

SOCIEDAD CHILENA DE INGENIERÍA HIDRÁULICA
XXV CONGRESO CHILENO DE INGENIERÍA HIDRÁULICA

**ASIMILACIÓN DE DATOS APLICADA A ESTIMACIONES DE BALANCE
HÍDRICO EN CUENCAS CON APORTE NIVAL DE CHILE CENTRAL**

EDUARDO MUÑOZ CASTRO¹
PABLO A. MENDOZA²
NICOLÁS VÁSQUEZ³
GONZALO CORTÉS⁴
XIMENA VARGAS⁵

RESUMEN

Este estudio evalúa el beneficio de implementar asimilación de datos en estimaciones de flujos y estados hidrológicos en Chile continental. Para ello, se implementó el Filtro de Kalman de Conjuntos (EnKF) acoplado al modelo hidrológico *Variable Infiltration Capacity* (VIC), configurado a una resolución horizontal de 0,05° lat-lon en 12 cuencas ubicadas a lo largo de un gradiente hidroclimático en Chile central. Los experimentos de asimilación perturban simultáneamente las forzantes del modelo y estados iniciales de humedad de suelo y equivalente de nieve en agua, generándose cuatro configuraciones diferentes. Los resultados obtenidos muestran que i) asimilar sólo caudal es insuficiente para asegurar la disminución en la dispersión de otros flujos y estados del sistema, ii) asimilar información satelital, como cobertura de nieve o equivalente de nieve en agua, restringe la variación de la variable asociada y condiciona la dinámica de otros procesos simulados, y iii) los impactos de asimilar datos en estimaciones de balance hidrológico dependen de la cuenca de estudio. Finalmente, se concluye que la configuración de asimilación impacta en el balance hídrico a la escala de cuenca.

¹ Investigador, Departamento de Ingeniería Civil, U. de Chile - email: eduardo.munoz@uchile.cl

² Profesor Asistente, Departamento de Ingeniería Civil, U. de Chile – email: pamendoz@uchile.cl

³ Investigador, Departamento de Ingeniería Civil, U. de Chile – email: nicolas.vasquez.pl@uchile.cl

⁴ Cofund., KLIK, Mr. Scoot, Red Energy, Huerti, The Fresh Box – email: gonzalocortessoruco@gmail.com

⁵ Profesora Asociada, Departamento de Ingeniería Civil, U. de Chile – email: xvargas@uchile.cl

1. INTRODUCCIÓN

Los modelos hidrológicos son herramientas que sirven de apoyo para el estudio del ciclo del agua, permitiendo la estimación de flujos y variables de estado. Sin embargo, los modelos tienen asociadas diferentes fuentes de incertidumbre, que impactan directamente en su desempeño (e.g., Döll et al., 2016; Müller Schmied et al., 2014). Así, para cuantificar y reducir la incertidumbre en estimaciones hidrológicas se requiere una revisión rigurosa de las estimaciones del modelo y una mayor cantidad y calidad de datos hidrológicos (e.g., Clark et al., 2021; Klemeš, 1986; Refsgaard et al., 1996).

Pese a que los modelos hidrológicos se encuentran en constante actualización, incorporando mejoras en la representación de los procesos hidrológicos (e.g., Gnann et al., 2021), interacción con la atmósfera (e.g., Renner et al., 2020), entre otros (e.g., Clark et al., 2015; Hamman et al., 2018), siguen siendo representaciones imperfectas de la realidad. Sumado a lo anterior, existe una escasez de observaciones hidrológicas in-situ en Chile (Alvarez-Garreton et al., 2018; Vásquez et al., 2021) que, si bien se ha logrado sortear con los avances en percepción remota (e.g., Alvarez-Garreton et al., 2018; Zambrano-Bigiarini et al., 2017), continua siendo una limitación. La inclusión de distintas fuentes de información en la modelación hidrológica se puede realizar mediante la asimilación de datos (e.g., Liu and Gupta, 2007; Reichle, 2008), que busca mejorar la simulación hidrológica mediante actualizaciones de flujos y variables de estado en función de la disponibilidad de observaciones (e.g., Pauwels and De Lannoy, 2009).

Durante la última década, varios autores han explorado la posibilidad de asimilar observaciones en modelos hidrológicos con diferentes propósitos (e.g., Girotto et al., 2016; Zhang et al., 2017). No obstante, el impacto de la asimilación de datos en la partición de precipitación (P) entre evapotranspiración (ET) y escorrentía (Q) no se ha estudiado profundamente. En el caso de Chile no se ha estimado el impacto de asimilación de datos en el balance hídrico a escala de cuenca.

En base a lo anterior, se definen las siguientes interrogantes: (i) ¿Cuál es el impacto de la asimilación de datos en el balance hidrológico de una cuenca?, y (ii) ¿Cuál es el beneficio de asimilar información satelital versus observaciones *in situ*? La metodología considera la simulación hidrológica mediante el modelo *Variable Infiltration Capacity* (VIC; Liang et al., 1996, 1994), el cual se configura para 12 cuencas andinas de Chile Central. Las variables que se asimilan en el modelo corresponden a escorrentía media diaria registradas en los controles fluviométricos de la red hidrométrica nacional, el reanálisis de equivalente de nieve en agua desarrollado por Cortés y Margulis (2017), y el productos de cobertura de nieve de MODIS Terra y Aqua (Hall et al., 2002). En consecuencia, el presente estudio busca explorar potenciales mejoras en la estimación del balance hídrico mediante asimilación de datos.

2. DOMINIO DE ESTUDIO, OBSERVACIONES Y BASES DE DATOS

Tomando en cuenta el nivel de intervención humana en las aguas superficiales y la extensión y continuidad de los registros fluviométricos disponibles, se seleccionan 12 cuencas ubicadas Figura 1 entre las regiones de Coquimbo y Maule. Las cuencas se agrupan según su respuesta

hidrológica (panel A en Figura 1). En términos hidroclimáticos (Tabla 1), las cuencas presentan un marcado gradiente latitudinal, con precipitaciones y escorrentía (índice de aridez) que aumentan (disminuye) a medida que se avanza latitudinalmente. Por ejemplo, en las cuencas ubicadas en la zona B (Figura 1) predomina el clima de tundra y semiárido de lluvia invernal, con montos anuales promedio en torno a los 280 mm/año, mientras que, a medida que se avanza latitudinalmente (zonas C y D; Figura 1) cambia a mediterráneo frío y/o de altura con lluvia invernal, con montos de precipitación de hasta 1.928 mm/año.

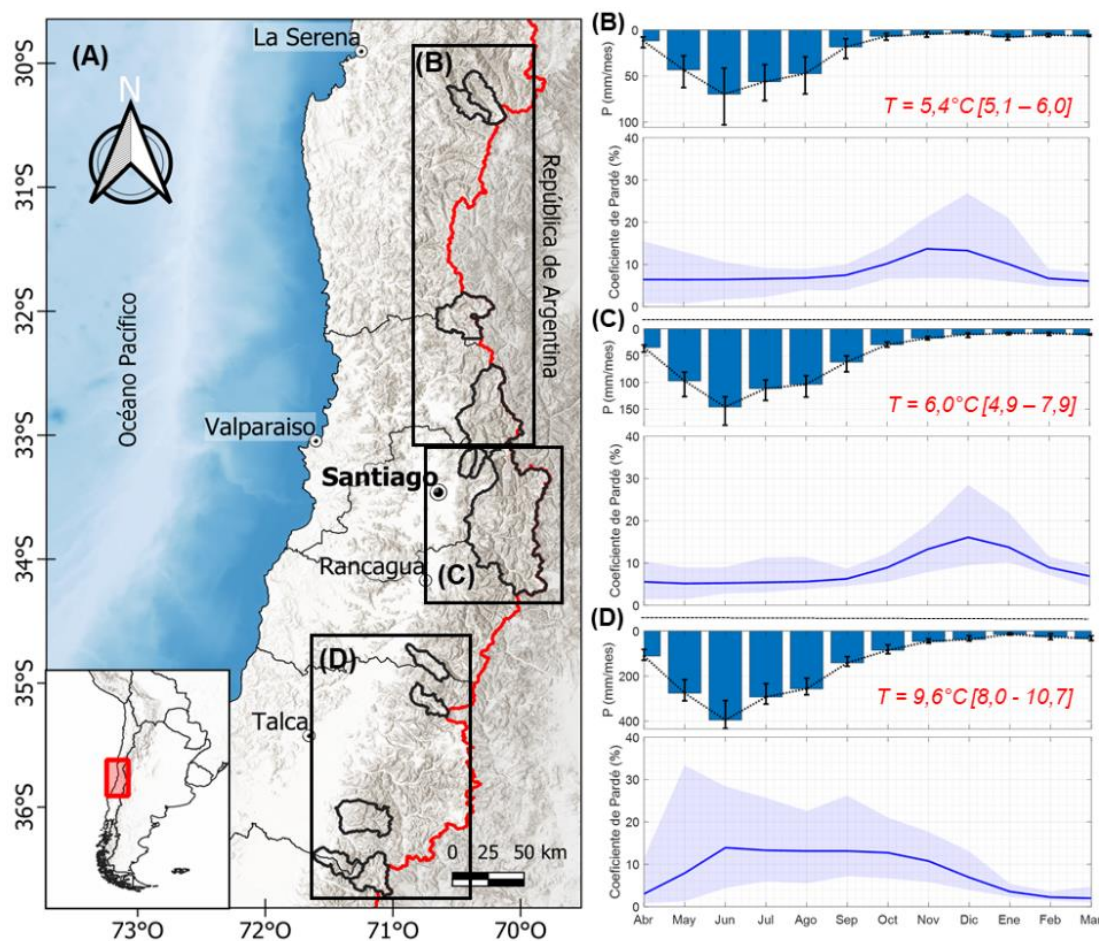


Figura 1: Dominio de estudio y patrones hidroclimáticos estacionales promedio en el periodo 1985-2015 agregados por subdominio espacial.

Para la caracterización del dominio de estudio, la configuración del modelo y la asimilación de se utilizan cuatro bases de datos:

- **Registros fluiométricos DGA.** Registros de caudal reportados por la Dirección General de Aguas (DGA), obtenidos a partir de curvas de descarga ajustadas en cada estación de la red de monitoreo nacional. Son de carácter público y pueden ser adquiridos directamente a través de la plataforma web de la DGA (<https://dga.mop.gob.cl/>).
- **Cobertura de nieve fraccional MODIS.** Producto de fracción de cobertura nival (SC, SCA o fSCA) diaria a resolución horizontal de 500 metros de los satélites MODIS Terra

y Aqua (Hall et al., 2002), disponibles para su descarga libre en la base de datos del *National Snow and Ice Data Center* (NSIDC; <https://nsidc.org/>).

- **Producto de SWE.** Producto de equivalente de nieve en agua (SWE) desarrollado por Cortés y Margulis (2017), disponible en la base de datos CAMELS-CL (Alvarez-Garreton et al., 2018) a la escala de cuenca y a resolución temporal diaria. Se utiliza como referencia observacional para evaluar el valor que aportaría la disponibilidad de datos de SWE.
- **Delimitaciones de cuencas.** Se obtienen los contornos de la red de drenaje delimitada por los controles fluviométricos de la DGA, los cuales se adquieren a través del explorador de cuencas CAMELS-CL (<https://camels.cr2.cl/>; Alvarez-Garreton et al., 2018).

Las forzantes del modelo VIC consideran series tri-horarias de precipitación, velocidad del viento, temperatura media y humedad relativa. Dichas variables se derivan de los siguientes productos:

- **Producto CR2MET.** Para obtener los valores precipitación y temperatura, se utiliza como base el producto grillado CR2Met versión 2.0 (DGA, 2019a), el cual incluye precipitación y temperaturas extremas (i.e., máxima y mínima) a resolución temporal diaria y espacialmente distribuida a 0,05° latitud-longitud para Chile continental.
- **Reanálisis ERA5.** Para obtener la distribución intradiaria de la precipitación y temperatura, así como la magnitud del viento y estimar la humedad relativa, se utilizan los datos horarios de precipitación, temperatura a 2 m, velocidad del viento zonal y meridional a 10 m, y temperatura en el punto de rocío a 2 m recuperados del reanálisis atmosférico ERA5 (C3S, 2017), los cuales son agregados a escala tri-horaria.

Tabla 1: Nombre, código BNA, y características hidroclimáticas promedio en periodo 1985-2015 de las 12 cuencas de estudio. Se incluyen los valores anuales promedio de temperatura media (T), precipitación (P), índice de aridez (IA), y escorrentía (Q).

Subdominio en Fig. 1	Nombre de la cuenca	BNA	T (°C)	P (mm/año)	IA (-)	Q (mm/año)
(B)	Estero Derecho En Alcohuz	4311001	5,1	213	4,72	84
(B)	Río Cochiguaz En El Peñón	4313001	5,4	175	6,10	117
(B)	Río Choapa En Cuncumén	4703002	6,0	326	3,38	260
(B)	Río Aconcagua En Chacabuquito	5410002	5,2	408,7	2,30	459
(C)	Río Maipo En El Manzano	5710001	4,9	796	1,08	683
(C)	Estero Yerba Loca Antes Junta San Francisco	5721001	5,2	541	1,18	195
(C)	Estero Arrayan En La Montosa	5722001	7,9	583	1,74	210
(D)	Río Claro En El Valle	6027001	10,7	1.319	0,88	762
(D)	Río Claro En Los Queñes	7103001	8,9	1.623	0,67	1.481
(D)	Río Perquilauquén en San Manuel	7330001	10,3	1.928	0,63	1.782
(D)	Río Achibueno En La Recova	7354002	10,0	1.877	0,64	1.351
(D)	Río Saucos Antes Junta Con Ñuble	8104001	8,0	1.813	0,65	1.560

3. METODOLOGÍA

3.1 Modelo hidrológico *Variable Infiltration Capacity* (VIC)

El modelo hidrológico *Variable Infiltration Capacity* (VIC) es un modelo de macro-escala, motivación física y semi-distribuido, que permite simular flujos de energía (e.g., radiación solar, intercambio de calor sensible) y agua (e.g., transpiración, escorrentía). Requiere de coberturas y propiedades de suelo, las cuales son definidas según DGA (2019b) a través de la información de textura de suelos desarrollada por el Centro de Información de Recursos Naturales (CIREN) e información derivada de productos satelitales como el índice de vegetación mejorado (EVI) y el índice de área foliar (LAI) de MOD13A2 y MOD15A2.

La calibración considera parámetros homogéneos a nivel de cuenca, salvo para la conductividad hidráulica y la profundidad de la capa de suelo, donde se recurre a método de regularización (Pokhrel and Gupta, 2010). A partir de un análisis de sensibilidad (DGA, 2019a) se incluyen 14 parámetros en el proceso de calibración, el cual se realiza mediante el algoritmo “*Shuffled Complex Evolution*” (Duan and Gupta, 1992). La función objetivo corresponde a la eficiencia de Kling-Gupta (KGE, Gupta et al., 2009) la cual se estima a partir de caudales medios diarios simulados en el periodo 1985-2005. Asimismo, la simulación considera un periodo de calentamiento entre los años hidrológicos (abril-marzo) 1979-1985 y de evaluación entre 2005-2015. El detalle de los parámetros calibrados y su valor se detalla en Muñoz-Castro (2021).

3.2 Esquema de asimilación de datos – Filtro de Kalman de Conjuntos (EnKF)

La asimilación de datos o *data assimilation* (DA) busca mejorar la estimación de una variable mediante la combinación de modelos y observaciones. El problema se puede expresar, en su forma más general, según el teorema de Bayes aplicado al problema de estimación de los estados de un modelo, condicionado por las observaciones disponibles:

$$p(x_{1:t}|y_{1:t}, u_{1:t}, \theta^f, \theta^h) \propto p_h(y_{1:t}|x_{1:t}, \theta^h) p_f(x_{1:t}|u_{1:t}, \theta^f) \quad (1)$$

En la ecuación (1), p_h y p_f corresponden a la función de densidad de probabilidad del espacio de estados y observaciones respectivamente. Los filtros corresponden a uno de los enfoques para abordar el problema de asimilación de datos y, por lo tanto, una posible solución de la ecuación (1). Éstos realizan una aproximación Markoviana del sistema dinámico modelado, la cual se expresa explícitamente a través de la probabilidad a priori del estado “ p_f ” (ec.(2)).

$$p(x_{1:t}|y_{1:t}, u_{1:t}, \theta^f, \theta^h) \propto p_h(y_t|x_t, \theta^h) p_f(x_t|x_{t-1}, u_t, \theta^f, y_{1:(t-1)}) \quad (2)$$

Considerando la idoneidad para trabajar con modelos geofísicos no-lineales (e.g., Ibaceta et al., 2020; Rasmussen et al., 2015), se implementa el Filtro de Kalman de Conjuntos o *Ensemble Kalman Filter* (EnKF; Evensen et al., 1994). El esquema de asimilación se divide en dos pasos: 1) predicción del modelo, y 2) actualización de estados. El primero consta de la generación del vector de estados a priori a partir de la función de transferencia espacio-estado “ f ”, el error del modelo para estimar el estado (ω), y el vector de observaciones

simuladas, generado a partir del operador de observaciones “ h ”. Así, las ecuaciones (3) y (4) muestran la forma general de la ecuación de estado a priori y de medida, respectivamente:

$$\hat{x}_{t,i}^- = f(\hat{x}_{t-1,i}^+, u_{t,i}, \theta) + \omega_{t,i} \quad (3)$$

$$\hat{y}_{t,i} = h(\hat{x}_{t,i}^-, \theta) \quad (4)$$

Luego, dada una observación “ y_t ” y su error asociado “ $\varepsilon_{t,i}$ ” (ec. (5)), se actualiza la variable de interés según a partir de la varianza entre los múltiples miembros del conjunto (ecuación 6), contenida en la ganancia de Kalman “ K ”. Dicha ganancia se calcula según la ecuación (7), donde la notación del tipo “ C_{ab} ” hace referencia a la matriz de covarianza entre la variable “ a ” y “ b ”, y “ Σ_0 ” corresponde a la matriz de covarianza del error de la observación, la cual se supone diagonal (i.e., no hay correlación entre los errores de las observaciones).

$$y_{t,i} = y_t + \varepsilon_{t,i} \quad (5)$$

$$\hat{x}_{t,i}^+ = \hat{x}_{t,i}^- + K(y_{t,i} - \hat{y}_{t,i}) \quad (6)$$

$$K = C_{xy}(C_{yy} + \Sigma_0)^{-1} \quad (7)$$

3.3 Error del modelo y observación

Para estimar el error del modelo y las observaciones, se realizan pruebas con distintos tipos de perturbación y magnitudes obtenidas de la literatura (e.g., Clark et al., 2008; Li et al., 2014). Así, se definen los errores presentados en la Tabla 2, los cuales son aplicados por igual a las 12 cuencas de estudio.

Tabla 2: Resumen de errores definidos para la asimilación de datos con EnKF.

Fuente de Incertidumbre	Variable	Notación	Valor	Perturbación	Tipo*
Forzante	Precipitación	ε_P	0,50	Multiplicativa	$U(-\varepsilon_P, \varepsilon_P)$
	Velocidad del viento	ε_u	0,20	Multiplicativa	$U(-\varepsilon_u, \varepsilon_u)$
	Temperatura	ε_T	0,20	Aditiva	$U(-\varepsilon_T, \varepsilon_T)$
	Humedad relativa	ε_{HR}	0,20	Multiplicativa	$U(-\varepsilon_{HR}, \varepsilon_{HR})$
Condición Inicial	Humedad de suelo capa 1 (SM1)	ε_{SM1}	2,00	Aditiva	$U(-\varepsilon_{SM1} \Delta S , \varepsilon_{SM1} \Delta S)$
	Humedad de suelo capa 2 (SM2)	ε_{SM2}	2,00	Aditiva	$U(-\varepsilon_{SM2} \Delta S , \varepsilon_{SM2} \Delta S)$
	Humedad de suelo capa 3 (SM3)	ε_{SM3}	2,00	Aditiva	$U(-\varepsilon_{SM3} \Delta S , \varepsilon_{SM3} \Delta S)$
	Equivalente de nieve en agua (SWE)	ε_{SWE}	2,00	Aditiva	$U(-\varepsilon_{SWE} \Delta S , \varepsilon_{SWE} \Delta S)$
Flujo observado	Escorrentía	ε_Q	0,15	Aditiva	$N(0, [\varepsilon_Q Q]^2)$

*U() y N() denotan una distribución uniforme y normal, respectivamente. $|\Delta S|$ corresponde a la variación absoluta de la condición inicial del estado S (i.e., SM1, SM2, SM3 o SWE según sea el caso) entre el tiempo t y t-1.

3.4 Experimentos numéricos

El tamaño del conjunto en este estudio considera 50 miembros. Asimismo, se definió el año hidrológico 2000/2001 como periodo de prueba de asimilación. En la Tabla 3 se resumen los experimentos numéricos propuestos, los cuales se configuran perturbando todas las forzantes del modelo y perturbando/actualizando, en el Caso N°1 (C1) la humedad inicial de las capas de suelo (SMX) y en el Caso N°2 (C2) agregando el equivalente de nieve en agua (SWE) y cobertura de nieve (SC) iniciales. A su vez, se definen tres tipos de experimentos según las observaciones asimiladas. Cabe mencionar que los estados perturbados/actualizados se eligen considerando la influencia que tienen en la dinámica de las cuencas estudiadas.

Tabla 3: Resumen de experimentos de asimilación propuestos para el estudio.

ID	Color	Variable perturbada										Observación asimilada			Estado actualizado			
		Forzante					Estado inicial											
		P	u	T	HR	SM1	SM2	SM3	SWE	SC	Q	SWE	SC	SM1	SM2	SM3	SWE	
C1.E1		X	X	X	X	X	X	X			X			X	X	X		
C2.E1		X	X	X	X	X	X	X	X	X	X			X	X	X	X	
C2.E2		X	X	X	X	X	X	X	X	X		X	X	X	X	X	X	
C2.E3		X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	

3.5 Evaluación de la asimilación de datos

Las simulaciones sin asimilación (*Open-Loop*; OL) y con asimilación (DA) se comparan mediante error cuadrático medio (RMSE) promediado entre los miembros del conjunto. Así, para cuantificar el valor agregado de aplicar DA respecto a las simulaciones en OL, se utiliza la diferencia en la raíz del error cuadrático medio normalizado (NRMSE) definido en Álvarez-Garretón et al. (2015), el cual corresponde a la razón entre el RMSE obtenido en el caso DA y OL, y se expresa según la ecuación (8). Así, valores menores (mayores) a 1 indican mejoras (empeoramiento) en los resultados respecto al caso base (OL).

$$NRMSE = \frac{\frac{1}{N} \sum_{i=1}^N \sqrt{\frac{1}{T} \sum_{t=1}^T (S_t^{DA} - O_t)^2}}{\frac{1}{N} \sum_{i=1}^N \sqrt{\frac{1}{T} \sum_{t=1}^T (S_t^{OL} - O_t)^2}} = \frac{RMSE^{DA}}{RMSE^{OL}} \quad (8)$$

4. RESULTADOS Y DISCUSIÓN

4.1 Calibración modelo VIC

Los resultados de calibración del modelo muestran que todas las cuencas presentan resultados de KGE mayores o iguales 0,7 para el periodo de calibración (Figura 2). Del mismo modo, el índice NSE (Nash and Sutcliffe, 1970) asociado a los parámetros calibrados según KGE presenta valores mayores a 0,5 en todas las cuencas salvo en la cuenca 5721001, donde se obtiene un NSE igual a 0,48. En cuanto a la verificación del modelo en las diferentes cuencas (no presentado), se debe notar que si bien disminuye la eficiencia del modelo, los criterios se mantienen en el rango aceptable/satisfactorio.

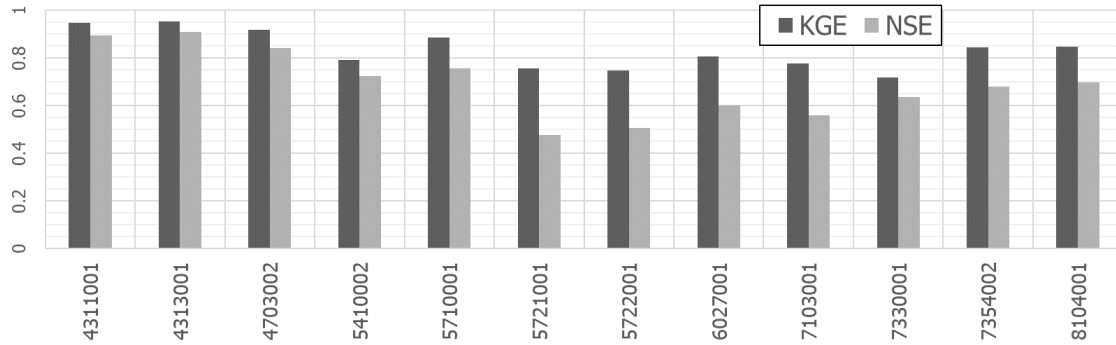


Figura 2: Resultados del proceso de calibración del modelo VIC usando el algoritmo SCE-UA y optimizando KGE como función objetivo.

4.2 Efectos de la asimilación sobre la observación

La simulación de escorrentía media diaria (Figura 3a), muestra que la configuración C2.E2 (i.e., sólo asimilación de fSCA y SWE) no tiene un valor agregado respecto al caso OL para las cuencas ubicadas en las zonas B y C (Figura 1), dominadas por la contribución nival. No obstante, se debe tener en cuenta que no se modificaron los errores del modelo respecto al caso C1.E1, agregándose un nuevo estado inicial perturbado (SWE) y el error asociado al mismo. El resto de las configuraciones (mencionarlas) mejora las estimaciones de escorrentía media diaria entre el 10% y 90% en todas las cuencas.

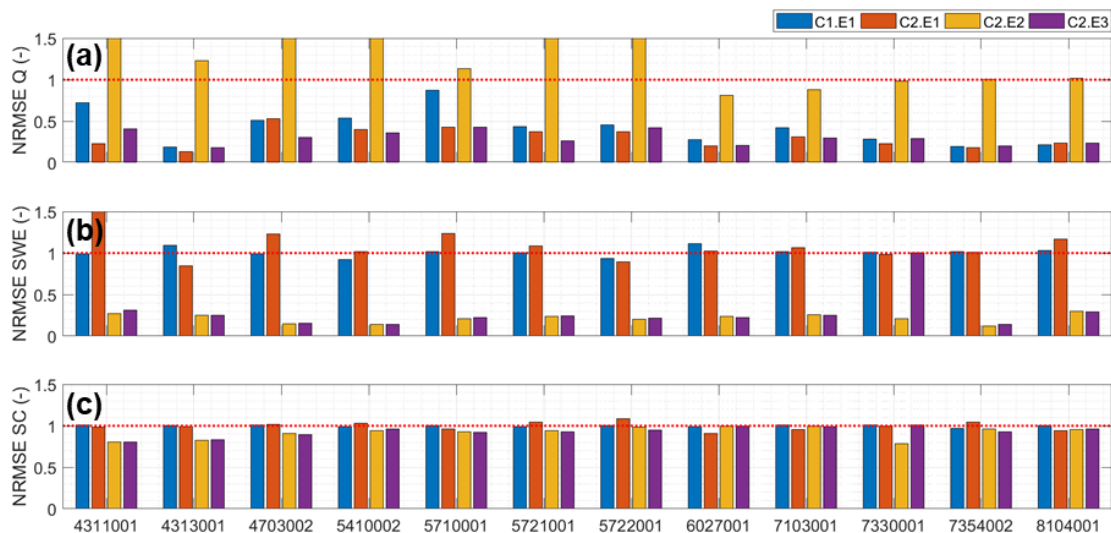


Figura 3: NRMSE para estimaciones de (a) escorrentía (Q), (b) equivalente de nieve en agua (SWE) y (c) cobertura de nieve (SC) con respecto a OL para el año hidrológico 2000/01. La línea roja segmentada representa el valor en que los resultados DA y OL son equivalentes. El eje Y se ha truncado en 1,5 para facilitar la comparación.

Los resultados indican que Figura 3asimilar sólo caudal (C1.E1) no ayuda a mejorar la estimación del SWE. Por otro lado, la configuración C2.E1 empeora las estimaciones de SWE para algunas cuencas (e.g, 4311001, 8104001) y las mejora en otras (e.g., 4313001, 5722001). En términos generales, se obtienen resultados similares para las configuraciones C2.E2 y C2.E3, con mejoras entre el 70-80% en el valor promedio de SWE a escala de

cuenca. En cuanto a la estimación de la cobertura de nieve (Figura 3c), se observa que, en general, los cambios no superan el 15%. La única excepción es la cuenca 7330001 donde, para la configuración C2.E2, el SC mejora un 20% respecto al caso OL.

4.3 Impactos de la asimilación en el balance hídrico

Bajo ciertos supuestos, el balance hidrológico a escala de cuenca se puede aproximar a cero en periodos de tiempo largos. Si bien en el presente estudio se modela sólo un año hidrológico, Figura 4 la estimación del balance hídrico permite evaluar los impactos de la asimilación. Los resultados muestran una mayor dispersión en los montos de ET anual en cuencas áridas respecto a cuencas húmedas ($Q/P > 0,5$, ver Figura 4). Se observa que, para la configuración C2.E2 (panel C), casi todas las cuencas pasan de un balance negativo a uno positivo al asimilar datos. No obstante, se debe considerar que los resultados presentados en secciones anteriores apuntan a que dicha configuración mantiene o incluso empeora el desempeño del modelo en las simulaciones de Q con respecto al caso OL. El resto de las configuraciones (paneles A, B y D), a diferencia del panel C, no muestran cambios evidentes en el signo de la variación anual de almacenamiento (i.e., existe acuerdo).

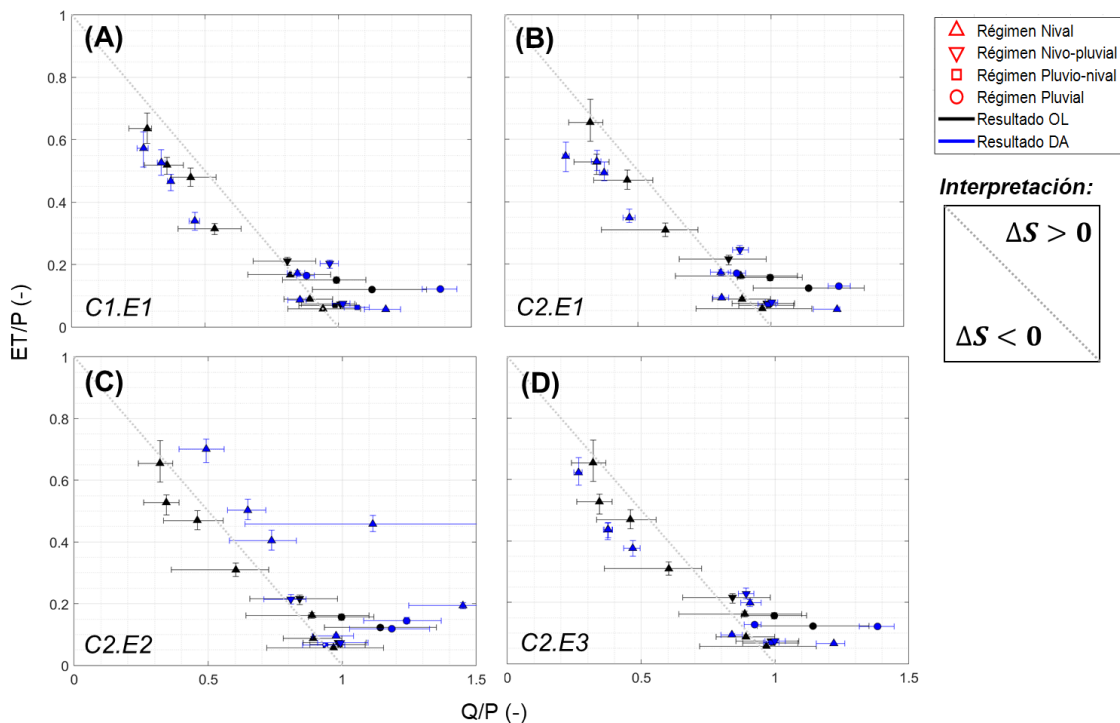


Figura 4: Estimación de balance hídrico para el año hidrológico 2000/01. Las barras de error corresponden al percentil 25% y 75% derivado de los miembros del conjunto asociado a cada cuenca, mientras que el símbolo corresponde a la mediana.

En algunas configuraciones y cuencas, se observan similitudes en las variaciones de almacenamiento obtenidas para una misma cuenca, pues si bien los montos cambian, el signo se mantiene para cada componente del balance (Tabla 4). Por ejemplo, para las cuencas 4311001 y 4313001, la partición ET/P y Q/P mantiene su signo en todos los casos, salvo C2.E2. Sin embargo, en otras cuencas (e.g., 5722001), no existe tal acuerdo entre todas las

configuraciones propuestas. Del mismo modo, resulta importante notar que la cuenca 8104001 no muestra cambios importantes (menores a $\pm 5\%$) y, en consecuencia, en el balance hídrico asociado. Algo similar ocurre para las cuencas 7103001 y 5710001, donde los cambios antes mencionados no superan, en valor absoluto, el 10% del valor OL.

Tabla 4: Cambios en la partición P-ET-Q al asimilar datos según las diferentes configuraciones propuestas respecto al caso sin asimilación, desagregados por cuenca. El color verde (rojo) indica aumento (disminución) en la magnitud de la partición, mientras que el color blanco representa valores cercanos a 0 (variaciones marginales).

BNA	Régimen hidrológico	C1.E1		C2.E1		C2.E2		C2.E3	
		$\Delta ET/P$	$\Delta Q/P$	$\Delta ET/P$	$\Delta Q/P$	$\Delta ET/P$	$\Delta Q/P$	$\Delta ET/P$	$\Delta Q/P$
4311001	Nival	-10%	-5%	-17%	-29%	7%	54%	-5%	-17%
4313001	Nival	8%	-14%	13%	-22%	31%	22%	22%	-22%
4703002	Nival	3%	3%	7%	-8%	19%	63%	22%	2%
5410002	Nival	-3%	25%	-4%	29%	18%	83%	15%	26%
5710001	Nival	-3%	-4%	5%	-9%	10%	10%	8%	-6%
5721001	Nival	1%	-6%	0%	0%	-5%	88%	-17%	9%
5722001	Nival	-3%	-17%	5%	-19%	-3%	142%	-7%	-18%
6027001	Nivo-pluvial	-4%	20%	14%	5%	-1%	-4%	6%	6%
7103001	Pluvio-nival	-9%	9%	4%	0%	-3%	-5%	2%	-1%
7330001	Pluvial	1%	23%	5%	10%	-3%	4%	-1%	21%
7354002	Pluvial	10%	-11%	9%	-13%	-8%	24%	-19%	-7%
8104001	Nivo-pluvial	3%	1%	5%	2%	-1%	1%	1%	1%

Para explicar las diferencias entre las distintas configuraciones, se relacionan los atributos de las cuencas con los efectos de DA en la partición de la precipitación, específicamente con $\Delta ET/P$ y $\Delta Q/P$ (Figura 5), que corresponden a variaciones de ET/P y Q/P con asimilación (DA) respecto al caso sin asimilación (OL). La incorporación de SWE y SC como variable asimilada (i.e., C2.E2 y C2.E3) cambia, en general, el signo de la correlación entre $\Delta ET/P$ respecto a los atributos analizados. Del mismo modo, se observa que la configuración C2.E2 cambia en la mayoría de los casos el signo de la correlación entre $\Delta Q/P$ respecto a los atributos. Descartando el caso C2.E2, identificado previamente como el de peor desempeño, se obtiene que, en cuencas húmedas (i.e., a menor PET/P y mayor Q/P), la asimilación aumenta la magnitud de la partición Q/P. Asimismo, en cuencas ubicadas más al sur con mayor cobertura de árboles nativos y pastizales, disminuye la partición ET/P y aumenta Q/P. Existe una correlación negativa (positiva) entre temperatura y precipitación (latitud) respecto a $\Delta ET/P$, lo cual se explica por la importancia del manto nival en la zona semiárida del dominio y cómo la asimilación permite corregir las estimaciones de SWE y, con ello, las tasas de evaporación que se explican por la sublimación desde el manto nival. No obstante, la mayoría de las correlaciones no resulta estadísticamente significativa (i.e., p-valor > 5%), por lo que, para robustecer el análisis, resulta necesario aumentar el tamaño de la muestra de cuencas y el periodo de asimilación.

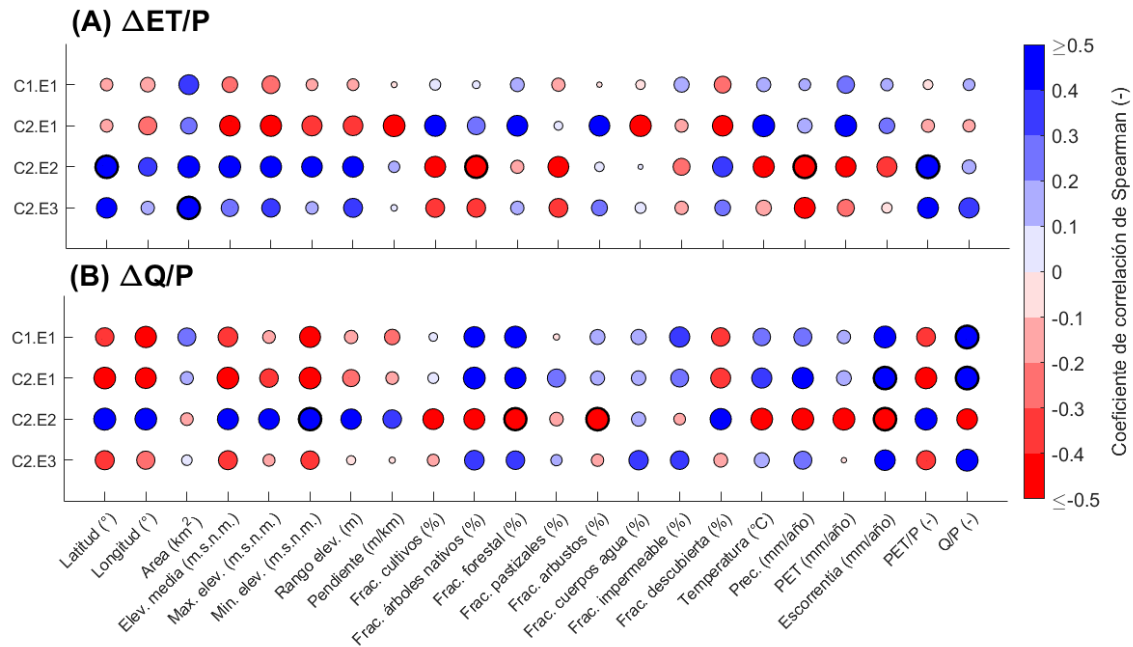


Figura 5: Coeficiente de correlación de Spearman entre impactos de DA en la partición P-ET-Q y atributos fisiográficos e hidroclimáticos. Espesor de línea en círculos indica la significancia estadística de la correlación (p -valor $\leq 5\%$).

En síntesis, la asimilación de datos tiene diferentes impactos en el balance hídrico anual dependiendo tanto de la cuenca como de la configuración (i.e., variable asimilada). Así, la elección de una configuración óptima del esquema de DA, entendiendo la dinámica del sistema y caracterizando apropiadamente las fuentes de incertidumbre, resulta clave para reducir la incertidumbre asociada a flujos y estados hidrológicos y, a su vez, ser consistentes en esas estimaciones. En este sentido, los resultados obtenidos indican que la mejor configuración corresponde a asimilar escorrentía, SWE y SC y actualizar la condición inicial de humedad del suelo por capa y SWE (i.e., C2.E3).

5. CONCLUSIONES

El presente estudio evalúa los beneficios de la asimilación de datos en estimaciones de flujos y estados hidrológicos en Chile continental. Para ello, se implementó el Filtro de Kalman de conjuntos (EnKF) acoplado al modelo hidrológico *Variable Infiltration Capacity* (VIC) en 12 cuencas con bajo nivel de intervención humana, ubicadas en Chile central. Los experimentos numéricos fueron diseñados en base a la perturbación de las forzantes del modelo y algunos estados iniciales, generándose cuatro configuraciones distintas. A partir de los resultados obtenidos, se concluye lo siguiente:

- **Impactos de DA en el balance hidrológico:** Cada configuración de DA tiene efectos diferentes en el balance hidrológico. Así, la elección de un esquema de DA, entendiendo la dinámica del sistema y caracterizando apropiadamente las fuentes de incertidumbre, resulta clave para reducir de manera consistente la incertidumbre asociada a flujos y estados hidrológicos. A su vez, los impactos (i.e., cuánto cambian las particiones P-ET-

Q) dependen de la cuenca de estudio, pudiendo existir una correlación entre dichos impactos y algunos atributos fisiográficos y/o hidroclimáticos.

- **Beneficio de asimilar información satelital:** Se concluye que asimilar SWE y SC mejora entre un 70-90% las estimaciones de SWE para el año 2000/01, lo cual se traduce en mejoras del RMSE de, en promedio, 80 ± 54 mm ($19 \pm 8\%$ del SWE máximo en la temporada). No obstante, lo anterior no asegura mejoras en la estimación de la escorrentía.
- **Observaciones in-situ vs información satelital:** Por sí sola, ninguna observación es capaz de asegurar la consistencia hidrológica en las estimaciones de flujos y estados simulados más allá de la variable asimilada. Así, la asimilación de múltiples observaciones permite restringir la variación de la variable asociada y, con ello, condicionar la dinámica de los procesos simulados.

Finalmente, el presente estudio abre el debate sobre las implicancias de la asimilación de datos en la estimación del balance hídrico en cuencas con componente nival. Más aún, se posiciona a la asimilación de datos como una alternativa para avanzar hacia una mejor representación del ciclo hidrológico, mejorando la gestión de los recursos hídricos y contribuyendo al cumplimiento de las metas 6.5 y 6.6 de los Objetivos de Desarrollo Sustentable (ODS). Asimismo, se evidencia la necesidad de evaluar la consistencia de los resultados de asimilación utilizando observaciones no asimiladas, y mejorar la comprensión de la relación estimación a priori – observación – estimación a posteriori.

AGRADECIMIENTOS

EMC agradece el financiamiento de ANID Beca Magister Nacional 2020 folio N°22200599. **Powered@NLHPC:** Esta investigación fue parcialmente apoyada por la infraestructura de supercómputo del NLHPC (ECM-02).

REFERENCIAS

- Alvarez-Garreton, C., Mendoza, P.A., Pablo Boisier, J., Addor, N., Galleguillos, M., Zambrano-Bigiarini, M., Lara, A., Puelma, C., Cortes, G., Garreaud, R., McPhee, J., Ayala, A., 2018. The CAMELS-CL dataset: Catchment attributes and meteorology for large sample studies-Chile dataset. *Hydrol. Earth Syst. Sci.* 22, 5817–5846. <https://doi.org/10.5194/hess-22-5817-2018>
- Alvarez-Garreton, C., Ryu, D., Western, A.W., Su, C.H., Crow, W.T., Robertson, D.E., Leahy, C., 2015. Improving operational flood ensemble prediction by the assimilation of satellite soil moisture: Comparison between lumped and semi-distributed schemes. *Hydrol. Earth Syst. Sci.* 19, 1659–1676. <https://doi.org/10.5194/hess-19-1659-2015>
- Clark, M.P., Nijssen, B., Lundquist, J.D., Kavetski, D., Rupp, D.E., Woods, R.A., Freer, J.E., Gutmann, E.D., Wood, A.W., Brekke, L.D., Arnold, J.R., Gochis, D.J., Rasmussen, R.M., 2015. A unified approach for process-based hydrologic modeling: 1. Modeling concept. *Water Resour. Res.* <https://doi.org/10.1002/2015WR017198>
- Clark, M.P., Rupp, D.E., Woods, R.A., Zheng, X., Ibbitt, R.P., Slater, A.G., Schmidt, J., Uddstrom, M.J., 2008. Hydrological data assimilation with the ensemble Kalman filter: Use of streamflow observations to update states in a distributed hydrological model. *Adv. Water Resour.* 31, 1309–1324. <https://doi.org/10.1016/j.advwatres.2008.06.005>
- Clark, M.P., Vogel, R.M., Lamontagne, J.R., Mizukami, N., Knoben, W.J.M., Tang, G., Gharari, S., Freer, J.E., Whitfield, P.H., Shook, K., Papalexioiu, S., 2021. The abuse of popular performance metrics in hydrologic modeling. *Water Resour. Res.* <https://doi.org/10.1029/2020wr029001>

- Copernicus Climate Change Service (C3S), 2017. ERA5: Fifth generation of ECMWF atmospheric reanalyses of the global climate [WWW Document]. Copernicus Clim. Chang. Serv. Clim. Data Store. URL <https://cds.climate.copernicus.eu/cdsapp#!/home>
- Cortés, G., Margulis, S., 2017. Impacts of El Niño and La Niña on interannual snow accumulation in the Andes: Results from a high-resolution 31 year reanalysis. *Geophys. Res. Lett.* 44, 6859–6867. <https://doi.org/10.1002/2017GL073826>
- DGA, 2019a. Aplicación de la metodología de actualización del balance hídrico nacional en las cuencas de la parte sur de la Macrozona Austral e Isla de Pascua, SIT N° 444.
- DGA, 2019b. Aplicación de la metodología de actualización del balance hídrico nacional en la macrozona sur y parte norte de la macrozona Austral, SIT N° 441.
- Döll, P., Douville, H., Güntner, A., Müller Schmied, H., Wada, Y., 2016. Modelling Freshwater Resources at the Global Scale: Challenges and Prospects. *Surv. Geophys.* 37, 195–221. <https://doi.org/10.1007/s10712-015-9343-1>
- Duan, Q., Gupta, V., 1992. Effective and Efficient Global Optimization for conceptual rainfall-runoff models. *Water Resour. Res.* 28, 1015–1031.
- Evensen, G., 1994. Inverse methods and data assimilation in nonlinear ocean models. *Phys. D Nonlinear Phenom.* 77, 108–129. [https://doi.org/10.1016/0167-2789\(94\)90130-9](https://doi.org/10.1016/0167-2789(94)90130-9)
- Giroto, M., De Lannoy, G.J.M., Reichle, R.H., Rodell, M., 2016. Assimilation of gridded terrestrial water storage observations from GRACE into a land surface model. *Water Resour. Res.* 52, 4164–4183. <https://doi.org/10.1002/2015WR018417>
- Gnann, S.J., McMillan, H.K., Woods, R.A., Howden, N.J.K., 2021. Including Regional Knowledge Improves Baseflow Signature Predictions in Large Sample Hydrology. *Water Resour. Res.* 57, 1–22. <https://doi.org/10.1029/2020WR028354>
- Hall, D.K., Riggs, G.A., Salomonson, V. V., Digirolamo, N.E., Bayr, K.J., 2002. MODIS snow-cover products 83, 181–194.
- Hamman, J.J., Nijssen, B., Bohn, T.J., Gergel, D.R., Mao, Y., 2018. The variable infiltration capacity model version 5 (VIC-5): Infrastructure improvements for new applications and reproducibility. *Geosci. Model Dev.* 11, 3481–3496. <https://doi.org/10.5194/gmd-11-3481-2018>
- Ibaceta, R., Splinter, K.D., Harley, M.D., Turner, I.L., 2020. Enhanced Coastal Shoreline Modeling Using an Ensemble Kalman Filter to Include Nonstationarity in Future Wave Climates. *Geophys. Res. Lett.* 47, 1–12. <https://doi.org/10.1029/2020GL090724>
- Klemeš, V., 1986. Operational testing of hydrological simulation models. *Hydrol. Sci. J.* 31, 13–24. <https://doi.org/10.1080/02626668609491024>
- Li, Y., Ryu, D., Western, A.W., Wang, Q.J., Robertson, D.E., Crow, W.T., 2014. An integrated error parameter estimation and lag-aware data assimilation scheme for real-time flood forecasting. *J. Hydrol.* 519, 2722–2736. <https://doi.org/10.1016/j.jhydrol.2014.08.009>
- Liang, X., Lettenmaier, D.P., Wood, E.F., Burges, S.J., 1994. A simple hydrologically based model of land surface water and energy fluxes for general circulation models. *J. Geophys. Res.* 99. <https://doi.org/10.1029/94jd00483>
- Liang, X., Wood, E.F., Lettenmaier, D.P., 1996. Surface soil moisture parameterization of the VIC-2L model: Evaluation and modification. *Glob. Planet. Change* 13, 195–206. [https://doi.org/10.1016/0921-8181\(95\)00046-1](https://doi.org/10.1016/0921-8181(95)00046-1)
- Liu, Y., Gupta, H. V., 2007. Uncertainty in hydrologic modeling: Toward an integrated data assimilation framework. *Water Resour. Res.* 43, 1–18. <https://doi.org/10.1029/2006WR005756>
- Müller Schmied, H., Eisner, S., Franz, D., Wattenbach, M., Portmann, F.T., Flörke, M., Döll, P., 2014. Sensitivity of simulated global-scale freshwater fluxes and storages to input data, hydrological model structure, human water use and calibration. *Hydrol. Earth Syst. Sci.* 18, 3511–3538. <https://doi.org/10.5194/hess-18-3511-2014>
- Muñoz-Castro, E., 2021. Asimilación de observaciones in situ e información satelital para estimaciones de balance hídrico en Chile continental. Universidad de Chile.
- Nash, J., Sutcliffe, J., 1970. River flow forecasting through conceptual models part I - A discussion of principles. *J. Hydrol.* 10, 282–290.
- Pauwels, V.R.N., De Lannoy, G.J.M., 2009. Ensemble-based assimilation of discharge into rainfall-runoff models: A comparison of approaches to mapping observational information to state space. *Water Resour. Res.* 45, W08428. <https://doi.org/10.1029/2008WR007590>
- Pokhrel, P., Gupta, H. V., 2010. On the use of spatial regularization strategies to improve calibration of

- distributed watershed models. *Water Resour. Res.* 46, W01505. <https://doi.org/10.1029/2009WR008066>
- Rasmussen, J., Madsen, H., Jensen, K.H., Refsgaard, J.C., 2015. Data assimilation in integrated hydrological modeling using ensemble Kalman filtering: evaluating the effect of ensemble size and localization on filter performance. *Hydrol. Earth Syst. Sci.* 19, 2999–3013. <https://doi.org/10.5194/hess-19-2999-2015>
- Refsgaard, J.C., Storm, B., Abbott, M.B., 1996. Comment on “A discussion of distributed hydrological modelling” by K. Beven. *Distrib. Hydrol. Model.* 279–287.
- Reichle, R.H., 2008. Data assimilation methods in the Earth sciences. *Adv. Water Resour.* 31, 1411–1418. <https://doi.org/10.1016/j.advwatres.2008.01.001>
- Renner, M., Kleidon, A., Clark, M., Nijssen, B., Heidkamp, M., Best, M., Abramowitz, G., 2020. How well can land-surface models represent the diurnal cycle of turbulent heat fluxes? *J. Hydrometeorol.* 22, 77–94. <https://doi.org/10.1175/JHM-D-20-0034.1>
- Vásquez, N., Cepeda, J., Gómez, T., Mendoza, P.A., Lagos, M., Boisier, J.P., Álvarez-Garretón, C., Vargas, X., 2021. Catchment-Scale Natural Water Balance in Chile, in: *Water Resources of Chile*. pp. 189–208. https://doi.org/10.1007/978-3-030-56901-3_9
- Zambrano-Bigiarini, M., Nauditt, A., Birkel, C., Verbist, K., Ribbe, L., 2017. Temporal and spatial evaluation of satellite-based rainfall estimates across the complex topographical and climatic gradients of Chile. *Hydrol. Earth Syst. Sci.* 21, 1295–1320. <https://doi.org/10.5194/hess-21-1295-2017>
- Zhang, H., Hendricks Franssen, H.-J., Han, X., Vrugt, J.A., Vereecken, H., 2017. State and parameter estimation of two land surface models using the ensemble Kalman filter and the particle filter. *Hydrol. Earth Syst. Sci.* 21, 4927–4958. <https://doi.org/10.5194/hess-21-4927-2017>