

Interreg

Sudoe

INUNDATIO



European Regional Development Fund

PROTOCOLO PARA REALIZAR ANÁLISIS DE INCERTIDUMBRE EN CUENCAS DE MONTAÑA

Proyecto INUNDATIO

1. Introducción

La presente entrega forma parte del proyecto financiado por fondos del programa europeo INTERREG SUDOE titulado “Automatización del modelado de riesgos de inundaciones en cabeceras de cuenca a través de técnicas de inteligencia artificial y Big Data” y cuyo acrónimo es “INUNDATIO” (Código SOE3/P4/E0929). Se expone el desarrollo de la actividad 2.3 **“Evaluación de los datos obtenidos en términos de probabilidad, vulnerabilidad, recurrencia, impacto”**.

Es pertinente mencionar que, dentro del Grupo de Tareas 1 “Desarrollo de la metodología de toma de datos de las cabeceras de cuenca”, se ha propuesto entregar un “conjunto de actuaciones que lleven a la recogida de datos de cuencas para su posterior análisis” conocido como producto 1. El objetivo de esta entrega consiste en disminuir la incertidumbre asociada a la falta de datos en cabeceras de montaña para mejorar la fiabilidad de los análisis de frecuencias por inundaciones. Este objetivo se abordará de diferentes perspectivas de acuerdo con cada cuenca piloto estudiada. No obstante, la estrategia principal consiste en emplear datos alternativos que ofrezcan información de calado y emplear modelos hidrodinámicos 2D para establecer la relación entre cotas y caudales y, posteriormente, analizar la incertidumbre del resultado (en términos de probabilidad, vulnerabilidad, recurrencia y daños) y la sensibilidad de los inputs del modelo.

De acuerdo con lo anterior, se ha elaborado un protocolo adaptable y aplicable a las cuencas piloto seleccionadas para el proyecto. Para reducir todo lo posible las altas incertidumbres asociadas a los modelos de montaña se ha desarrollado un protocolo con herramientas y métodos eficaces y generales basados en Monte Carlo y aplicados a metodologías, también generales, basadas en el cálculo de daños mediante curvas de calado-daños. No obstante, el análisis de incertidumbre y el análisis de sensibilidad pueden ser abordado de diversas maneras de acuerdo con los datos disponibles, el tamaño de la zona estudiada y el nivel de detalle necesario. Por esta razón se ha considerado añadir un breve marco conceptual en el entregable. Complementariamente se ha añadido al marco conceptual un resumen sobre el método de valoración económica que se aplica en la entrega.

Derivado de esta entrega se añade como anexo el script de Python encargado de hacer el análisis de incertidumbre y de sensibilidad.

2. Marco conceptual.

Los análisis relacionados con la incertidumbre deben ser entendidos como análisis integrales donde primero se seleccionen los parámetros de entrada que generan incertidumbre, específicos de cada caso, y posteriormente se evalúa tanto el resultado (resultados estocásticos que provienen del análisis de sensibilidad) y el comportamiento de los parámetros (estimado mediante el análisis de sensibilidad). Además, deben analizarse estos resultados para comprobar el éxito del análisis tanto en términos de convergencia como en términos de las funciones de distribución asignadas a cada parámetro de entrada (Merwade et al., 2008, Rajabi et al., 2015). Un buen esquema para definir este proceso integral lo presenta Merwade (Figura 1).

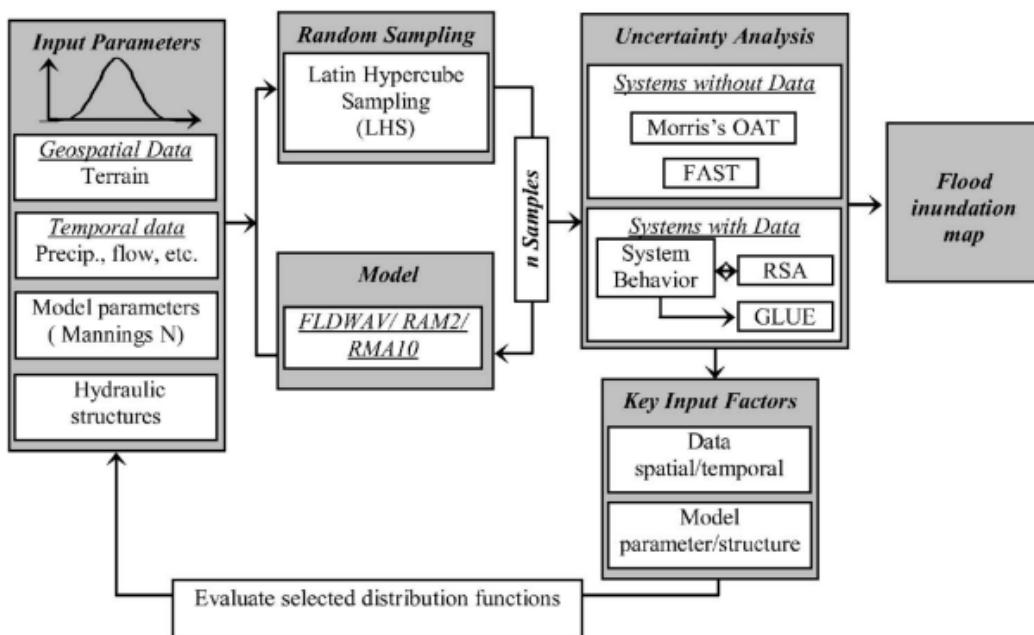


Figura 1 Esquema integral del análisis de incertidumbre. Fuente: Merwade et. al. 2008

Por un lado, por incertidumbre se entiende a todo aquello que no es conocido en su totalidad y que, por tanto, genera inseguridad. En el ámbito de las inundaciones esta inseguridad se traduce en el desconocimiento que se tiene sobre cómo será un evento en un futuro concreto. Este desconocimiento se puede desglosar en tres tipos de incertidumbre: incertidumbre epistémica, aleatoria y ontológica o profunda (Sun et al., 2012, Zarekarizi et al., 2020). De estas, se suelen considerar por sencillez y relevancia las dos primeras.

Por otro lado, por sensibilidad se entiende a la importancia que cada parámetro tiene sobre el resultado de una inundación, en otras palabras, pequeñas variaciones de un parámetro sensible pueden generar mayores cambios en el resultado que grandes variaciones en un parámetro no sensible (Rajabi et al., 2015).

Para realizar ambos procedimientos existen multitud de posibles métodos. Una buena revisión se encuentra en (Rajabi et al., 2015). A modo de resumen, los análisis de incertidumbres se pueden hacer: (1) con métodos posibilistas, que suponen el primer acercamiento al enfoque estocástico; (2) con métodos probabilísticos o estocásticos basados en Monte Carlo y sus derivaciones; (3) con métodos complejos que suponen una reducción significativa en el número

de simulaciones a realizar; y (4) con métodos de propagación estrictamente matemáticos basados en las ecuaciones de los modelos, los cuales no son aplicables en este ámbito por la complejidad de los modelos. Respecto a los análisis de sensibilidad, existen análisis locales y globales, siendo más interesantes estos últimos. Los análisis de sensibilidad global se clasifican según el parámetro en el que se base el proceso para obtener el resultado, existiendo análisis basados en la varianza, en el momento, análisis de regresión lineal o métodos derivativos.

Además, es necesario mencionar que, tanto los métodos basados en Monte Carlo como la mayoría de los métodos de análisis de sensibilidad requieren de realizar simulaciones y, consecuentemente, de haber realizado un muestreo específico de los inputs basado en sus rangos probables. También existen diferentes tipos de muestreo, como el muestreo aleatorio estándar, el muestreo basado en Hipercubos Latinos, el muestreo de Morris o de Sobol para realizar análisis de sensibilidad. La mayoría de estos muestreos no sirven para realizar simultáneamente análisis de sensibilidad y de incertidumbre.

También es necesario seleccionar el tipo de modelo hidráulico empleado para realizar estos procesos. Ambos procesos son procedimientos con un alto coste computacional, lo que obliga a utilizar modelos hidrológicos eficientes y con capacidad de paralelización. Por otro lado, conviene emplear modelos conocidos y fáciles de manejar para facilitar el uso del protocolo (Dysarz Tomasz, 2018).

Por último, también existen numerosos métodos de valoración económica, no obstante, en el ámbito de las inundaciones es una práctica muy estandarizada el uso de métodos basados en curvas de calado-dáño, que evalúan el daño en cada zona en función del calado y del tipo de uso de suelo presente en dicha zona. En España, este tipo de metodología presenta una guía estandarizada creada por el actual Ministerio para la Transición Ecológica y el Reto Demográfico (MAGRAMA, 2013). Esta guía integra la vulnerabilidad de los elementos afectados (usos del suelo) en un método fácilmente aplicable.

Métodos seleccionados.

Cada enfoque tiene ventajas y desventajas. En el presente protocolo se han seleccionado los enfoques que mayores ventajas tienen para el proyecto, siendo:

- **Análisis de incertidumbre.** Se ha seleccionado el método de Monte Carlo (MC) junto con el muestreo realizado con el método de *Latin Hipercuble Sampling* (LHS). De esta manera se logra una convergencia más rápida y el alto coste computacional del método de Monte Carlo se reduce. Posteriormente, estos resultados de incertidumbre de calado y velocidad se han propagado a través de las curvas de calado-dáño para analizar la incertidumbre en los daños.
- **Análisis de sensibilidad.** Se ha seleccionado el método de análisis de sensibilidad global conocido como *Random Balance Designs Fourier Amplitude Sensitivity Test* (RBD-FAST) que puede ser empleado con un muestreo basado en LHS, lo que permite utilizar los mismos modelos para ambos análisis.
- **Modelo hidráulico.** Se ha seleccionado HEC-RAS por sus eficacia y sus facilidades de automatización, al disponer de una API funcional en Python que permite manejar ciertos aspectos del programa de manera automática.

3. Protocolo

El protocolo aplicado consiste en 4 fases: (1) selección de datos de entrada (inputs); (2) muestreo de los inputs por el método de *Latin Hipercuble Sampling* (LHS); (3) adaptación de un código Python para el lanzamiento automático de modelos con el fin de realizar un análisis de incertidumbre empleando el método de Monte Carlo (MC) y un análisis de sensibilidad empleado el método *Random Balance Designs Fourier Amplitude Sensitivity Test* (RBD-FAST); y (4) propagación de resultados para la estimación de los daños económicos. A continuación, se describe cada apartado con sus correspondientes pasos y herramientas.

3.1. Selección de datos de entrada.

La falta de datos en las cuencas de montaña es el mayor limitante de estas. Se considera falta de datos cuando estos son tan escasos que generan grandes rangos posibles y, por lo tanto, grandes incertidumbres. Los datos básicos que podemos encontrar en cuencas de montaña aplicables a la modelización hidráulica e hidrológica son:

- **Caudales de entrada.** Estos datos son el input más importante y el objetivo es obtener un análisis de frecuencia para conocer la distribución de los distintos caudales a lo largo del tiempo.

Pueden obtenerse de manera directa, es decir, midiendo directamente en el curso de agua calado o caudal; o indirecta, donde se dispone de datos de lluvia que son transformados en caudal mediante procedimientos hidrológicos. En ambos casos se suele disponer de una escasa serie de datos que genera grandes incertidumbres o que, incluso, hace imposible su uso, obligando a trabajar con la mayor inundación probable.

- **N de Manning.** Es el parámetro de rugosidad con el que se representa el uso del suelo en el modelo. Este parámetro es desconocido en cierta medida, es decir, se conoce el rango de posibles valores, pero sin mediciones exactas se desconoce el valor exacto. No obstante, es un parámetro fácil de obtener mediante mapas oficiales de usos del suelo o fotointerpretación.
- **Evidencias indirectas.** La evidencia indirecta más estudiada son los daños generados en el arbolado (evidencias dendrocronológicas). Estos datos se presentan en aquellas cuencas vegetadas y aportan información importante en aquellos casos en los que la serie histórica de caudales es nula o escasa. Esta información puede ser empleada para definir el rango de caudales de entrada realizando un estudio previo de incertidumbre. Un ejemplo del uso de estos indicadores se puede ver en los artículos de Ballesteros et. al. (2010) y Ballesteros et. al. (2011) como se puede observar en la Figura 2.

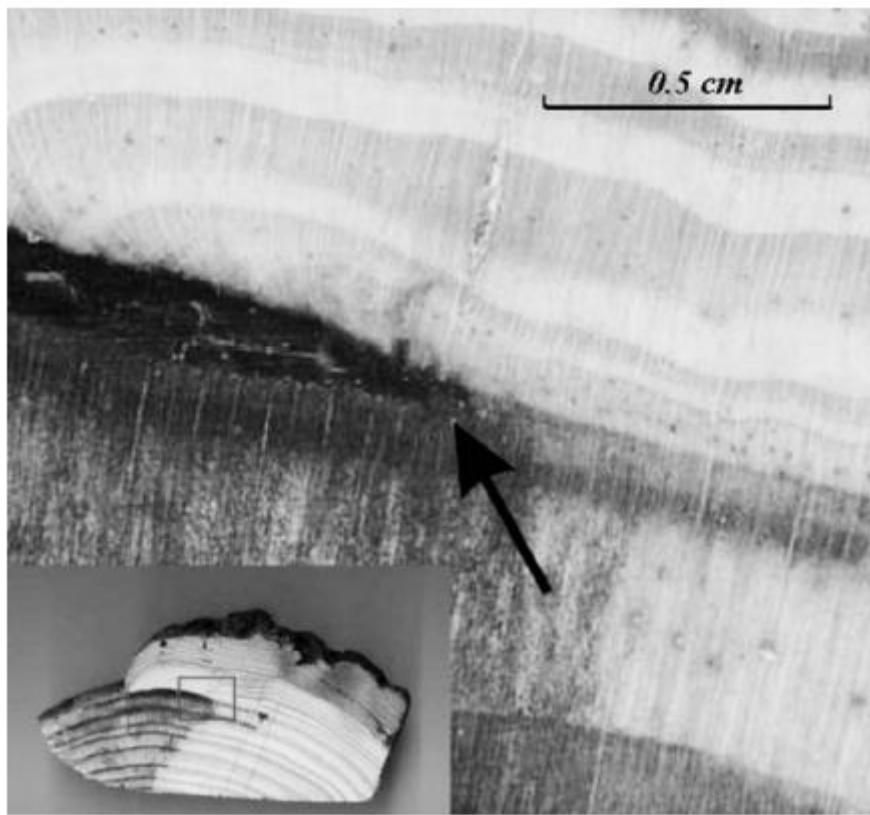


Figura 2. Daños en arbolado provocado por inundaciones. La altura del daño y el estudio del árbol pueden aportar información de las dimensiones de la avenida. Fuente: Ballesteros et. al. 2010.

Adicionalmente se pueden contemplar otros inputs que puedan aportar incertidumbre, siempre que se disponga de la información necesaria. Estos son:

- **Parámetros del modelo.**
- **Infraestructuras hidráulicas.**

3.2. Muestreo

El muestreo de todos los parámetros de entrada se ha realizado mediante el método *Latin Hipercubicle Sampling* (LHS). Este método de muestreo pseudo-aleatorio es una mejora al muestreo aleatorio estándar y permite representar la totalidad de las combinaciones posibles con una mayor eficiencia.

El método LHS consiste en dividir el rango de cada input en espacios equidistantes, generando tantos espacios como muestras se deseen tomar. Posteriormente se toma una muestra aleatoria en cada espacio.

Para aplicar este procedimiento en el script de Python se puede utilizar la librería *lhsmd*. Esta librería sirve para generar una muestra mediante este procedimiento entre 0 y 1. Posteriormente, con esta muestra se evalúa la función de distribución acumulada (PDF por sus siglas en inglés) de cada input. Al realizar este procedimiento con cada input se obtiene un muestreo realizado por LHS. Complementariamente, para poder tratar los datos, calcular las

funciones PDF y evaluar cada valor en dicha distribución son necesarias las librerías numpy, scipy.stats, os y pickle.

Como se ha mencionado y se muestra en la Figura 3 el muestreo LHS es más eficiente que un muestreo aleatorio estándar a la hora de representar la función de distribución de un input dado.

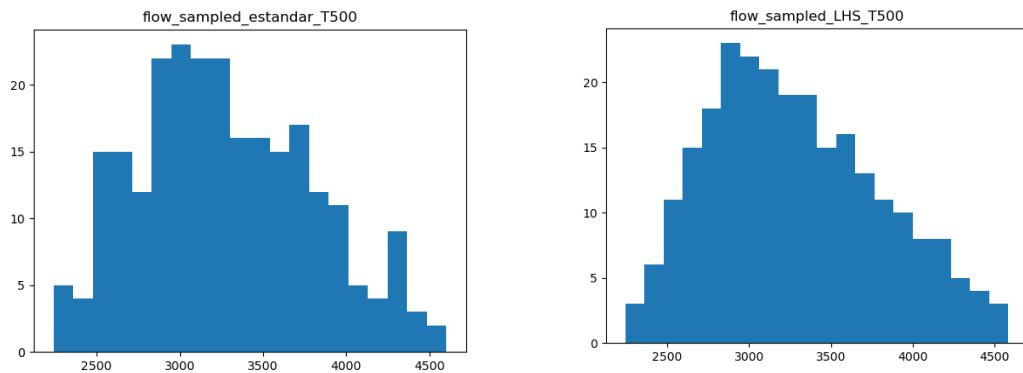


Figura 3. Comparativa entre muestreo aleatorio estándar y muestreo generado por LHS. Tamaño del muestreo de 250 y distribución de probabilidad triangular. Fuente: Elaboración propia.

Misma muestra aleatoria de tamaño 250 para la misma distribución triangular. Muestreo aleatorio estándar a la derecha y muestreo LHS izquierda.

3.3. Lanzamiento automático de modelos

Después de generar la muestra de cada input seleccionado mediante LHS se realiza un proceso que consta de tres fases: (A) Crear modelo de referencia funcional; (B) replicar el modelo funcional, editar cada simulación son los inputs correspondientes, lanzar el modelo y guardar sus resultados; y (C) comprobar la convergencia del modelo.

- Para crear el modelo funcional se emplea HEC-RAS. HEC-RAS es ampliamente conocido, utilizado y refutado, además de ser eficiente y de disponer de una herramienta (HECRASController) que permite su automatización con Python.

Se debe generar un modelo sólido, refinado y eficiente, con el fin de aumentar la velocidad del proceso. Después de comprobar su funcionalidad y velocidad se debe preparar este modelo como plantilla para el proceso.

La manera más adecuada de adaptarlo como plantilla es establecer valores clave para los diferentes inputs que sirvan para el procedimiento de automatización. Por ejemplo, el caudal de entrada se guarda en un archivo de HEC-RAS con la extensión “.u01”. Se puede establecer en la plantilla caudales con un valor de “7777.77” como valores clave para que el script de Python sea capaz de buscar estos valores y reemplazarlos por los caudales de entrada correspondientes a cada simulación (Figura 4).

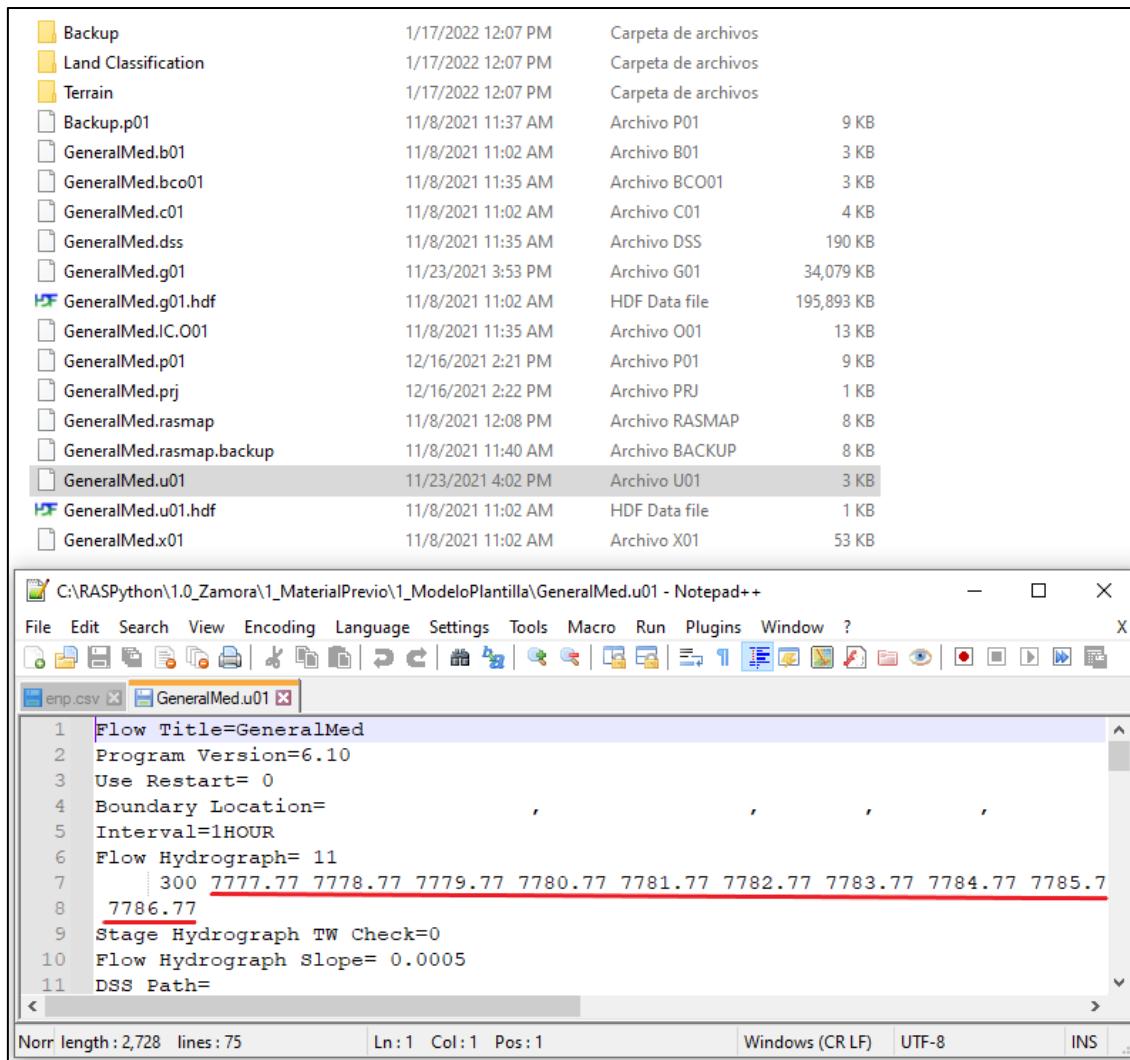


Figura 4. Ejemplo de adaptación de modelo hidráulico generado en HEC-RAS con carpetas y archivos del modelo en la zona superior y valores clave para los caudales de entrada en la zona inferior. Fuente: Elaboración propia.

- B. Este paso consiste en clonar la plantilla del modelo funcional con sus valores clave y adaptar el código para modificar cada input del modelo. Actualmente el caudal y los parámetros del modelo se pueden modificar directamente en los archivos. Sin embargo, la N de Manning en 2D no pueden cambiarse directamente en los archivos, pero si puede automatizarse su cambio tomando el control del ordenador (mouse y teclado) con la librería pyautogui.

Después de modificar cada plantilla clonada con los inputs correspondientes se lanza el modelo y, después de que este termine, se guardan sus resultados.

- C. Después de terminar todas las simulaciones se emplea el Teorema del Límite Central (CLT por sus siglas en inglés) para evaluar la convergencia de la mancha de inundación generada por el modelo. Empleando este teorema, el modelo se considera estable cuando se mantiene durante, al menos, 50 simulaciones seguidas dentro del margen especificado, tal y como se muestra en la **figx** extraída de Mustafa, Y. (2007).

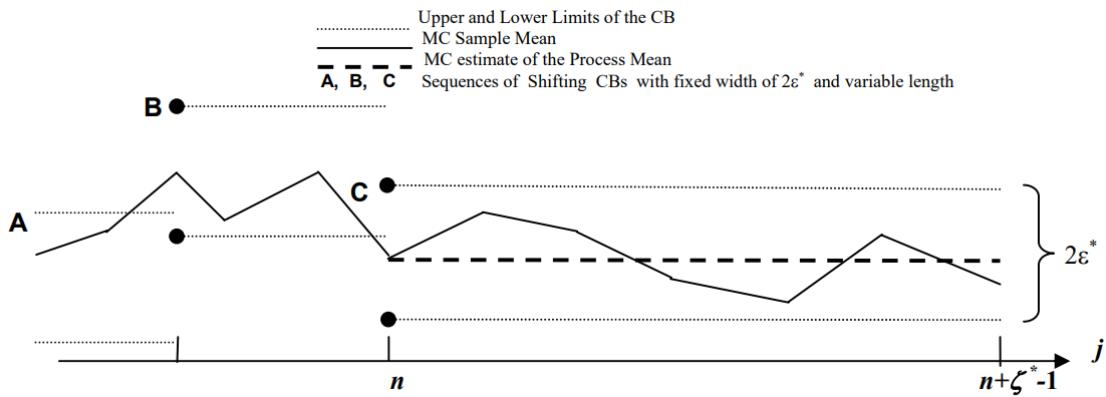


Figura 5. Esquema de convergencia de un modelo empleando el método de Monte Carlo. Cuando el promedio del parámetro se mantenga dentro de la banda de incertidumbre durante un número determinado de simulaciones el modelo a convergido. Fuente: Mustafa, Y. 2007.

Para realizar todo este proceso se necesitan librerías complementarias como os, pickle, win32.com.client, pyautogui, shutil, gdal o numpy.

Posteriormente se deben procesar los resultados, guardados en formato “.tiff”, para terminar el análisis de incertidumbre y de sensibilidad. Respecto al análisis de incertidumbre, se obtendrán como resultados mapas estocásticos de calado, de velocidad, de peligro y de daños. Respecto al análisis de sensibilidad, se obtendrán resultados de la importancia de cada parámetro a escala global (mancha de inundación) y a escala de detalle (celda por celda).

Para procesar todos los resultados de los modelos y obtener los citados resultados se puede trabajar con librerías como gdal, numpy, osr, SALib o scipy.stats.

3.4. Estimación de daños.

La última fase del protocolo consiste en propagar la incertidumbre al cálculo de daños económicos. Para lograr este fin se deben realizar los siguientes pasos, explicados en detalle en la metodología elaborada por el Ministerio para la Transición Ecológica y el Reto Demográfico (MAGRAMA, 2013). Cabe mencionar que el proceso explicado a continuación debe hacerse para cada resultado obtenido de calado y que debe adaptarse para cada cuenca piloto estudiada, ya que los tres pasos de la estimación de daños son propios de cada zona geográfica.

A. Recolección de datos de usos del suelo.

La recolección de datos de usos debería hacerse con información de detalle a nivel nacional, como es el caso del Sistema de información sobre Ocupación del Suelo de España (SIOSE). No obstante, se podrían utilizar en su defecto bases de datos internacionales o supranacionales sobre información del suelo como el Corine Land Cover (CLC).

B. Asignación de valor unitario por tipo de uso del suelo.

La asignación del valor unitario a cada tipo de suelo viene dada mayormente por la zona geográfica, por lo que es necesario adaptar estos valores a cada zona de estudio. Como referencia se indican en la siguiente tabla los valores unitarios

asignados a los tipos de uso del suelo presentes en España de acuerdo con la metodología citada.

Uso del suelo	Valor del riesgo inicial (€/m ²)
Urbano concentrado	
Edificación asociada a urbano concentrado (sin desagregar edificaciones)	350
Edificación asociada a urbano concentrado (edificios desagregados)	400
Urbano disperso	
Edificación asociada a urbano disperso (sin desagregar edificaciones)	170
Edificación asociada a urbano disperso (edificios desagregadas)	260
Asociado a urbano	150
Infraestructura social	200
Terciario	380
Industrial concentrado	
Industrial concentrado (sin desagregar edificaciones)	450
Industrial concentrado (edificios desagregados)	380
Industrial disperso	
Industrial disperso (sin desagregar edificaciones)	170
Industrial disperso (edificios desagregados)	380
Agrícola – Secano	1
Agrícola – Regadío	5
Agrícola indeterminado	3
Otros usos rurales	0.5
Forestal	0
Infraestructuras: carreteras	250
Infraestructuras: ferrocarriles	350
Infraestructuras: aeroportuarias	450
Infraestructuras: energía	500
Infraestructuras: comunicaciones	500
Infraestructuras: hidráulico - sanitarias	500
Infraestructuras: residuos	150
Masas de agua	0
Otras áreas sin riesgo económico	0

Figura 6. Tabla de valores unitarios de referencia establecida en la metodología de valoración de daños propuesta por el Ministerio. Fuente: MAGRAMA. 2013.

C. Definición de las curvas calado-dáño.

La definición de este tipo de curvas es específica de cada zona geográfica, por lo que es necesario comprobar la adecuación de la curva mediante revisión bibliográfica.

Este tipo de curvas representan el porcentaje del daño total que un calado dado puede generar basándose en un valor unitario de referencia que, en este caso, está determinado para cada uso del suelo. Esto quiere decir que el daño se evaluará en un porcentaje desde 0 hasta 100 con respecto al valor unitario.

Como referencia, se presenta la curva de daños de referencia empleada en territorio español para la estimación de daños.

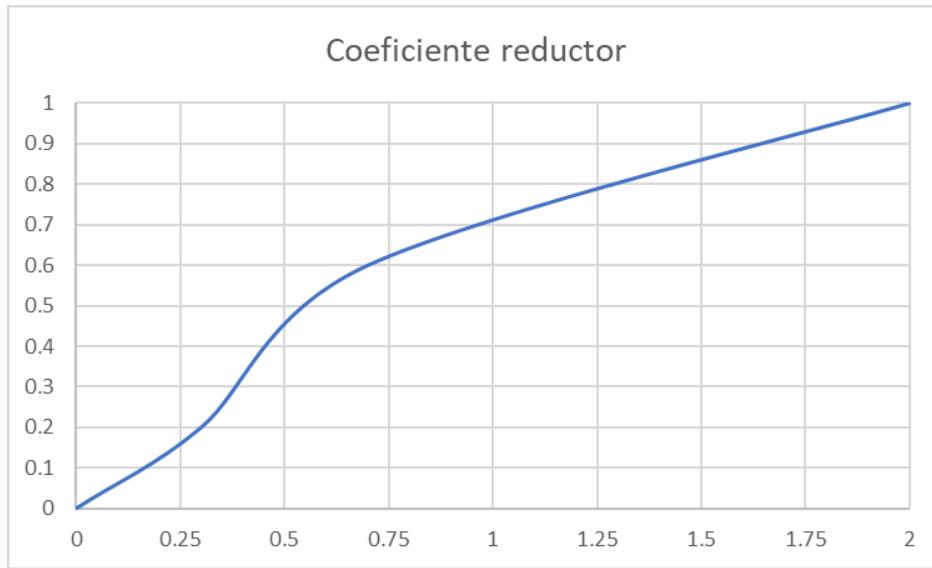


Figura 7. Curva de coeficientes reductores para la estimación de daños mediante valores unitarios de daño en función de los usos del suelo. Fuente: Adaptado de MAGRAMA. 2013.

D. Cálculo económico estocástico.

Para cada celda de cada resultado obtenido debe evaluarse el calado obtenido en la curva de daño y aplicarse al valor unitario del uso del suelo.

En este último paso se pueden obtener mapas estocásticos de daños económicos o se pueden integrar los cálculos para analizar el daño económico estocástico en conjunto o por tipo de uso de suelo.

4. Referencias

- Ballesteros, C., et al. 2010. Changes in wood anatomy in tree rings of *Pinus pinaster* Ait. following wounding by flash floods. *Tree-Ring Research*, vol. 66, no 2, p. 93-103.
- Ballesteros, C., et al. 2011. Estimating flash flood discharge in an ungauged mountain catchment with 2D hydraulic models and dendrogeomorphic palaeostage indicators. *Hydrological Processes*, vol. 25, no 6, p. 970-979.
- Dysarz, T. 2018. Application of python scripting techniques for control and automation of HEC-RAS simulations. *Water*, 10(10), 1382.
- MAGRAMA. 2013. Propuesta de mínimos para la metodología de realización de los mapas de riesgos de inundación. España: MAGRAMA Ministerio de Agricultura, Alimentación y Medio Ambiente. 46 pp.
- Merwade, V., Olivera, F., Arabi, M., & Edleman, S. 2008. Uncertainty in flood inundation mapping: current issues and future directions. *Journal of Hydrologic Engineering*, 13(7), 608-620.
- Mustafa Y. 2007. A convergence criterion for the Monte Carlo estimates. *Simulation Modelling Practice and Theory*, vol. 15, no 3, p. 237-246.
- Rajabi, M. M., Ataie-Ashtiani, B., Simmons, C. T. 2015. Polynomial chaos expansions for uncertainty propagation and moment independent sensitivity analysis of seawater intrusion simulations. *Journal of Hydrology*, 520, 101-122.
- Sun, S., Fu, G., Djordjević, S., Khu, S. T. 2012. Separating aleatory and epistemic uncertainties: Probabilistic sewer flooding evaluation using probability box. *Journal of hydrology*, 420, 360-372.
- Zarekarizi, M., Srikrishnan, V., Keller, K. 2020. Neglecting uncertainties biases house-elevation decisions to manage riverine flood risks. *Nature communications*, 11(1), 1-11.