Calibración del modelo de nieve con datos satelitales

Índice

[1 Introducción](#_Toc101717488)

[2 Zona de estudio](#_Toc101717489)

[3 Datos](#_Toc101717490)

[3.1 Cobertura de nieve](#_Toc101717491)

[3.2 Clima](#_Toc101717492)

[4 Métodos](#_Toc101717493)

[4.1 Modelo de nieve: método de grado-día](#_Toc101717494)

[4.2 Calibración](#_Toc101717495)

[5 Resultados](#_Toc101717496)

[6 Discusión](#_Toc101717497)

# Introducción

Ésta es una primera aproximación a la calibración espacial de un modelo hidrológico en la que se calibra un modelo de nieve muy simple, el método grado-día, frente a los datos de los satélites Terra y Aqua de la constelación MODIS.

# Zona de estudio

La cuenca de estudio es la cabecera del río Deva hasta la estación de aforo de Ojedo. Aunque para el modelo de nieve no es necesario que la zona de estudio sea una cuenca hidrográfica, ni que tenga una estación de aforo, he escogido esta cuenca pensando en poder calibrar más adelante el caudal.

Podría ser interesante ampliar la zona de estudio hasta la estación de aforo de Puentelles, justo antes de la confluencia con el río Cares. No lo he hecho por ahora por reducir el tiempo de cálculo.

# Datos

## Cobertura de nieve

He utilizado como base de datos de nieve los productos de MODIS de extensión máxima de la cobertura de nieve: MOD10A2 para el satélite Terra y MYD10A2 para el satélite Aqua. Esta constelación de satélites ofrece datos desde 2002 con una resolución espacial de 465 m y, originalmente, resolución temporal diaria. Puesto que en los mapas diarios hay numerosos huecos causados por nubes, imágenes nocturnas, etc., los productos MODIS anteriores no tienen resolución diaria, sino de 8 días. En concreto, los productos MOD10A2 y MYD10A2 son mapas clasificados con la cobertura máxima de nieve observada en los 8 días previos. En la Tabla 1 se muestran las clases que incluyen estos productos:

Tabla . Clases en los productos MODIS de máxima cobertura de nieve.

|  |  |
| --- | --- |
| **Código** | **Clase** |
| 0 | Dato faltante |
| 1 | Indeciso |
| 11 | Noche |
| 25 | Sin nieve |
| 37 | Lago |
| 39 | Océano |
| 50 | Nube |
| 100 | Lago helado |
| 200 | Nieve |
| 254 | Detector saturado |
| 255 | Relleno |

En los datos dentro de la cuenca del Deva únicamente hay celdas clasificadas como indeciso (1), sin nieve (440.000 aprox.), nube (55.000 aprox.) y nieve (120.000 aprox). Como sólo interesa tener un mapa binario de presencia-ausencia de nieve, he reclasificado como NaN todas las celdas que no fueran de las clases “sin nieve” o “nieve”, y a estas dos clases les he dado los valores 0 (ausencia) y 1 (presencia).

He comparado los datos de los dos satélites (Terra y Aqua) de cara a unificarlos y tener una sola base de datos de observación. Puesto que Aqua tiene mayor falta de datos, he tomado como referencia los datos de Terra, a los que he añadido los datos de Aqua cuando este satélite tiene dato y Terra no.

## Clima

El modelo de nieve de grado-día requiere tres variables climáticas de entrada: la precipitación diaria, la temperatura media (o mínima) diaria y la radiación.

### Precipitación y temperatura

He partido de los mapas diarios de precipitación y temperatura media que generé para IVERCAM. En su momento utilicé los datos de las estaciones de la AEMet y analicé mediante validación cruzada el método de interpolación que mejor se ajustaba a cada variable. Para la precipitación el método con el mejor rendimiento fue *universal kriging*, utilizando como deriva la altitud, mientras que para la temperatura fue la distancia inversa aplicada sobre los residuos de la regresión lineal entre temperatura y altitud.

Los mapas de IVERCAM tiene resolución espacial 1000 m, aproximadamente el doble que la resolución de MODIS. De cara a utilizar estos datos en este estudio, he remuestreado los mapas sobre la malla de MODIS en la cuenca del Deva mediante la distancia inversa al cuadrado de las 4 celdas más cercanas.

### Radiación

El método del grado-día incluye la posibilidad de utilizar la radiación incidente como un factor corrector del coeficiente de fusión diario, de manera que haya una mayor fusión en zonas con mayor soleamiento. Para ello, se surte al modelo con mapas mensuales de radiación adimensional, es decir, el cociente entre la radiación incidente teniendo en cuenta la orografía y la radiación incidente en el supuesto de que la zona de estudio fuera plana.

Para calcular estos mapas de radiación he remuestreado el modelo digital del terreno del IGN a la malla de MODIS en la cuenca del Deva, y he hecho una copia de este MDT remuestreado con cota cero constante. Con la herramienta de ArcGIS *area solar radiation* se calcula la radiación incidente en cada mes para el MDT real y el plano; el cociente de estos dos es el valor adimensional de la radiación requerido por el modelo de grado-día.

# Métodos

## Modelo de nieve: método de grado-día

He creado una función Python que replica la implementación del modelo grado-día de TETIS. El modelo tenía originalmente 3 parámetros, al que he añadido uno más. Los parámetros son:

* *Ts*: umbral de temperatura por debajo del cuál se asume que la precipitación es en forma de nieve. Éste es el parámetro que yo he añadido; en TETIS se asume un valor fijo de 0.
* *Tb*: umbral de temperatura por encima del cuál se produce la fusión.
* *DDF1*: índice grado día (mm/°C) para días sin precipitación.
* *DDF2*: índice grado-día (mm/°C) para días con precipitación (en concreto días de lluvia, porque se asume que *Tb* es mayor que *Ts*).

En un paso temporal cualquiera de la simulación, se calcula primero en qué celdas la precipitación () es en forma de lluvia () o de nieve ():

Esta aproximación es relativamente simplista, puesto que no se tiene en cuenta que hay un rango de temperaturas entre las que la precipitación es en forma de aguanieve. La conceptualización del modelo grado-día de HBV sí lo tiene en cuenta; para ello incluye un parámetro más que define la amplitud de dicho rango de temperaturas. Dentro de dicho rango, la distribución entre lluvia y nieve es lineal, de manera que para el valor mínimo del rango toda la precipitación es en forma de nieve, y para el valor máximo del rango toda la precipitación es en forma de lluvia.

Posteriormente se calcula el equivalente agua-nieve () disponible y la fusión potencial ():

Donde y son el equivalente agua-nieve en el paso temporal actual y anterior, respectivamente, y es la radiación adimensional. Por último, se calcula la fusión real () y se sustrae dicho valor para actualizar el equivalente agua-nieve al final del paso temporal:

El modelo grado-día de HBV es más complicado/realista, en cuanto que incluye un almacenamiento de agua líquida dentro del paquete de nieve que, antes de generar escorrentía, puede volver a congelarse en los días fríos.

El modelo aquí descrito genera dos series diarias de mapas como resultado: los mapas de equivalente agua-nieve y los de fusión.

Puesto que los datos observados son de cobertura máxima de nieve en periodos de 8 días, hay que tratar los resultados del modelo para semejarlos a la observación. En primer lugar, convierto los mapas de SWE en presencia-ausencia, asumiendo como presencia de nieve valores del SWE superiores a 1 mm. Posteriormente, agrego estos mapas a 8 días, tomando el máximo de cada ventana temporal para reproducir la máxima cobertura de nieve en ese periodo.

## Calibración

### Algoritmo

Como algoritmo de optimización he utilizado la implementación de SCEUA (*Shuffled Complex Evolution-University of Arizona*) en la librería *spotpy* de Python.

Dentro de los parámetros a ajustar en SCEUA, he definido únicamente el número de iteraciones máximas en 1000. El resto de los parámetros los dejo según el ejemplo en la documentación de *spotpy*.

### Función objetivo

Puesto que el modelo es un clasificador binario de presencia o ausencia de nieve, he optado por utilizar como función objetivo el *f1*, la media armónica entre precisión y exhaustividad. Como método de agregación utilizo *micro*, que calcula la métrica global a partir del total de positivos verdaderos, falsos negativos y falsos positivos.

Utilizo la función *f1\_score* de Scikit-Learn para calcular el *f1* entre cada mapa observado y su simulación correspondiente. Con ello genero una serie temporal de *f1*, de la que calculo la media como el valor de rendimiento global.

### Parámetros calibrados

Para simplificar la calibración y reproducir exactamente el funcionamiento del modelo de nieve de TETIS, sólo incluyo tres parámetros en la calibración (, , ) y fijo en 0 el valor de . El rango de búsqueda y el valor inicial de cada uno de los tres parámetros calibrados se muestran en la Tabla 2.

Tabla 2. Parámetros calibrados, su rango de búsqueda, valor inicial en la calibración y valor óptimo tras la calibración.

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Parámetro | Mínimo | Máximo | Valor inicial | Valor óptimo |
|  | -2 | 6 | 0 | 4.8 |
|  | 0 | 6 | 3 | 4.71 |
|  | 0 | 10 | 5 | 8.82 |

Calibro un periodo de 5 años (oct-2005 a oct-2010) y valido en otros 5 años (2010-2015).

# Resultados

En la Figura 1 se muestra la evolución de la calibración del modelo de nieve. La calibración requirió 416 iteraciones, obteniendo un f1 óptimo de 0.878. Los parámetros óptimos son Tb 4.8 °C, DDF1 4.71 mm/°C y DDF2 8.82 mm/°C (Tabla 2).

La serie de f1 indica que los parámetros del modelo grado-día no son muy sensibles, al menos para esta función objetivo, puesto que apenas hay diferencia en rendimiento entre la mejor y la peor iteración. Esta idea queda reforzada en los gráficos de dispersión de la Figura 2, donde se observa que DDF1 y DDF2 dan un rendimiento óptimo para cualquier valor dentro de su rango de búsqueda.

Interfaz de usuario gráfica

Descripción generada automáticamente

Figura . Evolución de la calibración del modelo de nieve. Se representa el rendimiento (f1) y los valores de los tres parámetros calibrados para las iteraciones de la calibración. La cruz roja representa la parametrización óptima.

La Figura 3 muestra los resultados de la validación. Se muestran las series simuladas de equivalente agua-nieve y de fusión (valores medios en la cuenca), la comparativa de celdas nevadas entre la simulación y la observación, y el rendimiento en cada paso de la simulación.

En el gráfico de número de celdas nevadas se constata que el modelo tiene un ligero desfase temporal frente a la observación. El inicio y el final de la época invernal se retrasa en la simulación. Los momentos con mayor superficie cubierta por nieve se captan relativamente bien.

En base a la serie de f1 se ve que incluir el verano en el cálculo del rendimiento puede ser contraproducente, puesto que por lo general no hay nieve y se obtiene un rendimiento de 1 que puede enmascarar la sensibilidad de los parámetros durante el invierno.

Gráfico, Gráfico de dispersión

Descripción generada automáticamente

Figura 2. Diagramas de dispersión del valor adoptado por los parámetros frente al rendimiento asociado. Cada punto de la nube es una iteración en la calibración; el punto naranja es la parametrización óptima.

Diagrama

Descripción generada automáticamente

Figura . Resultados de la validación. De arriba abajo: equivalente agua-nieve, fusión de nieve, número de celdas nevadas y rendimiento expresado en f1.

La Figura muestra los mapas simulados y observado para el 9 de enero de 2011 como un ejemplo de los resultados. El mapa de la izquierda representa el equivalente agua-nieve en mm; este mapa se reclasifica en un mapa binario de cobertura de nieve (centro), que se compara frente a la observación (derecha) para obtener un f1 de 0.769.

Gráfico

Descripción generada automáticamente

Figura . Ejemplo de los mapas simulados de equivalente agua-nieve y cobertura de nieve (izquierda y centro) y la cobertura de nieve observada por MODIS (derecha) para un día de la simulación.

He generado vídeos para comparar visualmente la simulación con la observación (ver vídeo *SC\_simulado\_vs\_observado(calib).mp4*).

# Discusión

* El modelo parece poco sensible a la calibración. Cualquier parametrización da un rendimiento aceptable. Pudiera deberse a que el criterio de rendimiento escogido, f1, es poco representativo.
* Utilizar el criterio de rendimiento f1 no tiene en cuenta la estructura espacial de los datos. Es decir, considera todas las celdas como muestras independientes, sin tener en cuenta que el mapa de nieve tenga una estructura espacial adecuada o no. Hice pruebas con otros criterios de rendimiento como el SPAEF o el KGE espacial, pero estas funciones no son aplicables a datos binarios. No sé si se podría aplicar EOFs.
* La veracidad de los datos de MODIS es dudosa: aparecen manchas en verano, la coherencia entre los datos de Aqua y Terra es, a veces, escasa (ver vídeo *SC\_Terra\_Aqua.mp4*). El uso de datos de cobertura y no de equivalente de agua nieve hace que se pierda buena parte de la información de lo que el modelo simula. Lo ideal sería tener datos de equivalente agua-nieve o espesor, pero no los he encontrado.
* El modelo de nieve podría mejorar, utilizando alguna versión más desarrollada del método de grado-día como la de HBV. Sin embargo, no me parece lo más prioritario porque se ve que el modelo en sí es poco sensible.
* Se podría trabajar sobre los parámetros de SCEUA, pero visto que no requiere ni 1000 iteraciones para llegar a la solución óptima, no lo vi necesario.