

## Data Assimilation for Calibration-Prediction Using SWAT Model

M. Bayat<sup>1</sup>, H. Alizadeh<sup>2\*</sup> and B. Mojaradi<sup>3</sup>

### Abstract

This paper deals with parameter estimation of SWAT model by means of streamflow data assimilation and application of calibrated model for hydrological simulation of Mahabad River leading to Urmia Lake. Data assimilation algorithm is compared with SUFI2 algorithm. SUFI2 is an uncertainty-based optimization method first developed for auto-calibration of environmental and water resource models. Due to availability in SWAT-CUP package this is usually used for calibration of SWAT. To illustrate capabilities of data assimilation for calibration of the model and prediction of the river discharge, Ensemble Kalman Filter (EnKF) is utilized in a joint state-parameter estimation framework. Both coding EnKF and calling SWAT is done in MATLAB environment. Results showed that the joint state-parameter estimation using EnKF for SWAT, lead to improvement of accuracy of simulation and prediction of Mahabad River's monthly discharge at Bitass hydrometry gauge compared to parameter estimation of the model using SUFI2.

**Keywords:** Calibration, Prediction, SWAT, Data Assimilation, Ensemble Kalman Filter.

Received: April 18, 2017

Accepted: June 11, 2017

## جذب داده برای واسنجی-پیش‌بینی با استفاده از مدل SWAT

مهرداد بیات<sup>۱</sup>، حسین علیزاده<sup>۲\*</sup> و برات مجردی<sup>۳</sup>

### چکیده

در این پژوهش به منظور تخمین پارامترهای مدل SWAT و بکارگیری آن برای شبیه‌سازی هیدرولوژیکی حوضه آبریز مه‌آبادچای (از رودخانه‌های منتهی به دریاچه ارومیه) و پیش‌بینی آبدهی رودخانه، از جذب داده استفاده می‌شود. در ابتدا پس از مدل‌سازی با استفاده از SWAT، روش SUFI2 که جزء روش‌های واسنجی خودکار مبتنی بر عدم قطعیت و بهینه‌سازی بوده و درون بسته نرم‌افزاری SWAT-CUP در دسترس است، برای واسنجی مدل SWAT استفاده گردید. سپس برای نشان دادن قابلیت‌های جذب داده متوالی در تخمین پارامترهای مدل SWAT و پیش‌بینی آبدهی رودخانه به صورت همزمان، از فیلتر آنسامبل کالمن (EnKF) به صورت تخمین متصل متغیر حالت-پارامتر با برنامه‌نویسی درون محیط نرم‌افزار MATLAB و فراخوانی مدل در همانجا، استفاده شد. نتایج نشان می‌دهد که روند همزمانی واسنجی-پیش‌بینی با استفاده از الگوریتم EnKF برای مدل SWAT باعث افزایش دقت پیش‌بینی و شبیه‌سازی آبدهی ماهانه رودخانه مه‌آباد در ایستگاه بیطاس در مقایسه با نتایج حاصل از واسنجی مدل با استفاده از الگوریتم SUFI2 می‌شود.

**کلمات کلیدی:** واسنجی، پیش‌بینی، SWAT، جذب داده، فیلتر آنسامبل کالمن.

تاریخ دریافت مقاله: ۹۶/۱/۲۹

تاریخ پذیرش مقاله: ۹۶/۳/۲۱

1- MSc Graduate in Water Resources Engineering and Management, School of civil engineering, Iran University of Science and Technology, Tehran, Iran.

2- Assistant Professor, School of civil engineering, Iran University of Science and Technology, Tehran, Iran. Email: alizadeh@iust.ac.ir

3- Assistant Professor, School of civil engineering, Iran University of Science and Technology, Tehran, Iran.

\*- Corresponding Author

۱- دانش‌آموخته کارشناسی ارشد مهندسی عمران مهندسی و مدیریت منابع آب، دانشگاه علم و صنعت ایران-تهران.

۲- استادیار دانشکده مهندسی عمران، دانشگاه علم و صنعت ایران-تهران.

۳- استادیار دانشکده مهندسی عمران، دانشگاه علم و صنعت ایران-تهران.

\*- نویسنده مسئول

بحث و مناظره (Discussion) در مورد این مقاله تا پایان تابستان ۱۳۹۷ امکانپذیر است.

از آنها برای واسنجی مدل و همچنین پیش‌بینی به صورت همزمان و متوالی استفاده می‌شود و نیاز به فراهم بودن مجموعه اولیه از داده‌ها نیست (Moradkhani and Sorooshian, 2008) و بدین ترتیب داده‌ها به صورت تدریجی (و نه یکجا) به مدل داده می‌شوند و این خاصیت باعث بهبود در روند واسنجی-پیش‌بینی می‌گردد. این رویکرد در مدل‌سازی بواسطه مشابهت با فرآیندهای بیولوژیکی تدریجی جذب مواد مغذی در بدن جانداران، تحت عنوان جذب داده (درون مدل) نامگذاری شده است. در این زمینه (Sun et al., 2015) نشان دادند که واسنجی یک مدل هیدرولوژیکی با استفاده از الگوریتم‌های متداول واسنجی خودکار لزوماً دقت مناسبی در پیش‌بینی متغیرهای هیدرولوژیکی در دوره‌ی اعتبارسنجی بوجود نخواهد آورد و همچنین قابلیت‌های روش جذب داده در رفع مشکلات اینچنین را نمایش دادند. (DeChant, 2010) دقت دو روش واسنجی خودکار شامل بهینه‌سازی سراسری و روش مبتنی بر عدم قطعیت MCMC را با روش‌های جذب داده شامل Particle Filter، EnKF و Particle Smoother در واسنجی دو مدل مفهومی هیدرولوژیکی مقایسه نموده و نشان داد که روش‌های جذب داده بر روش‌های واسنجی خودکار برتری دارند.

از مهمترین روش‌های جذب داده‌ی متوالی می‌توان به فیلتر کالمن (KF)، فیلتر توسعه یافته‌ی کالمن (EKF) و فیلتر (مبتنی بر) آنسامبل کالمن (EnKF) اشاره کرد. با فرض نرمال بودن خطا و خطی بودن سیستم، روش فیلتر کالمن می‌تواند بهترین تخمین را ارائه نماید (Evensen, 2009)؛ در اینجا بهترین تخمین آن است که کمترین واریانس را دارد. در سیستم‌های غیرخطی، مانند سیستم‌های هیدرولوژیکی، و در شرایط غیرنرمال بودن خطا، فیلتر کالمن مناسب نیست و بدین جهت روش EKF پیشنهاد شد که ایده‌ی اصلی آن خطی‌سازی روابط سیستم با استفاده از مشتق محلی (بسط تیلور) و سپس بکارگیری KF بود. دشواری‌ها و هزینه محاسبه‌ی مشتق محلی (Da Ros and Borga, 1997) و همچنین امکان رسیدن به جواب‌های ناپایدار (Evensen, 1994) نقاط ضعف اصلی روش EKF است. (Reichle et al., 2007) نشان دادند که در مسائل جذب داده بزرگ مقیاس با متغیرهای زیاد، نظیر مدل‌های هیدرولوژیکی توزیعی، EKF قابلیت مناسب ندارد. با توجه به محدودیت‌های EKF، Evensen (1994) برای نخستین بار روش EnKF را به منظور جذب داده در مدل‌های غیرخطی هواشناسی و اقیانوسی پیشنهاد داد. در EnKF همچون سایر روش‌های مبتنی بر مجموعه آماری یا آنسامبل، از تکنیک نمونه‌گیری استفاده می‌شود. همچنین EnKF به دلیل عدم نیاز به خطی‌سازی روابط، دارای بار محاسباتی کمتری نسبت به روش EKF است.

یکی از موضوعات مهم در زمینه‌ی مدیریت منابع آب، پیش‌بینی دقیق متغیرهای هیدرولوژیکی همچون آبدهی رودخانه‌ها، تراز آب‌های زیرزمینی و رطوبت خاک می‌باشد. در این رابطه استفاده از مدل‌های مبتنی بر فرآیند جهت پیش‌بینی متغیرهای هیدرولوژیکی، مستلزم واسنجی مدل‌ها یا به عبارت دیگر تخمین پارامترهای آنها است. با این توجه روش‌های واسنجی خودکار مدل‌های هیدرولوژیکی در دهه‌های اخیر بسیار مورد توجه محققین قرار گرفته و در این زمینه دو دسته مهم روش‌های مبتنی بر عدم قطعیت و روش‌های مبتنی بر بهینه‌سازی توسعه و بکار رفته‌اند. اجرای این روش‌ها بدون وجود رکوردهای تاریخی نسبتاً طولانی میسر نبوده چرا که در صورت بکارگیری دوره‌های کوتاه مدت امکان فرارگیری دوره‌ی واسنجی در شرایط خاص مانند خشکسالی و دوره پیش‌بینی در شرایط کاملاً متفاوت مانند ترسالی و یا بالعکس و در نتیجه امکان به خطا افتادن مدل‌ها و همچنین پیش‌بینی‌ها وجود خواهد داشت. علاوه بر آن بر اثر پدیده‌ی تغییر اقلیم، امکان بروز سیلاب‌ها و خشکسالی‌هایی که نمونه‌ای مانند آنها در تاریخچه‌ی مشاهداتی وجود نداشته نیز محتمل است. با انگیزه فائق آمدن بر این مشکلات در سال‌های اخیر پژوهشگران ایده‌های مربوط به واسنجی متوالی مانند روش جذب داده متوالی<sup>۱</sup> را بکار گرفتند.

جذب داده در هیدرولوژی به معنای ترکیب کردن مقادیر مشاهداتی (اندازه‌گیری شده) متغیرهای هیدرولوژیکی با مقادیر متناظر شبیه‌سازی شده توسط مدل هیدرولوژیکی در چهارچوبی قاعده‌مند و هدفمند، به منظور کاهش عدم قطعیت و افزایش اعتمادپذیری در نتایج است (Moradkhani, 2008; Xie and Zhang, 2010). جذب داده می‌تواند جهت تخمین پارامترها و پیش‌بینی متغیرها به صورت همزمان و همچنین متوالی در زمان و مکان استفاده گردد. این دیدگاه که از فضای پردازش سیگنال برخاسته، در زمینه‌های مرتبط با آب ابتدا در علوم هواشناسی و اقیانوس‌شناسی بکار رفته است (Xie and Zhang, 2010) و با گذشت زمان استفاده از روش‌های جذب داده در هیدرولوژی نیز مورد توجه واقع شده است (Liu et al., 2012).

روش‌های متداول واسنجی خودکار اصطلاحاً برون‌خطی‌اند، بدین معنا که در مرحله اول می‌بایست مجموعه‌ای نسبتاً بزرگ از سری زمانی مشاهدات در دسترس باشد. در مرحله بعدی با استفاده یکجا از مجموعه داده، مدل واسنجی شده و در مرحله نهایی برای پیش‌بینی از آن استفاده می‌گردد. در مقابل آن، روش‌های جذب داده اصطلاحاً درون‌خطی‌اند، بدین معنا که با اضافه شدن مشاهدات در هر گام زمانی،

یکی از نرم‌افزارهای هیدرولوژیکی مهم و پرکاربرد در جهان SWAT است. در سالهای اخیر بسته نرم‌افزاری ای به نام SWAT-CUP توسعه یافت که بوسیله آن امکان واسنجی مدل SWAT با استفاده از روشهای واسنجی خودکار شامل SUFI2، GLUE، MCMC، PSO و Parasol فراهم می‌آید. این بسته نرم‌افزاری امروزه توسط بسیاری از کاربران SWAT برای واسنجی بکار می‌رود.

تحقیقات معدودی در سالهای اخیر برای فراهم‌سازی امکان جذب داده در مدل SWAT انجام گردیده است. برای نخستین بار Xie and Zhang (2010) با استفاده از EnKF در یک مورد مطالعاتی مصنوعی به تخمین پارامتر و متغیر حالت مدل SWAT اقدام نمودند. همچنین Chen et al. (2011) به منظور پیش‌بینی (تخمین) متغیرهای تبخیر-تعرق، رواناب سطحی و رطوبت خاک با استفاده از مدل SWAT، مشاهدات رطوبت سطحی خاک را بوسیله EnKF جذب کردند. Han et al. (2012) با استفاده از EnKF به جذب مقادیر رطوبت سطحی خاک در مدل SWAT پرداختند. همچنین Sun et al. (2015) با استفاده از روش EKF به جذب مشاهدات رواناب، برای تخمین پارامتر CN2 در مدل SWAT و پیش‌بینی رواناب پرداختند.

در این تحقیق روند واسنجی-پیش‌بینی در مدل‌سازی هیدرولوژیکی با استفاده از SWAT برای رودخانه مهابادچای واقع در حوضه آبریز دریاچه ارومیه مورد توجه خواهد بود. در این راستا مشاهدات رواناب جهت تخمین ۱۴ پارامتر و پیش‌بینی آبدهی ماهانه با بکارگیری EnKF در مدل SWAT جذب می‌شود. در مقایسه با کارهای مشابه قبلی، بطور مشخص مقاله Sun et al. (2015) و مقاله Xie and Zhang (2010)، نوآوری این تحقیق در بکارگیری EnKF که قطعاً بر EKF برتری دارد در یک مطالعه موردی واقعی برای جذب داده‌های رواناب در مدل SWAT است. همچنین با توجه به گسترده‌ی کاربرد روشهای جذب داده در پژوهشهای بین‌المللی و از طرفی عدم استفاده از روش مذکور در مدل‌سازی‌های هیدرولوژیکی و منابع آب کشور چه در تحقیق و چه در کاربرد، این مقاله به بررسی و معرفی روشهای مذکور و قابلیت‌های آنها نیز می‌پردازد.

## ۲- جذب داده

با توجه به افزایش دسترسی به داده‌ها خصوصاً مشاهدات سنجش از دور و همچنین با توجه به آنکه حجم این داده‌ها معمولاً زیاد بوده و از منابع مختلف با ویژگی‌ها و دقت‌های کاملاً متفاوت تهیه می‌شود یک رویکرد هوشمند برای مواجهه و بکارگیری داده در روند مدل‌سازی و

خصوصاً مدل‌سازی مبتنی بر فرآیند در هیدرولوژی مورد نیاز است. جذب داده یک رویکرد جدید در این راستاست که در آن مشاهدات جدید در هر زمان که در دسترس قرار می‌گیرد برای بهبود مدل و پیش‌بینی بکار می‌روند و به عبارت دیگر مرحله واسنجی مدل و بکارگیری مدل برای پیش‌بینی (به نوعی اعتبارسنجی مدل) به صورت توأمان صورت می‌پذیرد. در ادامه ضمن بیان مسأله، رویکرد جذب داده شامل مفاهیم مربوطه و روش فیلتر کالمن و نسخه توسعه یافته آن تحت عنوان فیلتر آنسامبل کالمن مبتنی بر کار Evensen (2009) تشریح می‌گردد.

## ۲-۱- طرح تحلیلی مسأله

فرض کنید هدف پیش‌بینی یک متغیر (حالت) همچون رواناب، سطح آب زیرزمینی، تراز آب و یا رطوبت خاک باشد و در کنار بکارگیری یک مدل، از مشاهدات (داده‌ها) نیز برای بهبود پیش‌بینی‌ها استفاده شود. این بدان معناست که تخمین یک متغیر حالت همچون آبدهی رودخانه و یا رطوبت خاک از دو روش ممکن است. در روش اول می‌توان از یک مدل برای پیش‌بینی استفاده کرد، برای مثال با داشتن مقدار پارامترهای هواشناسی همچون بارش، دما و غیره، دانستن شرایط اولیه و مرزی و بکارگیری یک مدل بارش-رواناب می‌توان آبدهی رودخانه و یا رطوبت خاک را پیش‌بینی نمود. در روش دوم می‌توان با استفاده از اندازه‌گیری میدانی با استفاده از یک ایستگاه هیدرومتری برای آبدهی یا ایستگاه لایسیمتری برای رطوبت خاک یا سنجش از دور برای رطوبت خاک، متغیر حالت مورد نظر (آبدهی یا رطوبت خاک) را تخمین زد. همچنین می‌توان از ترکیب هر دو روش، که طبقاً دقت‌های متفاوت دارند، برای دستیابی به یک تخمین بهتر استفاده نمود.

در رویکرد ترکیبی بالا دو موضوع مهم شامل مدل و مشاهدات (اندازه‌گیری) پیش‌رو است و آنها را به ترتیب با معادلات زیر می‌توان نمایش داد:

$$\varphi^f = \varphi^i + p^f \quad (1)$$

$$d = \varphi^f + \varepsilon \quad (2)$$

که در آنها  $\varphi^f$  مقدار درست متغیر حالت مورد نظر است که معمولاً نامعلوم می‌باشد. همچنین  $\varphi^f$  پیش‌بینی مدل برای مقدار درست با خطای نامعلوم  $p^f$  و  $d$  اندازه‌گیری (مشاهده) از مقدار درست با خطای نامعلوم  $\varepsilon$  است. همچنین جملات خطای  $p^f$  و  $\varepsilon$  بصورت متغیرهای تصادفی نرمال هر دو با میانگین صفر، واریانس‌های به ترتیب برابر  $C_{\varphi\varphi}^f$  و  $C_{\varepsilon\varepsilon}$ ، و کوواریانس برابر صفر فرض می‌شوند. مقادیر واریانس مبتنی بر تحلیل حساسیت و با دانستن جزئیات مدل و روش اندازه‌گیری و عوامل غیرقطعی تأثیرگذار بر آنها تعیین می‌شوند. همانطور که

مشاهده می‌شود در اینجا با انگیزه بیان ساده و تحلیلی مسأله، فرض شده که معادله‌های مربوط به مدل و اندازه‌گیری بصورت خطی‌اند.

هدف یافتن بهترین تخمین تحلیل‌شده،  $\varphi^a$ ، از مقدار درست متغیر حالت با در نظر گرفتن توامان پیش‌بینی مدل و اندازه‌گیری است. اگر تخمین تحلیل‌شده به صورت یک ترکیب خطی از پیش‌بینی مدل و اندازه‌گیری مانند رابطه زیر در نظر گرفته شود:

$$\varphi^a = \alpha_1 \varphi^f + \alpha_2 d \quad (3)$$

و مشابه آنچه در مورد پیش‌بینی مدل و اندازه‌گیری ارائه شد، یک جمله خطای جمعی،  $p^a$ ، با توزیع نرمال، میانگین صفر و واریانس  $C_{\varphi\varphi}^a$  برای این تخمین تحلیل‌شده در نظر گرفته شود، از آنجا که بهترین تخمین همیشه باید نأریب و دارای حداقل واریانس باشد می‌توان بهترین مقادیر ضرایب خطی رابطه بالا را با هدف کمینه نمودن واریانس تخمین تحلیل‌شده محاسبه نمود و در نهایت بهترین تخمین تحلیل‌شده بصورت رابطه (۴) با واریانس بصورت رابطه (۵) خواهد بود (Evensen, 2009):

$$\varphi^a = \varphi^f + \frac{C_{\varphi\varphi}^f}{C_{\varphi\varphi}^f + C_{\varepsilon\varepsilon}} (d - \varphi^f) \quad (4)$$

$$C_{\varphi\varphi}^a = C_{\varphi\varphi}^f \left( 1 - \frac{C_{\varphi\varphi}^f}{C_{\varphi\varphi}^f + C_{\varepsilon\varepsilon}} \right) \quad (5)$$

## ۲-۲- فیلتر کالمن

اگر تحلیل و مساله پیش‌رو وابسته به زمان باشد، حل مساله نیازمند جذب داده متوالی خواهد بود؛ بدین صورت که در هر گام زمانی با استفاده از مدل، پیش‌بینی متغیر حالت انجام می‌شود و پس از دسترسی به مشاهدات جدید از متغیر حالت در همان گام زمانی، مقدار متغیر حالت بروزرسانی خواهد شد. بنابراین در هر گام زمانی دو مرحله پیش‌بینی و بروزرسانی باید انجام گردد. مرحله بروزرسانی در بخش قبل ضمن رابطه‌های (۴) و (۵) تشریح گردید و همانطور که مشخص است شناسایی و تحلیل خطای پیش‌بینی مدل و همچنین خطای اندازه‌گیری برای بروزرسانی مهم است. خطای اندازه‌گیری وابسته است به دقت روش و دستگاه اندازه‌گیری و معمولاً پیش‌فرض‌های ساده‌شده‌ای در مورد آن لحاظ می‌شود. اما خطای پیش‌بینی مدل موضوع مهمی است که تحت تأثیر عوامل متعدد است. در جذب داده این منابع مهم خطا و عدم قطعیت مدل به تفکیک در نظر گرفته می‌شوند.

آنچه Kalman (1960) برای اولین بار ارائه نمود یک معادله تحلیلی بیانگر خطای مدل به تفکیک منابع مختلف عدم قطعیت بود. وی مانند آنچه در بخش قبل ارائه شد، فرض کرد که خطای یک متغیر تصادفی نرمال است و از آنجا که تخمین باید نأریب باشد، میانگین خطا برابر

صفر است و تنها پارامتر نامعلوم این متغیر تصادفی که باید در مورد آن اطلاعات کسب شود (کو) واریانس است و با این انگیزه Kalman (1960) کوواریانس خطای پیش‌بینی مدل را به صورت یک معادله تحلیلی با شکل بسته ارائه نمود. بعدها روش وی تحت عنوان فیلتر کالمن نامگذاری شد. شایان ذکر است که فیلترینگ به صورت عام به معنای حذف یک ماده نامطلوب است و به صورت خاص در ادبیات پردازش سیگنال به معنای استفاده از مشاهدات برای بهبود وضعیت (متغیر حالت) یک سیستم (مدل) است. آنچه در واقع فیلتر می‌شود خطاست یعنی با در نظر گرفتن خطای پیش‌بینی مدل و مشاهدات بصورت جداگانه، تخمینی زده می‌شود که واریانس (خطای) آن از هر دوی تخمین بدست آمده از پیش‌بینی مدل و مشاهدات کمتر باشد و مطابق با معنای عام فیلترینگ، واریانس (خطا) کاهش می‌یابد و بصورت نسبی فیلتر یا حذف می‌شود.

Kalman (1960) عدم قطعیت پیش‌بینی مدل را به دو منبع مهم مرتبط دانست. منبع اول مربوط می‌شود به خطای ذاتی مدل و منبع دوم مربوط می‌شود به خطای استفاده از مقادیر تخمینی (و احتمالاً نادرست) متغیر حالت به جای مقادیر درست آن. در مورد منبع اول نکته مهم آن است که خطای ذاتی مدل از عدم قطعیت داده‌های ورودی، خطای ساختاری مدل، عدم قطعیت پارامترهای مدل و عدم قطعیت در شرایط مرزی و اولیه نشأت می‌گیرد که نسخه اولیه فیلتر کالمن آنها را به تفکیک در نظر نمی‌گیرد و همه آنها را در قالب یک جمله خطا جمع می‌نماید. شاید انگیزه کالمن برای تفکیک منبع دوم خطا از سایر منابع عدم قطعیت مدل، که در قالب منبع اول در نظر گرفته می‌شوند، را بتوان به جلوگیری از بروز خطای سیستماتیک حاصل از جایگزین نمودن مقادیر نادرست با مقادیر واقعی و درست متغیر حالت در گام زمانی قبل برای پیش‌بینی مقدار جدید متغیر حالت در گام زمانی مورد نظر مرتبط دانست؛ چرا که تکرار این خطا در طی فرآیند مدلسازی دینامیک می‌تواند باعث انباشت خطاهای از این نوع شده و در نهایت یک انحراف بزرگ در مدلسازی نسبت به واقعیت را بوجود آورد.

برای بیان تحلیلی معادله‌ای که کالمن ارائه کرد، در ابتدا مدل پیش‌بینی کننده با یک معادله دینامیکی خطی گسسته شده در زمان به صورت رابطه‌های زیر نمایش داده می‌شود:

$$\vec{\varphi}_k^t = G\vec{\varphi}_{k-1}^t + \vec{q}_{k-1} \quad (6)$$

$$\vec{\varphi}_0^t = \vec{\Psi}_0 + \vec{a} \quad (7)$$

که در آنها  $\vec{\varphi}_k^t$  بردار مقادیر درست متغیرهای حالت در گام زمانی  $k$ -ام،  $G$  ماتریس عملگر خطی مدل،  $\vec{q}_k$  بردار خطای ذاتی مدل در گام زمانی  $k$ -ام،  $\vec{\Psi}_0$  بردار شرایط اولیه متغیر حالت و  $\vec{a}$  بردار خطای مربوط به شرایط اولیه می‌باشد. همانطور که در رابطه (۶) معلوم است، بردار

مقادیر درست متغیر حالت در گام زمانی قبل درون مدل بکار گرفته می‌شوند تا بردار مقادیر درست متغیر حالت در گام زمانی مورد نظر بدست آید. از آنجا که دسترسی به مقادیر درست محتمل نیست، عملاً بهترین تخمین از بردار متغیر حالت در گام زمانی قبل بکار گرفته می‌شود تا بردار متغیر حالت در گام زمانی مورد نظر بصورت رابطه زیر پیش‌بینی شود:

$$\bar{\varphi}_k^f = G\bar{\varphi}_{k-1}^a \quad (8)$$

$$\bar{\varphi}_0^a = \bar{\varphi}_0^f \quad (9)$$

که در آنها  $\bar{\varphi}_k^f$  و  $\bar{\varphi}_k^a$  به ترتیب بردار مقادیر بهترین تخمین تحلیل شده و بردار مقادیر پیش‌بینی متغیرهای حالت در گام زمانی  $k$ -ام می‌باشد. با ادغام رابطه‌های (۶) و (۸) خطای پیش‌بینی مدل به صورت رابطه‌ی زیر محاسبه می‌شود:

$$\bar{\varphi}_k^t - \bar{\varphi}_k^f = G(\bar{\varphi}_{k-1}^t - \bar{\varphi}_{k-1}^a) + \bar{q}_{k-1} \quad (10)$$

رابطه بالا بیان می‌کند که خطای پیش‌بینی مدل از مجموع خطای ذاتی مدل و خطای حاصل از جایگزین نمودن مقدار تخمین تحلیل شده به جای مقدار درست بدست می‌آید. با استفاده از این رابطه ماتریس کوواریانس خطای پیش‌بینی مدل به صورت رابطه زیر است:

$$C_{\varphi\varphi,k}^f = GC_{\varphi\varphi,k-1}^a G^T + C_{qq,k-1} \quad (11)$$

که در آن ماتریس کوواریانس خطای پیش‌بینی مدل در گام زمانی  $k$ -ام،  $C_{qq,k-1}$  ماتریس کوواریانس خطای ذاتی مدل در گام زمانی قبل و  $C_{\varphi\varphi,k-1}^a$  ماتریس کوواریانس خطای تخمین مربوط به گام زمانی قبل است.

همانطور که پیش‌تر بیان شد الگوریتم‌های جذب داده مانند فیلتر کالمن شامل دو مرحله مهم پیش‌بینی و بروزرسانی‌اند. رابطه‌های (۱۰) و (۱۱) درحقیقت مربوط می‌شوند به مرحله پیش‌بینی در فیلتر کالمن. اما بروزرسانی در فیلتر کالمن که منتج می‌شود به بهترین تخمین‌های تحلیل شده متغیر حالت مبتنی است بر تئوری بیان شده در بخش قبل و بنابراین با استفاده از رابطه‌های (۴) و (۵) انجام می‌گردد. اگر بخواهیم رابطه‌های اشاره شده را در شکل برداری ارائه کنیم باید رابطه (۲)، مربوط به اندازه‌گیری، را به صورت زیر بازنویسی کنیم:

$$\bar{d}_k = M\bar{\varphi}_k^t + \bar{\varepsilon}_k \quad (12)$$

که در آن  $\bar{d}_k$  بردار مشاهدات گام زمانی  $k$ -ام،  $M$  ماتریس تبدیل بردار متغیر حالت به بردار مشاهدات و  $\bar{\varepsilon}_k$  بردار خطای اندازه‌گیری در گام زمانی  $k$ -ام است. در مورد ماتریس  $M$  شایان ذکر است که در جذب داده ممکن است که مشاهداتی که جذب مدل می‌شود هم‌جنس و هم‌بعد با متغیر حالتی که می‌خواهد پیش‌بینی و در نهایت بروزرسانی شود، نباشد. برای مثال ممکن است در مدل‌سازی بارش-رواناب مشاهدات رواناب برای پیش‌بینی رطوبت خاک و یا بالعکس جذب شوند. در چنین مواردی باید رابطه مشاهدات و متغیرهای حالت در

مدل‌سازی در نظر گرفته شود. این رابطه در حالت کلی می‌تواند بصورت یک تابع غیرخطی باشد اما آنچه در فیلتر کالمن فرض می‌شود آن است که این ارتباط به صورت یک تابع خطی قابل بیان است که البته جزء محدودیت‌ها و ملاحظات مهم مربوط به فیلتر کالمن است. در نهایت می‌توان مرحله بروزرسانی را مشابه قبل با استفاده از رابطه‌های زیر بیان نمود:

$$\bar{\varphi}_k^a = \bar{\varphi}_k^f + K_k(\bar{d}_k - M\bar{\varphi}_k^f) \quad (13)$$

$$C_{\varphi\varphi,k}^a = C_{\varphi\varphi,k}^f(I - K_kM) \quad (14)$$

$$C_{\varphi\varphi,0}^a = C_{aa} \quad (15)$$

$$K_k = C_{\varphi\varphi,k}^f M^T (MC_{\varphi\varphi,k}^f M^T + C_{\varepsilon\varepsilon,k})^{-1} \quad (16)$$

که در آنها  $C_{\varphi\varphi,k}^a$  ماتریس کوواریانس خطای بهترین تخمین تحلیل شده مربوط به گام زمانی  $k$ -ام،  $C_{\varepsilon\varepsilon,k}$  ماتریس کوواریانس خطای اندازه‌گیری مربوط به گام زمانی  $k$ -ام،  $C_{aa}$  ماتریس کوواریانس خطای مقدار اولیه متغیر حالت و  $K_k$  ماتریس بهره کالمن است.

فیلتر کالمن را به صورت خلاصه در قالب روال محاسباتی زیر می‌توان ارائه کرد:

- ۱- بهترین تخمین تحلیل شده متغیرهای حالت در گام زمانی صفرام و ماتریس کوواریانس خطای آن با استفاده از رابطه (۹) و رابطه (۱۵) محاسبه می‌شود.
- ۲- محاسبات از گام زمانی اول آغاز می‌شود و به صورت متوالی با زمان تا رسیدن به آخرین گام زمانی ادامه می‌یابد. در هر گام زمانی دو مرحله‌ی پیش‌بینی و بروزرسانی به ترتیب انجام می‌شوند.
- ۳- پیش‌بینی: بردار پیش‌بینی متغیر حالت و کوواریانس خطای آن با استفاده از رابطه‌های (۸) و (۱۱) محاسبه می‌شود.
- ۴- بروزرسانی: بردار بهترین تخمین تحلیل شده متغیر حالت و ماتریس خطای کوواریانس آن به ترتیب با استفاده از رابطه‌های (۱۳)، (۱۴) و (۱۶) محاسبه می‌شود.

### ۲-۳- فیلتر آنسامبل کالمن

فیلتر کالمن حاوی فرض‌های مهم و ساده‌کننده‌ای است. در این روش مدل دینامیکی مورد استفاده برای پیش‌بینی باید خطی باشد. همچنین معادله مورد استفاده برای ترکیب پیش‌بینی و اندازه‌گیری، خطی در نظر گرفته می‌شود. جملات خطای پیش‌بینی و اندازه‌گیری نیز به صورت حاصل جمعی و خطای ذاتی مدل به صورت مجزا از ساختار مدل لحاظ می‌گردد. در فیلتر کالمن، خطا به صورت متغیر تصادفی نرمال بوده و بین خطای اندازه‌گیری و خطای پیش‌بینی مدل استقلال خطی در نظر گرفته می‌شود. همچنین برای محاسبه خطای پیش‌بینی مدل فرض می‌شود که خطای جایگزینی مقدار تخمینی به جای مقدار درست از

خطای ذاتی مدل مستقل است. علاوه بر آن، برای محاسبه واریانس خطا از روابط تحلیلی استفاده می‌شود و حتی در فیلتر کالمن توسعه یافته برای آنکه بتوان روابط تحلیلی مربوطه را استخراج نمود از بسط تیلور به صورت تقریبی استفاده می‌شود، که این خود یک منشأ خطا و عدم قطعیت است. همچنین یکی از الزامات فیلتر کالمن آن است که باید روابط مربوط به مدل به صورت تحلیلی و در شکل بسته در دسترس باشند. فیلتر کالمن از پیش‌بینی مرکزی برای متغیرهای حالت استفاده می‌کند که تعبیر آماری خاصی ندارد. همچنین تغییرات این مقادیر پیش‌بینی مرکزی و همچنین مقادیر اندازه‌گیری را حول مقدار درست، که البته نامعلوم ولی قطعی است، در نظر می‌گیرد. به همین دلیل فیلتر کالمن و نسخه توسعه یافته آن را تحت عنوان روش‌های قطعی در ادبیات موضوع می‌شناسند.

با هدف مرتفع نمودن برخی از فرضیات ساده‌کننده اشاره شده، فیلتر آنسامبل کالمن (EnKF) اولین بار توسط (Evensen (1994 ارائه شد. در EnKF مدل پیش‌بینی دینامیکی می‌تواند غیرخطی باشد. همچنین جملات خطا لزوماً به صورت حاصل جمعی نبوده و می‌توان خطای ذاتی مدل را در درون ساختار مدل و نه لزوماً مجزای از آن در نظر گرفت و بدین جهت می‌توان اثر منابع مختلف عدم قطعیت بر خطای ذاتی مدل را لحاظ نمود. در EnKF لزومی ندارد که روابط تحلیلی مدل در شکل بسته در دسترس باشد و همچنین بجای آنکه با روابط تحلیلی کوواریانس کار کند با نمونه‌های آماری از متغیرهای حالت یا به تعبیر دیگر با آنسامبل آن کار می‌کند. برای کار کردن با آنسامبل نیاز به نمونه‌گیری یا همان شبیه‌سازی مونت-کارلو وجود دارد و در نتیجه خطای نمونه‌گیری وجود خواهد داشت که با افزایش اندازه آنسامبل میزان این خطا کاهش خواهد یافت. استفاده از شبیه‌سازی مونت-کارلو در EnKF در مقایسه با فیلتر توسعه یافته کالمن (EKF) باعث کاهش زمان حل شده و از پیچیدگی‌های حاکم بر مسأله نیز می‌کاهد. EnKF از میانگین نمونه، که دارای تعبیر آماری مهمی به صورت بهترین تخمین است، بجای مقدار پیش‌بینی مرکزی استفاده می‌کند و آنسامبل گسترش یافته حول میانگین را به صورت تعریف طبیعی خطای میانگین نمونه در نظر می‌گیرد. به همین دلیل فیلتر آنسامبل کالمن در ادبیات موضوع به صورت یک روش احتمالاتی شناخته می‌شود. بنابراین بجای نسخه اولیه ماتریس‌های کوواریانس حول مقدار درست نامعلوم یعنی  $C_{\varphi\varphi}^f = (\varphi^f - \varphi^t)(\varphi^f - \varphi^t)^T$  و آنسامبل ماتریس‌های کوواریانس حول آنسامبل از نسخه آنسامبل  $C_{\varphi\varphi}^a = (\varphi^a - \varphi^t)(\varphi^a - \varphi^t)^T$  و آنسامبل ماتریس‌های کوواریانس حول میانگین آنسامبل معلوم یعنی  $(C_{\varphi\varphi}^e)^f = (\varphi^f - \bar{\varphi}^f)(\varphi^f - \bar{\varphi}^f)^T$  و  $(C_{\varphi\varphi}^e)^a = (\varphi^a - \bar{\varphi}^a)(\varphi^a - \bar{\varphi}^a)^T$  استفاده می‌شود که در آن

$\varphi^f$  و  $\varphi^a$  ماتریس‌های آنسامبل،  $\bar{\varphi}^f$  و  $\bar{\varphi}^a$  بردارهای میانگین‌های آنسامبل و  $(C_{\varphi\varphi}^e)^f$  و  $(C_{\varphi\varphi}^e)^a$  ماتریس‌های کوواریانس آنسامبل به ترتیب مربوط به پیش‌بینی و تخمین تحلیل شده می‌باشد. برخی از فرضیات مهم فیلتر کالمن همچنان در نسخه آنسامبل آن نیز وجود دارد. EnKF نیز برای ترکیب پیش‌بینی و اندازه‌گیری از یک ترکیب خطی استفاده می‌کند و همچنان جملات خطا به صورت متغیر تصادفی نرمال در نظر گرفته می‌شود.

مراحل روش EnKF به صورت زیر است:

۱- با استفاده از شبیه‌سازی (مونت-کارلو) یک آنسامبل (نمونه تصادفی)  $n$  عضوی از مقدار اولیه بردار متغیر حالت با فرض معلوم بودن میانگین و کوواریانس تولید می‌شود.

۲- در هر گام زمانی در صورت وجود مشاهدات دو مرحله پیش‌بینی و برورسانی به صورت پشت سرهم و در صورت عدم وجود مشاهدات صرفاً پیش‌بینی انجام می‌گردد. در ادامه این دو مرحله توضیح داده شده‌اند.

۳- پیش‌بینی: در این مرحله آنسامبل مربوط به پیش‌بینی بردار حالت و کواریانس آن به صورت زیرمرحله‌های زیر محاسبه می‌گردد:

- فرض بر این است که آنسامبل تخمین تحلیل شده‌ی بردار متغیر حالت مربوط به گام زمانی قبل، که به صورت ماتریس  $[(\varphi_{k-1}^a)_{i,j}]_{m \times n}$  است، در ابتدا وجود دارد که در آن  $k-1$  اندیس گام زمانی قبل،  $i$  اندیس مربوط اعضای بردار متغیرحالت و در عین حال اندیس سطری درایه‌های ماتریس،  $j$  اندیس مربوط اعضای آنسامبل و در عین حال اندیس ستونی درایه‌های ماتریس،  $m$  تعداد متغیرهای حالت و  $n$  تعداد اعضای آنسامبل است. به عبارت دیگر هر ستون این ماتریس شامل یک ترکیب تصادفی تولید شده برای بردار متغیر حالت است.

- با استفاده از شبیه‌سازی و مفروض به داشتن خصوصیات آماری خطای ذاتی مدل، یک آنسامبل برای خطای ذاتی مدل،  $q_{k-1,j}$ ، به صورت تصادفی تولید می‌شود که در آن  $j$  اندیس مربوط به هر یک از اعضای آنسامبل است.

- با داشتن آنسامبل‌های مورد اشاره در بالا و با استفاده از مدل پیش‌بینی، یک آنسامبل برای پیش‌بینی بردار متغیر حالت در گام زمانی مورد نظر با استفاده از رابطه زیر بدست می‌آید که در آن  $\bar{\varphi}_{k-1,j}^a$  در حقیقت ستون  $j$ -ام از ماتریس  $[(\varphi_{k-1}^a)_{i,j}]_{n \times m}$ ،  $\bar{\varphi}_{k,j}^f$  یک عضو از آنسامبل بردار پیش‌بینی، یعنی ماتریس  $[(\varphi_k^f)_{i,j}]_{n \times m}$ ، یا به عبارت دیگر ستون  $j$ -ام این ماتریس و  $G(\cdot)$  مدل پیش‌بینی دینامیکی است:

$$\bar{\varphi}_{k,j}^f = G(\bar{\varphi}_{k-1,j}^a; q_{k-1,j}) \quad , \forall j \quad (17)$$

- آنسامبل خطای پیش‌بینی به صورت رابطه زیر است:

دارای پارامترهای زیادی است که قبل از کاربرد مدل باید تخمین زده شوند.

با توجه به رایگان بودن و همچنین امکان دسترسی آزاد به کدهای مرجع مدل SWAT از یک طرف و قابلیت‌های مناسب آن در مدل‌سازی فرآیندهای مرتبط با مدیریت حوضه آبریز از طرف دیگر، این مدل در دهه اخیر توسط محققین در سرتاسر جهان برای بررسی مسائل گوناگون در این حوضه از مطالعه بکار رفته است. در این پژوهش به منظور شبیه‌سازی هیدرولوژیکی از مدل SWAT استفاده شد. به منظور واسنجی مدل از الگوریتم واسنجی خودکار SUFI2 و الگوریتم جذب داده EnKF بطور جداگانه استفاده شد و نتایج حاصل از اجرای دو الگوریتم مقایسه گردید. در ادامه اجزاء مهم روش‌شناسی شامل معرفی مدل SWAT و همچنین نحوه جذب داده رواناب در مدل SWAT با استفاده از الگوریتم EnKF تشریح می‌شود. در انتهای این بخش ضمن معرفی مطالعه موردی برخی ملاحظات کاربرد مدل مطرح می‌شود.

### ۳-۲- اتصال SWAT به EnKF

در این تحقیق از روش تخمین متصل متغیر حالت-پارامتر برای پیش‌بینی و واسنجی به صورت توأمان استفاده می‌شود به این معنا که پارامترها به عنوان جزئی از بردار متغیر حالت در نظر گرفته شده که آن را بردار حالت تکمیل شده می‌خوانند. همچنین با توجه به آنکه هدف جذب مشاهدات رواناب برای انجام توأمان پیش‌بینی رواناب و واسنجی پارامترهای مدل است، متغیر رواناب به همراه  $m$  پارامتر مهم مدل در قالب بردار (متغیر) حالت تکمیل شده در نظر گرفته می‌شود و مشابه پژوهش Clark et al. (2008)، ماتریس تبدیل بردار متغیر حالت به بردار مشاهدات به صورت  $M = [1, 0_{1 \times m}]$  خواهد بود. در این بردار افقی تنها درایه‌ی با مقدار ۱ متناظر با درایه مربوط به رواناب در بردار تکمیل شده و درایه‌های با مقدار صفر متناظر با درایه‌های مربوط به پارامترها در بردار تکمیل شده می‌باشند.

در این تحقیق اندازه آنسامبل برابر با ۲۰۰ عضو در نظر گرفته شد. برای پخش آماری از توزیع نرمال با میانگین صفر استفاده شد. مقدار انحراف معیار جمله خطا به نحوی تعیین گردید که هیچکدام از اعضای آنسامبل پارامترها خارج از محدوده معنی‌دار فیزیکی قرار نگیرند. همچنین برای در نظر گرفتن خطای مشاهدات بارش و خطای ذاتی یا درونی مدل در این تحقیق مقادیر مشاهدات بارش به عنوان مهمترین ورودی مدل و همچنین منشأ عدم قطعیت درونی مدل، پخش آماری می‌شوند.

$$(\varphi_k^f - \overline{\varphi_k^f}) = [\varphi_{k,1}^f - \overline{\varphi_k^f}, \dots, \varphi_{k,n}^f - \overline{\varphi_k^f}]_{m \times n} \quad (18)$$

- کوواریانس خطای پیش‌بینی در گام زمانی پیش رو به صورت رابطه زیر است:

$$(C_{\varphi\varphi,k}^e)^f = \overline{(\varphi_k^f - \overline{\varphi_k^f})(\varphi_k^f - \overline{\varphi_k^f})^T} \quad (19)$$

۴- بروزرسانی: در این مرحله آنسامبل مربوط به (بهترین) تخمین تحلیل شده‌ی بردار متغیر حالت و ماتریس خطای کوواریانس آن به صورت زیر مرحله‌های زیر محاسبه می‌شود:

- با استفاده از شبیه‌سازی و مفروض به داشتن خصوصیات آماری یک آنسامبل برای خطای اندازه‌گیری،  $\varepsilon_k = [(\tilde{\varepsilon}_k)_z]$ ، به صورت تصادفی تولید می‌شود و با استفاده از آن یک آنسامبل برای اندازه‌گیری به صورت زیر بدست می‌آید که در آن  $\vec{d}_k$  بردار مشاهدات در گام زمانی مورد نظر است:

$$\vec{d}_{k,j} = \vec{d}_k + \tilde{\varepsilon}_{k,j}, \quad \forall j \quad (20)$$

- نسخه آنسامبل ماتریس کوواریانس خطای اندازه‌گیری به صورت زیر محاسبه می‌شود:

$$C_{\varepsilon\varepsilon,k}^e = \overline{\varepsilon_k \varepsilon_k^T} \quad (21)$$

- آنسامبل بهترین تخمین تحلیل‌شده‌ی بردار متغیر حالت به همراه میانگین و واریانس آن و همچنین ماتریس بهره کالمن به ترتیب با استفاده از رابطه‌های زیر محاسبه می‌شود:

$$\varphi_{k,j}^a = \varphi_{k,j}^f + K_{e,k} \times (d_{k,j} - M\varphi_{k,j}^f), \quad \forall j \quad (22)$$

$$\overline{\varphi_k^a} = \overline{\varphi_k^f} + K_{e,k} \times (\overline{d_k} - M\overline{\varphi_k^f}) \quad (23)$$

$$(C_{\varphi\varphi,k}^e)^a = (I - K_{e,k}M) \times (C_{\varphi\varphi,k}^e)^f \quad (24)$$

$$K_{e,k} = (C_{\varphi\varphi,k}^e)^f \times M^T \times (M(C_{\varphi\varphi,k}^e)^f M^T + C_{\varepsilon\varepsilon,k}^e)^{-1} \quad (25)$$

### ۳- روش‌شناسی

#### ۳-۱- مدل SWAT

SWAT یک مدل هیدرولوژیکی نیمه‌توزیعی و مبتنی بر فرآیند است که توانایی شبیه‌سازی فرآیندهای هیدرولوژیکی و بطور خاص بارش-رواناب را دارد و امروزه جزء مدل‌های پرکاربرد در هیدرولوژی، کشاورزی و مدیریت منابع آب محسوب شده و در تحقیقات گذشته از آن به منظور ارزیابی اثرات اقدامات مربوط به مدیریت حوضه آبریز بر سیستم‌های هیدرولوژیکی، کشاورزی و کیفیت آب استفاده شده است. مدل SWAT امکان در نظر گرفتن تغییرات مکانی متغیرهایی مانند کاربری زمین، نوع خاک، توپوگرافی و متغیرهای آب و هوایی را دارد. همچنین این مدل توانایی شبیه‌سازی پدیده‌ها را، به صورت پیوسته‌ی زمانی و برای یک دوره‌ی بلند مدت، دارد. مدل SWAT فرآیندهای هیدرولوژیکی را با استفاده از معادلات بیلانی ساده مدل می‌نماید و

بیطاس) در مدل SWAT عبارتند از: مدل رقومی ارتفاع (DEM) با تفکیک مکانی ۳۰ متر مطابق شکل ۱، نقشه کاربری اراضی حوزه مطابق شکل ۱، نقشه‌ی خاک‌شناسی حوزه مربوط به سازمان FAO مطابق شکل ۱، و داده‌های هواشناسی شامل سری زمانی بارش و دمای روزانه مربوط به ایستگاه‌های باران‌سنجی واقع در منطقه و همچنین داده‌های بازتحلیل شده‌ی مربوط به پروژه CFSR سازمان NCEP.

#### ۴- نتایج و بحث

تخمین پارامترهای مدل SWAT برای شبیه‌سازی در مقیاس زمانی ماهانه برای دوره ۱۵ ساله از سال ۱۹۹۷ تا ۲۰۱۱ میلادی انجام شد. چهار معیار  $NS$ ،  $PBIAS$ ،  $RSR$  و  $KGE$  برای بررسی دقت شبیه‌سازی‌ها استفاده شده است. معیار  $NS$  نشان‌دهنده‌ی میزان نزدیک بودن مقادیر شبیه‌سازی‌شده به مقادیر مشاهده‌ای است. حد بالای معیار مذکور برابر با یک بوده و حد پایینی برای این معیار وجود ندارد. در صورت تطابق کامل مشاهدات و نتایج شبیه‌سازی، مقدار  $NS$  برابر با یک خواهد بود. مقدار  $PBIAS$  بیانگر درصد اُریب بودن نتایج شبیه‌سازی نسبت به مشاهدات بوده و بازه تغییرات آن نامحدود  $(-\infty, \infty)$  است. مقادیر منفی این معیار نمایشگر کمتر بودن نتایج شبیه‌سازی از مشاهدات و در مورد مقادیر مثبت به عکس است؛ بنابراین بهترین مقدار این معیار برابر با صفر است.  $RSR$  شکل استاندارد شده‌ی  $RMSE$  بوده و بنابراین هرچه خطا کمتر باشد مقدار آن به صفر نزدیکتر است. معیار  $KGE$  در برگیرنده اثر میانگین و واریانس خطا و مقدار ضریب تعیین رگرسیون خطی بوده و هر چه مقدار آن به عدد یک نزدیکتر باشد، تطابق مشاهدات و نتایج شبیه‌سازی بیشتر است.

برای انجام روند جذب داده، الگوریتم EnKF در محیط نرم‌افزار MATLAB برنامه‌نویسی و مدل SWAT در همین محیط فراخوانی شد. از طرفی اجرای متوالی و گام به گام مدل SWAT درون حلقه تکرار الگوریتم EnKF زمانی میسر خواهد بود که متغیر رطوبت خاک و سایر متغیرهای مربوط به ذخیره (برف، مخزن، آبخوان و غیره) با گذر زمان و به صورت گام به گام به مدل معرفی شود. این درحالی است که چنین امکانی برای متغیر رطوبت خاک، که دارای نقش کلیدی است، در حال حاضر در مدل SWAT فراهم نیست. به این ترتیب اجرای متوالی یا گام به گام مدل SWAT با محدودیت مواجه است. برای رفع این مشکل و با هدف دسترسی به مقدار شبیه‌سازی شده‌ی رواناب در هر گام زمانی، مدل از ابتدا تا گام زمانی مورد نظر اجرا شده و خروجی مدل در آن گام زمانی به عنوان مقدار پیش‌بینی شده به الگوریتم EnKF معرفی گردیده و مقدار پیش‌بینی شده، در خلال مراحل اجرای الگوریتم بروزسانی می‌شود. این کار البته سبب افزایش بیش از اندازه زمان محاسبات می‌شود و پیشنهاد می‌شود که در آینده اصلاحاتی در کدهای مرجع SWAT انجام گردد.

#### ۳-۳- مطالعه موردی

حوضه‌ی آبریز رودخانه‌ی مه‌آباد چای، از زیر حوضه‌های دریاچه ارومیه منتهی به ایستگاه بیطاس به مساحت ۲۴۲ کیلومتر مربع، به عنوان مطالعه‌ی موردی این تحقیق انتخاب شد. دو شاخه اصلی رودخانه‌ی مه‌آباد چای از کوه‌های واقع در ۴۵ کیلومتری جنوب غربی شهر مه‌آباد سرچشمه گرفته و با طی مسیری، به دریاچه سد مه‌آباد وارد شده و رودخانه مه‌آباد را تشکیل می‌دهند. تأثیر برداشتهای انسان از رودخانه در بازه رودخانه منتهی به ایستگاه بیطاس ناچیز است. داده‌های مورد استفاده برای مدل‌سازی رودخانه مه‌آباد (در بازه منتهی به ایستگاه

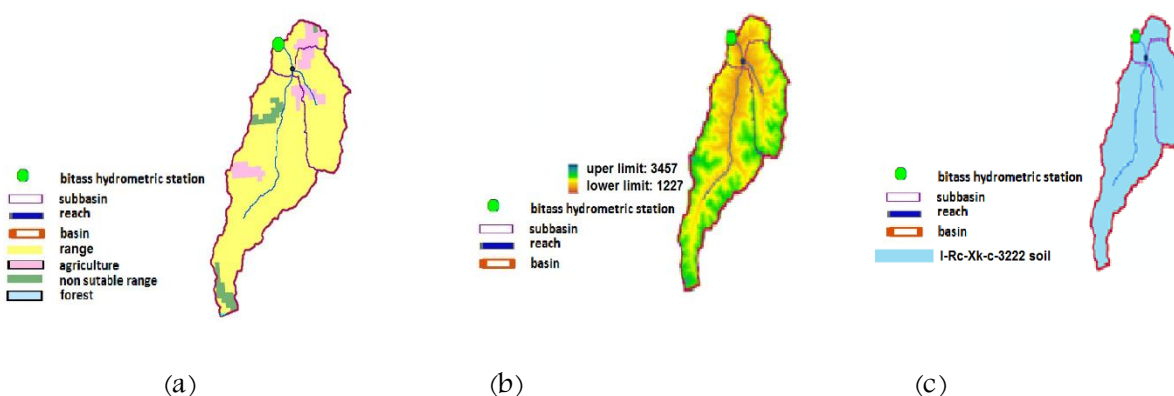


Fig. 1- Maps of Mahabad River Basin- upstream of Bitass hydrometry station: a) land use, b) DEM, c) soil types

شکل ۱- لایه‌های اطلاعاتی رودخانه مه‌آباد منتهی به ایستگاه بیطاس؛ (a) کاربری اراضی، (b) DEM، (c) خاک‌شناسی



مبتنی بر نتایج تحلیل حساسیت پارامترها با توجه به معیارهای آماری p-value و t-state و همچنین با توجه به اظهارات تحقیقات پیشین در خصوص حساسیت پارامترهای مدل SWAT در نهایت ۱۴ پارامتر به عنوان مهمترین پارامترهای واسنجی مشخص گردید. مشخصات این پارامترها شامل نام اختصاری بر طبق استاندارد ارائه شده توسط (Abbaspour et al. (2007)، بازه اولیه تغییرات مقدار آن و بازه نهایی (نتایج SUFI2) می‌باشد، در جدول ۱ ارائه شده است.

شایان ذکر است به دلیل آنکه در پژوهش حاضر تخمین پارامترها به صورت توده‌ای انجام پذیرفت، پارامترهایی که با توجه به لایه‌های اطلاعات مکانی مورد استفاده در فرآیند مدل‌سازی، دارای تغییرات مکانی بوده‌اند با پیشوند v تخمین زده شده‌اند. به عنوان نمونه پارامتر CN<sub>2</sub> به دلیل تغییرات مکانی با پیشوند v واسنجی شده است. در خصوص پارامترهای SOL\_BD و SOL\_AWC، با توجه به ثابت بودن نوع خاک در لایه اطلاعات مکانی مورد استفاده در فرآیند مدل‌سازی، انتخاب پیشوند v یا r تأثیری در نتایج واسنجی نخواهد داشت.

روند متوالی تخمین پارامترهای مدل با استفاده از الگوریتم EnKF در شکل ۲ قابل مشاهده است. توضیح مهم در مورد شکلها اینکه محدوده‌ای که با خط پررنگ (قرمز) و خط کم رنگ (سبز) نمایش داده شود به ترتیب نشان‌دهنده محدوده ۸۰٪ و ۹۵٪ آنسامیل پارامتر مورد نظر است. شایان ذکر است که در جذب داده، پارامترها متغیر با زمان در نظر گرفته نمی‌شوند و شکل‌هایی مانند آنچه به آن اشاره شد، صرفاً نمایش‌دهنده روند بروزرسانی مقدار پارامتر با استفاده از مشاهدات جدیدی است که در هر گام زمانی در دسترس قرار می‌گیرد.

در کنار معیارهای نکویی برازش مذکور که عمدتاً به منظور بررسی تطابق دو سری زمانی شبیه‌سازی و مشاهدات است، دو معیار p-factor و r-factor نمایشگر خصوصیات آنسامیل سری‌های زمانی حاصل از شبیه‌سازی در مقایسه با سری زمانی مشاهدات است.

به طور خلاصه می‌توان گفت p-factor درصد مشاهدات موجود در محدوده‌ی ۹۵PPU و r-factor نماینده میزان ضخامت ۹۵PPU است. ۹۵PPU هم محدوده‌ای است که ۹۵٪ نتایج شبیه‌سازی در آن قرار می‌گیرد. با این توضیح هر چه مقدار p-factor به عدد یک نزدیکتر و مقدار r-factor به عدد صفر نزدیکتر باشد مطلوب‌تر است.

الگوریتم SUFI2 چهار بار اجرا گردید و در هر بار یکی از چهار معیار NS و PBIAS و RSR و KGE به عنوان تابع هدف استفاده شد که نتایج مربوط به آن به ترتیب با علایم اختصاری SUFI2-NS، SUFI2-PBIAS، SUFI2-RSR و SUFI2-KGE نشان داده شده است. در مورد الگوریتم EnKF از معیارهای چهارگانه‌ی NS، PBIAS، RSR و KGE صرفاً به منظور ارزیابی نتایج استفاده شد. نتایج اجرای الگوریتم EnKF به صورت دو سری زمانی و با استفاده از چهار معیار نکویی برازش مذکور ارزیابی شد. سری زمانی اول که با EnKF-MUF نمایش داده شده، میانگین آنسامیل بروزرسانی شده خروجی مدل SWAT در هر گام زمانی است. سری زمانی دوم که با EnKF-MP نمایش داده شده است، مربوط می‌شود به شبیه‌سازی مدل SWAT با در نظر گرفتن میانگین آنسامیل پارامترهای بروزرسانی شده. هر دو سری زمانی مذکور با استفاده از یک اجرای الگوریتم EnKF به دست آمده است.

**Table 1- Parameter ranges before and after calibration**

**جدول ۱- محدوده پارامترها قبل و بعد از واسنجی**

Parameter	Before calibration	After calibration
v__CN2.mgt	[78, 98]	[74.21, 90.07]
r__SOL_AWC().sol	[-0.5, 0.5]	[-0.12, 0.5]
v__GW_REVAP.gw	[0.02, 0.2]	[0.03, 0.14]
v__REVAPMN.gw	[0, 800]	[209.92, 630.08]
v__GWQMN.gw	[0.0, 5000]	[1512.01, 4537.9]
v__RCHRG_DP.gw	[0, 1]	[0, 0.59]
v__CH_K1.sub	[0, 300]	[0, 160.08]
r__SOL_BD().sol	[-0.4, 0.4]	[-0.08, 0.4]
v__GW_DELAY.gw	[0, 500]	[41.7, 347.3]
v__SUB_SFTMP().sno	[-20, 20]	[-15.83, 8.07]
v__SUB_SMFMX().sno	[0, 20]	[0, 10.79]
v__SUB_SMFMN().sno	[0, 20]	[7.13, 14.17]
v__SUB_SMTMP().sno	[-20, 20]	[-8.5, 14.5]
v__SUB_TIMP().sno	[0, 1]	[0.03, 0.68]

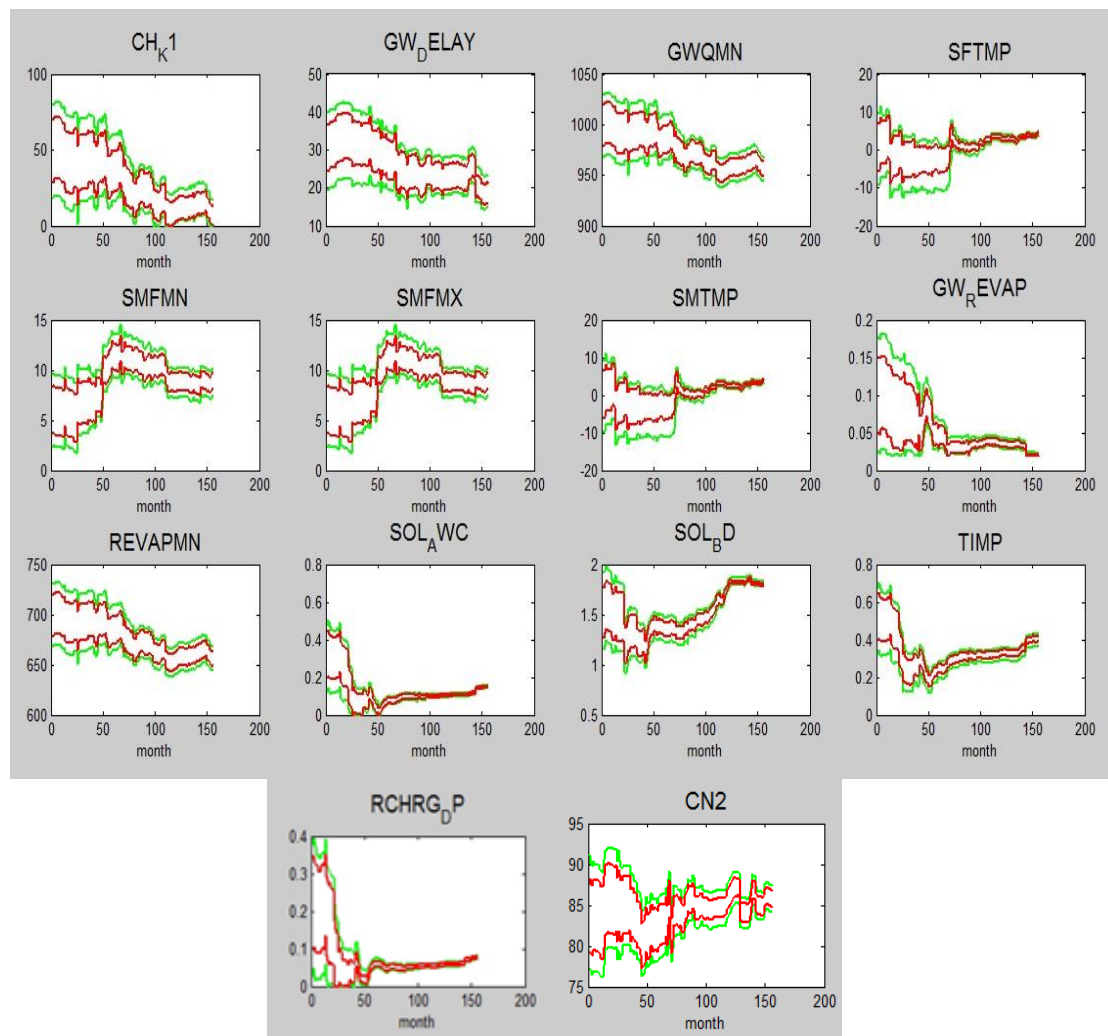


Fig. 2- Sequential parameters estimation using EnKF

شکل ۲- تخمین متوالی پارامترها با استفاده از EnKF

بهتر از SUFI2-NS عمل کرده است. بنابراین می‌توان نتیجه گرفت که اجرای SUFI2 با تابع هدف KGE نسبت به سه تابع هدف دیگر منجر به نتایج بهتری شده است. در مورد دو سری زمانی مربوط به الگوریتم EnKF اختلاف مقادیر معیارهای NS، PBIAS، RSR و KGE آنها در مقایسه کمتر از ۵ درصد بوده و بنابراین معنی‌دار نیست. سری‌های زمانی مشاهدات، بهترین تخمین و باند عدم قطعیت مربوط به دو الگوریتم در شکل ۳ قابل مشاهده است. توضیح آنکه، در بخش‌هایی از شکل که مربوط به اجرای الگوریتم SUFI2 است، صرفاً نتایج مربوط به بهترین روش از بین ۴ روش متناظر ارائه شده است. مقایسه نتایج دو الگوریتم مبتنی بر معیارهای NS و RSR حاکی از وجود تفاوت معنی‌دار (در حدود ۰/۲) میان نتایج SUFI2-KGE و EnKF-MUF بوده و از آنجا که نتایج این دو روش با توجه به معیارهای PBIAS و KGE تفاوت معنی‌داری ندارد، می‌توان نتیجه گرفت که در مجموع در شبیه‌سازی ماهانه عملکرد EnKF بهتر بوده

با این توجه انتظار می‌رود روند تخمین به صورتی باشد که تخمین پارامترها به مرور زمان بهبود یابد، یعنی واریانس مقادیر پارامترها کاهش یابد. بنابراین انتظار آن است که علاوه بر کاهش عرض محدوده‌های ۸۰٪ و ۹۵٪ نوسانات زمانی نیز کاهش یابد. با این توجه مشخص است که تخمین برخی از پارامترها قابل قبول‌تر است.

ارزیابی نتایج اجرای الگوریتم‌های SUFI2 و EnKF مبتنی بر معیارهای نکویی برازش مذکور در جدول ۲ ارائه شده است. همانطور که مشاهده می‌گردد نتایج مربوط به روش SUFI2-PBIAS از لحاظ معیارهای NS و RSR قابل قبول نیست و نتایج مربوط به دو روش SUFI2-NS و SUFI2-RSR یکسان است. با توجه به معیار NS نتایج دو روش SUFI2-KGE و SUFI2-NS دارای دقت بالایی بوده هر چند نتایج روش SUFI2-NS در مقایسه بهتر به نظر می‌رسد. همچنین مبتنی بر معیارهای PBIAS و KGE روش SUFI2-KGE

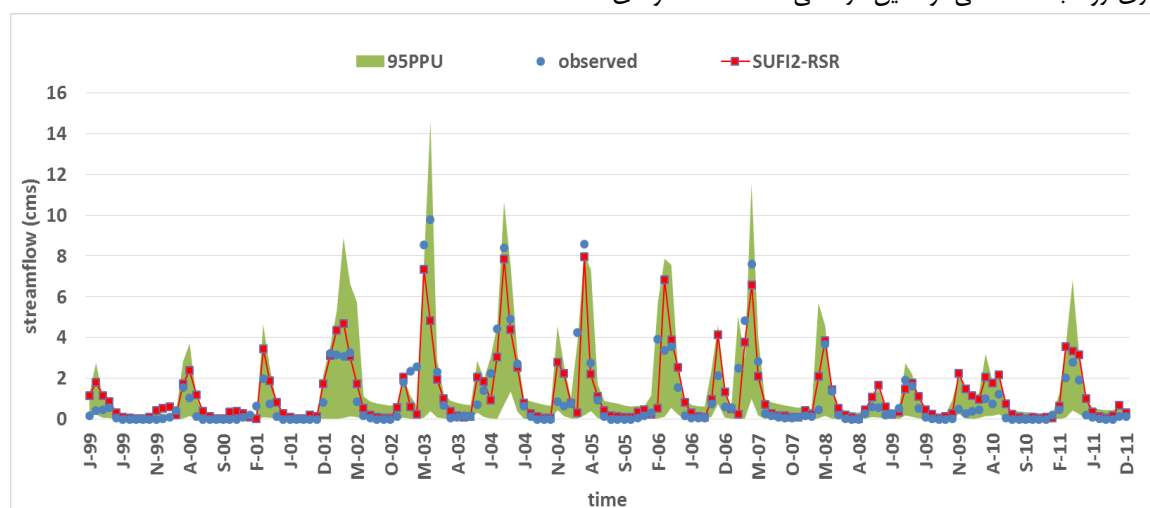
**Table 2- Evaluation of simulation results using goodness-of-fit measures**

جدول ۲- ارزیابی نتایج شبیه‌سازی مبتنی بر معیارهای نکویی برازش

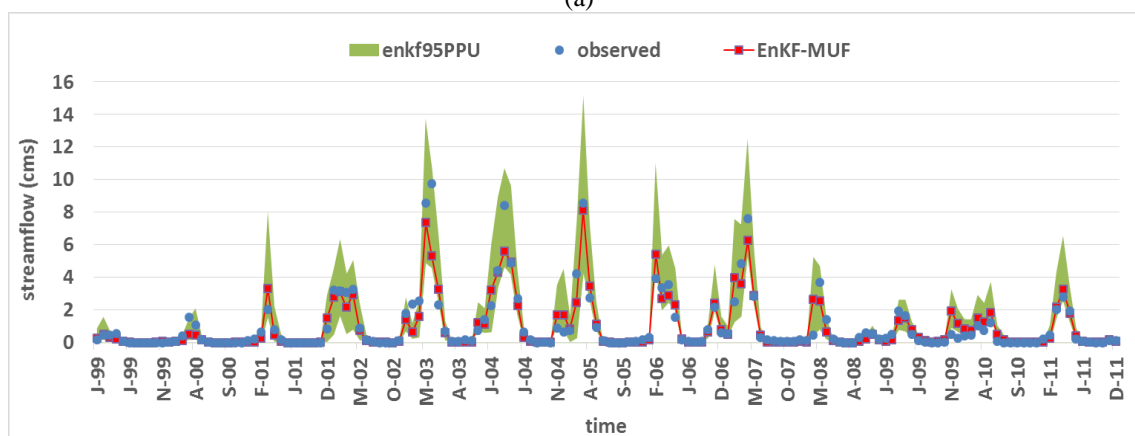
Method	NS	PBIAS	RSR	KGE	p-factor	r-factor
SUFI2-NS	0.72	-17.8	0.53	0.75	0.7	0.98
SUFI2-PBIAS	0	0	1	0.53	0.7	0.98
SUFI2-RSR	0.72	-17.8	0.53	0.75	0.7	0.98
SUFI2-KGE	0.66	-7.6	0.59	0.81	0.7	0.98
EnKF-MP	0.83	4.7	0.41	0.87	0.6	0.45
EnKF-MUF	0.86	5.79	0.37	0.84	0.6	0.45

پارامترهای الگوریتم SUFI2 در دوره‌ی زمانی شبیه‌سازی ثابت است. مقدار معیار p-factor در اجرای EnKF برابر ۰/۶ و در اجرای SUFI2 برابر ۰/۷ است. این در حالی است که اختلاف مقادیر r-factor در دو الگوریتم بسیار بیشتر از اختلاف p-factor است.

است. همچنین مقدار r-factor در اجرای EnKF برابر با ۰/۴۵ و در اجرای SUFI2 برابر ۰/۹۸ که نشانه از کمتر بودن عدم قطعیت نتایج EnKF نسبت به SUFI2 است. در اجرای الگوریتم EnKF با گذشت زمان و افزایش همگرایی پارامترها از عدم قطعیت محدودی شبیه‌سازی رواناب کاسته می‌شود، این در حالی است که محدودی



(a)



(b)

**Fig. 3- Observation, best estimation, and uncertainty band time series in monthly time scale; a) using SUFI2; b) using EnKF**

شکل ۳- سری‌های زمانی مشاهدات، بهترین تخمین و باند عدم قطعیت در مقیاس ماهانه؛ الف- با استفاده از SUFI2؛ ب- با استفاده از EnKF

update states in a distributed hydrological model. *Advances in Water Resources* 31(10):1309-1324

Da Ros D, and Borga M (1997) Adaptive use of a conceptual model for real time flood forecasting. *Hydrology Research* 28(3):169-188

DeChant CM (2010) Hydrologic Data Assimilation: State estimation and model calibration. Thesis, Portland State University

Evensen G (1994) Sequential data assimilation with a nonlinear quasi-geostrophic model using Monte Carlo methods to forecast error statistics. *Journal of Geophysical Research: Oceans* 99(C5):10143-10162

Evensen G (2009) Data assimilation: The ensemble Kalman filter. Springer Science & Business Media

Han E, Merwade V, and Heathman GC (2012) Implementation of surface soil moisture data assimilation with watershed scale distributed hydrological model. *Journal of Hydrology* 416:98-117

Kalman RE (1960) A new approach to linear filtering and prediction problems. *Journal of Basic Engineering* 82(1):35-45

Moradkhani H (2008) Hydrologic remote sensing and land surface data assimilation. *Sensors* 8(5):2986-3004

Moradkhani H, and Sorooshian S (2008) General review of rainfall-runoff modeling: Model calibration, data assimilation, and uncertainty analysis. *Hydrological modelling and the water cycle* p1-24

Reichle RH, Koster RD, Liu P, Mahanama SPP, Njoku EG, and Owe M (2007) Comparison and assimilation of global soil moisture retrievals from the Advanced Microwave Scanning Radiometer for the Earth Observing System (AMSR-E) and the Scanning Multichannel Microwave Radiometer (SMMR). *Journal of Geophysical Research: Atmospheres* 112(D9)

Sun L, Nistor I, and Seidou O (2015) Streamflow data assimilation in SWAT model using Extended Kalman Filter. *Journal of Hydrology* 531:671-684

Xie X, and Zhang D (2010) Data assimilation for distributed hydrological catchment modeling via ensemble Kalman filter. *Advances in Water Resources* 33(6):678-690

به بیان دیگر با وجود افزایش چشمگیر محدوده‌ی عدم قطعیت در الگوریتم SUFI2 نسبت به الگوریتم EnKF، تنها تعداد محدودی از مشاهدات بیشتر در داخل محدوده‌ی 95PPU قرار می‌گیرند که این نشانه توانایی بالقوه‌ی EnKF به منظور افزایش توأم معیارهای p-factor و r-factor است.

در نهایت با توجه به موارد فوق به عنوان یک نتیجه‌گیری کلی می‌توان گفت که استفاده از روشهای جذب داده (مانند الگوریتم EnKF) می‌تواند باعث افزایش دقت شبیه‌سازی فرآیندهای هیدرولوژیکی شود.

#### پی‌نوشت‌ها

- 1- Sequential Data Assimilation
- 2- Nash- Sutcliffe
- 3- Percent BIAS
- 4- Ratio of Standard Deviation of Observation to RMSE
- 5- Kling- Gupta Efficiency

#### ۵- مراجع

Liu Y, Weerts AH, Clark M, Franssen HJH, Kumar S, Moradkhani H, Seo DJ, Schwanenberg D, Smith P, van Dijk AIJM, van Velzen N, He M, Lee H, Noh SJ, Rakovec O, and Restrepo P (2012) Advancing data assimilation in operational hydrologic forecasting: Progresses, challenges, and emerging opportunities. *Hydrology and Earth System Sciences* 16:3863-3887

Abbaspour KC, Yang J, Maximov I, Siber R, Bogner K, Mieleitner J, Zobrist J, and Srinivasan R (2007) Modelling hydrology and water quality in the pre-alpine/alpine Thur watershed using SWAT. *Journal of hydrology* 333(2):413-430

Chen F, Crow WT, Starks PJ, and Moriasi DN (2011) Improving hydrologic predictions of a catchment model via assimilation of surface soil moisture. *Advances in Water Resources* 34(4):526-536

Clark MP, Rupp DE, Woods RA, Zheng X, Ibbitt RP, Slater AG, Schmidt J, and Uddstrom MJ (2008) Hydrological data assimilation with the ensemble Kalman filter: Use of streamflow observations to