. In Section 3, four different benchmark cases are used to validate the accuracy and analyse the computational efficiency of Iber+ when compared with other hydraulic models. Finally, in Section 4, Iber+ is applied to compute rainfall–runoff during a real flash-flood event in a mountain headwater catchment of 240 km^2.

2. Methods

2.1. Hydrodynamic Model

The 2D depth-averaged shallow water equations solved in Iber can be written as:

$$\frac{\partial h}{\partial t} + \frac{\partial h U_x}{\partial x} + \frac{\partial h U_y}{\partial y} = r - i \tag{1}$$

$$\frac{\partial h U_x}{\partial t} + \frac{\partial}{\partial x}\left(h U_x^2 + g\frac{h^2}{2}\right) + \frac{\partial}{\partial y}\left(h U_x U_y\right) = -gh\frac{\partial Z_b}{\partial x} - \frac{\tau_{b,x}}{\rho} + \frac{\partial}{\partial x}\left(v_t h\frac{\partial U_x}{\partial x}\right) + \frac{\partial}{\partial y}\left(v_t h\frac{\partial U_x}{\partial y}\right) \tag{2}$$

$$\frac{\partial h U_y}{\partial t} + \frac{\partial}{\partial y}\left(h U_y^2 + g\frac{h^2}{2}\right) + \frac{\partial}{\partial x}\left(h U_x U_y\right) = -gh\frac{\partial Z_b}{\partial y} - \frac{\tau_{b,y}}{\rho} + \frac{\partial}{\partial x}\left(v_t h\frac{\partial U_y}{\partial x}\right) + \frac{\partial}{\partial y}\left(v_t h\frac{\partial U_y}{\partial y}\right) \tag{3}$$

where h represents the water depth, U_x and U_y are the averaged horizontal velocities, g is the acceleration of the gravity, ρ is the density of the water, Z_b is the bed elevation, τ_b is the bed friction, v_t is the turbulent viscosity, r is the rainfall intensity and i is the infiltration rate. Iber also implements terms to account for wind surface friction, the Coriolis acceleration and baroclinic pressure, but those terms are not included here for the sake of simplicity. In Iber, the bed friction is computed with the Manning formulation as:

$$\tau_{b,x} = \rho g h \frac{n^2 U_x |U|^2}{h^{4/3}}, \quad \tau_{b,y} = \rho g h \frac{n^2 U_y |U|^2}{h^{4/3}} \tag{4}$$

2.2. Numerical Code Iber+

Iber+ is a new code that implements a parallelization of the hydraulic module of Iber [3], which is a hydraulic model that solves 2D shallow water equations using an unstructured finite volume solver. The software package Iber provides a user-friendly graphical user interface for pre- and post-processing and can be freely downloaded from http://iberaula.es. The Iber code is programmed in Fortran and it is partially parallelized with OpenMP. Even if the computation time is reduced by using this technique, the code is unable to efficiently use current multi-core processors, so the speedup that can be achieved is quite limited. Simulation time becomes critical when addressing certain kinds of problems such as real-time flood forecasting, Monte Carlo-based simulations, high-resolution rainfall–runoff modelling, or long term simulations.

The main aim of the Iber+ code is to significantly improve the computational efficiency of Iber, while being fully compatible with the Iber software. The new code has an object-oriented implementation programmed in C++. It is parallelized for shared memory systems with OpenMP and it also provides an Nvidia CUDA implementation for execution in GPUs.

OpenMP is an API (application programming interface) that allows easy parallelization of traditional loops using compiler directives. However, to achieve significant speedups, there are several aspects to be taken into account. The parallelization of a loop generates an overhead, so the parallelization of low-cycle loops could be counter-productive. It is also important to consider the granularity of loops. Combining simple loops into larger loops can reduce parallelism overhead, however this could affect the automatic vectorization. With vectorization enabled in the compiler, a program can use the SIMD (single instruction multiple data) instruction set, e.g., SSE (streaming SIMD extensions) or AVX (advanced vector extensions), that provides extra computing power in modern CPUs. However, loops that use branching or other complex structures may not be vectorized automatically. On the other hand, the principle of locality is fundamental. Modern CPUs have a complex memory hierarchy, including several levels of cache memory. Accessing a memory position that is not in cache implies a significant penalty. Therefore, memory access patterns should be studied and the design of suitable data structures is fundamental to achieve good performance. For instance, choices between using an array of structures or a structure of arrays should be evaluated for each algorithm.

GPUs use highly parallel architectures that confer them a high amount of computing power. This is needed to render complex 3D computer graphics scenes. Since most GPUs are programmable, the manufacturers provide APIs for GPGPU (general processing graphics processing unit) computing. This technology makes the processing power of GPUs available to problems not necessarily related to graphics. One of the most common GPGPU APIs used for scientific purposes is Nvidia CUDA, providing access to the GPUs with traditional programming languages like C/C++ or Fortran.

Nvidia GPUs are made of a large amount of processors organized in streaming multiprocessors (SMs) employing a single instruction multiple thread (SIMT) architecture. Each SM executes several threads in parallel, in groups of 32, called warps. All the threads start on the same program address but have their own registers. However, one warp can only execute one instruction at a time. If one of the threads branches to a different path than the rest, a divergence occurs. In that case, that path will be executed while the other threads are stalled, and later the other path, until the threads converge again. This procedure can produce a heavy performance penalty if it occurs frequently.

Another critical issue is reduction algorithms. In GPUs, global synchronization is expensive, so this kind of algorithm is not as trivial as in CPUs. To avoid heavy performance penalties, an alternative approach must be employed. In Iber+, the library Nvidia CUB (CUDA unbound) was employed to implement reduction algorithms. CUB is an open-source high performance library [19] developed by Nvidia that provides reusable pieces of software for CUDA programming.

The last issue to address when programming GPUs is data transfer. Unlike integrated GPUs of mobile devices where the GPU and CPU share the same memory, discrete high-performance GPUs are issued in separate cards with their own memories. The GPU memory is usually faster and smaller than

the main system memory and is located in a different address space. As the memory transfers from the system memory to the GPU memory are done via the PCI (peripheral component interconnect) bus, the bandwidth is limited and causes a bottleneck that needs attention. Even though the API can provide a unified address space, data transfers should be carefully made to avoid performance penalties. As data transfers are expensive, it could be more profitable to do certain computations (like reductions) on the GPU and being slower than on the CPU, rather than needing to transfer the data to system memory and run them on the CPU. Figure 1 shows the flow chart of the Iber+ execution. Once the simulation is started, most of the computations are performed on the GPU, minimizing the data transfers.
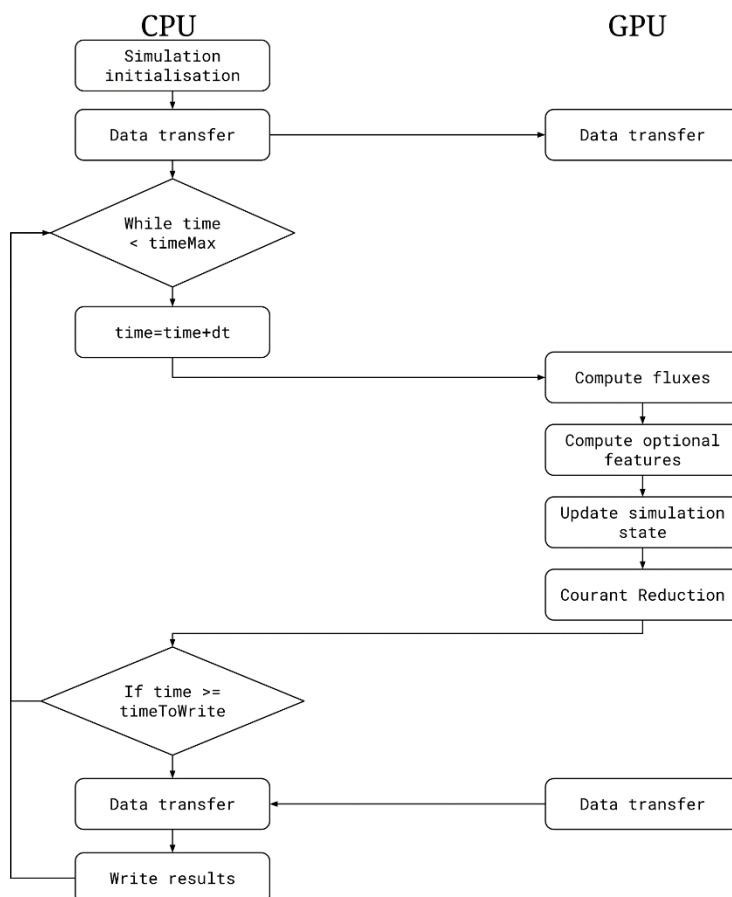
Figure 1. Flow chart of Iber+ GPU implementation.

The speedup achieved by using a GPU in respect to the CPU implementation increases as the problem size gets bigger. The speedup increases until the GPU computation capacity is saturated; this occurs when all GPU processors are effectively used.

Both the Iber and Iber+ CPU code uses only double precision computations. However, in GPU computing, using double precision could suppose a considerable performance penalty. Depending on the specific GPU model, the theoretical performance in double precision could be from two to more than ten times slower [20]. In this work, the GPU computations were performed in single precision because no significant differences were found in the cases analyzed. However, some applications may require the use of double precision arithmetic, so this option will be added in further versions.

The Iber+ implementation is based on the Iber code, but reimplemented from scratch in C++. No substantial modifications were made to the original algorithms. The computations are made using unstructured meshes and the fluxes are solved at the edges of the underlying graph data structure. However, data structures were revised and optimized to reduce the memory footprint and to improve the memory access patterns. Even though Iber has most of its loops parallelized

with OpenMP, these were redesigned in Iber+ to avoid critical regions and data dependencies that limited the speedup that could be achieved. These and other minor optimizations were done for the computation routines. Moreover, I/O (input/output) routines were optimized to avoid unnecessary data reading and writing.

For the GPU implementation, several drawbacks as mentioned previously should be considered. In order to achieve the highest speedup with GPU computing, memory transfers should be reduced as much as possible. Therefore, most of the computation routines were implemented in CUDA. Once the simulation is initialized, the data is transferred to the GPU. Then, most of the simulation is performed on the GPU and the only data transferred from the GPU are the results that should be written to disk or single variables like the time step. Additionally, as the CPU is free while the GPU is performing the computations, the CPU can run other tasks like writing the results to disk in parallel, hence, further improving the overall run time of the simulation.

3. Validation Tests

To validate the new code, four tests included in the benchmark of the hydraulic modelling packages published by the Environmental Agency of the U.K. government [21] were replicated. This benchmark consists of a set of nine tests that were run with 19 hydraulic models. The models were compared in terms of numerical accuracy and computational efficiency. Note that neither experimental nor theoretical solutions are provided as reference data, so the only way of comparing the model results is against other code. The three numerical models chosen to be compared with Iber and Iber+ are TUFLOW FV (v2012.000b) [22], InfoWorks ICM (v2.5.2) [23] and JFlow+ (v2.0) [14]. These models were chosen due their popularity [24] and similarity with Iber, since all of them solve 2D shallow water equations using finite volume solvers. Different hardware configurations were used in the validation of Iber+. The CPU simulations (both Iber and Iber+ CPU) were run in a server with an Intel Xeon E5-2695 v4 processor (18 cores). The GPU simulations (Iber+ GPU) were run in a workstation with an Intel Core i7-4770 CPU and two different GPUs: a modern GPU (Nvidia GTX 1080) and an older GPU (Nvidia GTX 480). The latter one was used to carry out a more objective comparison with JFlow+ and InfoWorks ICM, which used a Nvidia GTX 285 and a Nvidia Tesla C2050, respectively. The Nvidia GTX 480 offers a comparable performance [25,26] to those cards and shares the same architecture (Fermi) as the Tesla C2050. The main characteristics of the previously mentioned GPUs are shown in Table 1.

Table 1. Characteristics of the Nvidia GPUs used by the compared models. Memory units are expressed in gigabytes (GB). GDDR3, GDDR5 and GDDR5X refers to different specifications of Graphics Double Data Rate memory types.

Model	Release Date	Micro-Architecture	Code Name	CUDA Cores	Base Frequency	Memory
GTX 285	January 2009	Tesla	GT200-350-B3	240	1476 Mhz	1 GB GDDR3
GTX 480	March 2010	Fermi	GF100	480	1401 Mhz	1.5 GB GDDR5
Tesla C2050	July 2011	Fermi	GF100	448	1150 Mhz	3 GB GDDR5
GTX 1080	May 2016	Pascal	GP104-400-A1	2560	1607 Mhz	8 GB GDDR5X

All the validation tests presented in this section were run with the first order of the discretization scheme of Roe and a wet-dry tolerance of 0.0001 m. Notice that the nomenclature of the tests, as well as the numbering of the control points, are the same as those used in [21] for an easier comparison with the original benchmark document.

3.1. Test 1: Flooding a Disconnected Water Body

3.1.1. Case Description

The first test is designed to check the basic capabilities of the hydraulic models. The domain is shown in Figure 2. It consists of a rectangular flume of 700 m length and 100 m width with a

variable elevation (Figure 3a). The initial water level is 9.7 m in the whole flume. The spatial domain is discretized using a structured grid with an element size of 10 m. The Manning's coefficient is set to 0.03 s/m$^{1/3}$. The inlet boundary condition is defined at the left side of the domain (red line in Figure 2). The inlet hydrograph is shown in Figure 3b. The total physical time is 20 h. Time series of the water level were extracted at the two control points shown in Figure 2, in order to compare the results with the other software packages.
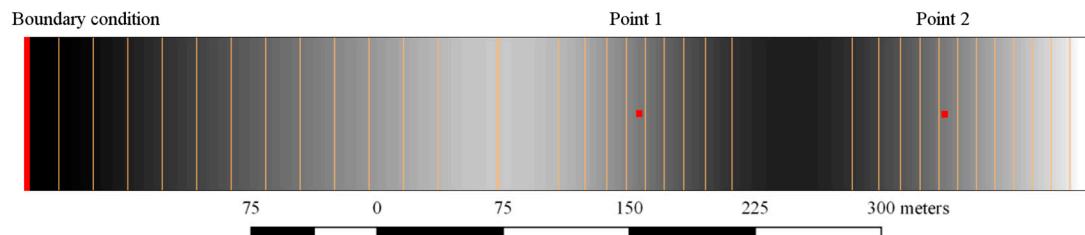
Figure 2. Simulated domain for Test 1. White areas represent high terrain elevation. To the left, there is an inlet boundary condition. In red, the measurement points.

(**a**) (**b**)
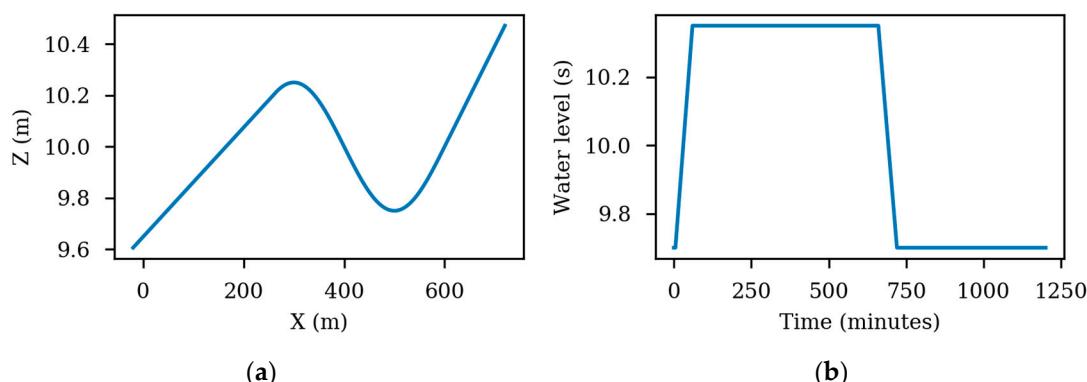
Figure 3. Elevations along the domain. Z stands for elevation in meters and X for horizontal distance from the origin of the channel (**a**). Hydrograph defined for the inlet boundary condition (**b**).

3.1.2. Results

Figure 4 shows the time series of the water level obtained at the two control points for the simulations carried out with TUFLOW FV, InfoWorks ICM, JFlow+ and Iber+. The water levels obtained with Iber and Iber+ using both the CPU and GPU are almost the same. The differences are due to round-off errors of the different implementations. These differences could not be appreciated in the plots. The coefficient of variation of the *RMSD* (root mean square deviation) is defined as follows:

$$CV(RMSD) = \frac{RMSD}{\overline{x}}, \quad RMSD = \sqrt{\frac{\sum_{i=1}^{N}(x_i - y_i)^2}{N}} \tag{5}$$

were x and y are the time series and N their number of elements. This was used to measure the error of Iber+ (including CPU and GPU implementation) regarding the Iber time series. The coefficients obtained were lower than 0.001%, so for the sake of clarity, only one plot has been used for the different Iber implementations. The rest of the numerical models performed similar to the Iber implementations and only minor differences have been found.
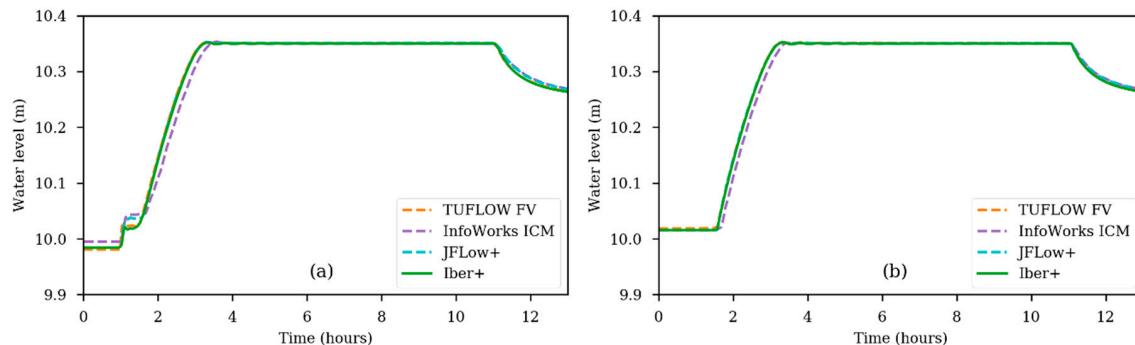
Figure 4. Time series of the water level for Test 1 obtained using all the numerical models at Point 1 (**a**) and at Point 2 (**b**).

Table 2 shows that the fastest code is Iber+ running on the Nvidia GTX 1080 GPU. It runs 16 times faster than Iber. When running on a GTX 480, a similar GPU to those used by InfoWorks and JFlow+, Iber+ is also faster than those codes. Respecting the execution on the CPU, Iber+ runs five times faster than Iber running both on a single thread (see Figure 5b), and 8.8 times faster using four threads. Note that in this test with 16 threads, Iber+ becomes slower compared to using just four threads. Creating a thread is a time-consuming process. Assigning a big workload to each thread will mitigate the overhead of creating the threads because the time needed to process the workload is much larger than the time used to create the threads. On the other hand, the less workload each thread has, the more significant the overhead in the overall run time. In this case, the mesh has only 700 elements, so if there are many threads involved, each thread has a very small workload. As shown in Figure 5b, there is almost no benefit in using four versus two threads. Moreover, using eight threads is slower than using two, and using 16 threads is even slower than a single thread.

Table 2. Performance measurements obtained for Test 1. Total run time of the simulation in seconds, average processing time per time step in milliseconds and the achieved speedups compared with Iber running in a single thread.

| Model	Hardware Configuration	Run Time (s)	Time per Step (ms)	Speedup vs. Iber 1 Thread
InfoWorks ICM	GPU	9	-	7.7
JFlow+	GPU	28	-	2.5
TUFLOW FV	12 Threads	4.4	-	15.7
Iber	1 Thread	69.2	0.77	1.0
	4 Threads	48.7	0.49	1.4
	16 Threads	43.5	0.42	1.6
Iber+	1 Thread	13.8	0.18	5.0
	4 Threads	7.9	0.10	8.8
	16 Threads	16.1	0.20	4.3
	GPU (GTX 480)	6.6	0.08	10.4
	GPU (GTX 1080)	4.3	0.05	16.0
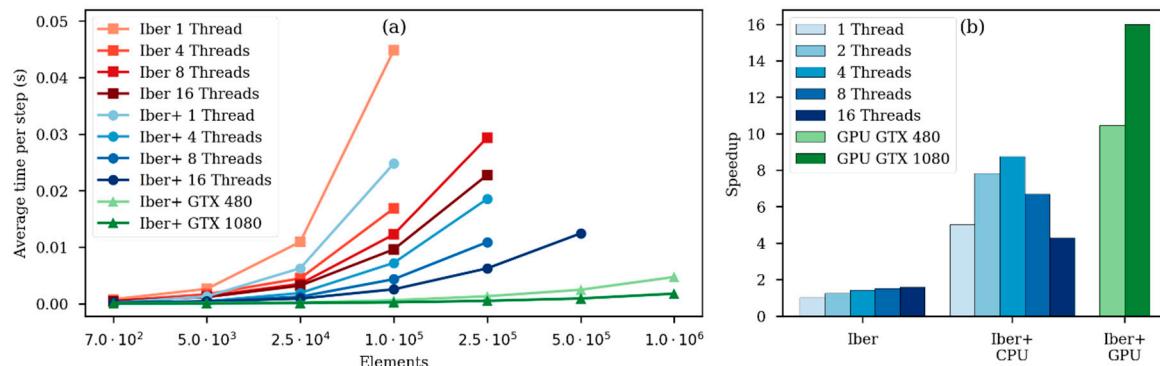
Figure 5. Average processing time per time step versus the number of elements of the simulated mesh using different configurations (**a**); Speedups obtained comparing Iber running in one thread with different configurations in Test 1 using a mesh of 700 elements (**b**).

This specific case will be studied in detail to show that the benefits of using parallel executions increases as the mesh size gets bigger. This case was simulated using mesh with different characteristic lengths. As we reduce the element size in the mesh, the Courant [27] increases so the number of steps needed to complete the simulation increases as well. Therefore, the total run-time is not an appropriate metric to analyze the effect of the mesh size on the advantages of parallelization. An alternative metric to analyze this effect is the average execution time per step, defined as the total run-time over the total number of time steps. As shown in Figure 5a, the use of more threads in CPU computations or GPU computing improves the runtime when the number of nodes (elements) increases, since the overhead due to parallelism becomes less significant.

The overhead produced by using parallelism can be measured with the method proposed in [28]. Therefore, the overhead of a parallel region of the code could be defined as $T_n - T_s/n$, where T_n is the execution time using n processors and T_s is the execution time of the sequential version of the code. For the case with a mesh of 700 elements, the overhead using four threads represents 52% of the total run time, while using 16 threads, it represents 92% of the total run time. On the other hand, in the case with a mesh of 100,000 elements, the overhead using four and 16 threads supposes 13% and, respectively, 37% of the total run time.

3.2. Test 2: Filling of Floodplain Depressions

3.2.1. Case Description

The second test evaluates how the models reproduce the flooding of several depressions that are interconnected. The final distribution of the water is essential to predict the extension of a flood. The topography is a 2000 m × 2000 m domain with a grid of 4 × 4 ground depressions and a slight descending slope in the down-right direction as shown in Figure 6a. The domain is discretized using a structured grid with an element size of 20 m, which gives a numerical grid of 10,000 elements. The Manning's coefficient is constant and equal to 0.03 s/m$^{1/3}$ in the entire domain. The initial condition is a completely dry bed. The inlet boundary (red line in Figure 6a) is defined at the top left corner of the domain. The inlet hydrograph is shown in Figure 6b. The total physical time is 48 h. The model outputs to be evaluated are the time series of the water level at the four control points shown in Figure 6a.
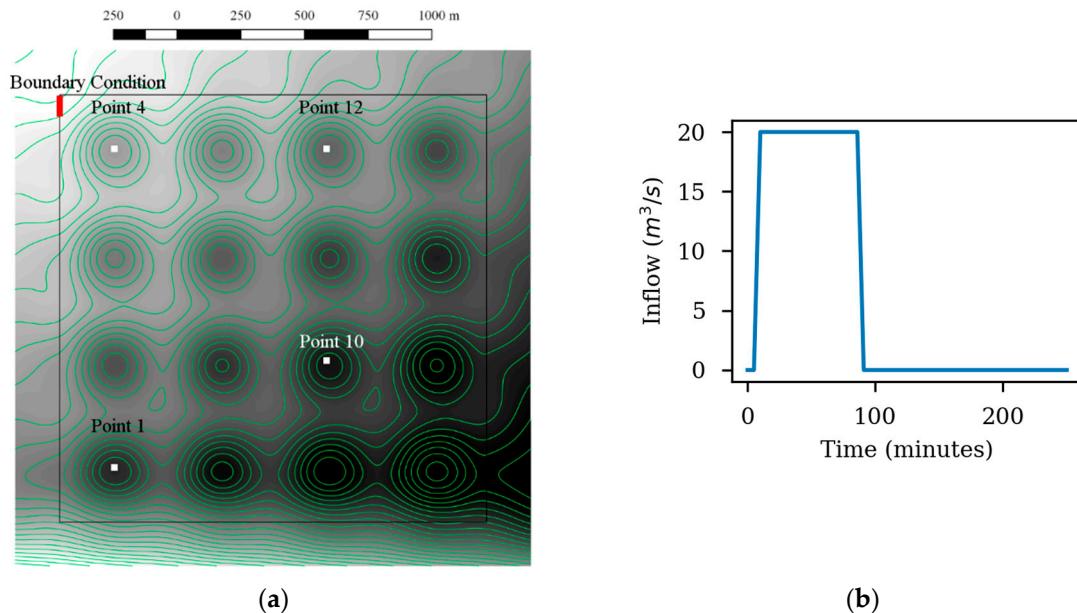
Figure 6. Setup of Test 2. (**a**) Shows the simulated domain. White areas represent higher elevations. The inlet boundary condition is marked in red. The measurements were performed in the points shown above; (**b**) Hydrograph used for the inlet boundary condition.

3.2.2. Results

In the original benchmark, the water level evolution was measured at 16 points; for this work and for the sake of clarity, more representative control points were selected. There are no noticeable differences between Iber and Iber+ (in both the CPU and the GPU implementations). Regarding the comparison with other software packages, some differences appear at Point 1 in the arrival time of the flood wave (Figure 7a). JFlow+ and InfoWorks ICM perform similarly, while TUFLOW FV and Iber+ predict a slightly faster propagation of the flood wave. In any case, the final water levels are almost the same for all models. At Point 4, all the models predict very similar results, with the exception of InfoWorks ICM, which shows slightly higher water levels (Figure 7b). At Point 10 (Figure 7c), the highest differences between the analyzed numerical models are shown. The final water level is similar but a bit lower in Iber+. The arrival times also differ, with Iber+ being faster but similar to TUFLOW FV and InfoWorks ICM. Finally, at Point 12 (Figure 7d), small differences in arrival time are shown, as previously reported for Point 1. In general, the models in which the flood arrives earlier also show higher final water levels.

Summarizing, in this case all the models produce very similar results. All of them predict the inundation of 11 out of 16 depressions. Nonetheless, there are some differences on the time of arrival of the water at some depressions. These differences were also noticed in [21] and were attributed to the weak flow between depressions. In these situations, the specific numerical schemes used to discretize the bed slope, the bed friction and the convective flux can have a significant impact on the model output.
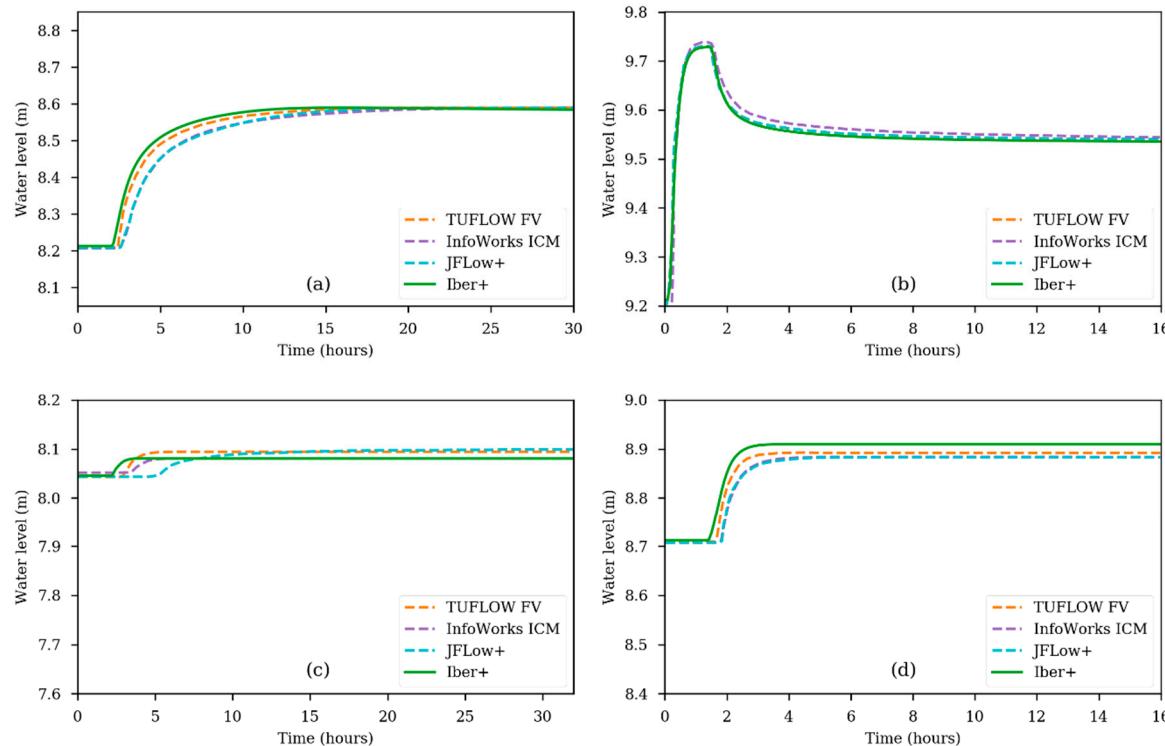
Figure 7. Water level time series for Test 2 obtained with different numerical models at Point 1 (**a**); Point 4 (**b**); Point 10 (**c**); Point 12 (**d**).

Table 3 shows the computational efficiency of the different models in this test case. Iber+ obtains a speedup of up to 25.7 over the single-threaded Iber. In this case, InfoWorks ICM and JFlow+ are slightly faster than Iber+. Although this case is more complex than Test 1, Iber+ is not able to get any improvement when scaling from eight to 16 threads (Figure 8). This means that for this case, the parallelism overhead is still significant when using 16 threads.

Table 3. Performance measurements for Test 2. Total run time of the simulation in seconds, average processing time per time step in milliseconds and the achieved speedups compared with Iber running in a single thread.

| Model	Hardware Configuration	Run Time (s)	Time per Step (ms)	Speedup vs. Iber 1 Thread
InfoWorks ICM	GPU	11	-	26.6
JFlow+	GPU	10	-	29.2
TUFLOW FV	12 Threads	26	-	11.2
Iber	1 Thread	292.1	1.61	1.0
	4 Threads	188.6	1.01	1.5
	16 Threads	170.6	0.90	1.7
Iber+	1 Thread	137.0	0.79	2.1
	4 Threads	50.8	0.29	5.7
	16 Threads	60.7	0.33	4.8
	GPU (GTX 480)	22.8	0.13	12.8
	GPU (GTX 1080)	11.4	0.06	25.7
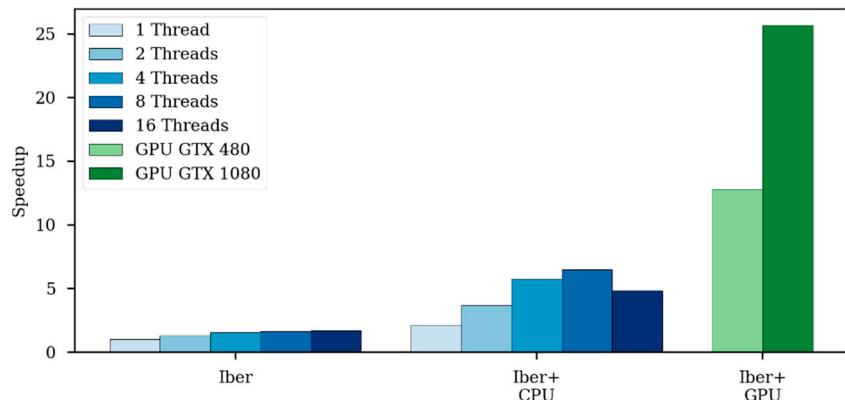
Figure 8. Speedups obtained comparing Iber running in one thread with different configurations in Test 2.

3.3. Test 5: Valley Flooding

3.3.1. Case Description

Test 5 consists of a dam break flow that originates a violent flood propagating through a river valley. This case is designed to check the capacity of the models to simulate large inundations. The spatial extension of this case is approximately 0.8 km × 17 km, as shown in Figure 9a. The spatial domain is discretized with an unstructured grid with a characteristic length of approximately 50 m, and a total number of mesh elements of 7753. The Manning's coefficient is set to 0.04 s/m$^{1/3}$. The initial condition is a fully dry bed. The inlet boundary condition is marked with a red line in Figure 9a, while the inlet hydrograph is shown in Figure 9b. The physical time in this case is 30 h. The time series of the water levels and velocities at the four control points shown in Figure 9a are used to evaluate and compare the different models.


(a)

(b)
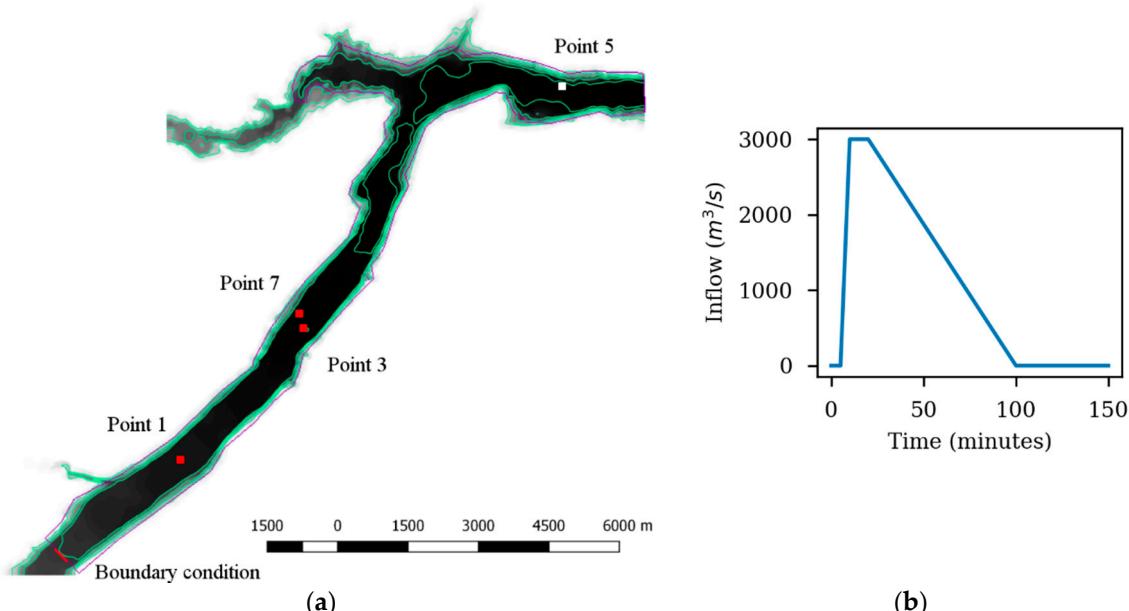
Figure 9. Setup of Test 5. (**a**) Shows the simulated domain. White areas represent higher elevations and black areas represent lower elevations. Light green lines are level curves. The inlet boundary condition is marked in red. The measurements were performed at the points shown above; (**b**) Hydrograph used as the upstream inlet boundary condition.

3.3.2. Results

The water levels given by all the models at Points 1, 3 and 7 are very similar (Figure 10). However, some differences can be found at the end of the valley (Point 5), with a delay in the order of 30 min on the time of arrival computed with the different models. Regarding the final water level at Point 5, Iber and Iber+ show values higher than the rest of the models, while InfoWorks ICM and JFlow+ show lower values than TUFLOW FV.
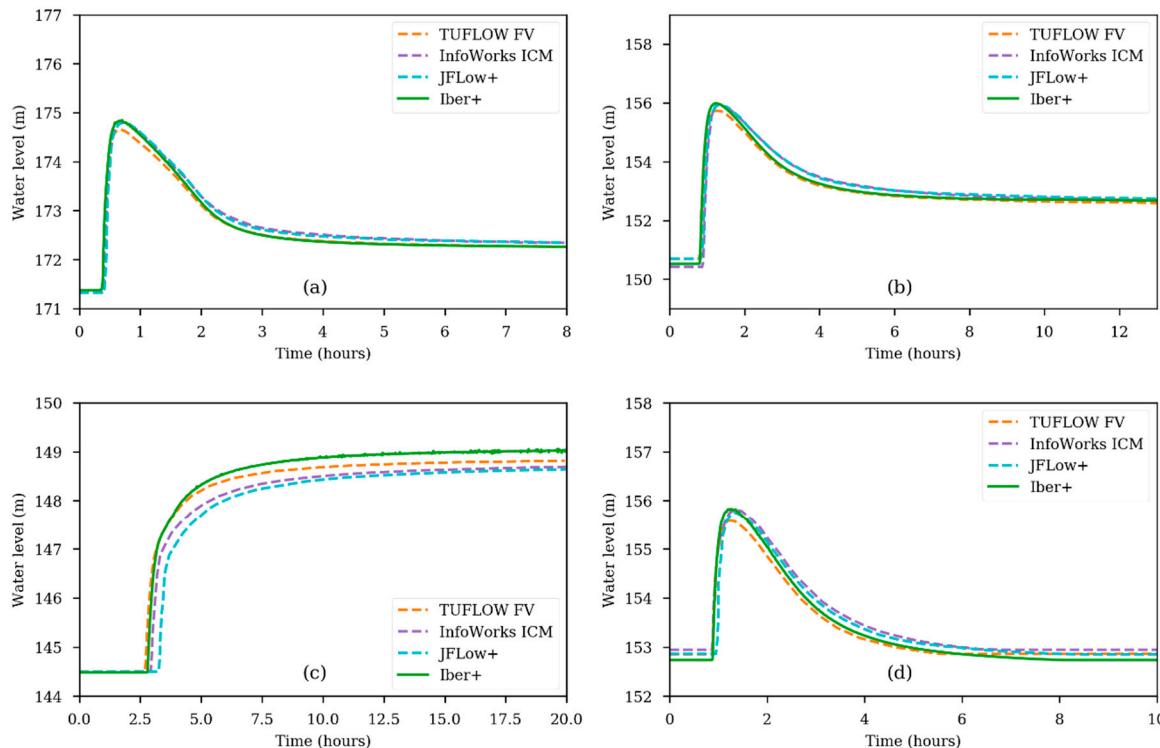

Figure 10. Time series of water level for Test 5 obtained with different numerical models at Point 1 (**a**); Point 3 (**b**); Point 5 (**c**); and Point 7 (**d**).

Figure 11 shows the velocity time series at three control points. Iber+ behaves very similar to TUFLOW FV, while InfoWorks ICM and JFlow+ show delays in the arrival times with respect to TUFLOW FV and Iber+ at Point 1 and Point 3.
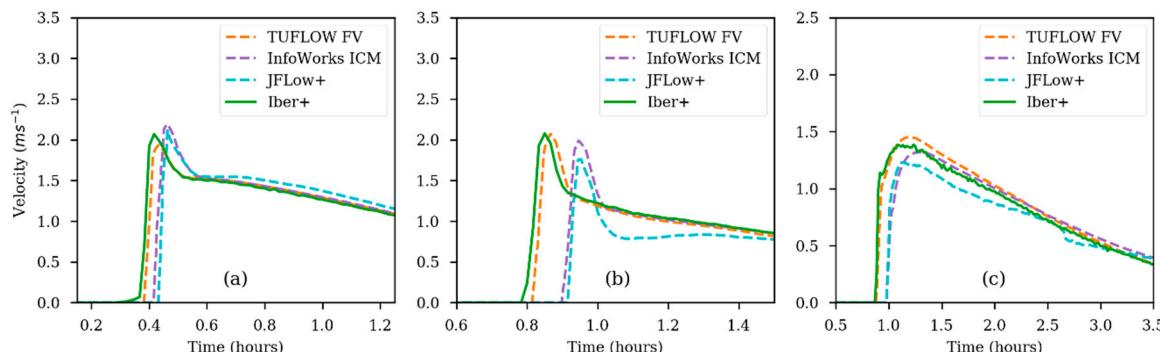

Figure 11. Time series of velocity for Test 5 obtained with different numerical models at Point 1 (**a**); Point 3 (**b**); and Point 7 (**c**).

Table 4 shows the computational efficiency of the analyzed models in this case. Iber+ reaches a speedup of 39.1 compared with Iber running in a single thread. In this case, Iber+ runs faster than InfoWorks ICM and JFlow+ using the GTX 1080, but is slower than InfoWorks ICM using the older

GTX 480 card. Figure 12 shows a similar behavior as previously observed in Figure 8. Using 16 threads results in worse run times than using just eight threads. This is due to the low number of elements in Test 2 and Test 5.

Table 4. Performance measurements for Test 5. Total run time of the simulation in seconds, average processing time per time step in milliseconds and the achieved speedups compared with Iber running in a single thread.

| Model	Hardware Configuration	Run Time (s)	Time per Step (ms)	Speedup vs. Iber 1 Thread
InfoWorks ICM	GPU	9	-	37.9
JFlow+	GPU	22	-	15.5
TUFLOW FV	12 Threads	67	-	5.1
Iber	1 Thread	340.8	3.02	1.0
	4 Threads	265.3	2.35	1.3
	16 Threads	243.5	2.16	1.4
Iber+	1 Thread	86.9	0.77	3.9
	4 Threads	37.4	0.33	6.4
	16 Threads	42.6	0.38	8.0
	GPU (GTX 480)	15.6	0.14	21.9
	GPU (GTX 1080)	8.7	0.08	39.1
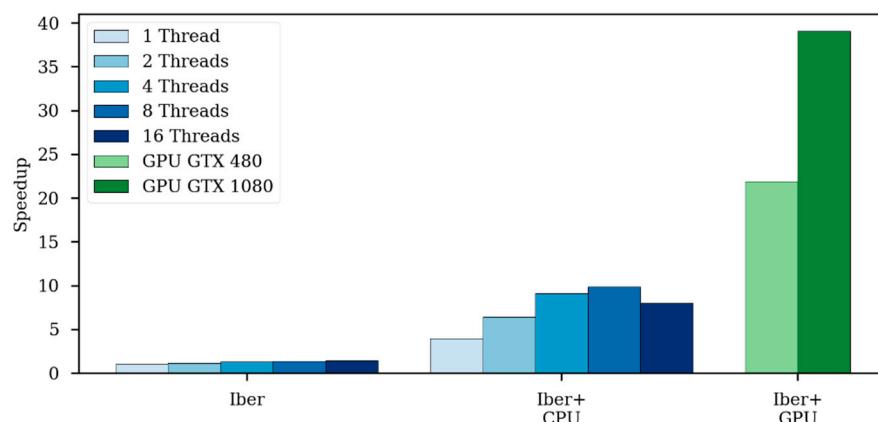
Figure 12. Speedups obtained comparing Iber running in one thread with different configurations in Test 5.

3.4. Test 8A: Urban Flood

3.4.1. Case Description

Test 8A simulates an inundation in an urban area of approximately 400 m × 960 m (Figure 13). The topography is defined from a DTM (digital terrain model) with a spatial resolution of 0.5 m. Buildings and vegetation are not included in the DTM and are ignored in the simulation. In this case, the flooding originates from two water sources. A short and intense rainfall event, with intensities up to 400 mm/h, is first defined for the entire domain, and followed by a point inflow hydrograph (blue dot in Figure 13). The hyetograph and the inlet hydrograph are shown in Figure 14a,b. The physical time of the simulation is 3 h. In order to compare the numerical models, the time series of the water levels are extracted at the four control points shown in Figure 13. The time series of velocity at Points 2 and 6 are also analyzed. In this case, the spatial domain is discretized with a structured mesh with a grid resolution of 2 m, which gives a mesh of 96,400 quadrilateral elements.
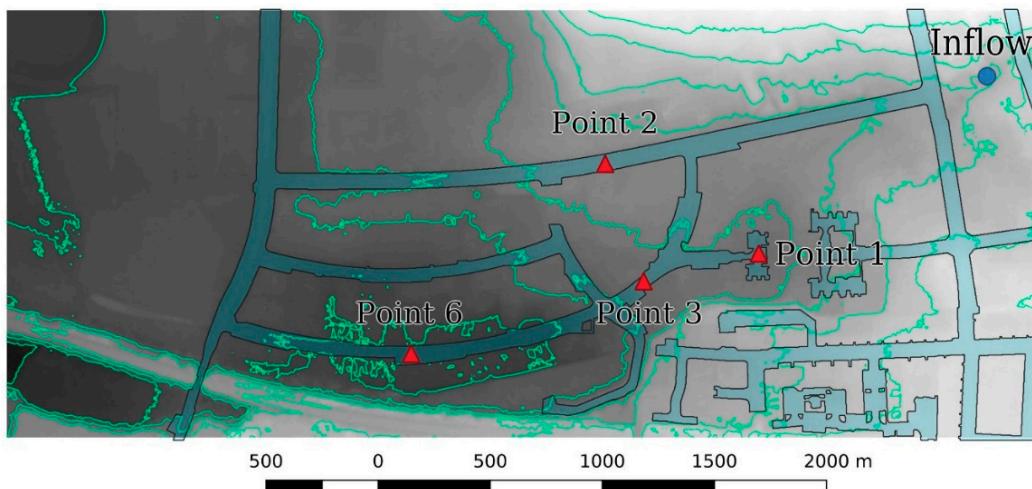
Figure 13. Setup of Test 8A. White areas represent higher elevations. Measurement points are marked in red. A point inflow is marked in blue. Streets are painted in a bluish tone and use a different Manning coefficient than the rest of the domain.


(**a**)	(**b**)
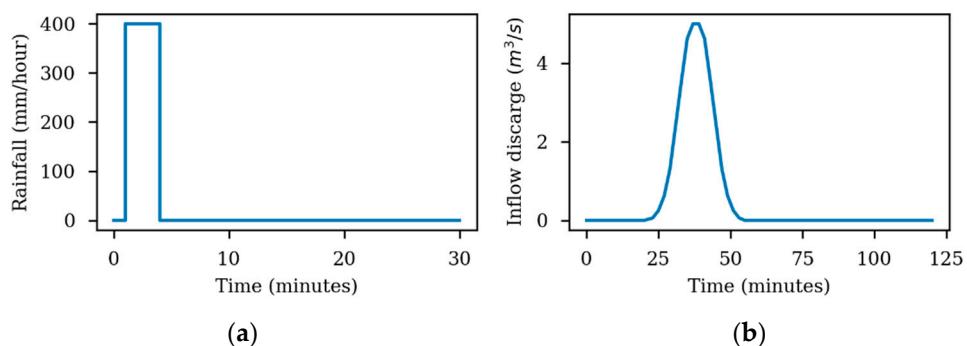
Figure 14. Water sources for Test 8A. (**a**) Hyetograph of the rain event in the simulation; (**b**) Hydrograph of the point inflow.

3.4.2. Results

The time series of the water level at Points 1, 2, 3 and 6 are shown in Figure 15. No significant differences were found between Iber and Iber+ CPU/GPU. At Point 1 (Figure 15a) and Point 2 (Figure 15b), the results obtained with Iber+ are very similar to those given by the rest of the models. However, the first peak, caused by the precipitation event, is more pronounced in Iber+ than in the rest of the numerical models. At Point 3 (Figure 15c), Iber+ provides similar results to TUFLOW FV. InfoWorks ICM and JFlow+ produce water levels slightly lower than those obtained using Iber+. At the last point (Point 6 in Figure 15d), Iber+ shows water levels between those provided by TUFLOW FV and InfoWorks ICM.
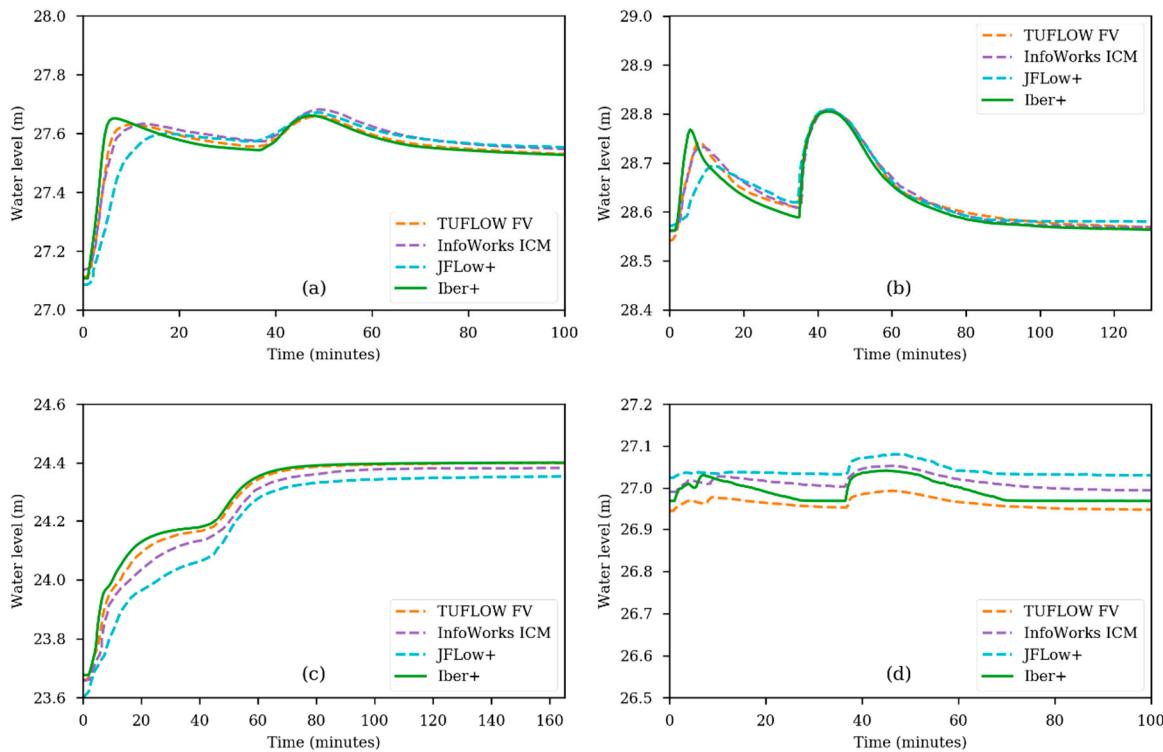
Figure 15. Time series of the water level for Test 8A obtained with different numerical models. Measurements for Point 1 (**a**); Point 2 (**b**); Point 3 (**c**); and Point 6 (**d**).

The velocity time series at Points 2 and 6 are shown in Figure 16. Velocities measured at Point 2 are similar for all the models, and the differences are in concordance with the differences seen in the water levels and arrival times for that point. However, at Point 2, with lower water depths, there are more differences among models.
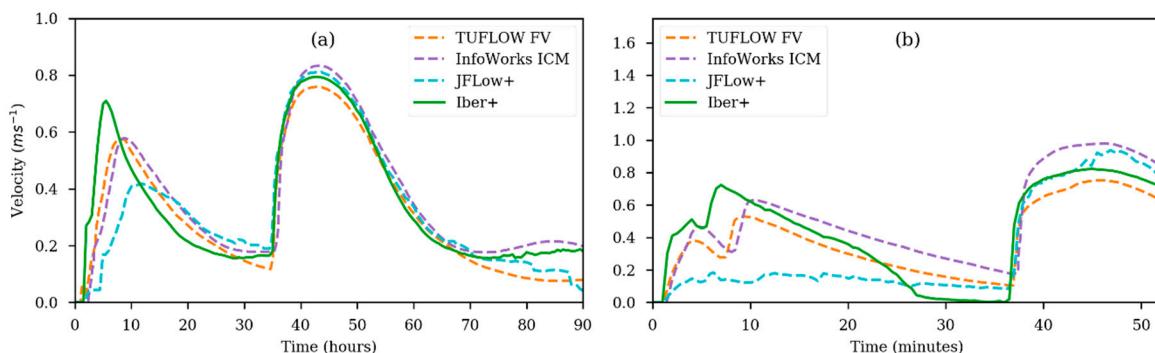

Figure 16. Time series of the velocities for Test 8A obtained with different numerical models. Measurements for Point 2 (**a**); and Point 6 (**b**).

Table 5 shows the computational efficiency of the different models in this test case. Since this case is more complex and has many more elements than the previous ones, it has a higher benefit from GPU implementation, running up to 94 times faster than Iber on a single CPU (Figure 17). Compared with other models from [21], in this case, the Iber+ GPU is slightly faster than InfoWorks and JFlow+ using a GPU with similar specifications (GTX 480).

Table 5. Performance measurements for Test 8A. Total run time of the simulation in seconds, average processing time per time step in milliseconds and the achieved speedups compared with Iber running in a single thread.

| Model	Hardware Configuration	Run Time (s)	Time per Step (ms)	Speedup vs. Iber 1 Thread
InfoWorks ICM	GPU	66	-	35.1
JFlow+	GPU	66	-	35.1
TUFLOW FV	12 Threads	410	-	5.7
Iber	1 Thread	2317.4	28.18	1.0
	4 Threads	1615.8	19.62	1.4
	16 Threads	1360.6	16.51	1.7
Iber+	1 Thread	850.1	10.40	2.7
	4 Threads	313.8	3.80	7.4
	16 Threads	150.9	2.01	15.4
	GPU (GTX 480)	51.1	0.61	45.4
	GPU (GTX 1080)	24.6	0.28	94.3
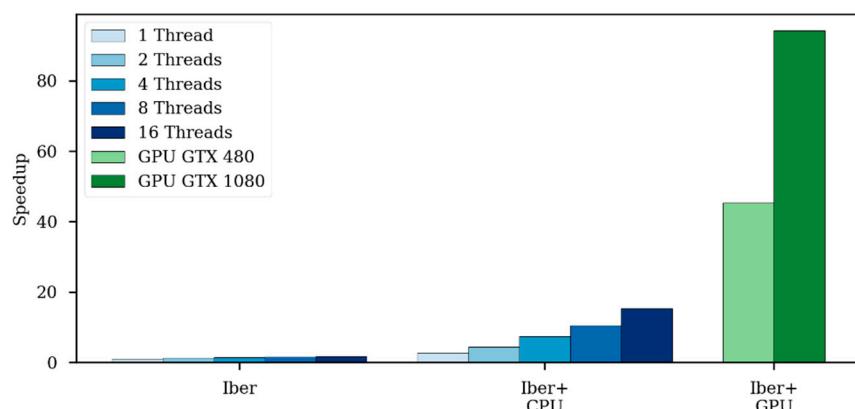
Figure 17. Speedups obtained comparing Iber running in one thread with different configurations in Test 8A.

4. Application Case: Surface Runoff Generation in a Mountain Basin during a Storm Event

In this section, a real fast flood event due to heavy precipitation in a mountain basin was used to verify the computational performance of Iber+ in large scale simulations. The discharge computed with Iber+ is compared with observed data registered during the real event at a gauging station located at the basin outlet. The run times of Iber and Iber+ are also compared. The simulations were run with a CPU Intel Core i7-920 and with a GPU NVIDIA GeForce GTX 1080.

4.1. Case Description

On 19–21 October 2012, an extreme rainfall event was observed in the central western Pyrenees, affecting mainly the headwater catchments of the Aragón river [29,30]. This basin is located in north-eastern Spain, near the Mediterranean Sea and in the Pyrenees mountain range, and is prone to flooding during autumn. Precipitation events of more than 200 mm in just a few hours have been reported by various authors in that area [18,19]. Figure 18 shows the location of the Upper Aragón basin, which has an area of around 240 km^2 and altitudes ranging from approximately 600 m to 3000 m. Figure 18 also shows the location of five rain gauges (Candanchu, Hecho, Formigal, Canfranc and Albarun) operated by the Ebro River Basin Authority (CHEbro), and the hyetographs recorded by these gauges on 19–21 October 2012. Three of these rain gauges (Candanchu, Canfranc and Albarun) are located inside the study catchment. The observed hyetographs show peak values close to 25 mm/h on 19 October. The rain gauges at Hecho and Canfranc registered peak values of

more than 30 mm/h during the first hours of 19 October. The rain gauge at Formigal, which is the eastern-most gauge, recorded peak values smaller than 20 mm/h. The accumulated precipitation of the rain gauges ranged from 200 mm at Formigal to near 250 mm at Candanchu. These values are more than three times higher than the monthly averaged accumulated precipitation of October over the period of 1981–2010 (60 mm) registered at the AEMET (Agencia Estatal de Meteorología) rain gauge located at Huesca airport (available online at http://www.aemet.es/es/serviciosclimaticos/ datosclimatologicos/). The discharge of the Aragón river registered at the town of Jaca was used for comparison with the numerical results.
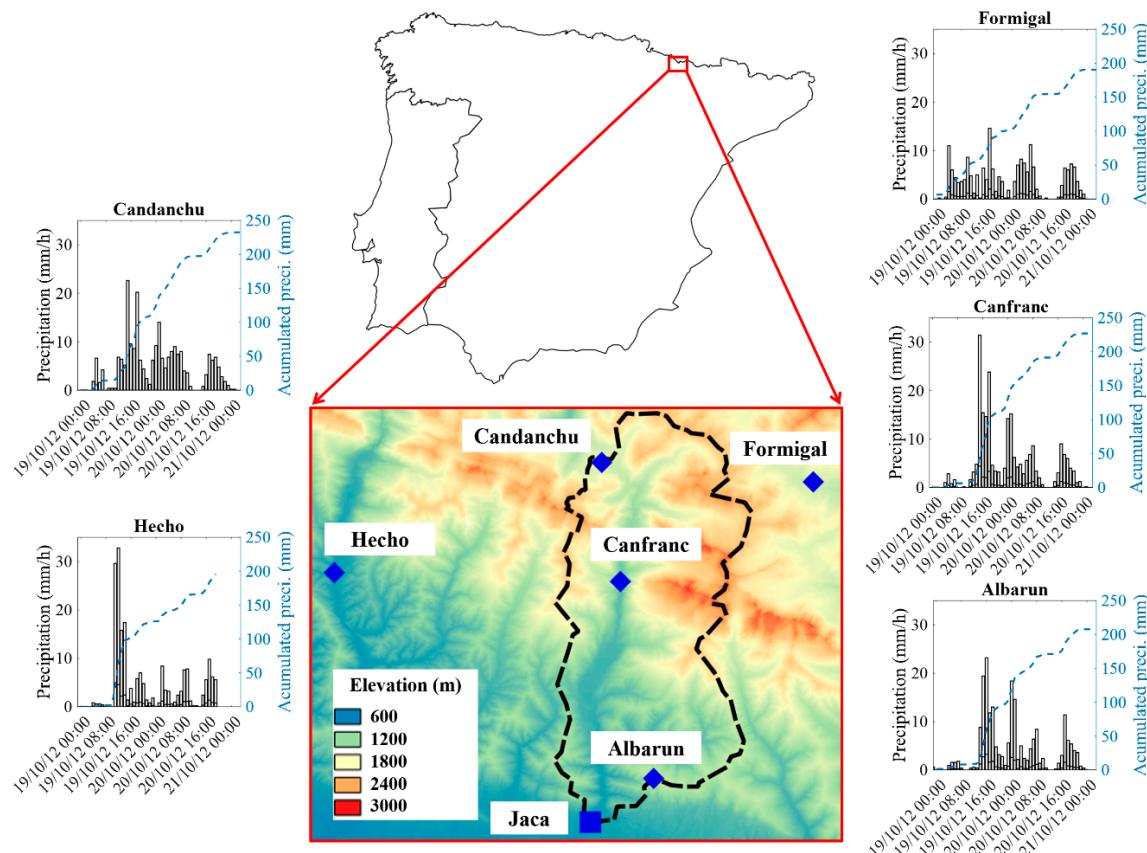

Figure 18. Location of the upper basin of the Aragon river (enclosed by a black dashed line) in the Iberian Peninsula. Blue diamonds represent the location of the available rain gauges. The blue square represents the gauging station at Jaca. The hyetographs (bar graphs) and the accumulated precipitations (blue dashed lines) recorded during the event are also shown.

A series of simulations were carried out with Iber+ in order to compute the discharge at the gauge station of Jaca during the first 24 h of this extreme event. The numerical domain was discretized using triangular elements with a characteristic length ranging from 5 m near the riverbed, to 40 m. This led to a mesh of 459,000 elements. The wet-dry threshold was set equal to 0.001 m. The hydrodynamic equations were discretized using the DHD scheme [6], which is especially designed for solving the shallow water equations in rainfall–runoff applications. The physical time of the simulation was 24 h. The hyetographs registered at the five rain gauges (Figure 18) were spatially interpolated to define precipitation rasters that were used as the rainfall input in the numerical model.

The Manning coefficient was also provided to Iber+ as a raster file. The raster file of the Manning coefficients was obtained starting from shape files of land uses of the SIOSE (Sistema de Información sobre Ocupación del Suelo de España) downloaded from the web of the CNIG (Centro Nacional de Información Geográfica). These shape files divide the entire area of study into small parcels ($j = 1$, ..., N), each of them with a combined land use, which was labelled according the rules of SIOSE by

the type of parcel, individual land uses defined in each parcel and the ratio of area of each individual land use to the total area of the parcel. Each individual land use is associated to a Manning coefficient obtained from [6]. The combined Manning coefficient of each parcel j was obtained using:

$$C_{comb}^{j} = \sum_{i=1}^{M_j} Per_i^j C_i \tag{6}$$

where Per_i^j is the ratio of land use i in parcel j ($\in [0,1]$), C_i is the Manning coefficient associated to the land use i, M_j is the number of land uses in the parcel j. Details of the methodology to compute the Manning coefficients can be found in [31].

Infiltration losses were defined using the curve number method of the SCS (Soil Conservation Service). The curve number (*CN*) was computed with the formulation proposed in [32]. First, the maximum surface retention (*S*) was obtained as:

$$S = \frac{P^2}{R_u} - P \tag{7}$$

where R_u is the mean value of surface runoff and P the precipitation depth both provided as raster files. From the maximum surface retention, the curve number was estimated as:

$$CN = \frac{25,400}{S + 254} \tag{8}$$

There are multiple sources of uncertainty in numerical models. Apart from the limitations in the model formulation, especially those related to the mathematical discretization of the equations, the models are very sensitive to the initial boundary conditions, as well as to the input parameters. In the particular case of hydraulic models like Iber, the parameterization of the Manning coefficients and *CN* are of utmost importance in order to determine the timing and extent of flooding events. In the present case, the reference simulation described above and a set of 16 additional simulations were considered. Among them, eight simulations were defined by the reference *CN* and varying Manning coefficients ($\pm15\%$, $\pm35\%$, $\pm55\%$ and $\pm75\%$) and eight simulations were defined by the reference Manning coefficients and varying curve number ($\pm3\%$, $\pm6\%$, $\pm9\%$ and $\pm12\%$).

4.2. Results

Figure 19 shows the averaged hydrograph computed with Iber+ and the experimental values observed at Jaca. The shaded area corresponds to the mean value plus/minus two times the standard deviation of the numerical simulations. The average peak discharge obtained with Iber+ is slightly higher (~25%) than the experimental one (349.5 m^3 s^{-1} and 279.5 m^3 s^{-1} respectively). The difference between both values (70.0 m^3 s^{-1}) is considerably lower than twice the standard deviation (shaded area). The timing of the numerical peak (20.0 h) is slightly delayed with respect to the observed data (19.5 h). The total volume of water also shows slight differences (~2%) between the experimental (4.6 Hm^3) and numerical (4.7 Hm^3) results. Therefore, the model reproduces the observed data with an acceptable accuracy, especially considering the uncertainty inherent to hydrological simulations.
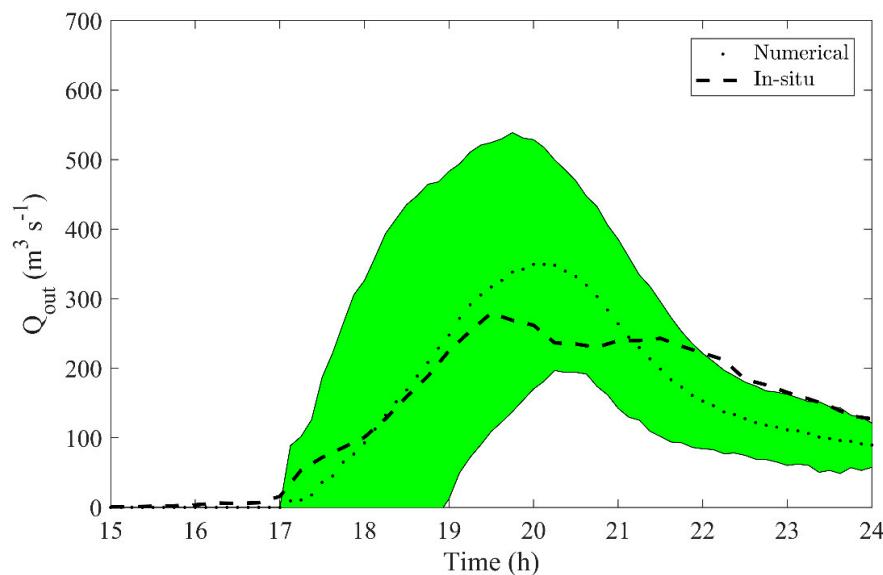
Figure 19. Numerical (dotted line) and in-situ (dashed line) outflow of the basin of the Alto Aragón river obtained at Jaca. The green shaded area represents twice the standard deviation of the numerical results. Q_{out} refers to the outflow of the basin.

In addition to the visual comparison derived from Figure 19, three non-dimensional quantitative statistics and two dimensional techniques have been used to evaluate, in an objective way, the accuracy of the numerical results compared to the field data. The non-dimensional quantitative statistics used are the Nash–Sutcliffe efficiency (*NSE*), percent bias (*PBIAS*) and the ratio of the root mean square error to the standard deviation of measured data (*RSR*). The used dimensional techniques are the correlation coefficient (*r*) and the coefficient of determination (R^2). The Nash–Shutcliffe coefficient (*NSE*) [33] defines the relation of the residual variance to the field data variance. It can range from $-\infty$ to 1 and its optimal value is 1. *NSE* was used according to the recommendation of [34] and its widespread use in the hydrology field [35,36]. Percent bias (*PBIAS*) [37] measures the tendency of the numerical results to be larger (*PBIAS* > 0) or smaller (*PBIAS* < 0) than the reference data and is expressed as a percentage. Therefore, the optimal value of this parameter is 0. This parameter is also recommended by [34] and, according to [37], it can identify numerical models with poor accuracy. The ratio of the root mean square error to the standard deviation of measured data (*RSR*) [38] is defined as the ratio of root mean square error (*RMSE*) to the standard deviation of the field data. It can range from 1 to a large positive number. In the former case, the *RMSE* would be equal to the standard deviation of the field data and in the latter case, the *RMSE* would be greater than the standard deviation of the field data.

Following [39], the range of values of the non-dimensional statistics is shown in Table 6. Note that these values are defined for monthly time steps. Usually as the time step decreases, the ratings are less stringent.

Table 6. Performance rating for monthly time step. *NSE* = Nash–Sutcliffe efficiency. *PBIAS* = percent bias. *RSR* = the ratio of the root mean square error to the standard deviation of measured data.

Rating	*NSE*	*PBIAS* (%)	*RSR*
Very good	$0.75 < NSE \leq 1.00$	$PBIAS < \pm 10$	$0.0 \leq RSR \leq 0.5$
Good	$0.65 < NSE \leq 0.75$	$\pm 10 \leq PBIAS < \pm 15$	$0.5 < RSR \leq 0.6$
Satisfactory	$0.50 < NSE \leq 0.65$	$\pm 15 \leq PBIAS < \pm 25$	$0.6 < RSR \leq 0.7$
Unsatisfactory	$NSE \leq 0.50$	$\pm 25 \leq PBIAS$	$0.7 < RSR$

Table 7 shows the values of the statistics parameters obtained for the case under study. The values of *NSE*, *PBIAS* and *RSR* clearly indicate, according to the reference values of Table 6, that the accuracy

of the numerical results is very good. In addition, the values of r and R^2 show the high accuracy of the numerical model.

Table 7. Performance rating calculated for the mean signal shown in Figure 19. r = correlation coefficient. R^2 = coefficient of determination.

NSE	PBIAS	RSR	r	R^2
0.89	5.72	0.34	1.00	0.90

Apart from the accuracy of the results, this case study highlights the need for computationally efficient models. The total runtime needed for the 17 numerical simulations was approximately two hours and twenty minutes (approximately eight minutes per simulation). The time needed to run the same case with the standard Iber ranged from 14 h 25 min with eight threads and 18 h 9 min with one thread. We should note that weather forecast systems can predict extreme rain events only a few hours in advance, which reduces the response time and highlights the need for fast and reliable hydraulic models to analyze the areas at risk of flooding. Iber+ can then be considered as a suitable tool to predict these type of extreme events.

5. Conclusions

We have presented a new parallel implementation of the 2D shallow water model Iber. The new implementation, named Iber+, takes advantage of different parallelization strategies both on CPUs and GPUs to speed up the computations, while keeping the same accuracy as the original model. The computational efficiency of the new code was analysed using four different benchmarks proposed by the Environmental Agency of the United Kingdom. The results obtained with Iber+ on those benchmarks was compared to those given by other hydraulic models, showing a similar level of accuracy.

The main limitation of the standard implementation of Iber is its lack of computational efficiency in large problems. This limitation is overcome in the new parallel implementation, where a speedup of two orders of magnitude, compared to the standard version, can be obtained depending on the case under study. This speedup is especially relevant on the GPU version of the code when modelling cases with a high workload.

Iber+ was applied to a real case, an extreme flash flood that took place in the Spanish Pyrenees in October 2012. Iber+ was able to simulate 24 h of physical time in less than 10 min using a numerical mesh of almost half a million elements. The same case run with the standard version needs more than 15 h of CPU time. This improvement in computational time is especially important to extend the potential application of the model to early warning systems, since extreme precipitation events can only be forecasted a few hours in advance by the meteorological agencies. Thus, the use of fast and accurate numerical tools such as Iber+ is crucial for a real-time analysis and deployment of flood protection measures by decision makers.

Although Iber+ offers a huge performance leap compared to the standard version, there is still room for further speedups by the use of ordering algorithms for data locality, the improvement in the efficiency to skip dry elements from the computation, or the implementation of asynchronous tasks in GPU programming.

Author Contributions: O.G.-F., J.G.-C., M.G.-G. and J.M.D. conceived the study; O.G.-F. developed the software; L.C., J.M.D. and A.F. supervised the code development; O.G.-F. and J.G.-C. performed the experiments; O.G.-F., J.G.-C. and M.G.-G. analysed the results; O.G.-F., J.G.-C., L.C., M.G.-G. and A.F. wrote the manuscript.

Abbreviations

AEMET	Agencia Estatal de Meteorología.
API	Application Programming Interface
AVX	Advanced Vector Extensions
CHEbro	Confederación Hidrográfica del Ebro.
CHMS	Confederación Hidrográfica Miño Sil.
CNIG	Centro Nacional de Información Geográfica
CPU	Central Processing Unit
CUDA	Compute Unified Device Architecture
CUB	CUDA Unbound
DHD	Decoupled Hydrological Discretization
DTM	Digital Terrain Model
GPGPU	General Processing Graphics Processing Unit
GPU	Graphics Processing Unit
I/O	Input/Output
MPI	Message Passing Interface
OpenMP	Open Multi-Processing
PCI	Peripheral Component Interconnect
RMSD	Root Mean Square Deviation
SCS	Soil Conservation Service
SIMD	Single Instruction Multiple Data
SIMT	Single Instruction Multiple Thread
SIOSE	Sistema de Información sobre Ocupación del Suelo de España.
SM	Streaming Multiprocessors
SWE	Shallow Water Equations
SSE	Streaming SIMD Extensions

References

1. Houghton, J.T.; Ding, Y.; Griggs, D.J.; Noguer, M.; van der Linden, P.J.; Dai, X.; Maskell, K.; Johnson, C.A. *Climate Change 2001: The Scientific Basis*; Cambridge University Press: Cambridge, UK, 2001; ISBN 9780521014953.
2. Hov, Ø.; Cubasch, U.; Fischer, E.; Höppe, P.; Iversen, T. *Extreme Weather Events in Europe: Preparing for Climate Change Adaptation*; Norwegian Meteorological Institute: Oslo, Norway, 2013.
3. Bladé, E.; Cea, L.; Corestein, G.; Escolano, E.; Puertas, J.; Vázquez-Cendón, E.; Dolz, J.; Coll, A. Iber: Herramienta de simulación numérica del flujo en ríos. *Rev. Int. Métodos Numéricos para Cálculo y Diseño en Ing.* **2014**, *30*, 1–10. [CrossRef]
4. Cea, L.; Bermudez, M.; Puertas, J.; Blade, E.; Corestein, G.; Escolano, E.; Conde, A.; Bockelmann-Evans, B.; Ahmadian, R. IberWQ: New simulation tool for 2D water quality modelling in rivers and shallow estuaries. *J. Hydroinform.* **2016**. [CrossRef]
5. Roe, P.L. Discrete Models for the Numerical Analysis of Time-Dependent Multidimensional Gas Dynamics. In *Upwind and High-Resolution Schemes*; Springer: Berlin/Heidelberg, Germany, 1986; pp. 451–469.
6. Cea, L.; Bladé, E. A simple and efficient unstructured finite volume scheme for solving the shallow water equations in overland flow applications. *Water Resour. Res.* **2015**, *51*, 5464–5486. [CrossRef]
7. Bladé Castellet, E.; Cea, L.; Corestein, G. Numerical modelling of river inundations. *Ingeniería del Agua* **2014**, *18*, 68. [CrossRef]

8. Bodoque, J.M.; Amérigo, M.; Díez-Herrero, A.; García, J.A.; Cortés, B.; Ballesteros-Cánovas, J.A.; Olcina, J. Improvement of resilience of urban areas by integrating social perception in flash-flood risk management. *J. Hydrol.* **2016**, *541*, 665–676. [CrossRef]

9. Cea, L.; French, J.R. Bathymetric error estimation for the calibration and validation of estuarine hydrodynamic models. *Estuar. Coast. Shelf Sci.* **2012**, *100*, 124–132. [CrossRef]

10. Fraga, I.; Cea, L.; Puertas, J.; Suárez, J.; Jiménez, V.; Jácome, A. Global Sensitivity and GLUE-Based Uncertainty Analysis of a 2D-1D Dual Urban Drainage Model. *J. Hydrol. Eng.* **2016**, *21*, 04016004. [CrossRef]

11. Cea, L.; Legout, C.; Grangeon, T.; Nord, G. Impact of model simplifications on soil erosion predictions: Application of the GLUE methodology to a distributed event-based model at the hillslope scale. *Hydrol. Process.* **2016**. [CrossRef]

12. Sopelana, J.; Cea, L.; Ruano, S. A continuous simulation approach for the estimation of extreme flood inundation in coastal river reaches affected by meso- and macrotides. *Nat. Hazards* **2018**, *93*, 1337–1358. [CrossRef]

13. Crespo, A.J.C.; Domínguez, J.M.; Rogers, B.D.; Gómez-Gesteira, M.; Longshaw, S.; Canelas, R.; Vacondio, R.; Barreiro, A.; García-Feal, O. DualSPHysics: Open-source parallel CFD solver based on Smoothed Particle Hydrodynamics (SPH). *Comput. Phys. Commun.* **2015**, *187*, 204–216. [CrossRef]

14. Crossley, A.; Lamb, R.; Waller, S. Fast Solution of the Shallow Water Equations Using GPU Technology. 2010. Available online: http://www.jflow.co.uk/sites/default/files/Crossley%20Lamb%20Waller%20-%20BHS%202010.pdf (accessed on 9 October 2018).

15. Vacondio, R.; Dal Palù, A.; Mignosa, P. GPU-enhanced finite volume shallow water solver for fast flood simulations. *Environ. Model. Softw.* **2014**. [CrossRef]

16. Lacasta, A.; Morales-Hernández, M.; Murillo, J.; García-Navarro, P. An optimized GPU implementation of a 2D free surface simulation model on unstructured meshes. *Adv. Eng. Softw.* **2014**. [CrossRef]

17. Liu, Q.; Qin, Y.; Li, G. Fast Simulation of Large-Scale Floods Based on GPU Parallel Computing. *Water* **2018**, *10*, 589. [CrossRef]

18. NVIDIA Corporation. CUDA C Programming Guide. Available online: https://docs.nvidia.com/cuda/pdf/CUDA_C_Programming_Guide.pdf (accessed on 25 September 2018).

19. Merry, B. A Performance Comparison of Sort and Scan Libraries for GPUs. *Parallel Process. Lett.* **2015**, *25*, 1550007. [CrossRef]

20. NVIDIA Corporation. GP100 Pascal Whitepaper. Available online: https://images.nvidia.com/content/pdf/tesla/whitepaper/pascal-architecture-whitepaper.pdf (accessed on 9 October 2018).

21. Neelz, S.; Pender, G. *Delivering Benefits through Evidence-Benchmarking of 2D Hydraulic Modelling Packages*; Environment Agency: Bristol, UK, 2010; ISBN 9781849111904.

22. TUFLOW FV Science Manual. Available online: https://www.tuflow.com/Download/TUFLOW_FV/Manual/FV_Science_Manual_2013.pdf (accessed on 10 September 2018).

23. InfoWorks ICM Product Information—Overview. Available online: http://www.innovyze.com/products/infoworks_icm/ (accessed on 10 September 2018).

24. Néelz, S.; Pender, G.; Britain, G. *Desktop Review of 2D Hydraulic Modelling Packages*; Environment Agency: Bristol, UK, 2009; ISBN 9781849110792.

25. Anzt, H.; Hahn, T.; Heuveline, V. *GPU Accelerated Scientific Computing: Evaluation of the NVIDIA Fermi Architecture*; Elementary Kernels and Linear Solvers; EMCL: Heidelberg, Germany, 2010.

26. Gregg, C.; Hazelwood, K. Where is the data? Why you cannot debate CPU vs. GPU performance without the answer. In Proceedings of the ISPASS 2011—IEEE International Symposium on Performance Analysis of Systems and Software, Austin, TX, USA, 10–12 April 2011.

27. Courant, R.; Friedrichs, K.; Lewy, H. Über die partiellen Differenzengleichungen der mathematischen Physik. *Math. Ann.* **1928**. [CrossRef]

28. Bull, J.M. Measuring Synchronisation and Scheduling Overheads in OpenMP. In Proceedings of the First European Workshop on OpenMP, 1999. Available online: http://cc.jlu.edu.cn/Download/4c09bb4a-8dae-4083-acd3-052c4eb8ad61.pdf (accessed on 15 October 2018).

29. Naverac, V.; Ferrer, D.; Curiel, P.; Zueco, S.D.; Gil, F.E.; Hidalgo, J.C.G.; García, D.G.; de Matauco, A.I.G.; Albero, C.M.; Mur, D.M.; et al. Sobre las Precipitaciones de Octubre de 2012 en el Pirineo Aragonés, su Respuesta Hidrológica y la Gestión de Riesgos. 2012. Available online: https://dialnet.unirioja.es/descarga/articulo/4138905.pdf (accessed on 15 October 2018).

30. Serrano-Muela, M.P.P.; Nadal-Romero, E.; Lana-Renault, N.; González-Hidalgo, J.C.C.; López-Moreno, J.I.I.; Beguería, S.; Sanjuan, Y.; García-Ruiz, J.M.M. An exceptional rainfall event in the central western pyrenees: Spatial patterns in discharge and impact. *Land Degrad. Dev.* **2015**. [CrossRef]

31. González-Cao, J.; García-Feal, O.; Crespo, A.J.C.; Gómez-Gesteira, M.; Cea, L. Predicción de Inundaciones Originadas por Precipitaciones Extremas Mediante el Módulo Hidrológico de Iber. 2017. Available online: http://geama.org/jia2017/wp-content/uploads/ponencias/tema_B/b1.pdf (accessed on 15 October 2018).

32. Patil, J.P.P.; Sarangi, A.; Singh, A.K.K.; Ahmad, T. Evaluation of modified CN methods for watershed runoff estimation using a GIS-based interface. *Biosyst. Eng.* **2008**. [CrossRef]

33. Nash, J.E.; Sutcliffe, J.V. River flow forecasting through conceptual models part I—A discussion of principles. *J. Hydrol.* **1970**, *10*, 282–290. [CrossRef]

34. American Society of Civil Engineers. Criteria for Evaluation of Watershed Models. *J. Irrig. Drain. Eng.* **1993**, *119*, 429–442. [CrossRef]

35. Servat, E.; Dezetter, A. Selection of calibration objective fonctions in the context of rainfall-ronoff modelling in a Sudanese savannah area. *Hydrol. Sci. J.* **1991**, *36*, 307–330. [CrossRef]

36. Legates, D.R.; McCabe, G.J. Evaluating the use of "goodness-of-fit" measures in hydrologic and hydroclimatic model validation. *Water Resour. Res.* **1999**. [CrossRef]

37. Gupta, H.V.; Sorooshian, S.; Yapo, P.O. Status of Automatic Calibration for Hydrologic Models: Comparison with Multilevel Expert Calibration. *J. Hydrol. Eng.* **1999**. [CrossRef]

38. Singh, J.; Knapp, H.V.; Arnold, J.G.; Demissie, M. Hydrological modeling of the Iroquois River watershed using HSPF and SWAT. *J. Am. Water Resour. Assoc.* **2005**. [CrossRef]

39. Moriasi, D.N.; Arnold, J.G.; Van Liew, M.W.; Bingner, R.L.; Harmel, R.D.; Veith, T.L. Model Evaluation Guidelines for Systematic Quantification of Accuracy in Watershed Simulations. *Trans. ASABE* **2007**, *50*, 885–900. [CrossRef]

5.2. Towards an automatic early warning system of flood hazards based on precipitation forecast: the case of the Miño River (NW Spain)

Información de la revista	
Revista	Nat. Hazards Earth Syst. Sci.
eISSN	1684-9981
País	Alemania
Editorial	Copernicus
Factor de impacto	3,102 (2019)
Cuartil	Q2

5.2 Towards an automatic early warning system of flood hazards based on precipitation forecast: the case of the Miño River (NW Spain)

63

Natural Hazards
and Earth System
Sciences

Towards an automatic early warning system of flood hazards based on precipitation forecast: the case of the Miño River (NW Spain)

José González-Cao, Orlando García-Feal, Diego Fernández-Nóvoa, José Manuel Domínguez-Alonso, and Moncho Gómez-Gesteira

Environmental Physics Laboratory (EPhysLab), CIM-UVIGO, Universidade de Vigo, Ourense, Spain

Correspondence: Diego Fernández-Nóvoa (diefernandez@uvigo.es)

Abstract. An early warning system for flood prediction based on precipitation forecast is presented. The system uses rainfall forecast provided by MeteoGalicia in combination with a hydrologic (Hydrologic Modeling System, HEC-HMS) and a hydraulic (Iber+) model. The upper reach of the Miño River and the city of Lugo (NW Spain) are used as a study area. Starting from rainfall forecast, HEC-HMS calculates the streamflow and Iber+ is automatically executed for some previously defined risk areas when a certain threshold is exceeded. The analysis based on historical extreme events shows that the system can provide accurate results in less than 1 h for a forecast horizon of 3 d and report an alert situation to decision makers.

1 Introduction

According to Noji (2000), floods are one of the most dangerous natural hazards in the world. Jonkman (2005) estimated that more than 100 000 deaths in the last century were caused by floods. From 1940 to 2018 the number of deaths related to flood events (8138) is only surpassed by lightning fatalities (9386) in the US (https://www.weather.gov/hazstat/, last access: September 2019). Furthermore, the effect of climate change will increase the number of flood events and their negative impact on people and properties (Dankers and Feyen, 2008; Alfieri et al., 2017). Therefore, the ability to predict these extreme events and prevent their consequences is a challenge for the scientific community worldwide.

In this context early warning systems (EWSs) play a key role. UNISDR (2009) defines early warning systems as the set of capacities needed to generate and disseminate timely and meaningful warning information to enable individuals, communities and organisations threatened by a hazard to prepare and to act appropriately and in sufficient time to reduce the possibility of harm or loss.

A complete EWS is divided in four steps: (1) risk knowledge, (2) monitoring, forecasting and warning, (3) communication of an early warning system, and (4) response capability (UN, 2006). The first two steps are related to the field of physical sciences while the two last steps are associated with social science aspects. There are several works related to the impact of early warning systems on the prevention of floods. Baudoin et al. (2014) and UNISDR (2015) show some interesting examples on how early warning systems can save lives and reduce the damage to people. Borga et al. (2011) developed an early warning system methodology for flash floods in Europe through the HYDRATE project. The authors enhanced the capability of flash flood forecasting in ungauged basins by exploiting the extended availability of flash flood data and the improved process understanding. Alfieri et al. (2012) analysed several early warning systems applied to detect surface water flooding, flash floods, debris flows, landslides induced by extreme rainfall events, a river and coastal floods. The authors proposed several tasks to palliate the main drawbacks of some of these systems. Also, Hossain et al. (2014) developed a system to measure the water depth of the river at the "Valley of Death" and Cools et al. (2012) developed an early warning system to detect flash floods in the Sinai Peninsula, both based on a satellite-based forecast system. In Europe a very interesting example of an early warning system is the EWS

applied to the region of Flanders (Schelfaut et al., 2011 and CIW, 2011). In this work, the different steps are analysed under the FREEMAN project (Flood Resilience Enhancement and Management). The European Flood Awareness System (EFAS) is also another example of an EWS developed under the sponsorship of the European Commission. This system provides daily streamflow forecasts for Europe starting from up to 10 d weather forecasts (medium-term forecast). More details of this model can be shown in Thielen et al. (2009), Pappenberger et al. (2011), Cloke et al. (2013) and Alfieri et al. (2014). Using this model Dottori et al. (2017) develop a methodology to adapt EFAS to real-time forecasting. Demerit et al. (2013) analyse the problems derived from the use of the early warning system for medium- and long-term flood forecasts, mainly the dissemination of the information to people potentially affected by these events. They reveal that flood forecasters usually wait for the confirmation from local institutions (hydrologic confederations) instead of acting following the information provided by the early warning systems. These local systems are focused on short-term forecasts (0 to 48 h) that are more suitable to evacuation than damage mitigation. Some examples of these short-term local systems focused on river floods are the River Forecast Centers (https://water.weather.gov/ahps/rfc/rfc.php, last access: September 2019) in the United States of America or "Sistema de Ayuda a la Decisión" (http://www.chebro.es/contenido.visualizar.do?idContenido=12789&idMenu=2902, last access: November 2019) developed by the Hydrographic Confederation of the Ebro River (Spain). In Europe Meteoalarm (http://www.meteoalarm.eu/?lang=en_UK, last access: September 2019) provides advice on exceptional weather events including floods with a temporal window of 48 h. There are mainly two kind of floods derived from precipitation events: flash floods and river floods. On the one hand, flash floods are characterised by a time delay, from the peak precipitation time to the peak of flood, from 3 to 6 h. These floods are usually registered in a dry climate and rocky terrain due to the lack of vegetation to filtrate the precipitation into the ground. These kinds of floods have a very high associated level of risk due to their velocity of propagation. On the other hand, river floods are generally registered in larger rivers in areas with a wet climate and the delay time is greater than 6 h. The consequences associated with river floods can also be dramatic for the people and their property. This makes it necessary to develop an EWS to improve the security of the areas exposed to these events. The area of study analysed in this work is mainly affected by river floods.

In this paper, a flood early warning system based on precipitation forecasts is presented. The system, which is being developed in collaboration with the Hydrographic Confederation of Miño–Sil, consists of three steps: (i) precipitation forecast, (ii) use of a hydrologic model to predict extreme flows and (iii) use of a hydraulic model that is applied at certain areas only under extreme flows. Starting from 1, 2

and 3 d precipitation forecast windows provided by the regional meteorological office (MeteoGalicia), the outflows associated with the catchment of the Miño River (NW Spain) were obtained using the Hydrologic Modeling System (HEC-HMS) model (U.S. Army Corps of Engineers, 2018). This model was calibrated for the area of study by means of a series of historical flood events detected over the last decade. The numerical model Iber (Bladé et al., 2014) was used to obtain water depth and velocity under extreme flow conditions for some risk areas where previous events have caused damage or material losses. Both models (i.e. HEC-HMS and Iber) are freely available software so the system can be applied at any location without costs derived from the licences of commercial codes. The main contribution of the EWS presented in this work with respect to the systems shown in the bibliography is that all the components are freely available and easily adaptable to different areas of the world.

The paper, which aims to describe the steps followed to develop the EWS, is organised as follows. First, a description of the area of study (the upper reach of Miño River and the city of Lugo, NW Spain) is shown. Then the methodology to obtain the weather forecast, the computation of the runoff and the hydraulic processes are briefly presented. Also, the communication among all the models (precipitation forecast–runoff–hydraulic processes) is explained. Next, the results of the precipitation and outflow forecast of a series of historical flood events are presented along with a statistic analysis of their accuracy. Finally, the numerical water depth obtained for a particular flood event at the city of Lugo is shown and compared with field data measured during the event.

2 Study area

The area of study is located in northwestern Spain (Fig. 1). It corresponds to the upper reach of the Miño River. This sub-catchment area is about $2200\,\mathrm{km}^2$ and the elevation ranges from 360 to 980 m a.s.l. The average annual precipitation ranges from 144 to $1300\,\mathrm{mm\,yr}^{-1}$. Miño River presents an annual hydrologic cycle characterised by a pluvial regime, with maximum river discharges during winter months, then descending to reach its minimum values during summer (Fernández-Nóvoa et al., 2017). Specifically, considering the period under study at Lugo station, Miño River reaches maximum flows of 114 and $128\,\mathrm{m}^3\,\mathrm{s}^{-1}$ in January and February and minimum flows of 7 and $8\,\mathrm{m}^3\,\mathrm{s}^{-1}$ in August and September, respectively.

Figure 1a shows the catchment of the upper reach of the Miño River, which is divided into three main sub-basins according to their topographic characteristics. Seven rain gauges operated by MeteoGalicia are located in the entire sub-catchment. Table 1 shows the location and the elevation of each of the rain gauges located in the upper reach of the Miño River. The outlet of this catchment is located in the city of Lugo (Fig. 1c). This area is usually flooded during events
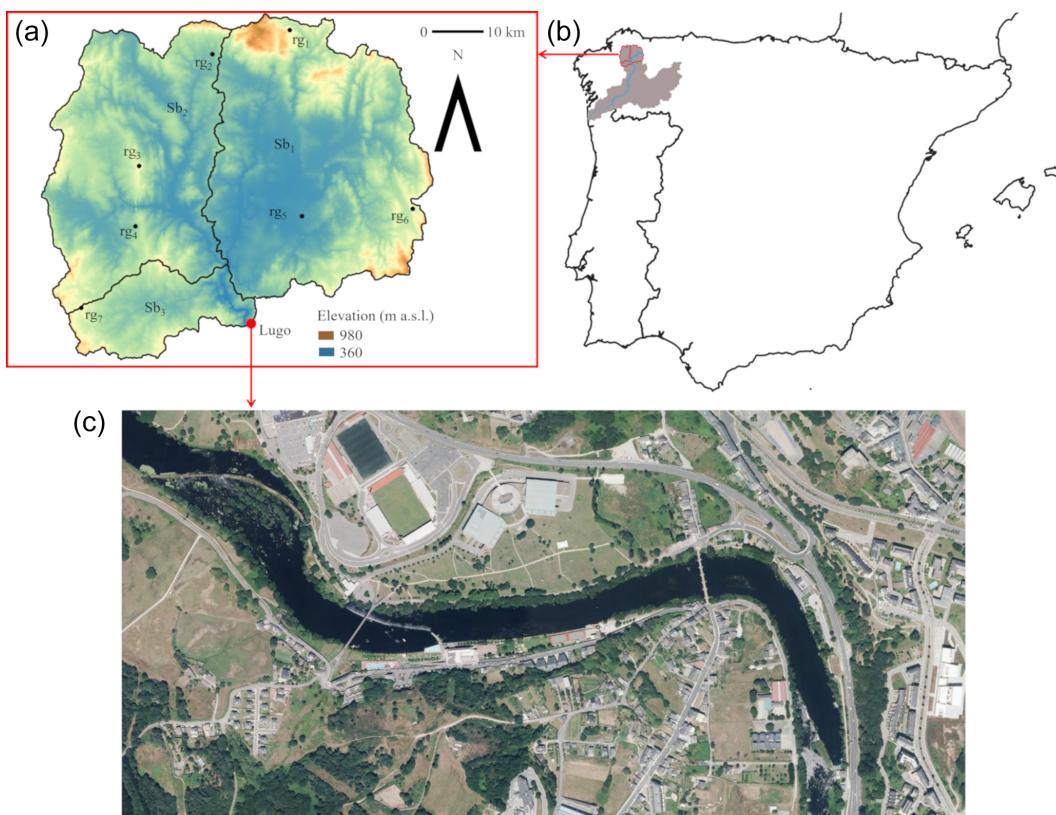
Figure 1. Area of study. In **(b)**, the location of the entire catchment of the shared Portuguese–Spanish river (shaded area) in the Iberian Peninsula and the riverbed of the Miño River (blue line) are shown. The rain gauges (rg1, ..., rg7) located in the catchment and the sub-basins (Sb$_1$, Sb$_2$ and Sb$_3$) of the domain **(a)** as well as the area of study in Lugo **(c)** are also shown (PNOA courtesy of © Instituto Geográfico Nacional).

of extreme precipitation in the upper reach of the Miño River. The absence of dams in the catchment to regulate the flow also affects the high frequency of these events.

3 Methodology

In this work, an automatic EWS is proposed. This system is composed of several elements as shown in Fig. 2. All these components are orchestrated by a Python script that is responsible for gathering and transforming the data properly in order to feed the models used in the system. First of all, the rainfall forecast performed with the Weather Research and Forecasting model is provided by the weather agency (MeteoGalicia). Details are provided in next section. Forecasted data are automatically downloaded, and the rainfall relative to each sub-basin is extracted to feed the hydrological model HEC-HMS. When the catchment outflow obtained with HEC-HMS surpasses the 90th percentile of historical data, it is considered a possible extreme event and the following steps will be applied. This outflow will be used as the inlet condition for the hydraulic simulation using the model Iber to provide flood maps with water depths and ve-

locities at certain risk areas (the city of Lugo in this particular case). Data provided by Iber are processed for hazard evaluation. At this stage the system checks if there is a risk condition in the areas accessible by pedestrians. These areas are user defined and can be changed depending on seasonal events. In order to emit a warning alert, the criteria of Cox et al. (2010) are used to define safety limits for children since they are the most vulnerable population group. Following this criterion, a warning will be emitted if there is a zone where any of the following thresholds are surpassed: the water depth (h) is higher than 0.5 m, the magnitude of water velocity (v) is higher than $0.2\,\mathrm{m\,s^{-1}}$ or the product ($h \times v$) in excess of $0.4\,\mathrm{m^2\,s^{-1}}$. This warning is sent in the form of report to a decision maker so an expert can validate the resulting data and discard false positives.

The details of the components of the EWS, the data sources and the calibration processes are described in the following sections.

Table 1. Location and elevation of the rain gauges located in the area of study (the system of reference for latitude and longitude is the EPSG: 4326).

Rain gauge ID	Name	Latitude	Longitude	Elevation (m a.s.l.)
rg_1	Labrada	43.4054	−7.50205	662
rg_2	Lanzós	43.3746	−7.64468	470
rg_3	Guitiriz-Mirador	43.2266	−7.78307	684
rg_4	Sanbreixo	43.1457	−7.79112	496
rg_5	Castro de Rei Lea	43.1559	−7.48588	428
rg_6	Pol	43.1626	−7.28258	647
rg_7	Corno do Boi	43.0374	−7.89265	731
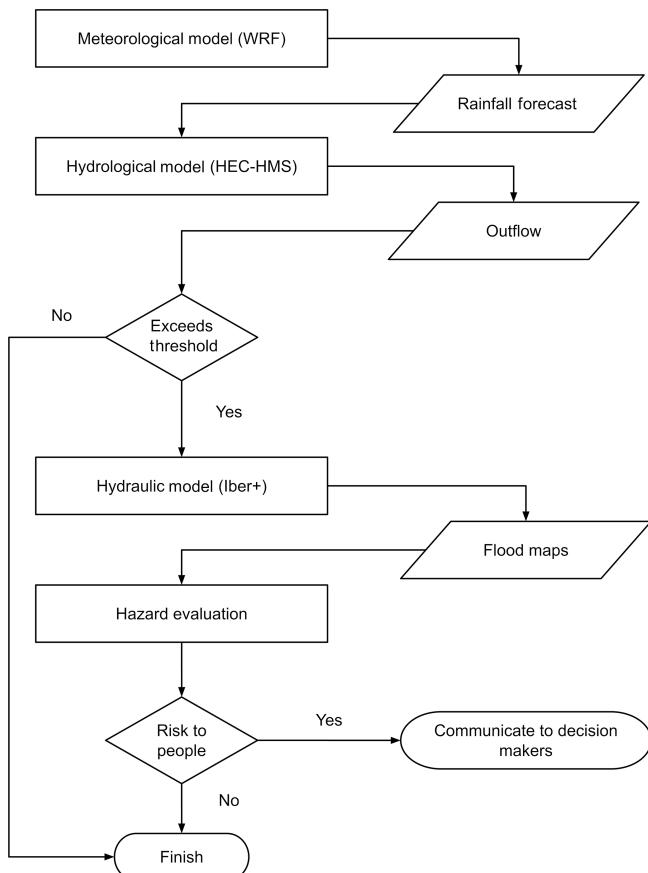
Figure 2. Flow chart of the proposed EWS.

3.1 Precipitation data

3.1.1 Forecasted precipitation data

Forecasted precipitation data were obtained from the regional meteorological office (MeteoGalicia, http://www.meteogalicia.gal/, last access: September 2019). MeteoGalicia publishes weather forecast results based on the Weather Research and Forecasting (WRF) model (Skamarock et al., 2005) (https://www.mmm.ucar. edu/weather-research-and-forecasting-model, last access: September 2019). The WRF model is a numerical weather prediction system at the regional mesoscale designed mainly for forecasting applications. WRF has been run operationally since 2008 providing daily data until the end of 2012 (00:00 UTC) and twice a day (00:00 and 12:00 UTC) from then on, with a 72 h forecast window, a temporal resolution of 1 h and maximum spatial resolution of 4 km (Sousa et al., 2013). Data provided by MeteoGalicia are freely available from the THREDDS (Thematic Realtime Environmental Distributed Data Service) server, also maintaining a historical archive of past forecasts since 2008. The model outputs provide several variables related to weather. In the case of this study, precipitation information was automatically obtained for the areas under interest at 00:00 UTC of each day during the period 2008–2018.

3.1.2 Measured precipitation data

Real precipitation data at an hourly scale were obtained from the rain gauges managed by MeteoGalicia, which is responsible for their maintenance and data quality control. Data from these rain gauges were used to assess the performance of the MeteoGalicia weather forecast to predict extreme rain events. The mentioned rain gauges are pictured in Fig. 1 and their location and elevation are detailed in Table 1.

3.2 River discharge data

Daily discharge data of the Miño River were provided by the corresponding river basin authority (Confederación Hidrográfica del Miño-Sil, https://www.chminosil.es, last access: September 2019). In this case of study, Miño flow data at Lugo station covering the period 2008–2018 were selected. River data were used to calibrate and validate the hydrologic model system used during the development of this study.

3.3 HEC-HMS & Iber+

Here the hydrological and hydraulic models used in the study will be briefly described along with the methods to analyse their accuracy.

The semi-distributed model HEC-HMS (Feldman, 2000 and U.S. Army Corps of Engineers, 2018) was used to analyse the rain–runoff processes, and the numerical model Iber (Bladé et al., 2014) was used to compute the hydraulic processes.

The HEC-HMS is a model developed by the U.S. Army Corps of Engineers that is applied to simulate continuous hydrological processes. The HEC-HMS model can be used to analyse various hydrological aspects, such as flooding events, reservoir capacity, storm water warnings and stream restoration (U.S. Army Corps of Engineers, 2008). HEC-HMS is divided into four components: (i) an analytical model: calculation of direct runoff and channel routing; (ii) a basin model: representation of hydrological elements in a watershed; (iii) a system to manage input data and store data; and (iv) a post-processing tool to report and illustrate simulation results. Two main processes were taken into account in the methodology developed in this case of study: loss (infiltration) and transformation methods. In the first case, the Soil Conservation Service (SCS) curve number was selected. This method implements the curve number methodology for incremental losses since it was designed to calculate the infiltration during periods of heavy rainfall and therefore is well suited to this type of study. With respect to the transform process, based on the method of converting the excess precipitation as runoff, the SCS unit hydrograph method was also selected for the reasons mentioned above. More information about the loss and transformation methods used in this work is detailed in NRCS (2007). Last, the Muskingum–Cunge routing method was selected for runoff propagation because it provides a good approach in basins with similar slopes. This method takes into account the conservation of mass as well as the diffusion representation of the conservation of momentum (U.S. Army Corps of Engineers, 2008). Other parameters like the base flow were not considered because they represent less than 3 % of the peak flow for this kind of event and can be neglected.

Taylor diagrams (Taylor, 2001) were used to compute the accuracy of the results obtained with HEC-HMS by means of the normalised standard deviation (Eq. 1), normalised centred root-mean-square difference (Eq. 2) and correlation (Eq. 3).

$$\sigma_{n,A} = \frac{\sqrt{\frac{\sum_{i=1}^{N} \left(A_i - \overline{A} \right)^2}{N}}}{\sigma_B} \tag{1}$$

$$E_{n,A} = \frac{\sqrt{\frac{\sum_{i=1}^{N} \left[\left(A_i - \overline{A} \right) - \left(B_i - \overline{B} \right) \right]^2}{N}}}{\sigma_B} \tag{2}$$

$$R_A = \frac{\sum_{i=1}^{N} \left[\left(A_i - \overline{A} \right) \left(B_i - \overline{B} \right) \right]}{N \sigma_A \sigma_B} \tag{3}$$

Here A is a numerical variable and B a reference variable. The subscript n refers to the normalised parameter, subscript i refers to the different samples, N is the number of samples, barred variables refer to mean values and σ is the standard deviation.

The hydraulic simulations were carried out using the numerical model Iber (Bladé et al., 2014). Iber is a numerical code that solves the 2-D (two-dimensional) shallow water equations by means of finite volume schemes (FVSs). The software package is formed by three elements: a preprocessing tool, numerical model and post-processing tool. The first and the last modules are based in the software GID (GID, 2018). It provides a user-friendly graphical interface (GUI) to create the case and edit the parameters that define the problem to solve. It also provides tools to analyse the results of the numerical simulations. The preprocessing and post-processing tools were used only during the modelling and testing of the study area. However, the automatic EWS runs the model in batch mode without user interaction. Iber was recently improved in terms of efficiency, becoming Iber+ (García-Feal et al., 2018). This new parallel implementation of the Iber model takes advantage of GPU (graphics processing unit) computing using the Nvidia CUDA (NVIDIA Corporation, 2019) platform. Using this technology, the new implementation is able to run up to 100 times faster. This fact makes Iber+ especially suitable for the implementation of an EWS where the response times can be crucial to issuing an early alert. The accuracy of the water depth results computed with Iber+ at five control points was assessed by means of the bias and the RMSE (root-mean-square error) for the extreme event recorded on January 2013.

$$\text{RMSE} = \sqrt{\frac{\sum_{i=1}^{N} (A_i - B_i)^2}{N}} \tag{4}$$

$$\text{Bias} = \frac{\sum_{i=1}^{N} (A_i - B_i)}{N} \tag{5}$$

Here A is the numerical value, B the measured value and N the number of control points.

4 Results and discussion

4.1 Accuracy of MeteoGalicia precipitation forecast

The capability of the MeteoGalicia weather forecast system to predict rain events was evaluated by means of the comparison with real precipitation data provided by the rain gauges in the area of study. For that purpose, the predicted (numerical) precipitation was obtained at the grid points closest to the location of the rain gauges. The correlation between predicted and measured precipitation was calculated for each rain gauge during the available period (2008–2018). Spearman rank correlation was used due to its robustness of deviations from linearity as well as its strength against the influence of outliers. This procedure was carried out for three forecast windows (1–24, 25–48 and 49–72 h; 1, 2 and 3 d forecast from now on) to determine the accuracy of the

Table 2. Values of the correlation (Spearman's r) and normalised standard deviation (σ_n) of the precipitation forecast using the measured data as reference at each rain gauge, considering the complete time series of precipitation. The averaged values for each precipitation forecast are also shown.

| Rain gauge | Forecast window (h) | | | | | |
| | 1–24 | | 25–48 | | 49–72 | |
	r	σ_n	r	σ_n	r	σ_n
rg_1	0.84	0.80	0.82	0.81	0.77	0.80
rg_2	0.84	1.09	0.82	1.07	0.79	1.07
rg_3	0.83	1.00	0.81	0.96	0.77	0.99
rg_4	0.81	0.97	0.79	0.96	0.75	0.98
rg_5	0.81	1.13	0.80	1.10	0.76	1.12
rg_6	0.84	1.16	0.83	1.07	0.79	1.07
rg_7	0.83	1.05	0.81	1.06	0.77	1.10
Mean value	0.83	1.03	0.81	1.00	0.77	1.02

Table 3. Values of the correlation (Spearman's r) and normalised standard deviation (σ_n) of the precipitation forecast using the measured data as reference at each rain gauge, considering only rainy events (above the 75th percentile). The averaged values for each precipitation forecast are also shown.

| Rain gauge | Forecast window (h) | | | | | |
| | 1–24 | | 25–48 | | 49–72 | |
	r	σ_n	r	σ_n	r	σ_n
rg_1	0.66	0.72	0.61	0.70	0.53	0.72
rg_2	0.71	1.00	0.63	0.98	0.56	0.99
rg_3	0.70	0.98	0.61	0.93	0.59	0.98
rg_4	0.73	0.93	0.65	0.90	0.60	0.93
rg_5	0.68	1.02	0.63	1.01	0.54	1.04
rg_6	0.69	1.14	0.65	0.98	0.56	1.00
rg_7	0.74	1.03	0.68	1.02	0.63	1.10
Mean value	0.70	0.97	0.64	0.93	0.57	0.97

forecast at different temporal scales. The comparison is carried out for an aggregation time of 24 h, which matches the recording frequency of rain data provided by MeteoGalicia and is compatible with the kind of flood events (mainly river floods) of the area.

The values of the correlation and the normalised standard deviation for each rain gauge are shown in Tables 2 and 3. Table 2 shows the analysis for the complete series and Table 3 shows the results considering only rainy events (precipitation above the 75th percentile). In general, considering the complete series, precipitation prediction offers a good representation of the registered values and the variability of precipitation. In fact, correlations above 0.8 were obtained for the first two windows (1 and 2 d forecast), although with a higher correlation for the first one. The correlation is slightly lower for the 3 d forecast, although it is still close to 0.8. When only rainy events are considered mean correlation values are slightly lower than considering the complete series, although they show a good representation of the registered data. The high correlation shown under a 1 d forecast window with a mean value above 0.7 is especially remarkable (Table 3). With respect to the normalised standard deviation, most cases in both series are similar to 1, which shows a good agreement between forecast and real precipitation. Therefore, it can be concluded that the precipitation forecast provided by Meteo-Galicia offers results very close to the real rain events for the entire time series of precipitation data (2008–2018). This shows the accuracy of MeteoGalicia models to forecast precipitation events up to 3 d in advance.

4.2 Calibration and validation of hydrological processes using HEC-HMS

A set of 15 extreme flood events registered during the period 2008–2018 were used to calibrate and validate the rain–

runoff model HEC-HMS (Table 4) by comparing the outflows measured at the gauge station located at Lugo with the flows obtained with HEC-HMS using the 1 d forecast of precipitation. Forecasted rain data were considered because they are used to feed the model in its forecast version. In situ data would be only valid for hindcast purposes. Calibration was carried out using the specific calibration tools implemented in HEC-HMS (Feldman, 2000) in order to choose two independent parameters, the curve number (CN) and lag time (L_g), for each sub-basin. The values of CN and L_g were computed using a particle swarm algorithm (Kennedy and Eberhart, 1995; Pedersen, 2010 and Mezura-Montes and Coello, 2011) to minimise the error between the measured streamflow and the numerical one. No empirical formulas were used for CN and L_g due the uncertainty associated with their definition (Fang et al., 2008; Upegui and Gutiérrez, 2011; Grimaldi et al., 2012). A total of 11 flood events were used for calibration purposes and the rest of the cases were used to validate the model. Table 5 shows the values of the CN and L_g for each sub-basin obtained for each event used in the calibration step.

The mean values of CN and L_g of each sub-basin were used to validate the model in four flood events (January 2013, January 2014, February 2016 and March 2018) by means of a Taylor diagram (Fig. 3).

The values of normalised standard deviation (σ_n) range from 0.8 to 1.2, the values of the root-mean-squared difference (RMSD) range from 0.3 to 0.6 and the correlation of the numerical results ranges from 0.85 to 0.95. The values of σ_n mean that the variability of the numerical results are quite similar to the variability of the reference time series (difference less than 20 %) and the values of E_n can be considered good values according to Moriasi et al. (2007). These values of σ_n, E_n and correlation show that the mean values of
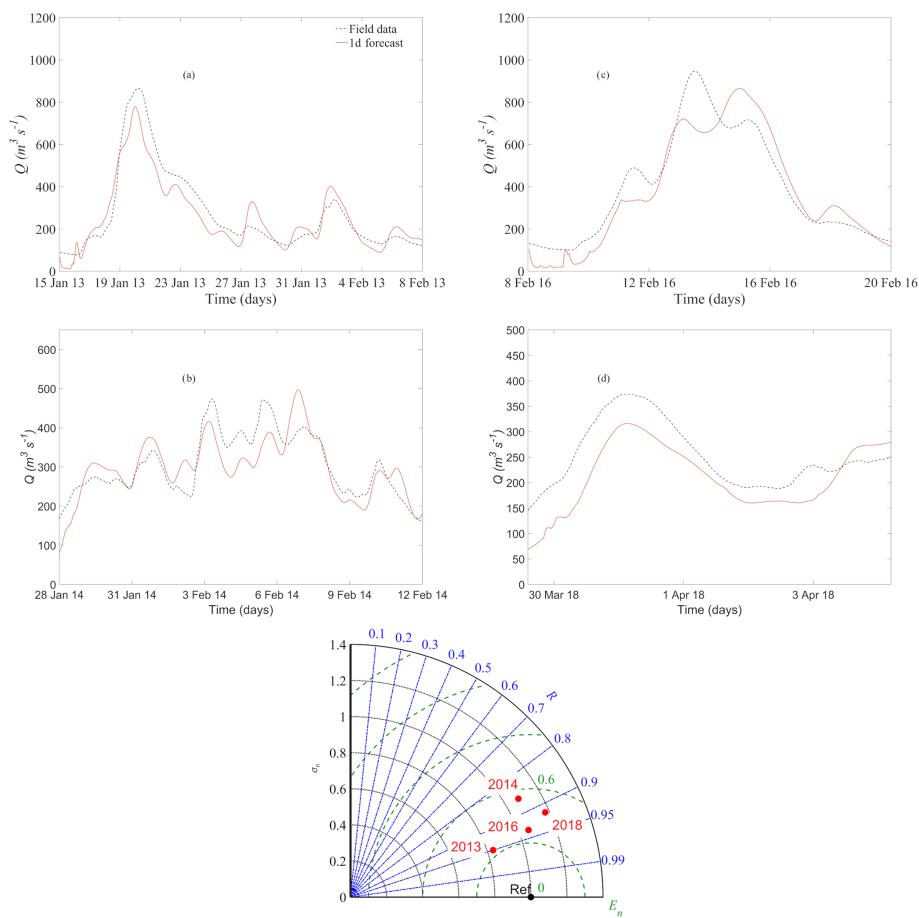
Figure 3. Time series of the registered streamflow (dashed line) and numerical streamflow (orange line) of the validation events: **(a)** January 2013, **(b)** January 2014, **(c)** February 2016 and **(d)** March 2018. Taylor diagram of the validation cases is also shown.

Table 4. Main characteristics of the analysed flood events.

Date of the flood event (dd/mm/yyyy)	Duration (days)	Initial flow ($m^3 s^{-1}$)	Initial depth (m)
28/12/09	4	52	1.3
17/11/10	5	116	1.7
17/01/13	10	164	1.9
11/03/13	5	179	2.0
05/11/13	7	234	2.3
14/01/14	10	165	1.9
28/01/14	15	202	2.1
01/03/14	4	134	1.8
30/01/15	3	184	2.0
01/03/15	3	134	1.8
10/02/16	7	216	2.1
26/02/16	3	137	1.8
05/03/16	4	175	2.0
10/03/18	6	154	1.9
30/03/18	4	201	2.1

CN and L_g obtained in the calibration step characterise the behaviour of the basin with a high accuracy.

Figure 4 compares the numerical and measured streamflow for the event that happened in January 2013 using the three forecast windows. Panel a shows that time series of the flows predicted by the model are similar to those measured at the gauge station. Panel b is the Taylor diagram corresponding to the three forecast windows. The standard deviation is observed to range from 0.8 to 1.2 for the three forecasts. RMSD values for 1 and 2 d forecasts are around 0.3 and around 0.6 for the 3 d forecast. Finally, the correlation coefficient for 1 and 2 d forecasts is close to 0.95 and around 0.85 for the 3 d forecast.

4.3 Case of study

Once the predicted water flow was shown to reproduce the real events with a high accuracy ($E_n \sim 0.8$, $\sigma_n \sim 0.3$ and $R \sim 0.95$), the water depth and velocity during the flood event that affected Lugo on 20 January 2013 were computed using the numerical code Iber+ (García-Feal et al., 2018). Figure 5 shows the numerical domain at Lugo, where seven

Table 5. Curve number (CN) and lag time (L_g) values for each sub-basin for different flood events. The mean value and the standard deviation are provided in lower rows.

Date of the flood event (mm/yy)	Sb$_1$		Sb$_2$		Sb$_3$	
	CN	L_g (min)	CN	L_g (min)	CN	L_g (min)
12/09	92	1154	97	2700	98	2770
11/10	80	1140	84	2702	80	2781
03/13	79	1157	96	2701	99	2774
11/13	80	1148	86	2685	83	2778
01/14	78	1155	96	2700	98	2767
03/14	81	1153	88	2706	92	2764
01/15	96	1153	99	2701	99	2773
03/15	81	1151	91	2700	98	2771
02/16	81	1155	88	2700	98	2767
03/16	82	1153	80	2711	84	2764
03/18	80	1152	82	2691	93	2769
Mean	85	1152	90	2700	93	2771
σ	6	4	6	7	7	5
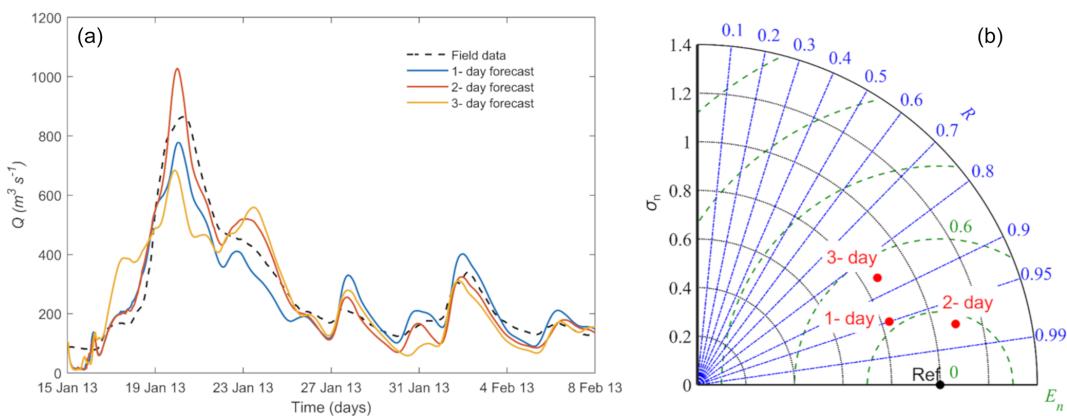
Figure 4. Time series of the outflow at the control point obtained in the gauge station (dashed line) and calculated using the three forecast windows **(a)** and Taylor diagram for the same cases **(b)**.

Table 6. Manning's coefficients of the numerical domain.

Land uses	Manning's coefficient ($\mathrm{s\,m^{-1/3}}$)
River	0.025
Brush	0.050
Trees	0.120
Sparse vegetation	0.080
Infrastructure	0.020
Industrial	0.100
Residential	0.150

land uses were defined to model the characteristics of the terrain. Manning's coefficient associated with each land use is shown in Table 6. Figure 5 also shows the location of the inlet

and outlet boundary conditions. The initial water depth was obtained from data provided by the gauge station located at Lugo. The inlet condition was defined by means of the input hydrograph (critical–subcritical) and the outlet condition was defined using a supercritical–critical outflow. Turbulence was not taken into account as suggested by SNCZI (2011) and in accordance with similar works (Erpicum et al., 2010; Liu et al., 2013; Segura-Beltrán et al., 2016).

The topography of the area of study was obtained from raster files freely downloaded from the Instituto Geográfico Nacional website (https://www.ign.es/web/ign/portal, last access: September 2019). The computational domain was discretised using a mesh with nearly 200 000 unstructured triangular elements, with an average area of $2\,\mathrm{m^2}$.

Five control points were defined at the area of study (see Fig. 6) to analyse the accuracy of the numerical results.
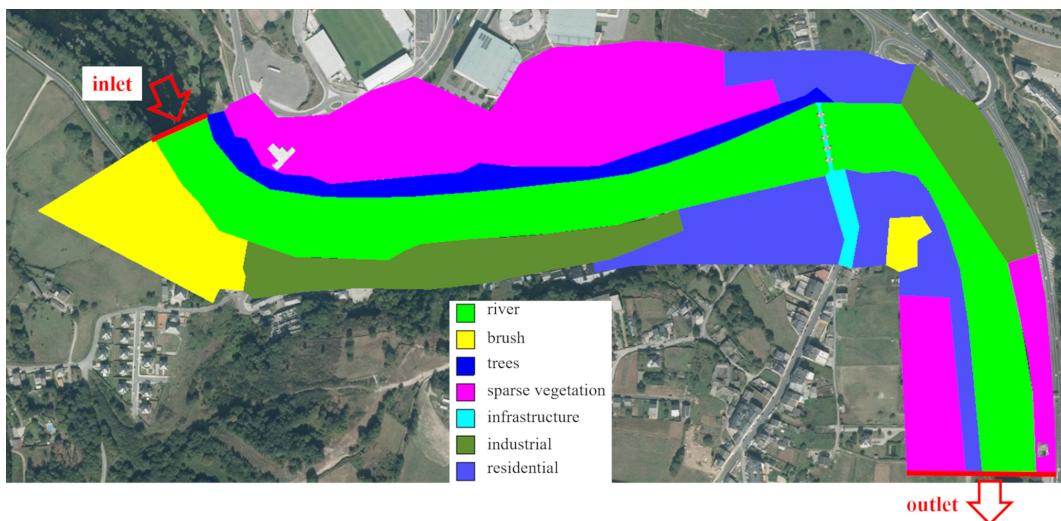
Figure 5. Numerical domain at Lugo. The land uses and the location of the boundary conditions (red lines) are also shown (PNOA courtesy of © Instituto Geográfico Nacional).

Points from 1 to 4 are located in places next to the riverbank usually frequented by pedestrians while the last one is located in the riverbed. Therefore, the first four points are of special interest to issue an alert.

Figure 7 shows the values of the water depth obtained in the numerical simulations along with the water depth obtained at the control points during the flood event. These field values were obtained from photographs provided by volunteers and local media and taken within the interval 12:00–16:00 UTC+1 (for all times throughout the paper) on 20 January. The numerical water depth is expressed in terms of a mean value and a range that corresponds to 3 times the standard deviation of the values within that interval. Visually, the numerical results are quite similar to the field data when considering the 1 d forecast, especially if one considers that the accumulation of the small inaccuracies of the three models involved can give rise to biases. The values are slightly less accurate when considering the 2 d forecast and worse for the 3 d forecast due to lower accuracy in rainfall forecast. Finally, it must be mentioned that the depicted values do not correspond to the peak flow that took place on 21 January 2013 (at approximately 04:00).

Apart the visual comparison, the accuracy of the model to calculate water elevation was analysed in terms of two estimators (RMSE and bias) computed using the three forecast windows. The minimum values of RMSE and bias are obtained with the 24 h forecast window (21 and 0 cm, respectively). The RMSE is satisfactory when compared with the mean upward displacement of water during the event, which is about 2.5 m. In addition, the bias is null, showing that the model (in average) neither overestimates nor underestimates real water elevation. The accuracy decreases with the forecast window, although results are still good for a 2 d forecast (RMSE = 28 cm and bias = 4 cm). Finally, the accu-

racy is acceptable for a 3 d forecast (RMSE = 41 cm and bias = −35 cm), although with limitations in terms of bias since the model clearly tends to underestimate field measurements. In summary, the agreement between measured and computed values indicates that the system can be used to issue an alert up to 3 d in advance.

Figure 8 shows the maximum water depth and maximum velocity obtained for the 1 d forecast. Hazard maps (Fig. 9) can be computed from these data according to the criterion of Cox et al. (2010). Several recreation areas near the riverbanks are shown to have surpassed the aforementioned hazard threshold. Therefore, decision makers can use the map to restrict activities in these areas, in order to mitigate the consequences of floods.

5 Conclusions

In this paper an early warning system for flood prediction using precipitation forecast was presented. This system starts automatically using rain forecast data retrieved from the regional meteorological office (MeteoGalicia) and concatenates two freely available software packages (HEC-HMS and Iber+). The upper reach of the Miño River (NW Spain) and, in particular, the city of Lugo were used as benchmarks.

A Python script was developed to deal with all the components involved in the system without user interaction. First, the precipitation forecast provided by MeteoGalicia is automatically obtained for the area of study. Second, the rain forecast is provided to HEC-HMS as an input to compute the streamflow in the catchment area. When the streamflow obtained with HEC-HMS surpasses the 90th percentile of the historical percentile at some previously selected risk area (the city of Lugo in this particular case), the possibility of an ex-

Figure 6. Location of the five control points at the area of study in Lugo (PNOA courtesy of © Instituto Geográfico Nacional).
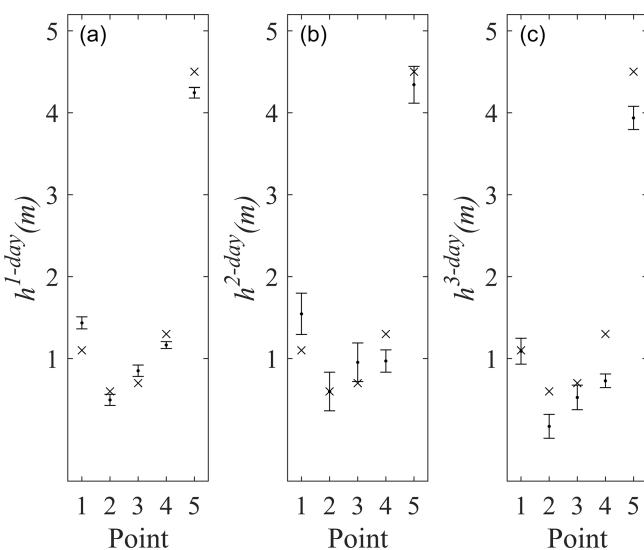

Figure 7. Comparison between water depth (h in metres) between the numerical model ($\cdot$) and the field data ($\times$) for the three forecast windows 1 d (**a**), 2 d (**b**) and 3 d (**c**). The range of the numerical values corresponds to 3 times the standard deviation of the elevations obtained from 12:00 to 16:00 on 20 January 2013.

treme event is detected and that streamflow is automatically defined as an inlet condition for Iber+. Finally, data obtained from Iber+ are processed for risk assessment and, if applicable, decision makers are notified.

The accuracy of the different models was assessed to analyse the capability of the system to provide reliable results. First, the accuracy of the precipitation forecast provided by MeteoGalicia was analysed for the period 2008–2018 showing that the 1 d forecast is slightly more accurate than the 2 d forecast, with the 3 d forecast being slightly worse, although the three forecast windows showed a reasonable agreement with field data. As a second step, the accuracy of HEC-HMS

to reproduce extreme flows was assessed by means of 15 flood events recorded over for the period 2008–2018. Taylor diagrams were used to compute the accuracy of the numerical streamflow compared with field data obtained at the control station located near Lugo. Once again, results were satisfactory for the three forecast windows, especially for the 1 and 2 d forecasts. Finally, a historical flood event recorded in January 2013 was used to assess the accuracy of Iber+ to reproduce real water elevation at five control points located at the riverbank and riverbed. Both the RMSE and the bias between the measured and computed elevations were satisfactory, especially for the 1 d forecast.

The system needs less than 1 h to run the models for a 3 d forecast horizon. While data can be downloaded in a few seconds and the hydrologic model can be run in less than a minute, no matter the extent of the area, the real bottleneck in the system is the hydraulic model. Fortunately, the execution time does not necessarily increase with the number of risk areas since different areas can be run concurrently when the available hardware resources allow it. Taking into account that meteorological data are available every day at 05:00, the system can provide an alert report to decision makers before 06:00. Additional improvements can be applied without additional cost in terms of runtime. For example, an ensemble approach can be applied when rain forecasts from different sources are used as an input condition for HEC-HMS, in such a way that Iber+ is only executed when at least one of the hydrological realisations indicates a possible extreme event.

Additional research is still needed to cover the entire Miño River basin, where other problems may arise from the presence of dams. The system, when fully developed, can even help to manage dams intelligently, maximising energy production and dampening floods at the same time.

The early warning system can be easily adapted for any area of the world since the required input data can be obtained freely from public institutions and the models to com-
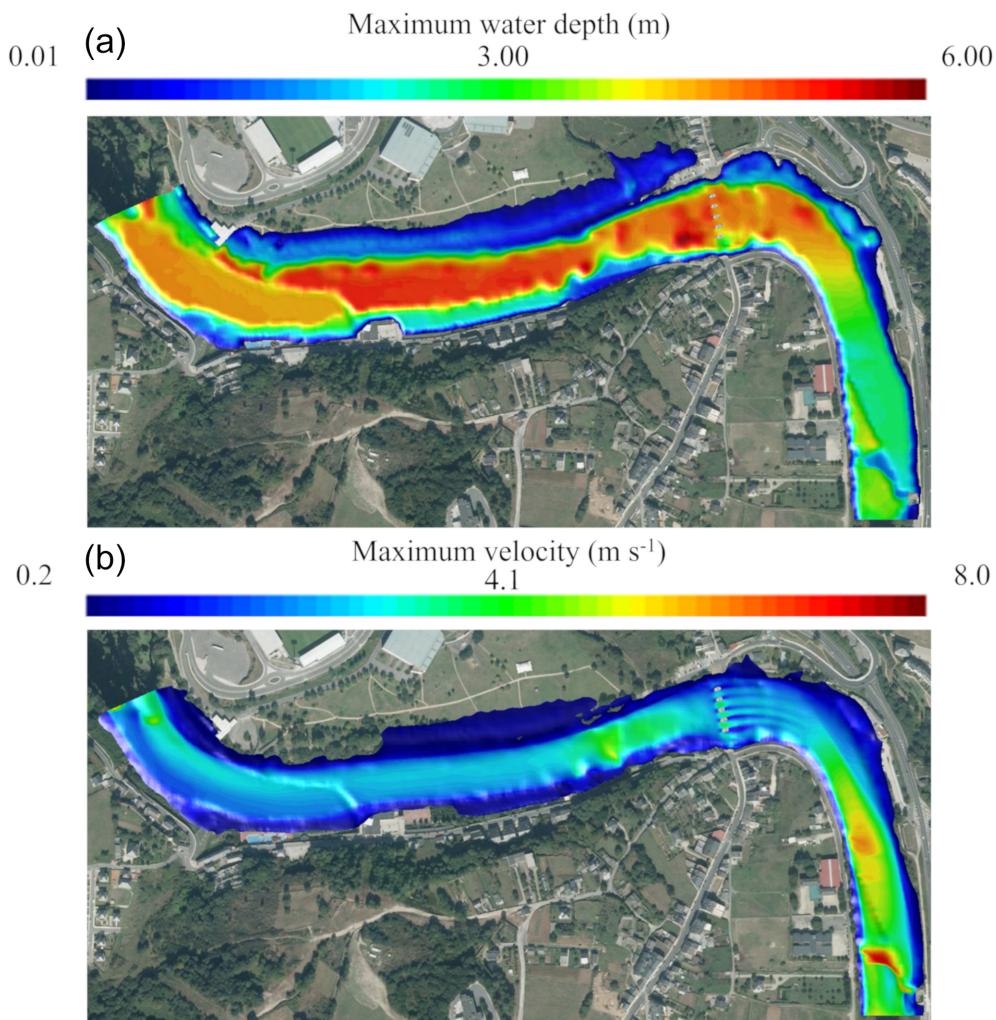
Figure 8. Maximum water depth (**a**) and maximum velocity (**b**) obtained with Iber+ for the 1 d precipitation forecast (PNOA courtesy of © Instituto Geográfico Nacional).

Figure 9. Areas where hazard criterion is surpassed (PNOA courtesy of © Instituto Geográfico Nacional).

pute the hydrological and the hydraulic processes (HEC-HMS and Iber+, respectively) are both freely available. Therefore, the EWS is especially interesting for developing countries where the acquisition of commercial software is not sustainable.

References

Alfieri, L., Salamon, P., Pappenberger, F., Wetterhall, F., and Thielen, J.: Operational early warning systems for water-related hazards in Europe, Environ. Sci. Policy, 21, 35–49, https://doi.org/10.1016/j.envsci.2012.01.008, 2012.

Alfieri, L., Pappenberger, F., Wetterhall, F., Haiden, T., Richardson, D., and Salamon. P.: Evaluation of ensemble streamflow predictions in Europe, J. Hydrol., 517, 913–922, https://doi.org/10.1016/j.jhydrol.2014.06.035, 2014.

Alfieri, L., Bisselink, B., Dottori, F., Naumann, G., de Roo, A., Salamon, P., Wyser, K., and Feyen, L.: Global projections of river flood risk in a warmer world, Earth's Future, 5, 171–182, https://doi.org/10.1002/2016EF000485, 2017.

Baudoin, M., Henly-Shepard, S., Fernando, N., Sitati, A., and Zommers, Z.: Early warning systems and livelihood resilience: exploring opportunities for community participation, UNU-EHS Working Paper Series, No. 1, United Nations University Institute of Environment and Human Security (UNU-EHS), Bonn, 2014.

Bladé, E., Cea, L., Corestein, G., Escolano, E., Puertas, J., Vázquez-Cendón, E., Dolz, J., and Coll, A.: Iber -River modelling simulation tool [Iber: herramienta de simulación numérica del flujo en ríos], Revista Internacional de Metodos Numericos para Calculo y Diseno en Ingenieria, 30, 1–10, https://doi.org/10.1016/j.rimni.2012.07.004, 2014.

Borga, M., Anagnostou, E. N., Blöschl, G., and Creutin, J. D.: Flash flood forecasting, warning and risk management: The HYDRATE project, Environ. Sci. Policy, 14, 834–844, https://doi.org/10.1016/j.envsci.2011.05.017, 2011.

CIW: Rapport Globale Evaluatie Overstromingen, [Evaluation Report for Floods] Committee for Integrated Water Resources Management (CIW), Flemish Authority, Brussels, Belgium, 2011 (in Dutch).

Cloke, H., Pappenberger, F., Thielen, J., and Thiemig, V.: Operational European flood forecasting, in: Environmental Modelling: Finding Simplicity in Complexity, 2nd ed., edited by: Wainwright, J. and Mulligan, M., John Wiley and Sons, Ltd, Chichester, UK, https://doi.org/10.1002/9781118351475.ch25, 2013.

Cools, J., Vanderkimpen, P., El Afandi, G., Abdelkhalek, A., Fockedey, S., El Sammany, M., Abdallah, G., El Bihery, M., Bauwens, W., and Huygens, M.: An early warning system for flash floods in hyper-arid Egypt, Nat. Hazards Earth Syst. Sci., 12, 443–457, https://doi.org/10.5194/nhess-12-443-2012, 2012.

Cox, R. J., Shand, T. D., and Blacka, M. J.: Australian Rainfall and Runoff revision project 10: appropriate safety criteria for people, Water Res., ISBN 978-085825-9454, 2010.

Dankers, R. and Feyen, L.: Climate change impact on flood hazard in Europe: An assessment based on high-resolution climate simulations, J. Geophys. Res., 113, D19105, https://doi.org/10.1029/2007JD009719, 2008.

Demeritt, D., Nobert, S., Cloke, H. L., and Pappenberger, F.: The European Flood Alert System and the communication, perception, and use of ensemble predictions for operational flood risk management, Hydrol. Process., 27, 147–157, https://doi.org/10.1002/hyp.9419, 2013.

Dottori, F., Kalas, M., Salamon, P., Bianchi, A., Alfieri, L., and Feyen, L.: An operational procedure for rapid flood risk assessment in Europe, Nat. Hazards Earth Syst. Sci., 17, 1111–1126, https://doi.org/10.5194/nhess-17-1111-2017, 2017.

Erpicum, S., Dewals, B., Archambeau, P., Detrembleur, S., and Pirotton, M.: Detailed inundation modelling using high resolution DEMs, Eng. Appl. Comp. Fluid., 4, 106–208, https://doi.org/10.1080/19942060.2010.11015310, 2010.

Fang, X., Thompson, D. B., Cleveland, T. G., Pradhan, P., and Malla, R.: Time of concentration estimated using watershed parameters determined by automated and manual methods, J. Irrig. Drain. Eng., 134, 202–211, https://doi.org/10.1061/(ASCE)0733-9437(2008)134:2(202), 2008.

Feldman, A. D.: Hydrologic Modeling System HEC-HMS, Technical Reference Manual, Institute for Water Resources Davis, USA, p. 157, 2000.

Fernández-Nóvoa, D., deCastro, M., Des, M., Costoya, X., Mendes, R., and Gómez-Gesteira, M.: Characterization of Iberian turbid plumes by means of synoptic patterns obtained through MODIS imagery, J. Sea Res., 126, 12–25, https://doi.org/10.1016/j.seares.2017.06.013, 2017.

García-Feal, O., González-Cao, J., Gómez-Gesteira, M., Cea, L., Domínguez, J. M., and Formella, A.: An accelerated tool for flood modelling based on Iber, Water, 10, 1459, https://doi.org/10.3390/w10101459, 2018.

GID: Reference Manual: available at: https://www.gidhome.com/ (last access: September 2019), 2018.

Grimaldi, S., Petroselli, A., Tauro, F., and Porfiri, M.: Time of concentration: a paradox in modern hydrology, Hydrolog. Sci. J., 57, 217–228, https://doi.org/10.1080/02626667.2011.644244, 2012.

Hossain, F., Siddique-E-Akbor, A. H. M., Yigzaw, W., Shah-Newaz, S., Hossain, M., Mazumder, L. C., Ahmed, T., Shum, C. K., Lee, H., Biancamaria, S., Turk, F. J., and Limaye, A.: Crossing the "valley of Death": Lessons learned from implementing an operational satellite-based flood forecasting system, B. Am. Meteorol. Soc., 95, 1201–1207, https://doi.org/10.1175/BAMS-D-13-00176.1, 2014.

Jonkman, S. N.: Global perspectives on loss of human life caused by floods, Nat. Hazards, 34, 151–175, https://doi.org/10.1007/s11069-004-8891-3, 2005.

Kennedy, J. and Eberhart, R.: Particle swarm optimization. IEEE International Conference on Neural Networks – Conference Proceedings, 4, 1942–1948, 1995.

Liu, Y., Zhou, J., Song, L., Zou, Q., Liao, L., and Wang, W.: Numerical modelling of free-surface shallow flows over irregular topography with complex geometry, Appl. Math. Model, 37, 9482–9498, https://doi.org/10.1016/j.apm.2013.05.001, 2013.

Mezura-Montes, E. and Coello, C. A. C.: Constraint-handling in nature-inspired numerical optimization: Past, present and future, Swarm Evol. Comput., 1, 173–194, 2011.

Moriasi, D. N., Arnold, J. G., Van Liew, M. W., Bingner, R. L., Harmel, R. D., and Veith, T. L.: Model evaluation guidelines for systematic quantification of accuracy in watershed simulations, T. ASABE, 50, 885–900, 2007.

NRCS: Natural Resources Conservation Service: National Engineering Handbook: Chapter 16 Hydrographs, Washington, DC, 2007.

Noji, E. K.: The public health consequences of disasters, Prehospital and disaster medicine: the official journal of the National Association of EMS Physicians and the World Association for Emergency and Disaster Medicine in association with the Acute Care Foundation, 15, 147–157, 2000.

NVIDIA Corporation: CUDA C Programming Guide, available at: https://docs.nvidia.com/cuda/pdf/CUDA_C_Programming_Guide.pdf (last access: September 2019), 2019.

Pappenberger, F., Thielen, J., and Del Medico, M.: The impact of weather forecast improvements on large scale hydrology: analysing a decade of forecasts of the European Flood Alert System, Hydrol. Process., 25, 1091–1113, https://doi.org/10.1002/hyp.7772, 2011.

Pedersen, M. E. H.: Good parameters for particle swarm optimization, Tech. Rep. HL1001, Hvass Lab Copenhagen, Denmark, 2010.

Schelfaut, K., Pannemans, B., van der Craats, I., Krywkow, J., Mysiak, J., and Cools. J.: Bringing flood resilience into practice – the FREEMAN project, Environ. Sci. Policy, 14, 825–833, https://doi.org/10.1016/j.envsci.2011.02.009, 2011.

Segura-Beltrán, F., Sanchis-Ibor, C., Morales-Hernández, M., González-Sanchís, M., Bussi, G., and Ortiz, E.: Using post-flood surveys and geomorphologic mapping to evaluate hydrological and hydraulic models: The flash flood of the Girona River (Spain) in 2007, J. Hydrol., 541, 310–329, https://doi.org/10.1016/j.jhydrol.2016.04.039, 2016.

Skamarock, W. C. , Klemp, J. B., Dudhia, J., Gill, D. O., Barker, D. M., Wang, W., and Powers, J. G.: A Description of the Advanced Research WRF Version 2, Tech. Rep., National Center for Atmospheric Research, 2005.

SNCZI: Guía Metodológica para el Desarrollo del Sistema Nacional de Cartografía de Zonas Inundables, Ministerio de Medio Ambiente y Medio Rural y Marino, Gobierno de España, 2011.

Sousa, M. C., Alvarez, I., Vaz, N., Gomez-Gesteira, M., and Dias, J. M.: Assessment of wind pattern accuracy from the QuikSCAT satellite and the WRF model along the Galician coast (Northwest Iberian Peninsula), Mon. Weather Rev., 141, 742–753, https://doi.org/10.1175/MWR-D-11-00361.1, 2013.

Taylor, K. E.: Summarizing multiple aspects of model performance in a single diagram, J. Geophys. Res., 106, 7183–7192, 2001.

Thielen, J., Bartholmes, J., Ramos, M.-H., and de Roo, A.: The European Flood Alert System – Part 1: Concept and development, Hydrol. Earth Syst. Sci., 13, 125–140, https://doi.org/10.5194/hess-13-125-2009, 2009.

UNISDR: UNISDR terminology on disaster risk reduction, United Nations Office for Disaster Risk Reduction, 30 pp., 2009.

UNISDR: Making development sustainable: the future of disaster risk management. Global Assessment Report on Disaster Risk Reduction, United Nations Office for Disaster Risk Reduction (UNISDR), Geneva, Switzerland, 2015.

UN: Global survey of early warning systems, A report prepared at the request of the Secretary-General of the United Nations, Platform for the Promotion of Early Warning (UNDRR PPEW), 2006.

Upegui, J. J. V. and Gutiérrez, A. B.: Estimation of the time of concentration and the lag time at san Luis creek basin, Manizales [Estimación del tiempo de concentración y tiempo de rezago en la cuenca experimental urbana de la quebrada san luis, manizales], DYNA (Colombia), 78, 58–71, 2011.

U.S. Army Corps of Engineers: Hydrologic Modeling System (HEC-HMS) Applications Guide: Version 3.1.0. Davis: Institute for Water Resources, Hydrologic Engineering Center, 2008.

U.S. Army Corps of Engineers: Hydrologic Modeling System (HEC-HMS) User's Manual: Version 4.3. Institute for Water Resources Davis: Hydrologic Engineering Center, 2018.

5.3. IberWQ: A GPU Accelerated Tool for 2D Water Quality Modeling in Rivers and Estuaries

García-Feal, O., Cea, L., González-Cao, J., Domínguez, J. M. & Gómez-Gesteira, M. (2020). IberWQ: A GPU Accelerated Tool for 2D Water Quality Modeling in Rivers and Estuaries [Publisher: Multidisciplinary Digital Publishing Institute]. *Water, 12*(2), 413

Información de la revista	
Revista	Water
ISSN	2073-4441
País	Suíza
Editorial	MDPI
Factor de impacto	2,544 (2019*)
Cuartil	Q2

IberWQ: A GPU Accelerated Tool for 2D Water Quality Modeling in Rivers and Estuaries

Orlando García-Feal [1,*] , **Luis Cea** [2] , **José González-Cao** [1], **José Manuel Domínguez** [1] and **Moncho Gómez-Gesteira** [1]

1   Environmental Physics Laboratory (EPhysLab), CIM-UVIGO, Universidade de Vigo, Campus As Lagoas s/n, 32004 Ourense, Spain; jgcao@uvigo.es (J.G.-C.); jmdominguez@uvigo.es (J.M.D.); mggesteira@uvigo.es (M.G.-G.)
2   Environmental and Water Engineering Group, Department of Civil Engineering, Universidade da Coruña, Elviña, 15071 A Coruña, Spain; luis.cea@udc.es
*   Correspondence: orlando@uvigo.es; Tel.: +34-988-387-270

Abstract: Numerical models are useful tools to analyze water quality by computing the concentration of physical, chemical and biological parameters. The present work introduces a two-dimensional depth-averaged model that computes the most relevant and frequent parameters used to evaluate water quality. High performance computing (HPC) techniques based on graphic processing unit (GPU) parallelization have been applied to improve the efficiency of the package, providing speed-ups of two orders of magnitude in a standard PC. Several test cases were analyzed to show the capabilities and efficiency of the model to evaluate the environmental status of rivers and non-stratified estuaries. IberWQ will be freely available through the package Iber.

1. Introduction

The application of Environmental Quality Standards (EQS) has become a widely used technique to assess the status of surface water bodies. These standards are based on the spatial and temporal evolution of certain variables and pollutants that might pose a significant risk to the environment (temperature, salinity, fecal contamination, nutrients, dissolved oxygen, pH, etc.). The concentration of these variables, which depends on the receiving waters, must be within a range of values to protect the environment and the human health. More generally, not only the exceedance of a given threshold should be controlled, but also the frequency and duration of those exceedance events. Concentration-duration-frequency curves can be used to define the thresholds that should not be exceeded to guarantee a correct environmental status [1–3].

In this context, numerical models have become valuable tools to help decision makers to evaluate alternative measures and solutions for the control of sewage spills. Some of the most known and used models for this purpose are the Water Quality Analysis Simulation Program (WASP) originally developed by [4,5], the QUAL2K model [6] and the CE-QUAL-W2 model [7,8]. Since these models assume different hydrodynamic approximations, their suitability is strongly dependent on the case under study. QUAL2K is a 1D steady flow model for rivers that assumes that the flow is well-mixed in the whole cross section. CE-QUAL-W2 is a 2D laterally averaged model that solves the hydrodynamics in the longitudinal and vertical direction, being therefore appropriate for rivers, reservoirs and estuaries well-mixed in the lateral direction. WASP can solve 1D, 2D and 3D problems with a large variety of pollutant types. However, this model needs to be linked with a hydrodynamic model to provide

accurate flow velocities and water elevation. Many other codes have been used in the scientific literature to model different water quality parameters in rivers and estuaries [9–13].

The model IberWQ was presented in [14] as a numerical tool for 2D water quality modeling in rivers and non-stratified estuaries. Its use is especially adequate to simulate the mixing of effluents when the flow is well mixed in the vertical direction. This model is integrated into the software package Iber and is freely available at http://www.iberaula.com.

IberWQ is not a suitable tool to deal with flows with vertical stratification or with relevant vertical gradients of velocity and concentration. In this context, it might be used as a complementary tool to 3D near-field models as VisualPLUMES [15] or CORMIX [16]. It could also be integrated in a hydrological-hydraulic modeling cascade approach [17] as a complementary tool of a basin-scale water quality hydrological model as SWAT [18].

In its original version, IberWQ could be used to compute the spatial and temporal evolution of: *Escherichia coli*, dissolved oxygen (DO), carbonaceous biochemical oxygen demand (CBOD), organic nitrogen (org-N), ammoniacal nitrogen (NH_3-N), nitrate-nitrite nitrogen (NO_3-N), water temperature (T) and salinity (S). The model solved one 2D depth-averaged advection-diffusion-reaction equation for each water quality variable. The diffusion and advection by the mean flow were computed using the velocity and water depth fields obtained from the hydrodynamic module of Iber, which solves the 2D shallow water equations using an explicit finite volume solver [19].

An important limitation of IberWQ was the fact that it did not simulate variables as phosphorus, phytoplankton and pH. However, these variables are relevant parameters that are often used to evaluate the environmental status of water bodies. Phosphorus is a necessary nutrient for algae and phytoplankton, which can be a limiting factor for plant growth. Neglecting phosphorus and phytoplankton prevented the application of the model to evaluate eutrophication due to the excessive concentration of nutrients (mainly nitrates and phosphates). pH is also an important water quality indicator that should be inside a certain range to preserve the ecosystems. This parameter determines the solubility of chemicals (e.g., metals are more soluble, and thus toxic, at lower pH values) and the amount of nutrients that can be utilized by aquatic life [20].

Another practical limitation of IberWQ was the computational burden required to compute real cases. Most studies imply the simulation of several water quality parameters over large spatial domains and rather long time intervals. The implementation of water quality early warning systems also requires very fast computational codes to immediately and effectively respond to hazardous events. As IberWQ solves the conservation and water quality transport equations on a fine numerical grid, it requires a high computational time (hours or days depending on the specific case). A natural way to overcome this limitation is the implementation of HPC (high performance computing) techniques. This approach had already been implemented for the hydrodynamic module of Iber [21] using GPU (graphical processing unit) parallelization techniques. Due to these developments, the new implementation is able to achieve speed-ups up to nearly two orders of magnitude.

This work presents an improved version of IberWQ model that addresses the previously mentioned limitations. The paper is organized as follows. First, the implementation of new water quality parameters and the parallelization of the code are briefly described in Section 2. New variables include organic phosphorus (org-P), phosphates (PO_4-P), phytoplankton, inorganic carbon (inorg-C), alkalinity and pH. Several examples showing the capabilities of the code and the speed-ups achieved in problems with different computational burden are presented in Section 3. Finally, the conclusions are reported in Section 4.

2. Model Structure and Equations

2.1. Model Structure

IberWQ consists of a set of routines that solve a series of unsteady 2D depth-averaged transport equations for different water quality parameters. This includes advection by the mean flow, diffusion

due to concentration gradients and reaction with other variables. Those routines are fully integrated within the shallow water model Iber [19], which solves the 2D Saint-Venant equations including turbulence. Thus, the resulting model for water quality and hydrodynamics is fully coupled.

Figure 1 shows the different species and biochemical reactions considered in the extension of the IberWQ module presented in this paper. The new components considered in relation to [14] are organic and inorganic phosphorus, phytoplankton, alkalinity and inorganic carbon. In addition, the pH is computed as a function of inorganic carbon and alkalinity. Phytoplankton is represented as the concentration of Chlorophyll-A (Chl-A), as commonly done in other water quality models [22]. Thus, in the case of computing phytoplankton the user must introduce the ratios between nitrogen, phosphorus, carbon and Chl-A in order to evaluate the corresponding biochemical reactions.
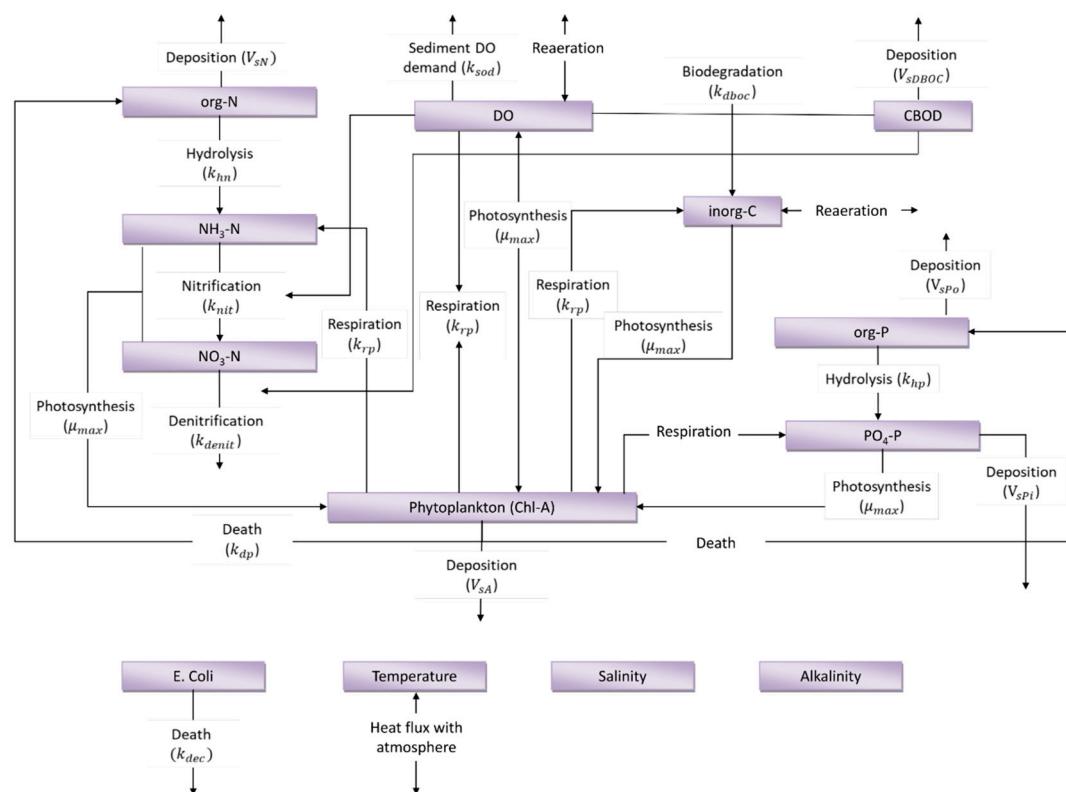
Figure 1. Model structure. Solid arrows indicate interactions between variables due to different biochemical processes.

Inorganic carbon includes carbonates (CO_3^{2-}), bicarbonates (HCO_3^{-}) and carbon dioxide (CO_2). The three species were modelled together as total inorganic carbon dissolved in water. A posteriori estimation of the concentration of each species was done as a function of pH (specific equations are presented in Appendix B).

The whole model included therefore 14 water quality parameters that are commonly used to evaluate the environmental status of rivers and estuaries (the 13 variables shown in Figure 1 plus pH). The different model components can be activated or deactivated depending on the problem to be solved. The kinetic constants of most of the biochemical reactions depend on water temperature and salinity. In case that the temperature field is computed by the model, additional atmospheric input data (time series of net solar and atmospheric radiation, air temperature, atmospheric relative humidity and wind velocity) must be introduced by the user.

2.2. Model Equations

A generic 2D depth-averaged advection-diffusion equation with reaction terms is solved for each of the species considered in the model. In order to compute the advection terms, the water quality module is linked to a hydrodynamic module that solves the 2D shallow water equations, given by:

$$\frac{\partial h}{\partial t} + \frac{\partial q_x}{\partial x} + \frac{\partial q_y}{\partial y} = 0 \tag{1}$$

$$\frac{\partial q_x}{\partial t} + \frac{\partial}{\partial x}\left(\frac{q_x^2}{h}\right) + \frac{\partial}{\partial y}\left(\frac{q_x q_y}{h}\right) = -gh\frac{\partial(h + z_b)}{\partial x} - g\frac{n^2}{h^{\frac{7}{3}}}|q|q_x - \frac{1}{2}\frac{gh^2}{\rho}\frac{\partial\rho}{\partial x} \tag{2}$$

$$\frac{\partial q_y}{\partial t} + \frac{\partial}{\partial x}\left(\frac{q_x q_y}{h}\right) + \frac{\partial}{\partial y}\left(\frac{q_y^2}{h}\right) = -gh\frac{\partial(h + z_b)}{\partial y} - g\frac{n^2}{h^{\frac{7}{3}}}|q|q_y - \frac{1}{2}\frac{gh^2}{\rho}\frac{\partial\rho}{\partial y} \tag{3}$$

$$\frac{\partial}{\partial t}(h\,C_i) + \frac{\partial}{\partial x}(q_x C_i) + \frac{\partial}{\partial y}\left(q_y C_i\right) = \frac{\partial}{\partial x}\left(h\,v_e\frac{\partial C_i}{\partial x}\right) + \frac{\partial}{\partial y}\left(h\,v_e\frac{\partial C_i}{\partial y}\right) + S_i h \tag{4}$$

where h is the water depth, z_b is the bed elevation (topography), (q_x, q_y) are the two components of the unit discharge, $|q|$ is the modulus of the unit discharge, C_i is the depth-averaged concentration of the species i (the available species are shown in Figure 1), S_i is a generic reaction term for the species i, ρ is the density of water, g is the gravity acceleration, n is the Manning coefficient and v_e is the effective viscosity computed using a depth-averaged turbulence model. The available turbulence models include the $k - \varepsilon$ model, a mixing-length model and a parabolic model, and are described in detail in [23]. If the baroclinic pressure is activated in the simulation, the water density is computed from the salinity and temperature fields. For this, the International One Atmosphere Equation of State of Seawater [24] is employed. The total number of partial differential equations solved (n_{eq}) depends on the number of species considered on each specific test case (n_i), and it is equal to $n_{eq} = 3 + n_i$.

The biochemical reactions considered for each species were modeled with the source terms S_i. Each solid arrow in Figure 1 represents a biochemical reaction and was modeled as a first order reaction [22,25]. The formulations used to compute the different source terms are basically the same as those used in well-known and extendedly used models such as WASP, QUAL-2K and CE-QUAL-W2, which follow [22]. The details of the equations implemented in IberWQ are given in Appendices A and B, while the model constants to be introduced by the user are detailed in Appendix C.

2.3. Numerical Solver

The 2D depth-averaged transport equations for each species were solved with an unstructured finite volume solver, using a computational grid formed by triangular and quadrilateral elements. The same finite volume mesh was used to solve the transport equations of the water quality species and the hydrodynamic equations. Thus, the water depth (h) and the unit discharges (q_x, q_y) needed to compute the advective fluxes in the scalar transport equations were obtained from the hydrodynamic module of Iber. This module solved the 2D Saint Venant equations using an upwind Godunov scheme with the approximate Riemann solver of Roe. In order to transfer the water velocity and depth from the hydraulic module to the water quality advection–diffusion equations, the mass conservative scheme detailed in [26] was used.

The advective terms can be discretized with either a first order upwind scheme [27] or a second order upwind scheme. In particular, the Gamma scheme proposed in [28] was implemented in the solver to obtain second order accuracy in space. Using this approach only implies an increase on the CPU (Central Processing Unit) time of approximately 5–10%.

2.4. Parallelization

In order to address the limitations in terms of efficiency of the Iber model, a new implementation was developed. The new implementation, Iber+ onwards, was first presented in [21]. It was developed in C++ using the object-oriented paradigm to improve the modularity and the maintenance of the source code. Iber+ was parallelized for shared memory systems using OpenMP and for GPUs using the Nvidia CUDA (Compute Unified Device Architecture) [29] platform. The GPUs are a cost-effective solution to accelerate numerical models without the necessity of expensive HPC facilities. GPUs are a solution available for servers, workstations and laptops. GPU computing is attractive not only to reduce the computational time of the simulations but also to improve the time spent on design and test cases. There are numerous cases of success in the literature for both meshless [30] and mesh-based [31–34] codes. Iber+ features both CPU and GPU parallelization, however the focus of the implementation is on GPU computing, so in this study only the GPU implementation is analyzed. Iber+ is included in the official package of Iber since version 2.5 and can be downloaded without any cost at http://www.iberaula.com. The new developments presented in this paper will be included in the next release.

Originally developed exclusively for computer graphics, the GPUs feature a highly parallel architecture that provides several TFLOPS (10^{12} floating point operations per second) of computing power in a single card [35]. This high-throughput hardware can be employed for general applications due to the GPGPU (general processing graphics processing unit) APIs (application programming interface), being Nvidia CUDA one of the most popular in scientific applications. CUDA exposes the GPUs capabilities to programmers due to an extension to common programming languages like C, C++ or Fortran. These extensions provide three abstractions to programmers: the thread hierarchy, shared memory and barrier synchronization [29].

The Nvidia GPUs feature a SIMT (single instruction multiple thread) architecture that issues threads in groups of 32 named warps. Each warp is executed in parallel in a SM (stream multiprocessor). All the threads in a warp execute the same instruction. If a branch instruction is processed, a divergence starts in this case, the threads that follow one of the paths are stalled while the others are executed and vice versa. Once the branch is finished, all the threads converge again. It is essential to consider this peculiarity of the GPUs in order to achieve a high throughput. The algorithms should be revised to avoid unnecessary branching and reorganize data to avoid divergence.

Discrete GPUs have their own memory, independent of the system memory. Data transfers should be made from the system memory to the GPU's memory, usually through the PCI-Express bus. These data transfers can be a potential bottleneck due to the limited bandwidth and higher latency compared with the system memory accesses. Figure 2 shows the flowchart of Iber+, the memory transfers have been reduced as much as possible. The problem data is transferred before the simulation starts and when it is necessary to write it to the hard disk. Only the current timestep of the simulation is transferred on every loop iteration of the simulation. The write of the simulation state to the hard disk can suppose an important part of the total run time, depending on the case and system configuration. To reduce these times, data is written to disk in background by a thread separated from the main execution thread. On the other hand, the data locality is essential in modern architectures to reduce the cache miss rate and thus improve performance. For this, the elements of the mesh are reordered using a space-filling curve, more specifically the Hilbert curve [36] was used for this purpose.
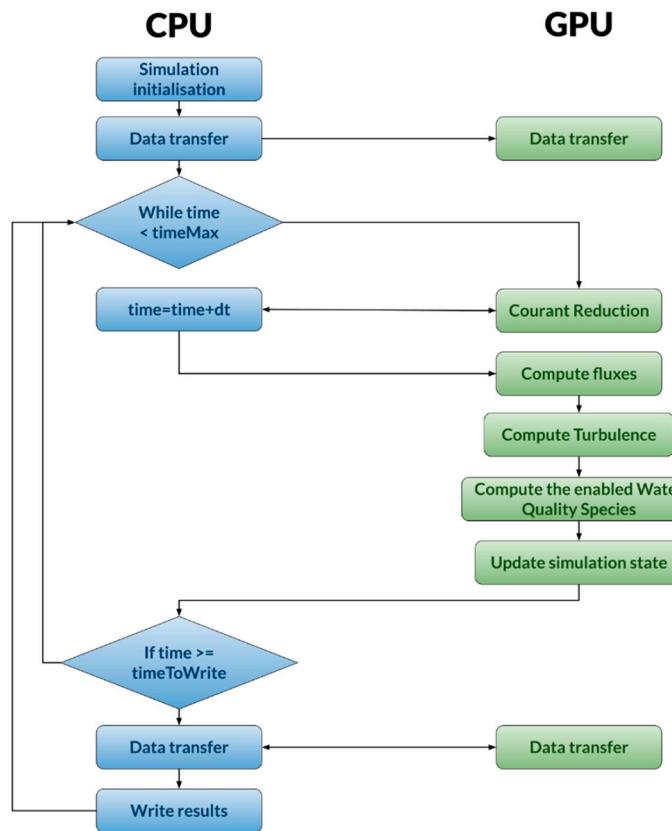
Figure 2. Flowchart of the Iber+ implementation.

Global synchronization in GPUs is expensive, making some algorithms like reductions more complicated than their CPU counterpart. In order to implement reductions efficiently, Iber+ employs the library Nvidia CUB (CUDA unbound) [37].

Modern GPUs have significant performance differences when working with single or double precision. The performance ratio single/double precision may vary from 1/2 for the Nvidia Tesla V100 [35] targeted to the HPC market up to 1/32 for the Nvidia RTX 2080ti [38] targeted for the consumer market. For this reason, most of the computations are performed in single precision. The use of double precision is restricted to potentially sensitive variables like the simulation time or the computation of distances.

For the implementation of the water quality module, an object-oriented approach was used. In this way, new species can be easily added by implementing an abstract class without significantly affecting the existing code.

3. Test Cases

In this section, four test cases were analyzed to illustrate the capabilities and performance of IberWQ and its GPU implementation. For a detailed validation of the test cases the reader is addressed to previous publications from the authors [14]. More information about the data sources used for the cases analyzed can be found in Appendix F. All the cases were run in a workstation equipped with an AMD Ryzen 7 2700X processor, 32 GB of RAM (random access memory) and an Nvidia RTX 2080ti graphics card. The run time measurements were carried out using the execution time corresponding to the simulation loop.

3.1. Faecal Contamination in a Coastal Estuary

3.1.1. Description

The first test case simulated the concentration of fecal contamination in a coastal estuary. This is a typical application of 2D depth-averaged water quality models [9,13,39–43] in which only one species, *E. coli*, needs to be computed.

The model was applied to compute the concentration of *E. coli* in the coastal estuary of Ferrol, located in the NW of Spain (Figure 3). The estuary extends over 32 km^2 with a relatively narrow and elongated shape. The total length of the estuary is 17 km, while its width varies from 400 m to a couple of kilometers in its widest part. Tidal ranges vary from 0.9 (neap tides) to 4.6 m (spring tides), being freshwater inflows from rivers negligible. As we are using a 2D depth-averaged model, we assumed well-mixed conditions over the water depth, which are characteristic of the autumn, winter and spring seasons in this estuary [39].
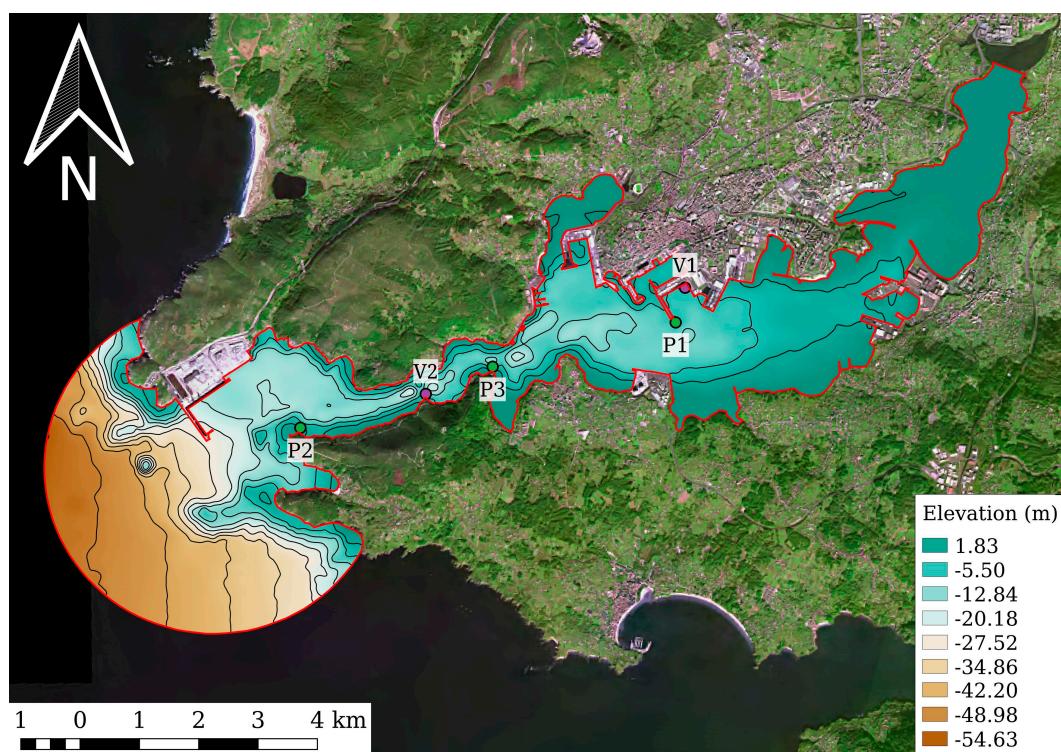
Figure 3. Bathymetry of the Ferrol estuary (NW, Spain). The sewage discharges are located at V1 and V2, whilst the control points are placed at P1, P2 and P3.

Approximately six spring semi-diurnal tidal cycles were simulated (3 days), with an approximate tidal range of 4.2 m. Two continuous sewage spills at different locations were considered. The first one (V2 in Figure 3) is located in the narrowest region of the estuary, where the high water velocities activate a rapid mixing with the receiving waters. The second one (V1 in Figure 3) is located in the inner part of the estuary, where mixing and dispersion were slower due to the lower water velocity. Both discharges were characterized by an *E. coli* concentration of 10^7 cfu/100 mL and a sewage discharge of 0.1 m^3/s, both values were constant in time. Even though the model included the possibility of using the formulation of Mancini to evaluate the degradation of *E. coli*, in this case we had fixed the value of the degradation constant to 5.5 days^{-1}, following previous studies performed in this estuary [44].

To perform the computations, the estuary was discretized with an unstructured computational grid with 143,993 elements (with an average element size of 222 m^2). The levels of *E. coli* were sampled every 600 s with Iber and Iber+ at control points P1, P2 and P3.

3.1.2. Results

The *E. coli* concentration was sampled at three control points (P1, P2 and P3 in Figure 3). Figure 4 compares the spatial distribution of *E. coli* in the estuary at different time steps for the CPU and GPU implementations. The results obtained with Iber and Iber+ were virtually identical. The time series of concentration at the three control points are shown in Figure 5 where no significant differences could be observed. Furthermore, the time series of concentration obtained with the GPU and CPU implementations were statistically compared in terms of the normalized root mean square deviation (in %):

$$\text{NRMSD} = 100 \times \frac{\text{RMSD}}{y_{\max} - y_{\min}} \quad \text{RMSD} = \sqrt{\frac{\sum_{i=1}^{N}(x_i - y_i)^2}{N}} \tag{5}$$

where x_i and y_i are the values computed with Iber+ and Iber respectively for any variable of interest.
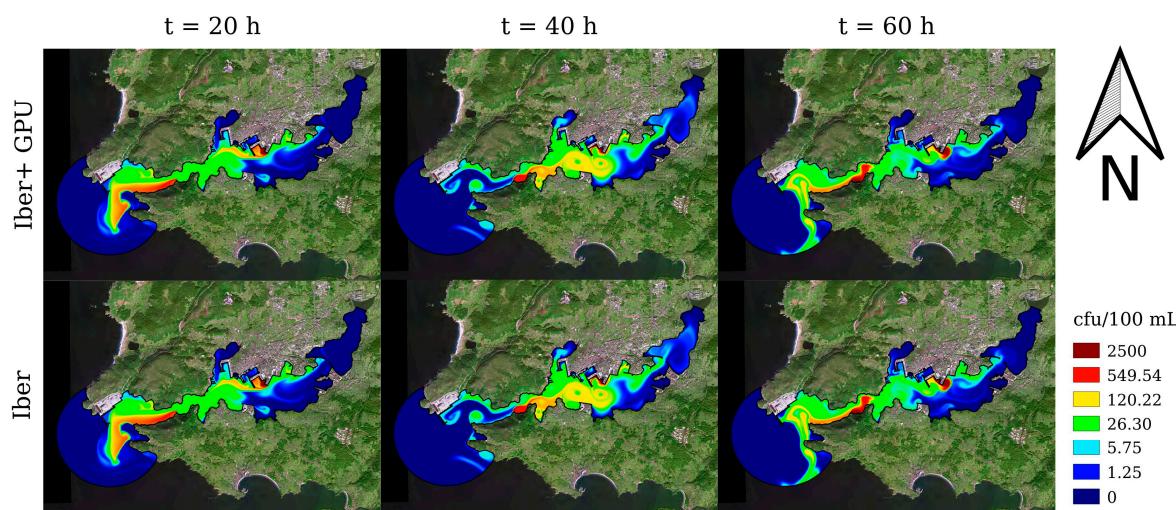
Figure 4. Comparison of *Escherichia coli* concentrations at different time steps obtained with Iber (CPU) and Iber+ (GPU).
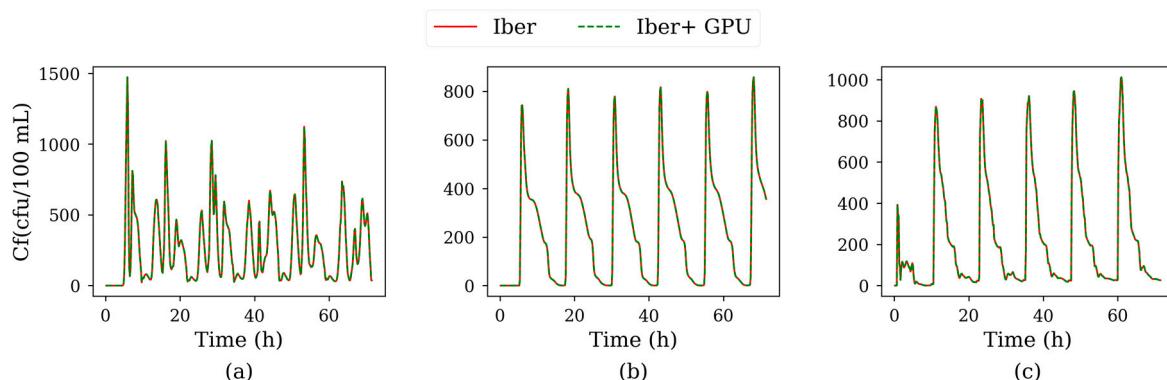
Figure 5. Time series of *E. coli* concentration at control points P1 (**a**), P2 (**b**) and P3 (**c**), for the first test case.

The average error, ⟨NRMSD⟩ = 0.07%, was calculated taking into account all variables at all control points, supporting the good resemblance between series suggested by the visual comparison.

Table 1 shows the performance measurements for the first test case. This was the best case for GPU implementation because it had the largest mesh (143,993 elements) among the cases under study. GPU computing implied a certain overhead, mainly due to memory transfers and kernel launch latency. The higher the GPU workload the less significant the overhead will be. In this case Iber took more

than 16 h to complete the simulation whilst Iber+ could run the same problem in less than 6 min. This implied a speedup of 181.

Table 1. Performance measurements obtained for the first test case (estuary of Ferrol).

Model	Run Time (s)	Time per Step (ms)	Millions of Cells per Second	Speedup vs. Iber
Iber	60,545	66.8	2.2	1
Iber+ GPU	335	0.4	400.0	181

3.2. Organic Matter Contamination in an Estuary

3.2.1. Description

In the second test case the concentration of dissolved oxygen (DO) and carbonaceous biochemical oxygen demand (CBOD) were computed in the estuary of A Coruña, located in the NW of Spain (Figure 6). The estuary has an area of 26 km^2. The outer part of the estuary is relatively deep (with depths of the order of 20 m and maximum values of circa 30 m in the mouth), while the inner part, where the sewage spills are located, is relatively shallow (with maximum water depths of the order of 5 m at high tide and many dry regions at low tide).
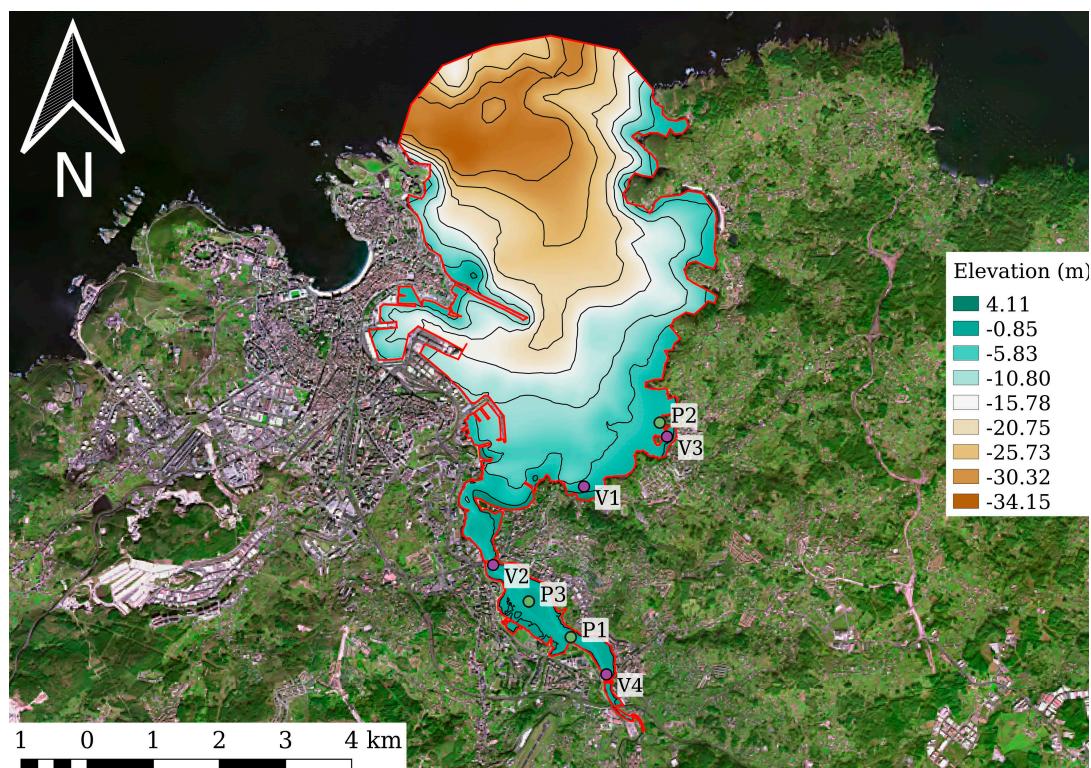
Figure 6. Bathymetry of the estuary of A Coruña (NW, Spain). The sewage spills are located at V1, V2, V3 and V4. The control points are placed at P1, P2 and P3.

The physical time simulated for this case extends over 8 tidal cycles (4 days) in which the tidal range moves from neap tides (with a tidal range of 1.2 m) to spring tides (with a tidal range of 3.5 m). The river discharge at the upstream boundary of the model is equal to 20 m^3/s, which corresponds to its annual average value. Four discontinuous sewage spills of CBOD were considered, all of them located in the inner estuary (Figure 6). The mean depth at the spills varied within 3 m and 4 m, which were relatively shallow depths. All the spills were characterized by the same uniform concentration of

CBOD and DO (511 mg/L of CBOD and 3.5 mg/L of DO). The water discharges were discontinuous and different from one spill to another, with maximum values ranging from 0.06 to 0.5 m^3/s. The CBOD degradation rate at 20 °C was set to $k_{dboc} = 0.23$ day^{-1}.

The domain was discretized using a computational grid of 50,329 elements. The mesh size was variable over the estuary. A finer resolution was used in the inner part of the estuary (element sizes of 10 m) since in this region the spills were located and unsteady wet-dry tidal fronts appeared. The size of the elements in the mouth of the estuary was approximately 180 m. The levels of DO and CBOD were sampled every 300 s using both Iber and Iber+ at control points P1, P2 and P3.

3.2.2. Results

In the second test case, the levels of CBOD and DO were sampled at four control points shown in Figure 6. In particular, the time series of CBOD computed at these control points are shown in Figure 7 with identical results for both implementations. Regarding DO concentration (Figure 8), very small differences could be appreciated at control point P2, where the peak values computed with Iber+ were slightly higher than those computed with Iber. Statistically, the differences were not significant with a negligible average error (⟨NRMSD⟩ = 0.32%, covering all points and variables), although higher than in previous case.
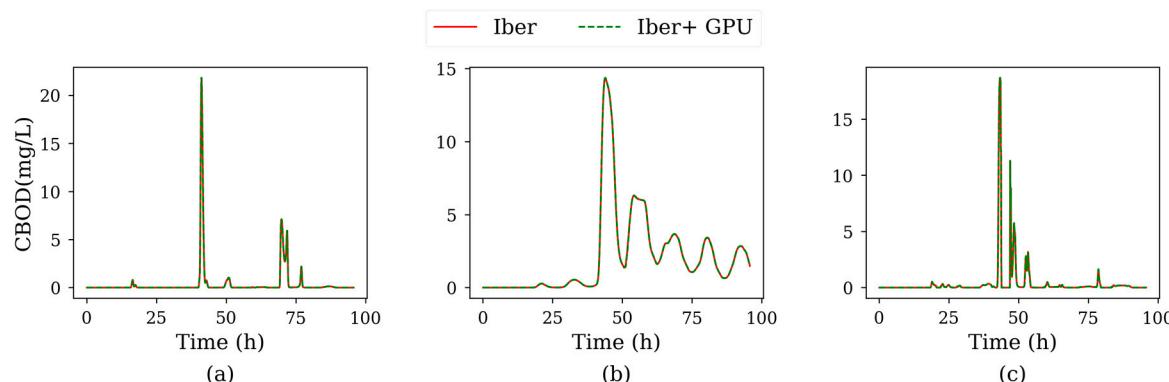
**Figure 7.** Time series of carbonaceous biochemical oxygen demand (CBOD) for Test 2 at control points P1 (**a**), P2 (**b**) and P3 (**c**).
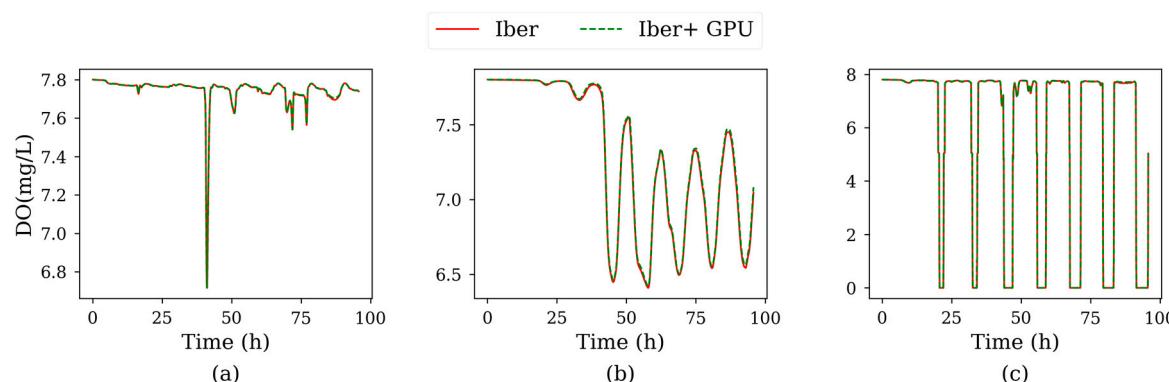
Figure 8. Time series of DO for Test 2 at control points P1 (**a**), P2 (**b**) and 3 (**c**).

This case employed a mesh with 50,329 elements, about a third of the mesh elements used in the first test case. As shown in Table 2, Iber+ was able to perform the simulation 61 times faster than Iber. In spite of having a lower number of elements, Table 2 shows that the real time employed by Iber+ to compute a simulation time step was longer than in the previous test. This can be explained by two reasons. First, this case had to compute two water quality species instead of one, causing a larger

kernel launch overhead. Second, the size of the problem was not big enough to saturate the GPU capacity. Thus, this case was not able to take full advantage of the GPU computing capacity, as it happened in the previous case.

Table 2. Performance measurements obtained for the second test case (estuary of A Coruña).

Model	Run Time (s)	Time per Step (ms)	Millions of Cells per Second	Speedup vs. Iber
Iber	31,615	28.8	1.8	1
Iber+ GPU	522	0.5	107.1	61

3.3. Combined Sewer Overflows in a River Miño Reach

3.3.1. Description

This test case was extracted from the study presented in [1]. In that work, a 2D water quality model is used as a fundamental part of an integrated modeling approach for the design of the sewer network of the city of Lugo (Spain). Here, the impact of combined sewer overflows (CSO) in the river Miño was evaluated by means of the Environmental Quality Standards (EQS) presented in the Urban Pollution Manual [44]. The application of EQS requires an efficient water quality model since usually a large number of simulations must be run over long periods of time.

The objective variables computed in [44] to apply the EQS are the concentrations of dissolved oxygen and ammonia. The evaluation of these variables requires the computation of the five following species: org-N, NH_3-N, NO_3-N, DO and CBOD.

In the example presented here, two days of discontinuous sewage discharges at three different locations along the river were modeled. All the spills were characterized by the following concentrations: 102 mg/L of CBOD, 5 mg/L of org-N, 1.5 mg/L of NH_3-N and 4 mg/L of DO. The concentration of nitrates (NO_3-N) in the spills is negligible. The spill discharges were discontinuous in time and different from each other, with maximum values of discharge ranging from 0.18 to 1.5 m^3/s. The ambient river concentration of the water quality species was negligible with the exception of the DO concentration, which was set to 9 mg/L. Those ambient values were imposed as initial and upstream boundary conditions in the model. The river discharge, which was obtained from a Water Quality Automatic Information System (SAICA) located upstream the river reach under study, varied from 40 to 60 m^3/s during the two days of computation. The same values of the reaction kinetic constants proposed in [45] were used in the simulations, namely (all values at 20 °C): CBOD degradation rate equal to $k_{dboc} = 0.35$ day^{-1}, org-N hydrolysis rate equal to $k_{hn} = 0.20$ day^{-1}, nitrification rate equal to $k_{nit} = 0.50$ day^{-1} and denitrification rate equal to $k_{denit} = 0.05$ day^{-1}.

Figure 9 shows the bathymetry of the river reach under study and the location of the sewage spills. The computational domain extends over 8 km of river, with an average width of 70 m and a total extension of 0.57 km^2. The domain was meshed with 8763 elements with an average size of 65 m^2. The levels of the simulated species were sampled every 600 s with both models at the control points P1, P2, P3 and P4.
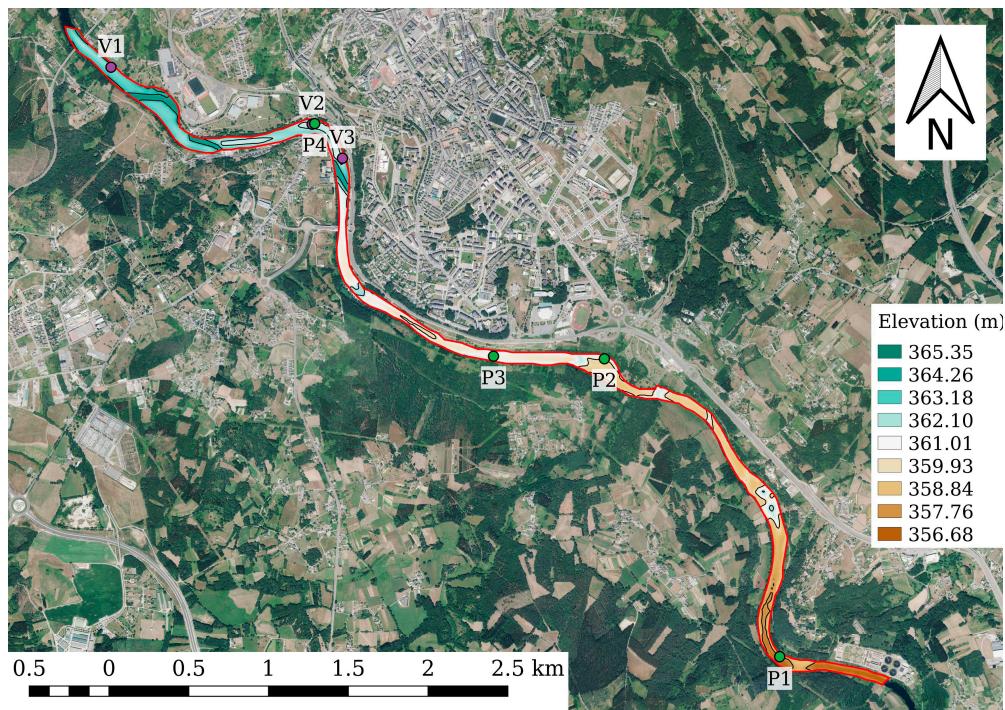
Figure 9. Bathymetry of the Miño River as it passes through the city of Lugo (NW, Spain). The sewage spills are located at points V1, V2 and V3. The control points are located at P1, P2, P3 and P4.

3.3.2. Results

In the third case, five species (org-N, NH$_3$-N, NO$_3$-N, DO and CBOD) were sampled at the four control points indicated in Figure 9. For the sake of clarity, only the results obtained at P3 are shown in Figure 10, while the rest of time series are shown in Appendix D. In this simulation, the average error ($\langle$NRMSD$\rangle$ = 0.04%) was negligible and smaller than in the previous two cases.
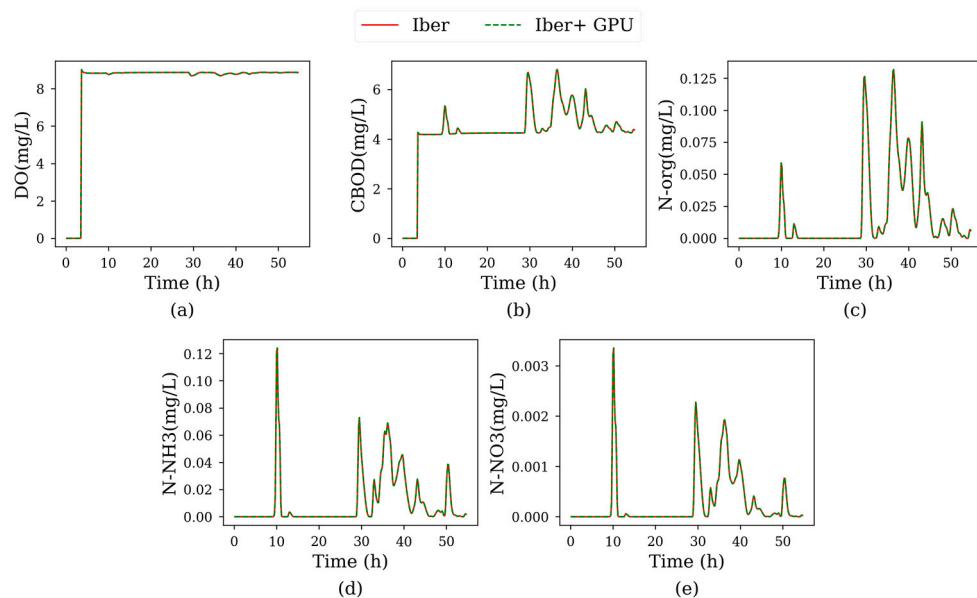
Figure 10. Time series sampled at control point P3 for dissolved oxygen (**a**), carbonaceous biochemical oxygen demand (**b**), organic nitrogen (**c**), ammoniacal nitrogen (**d**) and nitrite-nitrate nitrogen (**e**).

In this case, Iber+ completed the simulation 29 times faster than Iber (Table 3). This speedup was considerably lower than the one obtained in the previous two cases due to the relatively small size of the mesh, with just 8763 elements. Compared to the previous case, the time needed to process a single time step was lower whilst the number of cells processed per second was much lower. This indicates the difficulty of taking full advantage of the GPU resources in cases with a small number of elements, such as the present one. Nevertheless, even with a small mesh, Iber+ ran significantly faster (29 times) than the non-parallelized CPU version.

Table 3. Performance measurements obtained for the third test case (Miño River, NW Spain).

Model	Run Time (s)	Time per Step (ms)	Millions of Cells per Second	Speedup vs. Iber
Iber	3054	8.6	1.0	1
Iber+ GPU	105	0.3	30.2	29

3.4. Effluent Discharge from a Wastewater Treatment Plant

3.4.1. Description

In this case an effluent discharge from a wastewater treatment plant into a river located in the south of Spain was modeled. The water quality species computed in this case are DO, CBOD, org-N, NH_3-N, NO_3-N, org-P, PO_4-P, inorg-C, alkalinity, salinity and pH. Thus, 10 transport equations were solved in addition to the three Saint Venant equations (as mentioned in Section 2, the pH did not have an associated transport equation, instead its value was computed at each mesh element from the values of alkalinity and inorg-C).

The river reach modeled (Figure 11) is 18 km long and approximately 200 m wide, with a total area of 4 km^2. The downstream boundary of the reach modeled is located 50 km upstream the river mouth into the ocean and is affected by the ocean tidal elevation, although tide does not propagate all the way up to the upstream boundary. Thus, a tidal wave with a range of 1.0 m was imposed at the downstream boundary while a constant river discharge of 3.6 m^3/s was imposed at the upstream boundary. The ambient concentrations of the water quality components imposed at the upstream boundary were the following: 8 mg/L of DO, 0 mg/L of CBOD, 2.8 mg/L of org-N, 1.3 mg/L of NH_3-N, 4.2 mg/L of NO_3-N, 0.3 mg/L of org-P, 0.2 mg/L of PO_4-P, 90 mg/L of inorg-C, 346 mg/L CO_3Ca (alkalinity) and 0.8 mg/L of salt. These values correspond to mean winter conditions in this river reach.

A continuous spill from a waste water treatment plant (Figure 11) was modeled, with a discharge of 2.1 m^3/s and the following concentration of the water quality variables considered in this case: 5 mg/L of DO, 23.6 mg/L of CBOD, 23.9 mg/L of org-N, 28.4 mg/L of NH_3-N, 0.3 mg/L of NO_3-N, 0.8 mg/L of org-P, 0.4 mg/L of PO_4-P, 97 mg/L of inorg-C, 392 mg/L CO_3Ca (alkalinity) and 1 mg/L of salt.

The following kinetic constants were used in the simulations (all values at 20 °C): CBOD degradation rate ($k_{dboc} = 0.23$ day^{-1}), org-N hydrolysis rate ($k_{hn} = 0.20$ day^{-1}), nitrification rate ($k_{nit} = 0.20$ day^{-1}), denitrification rate ($k_{denit} = 0.05$ day^{-1}) and organic phosphorus hydrolysis rate ($k_{hp} = 0.20$ day^{-1}).

The simulated physical time was 3 days, using an unstructured mesh of 90,406 elements. At the end of the simulation a steady state was achieved for the whole reach. All the simulated species were sampled every 600 s with both Iber and Iber+ at control points P1, P2, P3 and P4.
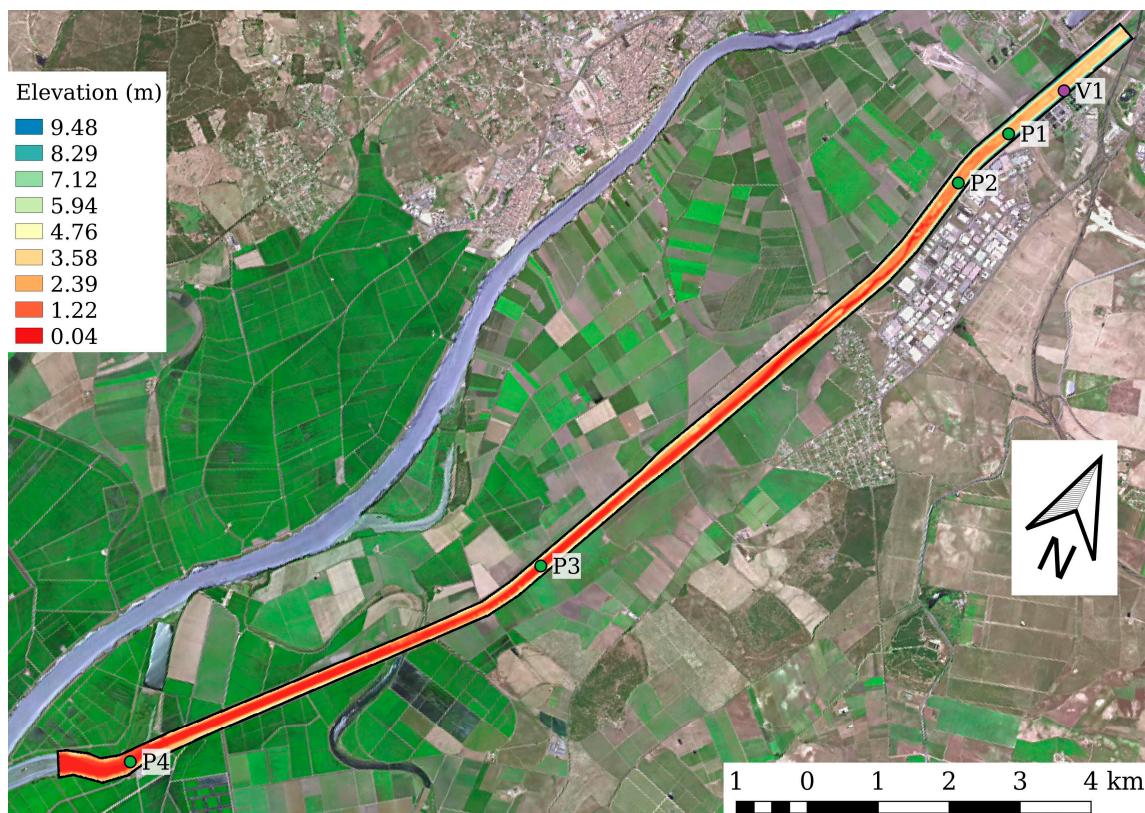
Figure 11. Bathymetry of the river reach in Test Case 4. The sewage spill of the Wastewater Treatment Plant (WWTP) is located at point V1. Control points are located at P1, P2, P3 and P4.

3.4.2. Results

In the last case, ten different species were simulated (DO, CBOD, org-N, NH_3-N, NO_3-N, org-P, PO_4-P, inorg-C, alkalinity, salinity and pH). Time series of concentrations were taken at Points P1 to P4 as indicated in Figure 11. The zone of the sewage spill is shown in detail in Figure 12. The time series of concentration at control point P3 are shown in Figure 13. For the sake of clarity, the results at the other control points were included in Appendix E. The results provided by Iber and Iber+ were almost identical in all cases, confirmed by the negligible average error ($\langle$NRMSD$\rangle$ = 0.16%) as obtained for the rest of the cases.
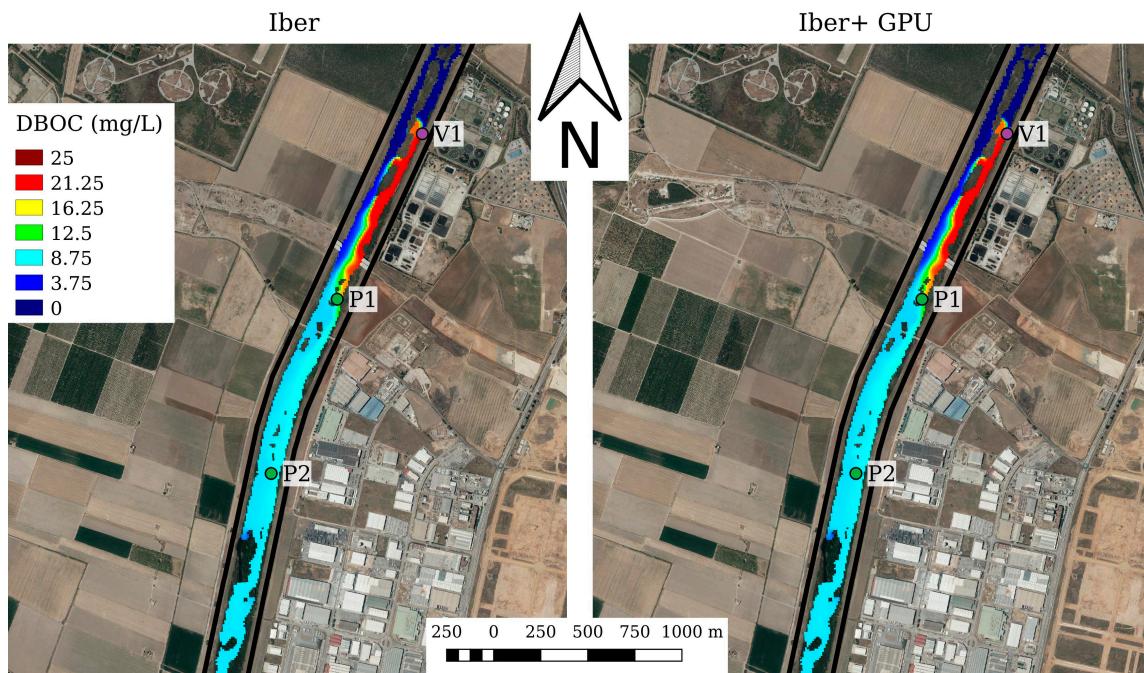
Figure 12. Detail of the concentration of CBOD downstream the sewage spill at t = 60 h using both Iber and Iber+ implementations.
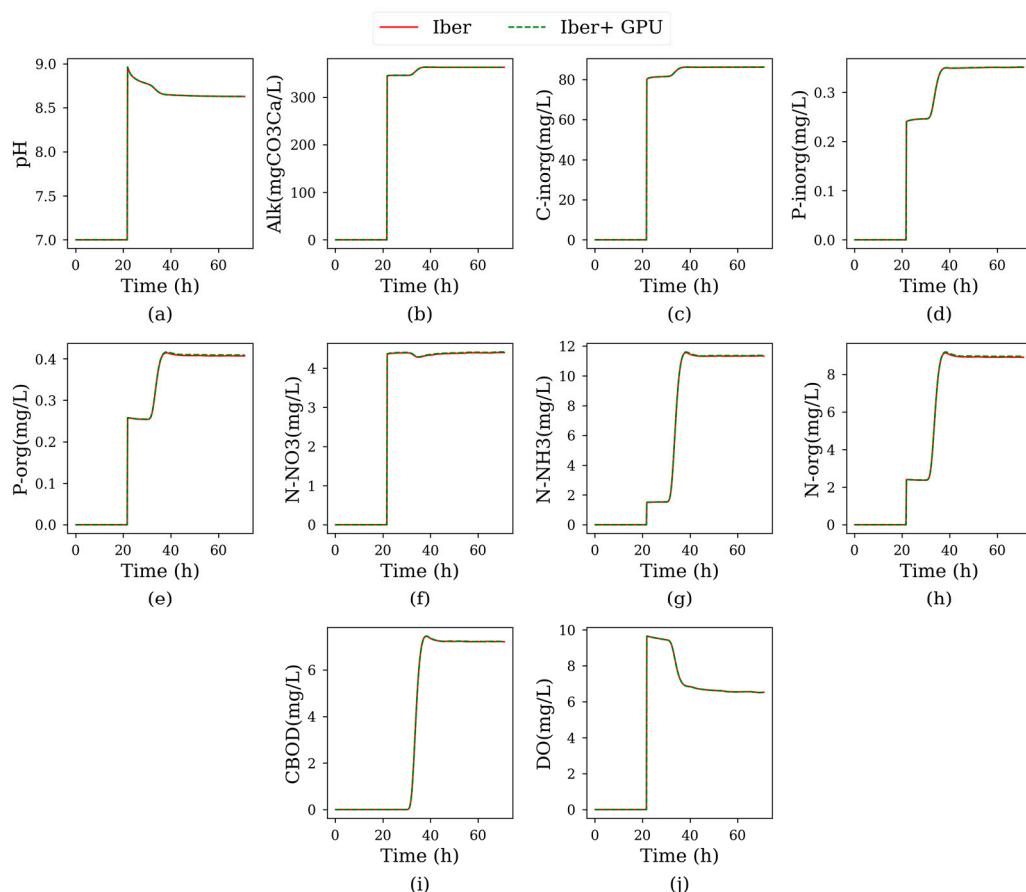
Figure 13. Time series for the analyzed species in Test 4, sampled at control point P3.

Table 4 shows the performance measurements for this case, where a mesh of 90,406 elements (halfway between Test 1 and Test 2) was considered. Iber+ reached a speedup of 92 with respect to Iber. Comparing with Test 2, the new mesh was almost double the size of the former one and ten species were simulated instead of two. As a result of this, the time required to process a single time step was almost 3.5 times longer. However, the number of mesh cells processed per second was only twice lower. In this case, Iber+ could take more advantage of the GPU processing capacity due to the bigger mesh size.

Table 4. Performance measurements obtained for the fourth test case.

Model	Run Time (s)	Time per Step (ms)	Millions of Cells per Second	Speedup vs. Iber
Iber	44,367	146.5	0.6	1
Iber+ GPU	482	1.6	56.9	92

4. Conclusions

Numerical models are useful tools for the evaluation of the environmental status of water bodies. However, it should be noted that these kinds of tools rely on a set of simplifications and are much simpler than the real systems, whose complexity cannot be fully reproduced by the models. Hence, they can provide wrong results, especially when the managers or the decision makers are not aware of those limitations. The parametrization of the models is then a crucial task in order to obtain precise results, which requires an exhaustive sensitivity analysis. In the particular case of Iber+, its 2D nature makes it only applicable to rivers and non-stratified estuaries.

On the other hand, the main advantage of 2D models like Iber+ is their affordable execution time, which makes them especially valuable for fast response purposes when compared with 3D models. However, its application is hindered by the high spatial resolution required for water quality studies in long river reaches or large estuaries.

This paper presented an improved version of a two-dimensional depth-averaged water quality model, whose efficiency was improved by implementing HPC techniques based on GPU parallelization. The model considered the water quality parameters most commonly used in the environmental assessment of receiving waters, including dissolved oxygen, CBOD, organic nitrogen, ammonia, nitrates, organic phosphorus, phosphates, pH, salinity and temperature. The implementation of the code ran in NVIDIA GPUs that are commonly installed in standard laptop and desktop PCs. In the test cases presented here, speedups in computational time between 29 and 181 were obtained when compared with the non-parallelized implementation, keeping the accuracy of the original model. The code will be integrated in the software package Iber, making it freely available.

In summary, the model presented three key features that make it very attractive and useful for the scientific and engineering community: simulation of the most common water quality parameters, high-performance computing in standard PCs and free availability.

Appendix A. Reaction Terms

The following source terms were included in the advection–diffusion equation to model the biochemical reactions between the water quality species considered in the model, as represented in Figure 1. Salinity and alkalinity were considered as conservative variables with no source terms.

Chlorophyll-A

$$S_{ChlA} = S_{ChlA}^{photo} - S_{ChlA}^{resp} - S_{ChlA}^{death} - S_{ChlA}^{dep} \tag{A1}$$

$$S_{ChlA}^{photo} = \mu_{max} F_L min(F_N, F_P, F_C) \, 1.047^{(T-20)} \, C_{ChlA}$$

$$S_{ChlA}^{resp} = k_{rp} \, F_{oxp} \, 1.047^{(T-20)} \, C_{ChlA}$$

$$S_{ChlA}^{death} = k_{dp} \, 1.047^{(T-20)} \, C_{ChlA}$$

$$S_{ChlA}^{dep} = \frac{V_{sA}}{h} C_{ChlA}$$

$$F_N = \frac{C_{NH_3\text{-}N} + C_{NO_3\text{-}N}}{C_{NH_3\text{-}N} + C_{NO_3\text{-}N} + k_{snp}} \quad F_P = \frac{C_{PO_4\text{-}P}}{C_{PO_4\text{-}P} + k_{spp}} \quad F_C = \frac{C_{inorg\text{-}C}}{C_{inorg\text{-}C} + k_{scp}}$$

$$F_L = \frac{1}{k_e \, h} Ln\left(\frac{k_{lp} + I_0}{k_{lp} + I \cdot e^{-k_e h}}\right)$$

$$F_{oxp} = \frac{C_{DO}}{C_{DO} + k_{sop}}$$

Nitrogen

$$S_{org\text{-}N} = -S_{org\text{-}N}^{hydro} - S_{org\text{-}N}^{dep} + r_{na} \, S_{ChlA}^{death} \tag{A2}$$

$$S_{NH_3\text{-}N} = S_{org\text{-}N}^{hydro} - S_{NH_3\text{-}N}^{nit} + r_{na} \, S_{ChlA}^{resp} - r_{na} \, P_{ap} \, S_{ChlA}^{photo}$$

$$S_{NO_3\text{-}N} = S_{NH_3\text{-}N}^{nit} - S_{NO_3\text{-}N}^{denit} - r_{na} \left(1 - P_{ap}\right) S_{ChlA}^{photo}$$

$$S_{org\text{-}N}^{hydro} = k_{hn} \, 1.047^{(T-20)} C_{org\text{-}N}$$

$$S_{org\text{-}N}^{dep} = \frac{V_{sN}}{h} C_{org\text{-}N}$$

$$S_{NH_3\text{-}N}^{nit} = k_{nit} \, 1.083^{(T-20)} F_n \, C_{NH_3\text{-}N}$$

$$S_{NO_3\text{-}N}^{denit} = k_{denit} \, 1.045^{(T-20)} F_{dn} \, C_{NO_3\text{-}N}$$

$$P_{ap} = \frac{k_{pa} C_{NH_3\text{-}N}}{k_{pa} C_{NH_3\text{-}N} + \left(1 - k_{pa}\right) C_{NO_3\text{-}N}}$$

$$F_n = \frac{C_{DO}}{k_{n1/2} + OD}, \, F_{dn} = \frac{k_{dn1/2}}{k_{dn1/2} + C_{DO}}$$

Phosphorus

$$S_{org\text{-}P} = -S_{org\text{-}P}^{hydro} - S_{org\text{-}P}^{dep} + r_{pa} \, S_{ChlA}^{death} \tag{A5}$$

$$S_{org\text{-}P}^{hydro} = k_{hp} \, 1.047^{(T-20)} \, C_{org\text{-}P}$$

$$S_{org\text{-}P}^{dep} = \frac{V_{sPo}}{h} C_{org\text{-}P}$$

$$S_{PO_4\text{-}P} = S_{org\text{-}P}^{hydro} - S_{PO_4\text{-}P}^{dep} + r_{pa} S_{ChlA}^{resp} - r_{pa} P_{ap} S_{ChlA}^{photo} \tag{A6}$$

$$S_{PO_4\text{-}P}^{dep} = -\frac{V_{sPi}}{h} C_{PO_4\text{-}P}$$

CBOD

$$S_{CBOD} = - S_{CBOD}^{deg} - S_{CBOD}^{dep} - 2.86 S_{NO_3\text{-}N}^{denit} \tag{A7}$$

$$S_{CBOD}^{deg} = k_{CBOD} \, 1.047^{(T-20)} \, F_{oxc} \, C_{CBOD}$$

$$S_{CBOD}^{dep} = \frac{V_{sCBOD}}{h} C_{CBOD}$$

$$F_{oxc} = \frac{C_{DO}}{k_{socf} + C_{DO}}$$

Dissolved Oxygen

$$S_{DO} = S_{DO}^{reair} - S_{CBOD}^{deg} - r_a S_{NH_3\text{-}N}^{nit} + r_{oa} S_{ChlA}^{photo} - r_{oa} S_{ChlA}^{resp} - S_{DO}^{DOS} \tag{A8}$$

$$S_{DO}^{reair} = k_{air} \, 1.024^{(T-20)} \, (\text{DO}_{\text{sat}} - C_{DO})$$

$$S_{DO}^{SOD} = \frac{k_{SOD}}{h}$$

$$k_{air} = k_{airh} + \frac{k_{airw}}{h}$$

$$k_{airh} = \begin{cases} 5.32 \dfrac{U^{0.67}}{h^{1.85}} & if & h < 0.61 \, m \\ 3.93 \dfrac{U^{0.5}}{h^{1.5}} & if & h > 3.45 \, U^{2.5} & \& & h > 0.61 \, m \\ 5.026 \dfrac{U}{h^{1.67}} & otherwise \end{cases}$$

$$k_{airw} = 0.728 \, V_{w10}^{0.5} - 0.317 \, V_{w10} + 0.0372 \, V_{w10}^2$$

Inorganic Carbon

$$S_{inorg\text{-}C} = S_{inorg\text{-}C}^{reair} + r_{cco} S_{CBOD}^{deg} + r_{cca} S_{ChlA}^{resp} - r_{cca} S_{ChlA}^{photo} \tag{A9}$$

$$S_{inorg\text{-}C}^{reair} = 0.923 \, k_{air} \, 1.024^{(T-20)} \left(CO2_{sat} - \alpha_0 \, C_{inorg\text{-}C} \right)$$

$$r_{cco} = \frac{1}{12} \frac{1}{r_{oc}} \quad r_{cca} = \frac{1}{12} r_{ca}$$

E. coli

$$S_{EC} = -k_{dec} \times EC \tag{A10}$$

$$k_{dec} = (0.8 + 0.02S) 1.07^{(T-20)} + 0.086 \frac{I_0}{k_e h} \left(1 - e^{(-k_e h)} \right)$$

Temperature

$$S_T = q_{rad} - q_{br} - q_{cond} - q_{evap} \tag{A11}$$

$$q_{br} = 0.97 \, \sigma \, T^4 \quad \sigma = 5.669 \times 10^{-8} \frac{W}{m^2 K^4}$$

$$q_{cond} = 0.47 \left(19 + 0.95 \, V_{w7}^2 \right) (T - T_{air})$$

$$q_{evap} = \left(19 + 0.95 \, V_{w7}^2 \right) (e_{water} - e_{air})$$

$$e_{water} = 4.596 \times exp\left(\frac{17.27 \, (T - 273.15)}{T - 35.7} \right) \quad e_{air} = RH \times 4.596 \times exp\left(\frac{17.27 \, (T_{air} - 273.15)}{T_{air} - 35.7} \right)$$

Appendix B. pH model

pH computations are based on the values of alkalinity and inorganic carbon. The pH model proposed by [45] was used. The model was based on the following equilibrium, mass balance and electroneutrality equations.

$$K_1 = \frac{[HCO_3^-][H^+]}{[H_2CO_3^*]} \quad K_2 = \frac{[CO_3^{2-}][H^+]}{[HCO_3^-]} \quad K_w = [H^+][OH^-]$$

$$C_{inorg\text{-}C} = [H_2CO_3^*] + [HCO_3^-] + [CO_3^{2-}]$$

$$Alk = [HCO_3^-] + 2[CO_3^{2-}] + [OH^-] - [H^+]$$

(A12)

where K_1, K_2 and K_w are equilibrium constants, Alk is the alkalinity (eq/L), $C_{inorg\text{-}C}$ is the concentration of inorganic carbon (mol/L), $[H_2CO_3^*]$ is the sum of carbon dioxide and carbonic acid dissolved in the water (mol/L), $[HCO_3^-]$ is the concentration of bicarbonate ions, $[CO_3^{2-}]$ is the concentration of carbonate ions, $[H^+]$ is the concentration of hydrogen ions and $[OH^-]$ is the concentration of hydroxyl ions. The equilibrium constants are computed as a function of water temperature as:

$$\log_{10}(K_w) = -\frac{4787.3}{T} - 7.1321 \log_{10}(T) - 0.010365 \times T + 22.8$$

$$\log_{10}(K_1) = -356.3094 - 0.0609196 \times T + \frac{21834.37}{T} + 126.8339 \log_{10}(T) - \frac{1684915}{T^2}$$

$$\log_{10}(K_2) = -107.887 - 0.0325285 \times T + \frac{5151.79}{T} + 38.9256 \log_{10}(T) - \frac{563713.9}{T^2}$$

(A13)

The values of $C_{inorg\text{-}C}$ and Alk are obtained from the solution of their respective depth-averaged transport equations. The previous system of five equations is solved as proposed in [44], by solving the following algebraic non-linear equation for $[H^+]$ at each mesh element:

$$(\alpha_1 + 2\alpha_2)\, C_{inorc\text{-}C} + \frac{K_w}{[H^+]} - [H^+] - Alk = 0$$

(A14)

with:

$$\alpha_0 = \frac{[H^+]^2}{[H^+]^2 + K_1\,[H^+] + K_1 K_2}$$

$$\alpha_1 = \frac{K_1\,[H^+]}{[H^+]^2 + K_1\,[H^+] + K_1\,K_2}$$

$$\alpha_2 = \frac{K_1\,K_2}{[H^+]^2 + K_1\,[H^+] + K_1\,K_2}$$

(A15)

The coefficients α_0, α_1 and α_2 represent respectively the fraction of inorganic carbon in form of carbon dioxide, bicarbonates and carbonates. Once the previous equation is solved at each mesh element, the pH can be computed as:

$$pH = -\log_{10}[H^+]$$

(A16)

and the concentration of carbon dioxide, bicarbonates and carbonates as:

$$C_{H_2CO_3^*} = \alpha_0 C_{inorg\text{-}C}$$

(A17)

Appendix C. Model Constants

Constant	Units	Suggested Values		Description
		Min	Max	
r_{na}	mg/mg	0.7	9.0	Ratio of Nitrogen to Chl-A in phytoplankton
r_{pa}	mg/mg	0.1	2.0	Ratio of Phosphorus to Chl-A in phytoplankton
r_{oa}	mg/mg	14	180	Ratio of Oxygen to Chl-A in phytoplankton
r_{ca}	mg/mg	-	-	Ratio of Carbon to Chl-A in phytoplankton
k_{rp}	1/day	0.04	0.8	Phytoplankton respiration rate
μ_{max}	1/day	1.0	3.0	Maximum photosynthesis rate
k_{dp}	1/day	0.05	0.5	Phytoplankton death rate
k_{pa}	-	0.0	1.0	Phytoplankton preference factor for ammonia
V_{sA}	m/day	-	-	Phytoplankton settling velocity
k_{snp}	mg/L	0.01	0.3	Nitrogen half-saturation constant for photosynthesis attenuation
k_{spp}	mg/L	0.001	0.05	Phosphorus half-saturation constant for photosynthesis attenuation
k_{scp}	mg/L	-	-	Carbon half-saturation constant for photosynthesis attenuation
k_{lp}	W/m^2	0.05	0.3	Light half-saturation constant for photosynthesis attenuation
k_{sop}	mg/L	-	-	Oxygen half-saturation constant for respiration attenuation
k_e	1/m	-	-	Light extinction coefficient in water
V_{sPo}	m/day	-	-	Organic phosphorus settling velocity
V_{sPi}	m/day	-	-	Inorganic phosphorus settling velocity
k_{hp}	1/day	0.01	0.7	Organic phosphorus hydrolysis rate at 20 °C
k_{nit}	1/day	0.01	1.0	Nitrification rate at 20 °C
k_{hn}	1/day	0.02	0.4	Organic nitrogen hydrolysis rate at 20 °C
k_{denit}	1/day	0.001	0.1	Denitrification rate at 20 °C
V_{sN}	m/day	0.001	0.1	Organic nitrogen settling velocity
$k_{n1/2}$	mg/L			Oxygen half-saturation constant for nitrification attenuation
$k_{dn1/2}$	mg/L			Oxygen half-saturation constant for denitrification attenuation
k_{dboc}	1/day	0.02	3.4	CBOD degradation rate at 20 °C
V_{sDBOC}	m/day	0.01	0.36	CBOD settling velocity
k_{socf}	mg/L			Oxygen half-saturation constant for CBOD degradation attenuation
k_{sod}	kg/m^2/day	0.0	0.01	Sediment oxygen demand rate
r_{oc}	mg/mg	-	-	Ratio of oxygen consumed per organic carbon oxidized to inorganic carbon
k_{dec}	1/day	Mancini		Degradation constant for *E. coli*

Appendix D. Time Series for Test 3
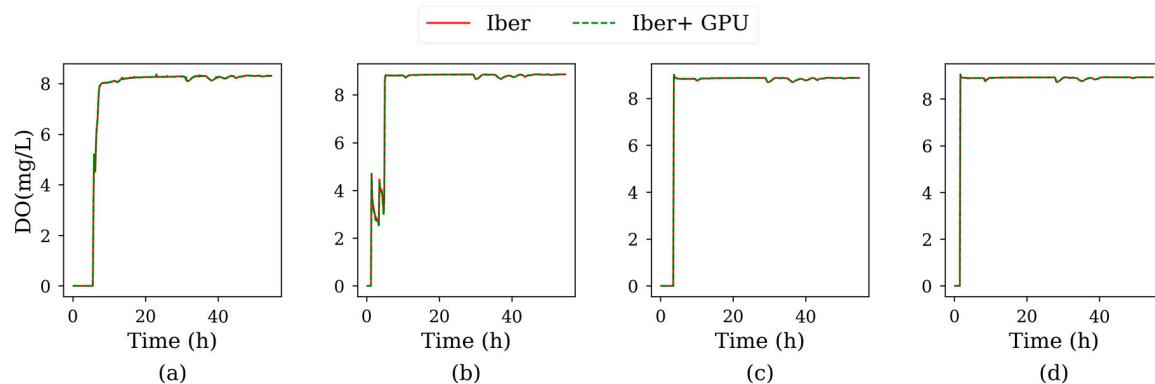

**Figure A1.** Time series of dissolved oxygen for Test 3, sampled at control points P1 (**a**), P2 (**b**), P3 (**c**) and 4 (**d**).
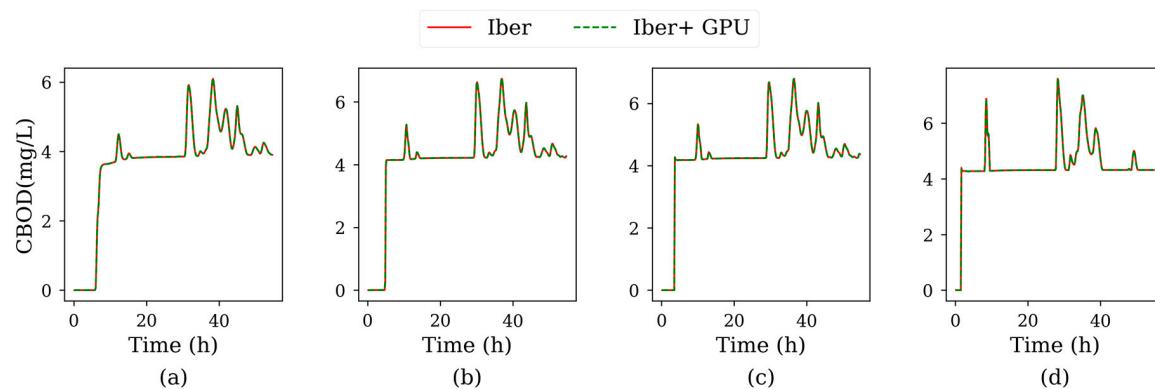
Figure A2. Time series of CBOD for Test 3, sampled at control points P1 (**a**), P2 (**b**), P3 (**c**) and P4 (**d**).
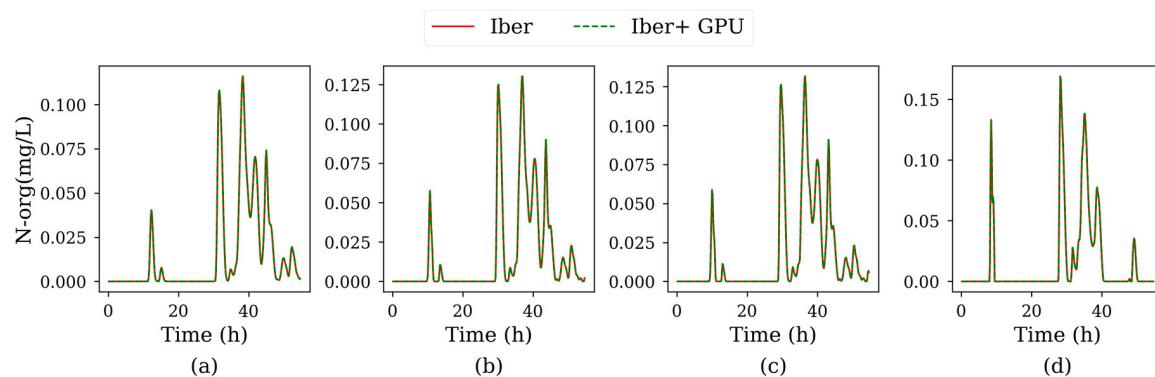
Figure A3. Time series of organic nitrogen for Test 3, sampled at control points P1 (**a**), P2 (**b**), P3 (**c**) and 4 (**d**).
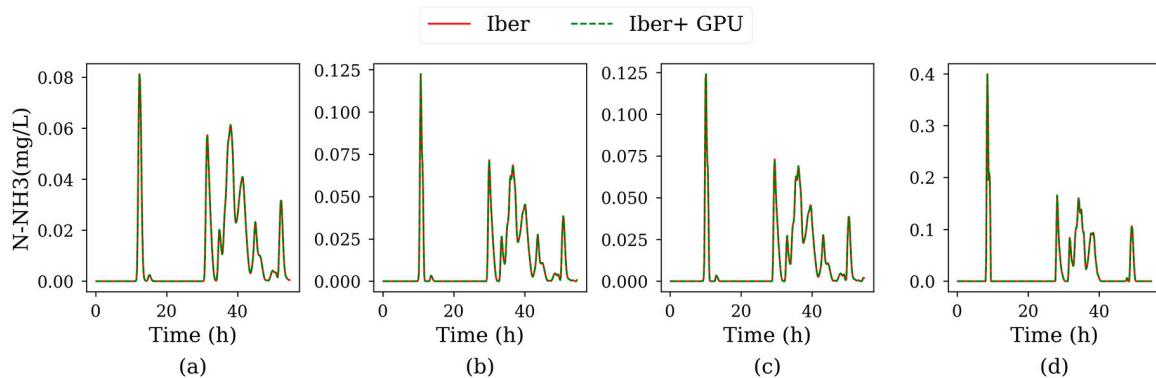
Figure A4. Time series of ammoniacal nitrogen for Test 3, sampled at control points P1 (**a**), P2 (**b**), P3 (**c**) and P4 (**d**).
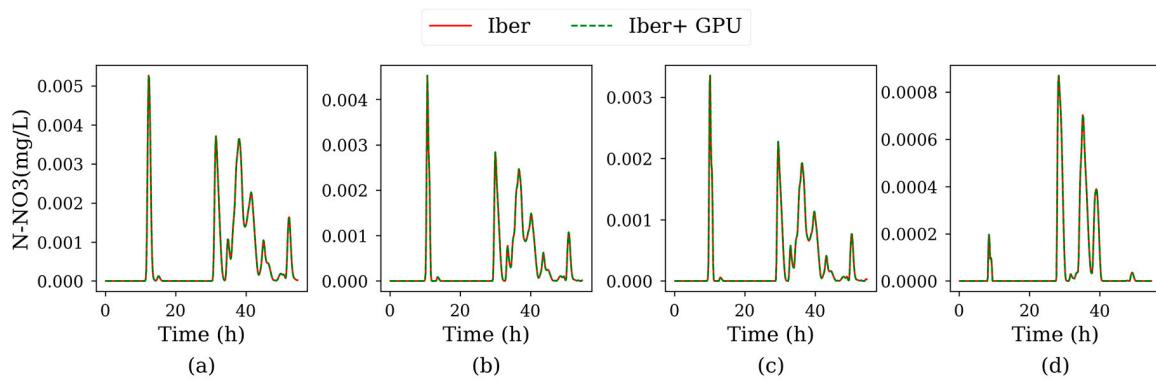
Figure A5. Time series of nitrate-nitrite nitrogen for Test 3, sampled at control points P1 (**a**), P2 (**b**), P3 (**c**) and P4 (**d**).

Appendix E. Time Series for Test 4
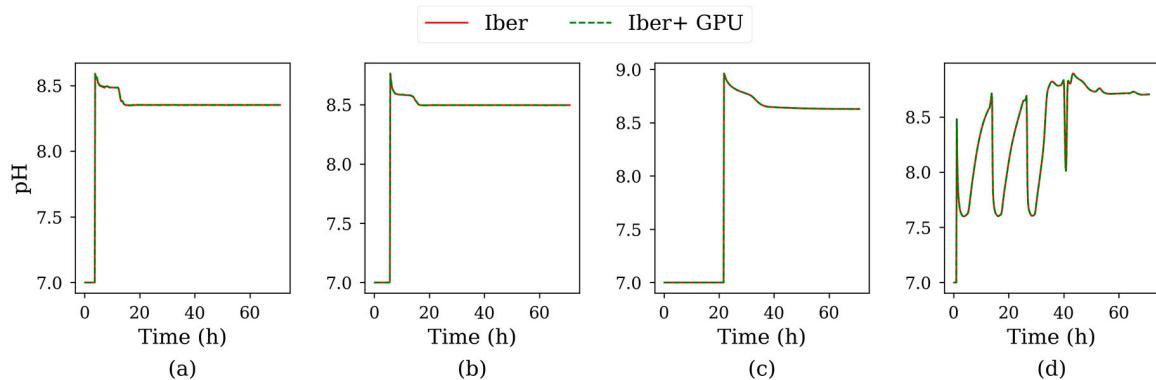
Figure A6. Time series of pH for Test 4, sampled at control points P1 (**a**), P2 (**b**), P3 (**c**) and P4 (**d**).
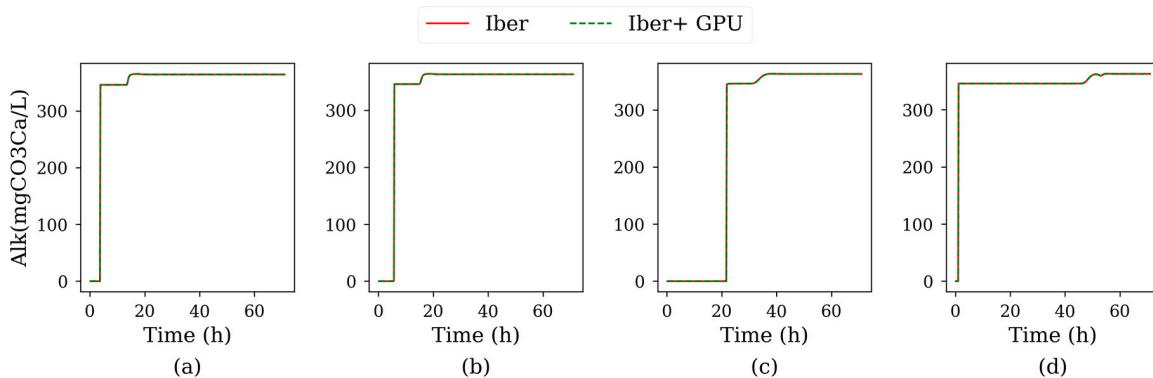
Figure A7. Time series of alkalinity for Test 4, sampled at control points P1 (**a**), P2 (**b**), P3 (**c**) and P4 (**d**).
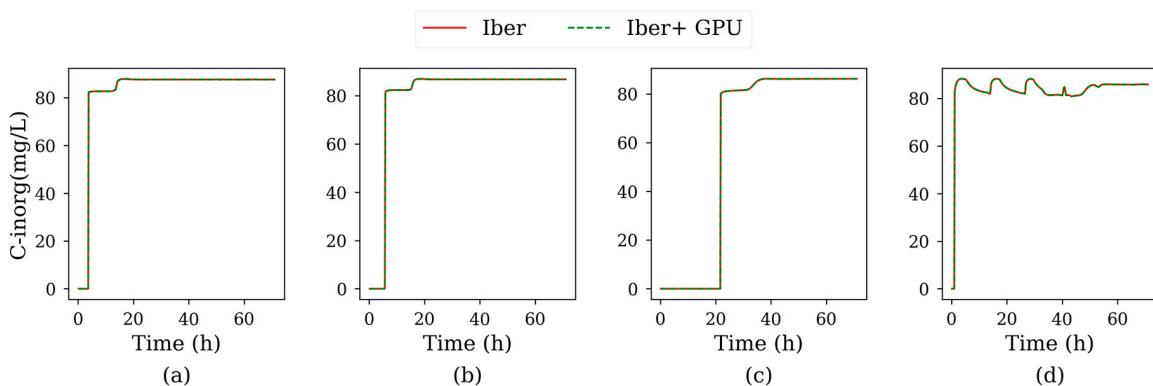

Figure A8. Time series of inorganic carbon for Test 4, sampled at control points P1 (**a**), P2 (b), P3 (**c**) and P4 (**d**).
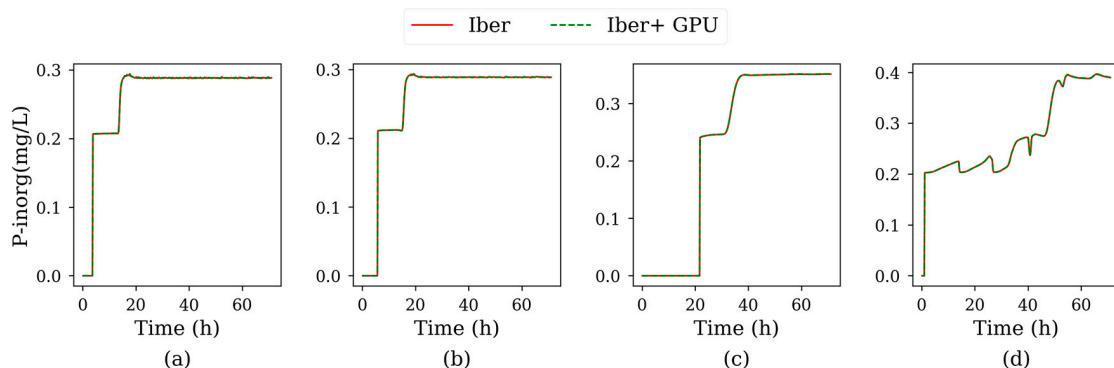

**Figure A9.** Time series of inorganic phosphorus for Test 4, sampled at control points P1 (**a**), P2 (**b**), P3 (**c**) and P4 (**d**).
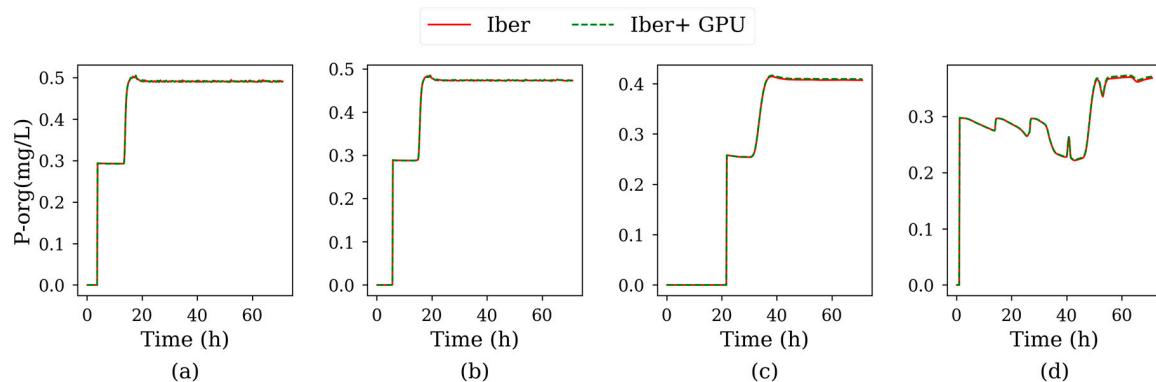
Figure A10. Time series of organic phosphorus for Test 4, sampled at control points P1 (**a**), P2 (**b**), P3 (**c**) and P4 (**d**).
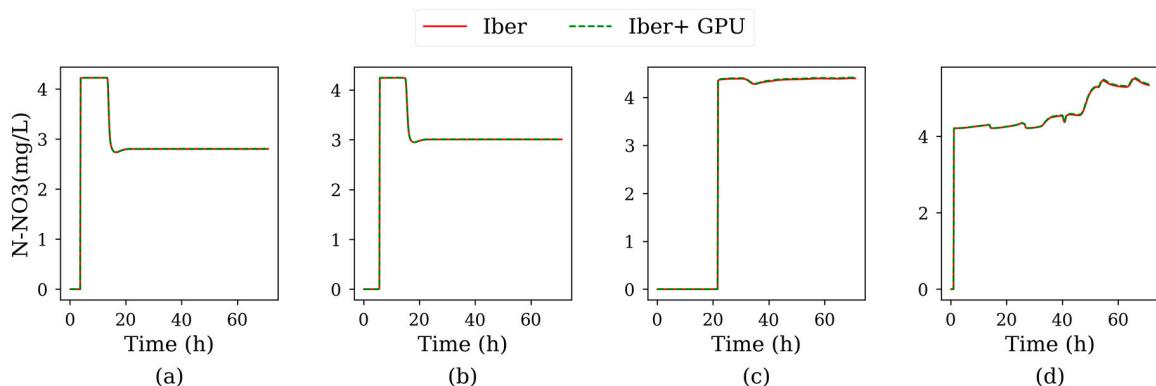

**Figure A11.** Time series of nitrite-nitrate nitrogen for Test 4, sampled at control points P1 (**a**), P2 (**b**), P3 (**c**) and P4 (**d**).
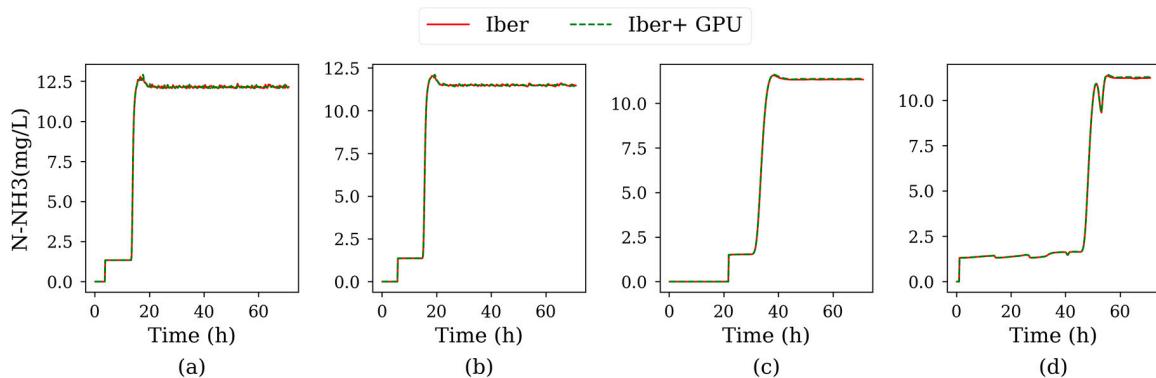
Figure A12. Time series of ammoniacal nitrogen for Test 4, sampled at control points P1 (**a**), P2 (**b**), P3 (**c**) and P4 (**d**).
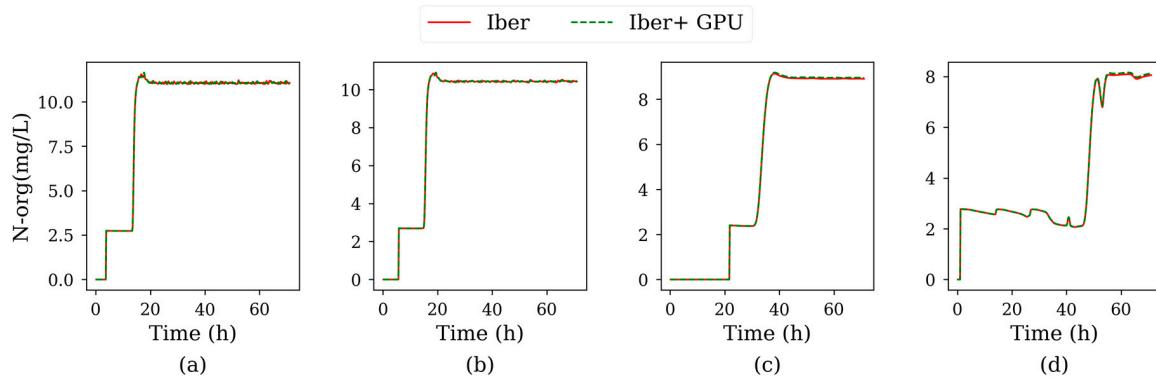
**Figure A13.** Time series of organic nitrogen for Test 4, sampled at control points P1 (**a**), P2 (**b**), P3 (**c**) and P4 (**d**).
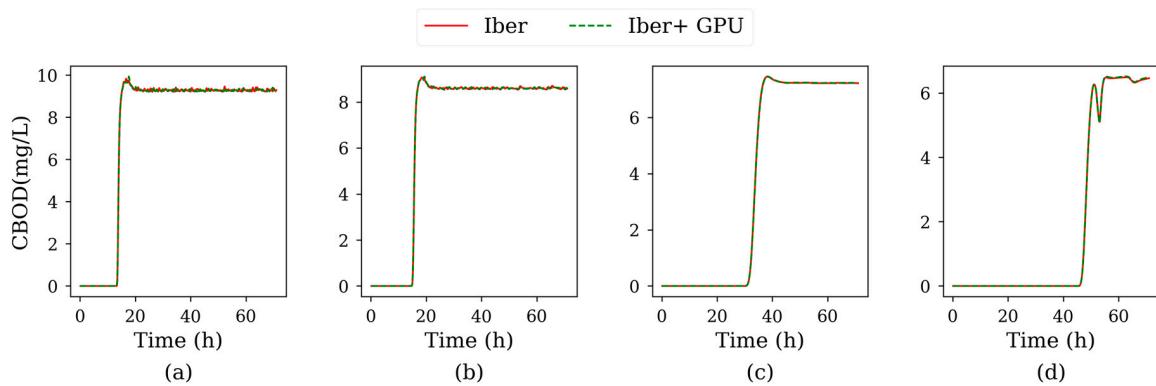

Figure A14. Time series of CBOD for Test 4, sampled at control points P1 (**a**), P2 (**b**), P3 (**c**) and P4 (**d**).
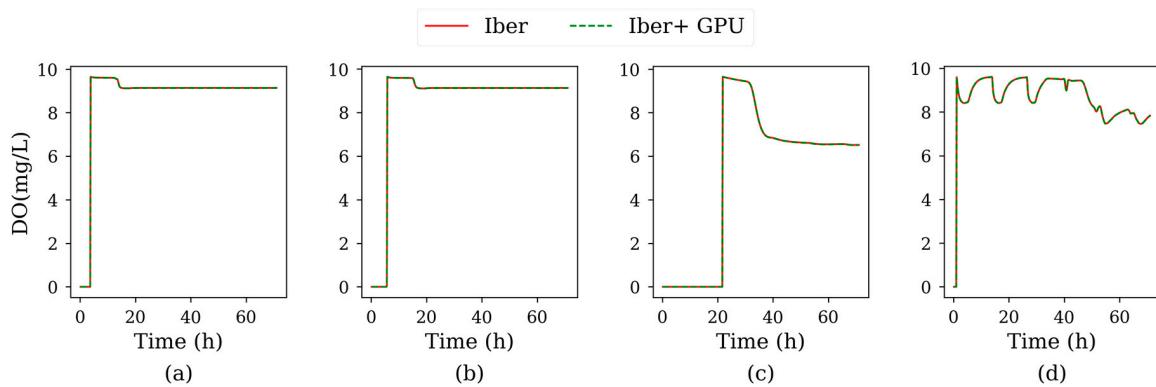

Figure A15. Time series of dissolved oxygen for Test 4, sampled at control points P1 (**a**), P2 (**b**), P3 (**c**) and P4 (**d**).

Appendix F. Data Sources for the Test Cases

Table A1. Data sources for the Test Case 1.

Bathymetry	Bathymetric survey carried out for previous studies. Spatial resolution of 30 m.
Effluent Discharge and Concentration	Virtual
Streamflow	Annual average flow from river Grande de Xubia. Available from the regional Meteorological Agency MeteoGalicia (www.meteogalicia.gal).
Tide	Tidal harmonics obtained from the tidal gauge of Ferrol. Available from Puertos del Estado (www.puertos.es).

Table A2. Data sources for the Test Case 2.

Bathymetry	Bathymetric survey carried out for previous studies. Spatial resolution of 30 m.
Effluent Discharge and Concentration	Sewer network model carried out in previous studies.
Streamflow	Annual average flow from river Mero. Available from the regional Meteorological Agency MeteoGalicia (www.meteogalicia.gal).
Tide	Tidal harmonics obtained from the tidal gauge of A Coruña. Available from Puertos del Estado (www.puertos.es).

Table A3. Data sources for the Test Case 3.

Bathymetry	Bathymetric survey carried out in [44].
Effluent Discharge and Concentration	Sewer network model carried out in [46] and [46].
Streamflow	River discharge obtained from a Water Quality Automatic Information System (SAICA) located upstream the river reach under study. Available from the regional water administration Confederación Hidrográfica del Miño-Sil (www.chminosil.es).

Table A4. Data sources for the Test Case 4.

Bathymetry	Digital terrain model at 2 m resolution, obtained from LiDAR data from the Spanish National Plan of Aerophotogrammetry (PNOA), available from the Spanish National Geographic Institute (www.ign.es).
Effluent Discharge and Concentration	Virtual.
Streamflow	Annual average flow obtained from the regional water administration Confederación Hidrográfica del Guadalquivir (www.chguadalquivir.es).

References

1. Foundation of Water Research. *Urban Pollution Management Manual: A Planning Guide for the Management of Urban Wastewater Discharges During Wet Weather*; Foundation of Water Research: Buckinghamshire, UK, 1998.
2. Robinson, R.B.; Roby, J.C. Concentration-duration-frequency curves for pH in a stream in the great smoky mountains. *J. Environ. Eng.* **2006**, *132*, 1600–1605. [CrossRef]
3. Schwartz, J.S.; Dahle, M.; Bruce Robinson, R. Concentration-duration-frequency curves for stream turbidity: Possibilities for assessing biological impairment. *J. Am. Water Resour. Assoc.* **2008**, *44*, 879–886. [CrossRef]
4. Di Toro, D.; Fitzpatrick, J.; Thomann, R. *Documentation for Water Quality Analysis Simulation Program (WASP) and Model Verification Program (MVP)*; United States Environmental Protection Agency: Washington, DC, USA, 1983.

5.   Ambrose, R. *WASP4, a Hydrodynamic and Water Quality Model: Model Theory, User's Manual and Programmer's Guide*; United States Environmental Protection Agency: Washington, DC, USA, 1988.

6.   Chapra, S.C.; Pelletier, G.J.; Tao, H. *QUAL2K: A Modeling Framework for Simulating River and Stream Water Quality*; Version 2.11: Documentation and Users Manual; Civil and Environmental Engineering Dept., Tufts University: Medford, MA, USA, 2008.

7.   Cole, T.M.; Buchak, E.M. *CE-QUAL-W2: A Two-Dimensional, Laterally Averaged, Hydrodynamic and Water Quality Model*; Version 2.0: User Manual; US Army Corps of Engineers: Vicksburg, MS, USA, 1995.

8.   Cole, T.M.; Wells, S.A. *CE-QUAL-W2: A Two-Dimensional, Laterally Averaged, Hydrodynamic and Water Quality Model*; Version 3.5; US Army Engineering and Research Development Center: Vicksburg, MS, USA, 2006.

9.   Kashefipour, S.M.; Lin, B.; Harris, E.; Falconer, R.A. Hydro-environmental modelling for bathing water compliance of an estuarine basin. *Water Res.* **2002**, *36*, 1854–1868. [CrossRef]

10.   Kay, D.; Stapleton, C.M.; Wyer, M.D.; McDonald, A.T.; Crowther, J.; Paul, N.; Jones, K.; Francis, C.; Watkins, J.; Wilkinson, J.; et al. Decay of intestinal enterococci concentrations in high-energy estuarine and coastal waters: Towards real-time T 90 values for modelling faecal indicators in recreational waters. *Water Res.* **2005**, *39*, 655–667. [CrossRef] [PubMed]

11.   Wu, Y.; Falconer, R.; Lin, B. Modelling trace metal concentration distributions in estuarine waters. *Estuarine Coast. Shelf Sci.* **2005**, *64*, 699–709. [CrossRef]

12.   Gao, G.; Falconer, R.A.; Lin, B. Numerical modelling of sediment–bacteria interaction processes in surface waters. *Water Res.* **2011**, *45*, 1951–1960. [CrossRef]

13.   Abu-Bakar, A.; Ahmadian, R.; Falconer, R.A. Modelling the transport and decay processes of microbial tracers in a macro-tidal estuary. *Water Res.* **2017**, *123*, 802–824. [CrossRef]

14.   Cea, L.; Bermúdez, M.; Puertas, J.; Bladé, E.; Corestein, G.; Escolano, E.; Conde, A.; Bockelmann-Evans, B.; Ahmadian, R. IberWQ: New simulation tool for 2D water quality modelling in rivers and shallow estuaries. *J. Hydroinf.* **2016**, *18*, 816–830. [CrossRef]

15.   Frick, W.E.; Roberts, P.J.W.; Davis, L.R.; Keyes, J.; Baumgartner, D.J.; George, K.P. *Dilution Models for Effluent Discharges*, 4th ed.; (Visual Plumes); U.S. Environmental Protection Agency, Office of Research and Development: Washington, DC, USA, 2003.

16.   Doneker, R.L.; Jirka, G.H. CORMIX-GI systems for mixing zone analysis of brine wastewater disposal. *Desalination* **2001**, *139*, 263–274. [CrossRef]

17.   Kiesel, J.; Schmalz, B.; Brown, G.L.; Fohrer, N. Application of a hydrological-hydraulic modelling cascade in lowlands for investigating water and sediment fluxes in catchment, channel and reach. *J. Hydrol. Hydromech.* **2013**, *61*, 334–346. [CrossRef]

18.   D'Ambrosio, E.; Gentile, F.; de Girolamo, A.M. Assessing the sustainability in water use at the basin scale through water footprint indicators. *J. Clean. Prod.* **2019**, *244*, 118847. [CrossRef]

19.   Bladé, E.; Cea, L.; Corestein, G.; Escolano, E.; Puertas, J.; Vázquez-Cendón, E.; Dolz, J.; Coll, A. Iber: herramienta de simulación numérica del flujo en ríos. *Rev. Int Metodos Numer. Para Calculo Diseno Ing.* **2014**, *30*, 1–10. [CrossRef]

20.   Michaud, J.P. *A Citizien's Guide to Understanding and Monitoring Lakes and Streams*; Washington State Department of Ecology, Publications Office: Olympia, WA, USA, 1991.

21.   García-Feal, O.; González-Cao, J.; Gómez-Gesteira, M.; Cea, L.; Domínguez, J.M.; Formella, A. An accelerated tool for flood modelling based on Iber. *Water (Switzerland)* **2018**, *10*, 1459. [CrossRef]

22.   Chapra, S.C. *Surface Water-Quality Modeling*; McGraw-Hill: New York, NY, USA, 1997; ISBN 1478608307.

23.   Cea, L.; Puertas, J.; Vázquez-Cendón, M.E. Depth averaged modelling of turbulent shallow water flow with wet-dry fronts. *Arch. Comput. Meth. Eng.* **2007**, *14*, 303–341. [CrossRef]

24.   Millero, F.J.; Poisson, A. International one-atmosphere equation of state of seawater. *Deep Sea Res. Part A* **1981**, *28*, 625–629. [CrossRef]

25.   Bowie, G.; Mills, W.; Porcella, D.; Campbell, C.; Pagenkopf, J.; Rupp, G.; Johnson, K.; Chan, P.; Gherini, S.; Chamberlin, C. *Rates, Constants, and Kinetics Formulations in Surface Water Quality Modeling*; EPA: Athens, GA, USA, 1985.

26.   Cea, L.; Vázquez-Cendón, M.E. Unstructured finite volume discretisation of bed friction and convective flux in solute transport models linked to the shallow water equations. *J. Comput. Phys.* **2012**, *231*, 3317–3339. [CrossRef]

27. Versteeg, H.K.; Malalasekera, W. *An Introduction to Computational Fluid Dynamics—The Finite Volume Method*; Pearson Education: New York, NY, USA, 2007.

28. Jasak, H.; Weller, H.G.; Gosman, A.D. High resolution NVD differencing scheme for arbitrarily unstructured meshes. *Int. J. Numer. Methods Fluids* **1999**, *31*, 431–449. [CrossRef]

29. Nvidia Corporation CUDA C Programming Guide. Available online: https://docs.nvidia.com/cuda/pdf/CUDA_C_Programming_Guide.pdf (accessed on 17 December 2019).

30. Crespo, A.J.C.; Domínguez, J.M.; Rogers, B.D.; Gómez-Gesteira, M.; Longshaw, S.; Canelas, R.; Vacondio, R.; Barreiro, A.; García-Feal, O. DualSPHysics: Open-source parallel CFD solver based on Smoothed Particle Hydrodynamics (SPH). *Comput. Phys. Commun.* **2015**, *187*, 204–216. [CrossRef]

31. Crossley, A.; Lamb, R.; Waller, S. Fast solution of the Shallow Water Equations using GPU technology. In Proceedings of the British Hydrological Society 3rd International Symposium, Newcastle, UK, 13–19 July 2010.

32. Vacondio, R.; Dal Palù, A.; Mignosa, P. GPU-enhanced finite volume shallow water solver for fast flood simulations. *Environ. Modell. Softw.* **2014**, *57*, 60–75. [CrossRef]

33. Lacasta, A.; Morales-Hernández, M.; Murillo, J.; García-Navarro, P. An optimized GPU implementation of a 2D free surface simulation model on unstructured meshes. *Adv. Eng. Softw.* **2014**, *78*, 1–15. [CrossRef]

34. Liu, Q.; Qin, Y.; Li, G. Fast Simulation of Large-Scale Floods Based on GPU Parallel Computing. *Water* **2018**, *10*, 589. [CrossRef]

35. NVIDIA Corporation NVIDIA Tesla V100 GPU Architecture. Available online: https://images.nvidia.com/content/volta-architecture/pdf/volta-architecture-whitepaper.pdf (accessed on 7 December 2019).

36. Moon, B.; Jagadish, H.V.; Faloutsos, C.; Saltz, J.H. Analysis of the Clustering Properties of Hilbert Space-filling Curve. *IEEE Trans. Knowl. Data Eng.* **2001**, *13*, 124–141. [CrossRef]

37. Nvidia CUB. Available online: https://nvlabs.github.io/cub/ (accessed on 7 December 2019).

38. NVIDIA Corporation NVIDIA Turing GPU Architecture. Available online: https://www.nvidia.com/content/dam/en-zz/Solutions/design-visualization/technologies/turing-architecture/NVIDIA-Turing-Architecture-Whitepaper.pdf (accessed on 2 February 2020).

39. Cea, L.; Bermúdez, M.; Puertas, J. Uncertainty and sensitivity analysis of a depth-averaged water quality model for evaluation of Escherichia Coli concentration in shallow estuaries. *Environ. Modell. Softw.* **2011**, *26*, 1526–1539. [CrossRef]

40. García-Barcina, J.M.; Oteiza, M.; de la Sota, A. Modelling the faecal coliform concentrations in the Bilbao estuary. *Hydrobiologia* **2002**, *475*, 213–219. [CrossRef]

41. Kashefipour, S.M.; Lin, B.; Falconer, R.A. Modelling the fate of faecal indicators in a coastal basin. *Water Res.* **2006**, *40*, 1413–1425. [CrossRef] [PubMed]

42. Manache, G.; Melching, C.S.; Lanyon, R. Calibration of a continuous simulation fecal coliform model based on historical data analysis. *J. Environ. Eng.* **2007**, *133*, 681–691. [CrossRef]

43. Bode, A.; Álvarez-Ossorio, M.T.; González, N.; Lorenzo, J.; Rodríguez, C.; Varela, M.; Varela, M.M. Seasonal variability of plankton blooms in the Ria de Ferrol (NW Spain): II. Plankton abundance, composition and biomass. *Estuarine Coast. Shelf Sci.* **2005**, *63*, 285–300. [CrossRef]

44. Anta Álvarez, J.; Bermúdez, M.; Cea, L.; Suárez, J.; Ures, P.; Puertas, J. Modelización de los impactos por DSU en el río Miño (Lugo). *Ingeniería del agua* **2015**, *19*, 105–116. [CrossRef]

45. Stumm, W.; Morgan, J. *Aquatic Chemistry: Chemical Equilibria and Rates in Natural Waters*; Wiley: New York, NY, USA, 1996.

46. Piñeiro, J.; Maestro, I.; Aguirre, F.; Ures, P.; Torres, D.; Anta, J.; Puertas, J.; Suárez, J. Análisis del funcionamiento de un depósito-aliviadero en el sistema de saneamiento unitario en la aglomeración de Lugo. In Proceedings of the Actas de las II Jornadas de Ingeniería del Agua, Barcelona, Spain, 5–6 October 2011.

5.4. MIDAS: A New Integrated Flood Early Warning System for the Miño River

Fernández-Nóvoa, D., García-Feal, O., González-Cao, J., de Gonzalo, C., Rodríguez-Suárez, J. A., Ruiz del Portal, C. & Gómez-Gesteira, M. (2020). MIDAS: A New Integrated Flood Early Warning System for the Miño River [Number: 9 Publisher: Multidisciplinary Digital Publishing Institute]. *Water*, *12*(9), 2319. https://doi.org/10.3390/w12092319

Información de la revista	
Revista	Water
ISSN	2073-4441
País	Suíza
Editorial	MDPI
Factor de impacto	2,544 (2019*)
Cuartil	Q2

MIDAS: A New Integrated Flood Early Warning System for the Miño River

Diego Fernández-Nóvoa [1], Orlando García-Feal [1,*] , José González-Cao [1], Carlos de Gonzalo [2], José Antonio Rodríguez-Suárez [2], Carlos Ruiz del Portal [3] and Moncho Gómez-Gesteira [1]

[1] Environmental Physics Laboratory (EPhysLab), CIM-UVIGO, Universidade de Vigo, Campus As Lagoas s/n, 32004 Ourense, Spain; diefernandez@uvigo.es (D.F.-N.); jgcao@uvigo.es (J.G.-C.); mggesteira@uvigo.es (M.G.-G.)

[2] Tragsatec, SAIH Miño-Sil, 32003 Ourense, Spain; c.gonzaloaranoa@gmail.com (C.d.G.); josearsuarez@hotmail.com (J.A.R.-S.)

[3] Confederación Hidrográfica Miño-Sil, 32003 Ourense, Spain; cgruiz@chminosil.es

* Correspondence: orlando@uvigo.es; Tel.: +34-988-368-791

Abstract: Early warning systems have become an essential tool to mitigate the impact of river floods, whose frequency and magnitude have increased during the last few decades as a consequence of climate change. In this context, the Miño River Flood Alert System (MIDAS) early warning system has been developed for the Miño River (Galicia, NW Spain), whose flood events have historically caused severe damage in urban areas and are expected to increase in intensity in the next decades. MIDAS is integrated by a hydrologic (HEC-HMS) and a hydraulic (Iber+) model using precipitation forecast as input data. The system runs automatically and is governed by a set of Python scripts. When any hazard is detected, an alert is issued by the system, including detailed hazards maps, to help decision makers to take precise and effective mitigation measures. Statistical analysis supports the accuracy of hydrologic and hydraulic modules implemented to forecast river flow and flooded critical areas during the analyzed period of time, including some of the most extreme events registered in the Miño River. In fact, MIDAS has proven to be capable of predicting most of the alert situations occurred during the study period, showing its capability to anticipate risk situations.

Keywords: early warning system; flood; hydrology; HEC-HMS; hydrodynamics; Iber+

1. Introduction

Flood events have increased both their frequency and intensity during the last few decades [1,2]. This has occurred for multiple reasons, but two are especially remarkable. On the one hand, changes in land uses, including the increase in urbanization, cause an increase in the run-off volume that can reach river flow systems [3,4]. On the other hand, changes in precipitation patterns have been induced by the impact of climate change [5,6]. In fact, one of the most dangerous consequences associated with climate change is the intensification of extreme rainfall events, increasing the hazard of the associated flood episodes [7–11].

Flood events have been estimated as one of the most important natural hazards in recent decades. These disasters have affected millions of people and caused billions of dollars in losses due to the damage caused [12,13]. Consequently, the development of early warning systems (EWSs) designed to forecast flood events has increased during the last few decades [14,15]. Several examples can be found in the literature. The Hydrometeorological Data Resources and Technology for Effective Flash Flood Forecasting (HYDRATE) project aims to develop a methodology for flash flood forecasting in Europe [16], the European Flood Awareness System (EFAS) provides streamflow forecasting with a

time horizon of 10 days for the European river network [17–20], and the Global Flood Awareness System (GloFAS) offers a global hydrological forecast [21]. Alfieri et al. [7] also describe several EWS developed at more local scale in Europe. The reader can find a detailed comparison between two EWSs applied to a regional scale in Corral et al. [22]. One of them is based on precipitation forecast, while the other one is based in rainfall-runoff models. Other EWS developed to a regional scale is EHIMI [22], applied by the Water Agency of Catalonia (ACA) in Spain. Liu et al. [6] shows the rapid advances in the flood early warning systems development in China, and Hirpa et al. [23] present a flood forecasting system for the major rivers in South Asia based on satellite data. Other examples can be found in Cloke and Pappenberger [24] and De Luca et al. [25]. Cloke and Pappenberger [24] show a review of different approaches of flood forecasting systems based on ensembles of weather prediction, and De Luca et al. [25] present several mathematical models operating into EWS functioning. The flood EWS focused on predicting the evolution of river floods, are essential because they allow to take measures to prevent and mitigate the dramatic consequences that arise in flood scenarios [15]. Therefore, EWSs suppose a useful tool with the potential to save lives, diminish the damage of fundamental infrastructures and enhance the resilience of the society [12,15,26]. An example of an EWS focused on the hazard and the personal and economic losses is shown in Ritter et al. [27]. Within this context, the base of a flood EWS is the hydrological model component, which transforms rainfall in runoff, which determines the extent of the flood in rain-dominated systems [28,29]. The hydrological component can be focused on continuous modelling or on single event applications. In the first case, the continuous schemes are able to reproduce correctly the rainfall-runoff processes along large time periods, however they depend on multiple variables that should be measured or/and calibrated [8,30]. This can be an important constraint, especially in poorly gauged basins, because the accurate development of the approach requires a complex data set [30]. The single event approach usually has a lower number of variables involved and it is only focused on individual and concrete events [29,31]. Although single-event approaches also need to be calibrated, the need for a lower number of parameters allows the approach to be applied in a greater number of situations. Taking also into account that EWSs are especially focused on extreme riverine floods, event-based models have been widely used for runoff calculation in several engineering applications worldwide, especially the soil conservation service curve number (SCS-CN) approach [32]. The main advantage of the SCS-CN model lies in its simplicity, which makes it applicable in a great number of basins, even when data availability is limited.

Most EWSs report flood hazard based only on the discharge provided by the hydrological model. However, some EWSs implement a step forward including a hydraulic module [8]. In this case, the river flow forecasted by the hydrological model is used to feed the hydraulic model. With this, it is possible to forecast the water depth and velocity throughout the area under interest. Thus, the alert reports are based on the hazard maps provided by the hydraulic model, which allows a detailed view of the impact of flood events [8,33]. However, hydraulic models usually require large computational times, which significantly limits their application in real-time flood forecasting. For this reason, the use of hydraulic models as part of an EWS is limited [8,33,34].

The aim of this study is to design an EWS that efficiently combines hydrological and hydraulic components. In this sense, a novel hydrological methodology was developed. This new technique allows us to use the SCS-CN method for continuous modelling, bringing together the main advantages of both approaches (continuous and single event). In addition, the new implementation of the 2D hydraulic model Iber+ [35] was integrated as part of the EWS. Iber+ is a GPU-parallelized version of the Iber model [36], which improves the efficiency of the former model in about two orders of magnitude. This implementation overcomes the limitations of hydraulic models to be used in EWSs and allows the user to simulate areas under flood risks at a reasonable time [33]. This is especially suitable for applications where a fast response is crucial.

The proposed methodology is applied and evaluated in the Miño River (Galicia, northwest Spain) whose flood events can cause damage in important urban areas [33]. This approach is particularly helpful for Galicia, where the intensity of extreme rain events is expected to increase in the next decades

due to climate change, leading to hazardous flood episodes [37]. This makes the development of systems able to accurately predict in advance the flood events occurring in the area under scope even more necessary.

The method offers numerous advantages, including ease of application in other areas, especially in poorly sampled river basins. In turn, the real-time information provided by the hydraulic model allows us to obtain a better representation and an in-depth knowledge of the impact of river floods. This information can help decision makers to take suitable measures to mitigate the damage.

This work is organized as follows: Section 2 presents a brief description of the area under study and describes in detail the hydrologic and hydraulic models applied in this work. The general architecture of the EWS and the statistic parameters used to define the accuracy of the EWS are also presented in this section. The results obtained with both hydrologic and hydraulic models are validated and analyzed in Section 3. Finally, Section 4 shows the conclusions of this work.

2. Study Area and Methodology

2.1. Study Area

Figure 1 shows the study area analyzed in this work. This area is located in the NW Iberian Peninsula occupying a total extent of around 5000 km^2, with an elevation ranging from 90 m.a.s.l. to 1200 m.a.s.l. (Figure 1a,b). It corresponds to the upper reach of the shared Portuguese–Spanish Miño River catchment. The Miño River presents a pluvial regime, with maximum river flows during the winter months and minimum flows during the summer [38]. This flow pattern is regulated by the precipitation characteristics of the area under scope, characterized by a seasonal evolution of Azores high and Iceland low, which provokes the occurrence of most of rainy events during the winter months [39]. Attending to its topographic features, the area under consideration was divided in six sub-basins for the hydrological procedure (Figure 1b). The outlet of the catchment is located at Ourense City (Figure 1c–e), which suppose a critical area affected by recurrent floods under extreme conditions. Attending to these reasons, Ourense is considered to be the test area to analyze the accuracy of the hydraulic model to reproduce flood events. The area delimited by the red line in Figure 1c corresponds to the domain of the hydraulic model. The input and output of the domain are also shown in Figure 1c. The section of the river is 3 km long and the total area is about 1 km^2. More than 50 land uses were defined attending to the characteristics of the terrain (Figure 1e).

2.2. Description of the Early Warning System

2.2.1. Hydrological Model: HEC-HMS

The semi-distributed model HEC-HMS [40,41] was used to process the main hydrological features of the area under scope. This modelling system, which is one of the most used for hydrological procedures, offers accurate results in locations close to the area under scope [8,31,33]. The designed methodology only requires the input of a few variables to resolve the hydrologic processes—curve number (CN), lag time (Tl), baseflow linear reservoir (BLR), and the routing coefficients along the river channel, as described below.

Rainfall infiltration is determined using the SCS-CN [42,43]. This procedure only requires the knowledge of the CN parameter. The standard CN (intermediate or average conditions) for each sub-basin under scope was firstly calculated following [44,45]:

$$\text{CN} = \frac{25,400}{S + 254} \tag{1}$$

where *S* is the potential maximum watershed site storage after runoff begins that is computed using:

$$S = 5\left(P + 2Q - \sqrt{4Q^2 + 5PQ}\right) \tag{2}$$

where Q is the runoff depth and P is the rainfall depth, both obtained from SIMPA (stands for integrated precipitation-contribution modelling system) model [46]. Equation (2) is also dependent of initial abstraction, which was simplified as $0.2 \cdot S$ following Stewart et al. [45], resulting in the equation presented above.
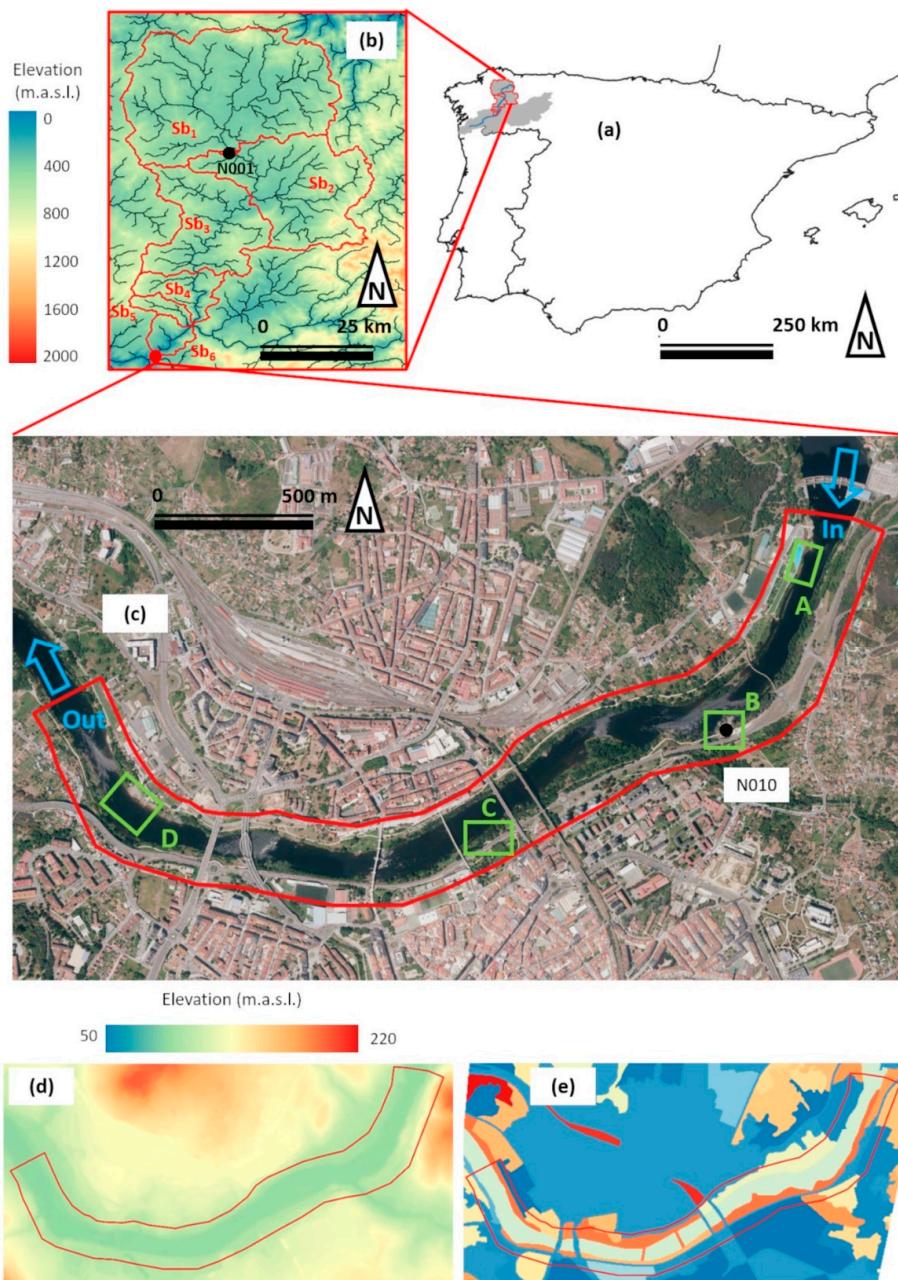

Figure 1. Area of study. (**a**) location of the entire catchment of the shared Portuguese–Spanish Miño River (shaded area) in the Iberian Peninsula and the riverbed (blue line); (**b**) subbasins (Sb1, Sb2, . . . , Sb6) of the domain; (**c**) area of study in Ourense (PNOA courtesy of © Instituto Geográfico Nacional). Area enclosed by the red line defines the domain of the hydraulic model. The inlet and outlet boundaries of the domain and the gauge stations N001 (Lugo) and N010 (Ourense) are also shown. Green rectangles define the control areas. Control area A: public sport facility; control area B: gauge station N010; control area C: public path; control area D: public thermal baths. The inlet (in blue arrow) and outlet (out blue arrow) boundaries are also depicted. (**d**) Topographic features and (**e**) land uses of the area under study in Ourense.

The theoretical CN value obtained at Lugo was calibrated using river flow data provided by the Confederacion Hidrografica del Miño-Sil (CHMS, https://www.chminosil.es). The Lugo location was selected for calibration purposes because the Miño River exists there under natural condition and is not affected by river structures like dams. This procedure was carried out in order to completely adjust the CNs characterizing the area under study. The Nelder–Mead algorithm [47] was used to obtain the optimum value with the Nash–Sutcliffe efficiency index (NSE) as an objective function, which has reported accurate results in surrounding areas [8]. The results obtained show a close relation between the theoretical CN value obtained from Equation (1) and the calibrated value (81 and 84, respectively), which suppose differences of less than 5%. The correction in CN value detected in calibration process was extrapolated to the rest of the sub-basins. First, theoretical values were calculated by means of Equation (1), and then, the correction (in percentage) was applied. In all cases, obtained CNs were rounded to the upper value (Table 1). CN values are similar to the ones obtained in other studies developed in close areas [31].

Table 1. Main hydrological characteristics of the area under scope.

Subbasin	Area (km^2)	CN$_{dry}$	CN$_{average}$	CN$_{moist}$	Lag Time (min)
Sb1	2172	69	84	93	1740
Sb2	1219	69	84	93	1500
Sb3	897	72	86	94	1620
Sb4	255	75	88	95	840
Sb5	93	72	86	94	540
Sb6	130	71	85	93	600

In addition, it is a well-known fact that this standard CN should be modified attending to the different moisture content of the soil [31,48,49]. This means that the antecedent moisture content (AMC) plays a key role in the rain-runoff processes. The criterion to determine the different conditions is based on [50], which defines three possible CN classes (dry, average, and moist), whose inter-connections are specified in the equations defined in [50].

$$CN_{dry} = \frac{4.2 \times CN_{average}}{10 - 0.058 \times CN_{average}} \tag{3}$$

$$CN_{moist} = \frac{23 \times CN_{average}}{10 + 0.13 \times CN_{average}} \tag{4}$$

The thresholds delimiting the corresponding CN class are also based on the SCS criterion [50], establishing two different annual seasons (non-humid and humid) and a dependence of the cumulative rainfall occurred in the antecedent days. The SCS approach only takes into account the antecedent five days for AMC calculations, however, some studies reveal that the rainfall of a larger number of antecedent days has also an important impact [51,52]. This was also tested in locations close to the area under scope, where [31] detected the important role of the previous 30 days in the moisture content. Slightly better results were obtained in the study area when the AMC of the previous 30 days was considered during the non-humid season and the AMC of the previous 5 days during the humid season (Table 2). In the present work, the non-humid season was calculated to last from April to November and the humid season from December to March. However, some false peaks of river discharge can be obtained in December under the first intense precipitations. To correct this issue, the transition from the non-humid to the humid season starts only when the 30-day AMC surpasses 216 mm (that corresponds with the dry conditions threshold for the non-humid season). Taking into account the considerations commented above, it can be concluded that the conditions that better represent the AMC of the area under scope are those described in Table 2. The obtained thresholds are in accordance with Cea and Fraga [31].

Table 2. Antecedent moisture content (AMC) conditions representative of the area under scope.

CN Classes	Non-Humid Season (April–November) (30-day Antecedent Rainfall) (mm)	Humid Season (December–March) (5-day Antecedent Rainfall) (mm)
CN_{dry}	<216	<13
$CN_{average}$	216–318	13–28
CN_{moist}	>318	>28

The SCS unit hydrograph was selected to model the transformation of rainfall excess into surface runoff [43,50]. This approach only requires the input of the Lag Time and has shown accurate results for nearby areas [8,33]. First, time of concentration was calculated following the equation proposed by spanish development ministry [53], which provides a good approximation for Spanish basins,

$$T_c = 0.3 \times L_c^{0.76} \times J_c^{-0.19} \tag{5}$$

where T_c is the time of concentration in hours, L_c is the length of the longest flow path in kilometers, and J_c is the slope of the longest flow path in m/m.

Then, Lag Time was estimated by means of the approximation proposed by the SCS,

$$T_l = 0.6 \times T_c \tag{6}$$

T_l was calibrated at Lugo station following the procedure explained above, obtaining larger values than the theoretical ones for our case of study. After the calibration process, the T_l equation which best represent the area under scope was

$$T_l = 0.37 \times L_c^{0.76} \times J_c^{-0.19} \tag{7}$$

This equation was used for all the sub-basins of the study. Lag time values were rounded at hourly scale according to the available time data resolution (Table 1).

Baseflow dynamics was also calibrated at the Lugo station. For that, periods of diminishing flow after precipitation events were evaluated in order to know the constants associated to baseflow routing. These periods were selected several days after the precipitation to ensure that most of the river flow is due to baseflow dynamics. According to previous studies analyzing nearby areas [8], the scheme, which better represents the baseflow dynamics of the area under study, was a linear reservoir with two layers, one of them with a faster response (180 h) than the other (600 h). The results obtained were used for the rest of the sub-basins. Also, it was observed that the baseflow response during the non-humid season was significantly attenuated compared with the humid season. This fact caused false peak values associated to precipitation events during dry season. To overcome this issue, a number between one and three sequential reservoirs was used depending on the average river flow in the previous 30 days. This approach achieves an attenuation in the baseflow response in concordance with the observed river flow.

Regarding the channel routing, the Muskingum–Cunge method [54] was selected, which has provided accurate results in previous studies in close areas [33]. The variables required (river length, slope, width, etc.) were determined according to data obtained from digital terrain models.

Finally, it is important to take into account that the most important tributary of Miño River, the Sil River, is not simulated. This simulation of Sil River flow is not straightforward since the river is highly regulated by a network of connections among dams. However, the Sil River debouches into the Miño River upstream from the city of Ourense, which makes it mandatory to forecast its contribution at that location. To overcome this limitation, the river flow predicted by the hydrological system in Ourense is reconstructed. For that, the system analyzes the relation between simulated river flow (which considers only Miño catchment) and the real flow observed at that location (which also includes

the contributions of Sil River) during the period preceding the day under study (24 h and 5 h for the non-humid and humid seasons, respectively). A scale factor between the observed and forecasted flow at Ourense is obtained for the previous day. Then, this factor is applied to the hydrological forecast for the day of interest to take into account the influence of Sil river at this location.

2.2.2. Hydraulic Model: Iber+

Iber [36] is a numerical tool that solves the 2D depth-averaged shallow water equations using the finite volume method. Iber+ [35] is a new implementation of the model in C++ and CUDA [55] to improve the efficiency of the simulations. The new code is able to achieve a two-order of magnitude speed-up while attaining the same precision by using graphical processing unit (GPU) computing high performance computing (HPC) techniques. These optimizations bring the possibility to employ the model in applications with large spatio-temporal domains [56] or time constrained applications [8,33]. The software package is freely available and can be downloaded from its official website (https://iberaula.es). It also includes a graphical user interface (GUI) with preprocessing and post-processing tools.

In the present study, Iber+ was applied to analyze flood events in the test area of the city of Ourense. Figure 1c shows the numerical domain at Ourense, where more than 50 land uses were defined to model the characteristics of the terrain (Figure 1e). Manning's coefficients were computed accordingly to Gonzalez-Cao et al. [57]. The inlet condition was defined by means of the input hydrograph (critical–subcritical), and the outlet condition was defined using a supercritical–critical outflow. Turbulence was not taken into account as suggested by [58] and in accordance with similar works [59–61]. The digital elevation model (DEM) that describes the topography of the area of study has a resolution of 5 m. The DEM data files were obtained from the Instituto Geográfico Nacional website (https://www.ign.es/web/ign/portal). The computational domain was discretized using a mesh with 91,216 unstructured triangular elements, with an average area of 5 m^2.

2.2.3. Automatic Early Warning System

The early warning system is governed by a set of Python scripts. Basically, the system is triggered automatically once the precipitation forecasts are published by the local meteorological agency MeteoGalicia (https://www.meteogalicia.gal). MeteoGalicia provides precipitation information with a 72 h forecast window under a temporal and spatial resolution of 1 h and 4 km, respectively, providing an adequately representation of rainy situations for the area under scope [33]. Then, the hydrological simulation is run with the HEC-HMS model, followed by the hydraulic simulations of Iber+ for each area of interest. The computational step and the results of HEC-HMS are on an hourly scale according to the precipitation data. The Iber+ model results are also on hourly scale, while the computational step is adaptive, using a Courant–Friedrichs–Lewy condition [62] value of 0.45. A 24-h forecast horizon is considered. If any hazard is detected, an alert is issued to the corresponding decision makers. The general architecture of the system is shown in Figure 2.

The data necessary to perform the simulations is automatically retrieved by the system. This is the precipitation forecast and rain gauge data for the basins analyzed (from MeteoGalicia) and the current river flow (from CHMS) to determine the initial baseflow of the reach. The separation of the baseflow from the total flow is performed with the Eckhardt method [63], which is applied recursively from an instant where all the flow could be considered as baseflow. It is also important to note that the existing dams were not considered in the present work. However, this is a reasonable approach since, under extreme events, dams have low potential of regulation.

The hydrological simulations are performed with HEC-HMS using an SCS-CN runoff model. Although it was meant for single event simulations, the proposed method was designed to overcome its limitations. For this, the simulation of each sub-basin is separated from the simulation of the river reach. Each sub-basin is simulated for several fixed precipitation events of 24 h (See Figure 3). Each simulated event will have an independent curve number based on the precedent conditions as described above.

Additionally, if the simulated event is under the CN_{moist} class, precipitation of the previous 6 h before the event is also analyzed, and if precipitation exceeds a threshold (8 mm), the CN will be increased to 97 in order to not interrupt the rain-runoff process. The number of forecasted events depends on the period that will be predicted and will take data from the precipitation forecast. The number of hindcast events depend on the characteristics of the basin, for the case of study we found that events prior to five days before the forecast do not have a significant influence on the flow. In the present analysis, the precipitation corresponding to the previous days was estimated from rain gauges, although the system can also run with forecasted precipitation. The second approach is especially useful for poorly sampled areas or in case of malfunction in the rain gauges. The simulations run with precipitation data of a single event (24 h), but the simulation period finishes at the end of the prediction period. Then, the outflow series of the sub-basin is reconstructed with data from each simulation. Starting from the flow series of the forecast simulations, we sum the run-off coming from the previous events and the increments in the baseflow. The resulting flow series represent the flow caused by the forecast events and also the previous ones. The outflow of each sub-basin is then used as input data for the simulation of the river reach. Although this method makes the simulation process more complex, the number of parameters needed is the same as for the regular SCS-CN method, making it especially suitable for those cases were the data needed to parametrize other methods is not available. The proposed method can be applied to simulate longer periods of time than SCS-CN without losing infiltration capacity and adjusting the best CN for each event of that period.
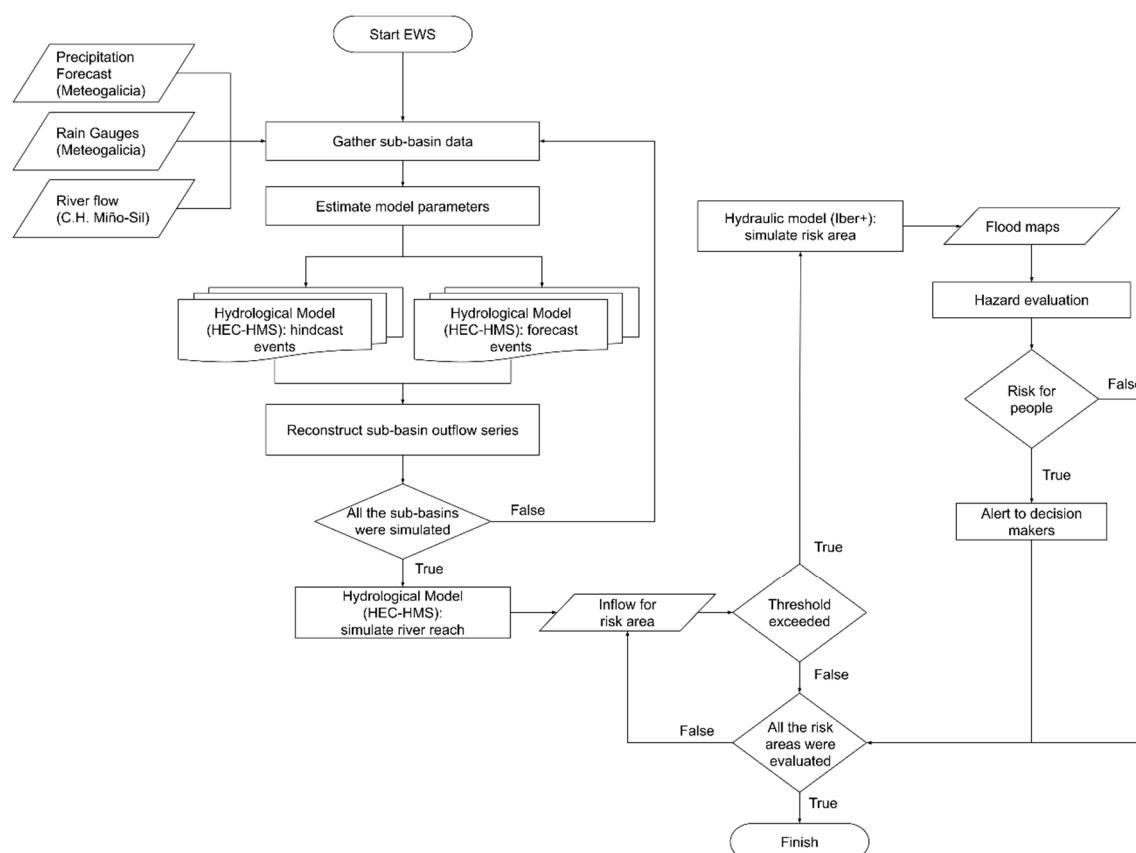

**Figure 2.** Flowchart of the Miño River Flood Alert System (MIDAS) early warning system.
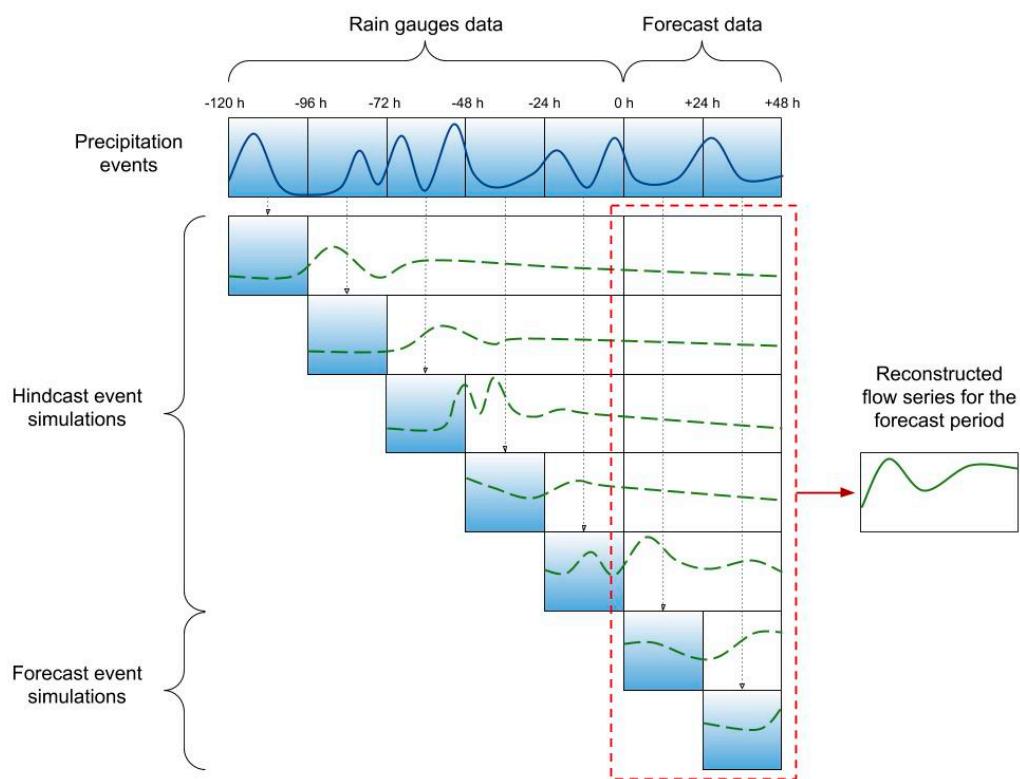
Figure 3. Procedure for the reconstruction of the sub-basin outflow series.

After the hydrological simulation, the hydraulic simulations with Iber+ are triggered for each zone of interest. Simulations can be launched on every EWS execution or only when a certain safety threshold is exceeded in the input flow. This second approach can be especially suitable when the computation resources are limited. Each of the simulations performed with Iber+ are executed in a GPU Nvidia RTX 2080ti and require a computational time in the order of 3–5 min for each 24 h of simulated time for meshes of the order of 70–200 k elements. If the modelled reach contains a high number of zones that require hydraulic simulations, these could be launched in parallel in different GPUs to maintain execution times in reasonable values. These simulations produce water depth and hazard maps for maximum values and hourly values. When some hazard is detected in the hydraulic simulations, an alert is issued to the corresponding decision makers.

2.3. Statistical Parameters to Analyze the Performance of the EWS

Some of the most used statistical parameters to evaluate the performance of this type of procedure were considered to analyze the accuracy of the EWS. In this sense, Spearman coefficient of correlation (r), the ratio of the root mean square error to the standard deviation of the observed data (RSR), Nash–Sutcliffe efficiency coefficient (NSE), percent bias (PBIAS), and Taylor diagrams (in terms of the normalized standard deviation (σ_n), normalized centered root-mean-square difference (E_n) and correlation (R)) [64], were evaluated following the equations presented below.

$$RSR = \frac{\sqrt{\Sigma_{i=1}^{N}\left(Q_i^{obs} - Q_i^{for}\right)^2}}{\sqrt{\Sigma_{i=1}^{N}\left(Q_i^{obs} - \overline{Q^{obs}}\right)^2}} \qquad (8)$$

$$\text{NSE} = 1 - \frac{\sum_{i=1}^{N}\left(Q_i^{obs} - Q_i^{for}\right)^2}{\sum_{i=1}^{N}\left(Q_i^{obs} - \overline{Q^{obs}}\right)^2} \tag{9}$$

$$\text{PBIAS} = \frac{\sum_{i=1}^{N}\left(Q_i^{obs} - Q_i^{for}\right)}{\sum_{i=1}^{N}\left(Q_i^{obs}\right)} \times 100 \tag{10}$$

$$\sigma_n = \frac{\sqrt{\dfrac{\sum_{i=1}^{N}\left(Q_i^{for} - \overline{Q^{for}}\right)^2}{N}}}{\sigma_{obs}} \tag{11}$$

$$E_n = \frac{\sqrt{\dfrac{\sum_{i=1}^{N}\left[\left(Q_i^{for} - \overline{Q^{obs}}\right) - \left(Q_i^{obs} - \overline{Q^{obs}}\right)\right]^2}{N}}}{\sigma_{obs}} \tag{12}$$

$$R = \frac{\sum_{i=1}^{N}\left[\left(Q_i^{for} - \overline{Q^{for}}\right)\left(Q_i^{obs} - \overline{Q^{obs}}\right)\right]}{N\,\sigma_{for}\,\sigma_{obs}} \tag{13}$$

where Q^{obs} is the observed value, Q^{for} is the forecasted value, N is the total number of observed values, barred variables refer to mean values, the subscript n refers to the normalized parameter, subscript *i* refers to the different samples, and σ is the standard deviation.

3. Results and Discussion

3.1. Hydrologic Module Validation

The accuracy of the hydrological module of the proposed EWS was evaluated by means of statistical analysis respect to measured river discharge data at Lugo and Ourense stations.

When considering the flow data for the studied period in a daily scale, the results show that the model is able to predict with a good degree of accuracy the annual cycle of river flow. The model also shows a clear contrast between dry and humid seasons and is able to reflect the high flow conditions (Figure 4). This is corroborated when the statistical parameters are analyzed (Table 3). It is considered a "very good" performance of a hydrological model when the NSE value surpasses 0.7 and the absolute value of PBIAS is lower than 25% [65–67], these parameters are widely fulfilled in the present case (NSE > 0.85 and |PBIAS| < 10%) (Table 3). Moreover, the high correlation detected (r > 0.90) along with the low RSR values obtained (RSR < 0.40) also indicate a good performance according to Moriasi et al. [65], showing the robustness of the methodology developed.

Table 3. Statistical analysis of the performance of forecasted river flow respect to measured data during the period 2012–2019.

Statistical Parameter	Lugo Station	Ourense Station
r	0.97	0.90
RSR	0.30	0.37
NSE	0.91	0.86
PBIAS	−5.25	−9.89

In addition, the most important flood events, whose prediction supposes the critical and main objective from the point of view of EWS, were also evaluated at hourly scale. In particular, those events where river flow surpassed the 99th percentile for the period under study, were taken into account. Analyzing the statistical parameters, model results show an accurate prediction of the real river flow under extreme conditions (Table 4). NSE (>0.7) and |PBIAS| (<7%) values indicate "very good" accuracy

of the model developed, which is corroborated by the high correlation values (r > 0.85) along with the low RSR values obtained (<0.55) (Table 4). The statistical analysis shows the high accuracy of the proposed EWS to detect and predict the most important flood events. This is also corroborated in Figure 5, which shows the comparison between the model forecast and the measured river flow under the most dangerous events occurring during the study period at Lugo and Ourense stations. Despite small deviations in the shape of the series, both events were well predicted both in terms of time and magnitude.
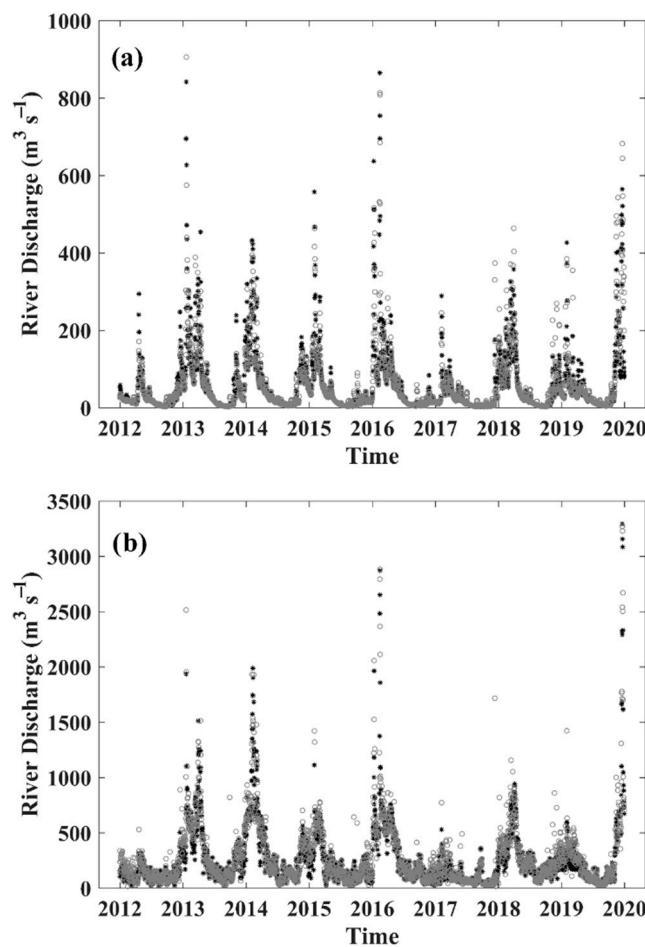
Figure 4. Forecasted (grey circles) and measured (black asterisks) river flow at (**a**) Lugo and (**b**) Ourense locations for the whole period 2012–2019.

Table 4. Statistical analysis of the performance of forecasted river flow respect to measured data considering only extreme river flow events above the 99th percentile over the period 2012–2019.

| Statistical Parameter	Lugo Station	Ourense Station
r	0.89	0.85
RSR	0.42	0.54
NSE	0.82	0.71
PBIAS	−4.08	−6.70
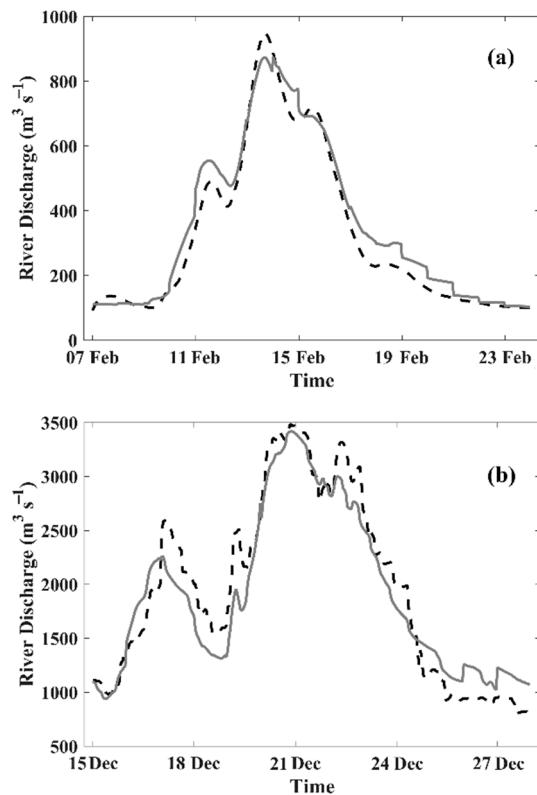
Figure 5. Forecasted (grey line) and measured (dotted black line) river flow for the most dangerous events occurred at (**a**) Lugo (2016) and (**b**) Ourense (2019) locations.

It is important to remark that PBIAS values obtained in Tables 3 and 4 indicate a slight overestimation of predicted river flows. However, this type of systems should be conservative, because the consequences of under-estimate flood events may cause severe damage [8]. In addition, statistical values show a better performance at the Lugo than at the Ourense location, both in the total series length and in the particular case of the extreme events. This is partly due to the fact that the Miño River at Lugo is under natural flow conditions, whereas several dams control the flow at Ourense, along with the fact that the Sil Tiver influence at Ourense is not directly simulated but extrapolated from Miño data as described above.

When the obtained results are compared with other systems developed in nearby areas, the hydrological module shows its potential as river flow predictor. Fraga et al. [8] also use the HEC-HMS hydrological model, but under a continuous simulation mode in nearby areas. This model is divided into the following two stages: hindcast and forecast. During the first stage (hindcast), the hydrological model assimilates several measured data of 30 previous days in order to reproduce the initial soil moisture content at the beginning of the second stage (forecast) by tuning the parameters associated to the infiltration of the terrain. In the forecast stage the river hydrograph is computed. Although it is important to note Fraga et al. [8] focused on small Galician catchments, where hydrological forecast is more sensitive to precipitation distribution, improved values of r and PBIAS are obtained in the present work, which reinforces the robustness of the methodology proposed. Vieira et al. [68] presented a flood forecast system developed for two sub-basins of Ave River catchment (northern Portugal) based on the Delft FEWS platform. In this case the meteorological data are obtained from an ensemble of precipitation predictions. The implementation of the hydrologic and hydraulic models was carried out using SOBEK software. The calibration of the model parameters was developed starting from values obtained in the literature and soil characteristics and using the RRL toolkit. Statistical values are only shown for historical validation period and daily time scale. The EWS developed in the present work achieves statistical values remarking a higher accuracy even considering

forecast simulations. A grid-based distributed rainfall-runoff model (GFWS) was implemented in the Guadalhorce basin (southern Spain) [69], which is a poorly gauged basin, in contrast to the basin analyzed in our work. The parameters of the model obtained for a small area of the catchment are transferred to the entire catchment of the Guadalhorce River. In this work, flood warnings obtained with the GFWS model were compared with those obtained with the low resolution EFAS model. Higher values of Nash efficiency for forecasted flood events were also obtained in the present work, showing the good performance of MIDAS.

Therefore, it can be concluded that the hydrological module of the system presented in this work is able to forecast with high accuracy both the annual cycle (continuous) river flow and the extreme flood events. The little discrepancies observed between forecasted numerical and measured river flow can be due to the limitations of these systems—mainly, the simplification of the modelled system and the uncertainty of the forecast precipitation data. In this sense, the hydrological model supposes an approximate representation of the reality; therefore, some discrepancies will be expected. Moreover, although precipitation forecasts provided by MeteoGalicia predict rainy events accurately as shown in González-Cao et al. [33], the precipitation forecast is, by itself, a source of uncertainty, especially under situations of extreme rainfall events [33]. Additionally, rain representation provided by rain gauges can suppose another source of uncertainty. Rain gauges can provide a biased view of the area under scope since they do not cover the subareas of the basins uniformly. Even with these limitations, the system proposed provides a great accuracy to predict river flow, especially under extreme flood events, which impacts the performance MIDAS.

3.2. Hydraulic Module Validation

Once the predicted water flow was shown to reproduce the actual flows with a high accuracy, the performance of the hydraulic model was also analyzed. For that, the extent and the water elevation of the flood event registered on December 2019 (from 15 to 31) were computed using Iber+ [17]. On the one hand, four control areas were defined at study area (green rectangles A, B, C, and D in Figure 1c) to analyze the accuracy of the numerical results of the flood extension. The control area A corresponds with a public sport facility (swimming pool) and its recreational zone, the control area B is located near the gauge station N010 of the CHMS, control area C corresponds with a public path near a national road, and control area D corresponds to a public thermal baths area. All these control areas are located near the riverbank and are usually frequented by pedestrians, therefore being of special interest to issue an alert. On the other hand, the water elevation registered at the gauge station located in control area B was used to compute the accuracy of the numerical water elevation obtained with Iber+ by means of Taylor and NSE vs. |PBIAS| diagrams.

Figure 6 shows the extent of the flood obtained in the numerical simulations along with photographs taken at the control areas during the flood event within the interval 16:00–16:30 UTC (for all times throughout the paper) on 20 December 2019. Images corresponding to non-flood conditions are also depicted, corresponding to a flow within the interval 262–267 $m^3 s^{-1}$ (data obtained from CHMS). The comparison between photographs taken during the flood and non-flood conditions reveals the magnitude of the flood event. Visually, the numerical results are quite similar to the field data. The image of the flood event at the control area A shows that the swimming pool is fully covered by the flood. The numerical result shows the same situation: the water covers the facility and reaches the wall near the road. The image at the control area B shows that the flood reaches the building of the gauge station. The results obtained with Iber+ shows an equivalent situation. Also, the numerical results show that the flood almost reaches the road. The image corresponding to the control area C shows the same situation as the numerical results: the flood extension covers the public path and almost reaches the road. Finally, the thermal baths are fully covered by the flood, and the water almost reaches the path behind them, as can be observed in the photograph at the control area D. The numerical results are equivalent to the image of real event.

Figure 6. Results of flood extent for each control area. First column depicts non-flood conditions (Q = 265 m³ s⁻¹ that corresponds to 66th percentile) at the control areas. The second column shows the images (photographs taken by one of the authors) of the flood event (20 December 2019, 16:00 to 16:30) and the third column shows the numerical results obtained with Iber+. Red arrows in the third column represent the direction of the photograph taken during the flood event.

The results of the time series of water elevation registered from 15 December 2019 to 31 December 2019 at gauge station N010 are depicted in Figure 7. This figure shows the time series of the water elevation registered at the gauge station (control area B, Figure 1) and the numerical time series obtained with Iber+. Visually, numerical results are similar to field data, which shows the capability of the model to forecast water elevation accordingly with the river discharge forecast (Figure 5b). The Taylor and the NSE vs. |PBIAS| diagrams are also depicted. Taylor diagram shows that the values of σ_n, E_n, and R are 1.00, 0.22, and 0.97, respectively. The values of NSE and |PBIAS| depicted in the NSE vs. PBIAS diagram are 0.94 and 0.15%. The circle marker in this diagram denotes a negative value of PBIAS. Shaded areas of the diagram refer to the accuracy level of the results according to the criterion of Yilmaz and Onoz [67]. In this case, the results are in the range corresponding to "very good" improving to some extent the results obtained by Fraga et al. [8] in nearby small catchments. In summary, both diagrams show high agreement between numerical and registered values of water elevation. This indicates that the system can reproduce this kind of events with high accuracy and can be used to issue an alert.
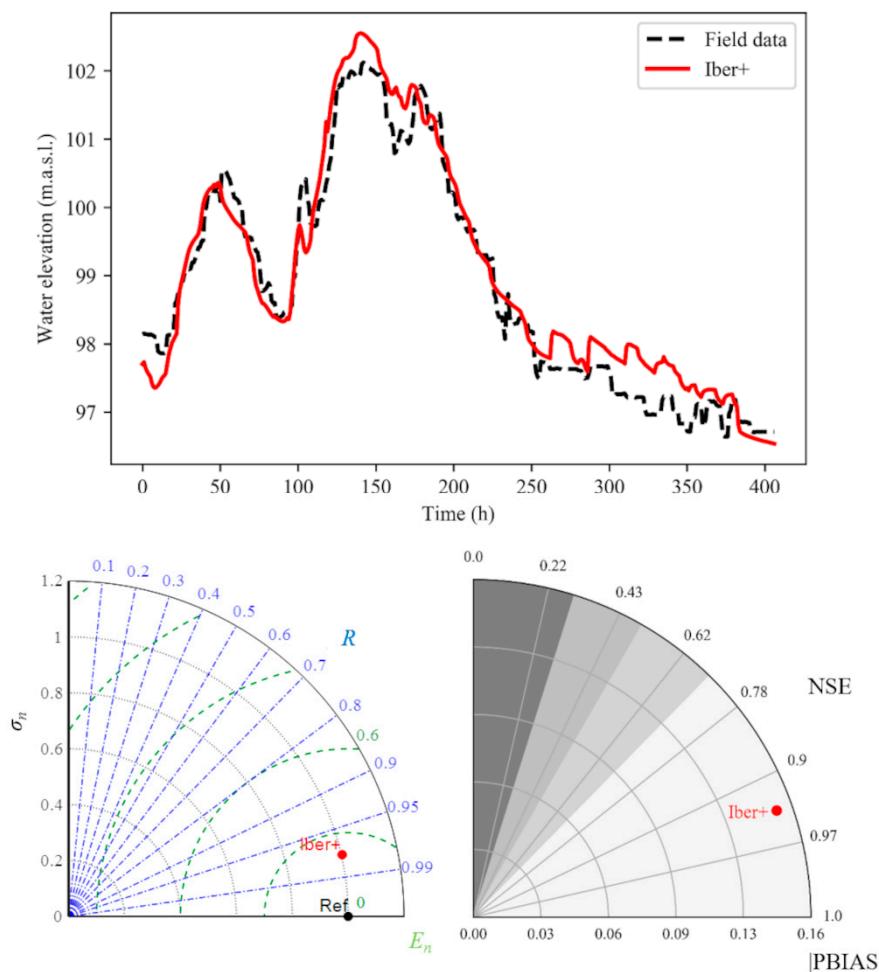
Figure 7. Upper panel: time series of water elevation obtained at gauge station (black dashed line) and using Iber+ (red line) at gauge station N010 (control area B). Lower left panel: Taylor diagram of the time series of water elevation obtained with Iber+ using the field data as reference. Lower right panel: NSE vs. |PBIAS| diagram of the numerical water elevation using the field data as reference. Marker denotes negatives values of PBIAS and shaded areas refer to the criterion of [44] to define the level of accuracy: very good, good, satisfactory, and unsatisfactory (from lighter to darker).

3.3. Analysis and Predictability of River Flood Risk Situations

The validation of the hydrologic and hydraulic modules has shown a good forecast of river flow under flood events as well as a good representation of the areas affected by these flood events. The final step is to focus the system on critical areas from the point of view of critical infrastructures damage. The ultimate goal is that the system can report dangerous situations well in advance so that decision makers can take appropriate mitigation measures. In this sense, the hazard in the city of Ourense is analyzed. The criterion employed to define a hazard area was the presented by Cox et al. [70]. First of all, after am in-depth study of the hazardous areas affected by different water flows, the alert status was divided in four levels (see Table 5 for specific thresholds). The first of them, pre-activation, defines the average conditions of the river reach. In the second level, activation, some risk can occur in areas accessible to but not frequented by pedestrians. The third state is pre-alert where hazard can occur in areas frequented by pedestrians. The last one is the alert state where the flood causes hazard in areas frequented by pedestrians and severe damage to infrastructures.

Table 5. Hazard thresholds for the Minho River at Ourense City.

Alert Status	Min Flow (m^3/s)	Max Flow (m^3/s)
Pre-activation	0	511
Activation	511	1253
Pre-alert	1253	2000
Alert	2000	-

Once the thresholds that define the different risk situations were evaluated, an analysis of the observed and predicted flow was carried out in order to assess the EWS's ability to accurately issue an alert. In this sense, 17 alerts should have been issued based on the flow observed during the eight years of the simulated period (Table 6). The EWS accurately predicted 13 of them, 4 were missed and 5 were false positives. This means that the proposed system is able to predict the 76% of the alerts and 28% of the alerts issued are false positives, which shows a good performance of the designed system in terms of predictability of risk situations.

The system is able to issue an alert based on the hydrological forecast and also to provide detailed water depth, water velocity, and hazard maps in terms of hydraulic model simulations. Both the maximum values and hourly evolution of these parameters are obtained for the forecasted period. Figure 8 shows the average number of days under hazard conditions per year. It can be observed that many zones frequented during outdoor activities can be affected by hazard conditions for more than 10 days per year on average, emphasizing the requirement of an accurate EWS. Figure 9 shows the maximum extension of the hazardous area for each of the warning levels defined in Table 5. This map shows that the differences between two warning levels are minimal in certain areas, illustrating the necessity of accompanying the alert reports with detailed flood maps. This information provides a detailed view of areas affected under the respective flood event. In this way, decision makers can have enough data to evaluate the situation correctly and take more precise and effective measures, diminishing the damage provoking by floods.
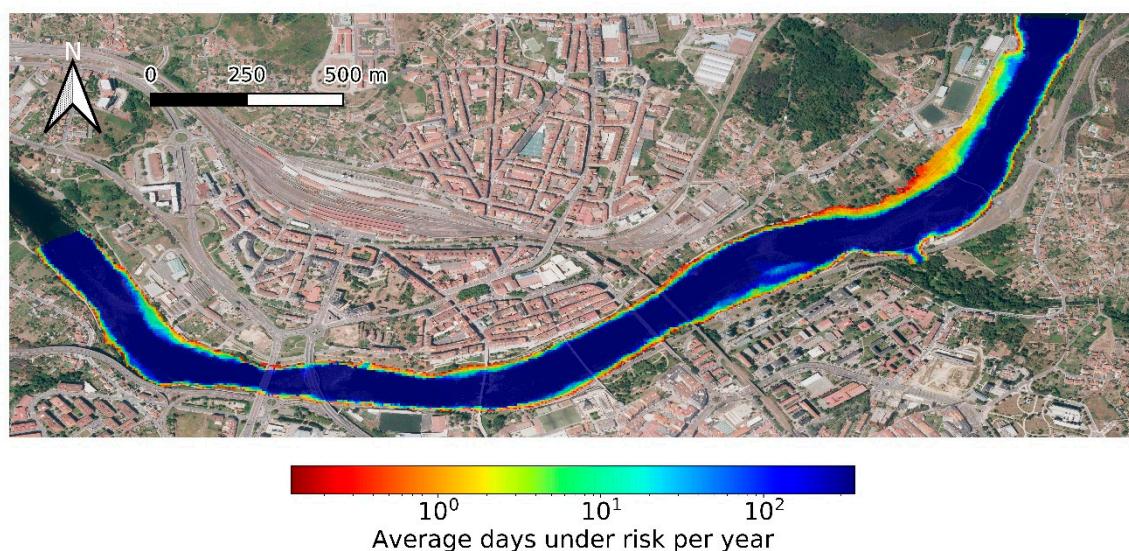

Figure 8. Average days under risk per year.

Table 6. Contingency table with the number of alert events observed and forecasted. From a total of 2922 events analyzed, 13 were correct positives, five false-positives, four missed alerts, and 2900 correct negatives.

		Observed		Total
		Alert	Non-Alert	
Forecasted	**Alert**	13	5	18
	Non-Alert	4	2900	2904
	Total	17	2905	2922
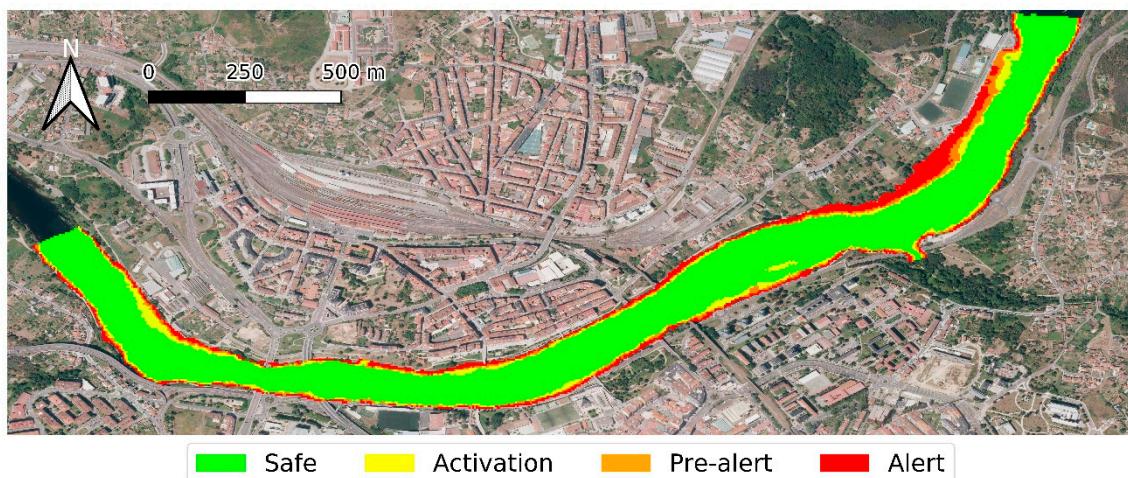
Figure 9. Maximum extents of the hazardous area under the different alert stages.

4. Conclusions

This paper presents MIDAS, a new flood early warning system based on integrated hydrologic and hydraulic models. The automatic operation of MIDAS is governed by a set of Python scripts. Starting from precipitation forecast, the rainfall-runoff process is simulated with the HEC-HMS model to obtain the predicted river flow, which is used to run Iber+ for each area of interest. An alert is issued when a hazardous situation is detected.

The results show the high capability of the system to accurately forecast river flows and are representative of the flooded areas under extreme events. In fact, MIDAS was able to predict 13 of the 17 alerts that took place over the study period, which is a good performance in terms of predictability of risk situations. This good performance of the system, together with the capability to provide detailed hazard maps, can provide information to decision makers to evaluate flood situations and to take precise and effective measures that mitigate the damage associated with these events.

Finally, it is important to remark that, despite MIDAS being applied to the Miño River, its implementation in other basins is straightforward. This approach is especially valuable for areas with low sampling resources, showing its potential in future applications.

Author Contributions: D.F.-N., O.G.-F., J.G.-C., and M.G.-G. conceived the study; D.F.-N., O.G.-F., and J.G.-C. conducted the research and developed the software; M.G.-G. supervised the research development; D.F.-N., O.G.-F., J.G.-C., C.d.G., J.A.R.-S., C.R.d.P., and M.G.-G. analyzed the results; D.F.-N., O.G.-F., J.G.-C., and M.G.-G. wrote the manuscript; C.d.G., J.A.R.-S., and C.R.d.P supervised the research, provided field data, and revised the manuscript. All authors have read and agreed to the published version of the manuscript.

Conflicts of Interest: The authors declare no conflict of interest. The funders had no role in the design of the study; in the collection, analyses, or interpretation of data; in the writing of the manuscript, or in the decision to publish the results.

References

1. Berghuijs, W.R.; Aalbers, E.E.; Larsen, J.R.; Trancoso, R.; Woods, R.A. Recent changes in extreme floods across multiple continents. *Environ. Res. Lett.* **2017**, *12*, 114035. [CrossRef]
2. Passerotti, G.; Massazza, G.; Pezzoli, A.; Bigi, V.; Zsótér, E.; Rosso, M. Hydrological Model Application in the Sirba River: Early Warning System and GloFAS Improvements. *Water* **2020**, *12*, 620. [CrossRef]
3. Rosburg, T.T.; Nelson, P.A.; Bledsoe, B.P. Effects of Urbanization on Flow Duration and Stream Flashiness: A Case Study of Puget Sound Streams, Western Washington, USA. *JAWRA J. Am. Water Resour. Assoc.* **2017**, *53*, 493–507. [CrossRef]
4. Booth, D.B.; Bledsoe, B.P. Streams and urbanization. In *The Water Environment of Cities*; Baker, L.A., Ed.; Springer: Boston, MA, USA, 2009; pp. 93–123. ISBN 978-0-387-84891-4.
5. Arnell, N.W.; Gosling, S.N. The impacts of climate change on river flood risk at the global scale. *Clim. Chang.* **2016**, *134*, 387–401. [CrossRef]
6. Liu, C.; Guo, L.; Ye, L.; Zhang, S.; Zhao, Y.; Song, T. A review of advances in China's flash flood early-warning system. *Nat. Hazards* **2018**, *92*, 619–634. [CrossRef]
7. Alfieri, L.; Salamon, P.; Pappenberger, F.; Wetterhall, F.; Thielen, J. Operational early warning systems for water-related hazards in Europe. *Environ. Sci. Policy* **2012**, *21*, 35–49. [CrossRef]
8. Fraga, I.; Cea, L.; Puertas, J. MERLIN: A flood hazard forecasting system for coastal river reaches. *Nat. Hazards* **2020**, *100*, 1171–1193. [CrossRef]
9. Groisman, P.Y.; Knight, R.W.; Easterling, D.R.; Karl, T.R.; Hegerl, G.C.; Razuvaev, V.N. Trends in Intense Precipitation in the Climate Record. *J. Clim.* **2005**, *18*, 1326–1350. [CrossRef]
10. Beniston, M. Trends in joint quantiles of temperature and precipitation in Europe since 1901 and projected for 2100. *Geophys. Res. Lett.* **2009**, *36*. [CrossRef]
11. Morss, R.E.; Wilhelmi, O.V.; Meehl, G.A.; Dilling, L. Improving Societal Outcomes of Extreme Weather in a Changing Climate: An Integrated Perspective. *Annu. Rev. Environ. Resour.* **2011**, *36*, 1–25. [CrossRef]
12. Wallemacq, P.; House, R.; Below, R.; McLean, D. *Economic Losses, Poverty & Disasters: 1998–2017*; Centre for Research on the Epidemiology of Disasters (CRED); United Nations Office for Disaster Risk Reduction (UNISDR): Brussels, Belgium, 2018.
13. Han, L.; Xu, Y.; Pan, G.; Deng, X.; Hu, C.; Xu, H.; Shi, H. Changing properties of precipitation extremes in the urban areas, Yangtze River Delta, China, during 1957–2013. *Nat. Hazards* **2015**, *79*, 437–454. [CrossRef]
14. Hallegatte, S. *A Cost Effective Solution to Reduce Disaster Losses in Developing Countries: Hydro-Meteorological Services, Early Warning, and Evacuation*; The World Bank: Washington, DC, USA, 2012.
15. Cools, J.; Innocenti, D.; O'Brien, S. Lessons from flood early warning systems. *Environ. Sci. Policy* **2016**, *58*, 117–122. [CrossRef]
16. Borga, M.; Anagnostou, E.N.; Blöschl, G.; Creutin, J.-D. Flash flood forecasting, warning and risk management: The HYDRATE project. *Environ. Sci. Policy* **2011**, *14*, 834–844. [CrossRef]
17. Thielen, J.; Bartholmes, J.; Ramos, M.-H.; de Roo, A. The European Flood Alert System—Part 1: Concept and development. *Hydrol. Earth Syst. Sci.* **2009**, *13*, 125–140. [CrossRef]
18. Pappenberger, F.; Thielen, J.; Medico, M.D. The impact of weather forecast improvements on large scale hydrology: Analysing a decade of forecasts of the European Flood Alert System. *Hydrol. Process.* **2011**, *25*, 1091–1113. [CrossRef]
19. Alfieri, L.; Pappenberger, F.; Wetterhall, F.; Haiden, T.; Richardson, D.; Salamon, P. Evaluation of ensemble streamflow predictions in Europe. *J. Hydrol.* **2014**, *517*, 913–922. [CrossRef]
20. Bartholmes, J.C.; Thielen, J.; Ramos, M.H.; Gentilini, S. The european flood alert system EFAS—Part 2: Statistical skill assessment of probabilistic and deterministic operational forecasts. *Hydrol. Earth Syst. Sci.* **2009**, *13*, 141–153. [CrossRef]
21. Alfieri, L.; Burek, P.; Dutra, E.; Krzeminski, B.; Muraro, D.; Thielen, J.; Pappenberger, F. GloFAS—global ensemble streamflow forecasting and flood early warning. *Hydrol. Earth Syst. Sci.* **2013**, *17*, 1161–1175. [CrossRef]

22. Corral, C.; Berenguer, M.; Sempere-Torres, D.; Poletti, L.; Silvestro, F.; Rebora, N. Comparison of two early warning systems for regional flash flood hazard forecasting. *J. Hydrol.* **2019**, *572*, 603–619. [CrossRef]

23. Hirpa, F.A.; Hopson, T.M.; De Groeve, T.; Brakenridge, G.R.; Gebremichael, M.; Restrepo, P.J. Upstream satellite remote sensing for river discharge forecasting: Application to major rivers in South Asia. *Remote Sens. Environ.* **2013**, *131*, 140–151. [CrossRef]

24. Cloke, H.L.; Pappenberger, F. Ensemble flood forecasting: A review. *J. Hydrol.* **2009**, *375*, 613–626. [CrossRef]

25. De Luca, D.L.; Biondi, D.; Capparelli, G.; Galasso, L.; Versace, P. Mathematical models for early warning systems. *IAHS-AISH Publ.* **2010**, *340*, 485–495.

26. Baudoin, M.-A.; Henly-Shepard, S.; Fernando, N.; Sitati, A.; Zommers, Z. *Early Warning Systems and Livelihood Resilience: Exploring Opportunities for Community Participation*; UNU-EHS Working Paper Series; United Nations University: Tokyo, Japan, 2014.

27. Ritter, J.; Berenguer, M.; Corral, C.; Park, S.; Sempere-Torres, D. ReAFFIRM: Real-time Assessment of Flash Flood Impacts–a Regional high-resolution Method. *Environ. Int.* **2020**, *136*, 105375. [CrossRef] [PubMed]

28. Collier, C.G. Flash flood forecasting: What are the limits of predictability? *Q. J. R. Meteorol. Soc.* **2007**, *133*, 3–23. [CrossRef]

29. de Gonzalo, C.; Robredo, J.C.; Mintegui, J.Á. Semidistributed hydrologic model for flood risk assessment in the Pejibaye River Basin, Costa Rica. *J. Hydrol. Eng.* **2012**, *17*, 1333–1344. [CrossRef]

30. De Silva, M.; Weerakoon, S.B.; Herath, S. Modeling of event and continuous flow hydrographs with HEC–HMS: Case study in the Kelani River Basin, Sri Lanka. *J. Hydrol. Eng.* **2014**, *19*, 800–806. [CrossRef]

31. Cea, L.; Fraga, I. Incorporating antecedent moisture conditions and intraevent variability of rainfall on flood frequency analysis in poorly gauged basins. *Water Resour. Res.* **2018**, *54*, 8774–8791. [CrossRef]

32. Wang, D. A new probability density function for spatial distribution of soil water storage capacity leads to the SCS curve number method. *Hydrol. Earth Syst. Sci.* **2018**, *22*, 6567–6578. [CrossRef]

33. González-Cao, J.; García-Feal, O.; Fernández-Nóvoa, D.; Domínguez-Alonso, J.M.; Gómez-Gesteira, M. Towards an automatic early warning system of flood hazards based on precipitation forecast: The case of the Miño River (NW Spain). *Nat. Hazards Earth Syst. Sci.* **2019**, *19*. [CrossRef]

34. Nguyen, P.; Thorstensen, A.; Sorooshian, S.; Hsu, K.; AghaKouchak, A.; Sanders, B.; Koren, V.; Cui, Z.; Smith, M. A high resolution coupled hydrologic–hydraulic model (HiResFlood-UCI) for flash flood modeling. *J. Hydrol.* **2016**, *541*, 401–420. [CrossRef]

35. García-Feal, O.; González-Cao, J.; Gómez-Gesteira, M.; Cea, L.; Domínguez, J.M.; Formella, A. An accelerated tool for flood modelling based on Iber. *Water* **2018**, *10*, 1459. [CrossRef]

36. Bladé, E.; Cea, L.; Corestein, G.; Escolano, E.; Puertas, J.; Vázquez-Cendón, E.; Dolz, J.; Coll, A. Iber: Herramienta de simulación numérica del flujo en ríos. *Revista Internacional de Métodos Numéricos para Cálculo y Diseño en Ingeniería* **2014**, *30*, 1–10. [CrossRef]

37. Lorenzo, M.N.; Alvarez, I. Climate change patterns in precipitation over Spain using CORDEX projections for 2021–2050. *Sci. Total Environ.* **2020**, 138024. [CrossRef] [PubMed]

38. Fernández-Nóvoa, D.; deCastro, M.; Des, M.; Costoya, X.; Mendes, R.; Gómez-Gesteira, M. Characterization of Iberian turbid plumes by means of synoptic patterns obtained through MODIS imagery. *J. Sea Res.* **2017**, *126*, 12–25. [CrossRef]

39. deCastro, M.; Lorenzo, N.; Taboada, J.J.; Sarmiento, M.; Alvarez, I.; Gomez-Gesteira, M. Influence of teleconnection patterns on precipitation variability and on river flow regimes in the Miño River basin (NW Iberian Peninsula). *Clim. Res.* **2006**, *32*, 63–73. [CrossRef]

40. Feldman, A.D. *Hydrologic Modeling System HEC-HMS: Technical Reference Manual*; USA Army Corps of Engineers: Washington, DC, USA; Hydrologic Engineering Center: Davis, CA, USA, 2000.

41. Scharffenberg, B.; Bartles, M.; Brauer, T.; Fleming, M.; Karlovits, G. *Hydrologic Modeling System (HEC-HMS). User's Manual: Version 4.3*; USA Army Corps of Engineers: Washington, DC, USA, 2018.

42. USA Department of Agriculture. *Soil Conservation Service (SCS) "Hydrology" National Engineering Handbook*; Section 4; USA Department of Agriculture: Washington, DC, USA, 1985.

43. USA Department of Agriculture. *NRCS: Natural Resources Conservation Service National Engineering Handbook*; Chapter 16 Hydrographs; USA Department of Agriculture: Washington, DC, USA, 2007.

44. Soil Conservation Service (SCS). *Technical Release 55: Urban Hydrology for Small Watersheds*, 2nd ed.; USA Department of Agriculture: Washington, DC, USA, 1986.

45. Stewart, D.; Canfield, E.; Hawkins, R. Curve number determination methods and uncertainty in hydrologic soil groups from semiarid watershed data. *J. Hydrol. Eng.* **2012**, *17*, 1180–1187. [CrossRef]

46. Ministerio para la Transición Ecológica y el Reto Demográfico, Spain Government Modelo SIMPA 2019. Periodo de simulación: 1940/41 a 2017/18. Available online: https://www.miteco.gob.es/es/agua/temas/evaluacion-de-los-recursos-hidricos/evaluacion-recursos-hidricos-regimen-natural/ (accessed on 22 June 2020).

47. Nelder, J.A.; Mead, R. A simplex method for function minimization. *Comput. J.* **1965**, *7*, 308–313. [CrossRef]

48. Brocca, L.; Melone, F.; Moramarco, T. Distributed rainfall-runoff modelling for flood frequency estimation and flood forecasting. *Hydrol. Process.* **2011**, *25*, 2801–2813. [CrossRef]

49. Massari, C.; Brocca, L.; Tarpanelli, A.; Moramarco, T. Data assimilation of satellite soil moisture into rainfall-runoff modelling: A complex recipe? *Remote Sens.* **2015**, *7*, 11403–11433. [CrossRef]

50. USA Department of Agriculture. *Soil Conservation Service (SCS) "Hydrology" National Engineering Handbook*; Section 4; USA Department of Agriculture: Washington, DC, USA, 1972.

51. Hope, A.S.; Schulze, R.E. Improved estimates of stormflow volume using the SCS curve number method. In Proceedings of the International Symposium on Rainfall-runoff Modeling, Mississippi State University, Starkville, MS, USA, 18–21 May 1982.

52. Schulze, R.E. *The Use of Soil Moisture Budgeting to Improve Stormflow Estimates by the SCS Curve Number Method*; University of Natal, Department of Agricultural Engineering: Pietermaritzburg, South Africa, 1982; p. 63.

53. Spain Government. *Ministerio de Fomento: Norma 5.2-IC Drenaje Superficial de la Instrucción de Carreteras*; Spain Government: Madrid, Spain, 2016.

54. Cunge, J.A. On the subject of a flood propagation computation method (Musklngum method). *J. Hydraul. Res.* **1969**, *7*, 205–230. [CrossRef]

55. NVIDIA Corporation CUDA C++ Programming Guide. Available online: https://docs.nvidia.com/cuda/pdf/CUDA_C_Programming_Guide.pdf (accessed on 22 June 2020).

56. García-Feal, O.; Cea, L.; González-Cao, J.; Domínguez, J.M.; Gómez-Gesteira, M. IberWQ: A GPU Accelerated Tool for 2D Water Quality Modeling in Rivers and Estuaries. *Water* **2020**, *12*, 413. [CrossRef]

57. Gonzalez-Cao, J.; García-Feal, O.; Cea, L.; Gómez-Gesteira, M. Preservation of the cultural heritage from floods using the numerical code Iber. In Proceedings of the First International Electronic Conference on the Hydrological Cycle, Ourense, Spain, 12–16 November 2017; MDPI: Basel, Switzerland, 2017; p. 4843.

58. SNCZI. *Guía Metodológica para el Desarrollo del Sistema Nacional de Cartografía de Zonas Inundables*; Ministerio de Agricultura, Alimentación y Medio Ambiente, Centro de Publicaciones: Madrid, Spain, 2011; ISBN 978-84-491-1136-5.

59. Erpicum, S.; Dewals, B.; Archambeau, P.; Detrembleur, S.; Pirotton, M. Detailed inundation modelling using high resolution DEMs. *Eng. Appl. Comput. Fluid Mech.* **2010**, *4*, 196–208. [CrossRef]

60. Liu, Y.; Zhou, J.; Song, L.; Zou, Q.; Liao, L.; Wang, Y. Numerical modelling of free-surface shallow flows over irregular topography with complex geometry. *Appl. Math. Model.* **2013**, *37*, 9482–9498. [CrossRef]

61. Segura-Beltrán, F.; Sanchis-Ibor, C.; Morales-Hernández, M.; González-Sanchis, M.; Bussi, G.; Ortiz, E. Using post-flood surveys and geomorphologic mapping to evaluate hydrological and hydraulic models: The flash flood of the Girona River (Spain) in 2007. *J. Hydrol.* **2016**, *541*, 310–329. [CrossRef]

62. Courant, R.; Friedrichs, K.; Lewy, H. Über die partiellen Differenzengleichungen der mathematischen Physik. *Math. Ann.* **1928**, *100*, 32–74. [CrossRef]

63. Eckhardt, K. How to construct recursive digital filters for baseflow separation. *Hydrol. Process.* **2005**, *19*, 507–515. [CrossRef]

64. Taylor, K.E. Summarizing multiple aspects of model performance in a single diagram. *J. Geophys. Res. Atmos.* **2001**, *106*, 7183–7192. [CrossRef]

65. Moriasi, D.; Arnold, J.G.; Liew, M.W.V.; Bingner, R.; Harmel, R.D.; Veith, T.L. Model Evaluation Guidelines for Systematic Quantification of Accuracy in Watershed Simulations. *Trans. ASABE* **2007**, *50*, 885–900. [CrossRef]

66. Kalin, L.; Isik, S.; Schoonover, J.E.; Lockaby, B.G. Predicting Water Quality in Unmonitored Watersheds Using Artificial Neural Networks. *J. Environ. Qual.* **2010**, *39*, 1429–1440. [CrossRef]

67. Yilmaz, M.U.; Onoz, B. A Comparative Study of Statistical Methods for Daily Streamflow Estimation at Ungauged Basins in Turkey. *Water* **2020**, *12*, 459. [CrossRef]

68. Vieira, J.M.P.; Pinho, J.L.S.; Vieira, B.F.V.; Vieira, L.M.V. Flood forecast technological platforms: An adaptive response to extreme events. In Proceedings of the WEC2019: World Engineers Convention 2019, Melbourne, Australia, 18–22 November 2019; Engineers Australia: Melbourne, Australia, 2019; p. 1966.

69. Versini, P.-A.; Berenguer, M.; Corral, C.; Sempere-Torres, D. An operational flood warning system for poorly gauged basins: Demonstration in the Guadalhorce basin (Spain). *Nat. Hazards* **2014**, *71*, 1355–1378. [CrossRef]

70. Cox, R.J.; Shand, T.D.; Blacka, M.J. *Australian Rainfall and Runoff Revision Project 10: Appropriate Safety Criteria for Peopl*; Water Research Laboratory, The University of New South Wales: New South Wales, Australia, 2010; ISBN 978-0-85825-945-4.

5.5. Numerical reconstruction of historical extreme floods: The Badajoz event of 1876

El artículo mostrado a continuación fue enviado a la revista revisada por pares "Journal of Hydrology" y se encuentra en revisión en el momento de escribir esta tesis.

Numerical reconstruction of historical extreme floods: The Badajoz event of 1876

José González-Cao[a,*], Diego Fernández-Nóvoa[a], Orlando García-Feal[a], Jose R. Figueira[b], José M. Vaquero[c], Ricardo M. Trigo[d], Moncho Gómez-Gesteira[a]

[a] *Environmental Physics Laboratory (EPhysLab), CIM-UVIGO, Universidade de Vigo, Campus As Lagoas s/n, 32004 Ourense, Spain*

[b] *Departamento de Expresión Gráfica, Universidad de Extremadura, Mérida, Spain*

[c] *Departamento de Física, Universidad de Extremadura, Mérida, Spain*

[d] *Instituto Dom Luiz (IDL), Faculdade de Ciências da Universidade de Lisboa, 1749-016 Lisbon, Portugal*

* Corresponding author: jgcao@uvigo.es

ABSTRACT: Since ancient times, flood events have had a direct impact on the development of civilizations. Some of these events have been studied from different scientific perspectives. However, in recent years the development of numerical tools has facilitated the ability of the numerical simulations to reproduce such events. These simulations allow a far better understanding of historical flood events, which can contribute to mitigate their negative effects under similar future scenarios. In this work the catastrophic flood occurred on December 7, 1876 in Badajoz (Spain) is analysed. For this purpose, the work was divided into three stages. First, an estimation of the precipitation recorded in the Guadiana river basin during the days prior to the event was carried out starting from the precipitation data recorded in two precipitation stations located in the western part of the basin; second, Guadiana river flow in Badajoz was estimated using the HEC-HMS code. The peak flow associated to this event was also estimated by means of the numerical code Iber+, obtaining similar values with both methodologies; third, the flood registered on December 1876 in Badajoz was reproduced using the numerical code Iber+. Numerical results also presented a good agreement with water marks and some descriptions of the event shown in the literature. Results obtained in this work demonstrate the ability of the methodology to reproduce historical extreme flood events even when the availability of input data is relatively scarce.

Keywords: *historical extreme flood; precipitation reconstruction; river discharge estimation; numerical simulation; Iber+.*

1. Introduction

Since ancient times, flood events have had a direct impact on the development of civilizations. Some of these events had positive societal impacts for the inhabitants of the affected areas, such as the seasonal flooding of the Nile River in the ancient Egypt. However, other events had negative effects, e.g., the recurrent floods of the Tiber River that destroyed the harbour of Ostia (near Rome) cited by the roman author Strabo (64–58 BC–21–25 AD) (Strabo, Geographica, V, 3, 5.). Some of these events are especially interesting due to their extraordinary influence on the population and have been studied from different scientific perspectives (e.g. historical, archaeological, meteorology and hydrology associated to early instrumentation). However, in recent years the development of numerical tools strongly supported on the increment of the computational power, has facilitated the ability of the numerical simulations to reproduce such events with increasing detail. These simulations allow a far better understanding of those historical flood events helping to fill in the many gaps of the characterization -in space and time- of the phenomenon. This deeper understanding of the flood's behaviour contributes with relevant information to mitigate their negative effects in similar future scenarios, an increasing likely possibility under climate change scenarios (Groisman et al., 2005; Beniston, 2009; Morss et al., 2011). According to the European directive (E.D. 2007/60/C), it is important to assess potential risks derived from flood events through *"a description of the floods which have occurred in the past and which have had significant adverse impacts on human health, the environment, cultural heritage and economic activity and for which the likelihood of similar future events is still relevant, including their flood extent and conveyance routes and an assessment of the adverse impacts they have entailed"*, which highlights the importance of reproducing historical extreme flood events with numerical tools.

There are many examples of historical flood events often resulting from extreme precipitation. Eulenstein and Kellerer-Pirklbauer (2020) provide a detailed description of the flood event registered in central Europe (catchments of Danube and Vltava) in 1572 using geographical sources. Similarly, references to a devastating flood in Rome can be found in Long (2008) where the author shows the destruction of the ancient aqueduct of Acqua Vergin in 1570. Elleder et al. (2020a) describe the catastrophic flood registered in the Bohemian-Moravian Highlands (Sázava River) in 1714. They estimated several hyetographs to reproduce the rain-runoff processes associated to this event. Elleder et al. (2020b) analysed the consequences of torrential rain event in the city of Prague (Czech Republic) in 1872. The reproduction of the event was carried out using only historical documentation due to the lack of hydrological data.

More recent flash flood event analysis in Iberia are described by Portugués-Molla et al. (2016) for Valencia (Spain) in 1957 and Trigo et al (2014) for the Lisbon region (Portugal) in 1967. Portugués-Molla et al. (2016) developed a methodology based on aerial photographs and pictures at street level to understand the consequences of the flood. One of the most important years concerning the number and magnitude of flood events is 1876. The bibliography shows several examples of floods registered all over the world that year. For example, Benito et al. (1996) and Font (1988) describe floods of several rivers in Spain. Portugal (González, 1995), France and Germany (Pfister et al., 1999) were also affected during that year by extreme floods. The catastrophic flood occurred on December 7, 1876 in Badajoz (Spain) and most of the southern sector of Portugal is described by Trigo et al. (2014), where the authors highlight the record precipitation values in the Iberian Peninsula during the weeks prior to the event and the exceptional synoptic situation. In Zamora-Cabanillas (1999), the reader can also find a detailed description of the 1876 event as well as other floods registered in the Guadiana river basin. Both works describe record rainfall values in late November and early December 1876. The estimated values of discharge flow corresponding to the flood registered on December, 1876 in Badajoz reach values greater than 10,000 m^3 s^{-1} (Ortega, 2007). Other flood events as floods registered in January, 1940 or March, 1947 reach values near 7,500 and 8,000 m^3 s^{-1}, respectively (Ortega, 2007). This highlights the magnitude of the event registered in December 1876 as a clear outlier.

The present work tries to reproduce the flood registered on December 7, 1876 in the city of Badajoz by means of the numerical model *Iber* (Bladé et al., 2014). For this purpose, the work was mainly divided into three stages. In the first one, an estimation of the precipitation recorded in the Guadiana river basin during the days prior to the event was carried out. Next, the rain-runoff process associated with the precipitation in the Guadiana basin was computed using the HEC-HMS code (Feldman, 2000, Scharffenberg et al., 2018) in order to obtain an estimation of the discharge flow near Badajoz during the event. Once this discharge flow was computed, *Iber* was used to reproduce the flood event registered in Badajoz. This document is organized as follows: in Section 2, the area of study, the evolution of the flood event and its consequences in Iberian Peninsula are described. Then, the precipitation dataset along with the hydrological and hydraulic models used in this paper are briefly described in Section 3. The estimation of the precipitation, the discharge flow and the numerical simulations of the flood event in Badajoz are shown in Section 4. Finally, Section 5 provides the main conclusions of this work.

2. Area of study and description of the event

The area of study is shown in Fig. 1. The basin of the international Guadiana River is located in the south-west of the Iberian Peninsula and covers an area of 67,000 km^2 with 83% located in Spain and 17% in Portugal (Agência Portuguesa do Ambiente, 2016). Guadiana River has an approximate length of 800 km and an elevation that ranges from 0 at the mouth situated in the Atlantic Ocean near Ayamonte (just in the border between Portugal and Spain) to near 1300 m.a.s.l. at the headwaters in Lagunas de Ruidera (Castilla-La Mancha autonomous region, Spain). At present, the natural regime of the river is highly modified by the construction of several large reservoirs that have increased the storage capacity of the river since 1954 to 14,000 hm^3 (SNIRH, 2019). The reader can find a detailed description of the Guadiana River in Canuto et al. (2019).

Several cities are bathed by the Guadiana River, including Badajoz, located close to the border with Portugal (Fig.1) and with a population of about 200.000 inhabitants. The city was founded more than a thousand years ago taking advantage of its strategic location, just when the Guadiana River passes between two important promontories. Originally, the city stood on the promontory to the SE, growing on the left bank of the river. Later, the medieval bridge (the Palmas Bridge), of great strategic importance, was built around the year 1460. To defend this bridge, the hill located to the NW was later populated building a fort. At the end of the 19th century, with the arrival of the railroad network, the city also grew on the right bank of the river. Old maps of the city of Badajoz corresponding to 1873 and 1946 (Fig. 1(e) and Fig. 1(g), respectively) reveal that the city has not been subjected to great human-induced changes until mid-twentieth century. Figure 1(e) shows a map of Badajoz made by the Army General Staff Corps and published by the "Depósito de la Guerra" in 1873 at a scale of 1: 5000. We have used the copy preserved in the Municipal Historical Archive of Badajoz. A digital copy is available on the Extremadura government's spatial data infrastructure website (http://www.ideex.es/Geoportal/). Fig. 1(g) shows the image corresponding to Badajoz of the photographic flight of mainland Spain carried out in 1946 by the US Army Map Service. It is an orthophoto also available on the Extremadura government's spatial data infrastructure website. During the last century, several modifications can be observed in current maps of the area (Fig. 1(f)). These changes mainly influence the land uses of the area of study. The elevation of the city of Badajoz ranges from 150 to 250 m.a.s.l. (Fig. 1(d)).

Badajoz suffered the greatest adverse effects derived from the floods registered in 1876. The last week of November and the first week of December 1876 correspond to one of the most extreme precipitation events in the western Iberian Peninsula (Fragoso et al. 2010). The consequences of this extreme precipitation event were observed in the riverbanks of several rivers including the three large southern Iberian rivers (Tagus, Guadiana and Guadalquivir). Newspapers in Portugal and Spain describe the effects of this event, including titles such as: *Diario de Noticias*, *Gazeta do Algarve*, *El Magisterio Extremeño*, *La Ilustración Española* or *La Crónica*. In Portugal, according to *Diario de Noticias,* almost 95% of the buildings in the village of Mértola were razed to the ground. *Gazeta do Algarve* published that some villages as Pomarão, Alcoutim, or Vila Real de Santo Antonio almost disappeared. In central-west Spain, the effects of the flood were similar to those registered in Portugal. The communications by road and train were destroyed in Badajoz. The Palmas Bridge, main communication between both sides of the river, collapsed for the first time since 1603 (González, 1995). Other cities like Mérida, Olivenza or Talavera la Real, were also affected by this event. Loureiro (2009) refers to the event registered in 1876 like the biggest one in the basin of the Tagus River in Portugal. Fig. 2 shows some of the main scenes derived by the flood in Badajoz, Mérida and Seville published by *La Ilustración Española y Americana*.

The evolution of this extreme precipitation event is related to successive frontal systems with Atlantic origin often with a SW to NE direction (Trigo et al., 2014). A detailed synoptic analysis of the event is also shown in Zamora-Cabanillas (1999) and Trigo et al. (2014). Therefore, this outstating flood cannot be considered as a flash-flood event as in that case the delay between the precipitation peak and the flood only takes a few hours (Yin et al., 2016).
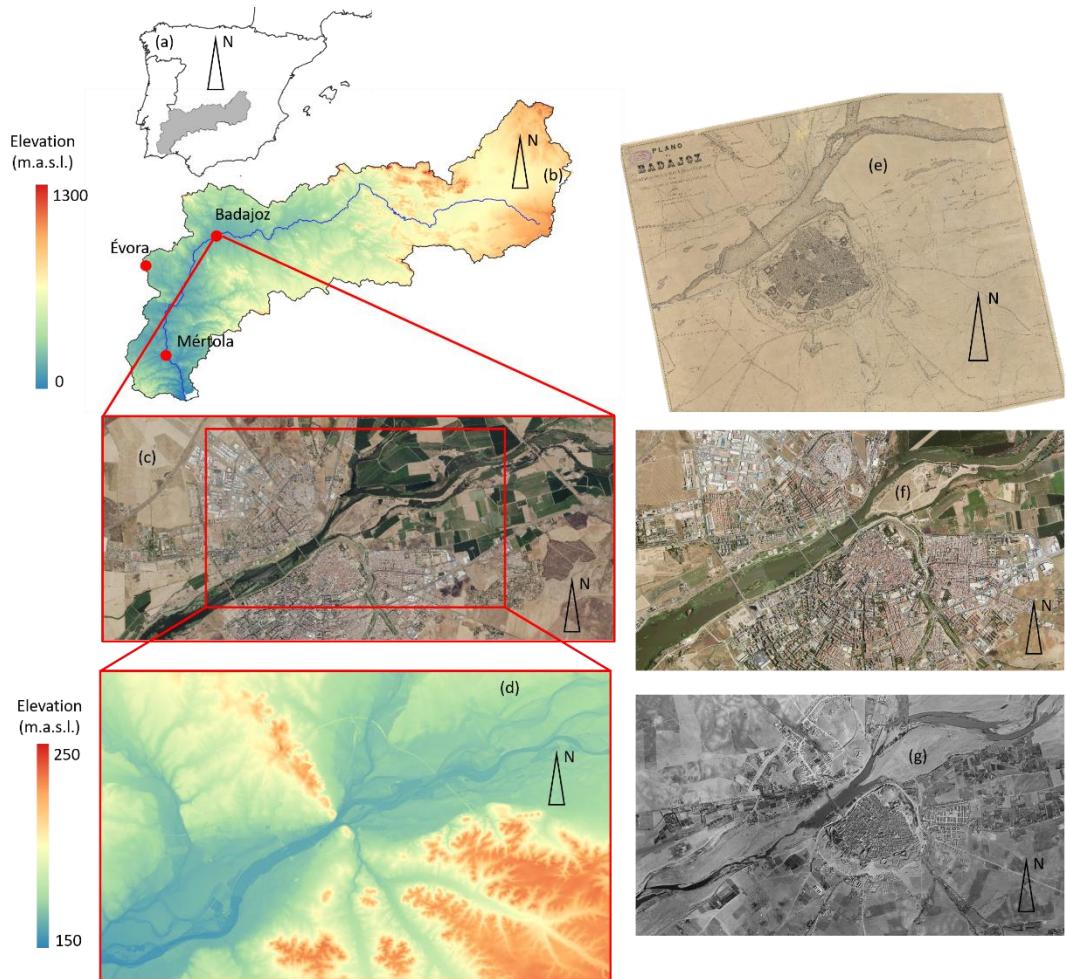
Figure 1. Area under study: (a) shows the location of the Guadiana river basin (shaded area) in the Iberian Peninsula; (b) shows the map of elevation of the basin and the location of the cities of Badajoz (Spain), Mértola (Portugal) and Évora (Portugal). The blue line refers to the river course; (c) shows a general view of the city of Badajoz; (d) elevation map of the city of Badajoz; (e), (f) and (g) shows a detail of the city of Badajoz obtained from a historical map of 1873, from a current view and from an American flight of 1946, respectively.
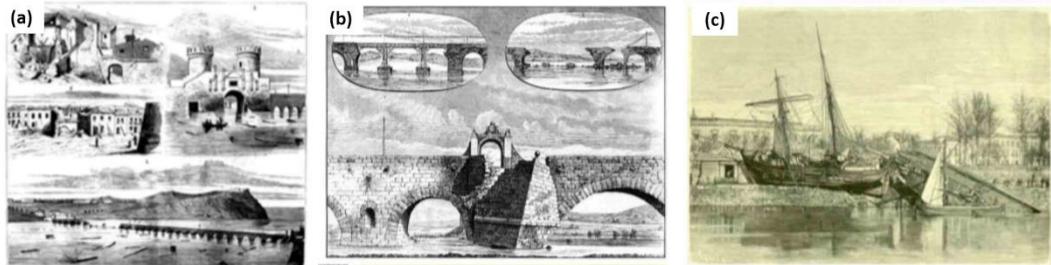
Figure 2. Effects of the flood event in Badajoz (a), Roman bridge of Mérida (b) and Seville (c). The pictures were published in the newspaper *La Ilustración Española y Americana*.

3. Material and methods

3.1. SPAIN02 precipitation dataset

The observational gridded data set Spain02 (Herrera et al., 2012; Herrera et al., 2016; Kotlarski, 2019) was employed in this work to estimate the precipitation recorded during the event of December, 1876. Spain02 represents a set of high-resolution daily precipitation and temperature gridded datasets developed for Spain. A high dense network of around 2500 control stations for the period 1950-2007 was selected from the Spanish Meteorological Agency (AEMET) in order to build the gridded dataset that divides the domain in pixels that act as virtual stations. In this work, the data were obtained from the version Spain02 v5, whose grid has a regular spatial resolution of 0.1º (~10 km).

3.2. Hydrologic model: HEC-HMS

The semi-distributed model HEC-HMS (Feldman, 2000, Scharffenberg et al., 2018) was used to model the basin of the Guadiana River to estimate the flow registered in Badajoz. HEC-HMS is mainly divided into four components: (i) an analytical model: calculation of direct runoff and channel routing; (ii) a basin model: representation of hydrological elements in a watershed; (iii) a system to manage input data and store data; (iv) a post-processing tool to report and illustrate simulation results. This modelling system is one of the most used for hydrological procedures (Du et al., 2012; Akter and Ahmed, 2015; Bai et al., 2019; Fraga et al., 2020). Two main processes were taken into account in the methodology implemented in this work: loss and transform methods. Loss method calculates the infiltration capability of the terrain and provides the excess rainfall which is subject to surface runoff whereas the transform method models the way of how to convert this excess rainfall into runoff. The Soil Conservation Service (SCS) curve number was selected to account for the infiltration process. This method implements the curve number (CN) methodology for incremental losses, since it was designed to calculate the infiltration during periods of heavy rainfall (NRCS, 2007). In the same way, the SCS unit hydrograph was also selected for the transformation method. More information about the loss and transform methods used in this work are detailed in NRCS (2007).

The basin of the Guadiana River that flows into Badajoz was divided in 35 sub-basins according to their topographic features and these were considered for the implementation of the HEC-HMS model (Fig.3).
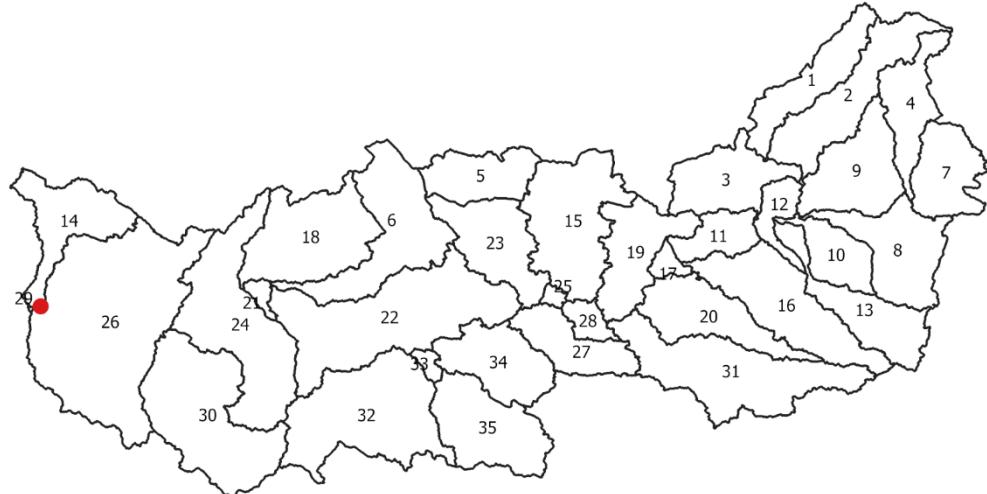

Figure 3. HEC-HMS model of the basin of Guadiana River to Badajoz (red dot). Numbers refer to id. of each sub-basin.

The lag time of the sub-basins and the reaches was calculated using data from Digital Terrain Models extracted by means of HEC-GeoHMS tool (HEC-GeoHMS Geospatial Hydrologic Modeling Extension: User's Manual, Version 5, 2010). The mean curve numbers (CN) defined for each sub-basin were obtained from the work of the Guadiana Hydrographic Confederation (CHG) (C.H.G., 2003). These curve numbers were adapted to the humid soil conditions corresponding to the specific wet state observed prior to the event of 1876 (Trigo et al., 2014) by applying the classic equations of the SCS (Eqs. 1-2). Lag time (T_l) values were obtained using Eq. (3) starting from the concentration time (T_c). T_c was computed according to Ministerio de Fomento (2016) and following the SCS recommendations according to which the T_l is about 0.6 times the T_c. In Eq. (3), L_c is the length of the longest flow channel (km) and J_c is its corresponding slope (m/m). The values of CNs range from 85 to 90 and the values of the T_l range from 53 to 1645 min.

$$CN_{dry} = \frac{4.2 \cdot CN_{average}}{10 - 0.058 \cdot CN_{average}} \tag{1}$$

$$CN_{moist} = \frac{23 \cdot CN_{average}}{10 - 0.13 \cdot CN_{average}} \qquad (2)$$

$$T_l = 0.18 \cdot L_c^{0.76} \cdot J_c^{-0.19} \qquad (3)$$

The spatial precipitation was averaged for each of the sub-basins of the Guadiana River defined in the HEC-HMS model (Fig. 3) in order to calculate the discharge flow in Badajoz, which constitute the input for *Iber+*.

3.3. Hydraulic model: *Iber+*

Iber is a numerical tool that solves the 2D depth-averaged shallow water equations using the finite volume method (Bladé et al., 2014). Recently some of the authors have developed (García-Feal et al., 2018) a new implementation of the model in C++, namely *Iber+*, and CUDA (NVIDIA, 2020). This new implementation of *Iber* improves the efficiency of the simulations by achieving a speed-up of two-orders of magnitude with the same precision by using GPU (graphical processing unit) computing HPC (high performance computing) techniques. These optimizations rise the possibility to employ the model in large spatio-temporal domains (González-Cao et al., 2019a, García-Feal et al., 2020; Fernández-Nóvoa et al., 2020) or time constrained applications (Fraga et al., 2020, González-Cao et al., 2019b). Also, some applications of hybridization of *Iber+* with other CFD models like DualSPHysics (Crespo et al., 2015) were developed to analyse the behaviour of dams in high precipitation scenarios (González-Cao et al., 2018). The software package is freely available and can be downloaded from its official website (https://iberaula.es). It also includes a GUI (graphical user interface) with pre-processing and post-processing tools. In the present study, *Iber+* was applied to analyse the flood events registered in Mértola (Portugal) and in Badajoz (Spain) in December 1876.

The numerical simulations in the area of Mértola, located in southern Portugal (Fig. 1), were carried out using a digital terrain model (DTM) with a spatial resolution of 30 m. The study area, along with the numerical domain and the topography of the area are shown in Fig. 4, with two control lines separated by 0.5 km. The light blue line in Figure 4 refers to an upstream cross section of the domain, while the red line refers to the downstream cross section where the historical reference water mark is located. Unfortunately, this second control line is located

in an area where several houses have been built on the left riverbank since the flood event. This modification that implies the narrowing of the area, together with the low spatial resolution of the DTM, led us to discard this control line. Therefore, the control line used to define the water elevation corresponds the upstream light blue line, where no significant landscape modifications were observed. The land uses defined in this case were estimated from georeferenced images. The physical simulation time was equal to 24 h and a 1st order numerical scheme was used, being the dry-wet limit was set equal to 0.01 m.
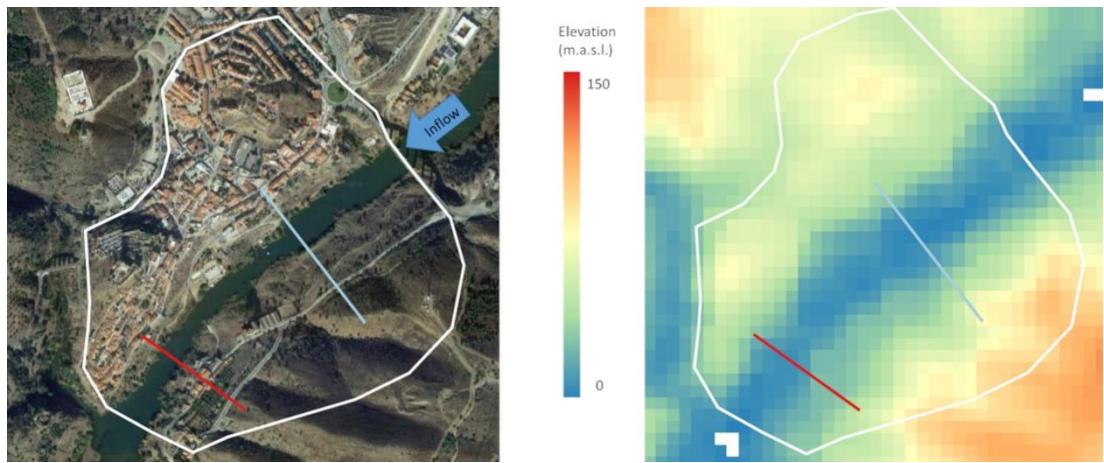

Figure 4. Numerical domain defined in Mértola (left) along with elevation map of the area (right). The blue line corresponds to the control line and the red line corresponds to the location of the historical reference water mark.

The domain defined in *Iber+* along with its inlet and outlet boundaries in the area of Badajoz is shown in Fig. 5. The domain was discretized using a mesh of unstructured triangles whose characteristic size varies from 0.5 to 20 m, with a total of 723,971elements. The georeferenced image corresponds to the so-called American Photographic Flight of 1946. This image will be used as a reference to define the different areas of the numerical domain in *Iber+* since it corresponds reasonably well with the map of the city of Badajoz of 1873 (Fig. 6). We acknowledge that the image from the American flight shows some built-up areas that did not appear in the 1873 map but, in general, the coincidence of most spatial features level is quite high. The American Flight image is also used as a reference since it covers a wider area and has a higher resolution.

Once the numerical domain was defined and discretized, it was adapted to the topography of the area by means of a DTM with a spatial resolution of 5 m that was retrieved from the Spanish National Geographic Institute (IGN- https://www.ign.es/web/ign/portal).
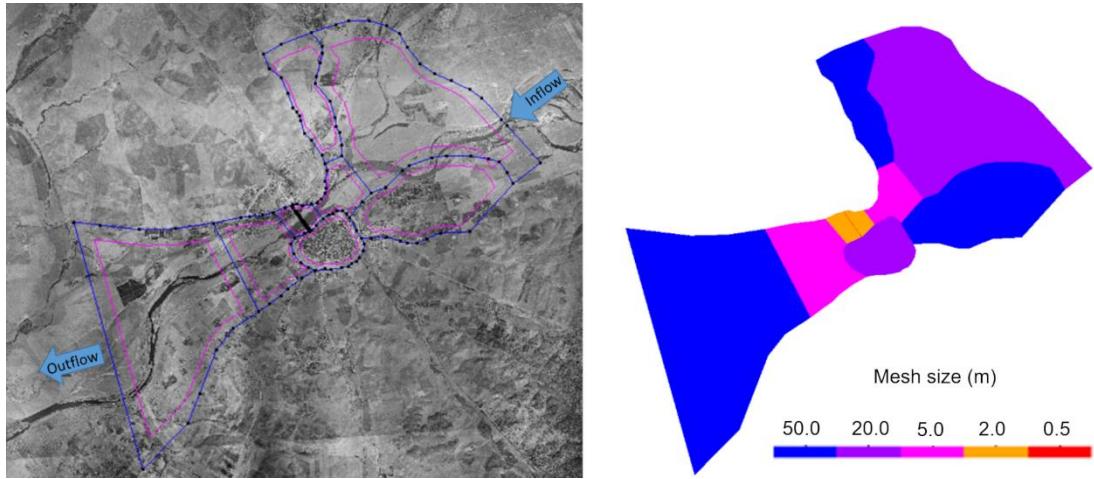
Figure 5. Numerical domain of Badajoz defined in *Iber+* (left) and characteristic sizes of the elements of the mesh (right). The image of the left panel corresponds to the American flight of 1946.

Figure 6. Details of the orthophoto obtained from the American Flight of 1946 (Left) and of the Badajoz plan of the year 1873 (Right).

There is also information regarding the topography of the area corresponding to the early and mid-20th century. A comparison study was carried out between the elevations provided by these sources and those provided by the 5 m DTM obtained from IGN. The differences obtained in this study are small and can be attributed to the lack of precision of the older data. Therefore, the 5 m DTM from IGN was taken as reference data. Land uses defined in the domain correspond to a constant generic value.

Special mention should be made to the Palmas Bridge located in Badajoz. This bridge has undergone various modifications over the past centuries. It is about 590 m long and had 28 semicircular arches in 1873, each of them with different dimensions. To model this bridge, the tools implemented in *Iber+* were used to design this type of structures. A series of simplification hypotheses were assumed given its constructive complexity. In particular, it was assumed that the height of the bridge is constant (in the plans it can be seen that it presents a slightly higher elevation in the centre than in the ends). Additionally, straight arches were considered since the definition of semicircular arches in *Iber+* would imply a significant increase in the complexity of the design that does not make sense in view of the existing uncertainty in the rest of input data. To study the effect of the bridge on the elevation of the water level, two simulations were carried out. The first one was made taking into account the bridge and the second one without the bridge. All the numerical simulations of *Iber+* in Mértola and Badajoz were performed using a GPU GeForce RTX 2080 Ti.

4. Results and discussion

4.1. Estimation of the precipitation in the basin of the Guadiana river during November and December 1876

The availability of data to estimate the precipitation in the entire basin is relatively scarce. This is a common caveat when dealing with historical floods, increasing the difficulty of reproducing such events. The most relevant available data are the daily rainfall registered in the stations of Évora and Badajoz (see Fig. 1) during the months of November and December 1876. These stations are located in the western part of the basin, so the spatial distribution precipitation in the entire Guadiana basin is difficult to determine. Fortunately, precipitation obtained in the western and central sectors of Iberia is highly dependent on synoptic circulation with a western component (Trigo and DaCamara, 2000; Cortesi et al., 2013). In fact, it has been shown objectively that almost all the winter precipitation that falls in the Guadiana basin is prevenient from the Atlantic Ocean (Gimeno et al., 2010, 2012). This fact implies that before reaching the basin upstream Badajoz the same synoptic systems have already affected the precipitation over both cities (Évora and Badajoz). This physical constrain provides the rationale to stablish a simple statistical model linking the daily precipitation observed in these two stations with the daily precipitation in the Guadiana basin (upstream Badajoz) where no stations were available in 1876. A multivariate linear adjustment was carried out for the months

of November and December of the years 1977-2006 between the rainfall registered in Évora and Badajoz with the rainfall associated to each of the pixels (571 virtual stations) belonging to the upstream Guadiana river basin according to Spain02. For each virtual station of the basin, a relationship was established with the two reference stations (Évora and Badajoz) as follows:

$$P_i = c_0^i + c_1^i P_B + c_2^i P_E \qquad (4)$$

where P_i is the precipitation at point i (virtual station), P_B and P_E are the precipitation at Badajoz and Évora, respectively. Three coefficients were obtained from the linear fitting depending on the point i.

The correlation between each virtual station and the two reference stations was calculated using Eq. 4 for each grid point in the Spanish Guadiana basin, obtaining a higher correlation (higher than 0.9) for the virtual stations near the border between both countries and decreasing slightly to correlation values around 0.7 for the virtual stations situated close to the Guadiana headwaters. The mean correlation, which was calculated by averaging the correlation at the 571 virtual stations, is equal to 0.9, with an average statistical significance greater than 99%.

Assuming that the relationship between the reference stations and the virtual stations derived from Spain02 has been maintained over time, the rainfall at each point of the basin in 1876 can be inferred by applying Eq. (4) to the rainfall data collected in Évora and Badajoz during the event under study in 1876. Thus, the spatial distribution of the precipitation corresponding to the event can be obtained and used as the input data of the HEC-HMS model in order to calculate the flow in Badajoz associated with this precipitation.

4.2. Estimation of the discharge flow in Badajoz in December 1876

Once the precipitation in the entire basin of the Guadiana River is obtained, the time series of the discharge flow at Badajoz can be estimated using the HEC-RAS model. The daily discharge flow at Badajoz between the 4th and 12th of December 1876 can be observed in Fig. 7. The average peak flow (solid line) reaches a value of 12,688 m^3 s^{-1} whilst the peaks obtained by varying +/- 5% the curve numbers of each sub-basin (dashed lines) range from 13,260 m^3 s^{-1} to 12,071 m^3 s^{-1}. These values are consistent with those estimated in the literature (Zamora-Cabanillas, 1999; Ortega, 2007; Trigo et al., 2014). In addition, the peak value obtained with HEC-HMS model is reached on day 7, which is also in full agreement with the references (Zamora-Cabanillas, 1999; Ortega, 2007; Trigo et al., 2014).
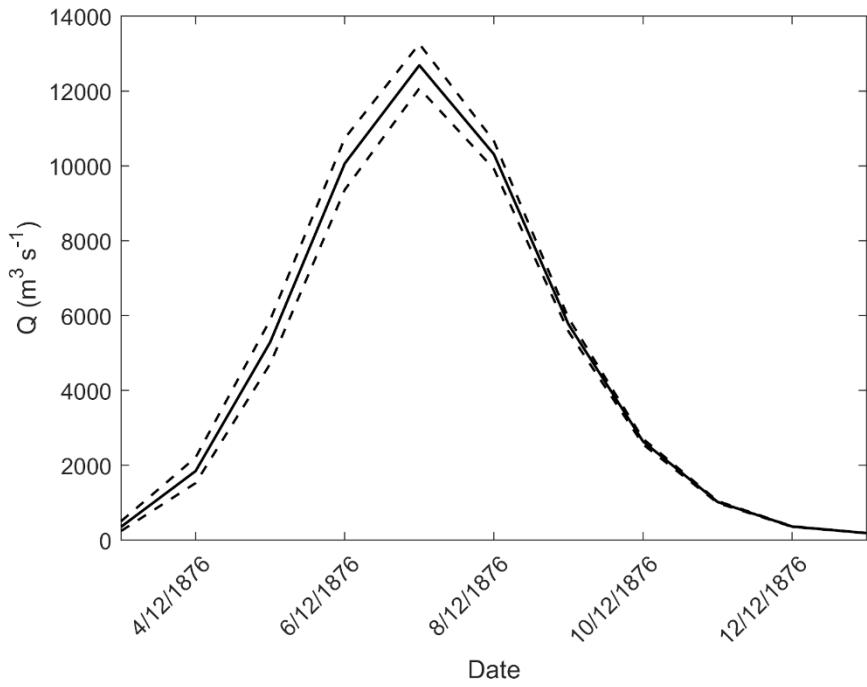
Figure 7. Discharge flow at Badajoz between 4 and 12 December 1876 obtained with the HEC-HMS model. Solid line refers to mean value and dashed lines refer to discharges obtained by changing the CN of the sub-basins +/- 5%.

Despite these promising results, we acknowledge that additional lines of research must be considered in order to confirm their overall validity. Thus, a series of numerical simulations were carried out using *Iber+* to reproduce the flood recorded during the event in the Portuguese town of Mértola, which is located more than 150 km downstream Badajoz (Fig. 1). There is a major reservoir between both cities but constructed in the early years of the twentieth first century. Local records indicate that the water level reached a value close to 30 meters above sea level. Furthermore, 5 km upstream from Mértola is located the Rocha da Galé gauge station. This station has flow data from 1978 to 1994, a period in which there were no important reservoirs built between Badajoz and Mértola, which allows establishing a relationship between the flow registered at Rocha da Galé and Badajoz during that period. Results confirm the very high one-to-one correspondence between average daily river flows of both time series as clearly observed in Fig. 8. This figure depicts the time series of flows registered in Badajoz and Rocha da Galé along with the scatter plot of the flow in Badajoz versus the flow in Rocha da Galé.
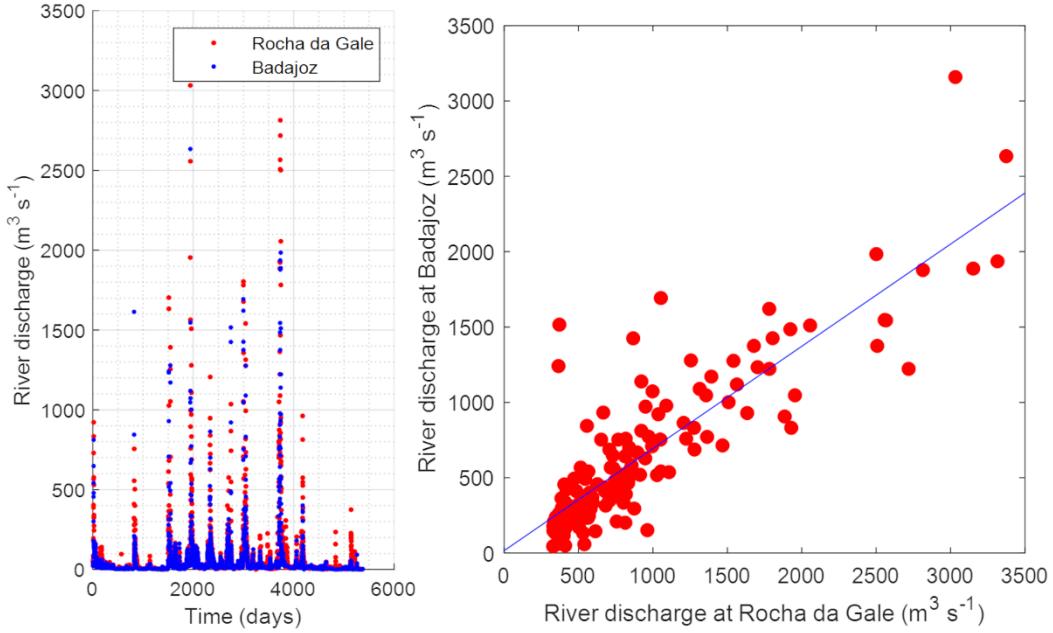
Figure 8. Time series of average daily river flow registered in Badajoz (blue dots) and Rocha da Galé (red dots) from October 1978 to October 1994 (Left) and scatter plot of flow in Badajoz vs. flow in Rocha da Galé for a percentile higher than 97th. The linear fit between both flows is also shown (blue solid line).

To define the relationship between the flows at both locations, the series were compared only considering high values of discharges to better analyse the flood conditions registered during the event of 1876. Specifically, the comparing analysis was restricted to flows registered in Badajoz higher than the 97th percentile. Fig. 8 (right panel) shows the scatter plot of the discharge flow in Badajoz versus the discharge flow in Rocha da Galé for a percentile greater than 97th and the linear fit between both flows.

Eq. (5) shows the linear fit between both variables.

$$Q_B = 0.6781 \cdot Q_R +\cdot 16.1 \tag{5}$$

where Q_B and Q_R refers to the discharge flow at Badajoz and Rocha da Galé, respectively.

Once the relationship between the two discharge flows was obtained, several numerical simulations with *Iber+* were carried out in the area of Mértola to define the value of discharge flow to raise the water level to approximately 30 m.a.s.l.

In the numerical simulations carried out with *Iber+*, different values of inlet flow ranging from 11,000 to 22,000 m^3s^{-1} were defined, allowing to obtain different water elevation levels for each value. The relationship between the elevation of the water level in Mértola and the inlet

15

flow defined in the *Iber+* model is practically linear, and therefore we applied a simple linear regression fit equation (Eq. 6).

$$D_M = 0.4648 \cdot Q_M + \cdot 21.57 \tag{6}$$

where D_M and Q_M refers to the water elevation and the discharge flow at Mértola, respectively. The Pearson correlation coefficient of the linear fit (r^2) is equal to 0.992.

Using Eq. (6), it can be estimated that the mean flow necessary to raise the water level to 30 ± 0.5 m.a.s.l. in Mértola is equal to $18,115 \pm 1,076$ m^{3}s^{-1}. The value of the mean flow in Mértola ($18,115$ m^{3}s^{-1}) corresponds to a flow in Badajoz equal to $12,300$ m^{3}s^{-1} using Eq. (5). This value is fully consistent with the value obtained by applying the rain-runoff methodology (Fig. 7).

4.3.   Numerical analysis of the flood event of 1876 in Badajoz

Once the time series of flow reaching Badajoz has been calculated, it can be used as an inlet condition in the numerical domain defined in *Iber+*. First, we perform a validation of the numerical results obtained through comparing the water levels attained in the simulations with the historical benchmarks available in the city of Badajoz and the references that can be consulted in the literature.

The maximum extension of the flooded area obtained with *Iber+* is shown in Fig. 9, together with an estimate of the flooded area obtained from the benchmarks and considering the topography of the area extracted from the Badajoz plan of the year 1873. Although both results show a similar extension, several discrepancies are clearly visible. These relatively confined discrepancies can be attributed to the following two facts: (i) the control points used for the georeferencing process of the 1873 map were located around the bastioned enclosure (on the left bank of the Guadiana River) and at the head of the river bridge and fort of San Cristóbal (on the right bank) being impossible to locate more exterior points. This causes a progressive increase in map deformations as we move away from the urban area (areas outlined in yellow in Fig. 10); and (ii) the variation in the surroundings of Rivillas stream is due to the intense anthropogenic occupation throughout the end of the 20th century, considerably modifying the topography of the stream and its banks.
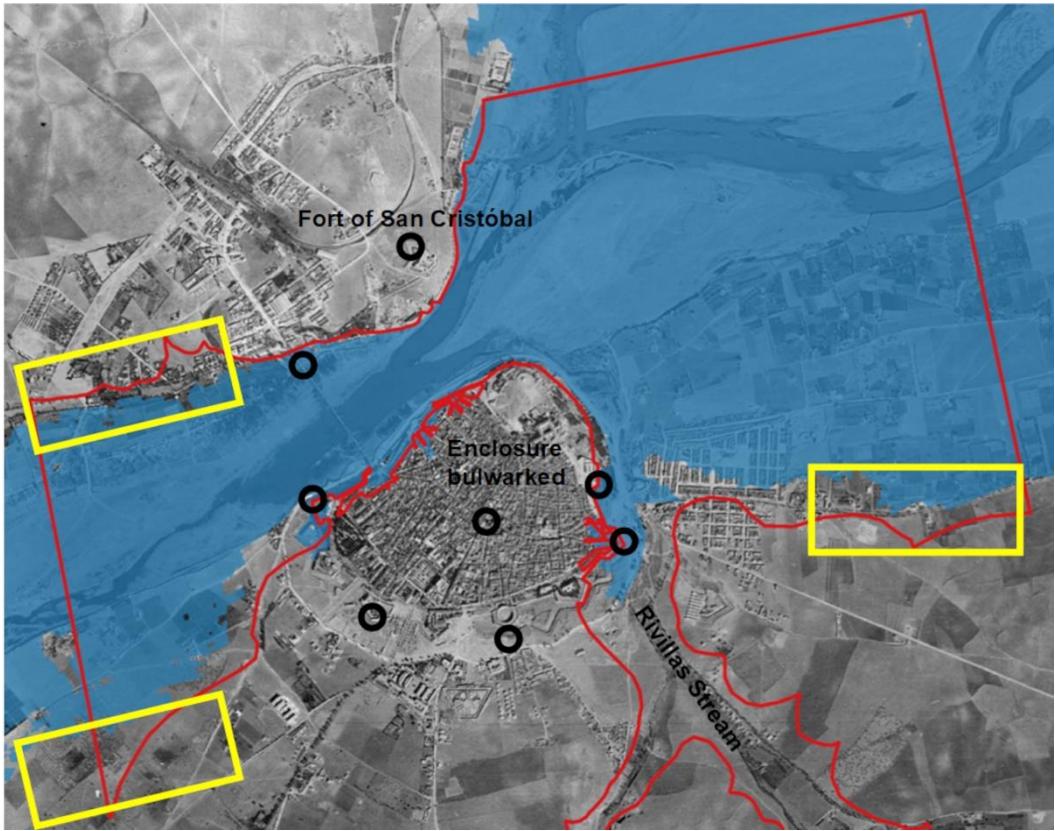
Figure 9. Maximum extension of the flooded area obtained with *Iber+* (blue shaded area) together with an estimate obtained from the benchmarks and considering the topography of the terrain preceding the Badajoz plan of the year 1873 (delimited by the red line), superimposed on the image of the American flight of 1947. The greatest differences obtained between the two procedures are delimited by the yellow boxes.

In Zamora-Cabanillas (1999), a reference to the 1876 event is mentioned, stating that water rose between 9 and 10 m in the vicinity of the Palmas Bridge. Fig. 10 shows the maximum (numerical) water depth of the water surface near the bridge obtained with the average flow. Results were obtained with (red line) and without the bridge (blue line). The maximum water level reached was 11.2 m when the bridge is considered and diminishing considerably to 8.6 m without the bridge. These results are compatible with the available records, if one considers that the bridge broke before the flow reached the potential maximum level of 11 m. It even realistically reproduces the sequence of events reported in the chronicle of the bridge breakage from the magazine "El Magisterio Extremeño": *"... we have spent fifteen days of copious and incessant rain ... in Guadiana, Rivillas and Gévora, you can see a considerable flood: its waters are penetrating through the streets ... shortly after two ... a sudden and considerable drop in the water level ... the bridge has broken !!! ... the sinking of many houses they increase panic... "*. In Zamora-Cabanillas (1999) the situation is described in a similar way: *"at three in*

the morning the flood suddenly subsided due to the collapse of the bridge…''. Indeed, once the bridge disappeared the water level descended as shown in numerical simulations. Since higher water depths were obtained when considering the bridge, the results shown in the following paragraphs, correspond to that case.
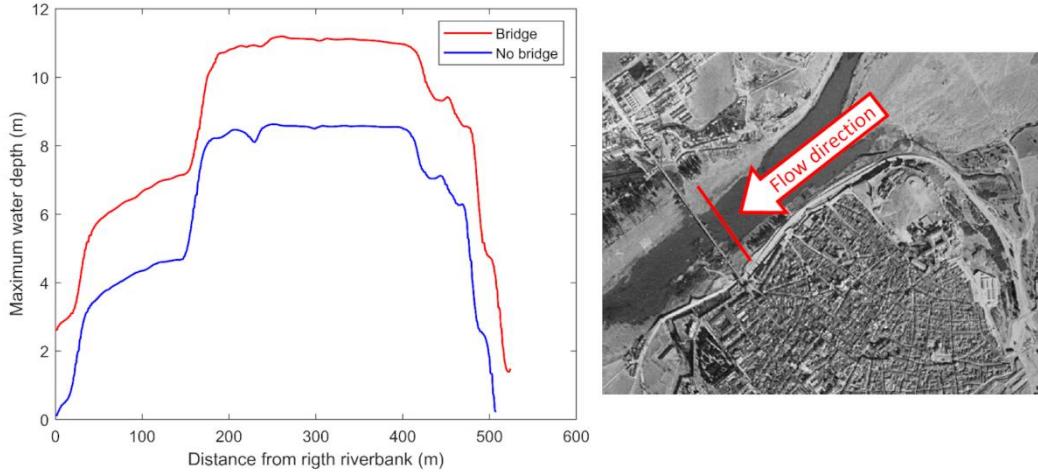

Figure 10. Maximum water depth obtained near the Palmas Bridge considering the bridge (red) and without the bridge (blue) (Left). Location of the cross section defined to determine the maximum water elevation (Right: red line).

There is a well-known historical benchmark placed in the Palmas Gate, near the Palmas Bridge, in the left riverbank (Fig., 11a), where a water depth around 1.33 meters can be estimated (Fig. 11b). The numerical simulation results obtained near that control benchmark imply a depth close to 1.5 m, a value that is in good agreement with field data (Fig. 11c).
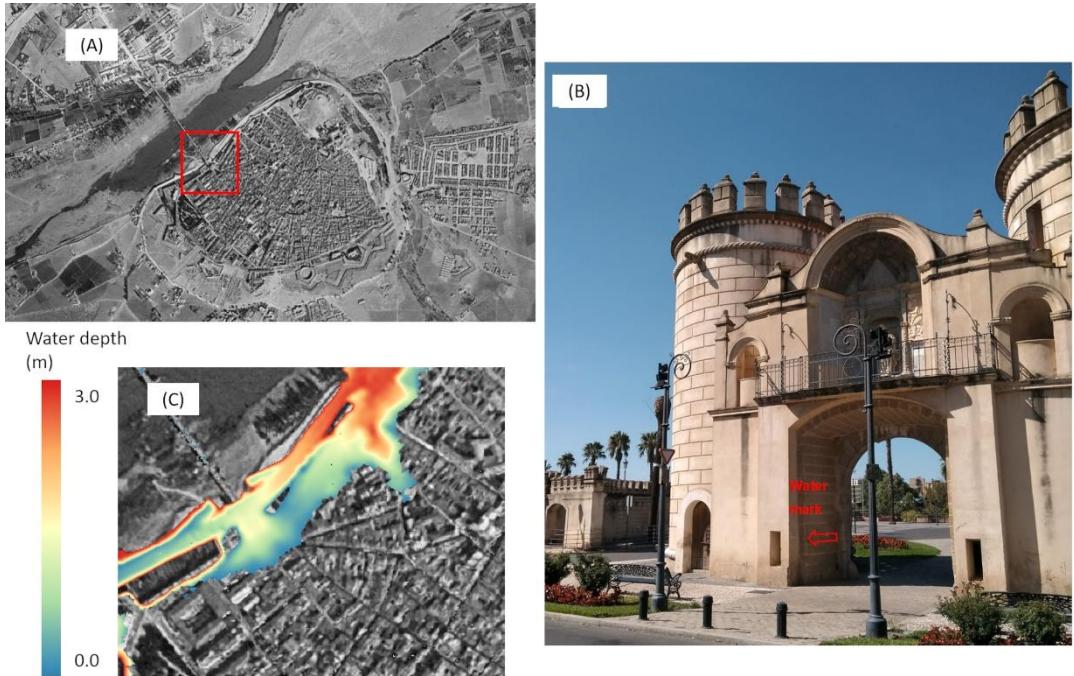
Figure 11. Maximum water depth obtained in the Palmas Gate. Panel (A): location of the control zone (area inside the red box); Panel (B): water mark in the Palmas Gate; Panel (C): numerical results of maximum water depth.

There are also some estimations on the extent of the water surface in the Trinidad Gate area on the other side of the city walls (Fig. 12a, 12b and 12c). Similarly, we also provide the extension of the water surface obtained in the Trinidad gate using the average flow (Fig. 12b). It can be verified that the extension of the water sheet obtained in the numerical simulations is very close to field estimations. There are certain discrepancies between the numerical results and the field measurements that are likely due to the uncertainties derived from the input data to *Iber+* and to the low spatial resolution of the DTMs. Some of the streets near the Trinidad Gate are very narrow and difficult to reproduce using DTMs with spatial resolution lower than 1 m. Nevertheless, the spatial extent of the flooded area in this sector shows a reasonable agreement. Finally, an additional evaluation assessment was applied to the maximum water depth registered at the Trinidad Gate, specifically in the intersection of Eugenio Hermoso and Ronda del Pilar streets. The maximum water depth can be estimated as 3 m (Fig 12e). The maximum water level obtained with *Iber+* using the average flow in this area is near 2.8 m (Fig 12d) and therefore the numerical results approximate reasonably well the reference data.
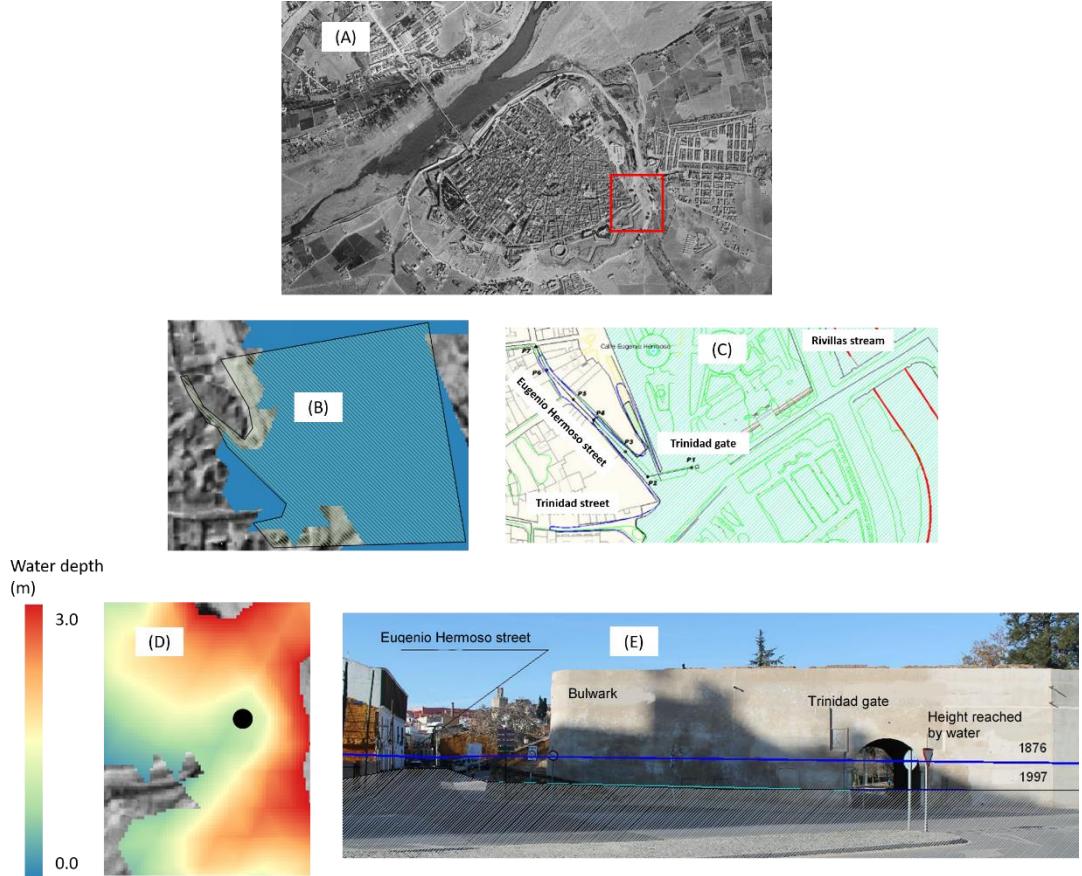
Figure 12 Results of the extension of the water surface and water depth at the Trinidad Gate area. Panel (A): location of the control zone (delimited by the red square); Panel (B): numerical results of maximum extension of the water surface (blue area) along with the maximum extension obtained from field measurements made by authors (delimited by the black line); Panel (C): extension of the water sheet from field measurements made by authors (shaded green area); Panel (D): numerical results of maximum extension of the water depth; Panel (E): water depth estimated from field data (obtained by authors)] for the events of 1876 and 1997. Black dot in panel (D) corresponds to the location of the building named "Baluarte" (Bulwark) shown in panel (E)

Once the precision of the numerical results has been verified, the flood in early December 1876 can be analysed in more detail. The evolution of the flood extent and magnitude from 4 to 12 December is shown in Fig. 13. Results from 4 to 7 December were obtained considering the bridge still stood whilst results from 8 to 12 December were obtained considering that the bridge had collapsed. The situation corresponding to the 4 December shows that some areas near Badajoz were flooded due to the continuous precipitation recorded during the previous days. As the flow increased, the fields upstream Badajoz started to flood (days 5 and 6). The situation became critical on December 7 and 8. Water depth reached the maximum values and the extension of the water-flooded surface stretched the entire domain on December 7. On December 9, the water level decreased although the situation continued to be serious. Water

levels kept declining until reaching a "controllable" situation already on December 12. This evolution of the flood is consistent with the detailed description shown by Zamora-Cabanillas (1999), where it is clearly stated that the flood started on December 7, in the early morning. In fact, it is revealed in several references (Zamora-Cabanillas, 1999, Ortega, 2007) that at 2:00 AM the state of the river did not bode well. The data mentioned here have been taken from eyewitnesses who narrated their experience in the magazine "El Magisterio Extremeño" (issue nº 35, corresponding to December 15, 1876). González Rodríguez (1995) reproduced the most relevant texts.
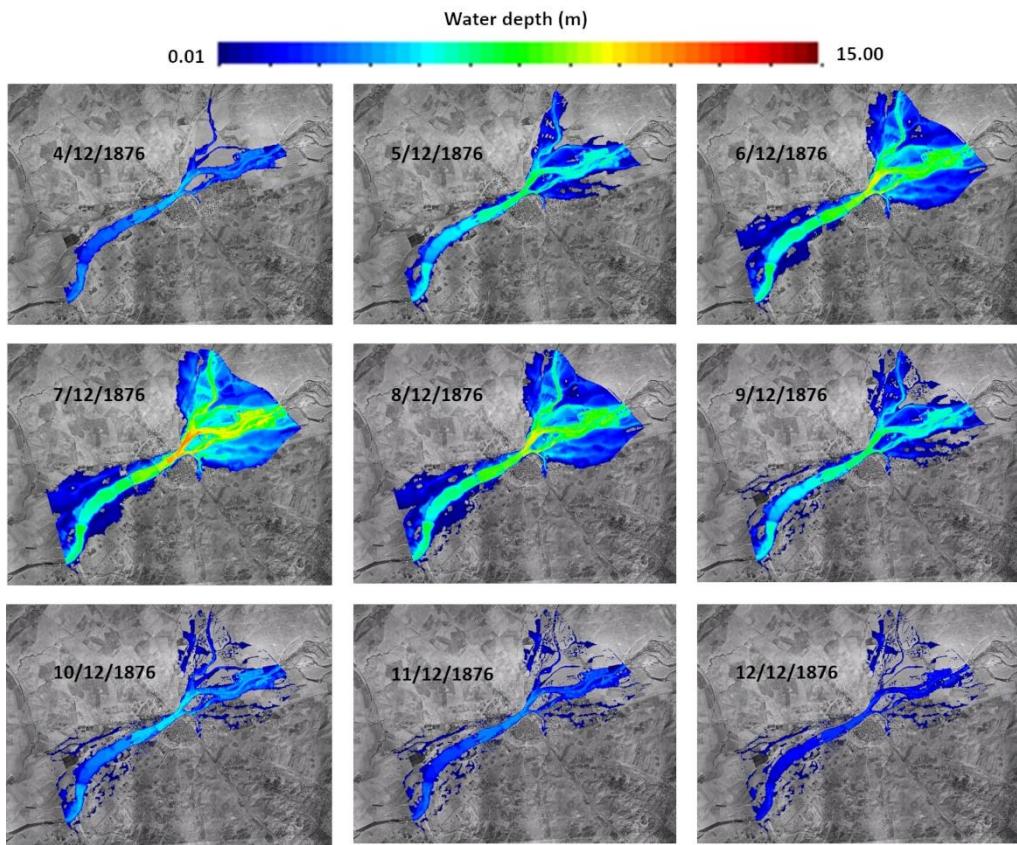

Figure 13. Evolution of the water depth from 4 to 12 December 1876.

The results of maximum water depths obtained for the entire numerical domain are shown in Fig. 14. In general, it can be seen that the fields upstream Badajoz were totally flooded and that the area downstream Badajoz was somewhat less affected, which is mainly due to the topography of the area.
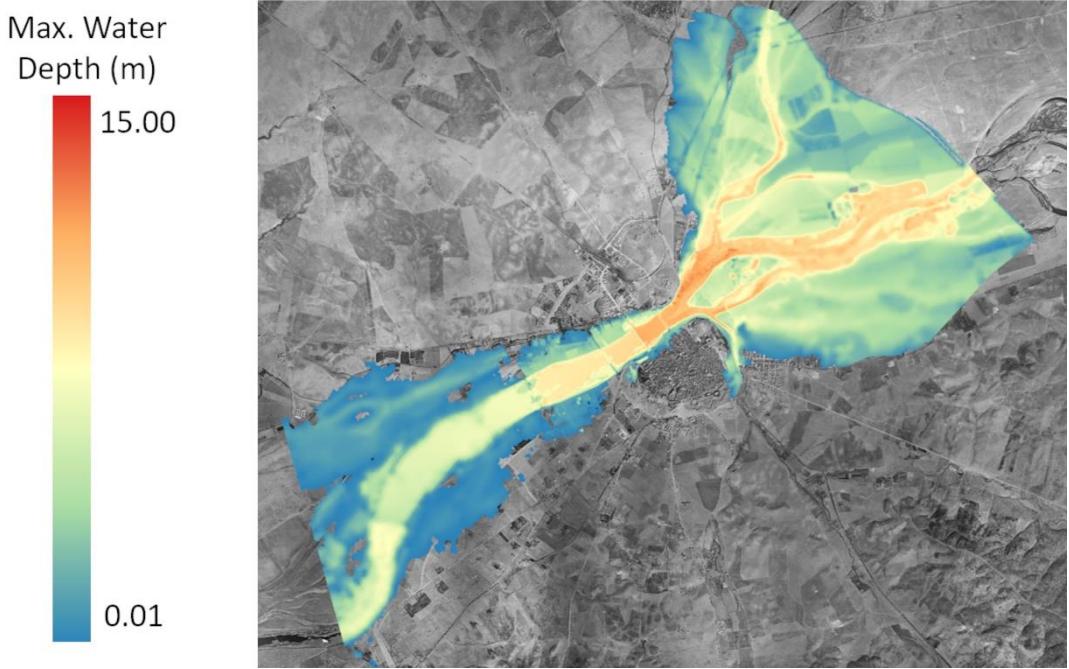
Figure 14. Maximum water depth obtained in the entire numerical domain.

More specifically, Fig. 15 shows the maps of maximum water depth and maximum velocity obtained near the Palmas Bridge. The maximum water depth near the bridge reached values close to 12 m as mentioned above and the maximum speed reached values close to 10 m s⁻¹. These values are extraordinarily high and are probably the origin of the bridge collapse. Unfortunately, there is no data on water velocity during the event to confirm this point.
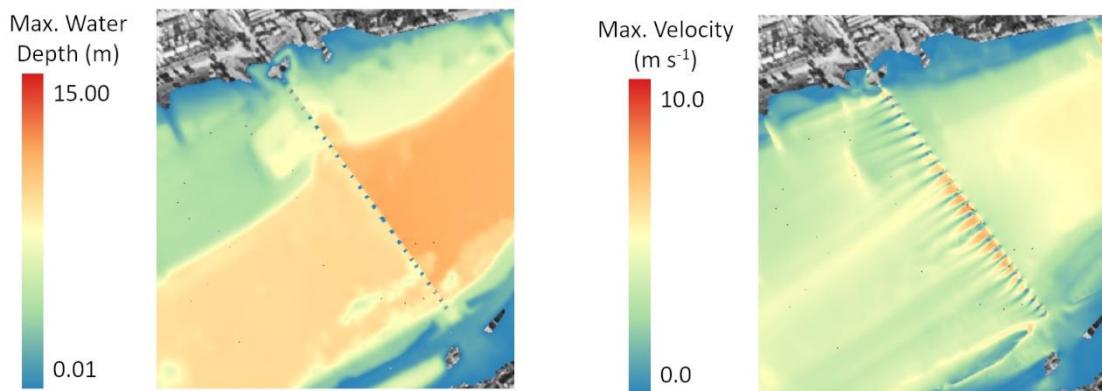

Figure 15. Maximum water depth (left panel) and maximum velocity (right panel) obtained in the numerical simulations near the Palmas Bridge.

5. Conclusions

In this work, the numerical reproduction of the flood event registered in Badajoz on 7 December 1876 was carried out by means of the *Iber+*. According to previous research, this is the greatest flood event in the Guadiana basin since 1603.

The task was divided into three stages. First, the determination of the spatial-temporal pattern of rainfall in the entire basin of the Guadiana River corresponding to the event was obtained from the daily rainfall records available at two meteorological stations (namely at the cities of Évora in Portugal and Badajoz in Spain) located in the extreme west of the basin. Fortunately, the vast majority of synoptic systems responsible for precipitation across the entire Guadiana basin in the winter months are also responsible for the precipitation events in Évora and Badajoz. This physical base connection is essential to guarantee the good representativeness that these two stations offer, as the validation of a simple model with recent data has confirmed. Thus, the quality of reconstruction of the daily precipitation field during this extreme event in 1876 was well assured and confirmed through a validation exercise for present climate. Then, once the pattern of rainfall was obtained, the Guadiana River basin was modeled by means of HEC-HMS code to estimate the river discharge in Badajoz. The peak discharge in Badajoz was also modelled by means of a series of numerical simulations carried out with *Iber+* in the area of Mértola (southern Portugal) and taking into account the relation between the time series of river flow obtained from two different gauge stations located in Badajoz and Mértola. The peak values obtained with both methodologies are quite similar, in the order of 12.500 m^3s^{-1}. Finally, the numerical reproduction of the flood event registered in Badajoz was performed using *Iber+*. To achieve this goal, we took into account a number of structural features, including the present and old topography of the area of study, the effects of land uses and structures (for example, the Palmas Bridge). The numerical simulation results obtained with *Iber+* were compared with reference data: benchmarks and data obtained from bibliography, showing a good agreement with reference data. Once the accuracy of the numerical results was assured, the reconstruction of the daily evolution of the flood and detailed analysis of the water depth and water velocities near the Palmas Bridge was carried out. The numerical simulations shown that the area near Badajoz was flooded during almost 4 days and the maximum water depths reached values near 10 m. Before the collapse of the Palmas bridge, maximum velocity in the vicinity of the bridge was about 10 m s^{-1}, a high value that undoubtedly influenced the collapse of some of the arches of the bridge.

The work has demonstrated the ability of *Iber+* to reproduce historical extreme flood events even when the availability of input data is relatively scarce. Finally, we are confident that the multidisciplinary methodology presented here can be adapted to characterize in detail other major historical floods with scarce meteorological and hydrological data available.

6. Acknowledgements

7. References

Agência Portuguesa do Ambiente, Plano De Gestão De Região Hidrográfica. Parte 2 – Caracterização E Diagnóstico. Região Hidrográfica Do Guadiana (RH7), 2016. Agência Portuguesa do Ambiente. Ministério do Ambiente e da Transição Energética, Lisboa, Portugal.

Akter, A., Ahmed, S., 2015. Potentiality of rainwater harvesting for an urban community in Bangladesh. Journal of Hydrology, 528, 84-93. DOI:10.1016/j.hydrol.2015.06.017.

Bai, Y., Zhang, Z., Zhao, W., 2019. Assessing the Impact of Climate Change on Flood Events Using HEC-HMS and CMIP5. Water, Air, & Soil Pollution, 230(6), 119. DOI:10.1007/s11270-019-4159-0.

Beniston, M., 2009. Trends in joint quantiles of temperature and precipitation in Europe since 1901 and projected for 2100. Geophysical Research Letters, 36, doi:10.1029/2008GL037119.

Benito, G., Machado, M. J. y Pérez González, A., 1996. Climate change and flood sensivity in Spain. En: Brarndson, J., Brown, A. G. y Gregory, K. L. (Eds). global continental Changes: the context of palaeohydrology. Geological Society Special Publication nº 115, 85-98. Londres.

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, 1–10.

Canuto, N., Ramos, T.B., Oliveira, A.R., Simionesei, L., Basso, M., Neves, R., 2019. Influence of reservoir management on Guadiana streamflow regime. Journal of Hydrology: Regional Studies, 25, 100628. DOI: 10.1016/j.ejrh.2019.100628.

C.H.G., 2003. Validación de Datos Hidrológicos para Estimación de Caudales de Avenidas para Diseños de Defensas en Encauzamientos, Junio.

Cortesi, N., Trigo, R.M., Gonzalez-Hidalgo, J.C., Ramos, A.M., 2013. Modelling monthly precipitation with circulation weather types for a dense network of stations over Iberia. Hydrology and Earth System Sciences, 17 (2), pp. 665-678. DOI: 10.5194/hess-17-665-2013

Crespo, A.J.C., Domínguez, J.M., Rogers, B.D., Gómez-Gesteira, M., Longshaw, S., Canelas, R., Vacondio, R., Barreiro, A., García-Feal, O., 2015. DualSPHysics: open-source parallel CFD solver on Smoothed Particle Hydrodynamics (SPH). Computer Physics Communications, 187: 204-216. DOI: 10.1016/j.cpc.2014.10.004.

Du, J., Qian, L., Rui, H., Zuo, T., Zheng, D., Xu, Y., Xu, C. Y., 2012. Assessing the effects of urbanization on annual runoff and flood events using an integrated hydrological modeling system for Qinhuai River basin, China. Journal of Hydrology, 464, 127-139. DOI:10.1016/j.hydrol.2012.06.057.

Elleder, L., Krejčí, J., Šírová, J., 2020(a). The 1714 flash flood in the Bohemian-Moravian Highlands – Reconstructing a Catastrophe. Quaternary International, 538, 14-28. DOI: 10.1016/j.quaint.2019.02.002.

Elleder, L., Krejčí, J., Racko, S., Daňhelka, J., Šírová, J., Kašpárek, L., 2020(b). Reliability check of flash-flood in Central Bohemia on May 25, 1872. Global and Planetary Change, 187, art. no. 103094, . DOI: 10.1016/j.gloplacha.2019.103094.

Eulenstein, J. and Kellerer-Pirklbauer, A., 2020. The central European flood of 1572 and its local scale effects as revealed by a damage inventory. Hydrological Sciences Journal. 65(6), 884-897, DOI: 10.1080/02626667.2020.1716980.

E.D. European Directive2007/60/C

Feldman, A.D., 2000. Hydrologic modeling system HEC-HMS: technical reference manual; US Army Corps of Engineers, Hydrologic Engineering Center.

Fernández-Nóvoa, D., García-Feal, O., González-Cao, J., de Gonzalo, C., Rodríguez-Suárez, J.A., del Portal, C.R., Gómez-Gesteira, M., 2020. MIDAS: A New Integrated Flood Early Warning System for the Miño River. Water (Switzerland), 12 (9), art. no. 3420, DOI: 10.3390/W12092319.

Font, I., 1988. Historia del clima en España. Cambios climáticos y sus causas. Instituto Nacional de Meteorología. 297p. Madrid.

Fragoso, M., Trigo, R. M., Zêzere, J. L., Valente, M. A., 2010. The exceptional rainfall event in Lisbon on 18 February 2008. Weather 65, 31–35. doi: 10.1002/wea.513.

Fraga, I., Cea, L., Puertas, J., 2020. MERLIN: a flood hazard forecasting system for coastal river reaches. Natural Hazards, 100, 1171–1193.7.

García-Feal, O., González-Cao, J.; Gómez-Gesteira, M., Cea, L., Domínguez, J.M., Formella, A., 2018. An accelerated tool for flood modelling based on Iber. Water, 10, 1459.

García-Feal, O., Cea, L., González-Cao, J., Domínguez, J.M., Gómez-Gesteira, M., 2020. IberWQ: A GPU Accelerated Tool for 2D Water Quality Modeling in Rivers and Estuaries. Water, 12, 413.

Gimeno, L., Drumond, A., Nieto, R., Trigo, R.M., Stohl, A., 2010. On the origin of continental precipitation. Geophysical Research Letters, 37 (13), art. no. L13804. DOI: 10.1029/2010GL043712.

Gimeno, L., Stohl, A., Trigo, R.M., Dominguez, F., Yoshimura, K., Yu, L., Drumond, A., Durn-Quesada, A.M., Nieto, R., 2012. Oceanic and terrestrial sources of continental

precipitation. Reviews of Geophysics, 50 (4), art. no. RG4003. DOI: 10.1029/2012RG000389.

González, A., 1995. Badajoz cara al Guadiana: Puerta de Palmas y el Puente Viejo (1460-1994). Badajoz: Caja Rural de Extremadura.

González-Cao, J., García-Feal, O., Domínguez, J.M., Crespo, A.J.C., Gómez-Gesteira, M., 2018. Analysis of the hydrological safety of dams combining two numerical tools: Iber and DualSPHysics. Journal of Hydrodynamics, 30 (1), 87-94. DOI: 10.1007/s42241-018-0009-6.

González-Cao, J., García-Feal, O., Fernández-Nóvoa, D., Domínguez-Alonso, J.M., Gómez-Gesteira, M., 2019a. Iber+: a new code to analyse dam-break floods. In book: Advances in Natural Hazards and Hydrological Risks: Meeting the Challenge (pp.165-169). DOI: 10.1007/978-3-030-34397-2_32.

González-Cao, J., García-Feal, O., Fernández-Nóvoa, D., Domínguez-Alonso, J.M., Gómez-Gesteira, M., 2019b. Towards an automatic early warning system of flood hazards based on precipitation forecast: the case of the Miño River (NW Spain). Natural Hazards & Earth System Sciences, 19.

González Rodríguez, A., 1995. La puerta de Palmas y el puente viejo (1460–1994): Badajoz cara al Guadiana. Badajoz: Caja Rural.

Groisman, P.Y., Knight, R.W., Easterling, D.R., Karl, T.R., Hegerl, G.C., Razuvaev, V.N., 2005. Trends in Intense Precipitation in the Climate Record. J. Climate, 18(9), 1326–1350, doi:10.1175/JCLI3339.1.

HEC-GeoHMS Geospatial Hydrologic Modeling Extension: User's Manual, Version 5 (2010) US Army Corps of Engineers, Hydraulic Engineering Center.

Herrera, S., Gutiérrez, J.M., Ancell, R., Pons, M.R., Frías, M.D., Fernández, J., 2012. Development and analysis of a 50-year high-resolution daily gridded precipitation dataset over Spain (Spain02). International Journal of Climatology, 32 (1), 74-85. DOI: 10.1002/joc.2256.

Herrera, S., Fernández, J., Gutiérrez, J.M., 2016 Update of the Spain02 gridded observational dataset for EURO-CORDEX evaluation: Assessing the effect of the interpolation

methodology. International Journal of Climatology, 36 (2). 900-908. DOI: 10.1002/joc.4391.

Kotlarski, S., Szabó, P., Herrera, S., Räty, O., Keuler, K., Soares, P.M., Cardoso, R.M., Bosshard, T., Pagé, C., Boberg, F., Gutiérrez, J.M., Isotta, F.A., Jaczewski, A., Kreienkamp, F., Liniger, M.A., Lussana, C., Pianko-Kluczyńska, K., 2019. Observational uncertainty and regional climate model evaluation: A pan-European perspective. International Journal of Climatology, 39 (9), pp. 3730-3749. DOI: 10.1002/joc.5249

Long, P.O., 2008. Hydraulic engineering and the study of antiquity: Rome, 1557-70. Renaissance Quarterly, 61 (4), pp. 1098-1138. DOI: 10.1353/ren.0.0320.

Loureiro, J. M. 2009. Rio Tejo: as grandes cheias: 1800–2007, 1st Edn. Tágides, Lisboa: ARH do Tejo I.P.

Morss, R.E., Wilhelmi, O.V., Meehl, G.A., Dilling, L., 2011. Improving societal outcomes of extreme weather in a changing climate: An integrated perspective. Annual Review of Environment and Resources, 36, 1-25. DOI: 10.1146/annurev-environ-060809-100145.

Ministerio de Fomento, 2016. Norma 5.2-IC Drenaje superficial de la Instrucción de Carreteras; Spain Government: Madrid.

NRCS (Natural Resources Conservation Service), 2007. National Engineering Handbook: Chapter 16 Hydrographs. Washington, DC.

NVIDIA Corporation CUDA C++ Programming Guide Available online: https://docs.nvidia.com/cuda/pdf/CUDA_C_Programming_Guide.pdf (accessed on Aug. 27, 2020).

Ortega, J., A., 2007. Paleocrecidas, Avenidas recientes e Hidroclimatología en la Cuenca Media y Baja del rio Guadiana. Departamento de Geodinámica, Facultad de Ciencias Geológicas, Universidad Complutense de Madrid, Spain (PhD thesis).

Pfister, C., Brázdil, R., Glaser, R., Barriendos, M., Camuffo, D., Deutsch, M., Dobrovolny, P., Enzi, S., Guidoboni, E., Kotyza, O., Militzer, S., Racz, L., Rodrigo, F.S., 1999. Documentary Evidence on Climate in the Sixteenth-Century Europe. Special Issue: Climatic Variability in Sixteenth-Century Europe and Its Social Dimensions, Climatic Change, 43:1, 55-110.

Portugués-Mollá, I., Bonache-Felici, X., Mateu-Bellés, J.F., Marco-Segura, J.B., 2016. A GIS-Based Model for the analysis of an urban flash flood and its hydro-geomorphic response. The Valencia event of 1957. Journal of Hydrology, 541, 582-596. DOI: 10.1016/j.jhydrol.2016.05.048.

Scharffenberg, B., Bartles, M., Brauer, T., Fleming, M., Karlovits, G., 2018. Hydrologic Modeling System (HEC-HMS). User's Manual:Version 4.3; U.S. Army Corps of Engineers.

SNIRH, 2019. Serviço Nacional De Informação Dos Recursos Hídricos. Agência Portuguesa do Ambiente, Lisboa, Portugal (2019). https://snirh.apambiente.pt/ (Last accessed 29/08/2020).

Trigo, R.M., DaCamara, C.C., 2000. Circulation weather types and their influence on the precipitation regime in Portugal. International Journal of Climatology, 20 (13), pp. 1559-1581. DOI: 10.1002/1097-0088(20001115)20:13<1559::AID-JOC555>3.0.CO;2-5

Trigo, R.M., Varino, F., Ramos, A.M., Valente, M.A., Zêzere, J.L., Vaquero, J.M., Gouveia, C.M., Russo, A., 2014. The record precipitation and flood event in Iberia in december 1876: Description and synoptic analysis Frontiers in Earth Science, 2 (April), art. no. 3, . DOI: 10.3389/feart.2014.00003.

Yin, J., Yu, D., Yin, Z., Liu, M., He, Q., 2016. Evaluating the impact and risk of pluvial flash flood on intra-urban road network: A case study in the city center of Shanghai, China. Journal of Hydrology, 537, 138-145. DOI: 10.1016/j.jhydrol.2016.03.037.

Zamora-Cabanillas, J.F., 1999. Consideraciones sobre la riada sucedida en la ciudad de Badajoz en Noviembre de 1997. Monografía: La Ciencia a las puertas del Tercer Milenio (I), Catedra Nova.

Conclusiones y trabajo futuro

6

> *Since all models are wrong the scientist must be alert to what is importantly wrong. It is inappropriate to be concerned about mice when there are tigers abroad.*

— **George E.P. Box**
Science and Statistics

Como resultado del trabajo realizado en esta tesis doctoral se ha desarrollado una nueva implementación del modelo hidráulico Iber que, mediante técnicas de computación de altas prestaciones, es capaz de hacer un uso eficiente de las capacidades de cálculo que ofrecen las CPU multinúcleo y especialmente las GPU actuales. Gracias a esto, Iber+ ha demostrado ser capaz de obtener mejoras en la velocidad de ejecución de las simulaciones de hasta dos órdenes de magnitud conservando la precisión de los resultados. Este hecho abre todo un abanico de nuevas aplicaciones que antes se encontraban limitadas por su excesivo tiempo de cálculo.

Aunque los casos de uso más comunes están cubiertos por la paralelización de Iber+, ciertas aplicaciones pueden requerir simular dominios todavía más grandes que utilicen mallas de decenas o cientos de millones de elementos. En estos casos, la memoria disponible en una GPU típica resulta un factor limitante al tamaño máximo de malla a utilizar, mientras que los sistemas CPU de memoria compartida pueden tener una gran cantidad de memoria disponible pero generalmente no disponen de la potencia de cálculo necesaria para realizar una simulación de esas dimensiones en un tiempo razonable. Para abordar estas limitaciones, puede utilizarse una paralelización para sistemas CPU de

memoria distribuida (LeVeque & George, 2008; Noh et al., 2018) utilizando generalmente MPI, o soluciones multi-GPU (de la Asunción et al., 2012; Viñas et al., 2013).

Además del módulo principal que implementa la hidrodinámica, también se ha realizado una paralelización completa para los módulos de turbulencia, hidrología y calidad de aguas, tanto para sistemas CPU de memoria compartida como para sistemas GPU. Siendo objeto de futuros desarrollos la paralelización de otros módulos presentes en la implementación original como el módulo de transporte de sedimentos.

Merece especial mención el desarrollo realizado para el módulo de calidad de aguas, siendo una de las pocas implementaciones en GPGPU de un modelo 2D de aguas someras totalmente acoplado con un modelo de calidad de aguas disponibles en la literatura. La mejora en la eficiencia resulta de especial interés en las aplicaciones típicas de este tipo de modelos debido al alto número de simulaciones (generalmente en dominios espacio-temporales grandes) necesarias para parametrizar el modelo.

Estos nuevos desarrollos se han mostrado efectivos a la hora de estudiar fenómenos extremos de inundación, siendo también aplicados con éxito en la reproducción de eventos históricos de inundación. Debido a limitada disponibilidad de datos en muchos de estos casos, son necesarias un elevado número de simulaciones para la calibración de parámetros iniciales. La disponibilidad de un modelo hidráulico eficiente permite realizar estas simulaciones en un tiempo razonable sin comprometer la precisión de los resultados.

Por otro lado, el desarrollo de Iber+ ha sido una pieza fundamental en el desarrollo del sistema de alerta temprana de inundaciones para el río Miño MIDAS. El sistema desarrollado es capaz predecir de forma precisa eventos de inundación en zonas urbanas a partir de datos de precipitación (modelos de predicción y/o pluviómetros) y caudal antecedente de estaciones de control.

Proporcionando además mapas de riesgo detallados a las correspondientes autoridades en un margen de tiempo reducido gracias a las optimizaciones introducidas en Iber+.

Aunque el sistema fue diseñado para el río Miño, éste puede ser adaptado de forma sencilla a otras regiones. Todos sus componentes están disponibles de forma gratuita y puede ser ajustado fácilmente a los datos disponibles. Además, la metodología presentada para la modelización hidrológica, permite simular de forma efectiva eventos de duración arbitraria con un número reducido de parámetros, lo cual puede ser especialmente conveniente en zonas poco estudiadas.

Existen diversas líneas de desarrollo del sistema de alerta temprana que pueden ser abordadas en trabajos futuros como puede ser el estudio de estrategias de actuación coordinadas por parte de los embalses ante previsión de un evento extremo, aplicación de *machine learning* en la predicción de caudales de salida de los embalses o la simulación de parámetros de calidad de aguas en base a datos del sistema automático de información de calidad de las aguas (SAICA) de la Confederación Hidrográfica del Miño-Sil.

Bibliografía

Abu-Bakar, A., Ahmadian, R. & Falconer, R. A. (2017). Modelling the transport and decay processes of microbial tracers in a macro-tidal estuary. *Water Research*. https://doi.org/10.1016/j.watres.2017.07.007 (vid. pág. 14)

Alfieri, L., Burek, P., Dutra, E., Krzeminski, B., Muraro, D., Thielen, J. & Pappenberger, F. (2013). GloFAS – global ensemble streamflow forecasting and flood early warning. *Hydrology and Earth System Sciences*, *17*(3), 1161-1175. https://doi.org/10.5194/hess-17-1161-2013 (vid. pág. 15)

Alfieri, L., Salamon, P., Pappenberger, F., Wetterhall, F. & Thielen, J. (2012). Operational early warning systems for water-related hazards in Europe [Publisher: Elsevier]. *Environmental Science & Policy*, *21*, 35-49 (vid. pág. 15).

Amdahl, G. M. (1967). Validity of the single processor approach to achieving large scale computing capabilities. *Proceedings of the April 18-20, 1967, spring joint computer conference*, 483-485 (vid. pág. 8).

Amirlatifi, A. (2013). Coupled geomechanical reservoir simulation. *Doctoral Dissertations*. https://scholarsmine.mst.edu/doctoral_dissertations/2146 (vid. pág. 4)

Arnell, N. W. & Gosling, S. N. (2016). The impacts of climate change on river flood risk at the global scale. *Climatic Change*, *134*(3), 387-401. https://doi.org/10.1007/s10584-014-1084-5 (vid. pág. 2)

Belikov, E., Deligiannis, P., Totoo, P., Aljabri, M. & Loidl, H.-W. (2013). *A survey of high-level parallel programming models* (inf. téc. HW-MACS-TR-0103). Heriot-Watt University. Edinburgh, UK. http://www.macs.hw.ac.uk/cs/techreps/docs/files/HW-MACS-TR-0103.pdf. (Vid. pág. 7)

Bennett, T. H. (1998). *Development and application of a continuous soil moisture accounting algorithm for the Hydrologic Engineering Center Hydrologic Modeling System (HEC-HMS)* (Tesis doctoral) [OCLC: 42712111]. University of California. Davis, California. (Vid. pág. 24).

Berghuijs, W. R., Aalbers, E. E., Larsen, J. R., Trancoso, R. & Woods, R. A. (2017). Recent changes in extreme floods across multiple continents [Publisher: IOP Publishing]. *Environmental Research Letters*, *12*(11), 114035. https://doi.org/10.1088/1748-9326/aa8847 (vid. pág. 1)

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 [Publisher: Elsevier]. *Revista Internacional de Métodos Numéricos para Cálculo y Diseño en Ingeniería*, *30*(1), 1-10 (vid. págs. 4, 21).

Bladé Castellet, E. (2005). *Modelación del flujo en lámina libre sobre cauces naturales. Análisis integrado con esquemas en volúmenes finitos en una y dos dimensiones* [Accepted: 2011-04-12T15:13:57Z Publication Title: TDX (Tesis Doctorals en Xarxa)]. Universitat Politècnica de Catalunya. Consultado 4 de diciembre de 2020, desde https://upcommons.upc.edu/handle/2117/93706. (Vid. págs. 4, 21)

Bladé Castellet, E., Cea, L. & Corestein, G. (2014). Numerical modelling of river inundations. *Ingeniería del agua*, *18*(1), 68. https://doi.org/10.4995/ia.2014.3144 (vid. pág. 5)

Bladé i Castellet, E., Sanz Ramos, M., Amengual, A., Romero, R., Roux, H., Savatier, J. & Cherriere, M. (2018). Gestión integrada del riesgo de inundación y de los recursos hídricos empleando modelización integrada meteorológica, hidrológica e hidráulica [Accepted: 2018-07-17T11:05:47Z], 15-23. Consultado 1 de diciembre de 2020, desde https://upcommons.upc.edu/handle/2117/119423 (vid. pág. 1)

Blake, G., Dreslinski, R. G. & Mudge, T. (2009). A survey of multicore processors [Conference Name: IEEE Signal Processing Magazine]. *IEEE Signal Processing Magazine*, *26*(6), 26-37. https://doi.org/10.1109/MSP.2009.934110 (vid. pág. 9)

Bodoque, J. M., Amérigo, M., Díez-Herrero, A., García, J. A., Cortés, B., Ballesteros-Cánovas, J. A. & Olcina, J. (2016). Improvement of resilience of urban areas by integrating social perception in flash-flood risk management. *Journal of Hydrology*, *541*, 665-676. https://doi.org/10.1016/j.jhydrol.2016.02.005 (vid. pág. 5)

Booth, D. B. & Bledsoe, B. P. (2009). Streams and Urbanization. En L. A. Baker (Ed.), *The Water Environment of Cities* (pp. 93-123). Springer US. https://doi.org/10. 1007/978-0-387-84891-4_6. (Vid. pág. 2)

Bradshaw, C. J. A., Sodhi, N. S., Peh, K. S.-H. & Brook, B. W. (2007). Global evidence that deforestation amplifies flood risk and severity in the developing world [_eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1111/j.1365-2486.2007.01446.x]. *Global Change Biology*, *13*(11), 2379-2395. https://doi. org/https://doi.org/10.1111/j.1365-2486.2007.01446.x (vid. pág. 2)

Brirhet, H. & Benaabidate, L. (2016). Comparison of two hydrological models (lumped and distributed) over a pilot area of the Issen watershed in the Souss basin, Morocco [Publisher: European Scientific Journal]. *European Scientific Journal*, *12*(18) (vid. pág. 14).

Cea, L., Bermúdez, M., Puertas, J., Bladé, E., Corestein, G., Escolano, E., Conde, A., Bockelmann-Evans, B. & Ahmadian, R. (2016). IberWQ: New simulation tool for 2D water quality modelling in rivers and shallow estuaries. *Journal of Hydroinformatics*. https://doi.org/10.2166/hydro.2016.235 (vid. págs. 21, 24, 33)

Cea, L. & Bladé, E. (2015). A simple and efficient unstructured finite volume scheme for solving the shallow water equations in overland flow applications. *Water Resources Research*, *51*(7), 5464-5486. https://doi.org/10.1002/2014WR016547 (vid. págs. 5, 21)

Cea, L. & French, J. R. (2012). Bathymetric error estimation for the calibration and validation of estuarine hydrodynamic models. *Estuarine, Coastal and Shelf Science*, *100*, 124-132. https://doi.org/10.1016/j.ecss.2012.01.004 (vid. pág. 5)

Cea, L. & Vázquez-Cendón, M. E. (2012). Unstructured finite volume discretisation of bed friction and convective flux in solute transport models linked to the shallow water equations. *Journal of Computational Physics*. https://doi.org/10.1016/j.jcp. 2012.01.007 (vid. pág. 24)

Cea, L. (2005). *An unstructured finite volume model for unsteady turbulent shallow water flow with wet-dry fronts: numerical solver and experimental validation* [Accepted:

2009-10-27T12:21:24Z]. Consultado 4 de diciembre de 2020, desde https://ruc.udc.es/dspace/handle/2183/805. (Vid. págs. 4, 21)

Cea, L., Bladé, E., Sanz-Ramos, M., Fraga, I., Sañudo, E., García-Feal, O., Gómez-Gesteira, M. & González-Cao, J. (2020). Benchmarking of the Iber capabilities for 2D free surface flow modelling [Accepted: 2020-01-09T13:04:30Z Publisher: Universidade da Coruña, Servizo de Publicacións]. Consultado 7 de diciembre de 2020, desde https://ruc.udc.es/dspace/handle/2183/24588 (vid. págs. 5, 32)

Cea, L., Legout, C., Grangeon, T. & Nord, G. (2016). Impact of model simplifications on soil erosion predictions: Application of the GLUE methodology to a distributed event-based model at the hillslope scale. *Hydrological Processes*. https://doi.org/10.1002/hyp.10697 (vid. pág. 5)

Cea, L., Puertas, J. & Vázquez-Cendón, M. E. (2007). Depth averaged modelling of turbulent shallow water flow with wet-dry fronts. *Archives of Computational Methods in Engineering*. https://doi.org/10.1007/s11831-007-9009-3 (vid. págs. 21, 23)

Chapra, S. C., Pelletier, G. J. & Tao, H. (2008). *QUAL2K: A Modeling Framework for Simulating River and Stream Water Quality, Version 2.11: Documentation and Users Manual*. Civil; Environmental Engineering Dept., Tufts University. (Vid. pág. 14).

Chow, V. T. (1964). *Handbook of applied hydrology: a compendium of water-resources technology*. McGraw-Hill. (Vid. pág. 25).

Clark, C. O. (1945). Storage and the unit hydrograph [Issue: 9]. *Proceedings of the American Society of Civil Engineers, 69*, 1333-1360 (vid. pág. 25).

Cole, T. M. & Buchak, E. M. (1995). *CE-QUAL-W2: A Two-Dimensional, Laterally Averaged, Hydrodynamic and Water Quality Model, Version 2.0. User Manual* (inf. téc.). (Vid. pág. 14).

Cole, T. M. & Wells, S. A. (2006). CE-QUAL-W2: A two-dimensional, laterally averaged, hydrodynamic and water quality model, version 3.5 (vid. pág. 14).

Corestein, G. & Bladé, E. (2013). Validación del módulo de transporte de sedimendos de fondo-Modelo Iber, 27-34 (vid. pág. 21).

Corral, C., Berenguer, M., Sempere-Torres, D., Poletti, L., Silvestro, F. & Rebora, N. (2019). Comparison of two early warning systems for regional flash flood hazard forecasting [Publisher: Elsevier]. *Journal of hydrology*, *572*, 603-619 (vid. pág. 15).

Cox, R., Shand, T. & Blacka, M. (2010). *Australian Rainfall and Runoff revision project 10: appropriate safety criteria for peopl*. Water Research Laboratory, The University of New South Wales. (Vid. pág. 34).

Crespo, A. J. C., Domínguez, J. M., Rogers, B. D., Gómez-Gesteira, M., Longshaw, S., Canelas, R., Vacondio, R., Barreiro, A. & García-Feal, O. (2015). DualSPHysics: Open-source parallel CFD solver based on Smoothed Particle Hydrodynamics (SPH). *Computer Physics Communications*, *187*, 204-216. https://doi.org/10.1016/j.cpc.2014.10.004 (vid. pág. 11)

Crossley, A., Lamb, R. & Waller, S. (2010). Fast solution of the Shallow Water Equations using GPU technology, 13-19 (vid. págs. 6, 11, 32).

Cutress, I. & Shilov, A. (2019). The Larrabee Chapter Closes: Intel's Final Xeon Phi Processors Now in EOL. Consultado 3 de enero de 2021, desde https://www.anandtech.com/show/14305/intel-xeon-phi-knights-mill-now-eol. (Vid. pág. 26)

Czarnul, P., Proficz, J. & Drypczewski, K. (2020). Survey of Methodologies, Approaches, and Challenges in Parallel Programming Using High-Performance Computing Systems [ISSN: 1058-9244 Pages: e4176794 Publisher: Hindawi Volume: 2020]. https://doi.org/https://doi.org/10.1155/2020/4176794. (Vid. pág. 7)

Dagum, L. & Menon, R. (1998). OpenMP: an industry standard API for shared-memory programming [Conference Name: IEEE Computational Science and Engineering]. *IEEE Computational Science and Engineering*, *5*(1), 46-55. https://doi.org/10.1109/99.660313 (vid. págs. 7, 28)

de la Asunción, M., Mantas, J. M., Castro, M. J. & Fernández-Nieto, E. D. (2012). An MPI-CUDA implementation of an improved Roe method for two-layer shallow water systems. *Journal of Parallel and Distributed Computing*, *72*(9), 1065-1072. https://doi.org/10.1016/j.jpdc.2011.07.012 (vid. pág. 160)

De St Venant, B. (1871). Theorie du mouvement non-permanent des eaux avec application aux crues des rivers et a l'introduntion des Marees dans leur lit. *Academic de Sci. Comptes Redus*, *73*(99), 148-154 (vid. pág. 22).

Eeckhout, L. (2017). Is Moore's Law Slowing Down? What's Next? [Publisher: IEEE Computer Society]. https://doi.org/10.1109/MM.2017.3211123. (Vid. pág. 9)

The European High Performance Computing Joint Undertaking - EuroHPC. (2018). Consultado 2 de enero de 2021, desde https://ec.europa.eu/digital-single-market/en/eurohpc-joint-undertaking. (Vid. pág. 26)

Eymard, R., Gallouët, T. & Herbin, R. (2000). Finite volume methods. *Handbook of Numerical Analysis* (pp. 713-1018). Elsevier. https://doi.org/10.1016/S1570-8659(00)07005-8. (Vid. pág. 2)

Feldman, A. D. (2000). *Hydrologic modeling system HEC-HMS: technical reference manual* (inf. téc.). US Army Corps of Engineers, Hydrologic Engineering Center. (Vid. págs. 24, 25).

Feldman, M. (2017). Intel Dumps Knights Hill, Future of Xeon Phi Product Line Uncertain | TOP500. Consultado 3 de enero de 2021, desde https://www.top500.org/news/intel-dumps-knights-hill-future-of-xeon-phi-product-line-uncertain/. (Vid. pág. 26)

Fernández-Nóvoa, D., García-Feal, O., González-Cao, J., de Gonzalo, C., Rodríguez-Suárez, J. A., Ruiz del Portal, C. & Gómez-Gesteira, M. (2020). MIDAS: A New Integrated Flood Early Warning System for the Miño River [Number: 9 Publisher: Multidisciplinary Digital Publishing Institute]. *Water*, *12*(9), 2319. https://doi.org/10.3390/w12092319 (vid. págs. 18, 35, 106)

Fraga, I., Cea, L., Puertas, J., Suárez, J., Jiménez, V. & Jácome, A. (2016). Global Sensitivity and GLUE-Based Uncertainty Analysis of a 2D-1D Dual Urban Drainage Model. *Journal of Hydrologic Engineering*, *21*(5), 04016004. https://doi.org/10.1061/(ASCE)HE.1943-5584.0001335 (vid. pág. 5)

Fragoso, M., Trigo, R. M., Zêzere, J. L. & Valente, M. A. (2010). The exceptional rainfall event in Lisbon on 18 February 2008 [Publisher: John Wiley & Sons, Ltd. Chichester, UK]. *Weather*, *65*(2), 31-35 (vid. pág. 36).

Gao, G., Falconer, R. A. & Lin, B. (2011). Numerical modelling of sediment–bacteria interaction processes in surface waters. *Water Research*, *45*(5), 1951-1960. https://doi.org/10.1016/j.watres.2010.12.030 (vid. pág. 14)

García-Feal, O., Cea, L., González-Cao, J., Domínguez, J. M. & Gómez-Gesteira, M. (2020). IberWQ: A GPU Accelerated Tool for 2D Water Quality Modeling in Rivers and Estuaries [Publisher: Multidisciplinary Digital Publishing Institute]. *Water, 12*(2), 413 (vid. págs. 18, 77).

García-Feal, O., Gonzalez-Cao, J., Cea, L., Formella, A. & Domínguez, J. M. (2017). Iber as a tool to analyse flooding scenarios. *Proceedings of First International Electronic Conference on the Hydrological Cycle*, 4841. https://doi.org/10.3390/CHyCle-2017-04841 (vid. págs. 5, 31)

García-Feal, O., González-Cao, J., Gómez-Gesteira, M., Cea, L., Domínguez, J. M. & Formella, A. (2018). An accelerated tool for flood modelling based on Iber [Publisher: Multidisciplinary Digital Publishing Institute]. *Water, 10*(10), 1459 (vid. págs. 17, 39).

Goldman, D. & Ely, P. (1998). *HEC-1 Flood Hydrograph Package User's Manual* (Computer Program Document CPD-1A). US ARMY CORPS OF ENGINEERSHY-DROLOGIC ENGINEERING CENTER (HEC). Davis, California. Consultado 2 de enero de 2021, desde https://www.hec.usace.army.mil/publications/ComputerProgramDocumentation/HEC-1_UsersManual_(CPD-1a).pdf. (Vid. pág. 24)

González-Cao, J., García-Feal, O., Fernández-Nóvoa, D., Domínguez-Alonso, J. M. & Gómez-Gesteira, M. (2019). Towards an automatic early warning system of flood hazards based on precipitation forecast: the case of the Miño River (NW Spain). *Natural Hazards & Earth System Sciences, 19*(11) (vid. págs. 18, 35, 63).

Green, W. H. & Ampt, G. A. (1911). Studies on Soil Phyics. [Publisher: Cambridge University Press]. *The Journal of Agricultural Science, 4*(1), 1-24. https://doi.org/10.1017/S0021859600001441 (vid. pág. 24)

Gropp, W., Gropp, W. D., Lusk, E., Skjellum, A. & Lusk, A. D. F. E. E. (1999). *Using MPI: Portable Parallel Programming with the Message-passing Interface*. MIT Press. (Vid. pág. 7).

Hallegatte, S. (2012). *A cost effective solution to reduce disaster losses in developing countries: hydro-meteorological services, early warning, and evacuation*. The World Bank. (Vid. pág. 1).

Harris, D. (2020). Green Light! Top500 Speeds Up, Saves Energy with NVIDIA. Consultado 16 de diciembre de 2020, desde https://blogs.nvidia.com/blog/2020/06/22/top500-isc-supercomputing/. (Vid. págs. 11, 29)

Henderson, F. M. (1966). *Open channel flow*. MacMillan Publishin Co. Inc. (Vid. pág. 25).

Herrera, S., Gutiérrez, J. M., Ancell, R., Pons, M. R., Frías, M. D. & Fernández, J. (2012). Development and analysis of a 50-year high-resolution daily gridded precipitation dataset over Spain (Spain02) [Publisher: Wiley Online Library]. *International Journal of Climatology*, *32*(1), 74-85 (vid. pág. 37).

Hoshino, T., Maruyama, N., Matsuoka, S. & Takaki, R. (2013). CUDA vs OpenACC: Performance Case Studies with Kernel Benchmarks and a Memory-Bound CFD Application. *2013 13th IEEE/ACM International Symposium on Cluster, Cloud, and Grid Computing*, 136-143. https://doi.org/10.1109/CCGrid.2013.12 (vid. pág. 12)

Hunter, N. M., Bates, P. D., Neelz, S., Pender, G., Villanueva, I., Wright, N. G., Liang, D., Falconer, R. A., Lin, B., Waller, S., Crossley, A. J. & Mason, D. C. (2008). Benchmarking 2D hydraulic models for urban flooding [Publisher: ICE Publishing]. *Proceedings of the Institution of Civil Engineers - Water Management*, *161*(1), 13-30. https://doi.org/10.1680/wama.2008.161.1.13 (vid. pág. 4)

InfoWorks ICM Product Information - Overview. (2018). Consultado 10 de septiembre de 2018, desde http://www.innovyze.com/products/infoworks_icm/. (Vid. págs. 6, 32)

The International Roadmap for Devices and Systems (IRDS 2020). (2020). Consultado 9 de diciembre de 2020, desde https://irds.ieee.org/editions/2020. (Vid. pág. 9)

Jonkman, S. N. (2005). Global Perspectives on Loss of Human Life Caused by Floods. *Natural Hazards*, *34*(2), 151-175. https://doi.org/10.1007/s11069-004-8891-3 (vid. pág. 1)

Kashefipour, S. M., Lin, B., Harris, E. & Falconer, R. A. (2002). Hydro-environmental modelling for bathing water compliance of an estuarine basin. *Water Research*. https://doi.org/10.1016/S0043-1354(01)00396-7 (vid. pág. 14)

Kay, D., Stapleton, C. M., Wyer, M. D., McDonald, A. T., Crowther, J., Paul, N., Jones, K., Francis, C., Watkins, J., Wilkinson, J., Humphrey, N., Lin, B., Yang, L., Falconer, R. A. & Gardner, S. (2005). Decay of intestinal enterococci concentrations in high-energy estuarine and coastal waters: Towards real-time T 90 values for modelling faecal indicators in recreational waters. *Water Research*. https://doi.org/10.1016/j.watres.2004.11.014 (vid. pág. 14)

Kuznetsov, E. & Stegailov, V. (2019). Porting CUDA-Based Molecular Dynamics Algorithms to AMD ROCm Platform Using HIP Framework: Performance Analysis. En V. Voevodin & S. Sobolev (Eds.), *Supercomputing* (pp. 121-130). Springer International Publishing. https://doi.org/10.1007/978-3-030-36592-9_11. (Vid. págs. 12, 29)

Lacasta, A., Morales-Hernández, M., Murillo, J. & García-Navarro, P. (2014). An optimized GPU implementation of a 2D free surface simulation model on unstructured meshes. *Advances in Engineering Software*. https://doi.org/10.1016/j.advengsoft.2014.08.007 (vid. pág. 11)

Lapedus, M. (2019). 5nm Vs. 3nm. Consultado 9 de diciembre de 2020, desde https://semiengineering.com/5nm-vs-3nm/. (Vid. pág. 9)

Lee, V. W., Kim, C., Chhugani, J., Deisher, M., Kim, D., Nguyen, A. D., Satish, N., Smelyanskiy, M., Chennupaty, S., Hammarlund, P., Singhal, R. & Dubey, P. (2010). Debunking the 100X GPU vs. CPU myth: an evaluation of throughput computing on CPU and GPU. *Proceedings of the 37th annual international symposium on Computer architecture*, 451-460. https://doi.org/10.1145/1815961.1816021 (vid. pág. 11)

LeVeque, R. J. (2002). *Finite Volume Methods for Hyperbolic Problems* [Google-Books-ID: O_ZjpMSZiwoC]. Cambridge University Press. (Vid. pág. 23).

LeVeque, R. J. & George, D. L. (2008). High-resolution finite volume methods for the shallow water equations with bathymetry and dry states. *Advanced Numerical Mo-*

dels for Simulating Tsunami Waves and Runup (pp. 43-73). WORLD SCIENTIFIC. https://doi.org/10.1142/9789812790910_0002. (Vid. pág. 160)

Liu, C., Guo, L., Ye, L., Zhang, S., Zhao, Y. & Song, T. (2018). A review of advances in China's flash flood early-warning system. *Natural Hazards*, *92*(2), 619-634. https://doi.org/10.1007/s11069-018-3173-7 (vid. pág. 2)

Liu, Q., Qin, Y. & Li, G. (2018). Fast Simulation of Large-Scale Floods Based on GPU Parallel Computing. *Water*, *10*(5), 589. https://doi.org/10.3390/w10050589 (vid. pág. 11)

Madsen, H. (2000). Automatic calibration of a conceptual rainfall–runoff model using multiple objectives. *Journal of Hydrology*, *235*(3), 276-288. https://doi.org/10.1016/S0022-1694(00)00279-1 (vid. pág. 13)

Manning, R., Griffith, J. P., Pigot, T. F. & Vernon-Harcourt, L. F. (1890). *On the flow of water in open channels and pipes*. (Vid. pág. 22).

MeteoGalicia. (2021). Consultado 5 de enero de 2021, desde https://www.meteogalicia.gal/web/inicio.action. (Vid. pág. 34)

Moore, G. E. (1965). *Cramming more components onto integrated circuits*. McGraw-Hill New York, NY, USA: (vid. pág. 6).

Moore, G. E. (1975). Progress in digital integrated electronics. *Electron devices meeting*, *21*, 11-13 (vid. pág. 6).

Munshi, A. (2009). The OpenCL specification. *2009 IEEE Hot Chips 21 Symposium (HCS)*, 1-314. https://doi.org/10.1109/HOTCHIPS.2009.7478342 (vid. págs. 7, 12)

N. Moriasi, D., G. Arnold, J., W. Van Liew, M., L. Bingner, R., D. Harmel, R. & L. Veith, T. (2007). Model Evaluation Guidelines for Systematic Quantification of Accuracy in Watershed Simulations [Place: St. Joseph, MI Publisher: ASABE]. *Transactions of the ASABE*, *50*(3), 885-900. https://doi.org/10.13031/2013.23153 (vid. págs. 5, 13, 35, 36)

Najafi, M., Zhang, K., Sadoghi, M. & Jacobsen, H. (2017). Hardware Acceleration Landscape for Distributed Real-Time Analytics: Virtues and Limitations [ISSN: 1063-6927]. *2017 IEEE 37th International Conference on Distributed Computing*

Systems (ICDCS), 1938-1948. https://doi.org/10.1109/ICDCS.2017.194 (vid. pág. 9)

Naverac, V. A., Ferrer, D. B., Curiel, P. B., Zueco, S. D., Gil, F. E., Hidalgo, J. C. G., García, D. G., Matauco, A. I. G. d., Albero, C. M., Mur, D. M., Romero, M. E. N., Ojeda, A. O., Fabre, M. S., Sánchez, M. Á. S. & Notivoli, R. S. (2012). Sobre las precipitaciones de octubre de 2012 en el Pirineo aragonés, su respuesta hidrológica y la gestión de riesgos [Publisher: Departamento de Geografía y Ordenación del Territorio Section: Geographicalia]. *Geographicalia*, (61), 101-128. Consultado 4 de enero de 2021, desde https://dialnet.unirioja. es/servlet/articulo?codigo=4138905 (vid. pág. 33)

Neelz, S. & Pender, G. (2010). *Delivering benefits through evidence-Benchmarking of 2D Hydraulic Modelling Packages*. (Vid. págs. 5, 31).

Nickolls, J. & Dally, W. J. (2010). The GPU Computing Era [Conference Name: IEEE Micro]. *IEEE Micro, 30*(2), 56-69. https://doi.org/10.1109/MM.2010.41 (vid. pág. 10)

Noh, S. J., Lee, J.-H., Lee, S., Kawaike, K. & Seo, D.-J. (2018). Hyper-resolution 1D-2D urban flood modelling using LiDAR data and hybrid parallelization [Publisher: Elsevier]. *Environmental Modelling & Software*, *103*, 131-145 (vid. pág. 160).

November 2020 | TOP500. (2020). Consultado 16 de diciembre de 2020, desde https://www.top500.org/lists/top500/2020/11/. (Vid. pág. 10)

November 2020 | TOP500. (2020). Consultado 8 de enero de 2021, desde https://www.top500.org/lists/green500/2020/11/. (Vid. pág. 10)

NVIDIA. (2020). CUDA C++ Programming Guide. Consultado 22 de junio de 2020, desde https://docs.nvidia.com/cuda/pdf/CUDA_C_Programming_Guide.pdf. (Vid. págs. 7, 9, 28)

Ortega Becerril, J. A. (2008). *Paleocrecidas, avenidas recientes e hidroclimatología en la cuenca media y baja del río Guadiana* (info:eu-repo/semantics/doctoralThesis) [ISBN: 9788466930260]. Universidad Complutense de Madrid. Madrid. Consultado 5 de enero de 2021, desde https://eprints.ucm.es/7664/. (Vid. pág. 36)

Paix, M. J. d. l., Lanhai, L., Xi, C., Ahmed, S. & Varenyam, A. (2013). Soil Degradation and Altered Flood Risk as a Consequence of Deforestation [_eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1002/ldr.1147]. *Land Degradation & Development*, *24*(5), 478-485. https://doi.org/https://doi.org/10.1002/ldr.1147 (vid. pág. 2)

Passerotti, G., Massazza, G., Pezzoli, A., Bigi, V., Zsótér, E. & Rosso, M. (2020). Hydrological Model Application in the Sirba River: Early Warning System and GloFAS Improvements [Number: 3 Publisher: Multidisciplinary Digital Publishing Institute]. *Water*, *12*(3), 620. https://doi.org/10.3390/w12030620 (vid. pág. 1)

Ponce, V. M. & Yevjevich, V. (1978). Muskingum-Cunge method with variable parameters [Publisher: ASCE]. *Journal of the Hydraulics Division*, *104*(12), 1663-1667 (vid. pág. 25).

Proposal for a Council Regulation on establishing the European High Performance Computing Joint Undertaking (inf. téc. COM(2020) 569 final). (2020). European Commission. Brussels, Belgium. Consultado 2 de enero de 2021, desde https://ec.europa.eu/newsroom/dae/document.cfm?doc_id=69379. (Vid. pág. 26)

Refsgaard, J. C. (1997). Parameterisation, calibration and validation of distributed hydrological models. *Journal of Hydrology*, *198*(1), 69-97. https://doi.org/10.1016/S0022-1694(96)03329-X (vid. pág. 13)

Ribó, R., Pasenau, M., Escolano, E., Ronda, J. S. P. & González, L. F. (1998). GiD reference manual. *CIMNE, Barcelona*, *27* (vid. págs. 4, 21).

Roe, P. L. (1986). Discrete models for the numerical analysis of time-dependent multidimensional gas dynamics. *Journal of Computational Physics*. https://doi.org/10.1016/0021-9991(86)90204-4 (vid. pág. 23)

Rosburg, T. T., Nelson, P. A. & Bledsoe, B. P. (2017). Effects of Urbanization on Flow Duration and Stream Flashiness: A Case Study of Puget Sound Streams, Western Washington, USA [_eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1111/1752-1688.12511]. *JAWRA Journal of the American Water Resources Association*, *53*(2), 493-507. https://doi.org/10.1111/1752-1688.12511 (vid. pág. 2)

Ross, P. E. (2008). Why CPU Frequency Stalled [Conference Name: IEEE Spectrum]. *IEEE Spectrum*, *45*(4), 72-72. https://doi.org/10.1109/MSPEC.2008.4476447 (vid. pág. 6)

Rotem, E., Ginosar, R., Mendelson, A. & Weiser, U. C. (2013). Power and thermal constraints of modern system-on-a-chip computer. *19th International Workshop on Thermal Investigations of ICs and Systems (THERMINIC)*, 141-146. https://doi.org/10.1109/THERMINIC.2013.6675226 (vid. pág. 10)

SAIH del Miño-Sil. Confederación Hidrográfica del Miño-Sil. (2021). Consultado 5 de enero de 2021, desde http://saih.chminosil.es/index.php?url=/datos/mapas/mapa:H1/area:HID/acc:. (Vid. pág. 34)

Saltelli, A., Ratto, M., Andres, T., Campolongo, F., Cariboni, J., Gatelli, D., Saisana, M. & Tarantola, S. (2008). *Global sensitivity analysis: the primer*. John Wiley & Sons. (Vid. pág. 13).

Sanz-Ramos, M., Castellet, E. B., Ibars, A. P., Querol, D. V. & Ramos-Fuertes, A. (2019). IberHABITAT: evaluación de la Idoneidad del Hábitat Físico y del Hábitat Potencial Útil para peces. Aplicación en el río Eume [Publisher: Taylor & Francis _eprint: https://doi.org/10.1080/23863781.2019.1664273]. *Ribagua*, *6*(2), 158-167. https://doi.org/10.1080/23863781.2019.1664273 (vid. pág. 21)

Scharffenberg, B., Bartles, M., Brauer, T., Fleming, M. & Karlovits, G. (2018). *Hydrologic Modeling System (HEC-HMS). User's Manual:Version 4.3* (Computer Program Document CPD-74A). U.S. Army Corps of Engineers. https://www.hec.usace.army.mil/software/hec-hms/documentation/HEC-HMS_Users_Manual_4.3.pdf. (Vid. págs. 24, 25)

(SCS), S. C. S. (1985). "Hydrology" National engineering handbook, Section 4. U.S. Dept. of Agriculture (USDA). (Vid. pág. 24).

(SCS), S. C. S. (1986). *Technical Release 55: Urban Hydrology for Small Watersheds* (inf. téc.). U.S. Dept. of Agriculture (USDA). Washington, DC. (Vid. pág. 25).

Serrano-Muela, M. P. P., Nadal-Romero, E., Lana-Renault, N., González-Hidalgo, J. C. C., López-Moreno, J. I. I., Beguería, S., Sanjuan, Y. & García-Ruiz, J. M. M. (2015). An exceptional rainfall event in the central western pyrenees:

Spatial patterns in discharge and impact. *Land Degradation and Development*. https://doi.org/10.1002/ldr.2221 (vid. pág. 33)

Service, N. N. R. C. (2007). National Engineering Handbook: Chapter 16 Hydrographs. (Vid. pág. 35).

Severance, C. & Dowd, K. (2010). *High Performance Computing* [Google-Books-ID: 9MaQzQEACAAJ]. OpenStax CNX. (Vid. pág. 26).

Skamarock, W. C., Klemp, J. B., Dudhia, J., Gill, D. O., Barker, D. M., Wang, W. & Powers, J. G. (2005). *A description of the advanced research WRF version 2* (inf. téc.). National Center For Atmospheric Research Boulder Co Mesoscale y Microscale ... (Vid. pág. 34).

Smith, G. D. (1985). *Numerical solution of partial differential equations: finite difference methods*. Oxford university press. (Vid. pág. 2).

Snir, M., Gropp, W., Otto, S., Huss-Lederman, S., Dongarra, J. & Walker, D. (1998). *MPI−the Complete Reference: The MPI core* [Google-Books-ID: x79puJ2YkroC]. MIT Press. (Vid. pág. 27).

Snyder, F. F. (1938). Synthetic unit-graphs [_eprint: https://agupubs.onlinelibrary.wiley.com/doi/pdf/10.1029/TR019i001p00447]. *Eos, Transactions American Geophysical Union*, *19*(1), 447-454. https://doi.org/https://doi.org/10.1029/TR019i001p00447 (vid. pág. 25)

Sopelana, J., Cea, L. & Ruano, S. (2018). A continuous simulation approach for the estimation of extreme flood inundation in coastal river reaches affected by meso- and macrotides. *Natural Hazards*, *93*(3), 1337-1358. https://doi.org/10.1007/s11069-018-3360-6 (vid. pág. 5)

Suggs, D., Subramony, M. & Bouvier, D. (2020). The AMD "Zen 2" Processor [Conference Name: IEEE Micro]. *IEEE Micro*, *40*(2), 45-52. https://doi.org/10.1109/MM.2020.2974217 (vid. pág. 27)

Sun, Y., Agostini, N. B., Dong, S. & Kaeli, D. (2019). Summarizing CPU and GPU Design Trends with Product Data. *arXiv preprint arXiv:1911.11313* (vid. pág. 10).

Sutter, H. (2005). The free lunch is over: A fundamental turn toward concurrency in software. *Dr. Dobb's journal*, *30*(3), 202-210 (vid. pág. 6).

Sutter, H. & Larus, J. (2005). Software and the Concurrency Revolution: Leveraging the full power of multicore processors demands new tools and new thinking from the software industry. *Queue*, *3*(7), 54-62. https://doi.org/10.1145/1095408.1095421 (vid. pág. 6)

Tallaksen, L. M. (1995). A review of baseflow recession analysis. *Journal of Hydrology*, *165*(1), 349-370. https://doi.org/10.1016/0022-1694(94)02540-R (vid. pág. 25)

Thielen, J., Bartholmes, J., Ramos, M.-H. & de Roo, A. (2009). The European Flood Alert System – Part 1: Concept and development [Publisher: Copernicus GmbH]. *Hydrology and Earth System Sciences*, *13*(2), 125-140. https://doi.org/https://doi.org/10.5194/hess-13-125-2009 (vid. pág. 15)

Trigo, R. M., Varino, F., Ramos, A. M., Valente, M. A., Zêzere, J. L., Vaquero, J. M., Gouveia, C. M. & Russo, A. (2014). The record precipitation and flood event in Iberia in December 1876: description and synoptic analysis [Publisher: Frontiers]. *Frontiers in Earth Science*, *2*, 3 (vid. pág. 36).

TUFLOW FV Science Manual. (2013). Consultado 4 de enero de 2021, desde https://www.tuflow.com/Download/TUFLOW_FV/Manual/FV_Science_Manual_2013.pdf. (Vid. págs. 6, 32)

US EPA, O. (2002). Guidance for Quality Assurance Project Plans for Modeling EPA QA/G-5M. Consultado 15 de diciembre de 2020, desde https://www.epa.gov/quality/guidance-quality-assurance-project-plans-modeling-epa-qag-5m. (Vid. pág. 5)

Vacondio, R., Dal Palù, A. & Mignosa, P. (2014). GPU-enhanced finite volume shallow water solver for fast flood simulations. *Environmental Modelling and Software*. https://doi.org/10.1016/j.envsoft.2014.02.003 (vid. pág. 11)

Varghese, B., Prades, J., Reaño, C. & Silla, F. (2015). Acceleration-as-a-Service: Exploiting Virtualised GPUs for a Financial Application. *2015 IEEE 11th International Conference on e-Science*, 47-56. https://doi.org/10.1109/eScience.2015.15 (vid. pág. 11)

Varghese, B. & Buyya, R. (2018). Next generation cloud computing: New trends and research directions. *Future Generation Computer Systems*, *79*, 849-861. https://doi.org/10.1016/j.future.2017.09.020 (vid. pág. 11)

Viñas, M., Lobeiras, J., Fraguela, B. B., Arenaz, M., Amor, M., García, J. A., Castro, M. J. & Doallo, R. (2013). A multi-GPU shallow-water simulation with transport of contaminants [_eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1002/cpe.2917]. *Concurrency and Computation: Practice and Experience*, *25*(8), 1153-1169. https://doi.org/https://doi.org/10.1002/cpe.2917 (vid. pág. 160)

Wallemacq, P., House, R., Below, R. & McLean, D. (2018). *Economic losses, poverty & disasters: 1998-2017* (inf. téc.) [Library Catalog: www.undrr.org]. Centre for Research on the Epidemiology of Disasters (CRED) y United Nations Office for Disaster Risk Reduction (UNISDR). Brussels, Belgium. Consultado 22 de junio de 2020, desde https://www.undrr.org/publication/economic-losses-poverty-disasters-1998-2017. (Vid. pág. 1)

Wu, Y., Falconer, R. & Lin, B. (2005). Modelling trace metal concentration distributions in estuarine waters. *Estuarine, Coastal and Shelf Science*, *64*(4), 699-709. https://doi.org/10.1016/j.ecss.2005.04.005 (vid. pág. 14)

Zoch, R. T. (1934). On the relation between rainfall and streamflow. *Monthly Weather Review*, *62*(9), 315-322 (vid. pág. 25).

Universida_{de}Vigo