

SOCIEDAD CHILENA DE INGENIERIA HIDRAULICA

XIX CONGRESO CHILENO DE HIDRAULICA

**APLICACIÓN DEL METODO DE DESCOMPOSICIÓN DE LA VARIANZA DE SOBOL
EN EL ANALISIS DE SENSIBILIDAD DE MODELOS DE SIMULACIÓN
HIDROLÓGICA CON COMPONENTE NIVAL**

**ALEXANDRA JACQUIN S.¹
CRISTIAN SÁNCHEZ U.²**

RESUMEN

En este trabajo se presenta la aplicación del método de Descomposición de la Varianza de Sobol en el análisis de sensibilidad de un modelo de simulación hidrológica con componente nival de uso extendido en el país. El método de Descomposición de la Varianza de Sobol permite evaluar en qué medida los distintos factores que determinan la respuesta de un modelo tienen influencia en sus variables de salida, tanto individualmente como incluyendo los efectos de interacciones con otros factores. En consecuencia, con el método DVS se obtienen estimaciones de sensibilidad válidas para cualquier tipo de modelo, lineal o no lineal, lo que representa una ventaja respecto a los métodos clásicos de análisis de sensibilidad. El análisis de sensibilidad desarrollado permitió determinar que la mitad de los parámetros del modelo estudiado no tenían gran influencia individual en la determinación de las variables de salida consideradas. Sin embargo, se comprobó que la mayoría de los parámetros sí resultaba influyente, si se consideraban los efectos de interacciones con otros parámetros.

¹Profesora, Carrera de Ingeniería Civil, Pontificia Universidad Católica de Valparaíso - mail: alexandra.jacquin@ucv.cl

²Ingeniero Civil, Departamento de Obras Civiles, Universidad Técnica Federico Santa María - mail: csanchezu@gmail.com

1. INTRODUCCIÓN

La estructura de un modelo usado para describir el funcionamiento de un sistema real está definida por una serie de ecuaciones y parámetros, que relacionan las variables de entrada con las variables de salida o respuesta del modelo. Dada una estructura, los factores que determinan la respuesta de un modelo incluyen tanto los valores asignados a sus parámetros como los valores que toman las variables de entrada. Un análisis de sensibilidad (AS) pretende evaluar en qué medida la variabilidad o incertidumbre en la respuesta de un modelo puede ser atribuida a cada uno de estos factores, estableciendo la importancia relativa de cada uno de ellos en la determinación de la respuesta. De esta forma, la aplicación de técnicas de AS permite detectar qué factores son los más sensibles y cuáles tienen importancia secundaria.

En un escenario real, con recursos económicos y tiempo disponible limitados, el modelador debiera preocuparse especialmente de determinar cuáles son los valores adecuados para los factores más importantes. Esto quiere decir que la recopilación de información de campo debiera centrarse en proveer observaciones confiables de las variables de entrada más relevantes. Asimismo, el esfuerzo invertido en el proceso de calibración del modelo debiera estar sobre todo orientado a encontrar valores adecuados para los parámetros de mayor influencia. En principio, si la aplicación de un método de AS apropiado para el tipo de modelo que se está estudiando señalara factores que no son realmente importantes en la determinación de la respuesta, sus valores podrían ser fijados a priori, eligiendo valores adecuados dentro de sus rangos esperados de variación.

Los métodos de AS disponibles en la literatura pueden ser agrupados en dos categorías: métodos de análisis de sensibilidad local (ASL) y métodos de análisis de sensibilidad global (ASG). Los métodos de ASL son métodos cuantitativos que buscan estimar la sensibilidad de la respuesta a pequeñas variaciones en los valores de los factores de entrada, respecto a valores nominales. La principal desventaja de estos métodos es que las estimaciones locales de sensibilidad no son informativas respecto al impacto que tendrían eventuales fluctuaciones de los factores del modelo a través de sus rangos factibles. Por otra parte, dado que sólo se varía un factor a la vez, los métodos de ASL no son útiles en casos en que las interacciones entre factores sean relevantes (Saltelli et al., 2004; Pappenberger et al., 2008). En consecuencia, la información provista por métodos de ASL no es completa si el modelo no es lineal.

Los métodos de ASG buscan estimar la importancia de cada factor de entrada con respecto a la variabilidad la respuesta del modelo, cuantificando el impacto de variaciones en el valor de cada factor a través de su rango factible (Saltelli, 2004). Adicionalmente, algunos métodos de ASG son capaces de cuantificar los efectos de interacciones con otros parámetros. Como resultado, los métodos de ASG tienen la enorme ventaja de proveer estimaciones de sensibilidad que son válidas para todo tipo de modelo (lineal o no lineal).

Este estudio presenta la aplicación del método Descomposición de la Varianza de Sobol (Sobol, 1993) en el análisis de sensibilidad de un modelo de simulación hidrológica con componente nival. Siendo un método ASG, la Descomposición de la Varianza de Sobol (DVS) permite determinar la importancia de los factores de entrada en la respuesta del modelo a lo largo de todo el rango de variación de cada uno de ellos. Además de evaluar el efecto de variaciones

unidimensionales en los valores de los factores, el método DVS es capaz de detectar sensibilidades provenientes de interacciones.

2. MÉTODO DE DESCOMPOSICIÓN DE LA VARIANZA DE SOBOL

En el método DVS, la variabilidad o incertidumbre en cada variable de salida de interés Y se mide en términos de su varianza $V[Y]$. Esta varianza es calculada mediante la exploración de todo el espacio factible del conjunto de factores θ_p . Sobol (1993) propuso la siguiente descomposición para la varianza $V[Y]$

$$V[Y] = \sum_p V_p + \sum_p \sum_{q>p} V_{pq} + \dots + V_{1,2,\dots,P}, \quad (1)$$

que es válida en casos en que los factores de entrada θ_p no están correlacionados. En la ecuación (1), los términos V_p representan la parte de la varianza $V[Y]$ que puede ser atribuida a cambios individuales de cada factor θ_p . Los términos de orden superior indican el efecto de cambios simultáneos de dos o más parámetros, excluyendo el efecto individual de cada uno.

El efecto de primer orden de θ_p en Y , definido por

$$S_p = \frac{V_p}{V[Y]}, \quad (2)$$

mide el efecto individual del factor θ_p en la variabilidad de Y . El hecho que el efecto de primer orden de un factor θ_p sea grande es condición suficiente para demostrar que θ_p es relevante. En el otro extremo, efectos de primer orden pequeños indican que los factores θ_p correspondientes no son influyentes por sí mismos. Sin embargo, no queda descartada la posibilidad de que éstos sean importantes en términos globales, es decir, incluyendo los efectos de interacciones con otros factores.

Por otra parte, el efecto total de θ_p en Y , definido por (Homma y Saltelli, 1996)

$$S_{Tp} = \frac{V_p}{V[Y]} + \frac{\sum_{q \neq p} V_{pq}}{V[Y]} + \frac{\sum_{q \neq p} \sum_{r > q} V_{pqr}}{V[Y]} + \dots, \quad (3)$$

mide la importancia del factor θ_p en términos globales, incluyendo tanto su efecto individual como el efecto de interacciones de θ_p con otros factores. Efectos totales muy pequeños indican que los factores correspondientes no son muy importantes para determinar el valor de la variable Y . Por el contrario, efectos totales grandes indican que los factores correspondientes son de gran relevancia. El caso en que un factor θ_p tiene un gran efecto total acompañado de un efecto de primer orden muy pequeño se produce cuando el factor θ_p afecta la respuesta del modelo principalmente a través de interacciones.

3. APLICACION

El método DVS fue aplicado en el análisis de sensibilidad de los parámetros del modelo CICA (CICA Ingenieros Consultores, 1982), modelo de simulación hidrológica con componente nival de uso extendido en Chile (e.g. 4C Ingenieros Ltda., 2000; Water Management Ltda., 2001; Arcadis Geotécnica, 2007). La versión del modelo CICA usada en este estudio corresponde a la desarrollada en el trabajo de Kamann (1998), que fue implementada en MATLAB® para su uso en el presente estudio. La información hidrológica necesaria para la operación del modelo CICA a nivel mensual incluye precipitación, número de días de lluvia, caudal, evaporación, temperatura, humedad relativa, velocidad media del viento y nubosidad. Para calibrar el modelo, se requiere además una estadística de caudales concurrente con las anteriores.

El modelo divide la cuenca en cinco bandas de altura, donde la 5ª banda corresponde a los glaciares de la cuenca. Dentro de cada banda, se realizan balances hídricos independientes, usando valores uniformes para las variables hidrometeorológicas. En particular, la precipitación en cada banda i se calcula como el producto entre la precipitación en la estación base y un factor de precipitación FP_i . La escorrentía superficial total se obtiene de la suma de la escorrentía superficial generada en las distintas bandas, suponiendo la existencia de un embalse superficial lineal en cada una de ellas. La percolación profunda total corresponde a la suma de la percolación profunda de las bandas individuales. Suponiendo la existencia de un embalse subterráneo lineal, la generación de escorrentía subterránea para la cuenca completa es calculada a partir de la estimación obtenida para la percolación profunda total. El modelo considera un balance hídrico diferente en la banda correspondiente a glaciares, considerados como fuentes inagotables de agua para efectos de la simulación. De esta forma, el modelo consta de 20 parámetros independientes, cuyas interpretaciones se indican en la tabla 1.

Tabla1. Parámetros del modelo.

Parámetro	Interpretación
CEN	Coefficiente de escorrentía de la nieve estacional
BB	Coefficiente de evapostublimación de la nieve
B	Coefficiente de evapotranspiración potencial
HNTOT	Equivalente en agua de la nieve a partir del cual la banda se cubre totalmente
CEP	Coefficiente de escorrentía de lluvia
CKS	Constante de tiempo de los embalses superficiales
CK	Constante de tiempo del embalse subterráneo
FKP	Coefficiente de exposición a la radiación
FK	Coefficiente de exposición al viento
GRT	Gradiente térmico de la atmósfera
CLOUDK	Coefficiente del tipo de nubes para el cálculo de la radiación solar
DRMIN	Tasa de derretimiento mínimo de la fórmula de derretimiento
COFDRT	Factor de corrección para la fórmula de derretimiento
FP1	Factor de precipitación en la banda 1
FP2	Factor de precipitación en la banda 2
FP3	Factor de precipitación en la banda 3
FP4	Factor de precipitación en la banda 4
FP5	Factor de precipitación en la banda 5
TU	Temperatura media mensual umbral que distingue entre precipitación líquida y sólida
FKPG	Coefficiente de exposición a la radiación en los glaciares

La cuenca de estudio corresponde a la del río Olivares antes de la junta con el río Colorado, con una superficie de 531[Km²]. Aproximadamente el 13% de superficie de la cuenca corresponde a glaciares. La sección de salida está ubicada a 1500[m.s.n.m.], mientras que las cumbres más altas alcanzan los 6000[m.s.n.m.]. Los datos disponibles se dividieron en dos períodos consecutivos. El período I constaba de 21 años hidrológicos (1962-1982) y el período II, 8 años (1983-1990).

La sensibilidad de los parámetros de modelo se evaluó respecto al error cuadrático medio (ECM), el error volumétrico relativo (EVR) y la suma ponderada de la altura equivalente en agua de la nieve acumulada en cada banda (NAB), todas calculadas en forma independiente para el período I y el período II. La sensibilidad de los parámetros FP1, FP2, FP3 y FP4 fue estimada en términos de índices de sensibilidad grupal. En este caso, el efecto de primer orden del grupo de parámetros FP1 a FP4 indica la importancia de los parámetros como grupo, excluyendo interacciones con parámetros fuera de éste. El efecto total del grupo de parámetros FP1 a FP4 incluye tanto el efecto de primer orden del grupo como el efecto de interacciones con otros parámetros.

4. RESULTADOS

La tabla 2 muestra los efectos de primer orden calculados para cada una de las variables de salida consideradas. Asimismo, la tabla 3 muestra los efectos totales de los parámetros del modelo en estas variables. Las celdas sombreadas en color gris oscuro indican parámetros muy importantes (valores mayores que 15%), mientras que celdas sombreadas en gris claro indican importancia moderada (valores entre 5% y 15%). Índices de sensibilidad menores que 5% son considerados como poco significativos

Según muestra la tabla 2, sólo 10 parámetros tienen sensibilidad al menos moderada, desde el punto de vista de sus efectos de primer orden, con respecto a una o más de las variables de salida consideradas. Además, sólo el parámetro FKP, el grupo FP1 a FP4 y el parámetro FKPG son muy importantes, en forma individual, para determinar los valores de al menos una de las variables de salida estudiadas. Estos parámetros son aquellos que debieran ser ajustados con mayor cuidado en un ejercicio de calibración que usara alguna de estas cantidades en su función objetivo.

Por otra parte, la tabla 3 muestra que sólo cinco parámetros del modelo obtienen efectos totales no significativos, a saber, B, HNTOT, CK, GRT y TU. En casos en que el esfuerzo computacional involucrado en el proceso de calibración fuera una limitante, los valores de estos parámetros podrían ser fijados, en principio, en algún valor conveniente de su rango de variación. Sin embargo, este resultado no debe ser considerado como general, pues sólo es válido para la cuenca en estudio y los datos hidrológicos considerados.

La comparación visual de las tablas 2 y 3 revela que, en el caso de varios parámetros, los efectos totales son significativamente mayores que los efectos de primer orden, con respecto a las variables ECM y NAB. Este resultado indica que estos parámetros están involucrados en interacciones que tienen un impacto significativo en el cálculo del error cuadrático medio y en la altura de nieve acumulada. Por ejemplo, el parámetros CKS (constante de tiempo de los embalses superficiales) obtuvo efectos de primer orden muy pequeños con respecto a todas las variables de salida consideradas, pero sus efectos de orden total con respecto al error cuadrático medio no son

despreciables. Por el contrario, no se aprecian diferencias importantes entre los efectos totales y los efectos de primer orden de los parámetros del modelo con respecto a la variable EVR.

Tabla 2. Efectos de primer orden de los parámetros del modelo.

Parámetro	ECM período I	ECM período II	EVR período I	EVR período II	NAB período I	NAB período II
CEN	0,1472	0,0609	0,0594	0,000465	2,35E-07	0,00000246
BB	0,0133	0,008	0,0465	0,0516	0,0679	0,0726
B	6,20E-04	0,0016	0,0029	0,0049	3,22E-07	3,77E-07
HNTOT	0,0000768	1,05E-04	0,001	0,000781	0,0000258	1,46E-05
CEP	0,0985	0,0401	0,011	0,002	7,09E-07	0,00000309
CKS	0,005	0,0422	3,76E-05	1,20E-03	1,03E-07	0,000001
CK	0,0044	0,0015	0,0111	0,0034	7,59E-07	1,87E-07
FKP	0,0088	0,000367	0,0481	0,0232	0,2093	0,163
FK	0,0333	0,0516	0,1136	0,1256	0,114	0,1068
GRT	8,21E-06	1,61E-05	1,29E-06	2,23E-05	1,24E-04	3,02E-05
CLOUDK	0,0214	0,0216	0,0553	0,045	0,0458	0,0391
DRMIN	0,00017	0,000632	0,0111	0,0062	0,0307	0,0278
COFDRT	0,0844	0,0347	0,1131	0,0831	0,0713	0,0564
FP1 a FP4	0,1825	0,2224	0,2587	0,3393	0,0896	0,069
FP5	0,00000212	7,46E-07	6,88E-08	6,55E-07	0,058	0,0496
TU	0,0099	0,0097	0,0000233	0,0006	0,0000139	2,85E-07
FKPG	0,092	0,1044	0,2264	0,2644	0,0314	0,0393

Tabla 3. Efectos totales de los parámetros del modelo.

Parámetro	ECM período I	ECM período II	EVR período I	EVR período II	NAB período I	NAB período II
CEN	0,404069	0,292439	0,076402	0,00546	0,000693	0,000894
BB	0,036472	0,037139	0,048843	0,055182	0,144133	0,183335
B	3,23E-03	0,00807	0,005104	0,007661	6,01E-04	0,000779
HNTOT	0,001182	1,42E-03	0,001884	0,001719	0,000668	8,40E-04
CEP	0,158123	0,065463	0,013777	0,004149	0,00071	0,001009
CKS	0,110055	0,11336	7,79E-04	2,56E-03	0,00059	0,000805
CK	0,010685	0,006634	0,023752	0,006292	0,000642	0,000914
FKP	0,091307	0,067645	0,064057	0,048882	0,328007	0,3275
FK	0,146097	0,213091	0,149052	0,161794	0,220981	0,275039
GRT	1,14E-03	1,54E-03	6,60E-04	8,50E-04	1,21E-03	1,66E-03
CLOUDK	0,072039	0,096519	0,070426	0,058983	0,084385	0,105287
DRMIN	0,0037	0,005105	0,014485	0,009274	0,067821	0,075358
COFDRT	0,221947	0,177416	0,144387	0,121346	0,120067	0,138373
FP1 a FP4	0,343293	0,453523	0,288869	0,375808	0,195111	0,204722
FP5	0,001323	0,001429	0,000627	0,000702	0,105055	0,130519
TU	0,036257	0,033805	0,002349	0,002552	0,000786	0,000983
FKPG	0,234455	0,355463	0,275154	0,310698	0,057108	0,083925

5. CONCLUSIONES

Este estudio mostró la aplicación del método DVS en el análisis de sensibilidad del modelo de simulación hidrológica CICA. La principal ventaja de este método consiste en permitir la estimación de la sensibilidad de los parámetros o variables de entrada de un modelo en términos globales, cuantificando su influencia como factores independientes y el efecto que podrían tener las interacciones en que estos factores están involucrados.

El análisis de sensibilidad desarrollado permitió identificar los parámetros de mayor influencia y aquellos de importancia secundaria. Se encontró que 10 de los 20 parámetros del modelo no tenían gran influencia individual en la determinación de las variables de salida consideradas. Sin embargo, se comprobó que la mayor parte de ellos sí resultaba influyente, si se consideraban los efectos de interacciones con otros parámetros. En varios casos, se encontró que los efectos totales del método DVS resultaron bastante mayores que los efectos de primer orden, con respecto a las variables ECM y NAB, lo que indica que existe un alto grado de interacción en la estructura del modelo.

AGRADECIMIENTOS

Esta investigación fue financiada por el proyecto FONDECYT 11070130. Los autores agradecen también a 4C Ingenieros Ltda., por proveer información respecto al modelo y los datos hidrológicos usados en este estudio.

REFERENCIAS

- 4C Ingenieros Ltda. 2000. Estudio Revisión Hidrología Río Mapocho y Esteros Ortiga y Riecillos. Compañía Minera Disputada Las Condes.
- Arcadis Geotécnica 2007. Ingeniería Conceptual Solución Ambiental ARD Proyecto Nueva Andina, CODELCO.
- CICA Ingenieros Consultores 1982. Estudio Integral de Riego de los Valles Aconcagua, Putaendo, Ligua y Petorca. Comisión Nacional de Riego.
- Homma, T. y Saltelli, A. 1996. Importance measures in global sensitivity analysis of model output. *Reliability Engineering and System Safety*, 52, 1-17, 1996.
- Kamann Chacana, P.G. 1998. Pronóstico estacional de deshielos en base a simulación. Memoria para Optar al Título de Ingeniero Civil, Universidad Técnica Federico Santa María, Valparaíso, Chile.
- Pappenberger, F., Beven, K.J., Ratto, M. y Matgen, P. 2008. Multi-method global sensitivity analysis of flood inundation models. *Advances in Water Resources*, 31(1), 1-14.
- Saltelli, A., Tarantola, S., Campolongo, F. y Ratto, M. 2004. *Sensitivity analysis in practice: A guide to assessing scientific models*. John Wiley & Sons, Chichester, 219 pp.
- Sobol, I.M. 1993. Sensitivity analysis for non-linear mathematical models. *Mathematical Modelling and Computational Experiments*, 1(4), 407-414.
- Water Management Ltda. 2001. Evaluación Recursos Hídricos Salar Punta Negra. Compañía Minera Escondida.