a) Objetivo del hito:

El objetivo del hito es generar un pronóstico hibrido para 45 cuencas de Chile Central, con diferentes tiempos de emisión a largo del año mediante un proceso automatizado.

b) Variables críticas estudiadas y rangos alcanzados:

Las variables críticas incluyen los índices climáticos mensuales estándar (SOI, ONI, PDO, NIÑO1,2,3,4), los datos horarios de reanálisis meteorológicos ERA5 de precipitación y temperatura, y las variables de estado de almacenamiento del modelo hidrológico como el manto nival, 2 estanques de agua subsuperficial y humedad del suelo. Como las bases de datos son de servicios externos (NOAA, ECMWF), los formatos y permisos de acceso pueden cambiar con el tiempo. El servicio de ERA5 tienen un retraso de aproximadamente 6 días respecto al presente, lo que debe considerarse en la ejecución.

Se probó el sistema retrospectivamente entre los años 1981 y 2019, incluyendo emisiones entre el 1 de mayo y el 1 de marzo (11 meses de emisión). Las métricas de evaluación nos indican que los primeros meses del año hidrológico (abril, mayo, junio) el pronóstico con los predictores actuales tiene coeficientes de correlaciones menores a 0.5 y el error probabilístico (CRPSS) medio es mayor 0.5 desde 1 de agosto. Por lo tanto, se puede ejecutar el pronóstico desde el 1 de agosto con mayor seguridad y precisión, los resultados en julio son variados y suelen ser más precisos en años húmedos (probabilidad de excedencia 25%).

c) Breve descripción de los experimentos, tratamientos y/o pruebas realizadas:

Se realizaron varias etapas de procesamiento y análisis de datos. Primero, se descargaron índices climáticos y datos meteorológicos. Estos datos se preprocesaron, escalando a nivel de cuenca y corrigiendo el sesgo (error sistemático) ocupando Shaake-Shuffle y DQM. Luego, se ejecuta un modelo hidrológico conceptual TUW utilizando datos diarios corregidos de precipitación y temperatura entre 1981 y el presente. A partir de estos datos, se ejecuta un modelo estadístico con predictores independientes descartando variables correlacionadas y seleccionando la mejor combinación de predictores de acuerdo con el criterio de Akaike (que minimiza el número de predictores con menor error). Finalmente, se utiliza el modelo de similitud KNN para distribuir el volumen en caudales medios mensuales dada la forma (caudal normalizado por el volumen) de años similares y se exportaran 2 archivos con los resultados para el año objetivo en formato csv. Este formato ha sido probado exitosamente para ser importado y mostrados en la página web [www.pronosticoscaudales.cl](http://www.pronosticoscaudales.cl) sección administración.

Para el preproceso meteorológico (agrupación mensual, promedio espacial, formateo) se ocupa el software CDO (Climate Data Operators) y para el resto de los procesos se ocupan códigos R (versión superior a 4.0). CDO se ejecuta internamente en R ocupando la línea de comando.

La ejecución de los procesos ha sido probada en estructuras Linux (Ubuntu) y Unix (MacOS), se han adaptado los códigos para que funcionen adecuadamente en ambos sistemas. En Windows la instalación de CDO es más inestable y puede variar dependiendo de la versión, desde Windows 10 se incluye una versión de Linux dentro del sistema por lo que se podría instalar ocupando la línea de comando (no probado).

La transferencia de los archivos a la página web no se encuentra automatizada dado que depende de la infraestructura del receptor (DGA). Por ejemplo, se puede correr el algoritmo con un cronometro en un servidor local que también contiene la página web o se pueden llevar todos los archivos a la nube y ejecutar remotamente desde internet. En general, se recomienda revisar los archivos antes ser publicados en la plataforma web para una mayor confiabilidad.

d) Principales resultados logrados y conclusiones:

Se logró la automatización del proceso de generación de pronósticos híbridos para 45 cuencas entre latitudes 27 ˚ S y 37 ˚ S. Dado los resultados de errores determinísticos y estadísticos se recomienda la ejecución cada mes comenzando desde agosto y finalizando en febrero, es decir 7 meses de emisión. La metodología permitió manejar un gran volumen de datos y los códigos permiten dar flexibilidad en la ejecución para mejorar o reparar cada proceso.