The PRONOS Hydrological Forecast System: Assessment And Learned Lessons From The First Year Of Operation At The Betania Hydropower Reservoir

Efraín Domínguez
Hector Angarita
Zulma Méndez
Gustavo Angulo
Diego Motavita

See next page for additional authors

Follow this and additional works at: http://academicworks.cuny.edu/cc_conf_hic

Part of the Water Resource Management Commons

Recommended Citation
Domínguez, Efraín; Angarita, Hector; Méndez, Zulma; Angulo, Gustavo; Motavita, Diego; and Rosmann, Thomas, “The PRONOS Hydrological Forecast System: Assessment And Learned Lessons From The First Year Of Operation At The Betania Hydropower Reservoir” (2014). CUNY Academic Works.
http://academicworks.cuny.edu/cc_conf_hic/96

This Presentation is brought to you for free and open access by CUNY Academic Works. It has been accepted for inclusion in International Conference on Hydroinformatics by an authorized administrator of CUNY Academic Works. For more information, please contact AcademicWorks@cuny.edu.
Authors
Efrain Dominguez, Hector Angarita, Zulma Mendez, Gustavo Angulo, Diego Motavita, and Thomas Rosmann
THE PRONOS HYDROLOGICAL FORECAST SYSTEM:
ASSESSMENT AND LESSONS FROM THE FIRST YEAR OF
OPERATION AT THE BETANIA HYDROPOWER RESERVOIR

EFRAÍN DOMÍNGUEZ (1), HECTOR ANGARITA (1), ZULMA MENDEZ (2), GUSTAVO
ANGULO (1), JUAN GABRIEL CEPEDA (2), DIEGO MOTAVITA(2), THOMAS ROSMANN (3)

(1): Departamento de Ecología y Territorio, Pontificia Universidad Javeriana, Colombia
(2): Departamento Hidrología, ICT, Emgesa S.A., Cra 13A 93-76, Bogotá, Colombia
(3): Maestría en Hidrosistemas, Pontificia Universidad Javeriana, Colombia

A feasible quantitative hydrological forecasting service is a combination of technological
elements, personnel and knowledge, working together to establish a stable operational cycle of
forecasts emission, dissemination and assimilation. The process for establishing such a system
usually requires significant resources and time to reach an adequate development and
integration in order to produce forecasts with acceptable levels of performance. Here we present
an operational assessment and lessons from the implementation and first year of operation of
the recently released Operational Forecast Service for the Betania’s Hydropower Reservoir –
PRONOS – located in the Upper-Magdalena River Basin (Colombia). PRONOS was developed
under the Flexible, Adaptive, Simple and Transient Time forecasting approach, or FAST-T, a
set of data structures, mathematical kernel, distributed computing and network infrastructure
designed to provide seamless real-time and operational forecast and automatic model
adjustment in case of failures in the data real-time transmission or assimilation. The PRONOS
service is specifically designed to support the hydropower operation, and therefore produces
forecasts of water levels and discharge for the three main streams affluent to the reservoir, for
lead times between +1 to +57 hours, and +1 to +10 days. At its current configuration, the
PRONOS performance objectives are fulfilled for 90% of the forecasts with lead times up to +2
days and +15 hours (using the S/σₘ predictability criteria) and the average accuracy is in the
range of 70-99% (r² criteria). However, results of longer lead times are at present not
satisfactory in terms of forecasts accuracy. System reliability was also evaluated in terms of
forecast performance consistency over time (65%), and the percentage of time offline (7%).

INTRODUCTION

The World Meteorological Organization [1], [2], states that a quantitative hydrological
forecasting system will be operational and feasible when: a) The necessary infrastructure for the
continuous measurement and reporting in real time of water levels and other hydro-
meteorological state variables; b) Information of physiographic, hydrological, meteorological,
topographical is available and updated for emission points of forecasts; c) Trained staff and
computing infrastructure is available for the assimilation of the real time data and operation of
hydrologic forecasting models; d) There is forecasting / modeling technology that can be used
in real time with acceptable levels of performance criteria; e) The channels of dissemination of
hydrological forecasting are formalized; and, f) It has a user community, officially registered,
able to assimilate the forecasts issued and provide feedback to the center of issuing forecasts.
This paper describes, from the above considerations, the recently developed and operationally deployed PRONOS hydrological forecast system (PRONOS-HFS), and presents results of its first year operational assessment. The PRONOS-HFS was developed as an end-to-end service to meet the specific requirements of the user base in charge of the Betania Hydropower Reservoir (BFR) operations. The system is intended for short-term inflows forecasting in order to optimize hydropower allocation, and to anticipate flow peak magnitudes and timing for flood management.

The selection of a modeling kernel is generally a key component of a hydrological forecast system. While there is a wide range of operators available for this purpose, from the most complex to the simplest [3], [4], its selection for operational regime is the central element upon which other aspects of the system architecture are determined. Considerations include: i. meeting the system objectives and the expectations of users, ii. Being as simple as possible in its use by forecasters on duty, iii. Being consistent with the available computing power and precision levels of the predictors recorded in real time; iv. Containing an optimization algorithm, which should be dynamic and adaptable to real time situations such as loss of reception of some of the monitoring signals, and v. Being applicable to the different physiographic conditions of forecasting points. Approaches for the selection of modeling kernels, range from single deterministic operators, such as based on ordinary differential equations [5], [6], Distributed (1D or 2D) based on conceptual or physically based models [4]; stochastic operators, such as based on Stochastic differential equations or auto-regressive models like AR, ARMA, VAR [7], [8], models based on optimal interpolation [9] and phase-space based operators such as in Casdagli et al. [10], Wagner [11], Hunt [12] and Friedrich et al. [13]. In some systems, multi-model combinations are used where no hydrological model can be identified as the “best” model in all circumstances [14], as well as run-time modifications [15], [16] to adjust boundary and initial conditions, as well as model parameters based on real-time stream-flow observations. In the PRONOS-HFS, the modeling kernel was developed to integrate several of the above mentioned elements. The system is based on an Optimal Adaptive Linear Combination (OALC) operator, a generalized numerical solution of locally linearized phase space equation [17], [18], and uses an automatic two layer run-time adaptive optimization that continuously updates and ranks a library of models for simulating the streamflow dynamics at different conditions, based on a predictability performance criteria. The approach was developed specifically to allow the system to autonomously adapt to changes in stream flow conditions or in data availability from the monitoring system, by continuously evaluating and detecting the most relevant information available to determine model ranking for the later configuration of forecasting models.

**PRONOS-HFS DESCRIPTION**

**What is it?**

The PRONOS hydrological forecast system is a short term, real time system that operationally predicts the total hourly affluences to the Betania hydropower reservoir. PRONOS-HFS is web based (Figure 1), autonomous, assimilates satellite transmitted hydro-climatological information, selects best predictors and forecast mathematical models and issues predictions at the required time. The Betania hydropower reservoir is located in the upper part of the Magdalena River basin and has three major tributary rivers: Magdalena, Páez and Yaguará, accounting for 55%, 40% and 5% of the reservoir inflows, respectively. The total drainage area is 13,297 km² with elevations in the range from 399 to 4500 m. From 35 available hydrological and meteorological stations (Figure 2), 7 have telemetric capabilities (quasi-real time) via NOAA-GOES satellites, for the rest of stations, the reading and recording of data is done manually twice per day. The hydrologic monitoring network covers 86% of the catchment area.
The PRONOS-HFS has three forecast locations at the reservoir’s main tributaries. The discharge flows of these locations are controlled by the hydrological stations 21047010-Pte Balseadero, (Magdalena river), 2105706-Paicol (Paez river) and 2108708-Hacienda Venecia (Yaguará river). From each of the forecast points, routing time is used to determine the total inflow at a given forecast time.

Figure 1. Screenshots of the PRONOS-HFS

Figure 2. The watershed of the Betania Hidropower Reservoir, its hydro-meteorological monitoring network and the forecast emission locations

Forecast requirements
Forecasts are required, in hourly resolution, for lead times of +57 hours, and, with daily resolution, for +10 days lead times to optimize the hydropower allocation, minimize the occurrence of low levels in the reservoir, and the anticipation of extreme events for minimization of spilling volumes.

PRONOS-HFS architecture:
PRONOS-HFS system logic is distributed into five modules: hydrometeorological information preprocessor, operational forecasting, models library builder/updater, post-processor, task manager and web-based interface. See Figure 3. The task manager is the application that controls the automatic execution of the system’s scheduled tasks, such as connecting to NOAA
servers to download monitoring network records, the execution of forecasting models and inflows integration, among others. The pre-processing module allows the information collected to be integrated into the database system. As part of the assimilation processes, the database is continuously updated and validated to include raw data from telemetric stations (15 minute resolution data for water levels and precipitation), time series of conventional stations (1 day resolution water levels and precipitation), aggregated/averaged time series of other variables of interest, including 1 hour/1 day resolution average water levels, total precipitation, and others that synthesize information from ground stations. The models library builder/updater module operates at off-forecast hours to optimize the mathematical models by determining the optimal model catalog at current stream conditions. This module is independent of the operational forecasting module, but is necessary to boot the forecasting system, since it provides the first catalog of optimized models. Following the system boot, this module will run every 12 to 48 hours (depending on the forecast horizon) to assimilate the most recent records in hydro meteorological forecasting models, analyze the historical results of the models and determine their performance and priority for the following operational forecast.

Figure 3. The PRONOS-HFS architecture. Communication operations between modules are shown as r: read, w: write.

Meanwhile, the operational forecasting module, is responsible for performing the calculations of water levels in the tributary rivers at the points closest to the reservoir. The process performed by this module is developed in four steps: i. Identification of stations whose data is available and updated, ii. Querying the model catalog to establish the list of viable models based on the stations’ availability, iii. Identification of the best model for the available records, and iv. Running the best available model and storing the results in the database. Finally, the post-processing module performs the integration of forecasts in time series of inflows to the reservoir. The system hardware is a XEON 4 core @ 2.4Ghz, 8Gb RAM, 1.000 TB (RAID 1). The software platform is Linux SUSE 7, using Python 2.7, Django application server and Oracle 10.

How does PRONOS-HFS forecast?

PRONOS-HFS uses an Optimum Adaptive Linear Combination - OALC approach that provides a computationally efficient and flexible mathematical kernel for the analyses of non-stationary time series with periodic components that occur at different frequencies. The full description of this method is described by Domínguez et al [18], [19]. The continuous adaptive approach also
provides flexibility for real-time automatic adjustment of models in case of failures in the process of transmission or data assimilation. The kernel is based on a two-step model identification process. The first step is intended to identify and rank potential models given recent watershed conditions, taking advantage of the available computing power in off-forecast hours. During this process, the system evaluates instances of given general model structures - or Metamodels, in terms of combination of predictor lags for each lead-time. The second step is done during forecast emission, and is intended to do a simple local optimization of the prioritized model. Therefore, the first step identifies the model structure, i.e. an optimal combination of predictors for a given model overall structure, while the second configures the model parameters for forecast emissions using the results from step 1. In both steps, the underlying operator is a linear vector function, that describes the observed transition of the water levels in a given point of interest \( X_i \), from time \( t \) to \( t + T \), as stated by the following equation (See notation in Table 1):

\[
\hat{X}_i(t + T) = \sum_{k=1}^{q} \sum_{j=1}^{\rho} c_{i,j,k}(t)W_{j,k}(t) + \gamma_i(t)
\]

Table 1. OALC approach conventions

<table>
<thead>
<tr>
<th>Symbol</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>( t )</td>
<td>Present time</td>
</tr>
<tr>
<td>( c_{i,j,k} ) and ( \gamma_i )</td>
<td>Time dependent parameters, obtained from step-wise calibration for the subset of real time available information</td>
</tr>
<tr>
<td>( T )</td>
<td>Forecast lead time</td>
</tr>
<tr>
<td>( \mathbb{X}(t) = {X_1, ..., X_i, ..., X_m}, X_i(t) \in \mathbb{R} )</td>
<td>Set of the ( m ) observable discrete signals, ( X_i )</td>
</tr>
<tr>
<td>( \hat{X}_i(t + T) )</td>
<td>Estimate of signal ( X_i ) at lead-time</td>
</tr>
<tr>
<td>( W(t) )</td>
<td>System state matrix at time ( t )</td>
</tr>
</tbody>
</table>

A parameterization window length – \( \theta \), is defined as the span of local states of the system in the interval \([t - \theta, t] : W_f(t), W_f(t - 1/f), ..., W_f(t - \theta/f)\) to determine \( c_{ijk} \) and \( \gamma_i \) at instant \( t \), by solving the following least squares optimization problem:

\[
\min \rightarrow \sigma^2(c_{ij}) = E[(\hat{X}_{i, t+T} - X_{i,t+T})^2] \\
\frac{\partial \sigma^2}{\partial c_{ijk}} = 0 \\
\frac{\partial \sigma^2}{\partial \gamma_i} = 0 \\
\text{With } k = 1, ..., m \text{ and } j = 1, ..., \rho.
\]

HOW DOES PRONOS-HFS PERFORM?

In order to evaluate the PRONOS-HFS performance, three elements were analyzed: 1) The usability of the web interface, 2) The availability of real time information during the first year of exploitation and 3) The quality of issued forecasts. Starting from scratch, the design of PRONOS-HFS was user oriented. Several workshops helped to understand what the users wanted to include, what they really needed and how they wished this would look like. A methodology was implemented in these workshops allowing to have mock-ups (Figure 4) as the initial approximation to the PRONOS-HFS user interface. The programming language python 2.7 with the scientific modules numpy, scipy and django [20]–[23], was selected as the development platform.
The reliability of the data acquisition/transmission varies from node to node, with a record availability between 20-99% for telemetric stations (average 78%), and between 40-99% for conventional stations (average 91%). However, monitoring is clustered around the main streams and some extensive areas of the basin do not have a good coverage of hydrological or meteorological monitoring, especially in the highest parts of the basin. Data reliability is a key consideration in the forecast quality. In the case of the PRONOS–HFS, the system was designed to be able to adapt to failures or delays in the transmission from some of the telemetric of conventional nodes, but nevertheless as the information availability becomes lower, forecast performance quality decreases.

![Figure 4. User oriented design of the PRONOS-HFS user interface](image)

![Figure 5. Average data availability from real time stations (WL: water levels, PR: Precipitation gauges)](image)

To assess the quality of issued forecasts, an orthogonal set of hydrological model performance criteria were used as recommended in Domínguez et al [24]. For this work, the performance criteria of the Russian Hydrometeorological Center ($\frac{S}{\sigma_d}$), the so called IRMSE at the hydrotest website for hydrological model performance assessment, together with the coefficient of determination and also the percentage of successful forecasts were used to evaluate the quality of broadcasted hydrological forecasts. In this sense, 87% of forecasts issued with 14 hours lead times should be considered as of good quality when its IRMSE holds values of $0.5 < \frac{S}{\sigma_d} \leq 0.8$ [24], [25]. At the same time, forecasts with lead times lower than 12 hours showed determination coefficients higher than 0.79. For forecast lead times lower than 5 hours the mean relative bias is always lower than 15% but the system performance decreased heavily for forecasts with lead times greater than 15 hours. At the moment this paper was written, PRONOS-HFS was operating with 54% of the stations with real time transmission capability,
which means that there is an important opportunity to improve the forecast performance by simply setting up and keeping online the unavailable stations.

![Figure 6. Forecast Performance](image)

**LESSONS LEARNED AS CONCLUDING REMARKS**

Operational use of PRONOS-HFS for daily and hourly water level and discharge forecasts shows the feasibility of such predictions in real time mode for Colombian water basins and also the advantage this system gives for the operation of hydropower reservoirs. The current level of hydrological monitoring with near to real time transmission capabilities is enough to implement the OALC forecast approach as the kernel for the operational hydrological forecast system. The design of such systems should be based on a user oriented approach. The system design time should last at least 30% of the project duration. The forecast requirements should be established with special attention. Forecast base knowledge transfer should be provided to the system users in order to provide a better comprehension of broadcasted forecasts. A user friendly interface is a key factor to increment the system usability. Keeping the monitoring system in its best condition for measurements and transmission of hydro meteorological information is a major requirement to hold a good quality of issued forecasts. At the same time, optimal density of the monitoring network is a factor that should be revisited recurrently. Forecast uncertainty assessment modules should become mandatory elements for the PRONOS-HFS.

**BIBLIOGRAPHY**


