DATE October 21, 2020

TO John Rygh
Save the South Fork
Pete Dronkers
Earthworks

FROM Betsy Semmens, RG **EMAIL betsy@basgc.com**

Dear John and Pete,

I am a hydrogeologist with over 20 years of experience in consulting with an emphasis in numerical groundwater modeling. I have a Bachelor of Science in Geology from the University of Minnesota and a Master of Science in Geology from Northern Arizona University. I am the owner of BAS Groundwater Consulting, Inc., located in Colorado, and am a Past-President of the Colorado Groundwater Association. I am a Registered Geologist in Idaho, Arizona, Utah, and Wyoming. I have personally constructed and calibrated dozens of numerical groundwater flow models, including for the mining industry, and used those models for predictions of impacts from mining activities and system recovery post-closure. I also have experience providing third-party reviews of groundwater models.

I have reviewed portions of the Draft Environmental Impact Statement (*Stibnite Gold Project, Draft Environmental Impact Statement*, United States Department of Agriculture, August 2020 (DEIS, 2020, referred to hereafter as "the DEIS")) and the following groundwater modeling reports by Brown and Caldwell (BC) to evaluate potential impacts to groundwater from proposed mining activities at the Midas Gold Stibnite mine in Idaho:

- *Final Stibnite Gold Project, Hydrologic Model Existing Conditions*, *April 27, 2018* (BC 2018a),
- *Revised Final Stibnite Gold Project, Hydrologic Model Proposed Action, October 5, 2018* (BC 2018b),
- *Stibnite Gold Project, Hydrologic Model Sensitivity Analysis, December 2019* (BC 2019a), and
- *Final Stibnite Gold Project Modified PRO Alternative Modeling Report, September 2019* (BC 2019b).

In addition, I briefly reviewed specific information in these supporting hydrogeologic characterization reports:

- *Golden Meadows Project Overburden Geotechnical Investigation, June 2012* (SRK 2012),
- *Stibnite Gold Project Water Resources Summary Report, June 30, 2017* (BC 2017),
- *Groundwater Hydrology Baseline Study, Stibnite Gold Project, Final. June 2017* (SPF 2017), and

BAS Groundwater Consulting, Inc.
3649 Evergreen Parkway, #1510, Evergreen, Colorado 80437 T: +1 720 334-8249 F: +1 303 670-6966

www.basgroundwater.com

- *Final Workplan: Hydrologic Model of the Upper Watershed of the East Fork of the South Fork of the Salmon River, Stibnite, Idaho. October 23, 2017* (JSAI 2017).

Lastly, I reviewed the groundwater modeling computer files provided by Midas Gold and Brown and Caldwell:

- Existing Conditions model – file name: *Midas_ExistingCond.gwv,* and associated MODFLOW-NWT input files,
- Life of Mine model without mining – file name: *Midas_NoAction_AvgPrecip.gwv* and associated MODFLOW-NWT input files,
- Life of Mine model with mining – file name: *Midas_MineOps_AvgPrecip.gwv* and associated MODFLOW-NWT input files, and
- Instructions provided with the model files – file name: *Midas MODFLOW Hydro Model Readme.doc*

My comments on the groundwater modeling effort are provided below. The sections of the DEIS to which the comments pertain are indicated with each comment.

1.0 OVERALL COMMENTS

I have the following overall, general comments on the modeling effort:

1. Bedrock groundwater system, *DEIS Sections 4.8.1.1.1, 4.8.2*

A fundamental assumption of the conceptual model is that groundwater moves mainly within a local, relatively shallow alluvial aquifer system contained largely along the river channels, yet little supporting data are provided to characterize the hydrogeology of the underlying or upgradient bedrock. While the alluvial materials most certainly create an aquifer that is significantly more permeable than the unfractured bedrock, the bedrock system may transmit enough groundwater to influence the resulting conclusions about the impact of mining, particularly where assumptions were made about seepage to groundwater beneath bedrock diversion channels and unlined facilities that will be placed on "impermeable" bedrock. It does not appear that the model sensitivity analysis tested alternative conceptualizations of recharge to, or groundwater flow in, bedrock, nor that sufficient data exist to characterize the hydraulic properties of the bedrock, particularly away from the immediate project site.

2. Sensitivity Analysis, *DEIS Sections 4.8.1.1.1, 4.8.8.2*

The sensitivity analysis performed on the model calibration was not designed to identify the bounds of modeled parameters that result in a set of comparably calibrated groundwater models. The sensitivity analysis was not robust enough to cover the range in measured field values of hydraulic conductivity, the range of tested values for specific yield in bedrock was narrow, and many parameters and model features were not tested at all (this is discussed more specifically below). A sensitivity analysis should determine the extent to which the calibrated model parameters can be varied while maintaining a reasonable model calibration, for use in designing an uncertainty analysis of the model predictions.

3. Uncertainty Analysis, *DEIS Section 4.8.8*

The uncertainty in the predictive mining and post-closure models was tested by varying the simulated climatic conditions including average, wet, and dry scenarios. Additionally, the variations of parameters that defined the sensitivity analysis were used in the predictive simulations to represent an uncertainty analysis. The report emphasized that the alternative models used for the uncertainty analysis were uncalibrated models for which parameter modifications worsened the model calibration. Uncalibrated models should not be used as a basis for conducting an uncertainty analysis, and instead, the bounds of calibrated parameters that result in reasonable, alternate model calibrations should be used as the basis of an uncertainty analysis. Correlated parameters can be combined to demonstrate worst case scenarios, but from alternatively calibrated models. Using the results of the sensitivity analysis for the uncertainty analysis as they did, they excused the significant differences in model predictions stating that these models are uncalibrated and do not represent reasonable values of modeled parameters.

The general impact of modifying the model to address the comments in this letter could potentially change model predictions such as:

- groundwater flow directions and interaction with surface water during and after mining,
- the amount of excess water returned to the groundwater system in the rapid infiltration basins (RIBs),
- estimated depth to groundwater and impacts of groundwater mounding beneath or within facilities and generated geochemistry, and
- estimates of groundwater discharge to the open pits which could, in turn, influence estimates of ultimate pit lake levels, excess water to the RIBs, and the geochemistry of seepage during and after mining.

2.0 TECHNICAL COMMENTS – MODELING APPROACH

I have several technical comments on the approach BC undertook for their numerical groundwater modeling effort that may impact their predictive modeling results. These technical comments are listed below.

1. **Modeled Hydraulic Conductivity, *DEIS Sections 4.8.1.1.1, 4.8.2, 4.8.2.1.2.1***

A relatively high zone of hydraulic conductivity (K) with a value of 3 feet per day (ft/d) was delineated along each drainage channel in model layer 2 to represent localized, shallow, fractured bedrock. With a thickness of 500 feet, the fractured bedrock in model layer 2 within each stream channel creates areas of high transmissivity in the model, potentially equal to the alluvial aquifer where model layer 1 has its maximum thickness (200 to 250 feet) and about double where model layer 1 is thinner (50-100 feet). The modeled K value for these permeable bedrock channels is about three orders of magnitude higher than the rest of the bedrock in model layer 2, and have a modeled transmissivity about two orders of magnitude higher than the overlying material in model layer 1 outside of the area of mapped alluvium/overburden.

These bedrock channels therefore constitute a dominant feature in the model to route groundwater along specific paths, contrary to the conceptual model that the alluvial aquifer is the most groundwater-transmissive formation (DEIS Section 4.8.2). From the data sets, it appears that most or all of the data to characterize the bedrock is local to the mine site, and it is unclear what data were relied upon for the assumption that bedrock everywhere else in the model domain should be so much less permeable. It does not appear that this assumption was tested during model sensitivity or in an uncertainty analysis of the mining and post-mining model scenarios to determine how strong of an influence this assigned high permeability pathway has on the model predictions.

Figure 5-2 of BC 2018a shows that model layer 1 has saturation throughout the layer in the winter months, including in the thinner areas outside of the mapped alluvium/overburden materials. The lower K bedrock unit is beneath much of this saturated area, and much of this area is not well characterized with data or model calibration. One concern is that the lower K bedrock unit incorrectly limits the downward infiltration of seasonal recharge, and instead, potentially maintains artificially shallow groundwater flow paths that may impact model results that rely upon the modeled depth to water, such as infiltration of seepage, discharge to open pits, and restored stream flows in certain areas. This may affect the true impacts to seep-, spring-, and wetland groundwater dependent ecosystems (DEIS Section 4.8.2) and estimates of groundwater drawdown used to quantify impacts from the proposed mine (DEIS Section 4.8.2.1.2.1).

It is possible that the modeled distribution of K in the alluvium and bedrock is appropriate to represent groundwater flow in this area, but this was not proven with the dataset or the modeling approach. Also, DEIS Section 4.8.8.2.1.2 references long term pumping bedrock aquifer testing conducted in late 2019 and that these data will "likely improve characterization of hydraulic properties of the bedrock formations." These data should be compared against the modeled bedrock K values and the model calibration should be tested with these values.

2. Modeled Storage, *DEIS Section 4.8.2.1.1.1*

The modeled values of specific yield are quite low for the unfractured bedrock (0.1%) and should have been tested in a sensitivity analysis, particularly if aquifer test data are insufficient (or non-existent) to define storage.

Higher specific yield in the bedrock could result in higher discharge to the open pits as the mining plan is advanced with the model Drain boundary conditions, which in turn affects the conclusions regarding predictions of water available for mining operations, estimates of recharge through the RIBs, and the mitigating effect of the RIBs on streamflow.

3. **Model Layering, *DEIS Section 4.8.1.1.1, 4.8.2***

The model layering may be too coarse/thick to adequately simulate vertical hydraulic gradients that may be important for model predictions of water exchange between groundwater and the streams or infiltration of seepage from mining activities. Hydraulic gradients measured in monitoring wells were discussed in BC 2017, but were not discussed in the context of model calibration in BC 2018a. There was discussion, however, that the model did not calibrate well to a pump test in the Gestrin well because the model layering is a "poor analog of the test conditions" because the pumping well and observation wells have well screens that are only about 10 feet long and model layer 1 is "approximately 190 feet thick near the pumping well" (BC 2018a). Appendix A of SPF, 2017 provides monitor well construction information and most monitor wells have screen lengths of approximately 10 feet. Model layer 1 varies in thickness between 15 and greater than 250 feet and the upper bedrock layer (model layer 2) is 500 feet thick. The model-calculated groundwater elevation is averaged vertically over the thickness of each layer. It has not been shown that the model layering is sufficiently discretized to represent the vertical movement of groundwater typically found near open pits nor the hydraulic gradients needed to properly model the exchange of groundwater with surface water.

4. **Modeled Water Budget, *DEIS Section 4.8***

A calibrated water budget was not provided for the model calibration. Components of the conceptual water budget, derived from the meteoric water balance were provided in BC 2018a but the resulting calibrated water budget was not provided to allow confirmation that the model adequately represents the total expected amount of water and its movement between the various components of the model. Providing a calibrated water budget and presenting it in the context of how well, or not well, it represents the conceptual water budget is a basic standard of groundwater modeling (American Society for Testing and Materials (ASTM) D5490).

Similarly, while a large amount of modeled flow information is provided for various components of the Proposed Action predictive model (simulated streamflows, groundwater discharges to open pits by geologic unit, etc) (BC 2018b), a summarized water budget that can be easily compared to the calibrated model water budget and the conceptual model water budget was not provided. Additionally, the pit filling curves (Section 5.1, BC 2018b) did not provide the modeled fluxes for each component of the pit lake simulation (runoff, evaporation, and precipitation over time), hindering review and understanding of the modeled predictions.

5. Modeled Recharge, *DEIS Section 5.8.1.1.1*

The modeled amount of recharge was assigned in advance of the model calibration from the meteoric water balance, by estimating the total ‘available’ water (snowpack and precipitation minus sublimation and evapotranspiration) and distributing it between runoff and recharge to groundwater. This total water was assigned to recharge first, up to a maximum value based on assumed hydraulic properties of the surface material, and the remainder applied as runoff.

There were many assumptions made along the way with this multi-step approach to estimating the amount of recharge to groundwater, stemming from the method and results of the meteoric water balance and the assumption of hydraulic properties of the surface material. Much of their meteoric water balance was based on the PRISM precipitation dataset and not site-specific climate data, and large portions of the modeled area have no data from which to characterize the surface and near-surface materials. While this methodology is one approach to obtain an initial estimate of the recharge and runoff distribution and amount, it is not clear to what extent these assumptions were tested during model calibration. Perhaps the fundamental assumptions for recharge were calculated once and assumed correct, without further testing. The recharge distribution and amount were not varied in the model sensitivity analysis.

From review of the provided model files it appears that the amount of recharge simulated in the model is a significant proportion of the modeled water budget; therefore the amount of simulated recharge should have been tested during a sensitivity analysis to determine the range over which recharge could be varied and maintain model calibration. This range should have been used in an uncertainty analysis of the mining and post-mining period predictions.

6. Modeled Streamflow, *DEIS Section 4.8.2*

The model simulates streamflow using MODFLOW’s Streamflow Routing (SFR) package. A constant depth was assumed for the streams (ICALC = 0), and the modeling report indicates that the depth was assumed to equal 2 feet in all streams (Section 4.3.2, BC, 2018a). This depth may be too small to property represent the actual stream depth for some portions of the streams and this overall assumption may be too restrictive to allow for appropriate representation of the stream geometry because the model will hold the depth constant while exchanging streamflow with the underlying aquifer. Stream geometry is important because the stream stage is used to calculate the hydraulic gradient with the underlying aquifer and thus impacts the exchange of water between the aquifer and the stream, a major component of the model calibration and an important predictive result of the future models. At the very least the assumption of stream depth should have been tested in the model sensitivity analysis along with other options for defining stream characteristics, such as defining a Manning's coefficient from literature values for comparable settings, or conducting stream bathymetry surveys (ICALC = 1 or 4). Given that two main conclusions made from modeling of the mining and post-mining periods is the impact of mining on the amount of streamflow and the time, post-closure, until streamflow is restored, the stream characteristics should have been thoroughly tested in the sensitivity analysis.

7. Geologic Structures, *DEIS Sections 4.8.2.1.2.1, 4.8.8.2.1.3*

The model approach did not adequately consider the effects of geologic structures on the model calibration or in an uncertainty analysis for the predictive simulations. That faults are present and may locally convey groundwater is noted in the existing conditions modeling report, including the potential for permeable faults in the deeper bedrock (BC 2018a) but no attempt was reportedly made to represent these faults in the model, beyond the thick zone of higher permeability in the upper bedrock beneath all drainages. Statements are made that no significant permeable structures have been identified during drilling (JSAI 2017), but it does not appear that the drilling program was extensive away from the immediate project site. Reference is made in the DEIS to the lack of characterization of the hydraulic nature of major faults (e.g. Meadow Creek Fault Zone) that extend through the Hangar Flats and Yellow Pine pit locations and the East Fork of Meadow Creek and Sugar Creek (DEIS Section 4.8.8.2.1.3). No surficial geologic map is provided in the modeling reports for the area to show mapped structures, despite that such maps are available (*Geologic Map of the Stibnite Quadrangle, Valley County, Idaho*. Idaho Geological Survey, Geologic Map 51).

Additionally, the major faults were not tested as low permeability features. Extensive mineralization along the fault may impede groundwater flow and impact the groundwater flow directions and gradients. A series of low permeability geologic structures may create compartmentalized groundwater systems that respond differently to an advancing open pit than what was simulated.

An independent review of the modeling effort referenced in Section 4.8.8.2 of the DEIS also noted this lack of testing of faults in the model and the lack of characterization of the properties of the faults. The DEIS notes that hydraulic testing of faults is expensive and difficult, but the assumptions of permeable and impermeable faults on the groundwater system during mining and post-closure could easily be tested with the groundwater model.

8. Model Calibration, *DEIS Sections 4.8.1.1.1, 4.8.8.2.1.1*

The model calibration was limited to the area of available data, which is local to the mine site. Additional wells should have been installed, at least throughout the modeled area, to provide characterization data and additional calibration locations to improve confidence in the model's ability to represent the groundwater flow system.

Table 2-1 from the existing conditions modeling report (BC, 2018a) is titled Summary of Target Groundwater Elevations, and provides well names, coordinates, and water level elevations for 50 monitor wells. Given the title of this table, these wells should represent the calibration data set used to calibrate the groundwater model to measured groundwater elevations. However, Table 5-1, which presents calibration statistics, indicates that there are 55 calibration targets. No dates are provided for the water levels reported in Table 2-1, and Section 2.3.2 (BC 2018a) states that the water levels in Table 2-1 are "based on water levels measured in during fall months." The groundwater elevations provided in Table 2-1 are presented as whole integers.

The actual data that were used to calibrate the model should be presented in the modeling report, especially considering that the targets were stripped from the provided Groundwater Vistas calibration file (see Section 3 below). These data should have included the elevation of the water level measurement, the date of the measurement, and the actual water level measurement, which was presumably not measured from the field program to the nearest foot.

Furthermore, Table 2-1 categorizes each well as 'alluvial' or 'bedrock'. Seven wells are categorized as 'bedrock' but 20 wells are shown as calibration targets in model layer 2 on Figure 5-5. Given that model layer 1 was defined to represent alluvium and overburden with the bottom elevations defined from an assessment of the thickness of soil, colluvial, and alluvial overburden (DEIS Section 4.8.1.1.1), and model layer 2 represents near surface bedrock, it is not clear why 13 wells categorized as 'alluvial' wells were simulated in the bedrock model layer.

It appears that the model calibration to water levels was not performed as a transient calibration to time series of water level data. Water level data are provided in Appendix C of SPF 2017 and these transient, time-series data should have been used for the model calibration. A transient calibration to measured water level elevations would have strengthened the model calibration and is appropriate for a model that will be used to predict transient changes to the groundwater system.

Additionally, there is an existing pit lake in the Yellow Pine pit. There was no mention in the modeling report of any attempt to estimate a pit lake water balance including estimates of groundwater discharge during periods without recent precipitation. It also does not appear that the current pit lake levels were used as part of the model calibration, which would have strengthened the model calibration.

9. Assumptions of Seepage from Mining Facilities, *DEIS Section 4.8.1.1.2*

Assumptions were made in the Proposed Action predictive model regarding seepage from facilities. They assumed no seepage beneath the tailings storage facility (TSF) and the main diversions because they will be lined, but appear to have made the same assumption where the facility will be constructed in “relatively impermeable bedrock” to assume that the “affected portion of Meadow Creek” will not interact with the alluvial groundwater. Bedrock, even low permeability bedrock, is never impermeable and a reasonable amount of seepage where facilities are unlined over bedrock should be estimated, modeled, and tested.

Similarly, the stream conductance was defined with a very low width ($1x10^{-6}$ ft) for lined diversions and streams over bedrock to essentially prevent interaction with the bedrock. Stream conductance should be defined based on understood characteristics of the stream channel and the model should be allowed to define the interaction with the underlying bedrock based on the calibrated (and tested) values of bedrock K.

It was assumed that the Developmental Rock Storage Facilities (DRSF) and RIBs will be reclaimed and therefore background estimates of recharge and runoff can be applied during post-closure (BC 2018b). Unless the DRSF will be drained at the end of mining, a drain-down period should be assumed for the post-closure period. Drain down periods can be long, depending on the site and nature of the deposits, and should be estimated from the engineering designs and planned operations.

10. Data Used for Post-Closure Modeling Analysis, *DEIS Sections 4.8.1.1.1, 4.8.1.1.2, 4.8.5*

No site pan evaporation data were available. This is a basic and minimal data compilation requirement for a site at which predictions of pit lake formation will be made, and it is unacceptable that this is not available. Evaporation estimates are critical for the post-closure pit lake simulations.

11. Modeling Approach to Estimate Post-Closure Impacts, *DEIS Sections 4.8.1.1.2, 4.8.2, 4.8.5, 4.8.6*

The proposed action modeling was used to predict that the surface flows will return to a long-term stable regional pattern within 10 years after mining ceases. Use of the SFR package to quasi-simulate groundwater and surface water interactions is not rigorous enough to draw that conclusion and conclusions cannot be reasonably made without a sensitivity analysis on the stream parameters.

12. Sensitivity and Uncertainty Analyses, *DEIS Section 4.8.8.1*

As discussed in Section 1, the sensitivity and uncertainty analyses were inadequate to characterize the sensitivity of the model calibration and associated uncertainty in the model's predictions. The appropriate modeling approach is to a) perform a model calibration, b) define a sensitivity analysis to test the modeled parameters within potential ranges (i.e. measured data for K values, upper range of recharge based on percent of precipitation or site-wide water balance, etc) or until the model calibration is worsened, and c) devise an uncertainty analysis to test the predictive model to the range of parameters that produces comparably calibrated models. Parameters can be grouped together based on the results of the sensitivity analysis to create 'worst case' alternative models that can be used in additional predictive model runs. This overall approach is referenced in DEIS Section 4.8.8.1 as the standard approach outlined in the ASTM guidance yet was not performed for this evaluation.

For example, for this modeling effort, it may make sense to find how far the bedrock hydraulic conductivity can be raised without ruining the model calibration as an alternate calibrated model with greater permeability that may interact differently with the open pit or infiltrate recharge into the deeper model layers, creating a deeper groundwater flow system that interacts differently with the streams. Increasing the K values may lower the water levels, worsening the model calibration, but perhaps this can be balanced by decreasing the K of the higher bedrock zones beneath the washes or increasing recharge within a reasonable range.

Similar examples could be made for testing permeable and impermeable major geologic structures, the stream parameters, vertical anisotropy, etc. The sensitivity analysis that was performed individually tested a few of the model parameters but did not test the major assumptions of the conceptual model, nor did it combine parameters to find alternative calibrations.

The results of the proposed action future, predictive model simulations indicate that all the open pits will be flow through and will not create a passive containment capture zone. Therefore, impacted pit lake water will ultimately discharge to streams. Considering this, and because the predictive model was used to quantify volumes of water for the geochemical evaluation, an uncertainty analysis that adequately brackets reasonable uncertainty in the model predictions is critical.

3.0 TECHNICAL COMMENTS – MODEL FILES

Model files were made available for the existing conditions model, the no action and mining operations models for average, above average, and below average precipitation scenarios, and the post-mining model. For all models, the MODFLOW-NWT input files were provided along with the Groundwater Vistas file (the graphical user interface (GUI) model files).

I have reviewed the model files for the existing conditions model, and the no action and mining operations models for average precipitation conditions (files: *Midas_ExistingCond.gwv, Midas_NoAction_AvgPrecip.gwv, and Midas_MineOps_AvgPrecip.gwv).* My review consisted of opening the Groundwater Vistas GUI files, reviewing the definitions for model parameters and boundary conditions, translating the MODFLOW input files, modifying the name (.nam) files to include the entries needed to write the text files for the gage package (as outlined in the README file distributed with the model files), running MODFLOW-NWT, and evaluating the model results.

By comparing the predictive model results from the output from the provided model files to the graphs and discussion of simulated impacts from mining and post-closure presented in the Proposed Action modeling report (BC 2018b) and the DEIS, it appears that the provided model files are for the Proposed Action presented in BC 2018b (referred to as Alternative 1 in the DEIS), and not the alternatives presented in the Modified Plan of Restoration and Operations modeling report (BC 2019b).

Below are specific technical comments from my review of these model files. These comments document deficiencies and concerns about the execution of the groundwater model and by extension, concerns regarding the model's ability to properly represent the groundwater system and predict future impacts from the proposed Stibnite mining activities. Therefore, these comments pertain, overall, to DEIS Section 4.8.

1. The provided GUI files were not offset to relate the coordinates of the model output to real world coordinates, nor was the offset and rotation documented in the existing conditions modeling report (BC 2018a). As such, one cannot easily export the model results in coordinates that allow for evaluation in Geographic Information Systems (GIS) software, nor to import georeferenced information such as well locations. The model results are provided on georeferenced base maps in the modeling report so it can only be assumed that the geographical offset was removed from the GUI files to make review of the modeling effort more difficult.

2. The well locations used to calibrate the groundwater model in the existing conditions model, shown on Figures 5-4 and 5-5 of the existing conditions modeling report (BC 2018a), were removed from the existing conditions GUI file (*Midas_ExistingCond.gwv*). Table 5-1 of the existing conditions modeling report (BC 2018a) provided calibration statistics and indicates that 55 calibration target locations were used for calibration. These calibration statistics could not be verified because the calibration targets were removed from the existing

conditions GUI file (*Midas_ExistingCond.gwv*), again, presumably to make review of the modeling effort more difficult.

3. The existing conditions model has 396 monthly stress periods. Section 4.1 (BC 2018a) indicates that the model initially was set up to run for 122-years (1895 through 2016) including the "2011 through 2016 calibration period" and that "once general calibration from the 122-year simulation was achieved, the model simulation period was focused to simulate conditions between 1985 and 2016...". It is therefore assumed that the 396 monthly stress periods in the existing conditions model (*Midas_ExistingCond.gwv*) represent this focused period from 1985 to 2016. However, the stress periods and their lengths provided in the existing conditions GUI file (*Midas_ExistingCond.gwv*) do not equate to January 1, 1985 through December 31, 2016, and instead are about one-year too long. Assuming that the end of the model period is December 31, 2016 based on the calibration streamflow graphs provided in BC 2018a, the start of the model simulation period in the existing conditions model (*Midas_ExistingCond.gwv*) would be January 1, 1984. Additionally, the length of the monthly stress periods accounts for Leap Years, however, it appears that they erroneously defined a Leap Year about every 3rd year, not every 4th year. For the tests of the model discussed hereafter, it was assumed that the provided existing conditions model begins on January 1, 1984, and the erroneous definitions for Leap Years were ignored.

4. I ran the existing conditions model twice: once from the MODFLOW-NWT executable provided with the input files using the provided MODFLOW-NWT input files (i.e changing nothing), and once by translating the input files from the provided GUI file (*Midas_ExistingCond.gwv*) and modifying the name file per instructions in the provided readme file. The model did not converge for the stress periods in the transient simulation for either model run. The overall mass balance is typically low for most stress periods, although mass balance is between about 2 and 8 percent for some stress periods. Persistent lack of convergence may suggest a problem with the fundamental assumptions in the model, such as the boundary conditions, and care should have been taken to create a numerical model that converges. Although the overall mass balance error in the model is generally acceptable for most stress periods, having mass balance error greater than 2 percent also may indicate a problem with the model's boundary conditions, in this case, most likely with the interaction with the streams. The model should be revised to one that converges without mass balance error before it is used for predictive purposes.

5. To evaluate the model calibration, I imported the water level elevations provided in Table 2-1 as calibration targets in the existing conditions GUI file (*Midas_ExistingCond.gwv*). The 50 calibration targets were imported to model layers 1 and 2 based on the data shown on Figures 5-4 and 5-5 (BC, 2018a). The targets were assigned to stress period 393 at elapsed time 11962 days because Section 2.3.2 indicated that these water levels are representative of conditions during fall months, and stress period 393 should represent September

2016 (given previously stated assumptions about the time period that the model represents). If the statement made in Section 2.3.2 (BC 2018a) holds, it should not matter which "fall month" stress period the targets are assigned to, because the water levels in Table 2-1 are representative of "fall months."

Additionally, I used the provided georeferenced GIS file, ModelArea.kmz, to approximate the offset and rotation of the groundwater model in Idaho State Plane West, NAD83, feet coordinates: 2,709,836, 1,171,555, -39 degrees. This gives a georeferenced coordinate system that appears close enough to be able to import the calibration targets from the coordinates provided in Table 2-1.

The model calibration statistics are presented in Table 5-1 (BC, 2018a) for the entire model, model layer 1, and model layer 2. Focusing only on the root mean squared error (RMS) as a percentage of the total head change across the model domain (%RMS), the results provided in Table 5-1 were: 1.5%, 1.4%, and 1.7% respectively. It is not clear what data were used to derive these calibration statistics nor to what time period the calibration data were compared, but assigning the calibration targets to stress period 393 at time 11962 days, the existing model (*Midas_ExistingCond.gwv*) calibration %RMS were: 5.8%, 6.5%, and 3.8%, significantly worse than reported in the modeling report.

Matching the modeled streamflow to measurements of streamflow in United States Geological Survey (USGS) gaging stations was a major component of the model calibration. Graphs were provided showing a very close match between modeled streamflow and measured streamflow for the period 2011 to 2017 at five USGS gaging locations:

- East Fork of the South Fork of the Salmon River Above Meadow Creek near Stibnite (USGS #13310800),
- Meadow Creek near Stibnite (USGS #13310850),
- East Fork of the South Fork of the Salmon River at Stibnite (USGS #13311000),
- East Fork of the South Fork of the Salmon River Above Sugar Creek near Stibnite (USGS #13311250), and
- Sugar Creek near Stibnite (USGS #13311450).

I downloaded the gage data for these USGS gaging stations and replicated the graphs shown in the existing conditions modeling report (BC, 2018a) to compare the model output from the provided model files (*Midas_ExistingCond.gwv*). The gage data were averaged monthly to be consistent with the data shown in Section 5 of the existing conditions modeling report (BC, 2018a). In general, the graphed output from the provided MODFLOW-NWT files for existing conditions does not match the output shown on graphs in the existing conditions modeling report (BC, 2018a), and in some places, the graphed measured gage data also do not match. The difference between the modeled and measured streamflow is greater in the graphs that I produced from the model runs from the provided MODFLOW-NWT input files than what is shown in the modeling report (BC 2018a). In other words, the calibration to the measured streamflow is not as good when assessed from the replications of the graphs using the rerun model results.

Screen shots from the graphs in the existing conditions modeling report (BC, 2018a) are shown along with the new graphs created from rerunning the provided existing conditions model for all five gaging stations below. These side-by-side comparisons show that the match between measured and simulated streamflow is not as good as presented in the modeling report.

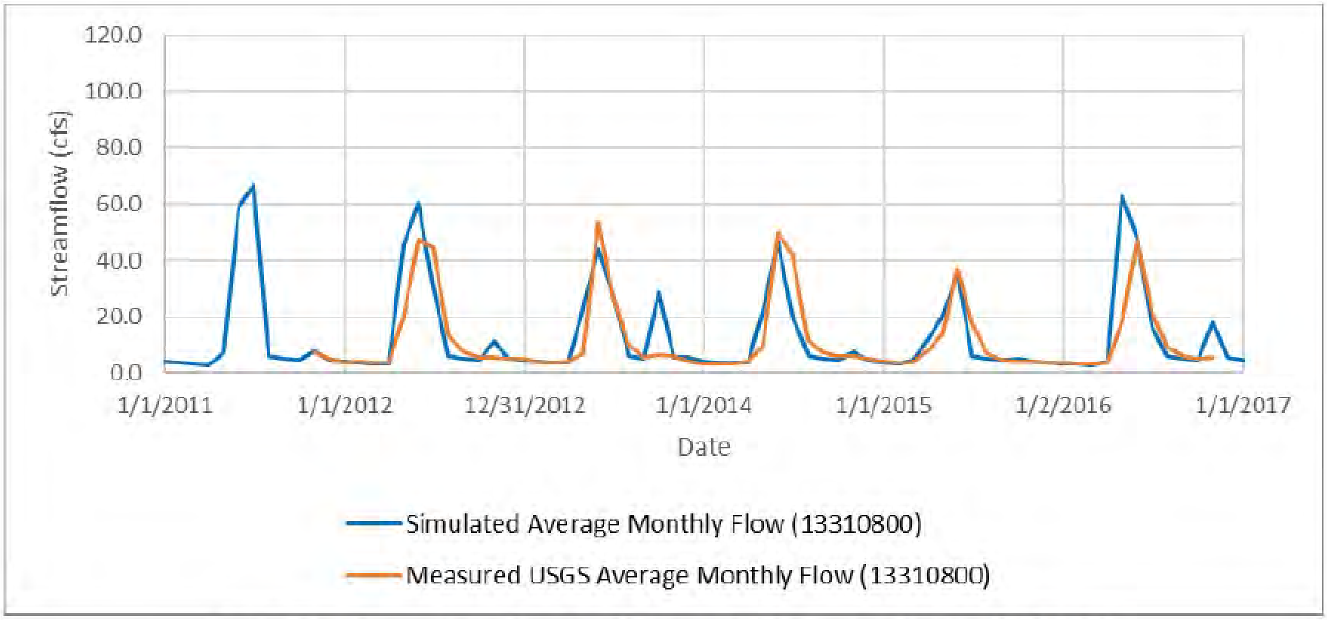
6. The transient model calibration should have been calibrated to transient groundwater elevations. As mentioned in an earlier comment, it is unclear what data were used for model calibration, however, time-series of groundwater elevations are provided in SPF 2017, Appendix C, Attachment A.

   The data from Appendix C, Attachment A (SPF 2017) were input to the existing conditions model using the well construction details provided in Appendix A (SPF 2017). These groundwater elevation data are transient and were input to the model assuming that the start of the model represents January 1, 1984, as discussed previously.

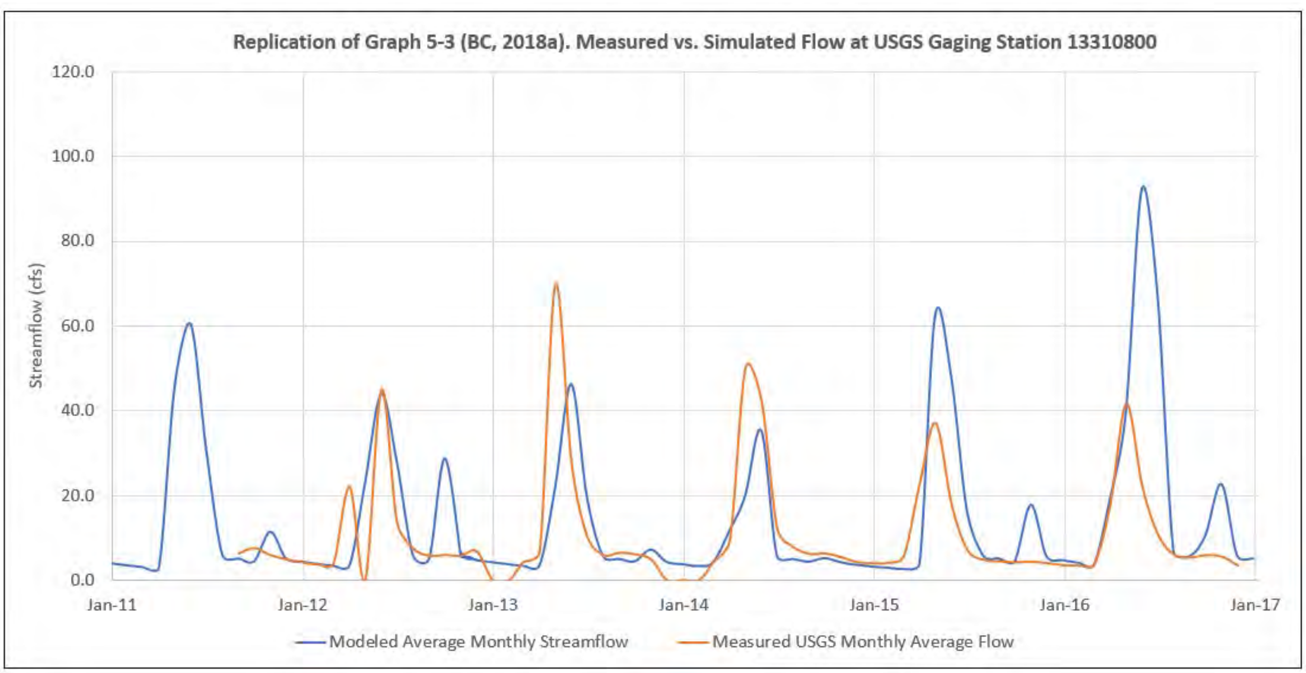
   The resulting calibration scaled RMS errors for the entire model, model layer 1, and model layer 2 were: 6.0%, 3.3%, and 7.7%, respectively. These statistics are significantly worse than the statistics provided in Table 5-1 and suggest that the model is not as well calibrated as indicated, particularly in the shallow bedrock. The residuals for the most recent groundwater elevation data value from this set of calibration targets is shown in Figure 1 below. A negative residual indicates that the modeled water level is higher than the measured water level, while a positive residual indicates that the modeled water level is lower than the measured water level. Figure 1 shows significant bias that the model overestimates groundwater elevations (red markers), particularly near Yellow Pine pit and upgradient from Hangar Flats. In other words, the model's groundwater elevations are generally too high, likely meaning that the overall K values or recharge are incorrect.

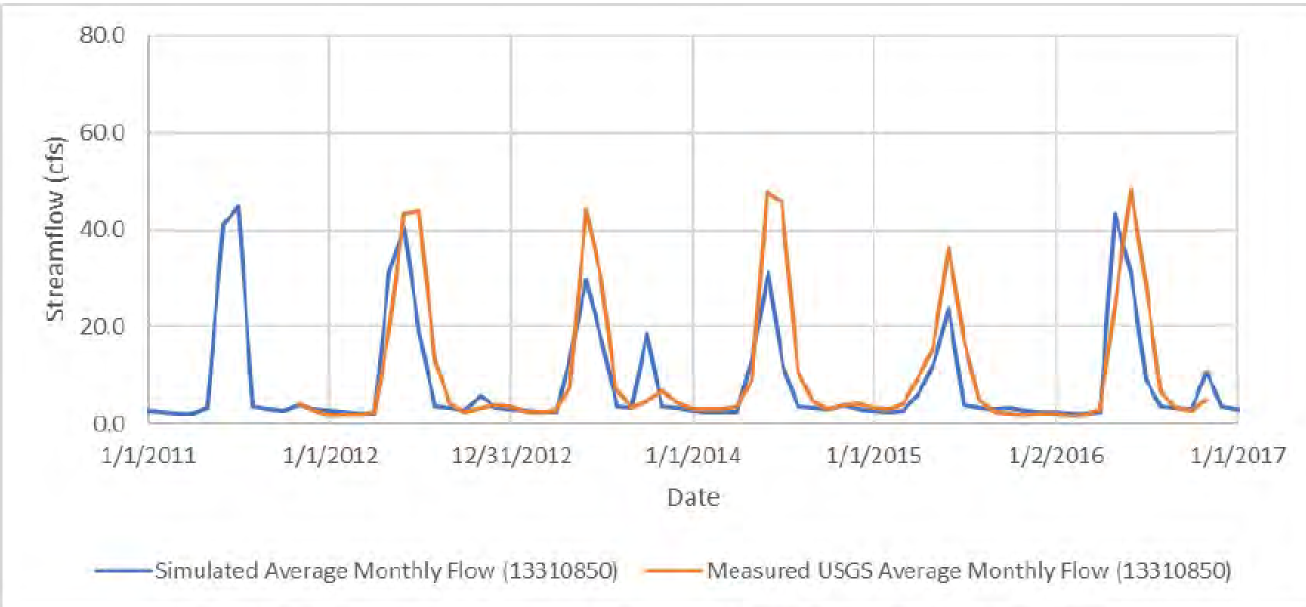
   Additionally, importing the calibration targets (monitor wells) using the well screen intervals from Appendix A (SPF 2017), several of the monitoring wells imported to model layers other than as indicated on Figures 5-4 and 5-5 of the modeling report (BC 2018a). Note that the model layering was not changed from that provided in the model files – the monitoring wells were simply imported to the provided model file using the coordinates and screen intervals provided in Appendix A (SPF 2017).

   This shows that the model calibration presented in the modeling report (BC 2018a) is a) not as good as presented, and b) suspect because it is not able to be verified. Therefore, the validity of using this model for predictive purposes is uncertain.

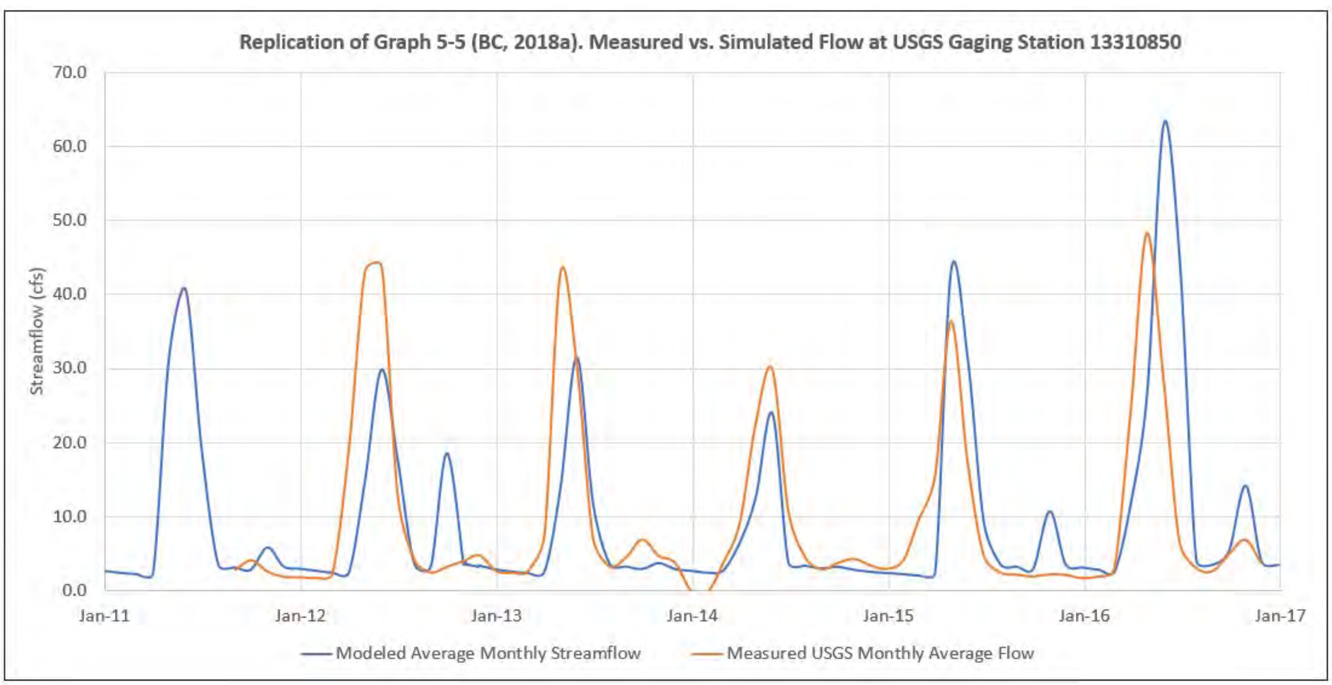


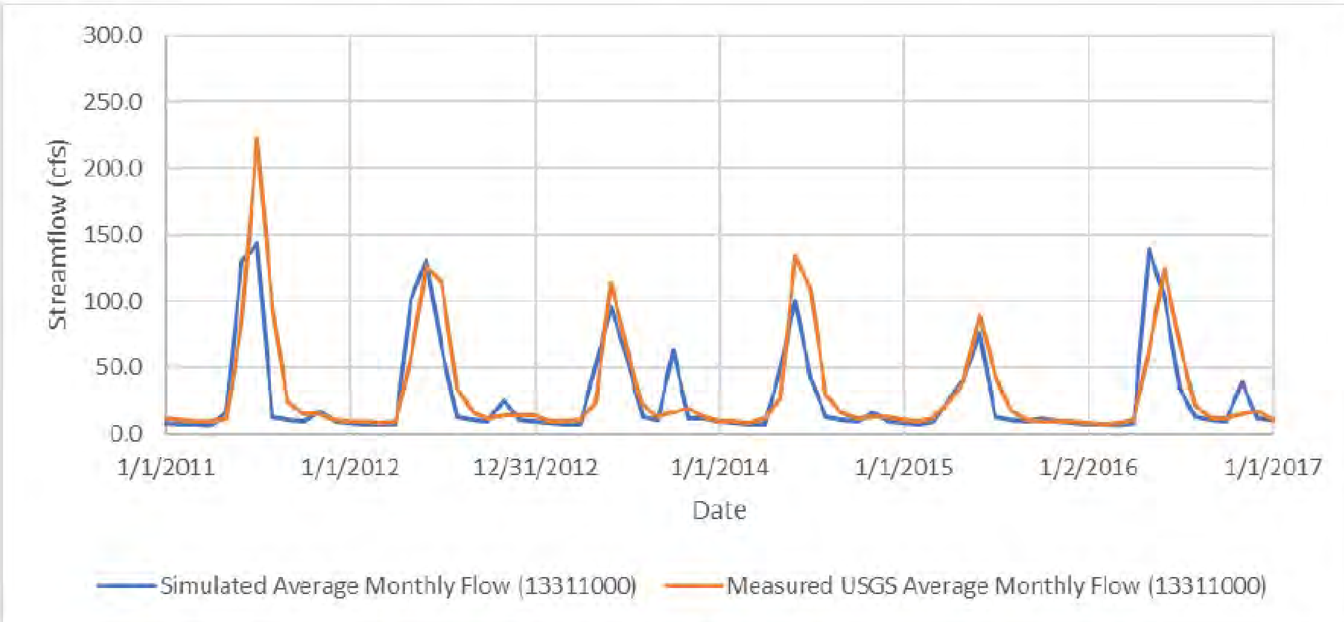
Graph 5-3. Measured vs. Simulated Flow at USGS Gaging Station 13310800, EFSFSR above Meadow Creek (Linear)



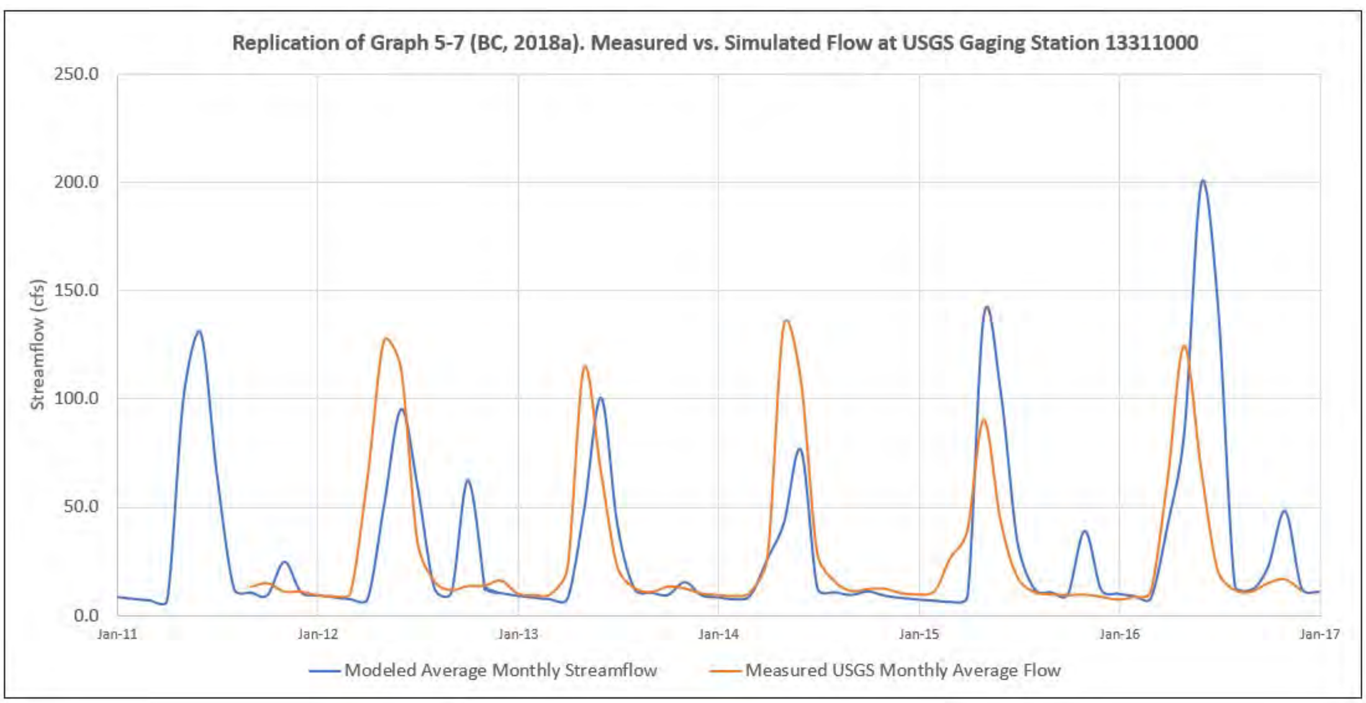


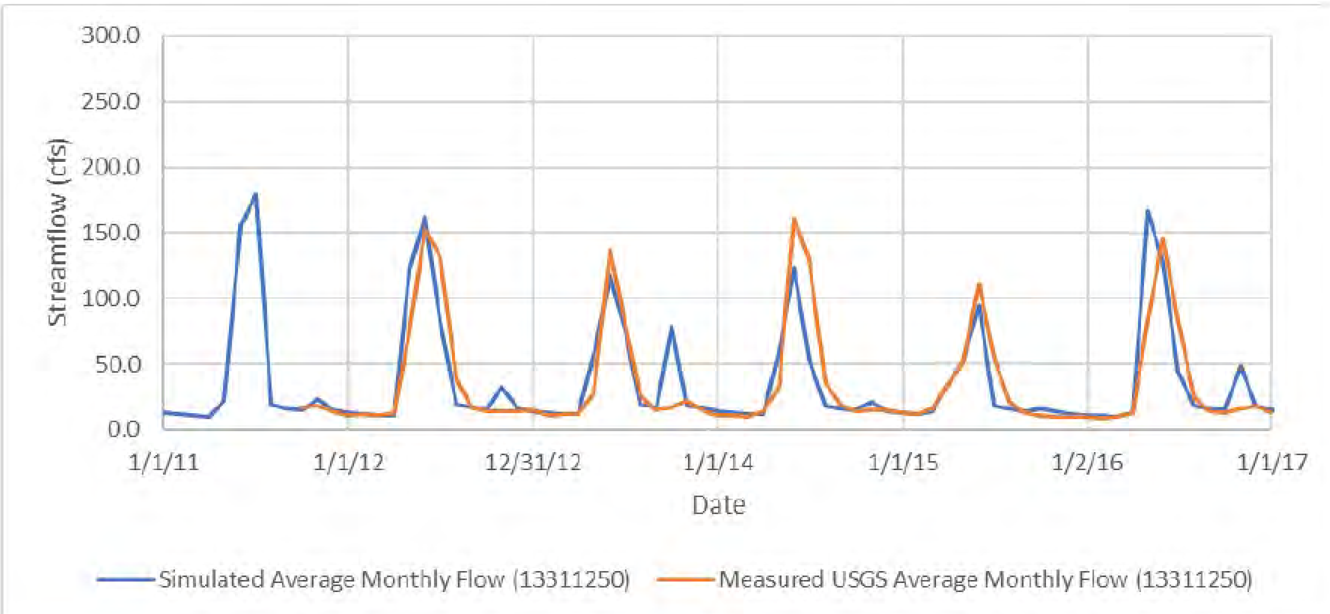
Graph 5-5. Measured vs. Simulated Flows at USGS Gaging Station 13310850, Meadow Creek near Stibnite, ID (Linear)



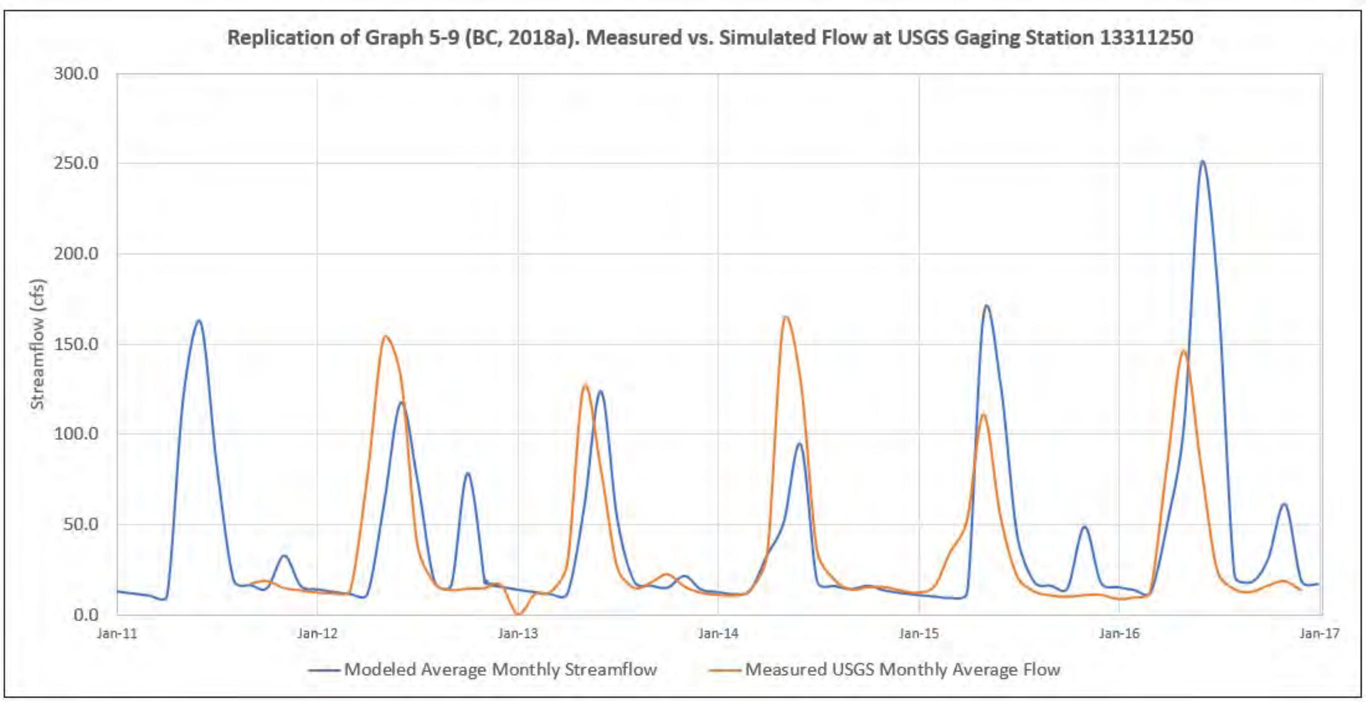


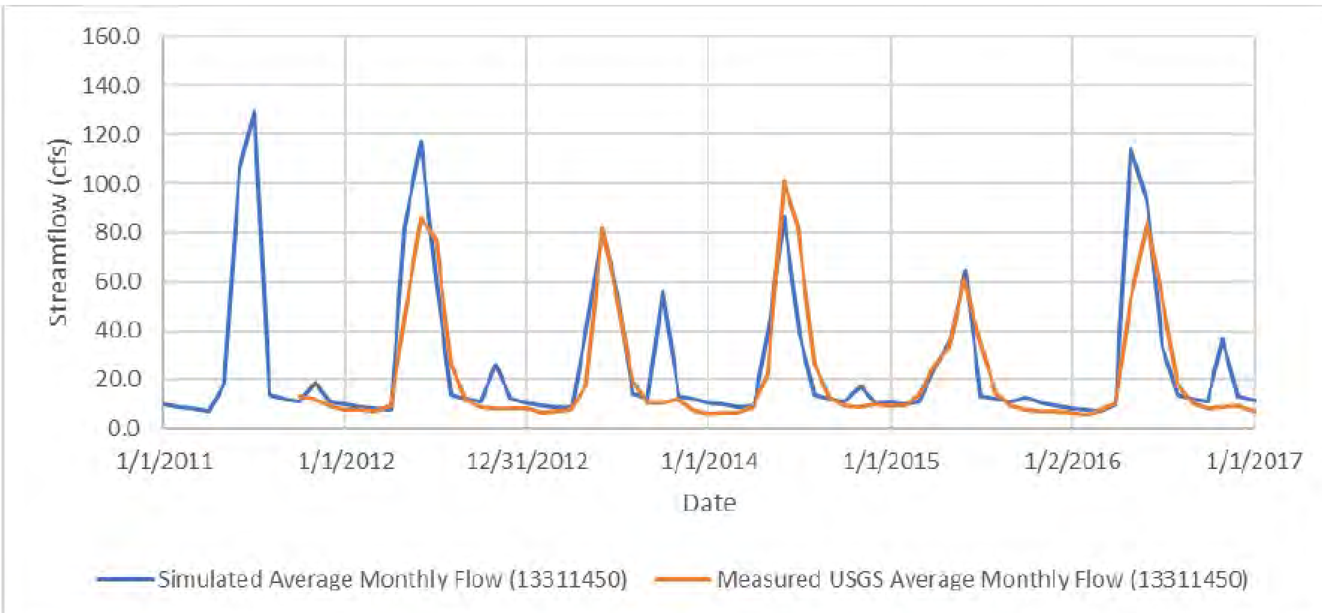
Graph 5-7. Measured vs. Simulated Flow at USGS Gaging Station 13311000, EFSFSR at Stibnite, ID (Linear)



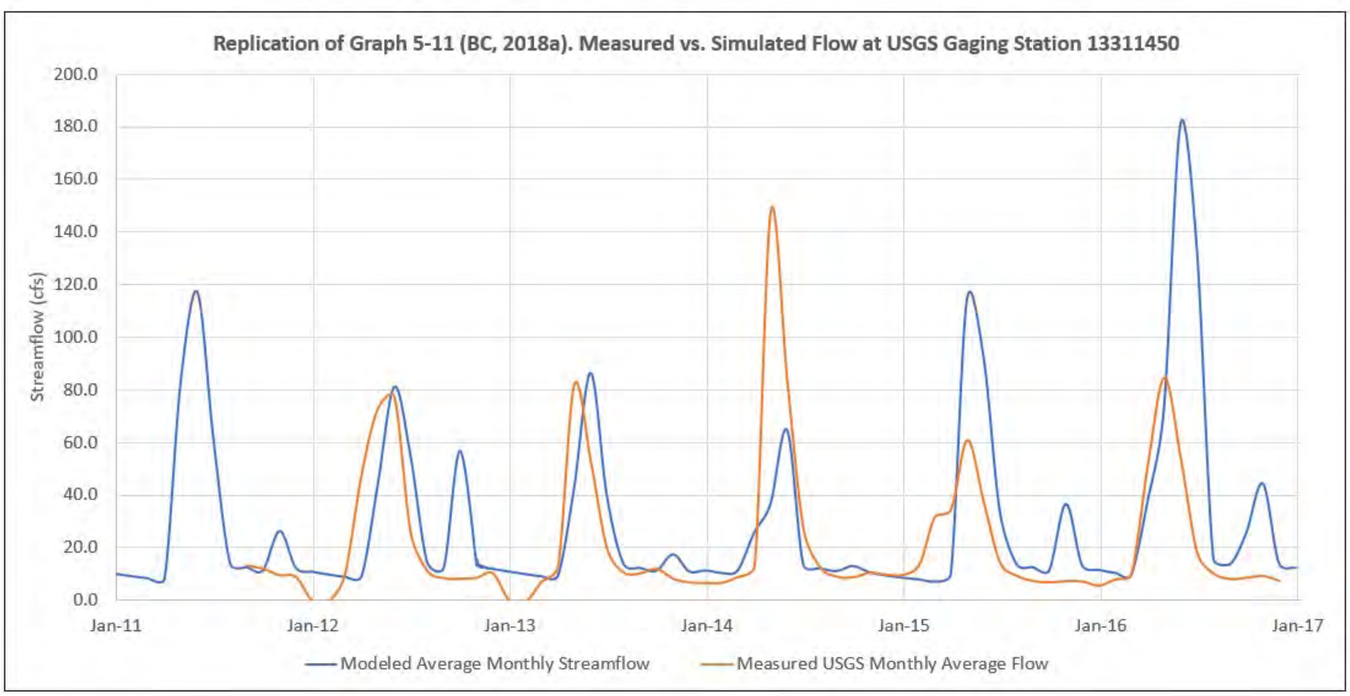


Graph 5-9. Measured vs. Simulated Flow at USGS Gaging Station 1331250, EFSFSR above Sugar Creek (Linear)





Graph 5-11. Measured vs. Simulated Flow at USGS Gaging Station 13311450, Sugar Creek near Stibnite, ID (Linear)



4.0 ALTERNATIVE MODEL TESTS

As noted previously, the sensitivity analysis conducted on the model calibration did not test enough parameters, nor the parameters to a wide-enough range, to characterize the bounds of parameters that produce equally calibrated models. In other words, the sensitivity analysis did not find alternative calibrated models from which to test the predictive models. Two sensitivity tests were performed using the provided model files to test the conceptual model assumptions for the shallow bedrock:

- Calibration Alternative (CA) Test 1 – All of model layer 2 was set to a uniform K value of 1 ft/d and specific yield in model layer 2 was increased to 3 percent.
- CA Test 2 – All of model layer 2 was set to a uniform K value of 0.5 ft/d, vertical anisotropy was removed from all model layers, and specific yield in model layer 2 was increased to 3 percent.

The purpose of these CA models was to test the conceptual model assumption that the shallow bedrock beneath the washes has a higher K than elsewhere and to test the very low value of specific yield that was assumed in the existing conditions model. These CA models were designed to potentially encourage recharge in upland areas to penetrate deeper into the groundwater system. The transient monitor well data from SPF 2017, Appendix C was input to these models to directly compare the calibrations to the provided existing conditions model calibration. No other changes were made to the models. It should be noted that these two CA tests of the model calibration do not represent the range to which modeled parameters can be adjusted while maintaining model calibration (i.e. they do not bracket the potential model calibrations), but instead are simply two alternative conceptualizations of the shallow bedrock system and vertical flow patterns.

The model calibration is improved, overall, for both CA tests compared to the provided model calibration, particularly upgradient of Hangar Flats pit and near Yellow Pine pit. The difference in the absolute calibration residual for CA Test 1, for the most recent data value for each monitor well, is shown on Figure 2, and for CA Test 2 on Figure 3. A green marker indicates an improvement to the calibration, black indicates worsened calibration, and white indicates that the residual is within 10 feet of the provided model calibration. Additionally, model layer 1 residuals are represented with circles and model layer 2 with squares, and the labels are the amount the absolute residual changed from the provided model calibration (positive = improved).

Calibration residuals are unchanged (within 10 feet) or improved for all wells for CA Test 2 (Figure 3), while they are improved for wells near Yellow Pine pit and upgradient of Hangar Flats for CA Test 1 (Figure 2). The resulting calibration statistics were improved for both CA tests compared to the provided model calibration, particularly in the shallow bedrock (model layer 2). These statistics are summarized here:

	Average Residual (ft)			% RMS Error		
	Model	Layer 1	Layer 2	Model	Layer 1	Layer 2
Midas_ExistingCond.gwv	-27.2	2.6	-53.6	6.0%	3.3%	7.7%
CA Test 1	-23.3	-0.4	-44.3	4.8%	3.1%	6.3%
CA Test 2	-4.7	4.8	-12.9	2.9%	3.0%	3.2%

The two CA test models were run through the 14-year mine operations model by making the same changes to K and specific yield in the provided No Action and Proposed Action model files (*Midas_NoAction_AvgPrecip.gwv, and Midas_MineOps_AvgPrecip.gwv)*. The Proposed Action predictive model that was presented in the modeling report was run twice, once without RIBs to determine the amount of groundwater discharge to the pit, and once with the RIBs defined as the difference between the groundwater discharge to the pit and the required water for mine operations (BC 2018b). Due to inadequate time available for public comment on the DEIS, long model run times, and difficult-to-implement changes to the RIB recharge with the provided information, this two-step process was not repeated here and RIB recharge was not altered from the Proposed Action model scenario. If the groundwater discharge to the open pits is higher than in the Proposed Action model, the RIB recharge will be underestimated. If the groundwater discharge to the open pits is lower than in the Proposed Action model, the RIB recharge will be overestimated.

The post-closure models were not simulated with CA Test models 1 and 2 because of inadequate available time associated with this review period.

Total discharges to the open pits increased by 33 percent in CA Test 1 (discharges up to 4,000 gpm compared to 3,000 gpm) and decreased by 33 percent in CA Test 2 (discharges as low as 1,000 gpm compared to 1,500 gpm), compared to the Proposed Action model. It should be noted again that these comparisons are made to the Proposed Action model discussed in BC 2018b (Alternative 1 in the DEIS), rather than to Alternative 2 discussed in BC 2019b, because the provided model files appear to be for DEIS Alternative 1 (DEIS Section 4.8.1.1.2). It should also be noted again that these changes in predicted pit dewatering rates do not represent a bracket of possible changes relative to the provided Proposed Action model results, but simply the range observed from two uncertainty tests using alternative model calibrations, and without altering recharge through the RIBs.

Streamflow was graphed in the same manner as in the Proposed Action modeling report (BC 2018b). Streamflow was compared between the provided No Action model and no-action models run for CA Tests 1 and 2. Streamflow for these No Action models generally falls from the seasonal peak flows to the seasonal low flows faster with CA Tests 1 and 2 than in the provided No Action model, resulting overall, in less streamflow during the low season. An example of this is provided below for the EFSFSR above Meadow Creek near Stibnite gage location (#13310800). The impact of mining on streamflow was compared for the provided model files (DEIS Alternative 1), and CA Tests 1 and 2. Streamflow for the No Action model was subtracted from the streamflow for the Proposed Action model and expressed as percent change; therefore negative values represent a decrease in streamflow during mining. Comparing CA Tests 1 and 2 to the provided models (DEIS Alternative 1), the greatest difference between the models is in the simulated impact of mining on streamflow at Sugar Creek near Stibnite (#13311450), as shown in the graph below. CA Test 1 shows decreased streamflow from mining of approximately 12 percent in mine years 5 and 6, compared to 4 percent in the provided model (DEIS Alternative 1). It should be noted again that the RIB recharge was not modified for CA Tests 1 and 2.

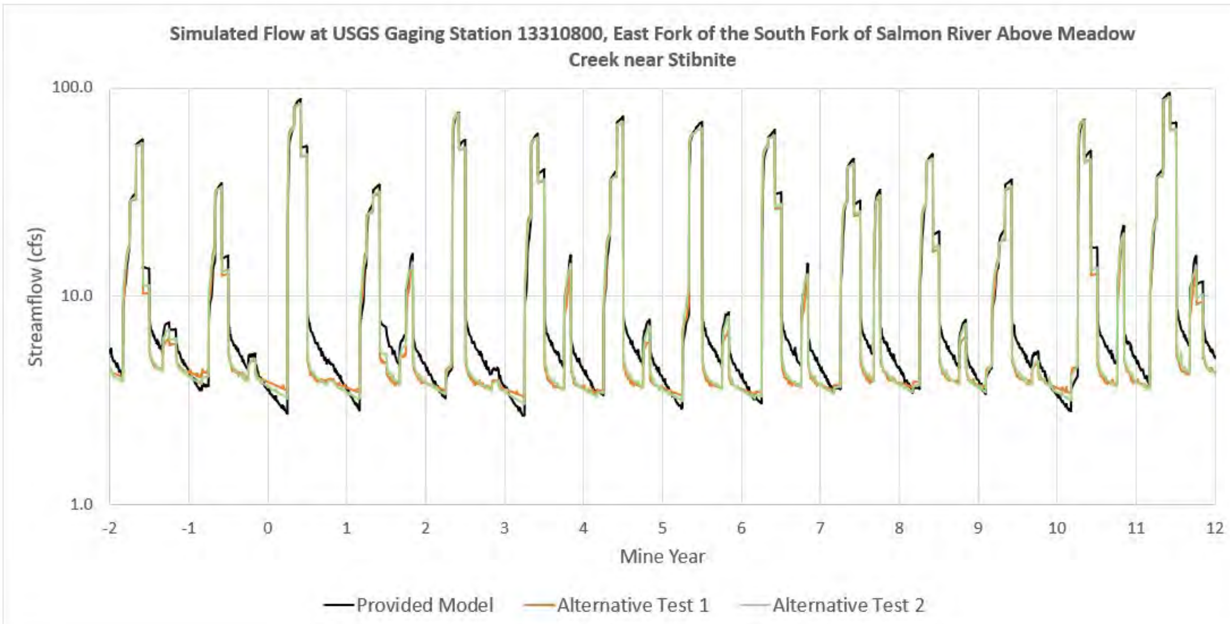
Sensitivity tests should be conducted on the assumed SFR package parameters including the assumption of a constant stream depth to determine if the simulated streamflow may show greater impacts from mining with less restrictive bounds on the way the streams are modeled. Given the long model run times and insufficient information to replicate all components of the predictive models (Modified PRO Alternative 2 model files and RIB information), these sensitivity tests were not attempted here.

Lastly, drawdown of the groundwater surface due to mining was compared between the provided Proposed Action model and CA Tests 1 and 2. Drawdown was calculated as outlined in the Proposed Action modeling report (BC 2018b) for alluvium (model layer 1) and shallow bedrock (model layer 2) at the same time intervals (mine years 6, 7, and 12) by subtracting groundwater elevations in the mine operations models from the no action models. The simulated extent of drawdown was greater within the shallow bedrock (model layer 2) in CA Tests 1 and 2 compared to the provided Proposed Action model. Near the Yellow Pine and West End pits, simulated drawdown in the shallow bedrock extends further across Sugar Creek, and extends further south of the pits toward Hangar Flats, with differences in the simulated extent of the 5-foot drawdown contour up to approximately 500 percent in CA Test 1 (Figure 4 below). Near Hangar Flats pit, simulated drawdown extends further in all directions by approximately 130 to 330 percent. The Proposed Action modeling report (BC 2018b) suggested that drawdown in the alluvium is more important than drawdown in the shallow bedrock because of the potential interaction with surface water streams. But if the predicted extent of drawdown in the shallow bedrock is greater than expected, seeps, springs, and groundwater dependent ecosystems could potentially be impacted, as well as downgradient streamflow (DEIS Section 4.8.6).

The simulated contours of drawdown for CA Tests 1 and 2 reached the model boundary downgradient of the Yellow Pine pit. This indicates that the model boundary is too close to the pits to allow proper testing of the impacts of dewatering the pit. Exterior model boundaries should be located far enough from the features of interest, so they do not impact the predictive results. The radial nature of groundwater discharge to these open pits cannot properly be simulated if the model boundary encroaches upon the pit. This impacts not only the CA

Tests 1 and 2 discussed here, but the predictive modeling presented in the Proposed Action modeling report and Modified PRO modeling report (BC 2018b, 2019b).

It should also be noted, again, that the differences shown for CA Tests 1 and 2 do not represent the maximum, potential variation in model predictions that may result from alternate calibrated models. These are simply two tests of the bedrock conceptual model that maintain model calibration and were used to illustrate that there is variability in the model predictions of the impacts from mining to the groundwater system. These tests with the model illustrate the need for a proper uncertainty analysis of the predictive models to a) better estimate potential impacts to seeps, springs, and groundwater dependent ecosystems (DEIS Section 4.8.6), b) better estimate groundwater dewatering rates and flow patterns during mine operations for water management and facility stability planning, and c) better estimate mining and post-closure geochemical impacts to groundwater and surface water that are influenced by groundwater flow patterns, interaction with mining facilities, and flow-through pit lakes.


Simulated Flow at USGS Gaging Station 13310800, East Fork of the South Fork of Salmon River Above Meadow Creek near Stibnite
100.0
10.0
1.0
Streamflow (cfs)
-2
-1
0
1
2
3
4
5
6
7
8
9
10
11
12
Mine Year
Provided Model
Alternative Test 1
Alternative Test 2

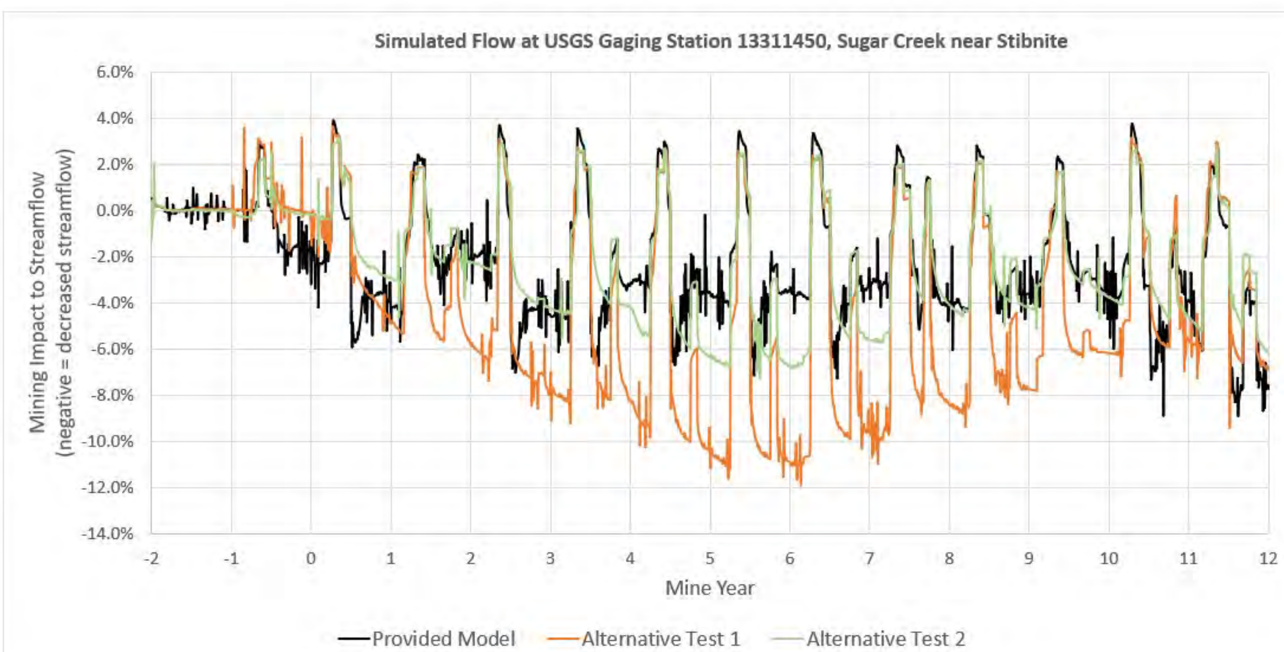

Simulated Flow at USGS Gaging Station 13311450, Sugar Creek near Stibnite
6.0%
4.0%
2.0%
0.0%
-2.0%
-4.0%
-6.0%
-8.0%
-10.0%
-12.0%
-14.0%
Mining Impact to Streamflow (negative = decreased streamflow)
-2
-1
0
1
2
3
4
5
6
7
8
9
10
11
12
Mine Year
Provided Model
Alternative Test 1
Alternative Test 2

Figure 1. Most recent calibration residual using groundwater elevation data from Appendix C (SPF 2017)

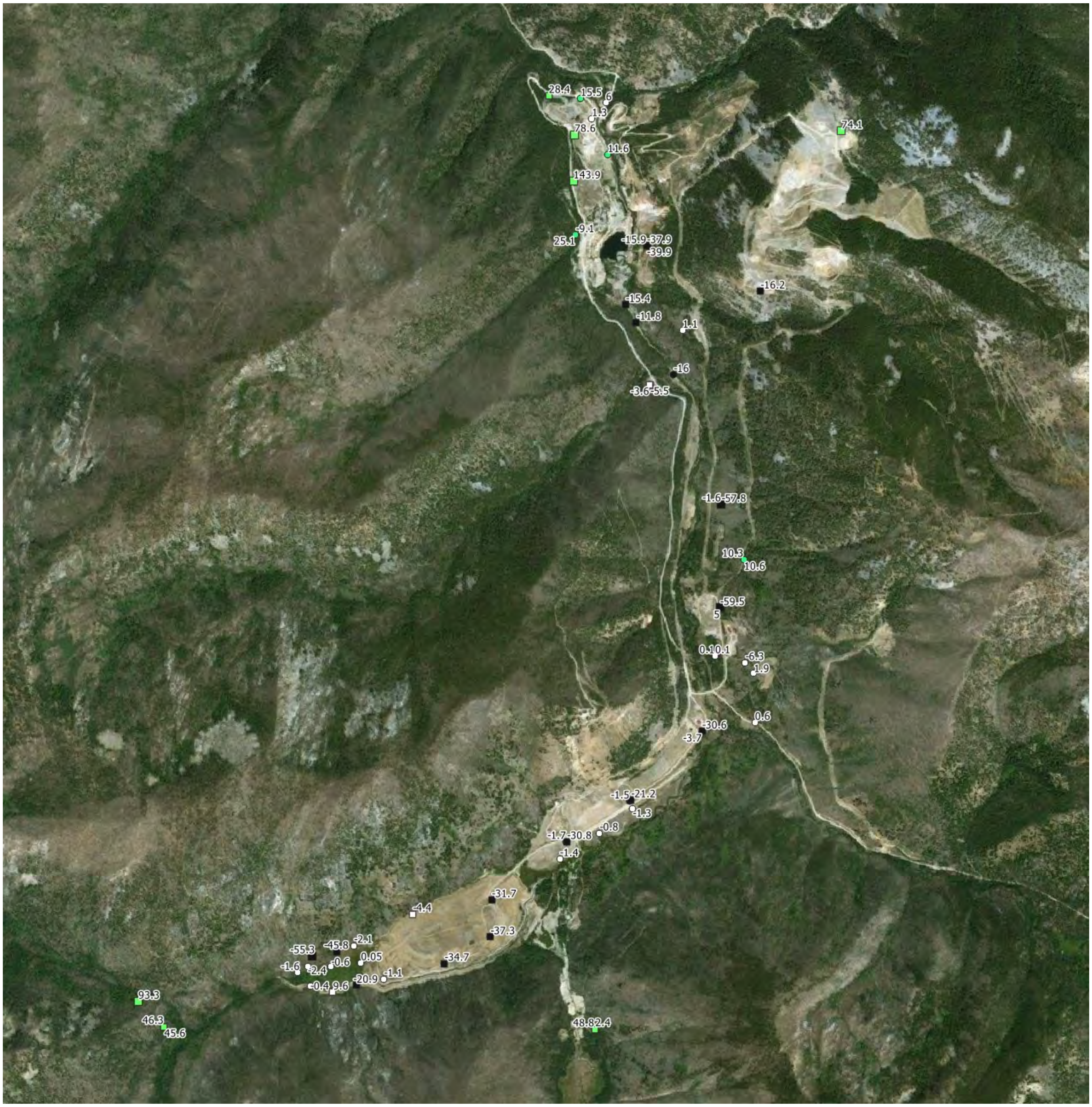


Figure 2. Change in absolute residual for CA Test 1, compared to the provided model calibration (*Midas_ExistingCond.gwv),* using transient groundwater elevations from Appendix C (SPF 2017). Green = improved absolute residual in CATest 1, Black = worsened absolute residual, White = unchanged within 10 feet.

Figure 3. Change in absolute residual for CA Test 2, compared to the provided model calibration (*Midas_ExistingCond.gwv),* using transient groundwater elevations from Appendix C (SPF 2017). Green = improved absolute residual in CA Test 2, Black = worsened absolute residual, White = unchanged within 10 feet.

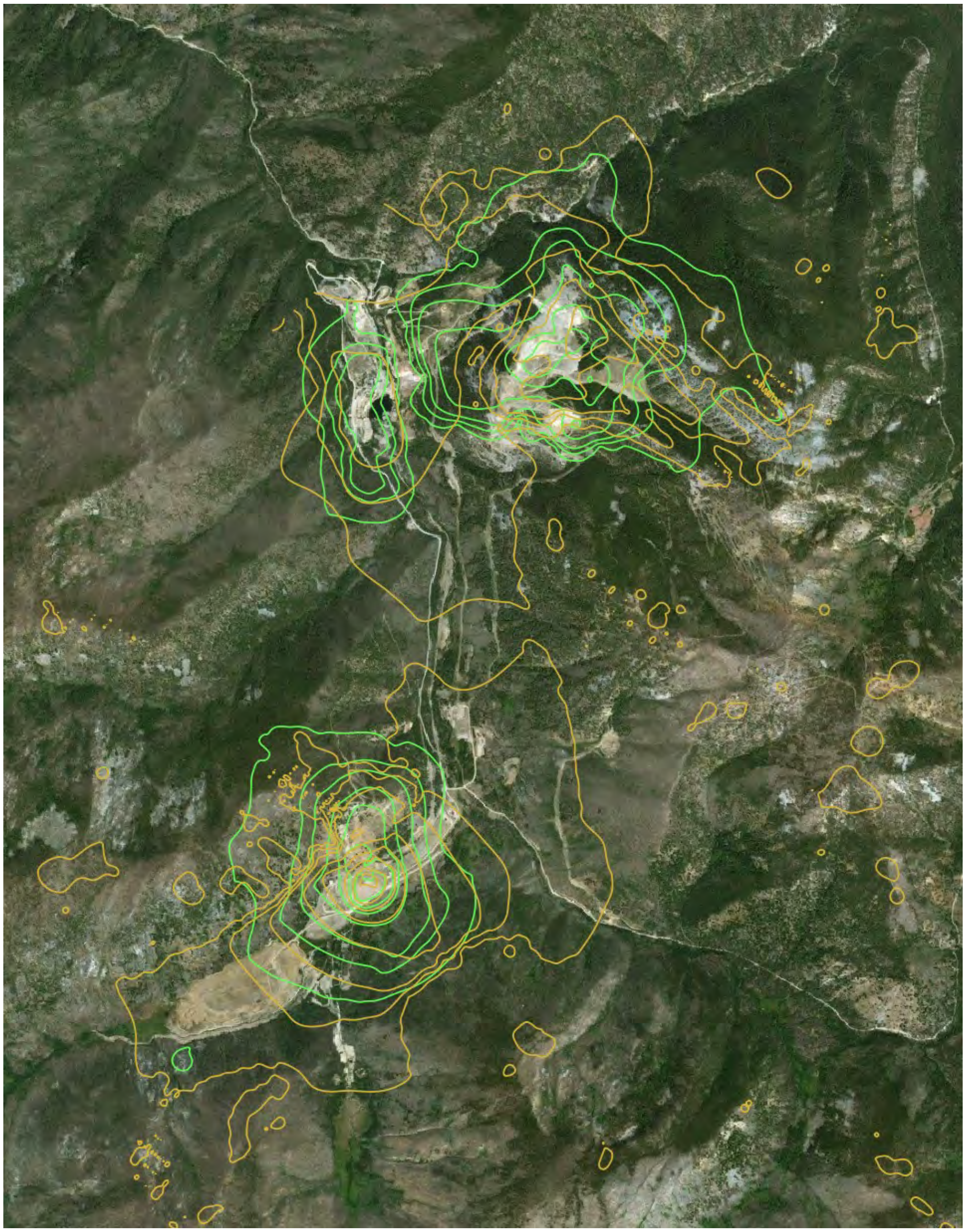


Figure 4. Simulated drawdown at mine year +12 (model time = 14 years), for the provided Proposed Action model (green), and CA Test 1 (yellow) for drawdown 5 feet and greater.

5.0 CONCLUSIONS

The modeling effort undertaken to estimate the impacts of proposed mining on the groundwater and surface water system near Stibnite, Idaho needs significant additional testing before predictions can be made and conclusions can be formed. Additionally, numerous inconsistencies were found between the provided model files and the reported model results, bringing the validity of the modeling effort into question. It was not possible to exactly correlate the provided model files with the information presented in the modeling reports, to verify that the model results are appropriate to use for predictive purposes of estimating impacts from mining operations. A simple sensitivity analysis to test assumptions of the bedrock system was conducted and two alternative test models were found that maintain the model calibration. Using these two models to test the mine operation period showed that simulated groundwater discharge to the pits, seasonal low streamflow, and drawdown of bedrock groundwater can vary substantially, illustrating the need for a proper uncertainty analysis.

Questions can be asked as to whether the modeling approach and methodology was appropriate, but a comprehensive sensitivity analysis and uncertainty analysis would go a long way to, at a minimum, address the reliability of the predicted impacts from mining. However, additional model testing would not satisfy all the uncertainty that stems from a poor characterization of the regional groundwater system. Additional, basic data compilation should have been undertaken not only to improve the quality of the groundwater model, but to better characterize the hydrologic and hydrogeologic systems, both for the purpose of predicting impacts and planning for restoration.

Sincerely,

Betsy Semmens

Betsy Semmens, R.G.