IMPROVING STREAMFLOW FORECASTS IN THE NOAA NATIONAL WATER MODEL USING OBSERVATIONAL CONSTRAINTS ON SNOWPACK ALBEDO AND SNOW-COVERED AREA FROM STC-MODSCAG  

Thomas L. Enzminger¹, Aubrey L. Dugger², Karl Rittger³, Edward H. Bair⁴, James L. McCreight⁵, Mark S. Raleigh⁶, and Mary J. Brodzik⁷

ABSTRACT

The NOAA National Water Model (NWM) is a physically-based modeling system which simulates major hydrologic processes across the conterminous United States (US). Difficulties in accurately simulating snowpack states hinder the NWM’s ability to provide high-quality streamflow forecasts, particularly in snow-dominated western US mountains. Errors in snowpack simulations propagate into streamflow time series, incurring errors in both magnitude and timing of peak streamflow in snow-dominated basins.

We imposed observation-based constraints on fractional snow-covered area (fSCA) and snowpack albedo in the NWM and investigated the impacts on simulated snow states and streamflow over the Upper Colorado River Basin. We identified a set of parameters that influence the relationship between snow depth and fSCA (the snow depletion curve) and seasonal snowpack evolution. For each parameter, we derived spatially-distributed values using 15 years of data from STC-MODSCAG (Spatially and Temporally Complete MODIS Snow-Covered Area and Grain Size), a remote sensing-based data product providing daily estimates of fSCA, snowpack albedo, and other variables at ~500 m spatial resolution. When implemented into the NWM’s snow model, the derived values tended to shift simulated streamflow peaks lower and earlier, often improving agreement with observed streamflow. Results from these experiments will help to improve streamflow forecasting for water management and inform NWM data assimilation strategies to accommodate model parameter uncertainties.

INTRODUCTION

Streamflow forecasting is central to water resource management strategies, particularly in regions prone to water scarcity such as the western United States. The accuracy of streamflow forecasts at lead times varying from hours to weeks impacts the efficacy of flood risk assessments, irrigation management decisions, drought severity predictions, and numerous other management applications. The NOAA (National Oceanic and Atmospheric Administration) National Water Model (NWM) is a hydrologic modeling system intended to meet this streamflow forecasting need at high spatial and temporal resolution across the continental United States. The NWM provides streamflow forecasts at multiple lead times ranging from 18 hours to 30 days for ~2.7 million stream reaches within this domain (ref). As a physically-based modeling system, the NWM simulates the terrestrial water and energy balances, outputting spatially-distributed estimates of soil moisture, evapotranspiration, snow water equivalent (SWE), and other quantities in addition to its streamflow forecasts.

As is the case for any method of streamflow prediction, NWM forecasts are subject to a significant error budget which is spatially and temporally variable. This error budget encompasses uncertainties in meteorological forcing data as well as errors related to model structure and parameterization in the NWM’s numerous land surface model modules (e.g. snow, soil moisture, groundwater, etc.). Structural and parametric errors in the NWM’s snow model (Noah-Multi-Parameterization (Niu et al., 2011)) are particularly impactful on the quality of streamflow predictions in the western United States, where on average >50% of annual streamflow originates as snowpack (Li...
et al., 2017). Errors in NWM snow simulations affect the magnitude and timing of peak simulated streamflow—key metrics used in water management decision making throughout the western United States.

In this study, we investigated the extent to which simulated streamflow could be improved by applying observational constraints to parameters within the NWM snow model. We chose two empirical model parameters which affect NWM fractional snow-covered area (fSCA) and snow albedo. We ran constrained simulations and quantified improvements in simulated streamflow by comparing to both observations and default, unconstrained simulations.

Parameter value constraints were derived using fSCA and snow albedo from STC-MODSCAG (Spatially and Temporally Complete MODIS Snow-Covered Area and Grain Size)—a suite of remote-sensing-based data products generated from MODIS (Moderate Resolution Imaging Spectroradiometer) observations (Rittger et al., 2016). We present results from the Upper Colorado River Basin (UCRB) over water years 2008-10.

**METHODS**

**Observational Data Products**

STC-MODSCAG fSCA and snowpack albedo are produced at ~500 m spatial and daily temporal resolution across the western United States. Both variables are generated using the MODSCAG spectral unmixing algorithm (Painter et al., 2009). The fSCA product has additionally been corrected for cloud and tree-canopy occultation of snowpack (Rittger et al., 2020). The snowpack albedo product is updated to solve for both snow grain size and light-absorbing impurity (e.g. dust, black carbon) content (Bair et al., 2019). Both variables have been validated at the small basin and/or study plot scale, where they have exhibited low root mean square error and minimal bias. We regridded STC-MODSCAG fSCA and snowpack albedo to the 1 km NWM grid for ease of comparison.

**Model Parameter Adjustments**

NWM simulations of fSCA and snowpack albedo with default snow parameters differed in key ways from their STC-MODSCAG equivalents. NWM fSCA exhibited nearly binary behavior: NWM grid cells tended to be either 0% or 100% snow covered, lacking the intermediate fSCA values common in STC-MODSCAG (Figure 1). Although disagreement between simulated and observed snowpack albedo was less pronounced compared to fSCA, discrepancies in the temporal variability of snowpack albedo indicated issues with the model’s snowpack aging component.

In order to use STC-MODSCAG fSCA and snowpack albedo values to constrain NWM simulated snow states and streamflow, we identified parameters which impact the maximum fSCA and snowpack aging rate in a given grid cell. For fSCA, we chose the parameter ‘scamax’, which directly imposes an upper bound on the model snow depletion curve—an empirically derived relationship between snow depth and fSCA (Niu and Yang, 2007) (Figure 2). In each grid cell within the model domain, we iteratively solved for the

**Figure 1.** (a) Observed (STC-MODSCAG) and (b) simulated (NWM) fSCA over the Upper Colorado River Basin for April 1, 2010.

**Figure 2.** Observed (orange), simulated (blue), and re-scaled simulated (green) relationship between snow depth and fSCA for an individual grid cell. Simulated points are re-scaled to match the approximate observed maximum (scamax).
scamax value which most closely matched the 99th percentile value from the observational data. The result was an optimized map of scamax constrained by STC-MODSCAG data. We used this optimized scamax distribution in an updated NWM simulation over the UCRB over water years 2008-10.

The NWM contains two snowpack albedo model options: the BATS (Biosphere-Atmosphere Transfer Scheme) (Yang et al., 1997) and the CLASS (Canadian Land Surface Scheme) (Verseghy, 1991). We used BATS snowpack albedo as it exhibited better baseline agreement with STC-MODSCAG. The BATS model employs an exponential decay-style aging routine, where snowpack albedo declines from a pre-set new snow value as the snowpack ages. We chose to vary the BATS parameter ‘swemx’, which determines the amount of new SWE required to reset the model’s snowpack age to zero and broadband snowpack albedo to the new snow value (~0.84) (Figure 3). The swemx parameter impacts snow albedo less directly than scamax impacts fSCA. As such, deriving a spatially heterogeneous, optimized swemx distribution is significantly more complex, and remains in progress. Instead, we varied swemx in a spatially homogeneous manner across a small test basin—the Animas River Basin in southwest Colorado. We present results using the default value of swemx, as well as a high swemx endmember (swemx_max).

RESULTS AND DISCUSSION

We found that the scamax simulation outperformed the default simulation throughout the UCRB with respect to both snow states and streamflow. The NWM default simulation exhibited pervasive high biases in fSCA, particularly at lower elevations. These high biases were largely corrected by constraining the model with observational data (Figure 4). Impacts on snowpack albedo and SWE were relatively small.

Streamflow in the scamax simulation was also markedly improved relative to the
default simulation. In the default NWM simulation, streamflow peaks tended to be too large in magnitude and occur too late in the year. In the scamax simulation, streamflow peaks shifted lower and earlier, yielding modest overall improvements in error metrics. These modest improvements were observed at stream gauges distributed across the UCRB (Figure 5).

The swemx simulations in the Animas River Basin yielded mixed results. Relative to the default simulation, the swemx_max simulation exhibited degraded performance with respect to snowpack albedo, but slightly improved performance with respect to both fSCA and streamflow. Specifically, the advent of snow melt occurred earlier in the swemx_max simulation, in better agreement with STC-MODSCAG observations (not shown). As in the scamax simulation, the earlier melt in the swemx_max simulation incurred lower, earlier peaks in simulated streamflow compared to the default simulation (Figure 6).

Results from this study underscore the importance of high-quality, observation-based data products such as STC-MODSCAG. These data offer vital insights into model behavior, such as the propensity of the NWM snow model to overestimate fSCA, particularly at low elevations. Our finding that streamflow performance can be improved by forcing model fSCA to more closely resemble STC-MODSCAG fSCA is encouraging, and suggests that future efforts to assimilate STC-MODSCAG data into the NWM may offer further improvements. We will institute a similar methodology for applying observational constraints to snowpack albedo in the near future.

Our results also signal a need to re-evaluate the interdependent components of the NWM snow model. Given our finding that increasing swemx_max improved model fSCA and streamflow but degraded albedo, it is clear that the interactions between the individual modules of the NWM snow model do not closely approximate real snowpack dynamics. One future avenue to improve the NWM snow model will be to re-derive or reformulate the empirical snow depletion curve relating snow depth to fSCA. A reformulation must address the tendency of the model to produce binary fSCA maps. The snowpack albedo models in the NWM likely require similar treatment, as both rely heavily on empirically-estimated parameters and relationships which have not been robustly re-evaluated in some time. Data products such as STC-MODSCAG offer a significant opportunity to improve upon the model structure and parameterizations currently in place in the NWM, and therefore to improve the predictive power of the model.

Figure 5. Change in Nash-Sutcliffe Efficiency of daily streamflow at UCRB stream gauges. Positive changes (purple) indicate improved performance in the scamax simulation relative to the default simulation.

Figure 6. Comparison of observed (blue), default (orange), and swemx (green) streamflow at U.S. Geological Survey gauge 09361500. Changes in error metrics indicate improvement in the swemx simulation relative to the default simulation.
REFERENCES


